Biol Cybern (2005) 93: 256–274
DOI 10.1007/s00422-005-0002-3

ORIGINAL PAPER

Biological Cybernetics

Paul C. Bressloff

Spontaneous symmetry breaking in self–organizing neural fields

Received: 1 December 2004 / Accepted: 6 June 2005 / Published online: 29 September 2005

Abstract We extend the theory of self-organizing neural fields in order to analyze the joint emergence of topography and feature selectivity in primary visual cortex through spontaneous symmetry breaking. We first show how a binocular one-dimensional topographic map can undergo a pattern forming instability that breaks the underlying symmetry between left and right eyes. This leads to the spatial segregation of eye specific activity bumps consistent with the emergence of ocular dominance columns. We then show how a 2-dimensional isotropic topographic map can undergo a pattern forming instability that breaks the underlying rotation symmetry. This leads to the formation of elongated activity bumps consistent with the emergence of orientation preference columns. A particularly interesting property of the latter symmetry breaking mechanism is that the linear equations describing the growth of the orientation columns exhibits a rotational shift-twist symmetry, in which there is a coupling between orientation and topography. Such coupling has been found in experimentally generated orientation preference maps.

1 Introduction

One of the striking features of the visual system is that the visual world is mapped on to the cortical surface in a topographic manner. This means that neighboring points in a visual image evoke activity in neighboring regions of visual cortex. Superimposed upon this topographic map are additional maps reflecting the fact that neurons respond preferentially to stimuli with particular features. Neurons in the retina, lateral geniculate nucleus (LGN) of the thalamus, and primary visual cortex (V1) respond to light stimuli in restricted regions of the visual field called their classical receptive fields (RFs). Patterns of illumination outside the RF of a given neuron cannot generate a response directly, although they can significantly modulate responses to stimuli within the RF via long–range cortical interactions (Fitzpatrick 2000). The RF is divided into distinct ON and OFF regions. In an ON (OFF) region illumination that is higher (lower) than the background light intensity enhances firing. The spatial arrangement of these regions determines the selectivity of the neuron to different stimuli. For example, one finds that the RFs of most V1 cells are elongated so that the cells respond preferentially to stimuli with certain preferred orientations (Hubel and Wiesel 1962). The RFs of retinal ganglion neurons and LGN neurons, on the other hand, are circularly symmetric and hence these neurons do not exhibit any stimulus orientation preference. Neurons in both the LGN and input layers of V1 are also segregated according to whether or not they respond preferentially to left-eye or right-eye stimuli (ocular dominance) (Hubel and Wiesel 1977). Neurons in V1 with similar feature preferences tend to arrange themselves in vertical columns so that to a first approximation the layered structure of cortex can be ignored (LeVay and Nelson 1991). The corresponding feature maps then describe the spatial distribution of these columns as one moves tangentially over the surface of cortex. In recent years much information has accumulated regarding the two–dimensional distribution of both orientation preference and ocular dominance columns using optical imaging techniques (Blasdel and Salama 1986; Bonhoeffer and Grinvald 1991). These experimental studies indicate that there is an underlying periodicity in the microstructure of V1 with a period of approximately 1mm (in cats and primates). The fundamental domain of this tiling of the cortical plane is the hypercolumn, which contains two sets of orientation preferences $\theta \in [0, \pi)$ per eye, organized around a pair of orientation singularities or pinwheels (Obermayer and Blasdel 1993).

It is generally accepted that the preference of cortical neurons for particular stimulus features such as orientation and ocular dominance arises primarily from the spatial arrangement of convergent feedforward afferents from the LGN (or from other layers of cortex). The experimental observation that stimulus deprivation can modify ocular dominance columns during a critical period of postnatal development in

P.C. Bressloff
Department of Mathematics,
University of Utah,
Salt Lake City
Utah

cats and primates provides strong evidence that the formation of these columns is activity–dependent (Hubel et al. 1977; LeVay et al. 1978; Stryker and Harris 1986). On the other hand, since orientation and OD columns are already present in newly born primates, and the segregation of OD columns occurs as early as one week after LGN axons enter layer 4 of V1 in ferrets (Crowley and Katz 2000), it has been suggested that activity–independent molecular cues could play a major role in the initial formation of columns. However, specific molecules have not yet been found. Moreover, it is possible that spontaneous retinal waves (Wong et al. 1993) or endogenous activity in the cortico-geniculate feedback loop could support an activity–dependent mechanism in the early stages of columnar development (Penn and Shatz 1999). A much more likely role for molecular cues is in the initial development of the topographic map, where geniculate axons are guided to targets in input layer 4 of V1 after reaching the cortical subplate (Ghosh and Shatz 1992). However, the resulting map is rather crude and some form of activity appears to be necessary for the subsequent refinement of the topographic map through the pruning of initially exuberant axonal arborizations (Catalano and Shatz 1998).

A large number of models have been proposed that describe activity–dependent development as a self–organizing Hebbian process (see the review of Swindale 1996). In the case of correlation–based developmental models (Linsker 1986; Miller et al. 1989; Miller 1994; Erwin and Miller 1998), the statistical structure of input correlations provides a mechanism for spontaneously breaking some underlying symmetry of the neuronal receptive fields leading to the emergence of feature selectivity. When such correlations are combined with intracortical interactions, there is a simultaneous breaking of translation symmetry across cortex leading to the formation of a spatially periodic cortical feature map. Correlation–based models are essentially linear, so that considerable insight into the developmental process can be obtained by solving an associated eigenvalue problem (Mackay and Miller 1990; Miller and MacKay 1994; Wimbauer et al. 1998). One of the possible limitations of this class of model is that a regular topographic map is assumed already to exist before feature–based columns begin to develop. In order to model the joint development of topography and cortical feature maps, it appears necessary to introduce some form of nonlinear competition for activation (Willshaw and von der Malsburg 1976; Kohonen 1982; Goodhill 1993; Piepenbrock and Obermayer 1999), neurotrophic factors (Elliott and Shadbolt 1999) or a combination of the two (Whitelaw and Cowan 1981).

An alternative mathematical formulation of topographic map formation has been developed by Amari using the theory of self–organizing neural fields (Takeuchi and Amari 1979; Amari 1980, 1983, 1989). The basic network model involves a form of non–competitive Hebbian learning in the presence of hard threshold, nonlinear firing rate functions. It is found numerically that starting from a crude topographic map, the system evolves to a more refined continuous map that is dynamically stable. In the simpler one–dimensional case, conditions for the existence and stability of such a map can be derived analytically. Moreover, it can be shown that under certain circumstances the continuous topographic map undergoes a pattern forming instability that spontaneously breaks continuous translation symmetry, and the map becomes partitioned into discretized blocks; it has been suggested that these blocks could be a precursor for the columnar microstructure of cortex (Takeuchi and Amari 1979; Amari 1989). Given that cortical columns tend to be associated with stimulus features such as ocular dominance and orientation, this raises the interesting question whether or not such features could also emerge through the spontaneous symmetry breaking of self–organizing neural fields. Some recent numerical studies support such a possibility (Woodbury et al. 2002; Fellenz and Taylor 2002). In this paper, we explore this issue from a mathematical perspective by extending Amari's original analysis to networks with distinct left–eye and right–eye afferents and to two–dimensional networks. Throughout the paper we emphasize the important role of symmetry.

2 Neural field theory

We begin by reviewing Amari's neural field theory for topographic map formation (Takeuchi and Amari 1979; Amari 1980, 1983, 1989), and introduce the basic notation that will be used throughout the paper. We also present an alternative derivation of the linear stability conditions for one–dimensional topographic maps, which is more easily extended to networks with distinct left/right eye afferents (Sect. 3) and to two–dimensional networks (Sect. 4).

2.1 Network model

A schematic diagram of the basic network model is shown in Fig. 1. The lateral geniculate nucleus (LGN) and the primary

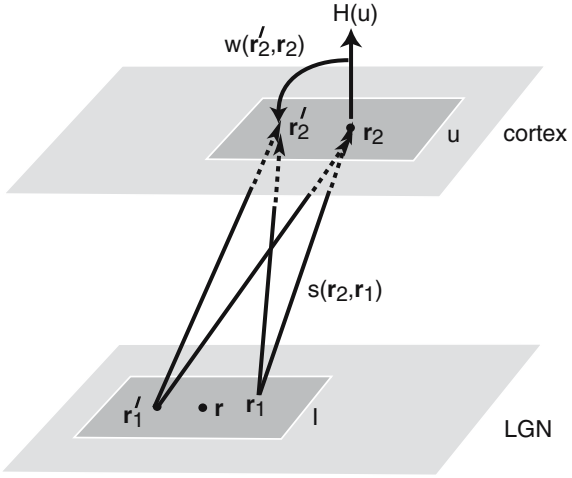

Fig. 1 Basic network architecture illustrating how a localized input I centered at position $\mathbf{r}$ in the LGN layer induces a corresponding response u in the cortical layer. (The global inhibition is not shown)

visual cortex are treated as two–dimensional continuous neural sheets. Let $\mathbf{r}_1 = (x_1, y_1) \in \Omega_1$ denote a point in the LGN layer and $\mathbf{r}_2 = (x_2, y_2) \in \Omega_2$ a point in the cortical layer. The strength of feedforward excitatory afferents connecting these two points is denoted by $s(\mathbf{r}_2, \mathbf{r}_1)$. Suppose, for the moment, that these feedforward afferents are fixed and that there exists presynaptic input activity $I(\mathbf{r}_1|\mathbf{r})$ centered about the point $\mathbf{r}$ in the LGN. This is supplemented by a global inhibitory input I_0 with associated feedforward synaptic density $s_0(\mathbf{r}_2)$ (Takeuchi and Amari 1979; Amari 1980). The total weighted input to a point $\mathbf{r}_2$ in cortex is then given by

$$v(\mathbf{r}_2|\mathbf{r}) = \int_{\Omega_1} s(\mathbf{r}_2, \mathbf{r}_1) I(\mathbf{r}_1|\mathbf{r}) \mathrm{d}\mathbf{r}_1 - s_0(\mathbf{r}_2) I_0. \tag{2.1}$$

Cortical neurons also receive synaptic inputs from recurrent connections within the layer, which are taken to be homogeneous and isotropic. Thus, the synaptic density from neurons at $\mathbf{r}_2'$ to neurons at $\mathbf{r}_2$ is of the form $w(|\mathbf{r}_2 - \mathbf{r}_2'|)$ for some prescribed function w. Given these two sources of input, the activity $u(\mathbf{r}_2, t|\mathbf{r})$ of neurons at $\mathbf{r}_2$ at time t in response to a stimulus centered at $\mathbf{r}$ satisfies the neural field equation (Takeuchi and Amari 1979; Amari 1980)

$$\eta \frac{\partial u}{\partial t} = -u(\mathbf{r}_2, t|\mathbf{r}) + \int_{\Omega_2} w(|\mathbf{r}_2 - \mathbf{r}_2'|) H(u(\mathbf{r}_2', t|\mathbf{r})) \mathrm{d}\mathbf{r}_2' + v(\mathbf{r}_2|\mathbf{r}) - h, \tag{2.2}$$

where η is a membrane time constant, $-h$ determines the background level of activity in the absence of stimuli, and $H(u)$ denotes the output firing rate function, which is taken to be a Heaviside function: $H(u) = 1$ if $u > 0$ and $H(u) = 0$ otherwise. It is assumed that each stimulus is presented to the network for a sufficiently long time, so the activity u converges to a stable equilibrium solution of the integral equation

$$u(\mathbf{r}_2|\mathbf{r}) = \int_{\Omega_2} w(|\mathbf{r}_2 - \mathbf{r}_2'|) H(u(\mathbf{r}_2'|\mathbf{r})) \mathrm{d}\mathbf{r}_2' + v(\mathbf{r}_2|\mathbf{r}) - h. \tag{2.3}$$

Now suppose that modifications in the strength of the feedforward afferents s, s_0 occur on a much slower time scale than both the relaxation time of the activity u and the time interval over which each input is sampled. This adiabatic condition implies that the equilibrium activity u is slaved to the slowly changing synaptic weights s, s_0. A Hebbian rule is assumed for the dynamics of the feedforward connections such that during the presentation of a single input centered at $\mathbf{r}$,

$$\eta \frac{\partial s}{\partial \tau} = -s(\mathbf{r}_2, \mathbf{r}_1, \tau) + c H(u(\mathbf{r}_2, \tau|\mathbf{r})) I(\mathbf{r}_1|\mathbf{r}) \tag{2.4}$$

and

$$\eta \frac{\partial s_0}{\partial \tau} = -s_0(\mathbf{r}_2, \tau) + \hat{c} H(u(\mathbf{r}_2, \tau|\mathbf{r})) I_0, \tag{2.5}$$

where $\tau = \varepsilon t$ for $0 < \varepsilon \ll 1$ and $c, \hat{c}$ are constants. Note that there is a separation of time–scales in which $u(\mathbf{r}_2, \tau|\mathbf{r}) = \lim_{t\to\infty} u(\mathbf{r}_2, t, \tau|\mathbf{r})$ is a stable equilibrium solution of Eqs. 2.2 and 2.1 with $s = s(\mathbf{r}_2, \mathbf{r}_1, \tau)$, $s_0 = s_0(\mathbf{r}_2, \tau)$ for fixed slow time variable τ.

The final step in the formulation of neural field theory is to assume that the the center $\mathbf{r}(\tau)$ of an input at time τ is generated at random from some probability density $\rho(\mathbf{r})$ (Takeuchi and Amari 1979; Amari 1980). This implies that Eqs. 2.4 and 2.5 become a set of stochastic differential equations. Given the above adiabatic condition, it is then possible to take an ensemble average over the distribution of inputs to obtain the deterministic equations

$$\eta \frac{\partial s}{\partial \tau} = -s(\mathbf{r}_2, \mathbf{r}_1, \tau) + c \left\langle H(u(\mathbf{r}_2, \tau|\mathbf{r}')) I(\mathbf{r}_1|\mathbf{r}') \right\rangle \tag{2.6}$$

and

$$\eta \frac{\partial s_0}{\partial \tau} = -s_0(\mathbf{r}_2, \tau) + \hat{c} \left\langle H(u(\mathbf{r}_2, \tau|\mathbf{r}')) \right\rangle I_0, \tag{2.7}$$

where $\langle\ \rangle$ denotes the ensemble average over $\mathbf{r}'$. These averaged equations involve the approximation that u depends on $\langle s\rangle$, $\langle s_0\rangle$ rather than s, s_0. For ease of notation, the averages $\langle s\rangle$, $\langle s_0\rangle$ are then simply denoted by s, s_0. The validity of such an approximation has been established analytically elsewhere (Geman 1979).

