THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Stability of Localized Patterns in Neural Fields

OPEN ACCESS

Stability of Localized Patterns in Neural Fields

Konstantin Doubrovinski
kdoubrov@princeton.edu
Princeton University, Department of Molecular Biology, Lewis Thomas Lab 343A,
Princeton, NJ 08544, U.S.A.

J. Michael Herrmann
michael.herrmann@ed.ac.uk
University of Edinburgh, Institute for Perception, Action and Behaviour,
Informatics Forum, 10 Crichton St., Edinburgh, EH8 9AB, U.K.

We investigate two-dimensional neural fields as a model of the dynamics of macroscopic activations in a cortex-like neural system. While the one-dimensional case was treated comprehensively by Amari 30 years ago, two-dimensional neural fields are much less understood. We derive conditions for the stability for the main classes of localized solutions of the neural field equation and study their behavior beyond parameter-controlled destabilization. We show that a slight modification of the original model yields an equation whose stationary states are guaranteed to satisfy the original problem and numerically demonstrate that it admits localized noncircular solutions. Typically, however, only periodic spatial tessellations emerge on destabilization of rotationally invariant solutions.

1 Introduction

Neural fields (Amari, 1977; Wilson & Cowan, 1973) describe the dynamics of distributions of activity on a layer of neurons. Neural fields have been suggested as models of internal representations in natural agents (Takeuchi & Amari, 1979; Gross, Stephan, & Krabbes, 1998) as well as in robots (Steinhage, 2000; Iossifidis & Steinhage, 2001; Erlhagen & Bicho, 2006; Giese, 1998). Various modalities are covered, such as spatial localization, viewing direction, attentional spotlight, the dynamics of decision making, elementary behaviors, and positions of other agents in the environment. More extensive studies in theoretical neuroscience (Suder, Wörgötter, & Wennekers, 2001; Mayer, Herrmann, & Geisel, 2002; Bressloff, 2005) concentrate on primary visual cortex (Lieke et al. 1989), superior culliculus

(Schierwagen & Werner, 1996), the representation of motoric primitives (Thelen, Schöner, Scheier, & Smith, 2001; Erlhagen & Schöner, 2002), and working memory in prefrontal cortex (Schutte, Spencer, & Schöner, 2003; Camperi & Wang, 1998). Generally neural fields serve as nonparametric representations of probability densities, and their dynamics may perform operations on the densities such as Bayesian computations (Herrmann, Pawelzik, & Geisel, 1999).

Recently neural fields have attracted attention in modeling and analysis of brain imaging data, because they are able to represent the dynamic interaction of an active medium with time-varying inputs and because the spatial and temporal scales in the data and neural field models are starting to become comparable. Especially if information about connectivity is available from the data (Jirsa, Jantzen, Fuchs, & Kelso, 2002), neural fields are of high explanatory power. Moreover the analysis has reached a level where applications directly benefit from the theoretical progress, and at the same time, computational power has become available that allows us to perform online simulations of two-dimensional neural fields.

One-dimensional problems were comprehensively studied in the 1970s (Amari, 1977; Kishimoto & Amari, 1979; Takeuchi & Amari, 1979). While for spatially extended (e.g., periodic) patterns, the transition to the more relevant two-dimensional case is nontrivial but fairly well understood (Ermentrout & Cowan, 1979; Ermentrout, 1998), localized activities in dimensions larger than one have not yet been treated with the same rigor. The situation, however, does not seem to be complex: a large body of numerical studies, together with theoretical considerations (Laing, Troy, Gutkin, & Ermentrout, 2002; Laing & Troy 2003), imply a general instability of multibump solutions if the interactions are excitatory at small and inhibitory at large distances (see also Laing & Chow, 2001, for stability analysis in one-dimensional models of spiking neurons). Further, there has been numerical evidence that single-bump solutions in two dimensions for radially symmetric interactions are essentially circular, which was exploited as an assumption in Taylor's early attack on the two-dimensional case (Taylor, 1999). Werner and Richter (2001) provided evidence for the existence of ring-shaped solutions that are possible for certain types of neural interactions. Along these lines, one may conjecture that finite mesh-like structures of higher genus exist as well.

The situation became more spirited recently when Herrmann, Schrobsdorff, and Geisel (2004), Bressloff (2005), and Doubrovinski (2005) tackled the stability problem of localized activations in two-dimensional fields. Although the generality that has been achieved in the one-dimensional neural fields cannot be expected in two dimensions, a number of interesting variants of the circular activity configurations have been analyzed.

Here we present a concise and reproducible scheme for the analysis of the stability of localized activity distributions in neural fields. We show the

applicability of the scheme to the special case of circular solutions and also study ring-shaped and bar-shaped solutions, which present the complete set of known localized solutions for simple kernels (Werner & Richter, 2001). In addition to the stability proofs based on the classical scheme (Pismen, 1999; Bressloff, 2005; Doubrovinski, 2005), we are interested in behavior beyond the phase transitions toward unstable regions. The destabilization of a circular solution is known to lead to a transient elongation of the activity bubble (Bressloff, 2005; Doubrovinski, 2005), which ultimately causes the localized solution to split or form meandering bands. Either case is unstable in a strict sense: the splitting into two continues toward a plane-filling hexagonal pattern, while the banded patterns develop a global stripe pattern or a quasiperiodic arrangement.

The destabilization is thus fundamental since for typical Mexican hat interaction kernels, there is no nearby stable state that is approached after the bifurcation, while the spatially extended patterns are not approached in finite time (unless a general criterion for convergence is drawn into consideration). Yet the destabilization, at least in neurobiological