It is convenient to determine the slow variation in the weighted input v induced by changes in the feeedforward afferents for a fixed input centered at $\mathbf{r}$. Differentiating Eq. 2.1 with respect to τ gives

$$\eta \frac{\partial v}{\partial \tau} = \eta \int_{\Omega_1} \frac{\partial s(\mathbf{r}_2, \mathbf{r}_1, \tau)}{\partial \tau} I(\mathbf{r}_1|\mathbf{r}) \mathrm{d}\mathbf{r}_1 - \eta \frac{\partial s_0}{\partial \tau}(\mathbf{r}_2, \tau) I_0.$$

Using Eqs. 2.6 and 2.7, this reduces to

$$\eta \frac{\partial v}{\partial \tau} = -v(\mathbf{r}_2, \tau|\mathbf{r}) + \int_{\Omega_1} \rho(\mathbf{r}') g(\mathbf{r}|\mathbf{r}') H(u(\mathbf{r}_2, \tau|\mathbf{r}')) \mathrm{d}\mathbf{r}', \tag{2.8}$$

where

$$g(\mathbf{r}|\mathbf{r}') = c \int_{\Omega_1} I(\mathbf{r}_1|\mathbf{r}) I(\mathbf{r}_1|\mathbf{r}') \mathrm{d}\mathbf{r}_1 - \hat{c} I_0^2. \tag{2.9}$$

Further simplification can be achieved by assuming that the inputs are homogeneous and isotropic, $I(\mathbf{r}_1|\mathbf{r}) = I(|\mathbf{r}_1 - \mathbf{r}|)$, so that the input kernel $g(\mathbf{r}|\mathbf{r}') = g(|\mathbf{r} - \mathbf{r}'|)$. This is only valid if we ignore boundary effects either by setting $\Omega_{1,2} = R^2$ or by using periodic boundary conditions. For example, taking the inputs to be Gaussians, $I(|\mathbf{r}|) = A\mathrm{e}^{-\mathbf{r}^2/2\sigma^2}$, then Eq. 2.9 ensures that the input kernel g is also a Gaussian:

$$g(|\mathbf{r}|) = c\sigma^2 \pi A^2 \mathrm{e}^{-\mathbf{r}^2/4\sigma^2} - \hat{c} I_0^2. \tag{2.10}$$

If we also take $\rho(\mathbf{r})$ to be a uniform distribution then we obtain the homogeneous equations

$$u(\mathbf{r}_2, \tau|\mathbf{r}) = \int_{\Omega_2} w(|\mathbf{r}_2 - \mathbf{r}_2'|) H(u(\mathbf{r}_2', \tau|\mathbf{r})) \mathrm{d}\mathbf{r}_2' + v(\mathbf{r}_2, \tau|\mathbf{r}) - h \tag{2.11}$$

and

$$\eta\frac{\partial v}{\partial\tau} = -v(\mathbf{r}_2,\tau|\mathbf{r}) + \int_{\Omega_1} g(|\mathbf{r}-\mathbf{r}'|)H(u(\mathbf{r}_2,\tau|\mathbf{r}'))\mathrm{d}\mathbf{r}'. \quad (2.12)$$

(Note that the normalization factor for the uniform distribution can be absorbed into the coefficients $c, \hat{c}$). Equations 2.11 and 2.12 are the basic neural field equations for topographic map formation (Takeuchi and Amari 1979; Amari 1980). Rescaling the LGN and cortical coordinates appropriately, that is, ignoring the effects of cortical magnification, one can then look for homogeneous steady–state solutions of the form $u(\mathbf{r}_2|\mathbf{r}) = U(|\mathbf{r}_2-\mathbf{r}|)$ and $v(\mathbf{r}_2|\mathbf{r}) = V(|\mathbf{r}_2-\mathbf{r}|)$, where U is a unimodal function satisfying the fixed point equation

$$U(|\mathbf{r}_2-\mathbf{r}|) = \int_{\Omega_2} w(|\mathbf{r}_2-\mathbf{r}_2'|)H(U(|\mathbf{r}_2'-\mathbf{r}|))\mathrm{d}\mathbf{r}_2' + \int_{\Omega_1} g(|\mathbf{r}-\mathbf{r}'|) H(U(|\mathbf{r}_2-\mathbf{r}'|))\mathrm{d}\mathbf{r}' - h, \quad (2.13)$$

and

$$V(|\mathbf{r}_2-\mathbf{r}|) = \int_{\Omega_1} g(|\mathbf{r}-\mathbf{r}'|)H(U(|\mathbf{r}_2-\mathbf{r}'|))\mathrm{d}\mathbf{r}'. \quad (2.14)$$

Such a solution represents a continuous topographic map in which the center of LGN input activity at $\mathbf{r}\in\Omega_1$ is mapped to the center of cortical output activity at the corresponding point $\mathbf{r}\in\Omega_2$. We now discuss the existence and stability of such solutions in the simpler one–dimensional case originally analyzed by Amari (Takeuchi and Amari 1979; Amari 1980, 1989).

2.2 One-dimensional topographic map

One–dimensional versions of Eqs. 2.11 and 2.12 take the form

$$u(x_2,\tau|x) = \int_{-\infty}^{\infty} w(x_2-x_2')H(u(x_2',\tau|x))\mathrm{d}x_2' + v(x_2,\tau|x) - h \quad (2.15)$$

and

$$\eta\frac{\partial v}{\partial\tau} = -v(x_2,\tau|x) + \int_{-\infty}^{\infty} g(x-x')H(u(x_2,\tau|x'))\mathrm{d}x' \quad (2.16)$$

with

$$g(x-x') = c\int_{-\infty}^{\infty} I(x_1-x)I(x_1-x')\mathrm{d}x_1 - \hat{c}I_0^2. \quad (2.17)$$

We will assume that $g(x)$ is a monotonically decreasing function of x. This will hold, for example, if the inputs $I(x)$ are Gaussians $I(x) = A\mathrm{e}^{-x^2/2\sigma^2}$ and hence

$$g(x) = c\sigma\sqrt{\pi}A^2\mathrm{e}^{-x^2/4\sigma^2} - \hat{c}I_0^2. \quad (2.18)$$

Let us consider an equilibrium solution of the form

$$u_0(x_2|x) = U(x_2-x), \quad v_0(x_2|x) = V(x_2-x) \quad (2.19)$$

with U, V satisfying the one–dimensional version of the fixed point Eqs. 2.13 and 2.14:

$$U(x_2-x) = \int_{-\infty}^{\infty} w(x_2-x_2')H(U(x_2'-x))\mathrm{d}x_2' + \int_{-\infty}^{\infty} g(x-x')H(U(x_2-x'))\mathrm{d}x' - h,$$

and

$$V(x_2-x) = \int_{-\infty}^{\infty} g(x-x')H(U(x_2-x'))\mathrm{d}x'.$$

In particular, we seek a unimodal solution U with

$$U(x) > 0,\ |x| < a, \quad U(x) = 0,\ |x| = a, \quad U(x) < 0,\ |x| > a, \quad (2.20)$$

where $2a$ is the width of the excited region (activity bump) in cortex. Then

$$U(x) = W(x+a) + W(x-a) + G(x+a) + G(x-a) - h \quad (2.21)$$

with

$$W(x) = \int_0^x w(x')\mathrm{d}x', \quad G(x) = \int_0^x g(x')\mathrm{d}x'. \quad (2.22)$$

The corresponding width of the activity bump is determined from the threshold conditions $U(\pm a) = 0$, which yields the implicit equation

$$W(2a) + G(2a) = h. \quad (2.23)$$

The stability of the bump with respect to fluctuations on the fast time–scale t can be determined by linearizing the equation

$$\eta\frac{\partial U}{\partial t} = -U(x,t) + \int_{-\infty}^{\infty} w(x-x')H(U(x',t))\mathrm{d}x' + \int_{-\infty}^{\infty} g(x-x')H(U(x',t))\mathrm{d}x' - h, \quad (2.24)$$

about the bump solution, and this leads to the stability condition (Amari 1977)

$$W'(2a) + G'(2a) \equiv w(2a) + g(2a) < 0. \quad (2.25)$$

As shown elsewhere (Amari 1977; Takeuchi and Amari 1979), if w consists of short–range excitation and long–range inhibition (the so–called Mexican hat profile) and g is a monotonically decreasing function then there exists a unique stable bump solution U for a range of threshold values h. We will assume that this holds in the following analysis.

Translation symmetry

The one–dimensional neural field Eqs. 2.15 and 2.16 are equivariant with respect to the product group $\mathcal{T}\times\mathcal{T}$ of translations acting on the space $\mathbf{R}\times\mathbf{R}$ according to

$$T_{s,s'}(x_2,x)=(x_2+s,x_2+s'),\quad T_{s,s'}\in\mathcal{T}\times\mathcal{T}.$$

The corresponding group action on the neural fields u, v is

$$T_{s,s'}(u(x_2|x),v(x_2|x))=(u(x_2-s|x-s'),\\ v(x_2-s|x-s')).$$

Equivariance means that if (u,v) is a solution of the neural field equations then so is $T_{s,s'}(u,v)$. This is a more formal way of expressing the fact that the homogeneous system has an underlying translation symmetry. It is also important to note that the homogeneous equilibrium solution $u_0(x_2|x)=U(x_2-x)$, $v_0(x_2|x)=V(x_2-x)$ explicitly breaks the symmetry group from $\mathcal{T}\times\mathcal{T}\to\mathcal{T}$ with resulting group action

$$T_s(u_0(x_2|x),v_0(x_2|x))=(u_0(x_2-s|x-s),\\ v_0(x_2-s|x-s)),\ T_s\in\mathcal{T}$$

that is, $T_s=T_{s,s}$. We will show below that the homogeneous equilibrium solution can undergo a pattern forming instability that spontaneously breaks the remaining translation symmetry.

2.3 Linear stability analysis

In order to investigate the stability of the topographic map solution with respect to fluctuations on the slow time–scale τ, we linearize Eqs. 2.15 and 2.16 by introducing small perturbations of the form

$$u(x_2,\tau|x)=U(x_2-x)+p(x_2,\tau|x),\\ v(x_2,\tau|x)=V(x_2-x)+q(x_2,\tau|x) \tag{2.26}$$

and expanding to first order in p,q. This leads to the equations (on setting $\eta=1$)

$$\frac{\partial q}{\partial\tau}=-q(x_2,\tau|x)\\ +\int_{-\infty}^{\infty}g(x-x')H'(U(x_2-x'))p(x_2,\tau|x')\mathrm{d}x'$$

and

$$p(x_2,\tau|x)=\int_{-\infty}^{\infty}w(x_2-x_2')H'(U(x_2'-x))\\ \times p(x_2',\tau|x)\mathrm{d}x_2'+q(x_2,\tau|x).$$

Using the result

$$H'(U(x))=\alpha^{-1}\left[\delta(x-a)+\delta(x+a)\right], \tag{2.27}$$

where $\alpha=|U'(\pm a)|$ and $\delta(x)$ is the Dirac delta function, we obtain the pair of linear equations

$$\frac{\partial q}{\partial\tau}=-q(x_2,\tau|x)+\alpha^{-1}\left[g(x-x_2+a)\right.\\ \times p(x_2,\tau|x_2-a)\\ \left.+g(x-x_2-a)p(x_2,\tau|x_2+a)\right] \tag{2.28}$$

and

$$p(x_2,\tau|x)=q(x_2,\tau|x)\\ +\alpha^{-1}\left[w(x_2-x+a)p(x-a,\tau|x)\right.\\ \left.+w(x_2-x-a)p(x+a,\tau|x)\right]. \tag{2.29}$$

Equations 2.28 and 2.29 involve nonlocal terms located at the boundaries $x_2=x\pm a$ of the unperturbed activity bump. This indicates why it is possible to analyze the stability of the topographic map by restricting attention to the effects of perturbations at the boundaries of the activity bump as originally formulated by Amari (Amari 1977; Takeuchi and Amari 1979; Amari 1989). In particular, if $u(x_2,\tau|x)=0$ at $x_2=x\pm a+\Delta_\pm(x,\tau)$, then

$$0=U(\pm a+\Delta_\pm(x,\tau))+p(x\pm a+\Delta_\pm(x,\tau),\tau|x)\\ =U(\pm a)+U'(\pm a)\Delta_\pm(x,\tau)+p(x\pm a,\tau|x)\\ +\mathcal{O}(\Delta^2)$$

that is,

$$\Delta_\pm(x,\tau)=\pm\alpha^{-1}p(x\pm a,\tau|x)$$

since $U(\pm a)=0$ and $U'(\pm a)=\mp\alpha$. Two particular examples of boundary perturbations are illustrated in Fig. 2: a uniform expansion of the bump for which $p(x+a,\tau|x)=p(x-a,\tau|x)$ and a shift in the center of the bump for which $p(x+a,\tau|x)=-p(x-a,\tau|x)$. In this paper we choose to work directly with the linear Eqs. 2.28 and 2.29, since these are more simply extended to the case of two–dimensional networks. Moreover, they take into account perturbations outside the boundary domain of the bump. However, the resulting stability conditions are equivalent to those derived following the boundary approach of Amari (Takeuchi and Amari 1979; Amari 1989): we show this explicitly in the case of one–dimensional topographic maps. Note that a similar approach to the one adopted here has previously been used to study the stability of activity bumps in single–layer networks with non–adapting synapses (Pinto and Ermentrout 2001; Folias and Bressloff 2004).

Defining

$$p_\pm(x,\tau)=p(x\pm a,\tau|x),\\ q_\pm(x,\tau)=q(x\pm a,\tau|x) \tag{2.30}$$

and setting $x_2=x\pm a$ in Eq. 2.29 gives the pair of equations

$$p_+(x,\tau)=q_+(x,\tau)+\alpha^{-1}[w(2a)p_-(x,\tau)\\ +w(0)p_+(x,\tau)], \tag{2.31}$$

$$p_-(x,\tau)=q_-(x,\tau)+\alpha^{-1}[w(0)p_-(x,\tau)\\ +w(2a)p_+(x,\tau)]. \tag{2.32}$$

Similarly, setting $x_2=x\pm a$ in Eq. 2.28 shows that

$$\frac{\partial q_+}{\partial\tau}=-q_+(x,\tau)+\alpha^{-1}\left[g(0)p_+(x,\tau)\right.\\ \left.+g(2a)p_-(x+2a,\tau)\right], \tag{2.33}$$

$$\frac{\partial q_-}{\partial\tau}=-q_-(x,\tau)+\alpha^{-1}\left[g(2a)p_+(x-2a,\tau)\right.\\ \left.+g(0)p_-(x,\tau)\right]. \tag{2.34}$$

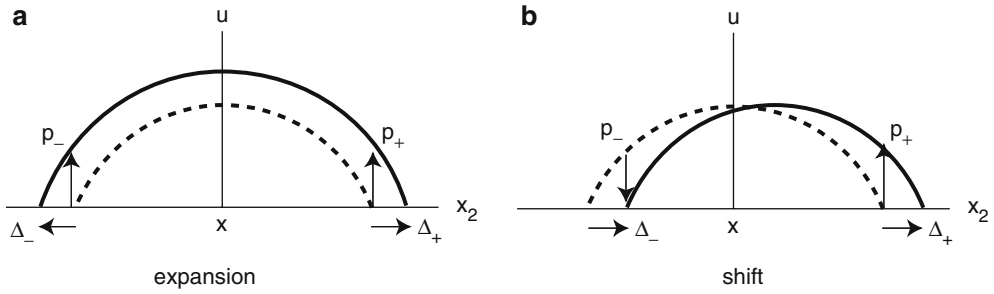

Fig. 2 Perturbations $p_\pm(x) = p(x \pm a|x)$ at the boundaries of a homogeneous bump solution centered about $x_2 = x$ and having width $2a$. Only the superthreshold part of the bump is shown. **a** Expansion of the bump such that $p_-(x) = p_+(x)$. **b** Shift in the position of the bump such that $p_-(x) = -p_+(x)$

We have used the fact that $w(x)$ and $g(x)$ are even functions. Equations 2.31–2.34 have eigensolutions of the form

$$p_\pm(x,\tau) = \mathrm{e}^{\lambda\tau}\mathrm{e}^{ikx}P_\pm(k),$$
$$q_\pm(x,\tau) = \mathrm{e}^{\lambda\tau}\mathrm{e}^{ikx}Q_\pm(k) \tag{2.35}$$

with the eigenvalue λ and eigenvectors $\mathbf{P} = (P_+, P_-)^{\mathrm{T}}$, $\mathbf{Q} = (Q_+, Q_-)^{\mathrm{T}}$ determined from the matrix equations

$$\lambda\mathbf{Q}(k) = -\mathbf{Q}(k) + \alpha^{-1}\mathbf{G}(k)\mathbf{P}(k),$$
$$\mathbf{P}(k) = \mathbf{Q}(k) + \alpha^{-1}\mathbf{W}\mathbf{P}(k), \tag{2.36}$$

where

$$\mathbf{W} = \begin{pmatrix} w_0 & w_2 \\ w_2 & w_0 \end{pmatrix}, \quad \mathbf{G}(k) = \begin{pmatrix} g_0 & g_2\mathrm{e}^{2ika} \\ g_2\mathrm{e}^{-2ika} & g_0 \end{pmatrix} \tag{2.37}$$

with $w_0 = w(0)$, $w_2 = w(2a)$, $g_0 = g(0)$, $g_2 = g(2a)$. It follows that

$$(\lambda + 1)\mathbf{Q}(k) = \mathbf{M}(k)\mathbf{Q}(k), \tag{2.38}$$

where

$$\mathbf{M}(k) = \mathbf{G}(k)\,[\alpha\mathbf{1} - \mathbf{W}]^{-1} \tag{2.39}$$
$$= \frac{1}{\Delta}\begin{pmatrix} (\alpha - w_0)g_0 + w_2g_2\mathrm{e}^{2ika} & (\alpha - w_0)g_2\mathrm{e}^{2ika} + w_2g_0 \\ (\alpha - w_0)g_2\mathrm{e}^{-2ika} + w_2g_0 & (\alpha - w_0)g_0 + w_2g_2\mathrm{e}^{-2ika} \end{pmatrix}$$

and

$$\Delta = (\alpha - w_0)^2 - w_2^2. \tag{2.40}$$

Thus

$$\lambda = \lambda_\pm(k) \equiv -1 + \mu_\pm(k), \tag{2.41}$$

where $\mu_\pm(k)$ are the eigenvalues of the matrix $\mathbf{M}(k)$,

$$\mu_\pm(k) = \frac{1}{\Delta}\left[\Sigma(k) \pm \sqrt{\Sigma(k)^2 - \Delta^2(g_0^2 - g_2^2)}\right] \tag{2.42}$$

with

$$\Sigma(k) = (\alpha - w_0)g_0 + w_2g_2\cos(2ka). \tag{2.43}$$

Note that we can express α in terms of the coefficients $g_{0,2}$, $w_{0,2}$ by differentiating Eq. 2.21 with respect to x,

$$\alpha \equiv |U'(\pm a)| = g_0 - g_2 + w_0 - w_2. \tag{2.44}$$

First consider the case $w_2 < 0$. Eqs. 2.40 and 2.44 then imply that

$$\Delta = (\alpha - w_0 + w_2)(\alpha - w_0 - w_2)$$
$$= (g_0 - g_2)(g_0 - g_2 - 2w_2) > 0,$$

since $g(x)$ is a monotonically decreasing function with $g_0 > g_2$, see Eq. 2.18. Define

$$\Sigma_{\min} = (\alpha - w_0)g_0 - |w_2g_2|,$$
$$\Sigma_{\max} = (\alpha - w_0)g_0 + |w_2g_2|$$

such that $0 < \Sigma_{\min} \le \Sigma(k) \le \Sigma_{\max}$ for all k. It follows from Eqs. 2.42 to 2.44 that

$$\Sigma_{\min}^2 - \Delta^2[g_0^2 - g_2^2]$$
$$= [|w_2|g_0 - (\alpha - w_0)|g_2|]^2 \ge 0$$

and, hence, $\mu_\pm(k)$ are real for all k. Combining this with the inequality $\Delta > 0$ shows that

$$\lambda_{max} \equiv \max_k \lambda_\pm(k) = -1 + \Delta^{-1}\left[\Sigma_+ + \sqrt{\Sigma_+^2 - \Delta^2[g_0^2 - g_2^2]}\right]$$
$$= -1 + \Delta^{-1}(g_0 + |g_2|) \times(\alpha - w_0 + |w_2|)$$
$$= \Delta^{-1}(\alpha - w_0 + |w_2|)(g_2 + |g_2|),$$

where we have used Eqs. 2.40 and 2.44. Finally, noting that $\Delta > 0$ and $\alpha - w_0 + |w_2| > 0$, we obtain the following stability conditions:

1. If $g_2 < 0$ then $\lambda_{\max} = \lambda_+(0) = 0$ and the topographic map is stable
2. If $g_2 > 0$ then $\lambda_{\max} = \lambda_+(\pi/2a) > 0$ and the topographic map is unstable. Moreover the fastest growing mode has a wavelength equal to $4a$, which is twice the width of an activity bump, and has vector components $\mathbf{P}(\pi/2a) = (1, -1)$. That is, the (real–valued) excited mode is of the form

$$p_\pm(x) = \pm\Gamma\cos\left(\frac{\pi(x - \bar{x})}{2a}\right),$$

where $\bar{x}$ is an arbitrary shift, reflecting hidden translation symmetry, and the amplitude Γ is arbitrary within the linear approximation.

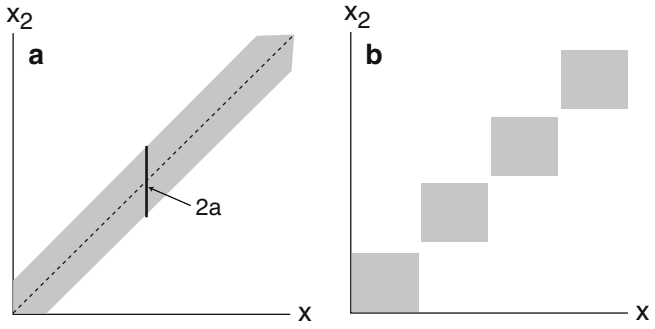

Fig. 3 **a** Homogeneous topographic map for $g(2a) < 0$. **b** Spatially periodic topographic map for $g(2a) > 0$ with a block–like microstructure. The *shaded* regions indicate where $u(x_2|x) > 0$

Note that the existence of a zero eigenvalue, $\lambda_+(0) = 0$, reflects the underlying translation symmetry of the homogeneous solution under simultaneous shifts $x \to x + \Delta$, $x_2 \to x_2 + \Delta$. In the above analysis we assumed that $w_2 < 0$. If $w_2 > 0$ then we require $g_2 < 0$ in order to satisfy the stability condition 2.25. Since the maximum eigenvalue is positive for $w_2 > 0$ and $g_2 < 0$,

$$\lambda_{max} = -1 + \Delta^{-1}(g_0 + |g_2|)(\alpha - w_0 + w_2)$$
$$= \frac{g_0 - g_2}{g_0 - g_2 - 2w_2} > 0,$$

it follows that the topographic map is unstable. Therefore, $g_2 < 0$ and $w_2 < 0$ are necessary and sufficient conditions for the stability of the one–dimensional topographic map, as previously shown by Takeuchi and Amari (Takeuchi and Amari 1979).

The above analysis establishes that the homogeneous equilibrium solution $u_0(x_2|x) = U(x_2 - x)$ undergoes a pattern forming instability as g_2 changes from a negative to a positive value induced, for example, by a reduction in the background inhibition $\hat{c}I_0^2$ or by an increase in the spread σ of the Gaussian inputs, see Eq. 2.18. Such an instability spontaneously breaks continuous translation symmetry, leading to the partitioning of the topographic map into discretized blocks (Takeuchi and Amari 1979). This is illustrated schematically in Fig. 3. The fact that the resulting pattern has a block–like structure can be understood from the observation that the dominant excited mode satisfies $p_+(x, \tau) = -p_-(x, \tau)$ and hence $\Delta_+(x, \tau) = \Delta_-(x, \tau)$. Thus the instability generates a leftward or rightward shift in an activity bump, depending on the location of the center x of its associated receptive field (see Fig. 2). It has been suggested that the blocks could be a precursor for the columnar microstructure of cortex (Takeuchi and Amari 1979; Amari 1989). As we mentioned in the introduction, cortical columns tend to be associated with a variety of stimulus features such as ocular dominance and orientation, which form spatially distributed feature maps that are superimposed upon the underlying topographic map (Swindale 1996). In the following sections we extend the stability analysis of one–dimensional topographic maps in order to investigate how such features could also emerge through the spontaneous symmetry breaking of self–organizing neural fields. First, in Sect. 3 we consider a one–dimensional network consisting of separate left and right eye afferents from the LGN. This introduces an additional Z_2 symmetry that can be spontaneously broken, resulting in the spatial segregation of eye specific activity bumps consistent with the emergence of ocular dominance columns. Second, in Sect. 4 we consider an isotropic two–dimensional network whose rotational symmetry can be spontaneously broken, leading to the formation of elongated activity bumps consistent with the emergence of orientation preference columns.

One final comment regarding Amari's model of topographic map formation is in order before proceeding with our analysis. This concerns the inclusion of feedforward inhibitory synapses that can also undergo Hebbian learning. Such inhibition is necessary in order to stabilize the smooth topographic map. However, as far as we are aware, there is no conclusive experimental support for the existence of Hebbian–like inhibitory synapses. On the other hand, most developmental models involving the Hebbian–like modification of excitatory synapses require additional constraints to ensure that an appropriate form of competition between synapses occurs and that a stable distribution of synaptic weights is generated (Miller and MacKay 1994). The constraints typically limit the sum of synaptic strengths received by a cell, or the mean activity of the cell. Although the constraints are not usually biophysically realistic, they are motivated by the idea that there exists some form of global intracellular signal controlling the synaptic weights. The modifiable inhibitory synaspses in Amari's model play an analogous role to these constraints. For example, in the binocular extension of Amari's model (see Sect. 3), feedforward inhibition ensures that the topographic map is stable (unstable) with respect to perturbations that are symmetric (anti-symmetric) under the exchange of left/right eye inputs. This should be compared with the use of subtractive normalization in correlation–based Hebbian models (Miller and MacKay 1994).

3 Spontaneous symmetry breaking in a binocular one dimensional network

Our first extension of the theory of self–organizing neural fields is to consider a one–dimensional network with distinct

left and right eye afferents. We derive conditions for the existence of a binocular topographic map, in which the response to a stimulus is independent of whether it is presented to the left or right eye. The resulting homogeneous solution is thus symmetric with respect to a discrete $\mathbf{Z}_2$ left/right exchange symmetry. We then generalize the linear stability analysis presented in Sec. 2, and show how the binocular state can undergo a pattern forming instability that spontaneously breaks the underlying $\mathbf{Z}_2$ symmetry. This leads to the spatial segregation of eye specific activity bumps consistent with the emergence of ocular dominance columns.

3.1 Binocular equilibrium state

Consider a one–dimensional version of the network model shown in Fig. 1, in which there are separate afferents from the left and right eye denoted by $s_{\rm L}(x_2, x_1, \tau)$ and $s_{\rm R}(x_2, x_1, \tau)$, respectively. The total input to cortical neurons at x_2 now becomes

$$v(x_2|x,\gamma) = \int_{-\infty}^{\infty} \big[s_{\rm L}(x_2,x_1) I_{\rm L}(x_1|x,\gamma){\rm d}x_1 + s_{\rm R}(x_2,x_1) I_{\rm R}(x_1|x,\gamma){\rm d}x_1\big] - s_0(x_2), \tag{3.1}$$

where γ is an additional stimulus label that takes into account differences in the statistical correlations between same eye and opposite eye inputs. We have also set the inhibitory input $I_0 = 1$. For concreteness, we choose Gaussian inputs of the form

$$I_L(x_1|x,\gamma) = A(1+\gamma){\rm e}^{-(x-x_1)^2/2\sigma^2},$$
$$I_R(x_1|x,\gamma) = A(1-\gamma){\rm e}^{-(x-x_1)^2/2\sigma^2}, \tag{3.2}$$

where γ is taken to be a binary random variable with ${\rm Prob}(\gamma = \gamma_0) = {\rm Prob}(\gamma = -\gamma_0) = 1/2$ for some constant γ_0, $0 < \gamma_0 < 1$. As in Sect. 2, the center of the input x is generated from a uniform random distribution. The derivation of the neural field equations proceeds in a similar fashion to the previous case, except now the Hebbian learning rules involve ensemble averages with respect to the left/right eye label as well:

$$\eta\frac{\partial s_{\rm L}}{\partial\tau} = -s_{\rm L}(x_2,x_1,\tau) + c\left\langle H(u(x_2,\tau|x',\gamma'))I_{\rm L}(x_1|x',\gamma')\right\rangle, \tag{3.3}$$

$$\eta\frac{\partial s_{\rm R}}{\partial\tau} = -s_{\rm R}(x_2,x_1,\tau) + c\left\langle H(u(x_2,\tau|x',\gamma'))I_{\rm R}(x_1|x',\gamma')\right\rangle \tag{3.4}$$

and

$$\eta\frac{\partial s_0}{\partial\tau} = -s_0(x_2,\tau) + \hat{c}\left\langle H(u(x_2,\tau|x',\gamma'))\right\rangle, \tag{3.5}$$

where $\langle\,\rangle$ denotes the ensemble average over x', γ', and

$$u(x_2,\tau|x,\gamma) = \int_{-\infty}^{\infty} w(x_2-x_2')H(u(x_2',\tau|x,\gamma)){\rm d}x_2' + v(x_2,\tau|x,\gamma) - h. \tag{3.6}$$

The corresponding equation for the weighted input v is obtained by differentiating Eq. 3.1 with respect to τ and using Eqs. 3.3–3.5:

$$\eta\frac{\partial v}{\partial\tau} = -v(x_2,\tau|x,\gamma) + \sum_{\gamma'=\pm\gamma_0}\int_{-\infty}^{\infty} g(x,\gamma|x',\gamma')H(u(x_2,\tau|x',\gamma')){\rm d}x' \tag{3.7}$$

with

$$g(x,\gamma|x',\gamma') = \left[1+\gamma\gamma'\right]\bar{g}(x-x') - \hat{c} \tag{3.8}$$

and

$$\bar{g}(x) = 2c\sigma\sqrt{\pi}A^2{\rm e}^{-x^2/4\sigma^2}. \tag{3.9}$$

Consider an equilibrium solution of the form

$$u_0(x_2|x,\gamma) = U(x_2-x,\gamma),$$
$$v_0(x_2|x,\gamma) = V(x_2-x,\gamma). \tag{3.10}$$

The corresponding fixed point equations are

$$U(x_2-x,\gamma) = \int_{-\infty}^{\infty} w(x_2-x_2')H(U(x_2'-x,\gamma))\,{\rm d}x_2' - \hat{c}\sum_{\gamma'}\int_{-\infty}^{\infty} H(U(x_2-x',\gamma'))\,{\rm d}x' - h + \sum_{\gamma'}(1+\gamma\gamma')\int_{-\infty}^{\infty}\bar{g}(x-x') \times H(U(x_2-x',\gamma'))\,{\rm d}x' \tag{3.11}$$

and

$$V(x_2-x,\gamma) = \sum_{\gamma'}\int_{-\infty}^{\infty}\left[(1+\gamma\gamma')\bar{g}(x-x') - \bar{c}\right]H(U(x_2-x',\gamma'))\,{\rm d}x'.$$

Taking $U(x,\gamma)$ to be an activity bump of width $2a(\gamma)$ and setting $x_2 = x \pm a(\gamma)$ then gives

$$W(2a(\gamma)) + \sum_{\gamma'}(1+\gamma\gamma')\bar{G}(2a(\gamma')) - 2\hat{c}\sum_{\gamma'}a(\gamma') = h \tag{3.12}$$

for $\gamma = \pm\gamma_0$. Defining $a_\pm = a(\pm\gamma_0)$ we finally obtain the pair of implicit equations

$$W(2a_\pm) + \left[\bar{G}(2a_+) + \bar{G}(2a_-) - 2\hat{c}(a_+ + a_-)\right] \pm\gamma_0^2\left[\bar{G}(2a_+) - \bar{G}(2a_-)\right] = h, \tag{3.13}$$

where W and $\bar{G}$ are defined as in Eq. 2.22 with g replaced by $\bar{g}$. We define a homogeneous binocular state to be one for which $a_+ = a_- = a$ with a satisfying the reduced equation

$$W(2a) + 2G(2a) = h \tag{3.14}$$

and $G(2a) = \bar{G}(2a) - 2\hat{c}a$. The associated activity bump will be stable with respect to fluctuations on the fast time–scale t provided that $W'(2a) + 2G'(2a) < 0$.

$\mathbf{Z}_2$ *symmetry*

The one–dimensional neural field Eqs. 3.6 and 3.7 are not only equivariant with respect to the product group of translations $\mathcal{T}\times\mathcal{T}$ described at the end of Sect. 2.2, but also have an additional $\mathbf{Z}_2$ symmetry. The latter group has elements ξ_0, ξ_1 where ξ_0 is the identity element and $\xi_1\cdot\xi_1=\xi_0$:

$$\xi_0.(u(x_2|x,\gamma),v(x_2|x,\gamma))=(u(x_2|x,\gamma),v(x_2|x,\gamma)),$$

$$\xi_1.(u(x_2|x,\gamma),v(x_2|x,\gamma))=(u(x_2|x,-\gamma),v(x_2|x,-\gamma)).$$

The homogeneous equilibrium solution $u_0(x_2|x,\gamma)=U(x_2-x,\gamma)$, $v_0(x_2|x,\gamma)=V(x_2-x,\gamma)$ then explicitly breaks the symmetry group from $\mathcal{T}\times\mathcal{T}\times\mathbf{Z}_2\to\mathcal{T}\times\mathbf{Z}_2$. We will show below that the homogeneous equilibrium solution can undergo a pattern forming instability that spontaneously breaks the remaining $\mathcal{T}\times\mathbf{Z}_2$ symmetry.

3.2 Linear stability analysis

Following along similar lines to Sect. 2.3, we linearize Eqs. 3.6 and 3.7 about the homogeneous binocular state by considering perturbations of the form

$$u(x_2,\tau|x,\gamma)=U(x_2-x,\gamma)+p(x_2,\tau|x,\gamma),$$
$$v(x_2,\tau|x,\gamma)=V(x_2-x,\gamma)+q(x_2,\tau|x,\gamma) \tag{3.15}$$

Using the identity (2.27) and setting $\eta=1$, this yield the linear equations

$$\begin{aligned}\frac{\partial q}{\partial\tau}=&-q(x_2,\tau|x,\gamma)-\alpha^{-1}\hat{c}\\&\sum_{\gamma'}\big[p(x_2,\tau|x_2-a,\gamma')+p(x_2,\tau|x_2+a,\gamma')\big]\\&+\alpha^{-1}\sum_{\gamma'}(1+\gamma\gamma')\,[\bar{g}(x-x_2+a)p(x_2,\tau|x_2\\&-a,\gamma')+\bar{g}(x-x_2-a)p(x_2,\tau|x_2+a,\gamma')]\end{aligned} \tag{3.16}$$

and

$$\begin{aligned}p(x_2,\tau|x,\gamma)=&\,q(x_2,\tau|x,\gamma)\\&+\alpha\big[w(x_2-x+a)p(x-a,\tau|x,\gamma')\\&+w(x_2-x-a)p(x+a,\tau|x,\gamma')\big].\end{aligned} \tag{3.17}$$

Defining

$$p_\pm(x,\gamma,\tau)=p(x\pm a,\tau|x,\gamma),$$
$$q_\pm(x,\gamma,\tau)=q(x\pm a,\tau|x,\gamma), \tag{3.18}$$

and setting $x_2=x\pm a$ in Eq. 3.17 gives the pair of equations

$$\begin{aligned}p_+(x,\gamma,\tau)=&\,q_+(x,\gamma,\tau)+\alpha^{-1}[w(2a)p_-(x,\gamma,\tau)\\&+w(0)p_+(x,\gamma,\tau)],\end{aligned} \tag{3.19}$$

$$\begin{aligned}p_-(x,\gamma,\tau)=&\,q_-(x,\gamma,\tau)+\alpha^{-1}[w(0)p_-(x,\gamma,\tau)\\&+w(2a)p_+(x,\gamma,\tau)].\end{aligned} \tag{3.20}$$

Similarly, setting $x_2=x\pm a$ in Eq. (3.16) shows that

$$\begin{aligned}\frac{\partial q_+}{\partial\tau}=&-q_+(x,\gamma,\tau)-\alpha^{-1}\hat{c}\sum_{\gamma'}\big[p_+(x,\gamma',\tau)\\&+p_-(x+2a,\gamma',\tau)\big]\\&+\alpha^{-1}\sum_{\gamma'}(1+\gamma\gamma')\big[\bar{g}(0)p_+(x,\gamma',\tau)\\&+\bar{g}(2a)p_-(x+2a,\gamma',\tau)\big],\end{aligned} \tag{3.21}$$

$$\begin{aligned}\frac{\partial q_-}{\partial\tau}=&-q_-(x,\gamma,\tau)-\alpha^{-1}\hat{c}\sum_{\gamma'}\big[p_+(x,\gamma',\tau)\\&+p_-(x-2a,\gamma',\tau)\big]\\&+\alpha^{-1}\sum_{\gamma'}(1+\gamma\gamma')\,[\bar{g}(2a)\\&\times p_+(x-2a,\gamma',\tau)+\bar{g}(0)p_-(x,\gamma',\tau)\big].\end{aligned} \tag{3.22}$$

Eqs. 3.19–3.22 have eigensolutions of the form

$$p_\pm(x,\gamma,\tau)=\mathrm{e}^{\lambda\tau}\mathrm{e}^{ikx}P_\pm(k,\gamma),$$
$$q_\pm(x,\gamma,\tau)=\mathrm{e}^{\lambda\tau}\mathrm{e}^{ikx}Q_\pm(k,\gamma) \tag{3.23}$$

with the eigenvalue λ and eigenvectors $\mathbf{P}=(P_+,P_-)^T$, $\mathbf{Q}=(Q_+,Q_-)^T$ determined from the matrix equations

$$\begin{aligned}\lambda\mathbf{Q}(k,\gamma)=&-\mathbf{Q}(k,\gamma)-\alpha^{-1}\hat{c}\mathbf{L}(k)\sum_{\gamma'}\mathbf{P}(k,\gamma')\\&+\alpha^{-1}\bar{\mathbf{G}}(k)\sum_{\gamma'}(1+\gamma\gamma')\mathbf{P}(k,\gamma'),\end{aligned} \tag{3.24}$$

$$\mathbf{P}(k,\gamma)=\mathbf{Q}(k,\gamma)+\alpha^{-1}\mathbf{W}\mathbf{P}(k,\gamma). \tag{3.25}$$

The matrices $\mathbf{W}$ and $\bar{\mathbf{G}}(k)$ are defined as in Eq. 2.37 with g replaced by $\bar{g}$, and

$$\mathbf{L}(k)=\begin{pmatrix}1 & \mathrm{e}^{2i\kappa a}\\ \mathrm{e}^{-2i\kappa a} & 1\end{pmatrix}. \tag{3.26}$$

The above matrix equations can be diagonalized with respect to the discrete label γ by introducing the symmetric and antisymmetric fields

$$p_\pm^S(x,\tau)=p_\pm(x,\gamma_0,\tau)+p_\pm(x,-\gamma_0,\tau),$$
$$q_\pm^S(x,\tau)=q_\pm(x,\gamma_0,\tau)+q_\pm(x,-\gamma_0,\tau), \tag{3.27}$$

$$p_\pm^A(x,\tau)=p_\pm(x,\gamma_0,\tau)-p_\pm(x,-\gamma_0,\tau),$$
$$q_\pm^A(x,\tau)=q_\pm(x,\gamma_0,\tau)-q_\pm(x,-\gamma_0,\tau) \tag{3.28}$$

with associated vector coefficients $\mathbf{P}^{S,A}(k)=\mathbf{P}(k,\gamma_0)\pm\mathbf{P}(k,-\gamma_0)$ and $\mathbf{Q}^{S,A}(k)=\mathbf{Q}(k,\gamma_0)\pm\mathbf{Q}(k,-\gamma_0)$. We then find that

$$(\lambda+1)\mathbf{Q}^S(k)=2\alpha^{-1}\left[\bar{\mathbf{G}}(k)-\hat{c}\mathbf{L}(k)\right]\mathbf{P}^S(k), \tag{3.29}$$

$$(\lambda+1)\mathbf{Q}^A(k)=2\alpha^{-1}\gamma_0^2\bar{\mathbf{G}}(k)\mathbf{P}^A(k), \tag{3.30}$$

$$\mathbf{P}^{S,A}(k)=\mathbf{Q}^{S,A}(k)+\alpha^{-1}\mathbf{W}\mathbf{P}^{S,A}(k). \tag{3.31}$$

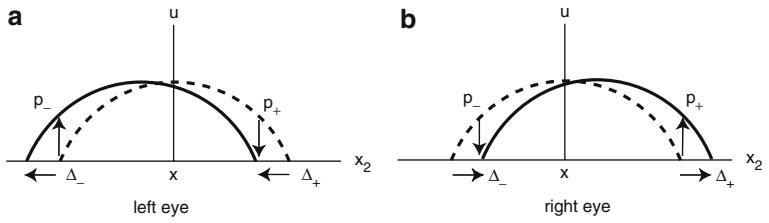

Fig. 4 Segregation of activity bumps generated by left and right eye dominated inputs, respectively. **a** A leftward shift due to perturbations $p_{\pm}(x, \gamma_0)$ at the boundaries of a homogeneous bump solution centered about $x_2 = x$ and having width $2a$. **b** Corresponding rightward shift due to perturbations $p_{\pm}(x, -\gamma_0) = -p_{\pm}(x, \gamma_0)$

Note that the basic structure of the eigensolutions reflects the fact that they form irreducible representations of the symmetry group $\mathcal{T} \times \mathbf{Z}_2$. In particular, the existence of symmetric and antisymmetric solutions under the exchange $\gamma \to -\gamma$ reflects the underlying $\mathbf{Z}_2$ symmetry.

The matrix equations for the symmetric and antisymmetric modes decouple and are identical in form to the monocular case considered in Sect. 2, see Eq. 2.36, with $\mathbf{G}(k)$ replaced by $\mathbf{G}^{S,A}(k)$:

$$\mathbf{G}^S(k) = 2\left[\bar{\mathbf{G}}(k) - \hat{c}\mathbf{L}(k)\right], \quad \mathbf{G}^A(k) = 2\gamma_0^2\bar{\mathbf{G}}(k) \tag{3.32}$$

The analysis of the corresponding symmetric and antisymmetric eigenvalues also proceeds along identical lines to Sect. 2. Therefore, assuming that $w_2 < 0$, the one–dimensional binocular topographic map will be stable with respect to the excitation of symmetric eigenmodes provided that

$$g_2^S \equiv 2(\bar{g}(2a) - \hat{c}) < 0, \tag{3.33}$$

and will be stable with respect to the excitation of antisymmetric eigenmodes provided that

$$g_2^A \equiv 2\gamma_0^2\bar{g}(2a) < 0. \tag{3.34}$$

A necessary condition for the formation of ocular dominance columns is that the binocular state should undergo an instability that breaks the underlying left/right $\mathbf{Z}_2$ symmetry. The latter will occur if the instability is associated with the growth of an antisymmetric mode, that is, if $g_2^A > 0$ and $g_2^S < 0$. This leads to the conditions

$$0 < \bar{g}(2a) < \hat{c}. \tag{3.35}$$

The first inequality is always satisfied, since Eq. 3.9 implies that $\bar{g}$ is a positive function. The dominant excited mode is given by (modulo an arbitrary phase)

$$\begin{aligned} p_{\pm}(x, \gamma_0) &= \pm\Gamma\cos(\pi x/2a), \\ p_{\pm}(x, -\gamma_0) &= \mp\Gamma\cos(\pi x/2a), \end{aligned} \tag{3.36}$$

which represents a state for which the center of the response to a left dominated input $(+\gamma_0)$ is shifted in the opposite direction to the center of the response to a right dominated input $(-\gamma_0)$, see Fig. 4. Moreover, the directions of the shifts periodically alternate across space according to the sign of $\cos(\pi x/2a)$. The form of the fastest growing mode suggests that ocular dominance columns will form, at least within the given linear approximation. Note that the joint development of a topographic map and ocular dominance columns has recently been demonstrated numerically using self–organizing neural fields with linear threshold nonlinearities (Woodbury et al. 2002).

The above analysis shows that a certain level of feedforward inhibition is needed in order to stabilize the topographic map with respect to perturbations that are symmetric under the exchange of left/right eye inputs. Indeed, if there were no inhibitory contribution ($\hat{c} = 0$), then the symmetric eigenmode would grow faster than the anti-symmetric mode (since $\gamma_0^2 < 1$) and no OD columns would form. As we commented at the end of Sect. 2, feedforward inhibition plays an analogous role to subtractive normalization in correlation–based Hebbian models (Miller and MacKay 1994). Although neither mechanism for stabilizing the symmetric eigenmode may be biophysically realistic, it is clear that some form of normalization is needed if the cortical development of ocular dominance columns occurs via Hebbian–like learning. Such a normalization will depend on properties of the inputs. In the case of the neural field model with feedforward inhibition, this is expressed by Eqs. 3.9 and 3.35, which show that the minimum level of inhibition $\bar{c}$ depends on the width σ and amplitiude A of the Gaussian inputs.

4 Spontaneous symmetry breaking in an isotropic two dimensional network

In this section, we extend the analysis presented in Sect. 2 to the case of two–dimensional topographic maps. We show how a dynamical instability of the topographic map can occur, in which there is a spontaneous breaking of continuous rotation symmetry, leading to the formation of elongated activity bumps; these are consistent with the emergence of orientation preference columns. Our analysis is based on a direct linearization of the neural field Eqs. 2.11 and 2.12 about a radially symmetric homogeneous solution.

4.1 Two–dimensional topographic map

Consider a radially symmetric, homogeneous equilibrium solution of Eq. 2.13 such that

$$U(r) > 0,\ 0 < r < a, \quad U(r) = 0,\ r = a,$$
$$U(r) < 0,\ r > a, \tag{4.1}$$

where a is the radius of the two–dimensional activity bump in cortex. Substituting into Eq. 2.13 gives

$$U(r) = F(a, r) - h, \tag{4.2}$$

where

$$F(a, r) = \int_0^{2\pi} \int_0^a f(|\mathbf{r} - \mathbf{r}'|) r' \, \mathrm{d}r' \, \mathrm{d}\theta \tag{4.3}$$

and we have defined $f(r) = w(r) + g(r)$. The radius of the bump is determined from the threshold condition $U(a) = 0$, which yields

$$F(a, a) = h. \tag{4.4}$$

As in the one–dimensional case, suppose that $w(r)$ is a Mexican hat function and the input $I(r)$ is a Gaussian so that $g(r)$ is a monotonically decreasing function of r, see Eq. 2.10. A unique stable bump solution then exists for a range of thresholds h. (The issue of stability will be addressed below). However, as has been pointed out elsewhere (Werner and Richter 2001), certain care has to be taken with regards the existence of two–dimensional bumps in the presence of short–range excitation and long–range inhibition. That is, in contrast to the one–dimensional case, the threshold condition may not be sufficient for existence, since the activity u could dip below threshold within the interior of the disc $r < a$. We will assume in the following that the stable bump solution is superthreshold for $r < a$.

It is possible to simplify the double integral in Eq. 4.3 using a Fourier transform, which for radially symmetric functions reduces to a Hankel transform (Folias and Bressloff 2004). To see this, consider the two-dimensional Fourier transform of the radially symmetric function f, expressed in polar coordinates,

$$\begin{aligned} f(r) &= \frac{1}{2\pi} \int_{\mathbf{R}^2} e^{i\mathbf{r}\cdot\mathbf{k}} \widetilde{\mathrm{f}}(\mathbf{k}) \mathrm{d}\mathbf{k} \\ &= \frac{1}{2\pi} \int_0^\infty \left(\int_0^{2\pi} e^{irk\cos(\theta-\phi)} \widetilde{\mathrm{f}}(k) \mathrm{d}\phi \right) k \mathrm{d}k, \end{aligned}$$

where $\widetilde{f}$ denotes the Fourier transform of f and $\mathbf{k} = (k, \varphi)$. Using the integral representation

$$\frac{1}{2\pi} \int_0^{2\pi} e^{irk\cos(\theta-\varphi)} \mathrm{d}\theta = J_0(rk),$$

where $J_\nu(z)$ is the Bessel function of the first kind, we express f in terms of its Hankel transform of order zero,

$$f(r) = \int_0^\infty \widetilde{f}(k) J_0(rk) k \, \mathrm{d}k \tag{4.5}$$

which, when substituted into Eq. 4.3, gives

$$\begin{aligned} &F(a, r) \\ &= \int_0^\infty \widetilde{f}(k) \left(\int_0^{2\pi} \int_0^a J_0(k|\mathbf{r} - \mathbf{r}'|) r' \mathrm{d}r' \mathrm{d}\, \theta' \right) k \, \mathrm{d}k. \end{aligned} \tag{4.6}$$

In polar coordinates,

$$\begin{aligned} &\int_0^{2\pi} \int_0^a J_0(k|\mathbf{r} - \mathbf{r}'|) r' \mathrm{d}r' \mathrm{d}\theta' \\ &= \int_0^{2\pi} \int_0^a J_0\left(k\sqrt{r^2 + r'^2 - 2rr'\cos(\theta - \theta')} \right) r' \mathrm{d}r' \mathrm{d}\theta' \end{aligned}$$

To separate variables, we use the addition theorem

$$\begin{aligned} &J_0\left(k\sqrt{r^2 + r'^2 - 2rr'\cos\theta'} \right) \\ &= \sum_{m=0}^\infty \epsilon_m J_m(kr) J_m(kr') \cos m\theta' \end{aligned}$$

where $\epsilon_0 = 1$ and $\epsilon_n = 2$ for $n \geq 1$. Since $\int_0^{2\pi} \cos\ m\theta' \, \mathrm{d}\theta' = 0$ for $m \geq 1$, it follows that

$$\begin{aligned} &\int_0^{2\pi} \int_0^a J_0(k|\mathbf{r} - \mathbf{r}'|) r' \mathrm{d}r' \mathrm{d}\theta' \\ &= 2\pi J_0(kr) \int_0^a J_0(kr') r' \mathrm{d}r' \\ &= \frac{2\pi a}{k} J_0(rk) J_1(ak). \end{aligned}$$

Hence, $F(a, r)$ has the integral representation

$$F(a, r) = 2\pi a \int_0^\infty \widetilde{f}(k) J_0(rk) J_1(ak) \mathrm{d}k. \tag{4.7}$$

Stability of two–dimensional bumps

The stability of a two–dimensional bump with respect to fluctuations on the fast time–scale t can be determined from linearizing the equation

$$\begin{aligned} \eta \frac{\partial U}{\partial t} &= -U(\mathbf{r}, t) \\ &\quad + \int_{\mathbf{R}^2} f(|\mathbf{r} - \mathbf{r}'|) H(U(\mathbf{r}', t)) \mathrm{d}\mathbf{r}' - h \end{aligned} \tag{4.8}$$

about the radially symmetric equilibrium solution. This particular problem has previously been studied in the restricted case of radially symmetric perturbations by Taylor (Taylor 1999). However, as recently shown by Folias and Bressloff (Folias and Bressloff 2004), it is also necessary to take into account non–radially symmetric perturbations in order to fully determine the stability of a two–dimensional activity bump. It is useful to review this latter analysis here before considering the stability of the associated topographic map. Consider the time-dependent perturbation $U(\mathbf{r}, t) = U(r) + p(\mathbf{r}, t)$ and expand to first order in p. This leads to the linearized equation

$$\begin{aligned} \frac{\partial p}{\partial t} &= -p(\mathbf{r}, t) \\ &\quad + \int_{\mathbb{R}^2} f(|\mathbf{r} - \mathbf{r}'|) H'(U(r') - \kappa) p(\mathbf{r}', t)) \, \mathrm{d}\mathbf{r}' \end{aligned} \tag{4.9}$$

which has solutions of the form $p(\mathbf{r},t)=p(\mathbf{r})e^{\lambda t}$. Introducing polar coordinates $\mathbf{r}=(r,\theta)$ and using the result

$$H'(U(r))=\delta(U(r))=\frac{\delta(r-a)}{|U'(a)|}$$

we obtain the eigenvalue equation

$$(\lambda+1)p(\mathbf{r})=\frac{a}{|U'(a)|}\int_0^{2\pi} f(|\mathbf{r}-\mathbf{a}'|)p(a,\theta')\,\mathrm{d}\theta', \tag{4.10}$$

where $\mathbf{a}'=(a,\theta')$.

If the eigenfunction p satisfies the condition

$$\int_0^{2\pi} f(|\mathbf{r}-\mathbf{a}'|)p(a,\theta')\,\mathrm{d}\theta'=0$$

for all $\mathbf{r}$ then the associated eigenvalue is $\lambda=-1$. This is part of the essential spectrum and does not cause instability. If p does not satisfy the above condition, then we must study the solutions of the integral equation

$$\mu p(r,\theta)=a\int_0^{2\pi}\mathcal{F}(a,r;\theta-\theta')p(a,\theta')\,\mathrm{d}\theta',$$

where $(\lambda+1)|U'(a)|=\mu$ and

$$\mathcal{F}(a,r;\phi))=f\left(\sqrt{r^2+a^2-2ra\cos\,\phi}\right).$$

It follows that $p(r,\theta)$ is determined completely by the restriction $p(a,\theta)$. Hence we need only consider $r=a$, yielding the integral equation

$$\mu p(a,\theta)=a\int_0^{2\pi}\mathcal{F}(a,a;\phi)p(a,\theta-\phi)\,\mathrm{d}\phi. \tag{4.11}$$

The solutions of this equation are exponential functions $e^{in\theta}$ where $n\in\mathbb{Z}$. Thus the integral operator with kernel $\mathcal{F}$ has a discrete spectrum given by

$$\begin{aligned}\mu_n&=a\int_0^{2\pi}\mathcal{F}(a,a;\phi)\mathrm{e}^{-in\phi}\mathrm{d}\phi\\&=a\int_0^{2\pi}f\left(\sqrt{a^2+a^2-2a^2\cos\phi}\right)\mathrm{e}^{-in\phi}\mathrm{d}\phi\\&=a\int_0^{2\pi}f\,(2a\sin\,(\phi/2))\,\mathrm{e}^{-in\phi}\mathrm{d}\phi\end{aligned}$$

(after rescaling ϕ). Note that μ_n is real since

$$\mathrm{Im}\{\mu_n(a)\}=-a\int_0^{2\pi}f(2a\sin(\phi/2))\sin(n\phi)\mathrm{d}\phi=0,$$

i.e. the integrand is odd-symmetric about π. Hence,

$$\begin{aligned}\mu_n(a)&=\mathrm{Re}\{\mu_n(a)\}\\&=a\int_0^{2\pi}f(2a\sin(\phi/2))\cos(n\phi)\mathrm{d}\phi\end{aligned} \tag{4.12}$$

with the integrand even-symmetric about π.

We conclude from the above analysis that an activity bump of radius a (assuming that it exists) will be stable provided that $\mu_n(a)\le|U'(a)|$ for all $n\in\mathbf{Z}$. This ensures that the corresponding eigenvalues are non–negative, $\lambda_n=-1+|U'(a)|^{-1}\mu_n(a)\le 0$ for all $n\in\mathbf{Z}$. Differentiating Eq. 4.3 with respect to r shows that

$$\begin{aligned}U'(a)&=\frac{\partial}{\partial r}F(a,r)\bigg|_{r=a}\\&=\int_0^{2\pi}\int_0^{a}\frac{f'\left(\sqrt{a^2+r'^2-2r'a\cos\,\phi}\right)}{\sqrt{a^2+r'^2-2r'a\cos\,\phi}}\\&\qquad\times(a-r'\cos\,\phi)r'\mathrm{d}r'\mathrm{d}\phi\\&=\int_0^{2\pi}\int_0^{a}\left[-\cos\phi\frac{\partial f}{\partial r'}+\frac{\sin\phi}{r'}\frac{\partial f}{\partial\phi}\right]_{r=a}r'\mathrm{d}r'\mathrm{d}\phi\\&=-a\int_0^{2\pi}f\left(\sqrt{2a^2-2a^2\cos\,\phi}\right)\cos\phi\,\mathrm{d}\phi\\&=-\mu_1(a).\end{aligned}$$

The final step in the above derivation involves integrating–by parts the term $-r'\cos\phi\partial f/\partial r'$ with respect to r' and the term $\sin\phi\partial f/\partial\phi$ with respect to ϕ. It follows that $\lambda_1=-1+|U'(a)|^{-1}\mu_1(a)=0$. The existence of a zero eigenvalue reflects the underlying translation symmetry of the system, which implies that the activity bump is marginally stable with respect to uniform shifts in space (see also Fig. 5 below). It follows that the bump will be stable if the zero eigenvalue is simple and all other eigenvalues are negative, that is, $\mu_n(a)<|U'(a)|$ for all $n\neq 1$. From Eqs. 4.2 and 4.3 we have

$$\begin{aligned}\mu_0(a)-|U'(a)|&=\frac{\partial}{\partial a}F(a,r)\bigg|_{r=a}+\frac{\partial}{\partial r}F(a,r)\bigg|_{r=a}\\&=\frac{\mathrm{d}}{\mathrm{d}a}F(a,a).\end{aligned} \tag{4.13}$$

Hence, a necessary condition for stability is $\mathrm{d}F(a,a)/\,\mathrm{d}a<0$, which was previously derived by Taylor (Taylor 1999) by considering only radially symmetric perturbations. However, our analysis shows that when one takes into account the full range of perturbations, this stability condition is not sufficient, since it does not ensure that $\mu_n(a)<|U'(a)|$ for all $n\neq 1$. We will assume in the following that on a fast time–scale (fixed weights), a given activity bump is stable with respect to both radially symmetric and non radially symmetric perturbations.

Euclidean symmetry

The two–dimensional isotropic and homogeneous neural field Eqs. 2.11 and 2.12 are equivariant with respect to the product Euclidean group $\mathbf{E}(2)\times\mathbf{E}(2)$ acting on the space $\mathbf{R}^2\times\mathbf{R}^2$ according to

$$\begin{aligned}&T_{\mathbf{s},\mathbf{s}'}\cdot(\mathbf{r}_2,\mathbf{r})=(\mathbf{r}_2+\mathbf{s},\mathbf{r}+\mathbf{s}'),\\&R_{\xi,\xi'}\cdot(\mathbf{r}_2,\mathbf{r})=(R_\xi\mathbf{r}_2,R_{\xi'}\mathbf{r}),\\&R_{\kappa,\kappa'}\cdot(\mathbf{r}_2,\mathbf{r})=(R_\kappa\mathbf{r}_2,R_{\kappa'}\mathbf{r}),\end{aligned} \tag{4.14}$$

where $R_\xi\mathbf{r}$ denotes the planar rotation of $\mathbf{r}$ through an angle ξ, and $R_\kappa=R_\pm$ with $R_\pm(x,y)=(x,\pm y)$. The corresponding group action on the neural fields u, v is

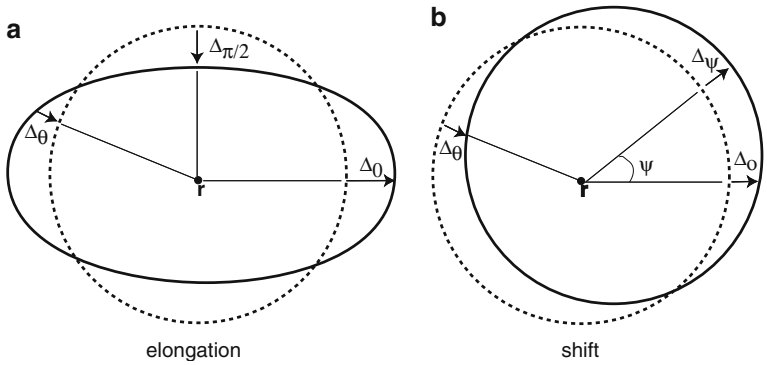

Fig. 5 Perturbations $\Delta_\theta(\mathbf{r})$ of the boundary of a radially symmetric bump solution (it dashed circle) centered about $\mathbf{r}_2 = \mathbf{r}$. **a** Elongation of the bump in the horizontal direction. **b** Shift of bump in ψ direction

$$T_{\mathbf{s},\mathbf{s}'}(u(\mathbf{r}_2|\mathbf{r}), v(\mathbf{r}_2|\mathbf{r})) = (u(\mathbf{r}_2 - \mathbf{s}|\mathbf{r} - \mathbf{s}'), v(\mathbf{r}_2 - \mathbf{s}|\mathbf{r} - \mathbf{s}')),$$

$$R_{\xi,\xi'}(u(\mathbf{r}_2|\mathbf{r}), v(\mathbf{r}_2|\mathbf{r})) = (u(R_{-\xi}\mathbf{r}_2|R_{-\xi'}\mathbf{r}), v(R_{-\xi}\mathbf{r}_2|R_{-\xi'}\mathbf{r})),$$

$$R_{\kappa,\kappa'}(u(\mathbf{r}_2|\mathbf{r}), v(\mathbf{r}_2|\mathbf{r})) = (u(R_\kappa\mathbf{r}_2|R_{\kappa'}\mathbf{r}), v(R_\kappa\mathbf{r}_2|R_{\kappa'}\mathbf{r})).$$

Equivariance means that if (u, v) is a solution of the neural field equations then so is $\eta \cdot (u, v)$ for all $\eta \in \mathbf{E}(2) \times \mathbf{E}(2)$. In other words, the two–dimensional network has both translation and rotation/reflection symmetries. An isotropic and homogeneous equilibrium solution of the form $u_0(\mathbf{r}_2|\mathbf{r}) = U(|\mathbf{r}_2 - \mathbf{r}|)$, $v_0(\mathbf{r}_2|\mathbf{r}) = V(|\mathbf{r}_2 - \mathbf{r}|)$ then explicitly breaks the symmetry group from $\mathbf{E}(2) \times \mathbf{E}(2) \to \mathbf{E}(2)$ with $\mathbf{E}(2)$ having the group elements $T_s = T_{s,s}$, $R_\xi = R_{\xi,\xi}$ and $R_\kappa = R_{\kappa,\kappa}$. We will show below that the homogeneous equilibrium solution can undergo a pattern forming instability that spontaneously breaks the remaining Euclidean symmetry.

4.2 Linear stability analysis

Following along analogous lines to Sect. 2.3, we investigate the stability of the two–dimensional topographic map by linearizing equations 2.11 and 2.12 about the homogeneous radially symmetric solution given by Eqs. 2.13, 2.14 and 4.1. That is, introducing the perturbations

$$u(\mathbf{r}_2, \tau|\mathbf{r}) = U(|\mathbf{r}_2 - \mathbf{r}|) + p(\mathbf{r}_2, \tau|\mathbf{r}),$$
$$v(\mathbf{r}_2, \tau|\mathbf{r}) = V(|\mathbf{r}_2 - \mathbf{r}|) + q(\mathbf{r}_2, \tau|\mathbf{r})$$

and expanding to first order in p, q leads to the linear equations (on setting $\eta = 1$)

$$\frac{\partial q}{\partial \tau} = -q(\mathbf{r}_2, \tau|\mathbf{r}) + \int_{R^2} g(|\mathbf{r} - \mathbf{r}'|)H'(U(|\mathbf{r}_2 - \mathbf{r}'|))p(\mathbf{r}_2, \tau|\mathbf{r}')\mathrm{d}\mathbf{r}',$$

and

$$p(\mathbf{r}_2, \tau|\mathbf{r}) = q(\mathbf{r}_2, \tau|\mathbf{r}) + \int_{R^2} w(|\mathbf{r}_2 - \mathbf{r}_2'|)H'(U(|\mathbf{r}_2' - \mathbf{r}|)) \times p(\mathbf{r}_2', \tau|\mathbf{r})\mathrm{d}\mathbf{r}_2'$$

Using the identity

$$H'(U(|\mathbf{r}_2' - \mathbf{r}|)) = \alpha^{-1}\delta(|\mathbf{r}_2' - \mathbf{r}| - a)$$

with $\alpha = |U'(a)|$, we can reduce the above linear equations to the form

$$\frac{\partial q}{\partial \tau} = -q(\mathbf{r}_2, \tau|\mathbf{r}) + \frac{a}{\alpha}\int_0^{2\pi} g(|\mathbf{r} - \mathbf{r}_2 + a\mathbf{e}_\phi|)\, p(\mathbf{r}_2, \tau|\mathbf{r}_2 - a\mathbf{e}_\phi)\mathrm{d}\phi \tag{4.15}$$

and

$$p(\mathbf{r}_2, \tau|\mathbf{r}) = q(\mathbf{r}_2, \tau|\mathbf{r}) + \frac{a}{\alpha}\int_0^{2\pi} w(|\mathbf{r}_2 - \mathbf{r} - a\mathbf{e}_\phi|) \times p(\mathbf{r} + a\mathbf{e}_\phi, \tau|\mathbf{r})\mathrm{d}\phi, \tag{4.16}$$

with $\mathbf{e}_\theta = (\cos\theta, \sin\theta)$. Defining

$$p_\theta(\mathbf{r}, \tau) = p(\mathbf{r} + a\mathbf{e}_\theta, \tau|\mathbf{r}),$$
$$q_\theta(\mathbf{r}, \tau) = q(\mathbf{r} + a\mathbf{e}_\theta, \tau|\mathbf{r}) \tag{4.17}$$

and setting $\mathbf{r}_2 = \mathbf{r} + a\mathbf{e}_\theta$ in Eq. 4.16, we find that

$$p_\theta(\mathbf{r}, \tau) = q_\theta(\mathbf{r}, \tau) + \frac{a}{\alpha}\int_0^{2\pi} w(a|\mathbf{e}_\theta - \mathbf{e}_\phi|)p_\phi(\mathbf{r}, \tau)\mathrm{d}\phi = q_\theta(\mathbf{r}, \tau) + \frac{a}{\alpha}\int_0^{2\pi} w(2a\sin([\theta - \phi]/2)) \times p_\phi(\mathbf{r}, \tau)\mathrm{d}\phi. \tag{4.18}$$

We have used the identity

$$|\mathbf{e}_\theta - \mathbf{e}_\phi|^2 = 2[1 - \cos(\theta - \phi)] = 4\sin^2([\theta - \phi]/2).$$

Similarly, setting $\mathbf{r}_2 = \mathbf{r} + a\mathbf{e}_\theta$ in Eq. 4.15 gives

$$\frac{\partial q_\theta}{\partial \tau} = -q_\theta(\mathbf{r},\tau) + \frac{a}{\alpha}\int_0^{2\pi} g(2a\sin([\theta-\phi]/2)) \times p_\phi(\mathbf{r}+a(\mathbf{e}_\theta-\mathbf{e}_\phi),\tau)\mathrm{d}\phi. \quad (4.19)$$

Equations 4.18 and 4.19 are the two–dimensional extensions of Eqs. 2.31–2.34.

As in the one–dimensional case (see Fig. 2), the neural field perturbations $p_\theta(\mathbf{r},\tau)$ can be related to perturbations of the boundary of the activity bump, which in the case of a radially symmetric bump is in the form of a circle. Let us write the perturbed threshold condition in the form $u(\mathbf{r}_2,\tau|\mathbf{r}) = 0$ at $\mathbf{r}_2 = \mathbf{r} + (a+\Delta_\theta(\mathbf{r},\tau))\mathbf{e}_\theta$. This yields

$$\begin{aligned} 0 &= U(a+\Delta_\theta(\mathbf{r},\tau)) + p(\mathbf{r}+(a+\Delta_\theta(\mathbf{r},\tau))\mathbf{e}_\theta,\tau|\mathbf{r}) \\ &= U(a) + U'(a)\Delta_\theta(\mathbf{r},\tau) \\ &\quad + p(\mathbf{r}+a\mathbf{e}_\theta,\tau|\mathbf{r}) + \mathcal{O}(\Delta^2) \end{aligned}$$

which implies that

$$\Delta_\theta(\mathbf{r},\tau) = \alpha^{-1} p_\theta(\mathbf{r},\tau)$$

since $U(a) = 0$ and $U'(a) = -\alpha$. Here $\Delta_\theta(\mathbf{r})$ represents the radial shift in the θ direction of the circular bump boundary centered at $\mathbf{r}$. In the special case of a small uniform shift δr of the bump in the ψ–direction, one can use a simple geometric argument to show that $\Delta_\theta = \delta r\cos(\theta-\psi) + \mathcal{O}(\delta r^2)$. Thus the generators of a uniform shift are the perturbations $\mathrm{e}^{\pm i\theta}$. Similarly, the perturbations $\mathrm{e}^{\pm 2i\theta}$ generate an elongation of the bump, whereas a θ–independent perturbation generates an expansion or contraction of the bump (see Fig. 5).

Eqs. 4.18 and 4.19 have solutions of the form

$$\begin{aligned} p_\theta(\mathbf{r},\tau) &= \mathrm{e}^{\lambda\tau}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}P_\theta(\mathbf{k}), \\ q_\theta(\mathbf{r},\tau) &= \mathrm{e}^{\lambda\tau}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}Q_\theta(\mathbf{k}) \end{aligned} \quad (4.20)$$

with

$$P_\theta(\mathbf{k}) = Q_\theta(\mathbf{k}) + \frac{a}{\alpha}\int_0^{2\pi} w\left(2a\sin\left(\frac{[\theta-\phi]}{2}\right)\right)P_\phi(\mathbf{k})\mathrm{d}\phi \quad (4.21)$$

and

$$\lambda Q_\theta(\mathbf{k}) = -Q_\theta(\mathbf{k}) + \frac{a}{\alpha}\int_0^{2\pi} g(2a\sin([\theta-\phi]/2))\mathrm{e}^{ia\mathbf{k}\cdot(\mathbf{e}_\theta-\mathbf{e}_\phi)} \times P_\phi(\mathbf{k})\mathrm{d}\phi. \quad (4.22)$$

Equations 4.21 and 4.22 can be analyzed further by introducing the Fourier series

$$\begin{aligned} P_\theta(\mathbf{k}) &= \sum_{n\in\mathbf{Z}} P_n(\mathbf{k})\mathrm{e}^{in\theta}, \\ Q_\theta(\mathbf{k}) &= \sum_{n\in\mathbf{Z}} Q_n(\mathbf{k})\mathrm{e}^{in\theta}. \end{aligned} \quad (4.23)$$

This leads to the discrete set of equations

$$(\lambda+1)Q_n(\mathbf{k}) = \alpha^{-1}\sum_{n\in\mathbf{Z}} G_{nn'}(\mathbf{k})P_{n'}(\mathbf{k}), \quad (4.24)$$

$$P_n(\mathbf{k}) = \frac{Q_n(\mathbf{k})}{1-\alpha^{-1}\nu_n} \quad (4.25)$$

with

$$G_{nn'}(\mathbf{k}) = a\int_0^{2\pi}\mathrm{e}^{-in\theta}\int_0^{2\pi}\mathrm{e}^{in'\phi}g\left(2a\sin\left(\theta-\phi/2\right)\right) \times\mathrm{e}^{ia\mathbf{k}\cdot(\mathbf{e}_\theta-\mathbf{e}_\phi)}\frac{\mathrm{d}\phi\mathrm{d}\theta}{2\pi} \quad (4.26)$$

and

$$\nu_n = a\int_0^{2\pi} w\left(2a\sin\left(\theta/2\right)\right)\cos(n\theta)\mathrm{d}\theta. \quad (4.27)$$

Note from Eqs 4.12 and 4.27 that

$$\nu_n = \mu_n - a\int_0^{2\pi} g\left(2a\sin\left(\theta/2\right)\right)\cos(n\theta)d\theta. \quad (4.28)$$

Moreover, the requirement that two–dimensional bumps are stable on the fast time–scale means that $\mu_n < \alpha$ for all $n \neq 1$ and $\mu_1 = \alpha$.

Calculation of eigenmodes: wide inputs

Determining the stability of the two–dimensional topographic map is reduced to the problem of finding the eigenvalues of the infinite–dimensional matrix $G_{nn'}(\mathbf{k})$ for $n, n' \in \mathbf{Z}$. It is not possible to do this analytically for general input kernel g. However, an explicit solution can be obtained in the limiting case of wide Gaussian inputs such that $\sigma \gg a$ in Eq. 2.10. We can then carry out a perturbation expansion in a^2/σ^2 by writing

$$\begin{aligned} g(2a\sin(\theta/2)) &= c\sigma^2\pi A^2\mathrm{e}^{-a^2\sin^2(\theta/2)/\sigma^2} - \hat{c}I_0^2 \\ &\approx g(0) - \bar{g}(0)a^2/2\sigma^2(1-\cos(\theta)) \\ &\quad + \mathcal{O}\left(a^4/\sigma^4\right), \end{aligned} \quad (4.29)$$

where $g(0) = c\sigma^2\pi A^2 - \hat{c}I_0^2$ and $\bar{g}(0) = c\sigma^2\pi A^2$. Keeping only lowest order terms we find that

$$G_{nn'}(\mathbf{k}) \approx ag(0)\int_0^{2\pi}\mathrm{e}^{-in\theta}\int_0^{2\pi}\mathrm{e}^{in'\phi} \times\mathrm{e}^{iak(\cos(\theta-\varphi)-\cos(\phi-\varphi))}\frac{\mathrm{d}\phi\mathrm{d}\theta}{2\pi}, \quad (4.30)$$

where $\mathbf{k} = (k,\varphi)$ in polar coordinates. The integrals over ϕ and θ may now be evaluated using the following Bessel function expansion:

$$\mathrm{e}^{ika\cos(\theta-\varphi)} = \sum_{m\in\mathbf{Z}}(-i)^m J_m(ka)\mathrm{e}^{im(\theta-\varphi)} \quad (4.31)$$

with $J_{-m} = J_m$. This gives

$$G_{nn'}(\mathbf{k}) \approx ag(0)\int_0^{2\pi} e^{-in\theta}\int_0^{2\pi} e^{in'\phi}\sum_{m\in\mathbf{Z}}(-i)^m J_m(ka)e^{im(\theta-\varphi)} \times \sum_{m'\in\mathbf{Z}}(i)^{m'} J_{m'}(ka)e^{-im'(\phi-\varphi)}\frac{d\phi d\theta}{2\pi} = 2\pi ag(0)(-i)^n(i)^{n'} J_n(ka)J_{n'}(ka)e^{i(n'-n)\varphi}. \tag{4.32}$$

Similarly, substituting Eq. 4.29 into 4.28 gives to lowest order

$$\nu_0 \approx \mu_0 - 2\pi ag(0)$$
$$\nu_1 \approx \mu_1 - \pi a\bar{g}(0)\frac{a^2}{2\sigma^2}$$
$$\nu_n \approx \mu_n, \quad n > 1 \tag{4.33}$$

Combining Eqs. 4.24, 4.25 and 4.32 yields a vector equation of the form

$$\mathbf{b}^*(\mathbf{k})(\mathbf{b}(\mathbf{k})\cdot\widehat{\mathbf{P}}(\mathbf{k})) = (1+\lambda)\widehat{\mathbf{P}}(\mathbf{k}), \tag{4.34}$$

where * denotes complex conjugate, and

$$\widehat{P}_n(\mathbf{k}) = \sqrt{\alpha-\nu_n}P_n(\mathbf{k}),$$
$$b_n(\mathbf{k}) = \sqrt{\frac{2\pi ag(0)}{\alpha-\nu_n}}(i)^n J_n(ka)e^{in\varphi}. \tag{4.35}$$

There are two classes of solution to Eq. 4.34. If $\mathbf{b}\cdot\widehat{\mathbf{P}} = 0$ then $\lambda = -1$ and the topographic map is stable with respect to excitation of the corresponding eigenmodes. On the other hand, if $\mathbf{b}\cdot\widehat{\mathbf{P}} \neq 0$ then $\widehat{\mathbf{P}} = \mathbf{b}^*$ (up to a constant multiplicative factor). Substituting into the Fourier series (4.23), the resulting eigenmode is of the form $P_\theta(\mathbf{k}) = P(k,\theta-\varphi)$ with $\mathbf{k} = (k,\varphi)$,

$$P(k,\theta) = \Gamma\left[\frac{J_0(ka)}{\alpha-\nu_0} + 2\sum_{n\geq 1}(-1)^n\frac{J_{2n}(ka)}{\alpha-\nu_{2n}}\cos(2n\theta) - 2\sum_{n\geq 1}(-1)^n\frac{J_{2n-1}(ka)}{\alpha-\nu_{2n-1}}\sin((2n-1)\theta)\right], \tag{4.36}$$

where Γ is an arbitrary amplitude. The corresponding eigenvalue is $\lambda = \lambda(k)$ with

$$\lambda(k) = -1 + |\mathbf{b}|^2 = -1 + \sum_{n\in\mathbf{Z}}\frac{2\pi ag(0)}{\alpha-\nu_n}J_n(ka)^2. \tag{4.37}$$

The Bessel functions J_n for $n = 0, 1, 2$ are plotted in Fig. 6.

For the sake of illustration, suppose that $\nu_n < \alpha$ for all $n \in \mathbf{Z}$. This is plausible given Eq. 4.33 and the conditions on μ_n. Equation (4.37) implies that if $g(0) < 0$ then the topographic map is stable since $\lambda(k) < 0$ for all k. On the other hand, if $g(0) > 0$ such that $\lambda(k_c) = \max_k \lambda(k) > 0$ then the topographic map is unstable and the fastest growing eigenmodes have the critical wavenumber k_c. Recall from Sect. 4.1 that $\alpha = \mu_1$. It then follows from Eq. 4.33 that $\nu_1 \approx \alpha$ and the dominant contribution to the sum in Eq. 4.37 will arise from the $n = 1$ term, at least distance from the zeros of $J_1(ka)$. Hence, k_c is approximately given by the point at which the first order Bessel function attains its global maximum, that is, $|J_1(ak_c)| = \max_k |J_1(ak)|$. Fig. 6 shows that $k_c \approx 3/a$. One of the major differences between the linear theory of one–dimensional and two dimensional topographic maps, is that in the latter case the eigenvalues $\lambda(k)$, $k \neq 0$, have an infinite degeneracy that reflects the additional rotation symmetry of the system. That is, all eigenmodes $P_\theta(\mathbf{k})$ with $|\mathbf{k}| = k$ have the same eigenvalue. It follows that the pattern forming instability will be dominated by some linear combination of eigenmodes lying on the critical circle $|\mathbf{k}| = k_c$:

$$p_\theta(\mathbf{r}) = \sum_{i=1}^{N}\left(z_i e^{i\mathbf{k}_i\cdot\mathbf{r}} + z_i^* e^{i\mathbf{k}_i\cdot\mathbf{r}}\right)P(k_c,\theta-\varphi_i), \tag{4.38}$$

where $\mathbf{k}_i = (k_c,\varphi_i)$ and z_i is a complex amplitude. Suppose that each eigenmode can be approximated by the first three terms of Eq. 4.36 so that

$$P(k_c,\theta) \approx \Gamma\left[\frac{J_0(k_c a)}{\alpha-\nu_0} + \frac{2J_1(k_c a)}{\alpha-\nu_1}\sin(\theta) - \frac{2J_2(k_c a)}{\alpha-\nu_2}\cos(2\theta)\right], \tag{4.39}$$

The first term generates an expansion of the bump, the second term generates a uniform shift of the bump and the third term generates an elongation of the bump (see Fig. 5). In general, we expect the eigenmode (Eq. 4.39) to be dominated by the first harmonic term $\sin(\theta)$, since $\nu_1 \approx \alpha$. However, if $\nu_2 \approx \alpha$ as well, then there could also be a significant contribution from the term $\cos(2\theta)$. Thus the spontaneous symmetry breaking mechanism has the potential for generating elongated receptive fields that are consistent with the formation of orientation columns. Moreover, since each eigenmode in the sum (Eq. 4.38) then represents an elongation in the direction φ_i or $\pi/2+\varphi_i$ (depending on the sign of its associated coefficient $z(\mathbf{r}) = z_i e^{i\mathbf{k}_i\cdot\mathbf{r}} + z_i^* e^{i\mathbf{k}_i\cdot\mathbf{r}}$), it follows that there is some complicated variation in the preferred orientation as $\mathbf{r}$ varies across the cortex. We note that the emergence of orientation selectivity in self–organizing neural fields has recently been demonstrated numerically (Fellenz and Taylor 2002). However, whether or not such a model can reproduce the detailed structure of orientation maps found experimentally remains to be established. For example, it might be necessary to develop a more detailed model that takes into account separate ON and OFF pathways as previously considered by Miller using correlation–based methods (Miller 1994).

Calculation of eigenmodes: narrow inputs

As in the one–dimensional case, if the excitatory inputs become sufficiently narrow then the topographic map is stable in the presence of feedforward inhibition. In the absence of such inhibition ($\hat{c} = 0$), it is possible to find an approximate

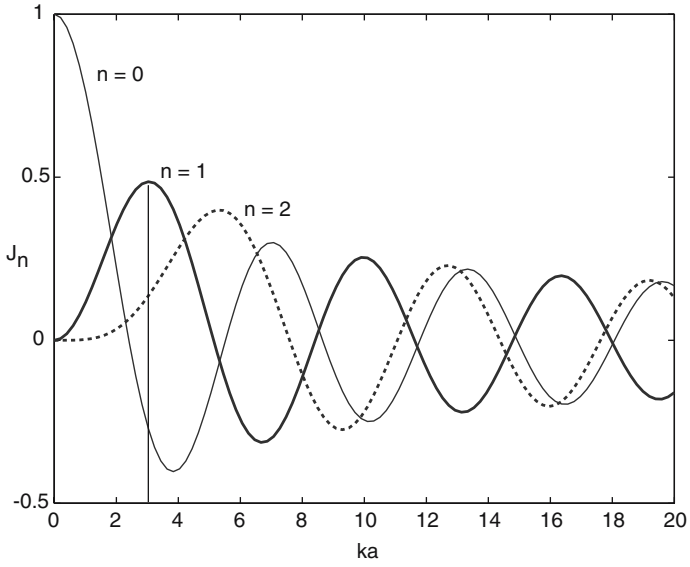

Fig. 6 Bessel functions J_n for $n = 0, 1, 2$

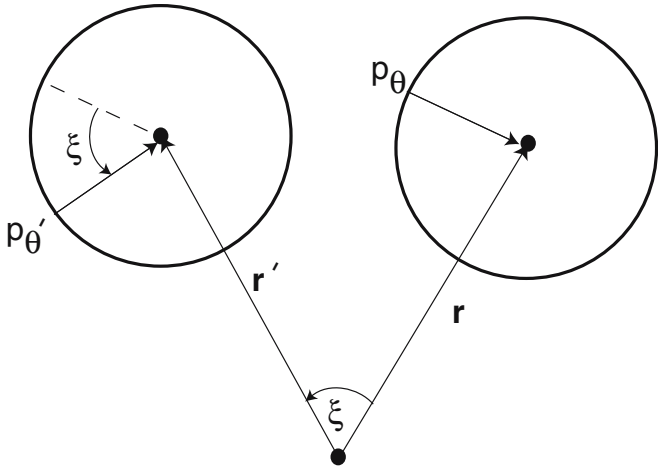

Fig. 7 Action of a rotation by ξ: $p_\theta(\mathbf{r}) \to p_{\theta'}(\mathbf{r}')$ where $(\mathbf{r}', \theta') = (R_\xi \mathbf{r}, \theta + \xi)$. Here $\mathbf{r}$ represents the position of the center of a two–dimensional bump and p_θ represents the perturbation of steady–state activity at a point θ on the boundary of the bump

solution for the eigenmodes in the limit of narrow inputs ($\sigma \ll a$) to determine the dominant eigenmode. For simplicity, we take the excitatory input kernel $\bar{g}$ to be a narrow step function rather than a Gaussian such that $\bar{g}(|\mathbf{r}|) = g_0$ if $|\mathbf{r}| < \sigma$ and zero otherwise. Under this approximation, $g(2a\sin([\theta - \phi]/2)) = g_0$ for $\phi \approx \theta \pm \sigma/a$ and is zero otherwise. Substitution into Eq. 4.26 shows that

$$
\begin{aligned}
G_{nn'}(\mathbf{k}) &\approx g_0 a \int_0^{2\pi} \mathrm{e}^{-in\theta} \int_{-\sigma/a}^{\sigma/a} \mathrm{e}^{in'(\theta+\psi)} \\
&\qquad \mathrm{e}^{iak[\cos(\theta-\varphi)-\cos(\theta+\psi-\varphi)]} \frac{\mathrm{d}\psi\,\mathrm{d}\theta}{2\pi} \\
&= g_0 a \int_0^{2\pi} \mathrm{e}^{-in\theta} \int_{-\sigma/a}^{\sigma/a} \mathrm{e}^{in'(\theta+\psi)} (1 \\
&\qquad + iak\sin(\theta+\varphi)\psi + \mathcal{O}(\psi^2)) \frac{\mathrm{d}\psi\,\mathrm{d}\theta}{2\pi} \\
&= 2g_0 a \frac{\sin(n\sigma/a)}{n} \delta_{n,n'} \\
&\qquad + g_0 a^2 k \left(\delta_{n',n-1}\mathrm{e}^{i\varphi} - \delta_{n',n+1}\mathrm{e}^{-i\varphi}\right) \\
&\qquad \times \frac{1}{in'}\left(\frac{\sigma\cos(n'\sigma/a)}{a} - \frac{\sin(n'\sigma/a)}{n'}\right) \\
&\qquad + \cdots \\
&= g_0 a \left[\frac{2\sigma}{a}\delta_{n,n'} + \mathcal{O}([\sigma/a]^3)\right]. \qquad (4.40)
\end{aligned}
$$

Note in particular that the $\mathcal{O}([\sigma/a]^2)$ term is zero. Substitution into Eq. 4.24 and using Eq. 4.25 implies that to lowest order in σ/a, the eigenmodes are $\mathbf{k}$–independent and of the form $\mathrm{e}^{in\theta}$ with corresponding eigenvalues

$$
\lambda_n = -1 + \frac{2g_0\sigma}{\alpha - \nu_n} \qquad (4.41)
$$

Given that $\nu_n = \mu_n + \mathcal{O}(\sigma/a)$ (see Eq. 4.28) and $\mu_1 = \alpha$, it follows that the dominant eigenmode is going to be $n = 1$, which represents a uniform shift in the topographic map. We conclude that even in the absence of feedforward inihibition, destabilization of the topographic map cannot generate elongated receptive fields nor an associated orientation map if the excitatory inputs are too narrow.

Euclidean shift–twist symmetry

The basic structure of the eigenmodes $p_\theta(\mathbf{k})$ can be understood from a more general group theoretic perspective by noting that the linear equations 4.18 and 4.19 are equivariant with respect to the so–called shift–twist action of the Euclidean group $\mathbf{E}(2)$ on the space $\mathbf{R}^2 \times S^1$ (Bressloff et al. 2001a,b):

$$\begin{aligned} T_{\mathbf{s}} \cdot (\mathbf{r}, \theta) &= (\mathbf{r} + \mathbf{s}, \theta), \\ R_\xi \cdot (\mathbf{r}, \theta) &= (R_\xi \mathbf{r}, \theta + \xi), \\ R_\kappa \cdot (\mathbf{r}, \theta) &= (R_\kappa \mathbf{r}, -\theta), \end{aligned} \tag{4.42}$$

where $R_\kappa(x, y) = (x, -y)$. The corresponding action on the fields $p_\theta(\mathbf{r})$ and $q_\theta(\mathbf{r})$ is

$$T_{\mathbf{s}} \cdot (p_\theta(\mathbf{r}), q_\theta(\mathbf{r})) = (p_\theta(\mathbf{r} - \mathbf{s}), q_\theta(\mathbf{r} - \mathbf{s})),$$

$$R_\xi \cdot (p_\theta(\mathbf{r}), q_\theta(\mathbf{r})) = (p_{\theta-\xi}(R_{-\xi}\mathbf{r}), q_{\theta-\xi}(R_{-\xi}\mathbf{r})),$$

$$R_\kappa \cdot (p_\theta(\mathbf{r}), q_\theta(\mathbf{r})) = (p_{-\theta}(R_\kappa \mathbf{r}), q_{-\theta}(R_\kappa \mathbf{r})).$$

It can be seen that the rotation operation comprises a translation or *shift* of the angle θ to $\theta + \xi$, together with a rotation or *twist* of the position vector $\mathbf{r}$ by the angle ξ. This is illustrated in Fig. 7. One of the consequences of the underlying Euclidean symmetry is that the associated eigenfunctions form irreducible representations of the shift–twist group action (Bressloff et al. 2001a,b). This explains why the eigenmodes $P_\theta(\mathbf{k})$ have the basic structure given by Eq. 4.36, with the angular variable θ coupled to the direction of the wavevector $\mathbf{k}$. Interestingly, there is growing evidence that there is a coupling between orientation and topography consistent with an underlying rotational shift–twist symmetry (Bosking et al. 1997; Lee et al. 2003), as highlighted in the discussion below.

5 Discussion

In this paper we have extended the theory of self–organizing neural fields in order to investigate from a mathematical perspective the possible joint emergence of topography and feature selectivity through spontaneous symmetry breaking. We first showed how a binocular one–dimensional topographic map can undergo a pattern forming instability that breaks the underlying $\mathbf{Z}_2$ symmetry between left and right eyes. This leads to the spatial segregation of eye specific activity bumps consistent with the emergence of ocular dominance columns. We then showed how a two–dimensional isotropic topographic map can undergo a pattern forming instability that breaks the underlying rotation symmetry. This leads to the formation of elongated activity bumps consistent with the emergence of orientation preference columns. A particularly interesting property of the latter symmetry breaking mechanism is that the linear equations describing the growth of the orientation columns exhibits a rotational shift–twist symmetry, in which there is a coupling between orientation and topography. A recent statistical analysis of orientation preference maps in primates indicates that there are correlations between the direction of the topographic axis joining pairs of columns with similar orientation preferences and their common orientation (Lee et al. 2003). Thus the orientation preference map exhibits a form of rotational shift–twist symmetry as predicted from our analysis of two–dimensional topographic maps. Numerical simulations of a feature–based dynamical spin mode has led to the suggestion that such a symmetry could help to stabilize the emerging orientation preference map with its associated set of pinwheels (Lee et al. 2003). As previously shown by Wolf and Geisel (Wolf and Geisel 1998), in the absence of such a coupling, the pinwheels typically annihilate in pairs. Hence, in order to maintain pinwheels, either development has to be stopped or one has to introduce inhomogeneities that trap the pinwheels. (Note that Thomas and Cowan (Thomas and Cowan 2004) have recently analyzed a spin model with a different form of rotational coupling between orientation and topography, and shown how dislocations in the topographic map can occur).

Another aspect of cortical structure that appears to exhibit shift–twist symmetry is the distribution of patchy horizontal connections found in superficial layers of cortex. Optical imaging combined with labeling techniques has established that these connections tend to link cells with similar feature preferences (Malach et al. 1993; Yoshioka et al. 1996). Moreover, in tree shrew and cat there is a pronounced anisotropy in the distribution of patchy connections, with iso–orientation patches preferentially connecting to neighboring patches in such a way as to form continuous contours along the topographic axis (Bosking et al. 1997). There is also a clear anisotropy in the patchy connections of owl (Sincich and Blasdel 2001) and macaque (Angelucci et al. 2002) monkeys. However, in these cases most of the anisotropy can be accounted for by the fact that V1 is expanded in the direction orthogonal to ocular dominance columns. It is possible that when this expansion is factored out, there remains a weak anisotropy correlated with orientation selectivity but this remains to be confirmed experimentally. Interestingly, the recently observed patchy feedback connections from extrastriate areas in macaque tend to be more strongly anisotropic (Angelucci et al. 2002); it is likely that the patchiness again signifies that feedback correlates cells with similar feature preferences (Shmuel et al. 1998). It has been shown elsewhere that the shift–twist symmetry of anisotropic horizontal connections has a nontrivial affect on the dynamics of neural activity in visual cortex (Bressloff et al. 2001a,b, 2002). It would be interesting to extend the analysis of this paper in order to determine how such connections self–organize through Hebbian learning. This will require treating both feedforward

and intracortical connections as adaptive (Bartsch and van Hemmen 2001), rather than keeping the latter fixed.

Acknowledgements This work was partially supported by NSF grant DMS 0515725.

References

Amari S-I (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. Biol Cybern 27:77–87

Amari S-I (1980) Topographic organization of nerve fields. Bull Math Biol 42:339–364

Amari S-I (1983) Field theory of self-organizing neural nets. IEEE Trans Man Cybern 13:741–748

Amari S-I (1989) Dynamical stability of formation of cortical maps. In: Arbib MA, Amari S-I (eds) Dynamical interactions in neural networks: models and data. Springer, Berlin

Angelucci A, Levitt JB, Walton EJS, Hupe JM, Bullier J and Lund JS (2002) Circuits for local and global signal integration in primary visual cortex. J Neurosci 22:8633–8646

Bartsch AP, van Hemmen JL (2001) Combined Hebbian development of geniculocortical and lateral connectivity in a model of primary visual cortex. Biol Cybern 84:41–55

Blasdel GG, Salama G (1986) Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. Nature 321:579–585

Bonhoeffer T, Grinvald A (1991) Orientation columns in cat are organized in pinwheel like patterns. Nature 364:166–146

Bosking WH, Zhang Y, Schofield B, Fitzpatrick D (1997) Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J Neurosci 17:2112–2127

Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener M (2001a) Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. Phil Trans Roy Soc Lond B 356:299–330

Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ (2001b) Scalar and pseudoscalar bifurcations: pattern formation on the visual cortex. Nonlinearity 14:739–775

Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener M (2002) What geometric visual hallucinations tell us about the visual cortex. Neural Comput 14:471–492

Catalano SM, Shatz CJ (1998) Activity-dependent cortical target selection by thalamic axons. Science 281:559–562

Crowley JC, Katz LC (2000) Early development of ocular dominance columns. Science 1271–1273

Elliott T, Shadbolt NR (1999) A neurotrophic model for the development of the retinogeniculocortical pathway induced by spontaneous retinal waves. J Neurosci 19:7951–7970

Erwin E, Miller KD (1998) Correlation-based development of ocularly matched orientation and ocular dominance maps: determination of required input activities. J Neurosci 18:9870–9895

Fellenz WA, Taylor JG (2002) Establishing retinotopy by lateral–inhibition type homogeneous neural fields. Neurocomputing 48:313–322

Fitzpatrick D (2000) Seeing beyond the receptive field in primary visual cortex. Curr Op Neurobiol 10:38–443

Folias SE, Bressloff PC (2004) Breathing pulses in an excitatory neural network. SIAM J Appl Dyn Syst 3:378–407

Geman S (1979) Some averaging and stability results for random differential equations. SIAM J Appl Math 36:86–105

Goodhill GJ (1993) Topography and ocular dominance: a model exploring positive correlations. Biol Cybern 69:109–118

Ghosh A, Shatz CJ (1992) Pathfinding and target selection by developing geniculocortical axons. J Neurosci 12:39–55

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Neurosci 3:1116–1133

Hubel DH, Wiesel TN (1977) Functional architecture of macaque monkey visual cortex. Proc Roy Soc Lond B 198:1–59

Hubel DH, Wiesel TN, LeVay S (1977) Plasticity of ocular dominance columns in monkey striate cortex. Phil Trans Roy Soc Lond B 278:377–409

Kohonen T (1982) Self–organized formation of topologically correct feature maps. Biol Cybern 43:59–69

Lee HY, Yahyanejad M, Kardar M (2003) Symmetry considerations and development of pinwheels in visual maps. Proc Natl Acad Sci (USA)

LeVay S, Stryker MP, Shatz CJ (1978) Ocular dominance columns and their development in layer IV of the cat's visual cortex. J Comp Neurol 179:223–244

LeVay S, Nelson SB (1991) Columnar organization of the visual cortex. In: Leventhal AP (ed) The neural basis of visual function. CRC Press, Boca Raton: pp. 266–315

Linsker R (1986) From basic network principles to neural architecture: emergence of spatial-opponent cells. Proc Nat Acad Sci (USA) 83:7508–7512

Mackay D, Miller KD (1990) Analysis of Linsker's application of Hebbian rules to linear networks. Network: Comput Neural Syst 1: 257–297

Malach R, Amir Y, Harel M, Grinvald A (1993) Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. Proc Natl Acad Sci (USA) 90:10469–10473

Miller KD, Keller JB, Stryker MP (1989) Ocular dominance column development: analysis and simulation. Science 245:605–614

Miller KD (1994) A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity dependent competition between on- and off-center inputs. J Neurosci 14:409–441

Miller KD, MacKay DJC (1994). The role of constraints in Hebbian learning. Neural Comput. 6:100–124

Obermayer K, Blasdel GG (1993) Geometry of orientation and ocular dominance columns in monkey striate cortex. J Neurosci 13:4114–4129.

Penn AA, Shatz CJ (1999) Brain waves and brain wiring: The role of endogenous and sensory-driven neural activity in development. Pedeatric Res 45:447–458

Piepenbrock C, Obermayer K (1999) The role of lateral cortical competition in ocular dominance development. Proc Adv Neural Inf Proc Syst

Pinto D, Ermentrout GB (2001) Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses. SIAM J Appl Math 62:226–243

Shmuel A, Korman M, Harel M, Grinvald A, Malach R (1998) Relationship of feedback connections from area V2 to orientation domains in area V1 of the primate. Soc Neurosci Abstr 24:767

Sincich LC, Blasdel GG (2001) Oriented axon projections in primary visual cortex of the monkey. J Neurosci 21:4416–4426

Stryker MP, Harris WA (1986) Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. J Neurosci 6:2117–2133

Swindale NV (1996) The development of topography in visual cortex: a review of models. Network 7:161–247

Takeuchi A, Amari S-I (1979) Formation of topographic maps and columnar microstructures in nerve fields. Biol Cybern 35:63–72

Taylor JG (1999) Neural 'bubble' dynamics in two dimensions: foundations. Biol Cybern 80:393–409

Thomas PJ, Cowan JD (2004) Symmetry induced coupling of cortical feature maps. Phys Rev Lett 92:188101

Werner H, Richter T (2001) Circular stationary solutions in two–dimensional neural fields. Biol Cybern 85:211–217

Whitelaw VA, Cowan JD (1981) Specificity and plasticity of retinotectal connections: a computational model. J Neurosci 1:1369–1387

Willshaw D, Price D (2003) Models for topographic map formation. In: Modeling Neural Development A. van Ooyen (editor) pp. 213–244. MIT Press, Cambridge MA.

Willshaw DJ, von der Malsburg C (1976) How patterned neural connections can be set up by self-organization. Proc Roy Soc Lond B 194:431–445

Wimbauer S, Gerstner W, van Hemmen JL (1998) Analysis of a correlation–based model for the development of orientation–selective fields in the visual cortex. Network: Comp Neural Syst 9:449–466

Wolf F and Geisel T (1998) Spontaneous pinwheel annihilation during visual development. Nature 395:73–78

Wong ROL, Meister M, Shatz CJ (1993) Transient period of correlated bursting activity during development of the mammalian retina. Neuron 11:923–938

Woodbury GA, van der Zwan R, Gibson WG (2002) Correlation model for joint development of refined retinotopic map and ocular dominance columns. Vis Res 42:2295–2310

Yoshioka T, Blasdel GG, Levitt JB, Lund JS (1996) Relation between patterns of intrinsic lateral connectivity, ocular dominance, and cytochrome oxidase–reactive regions in macaque monkey striate cortex. Cerebral Cortex 6:297–310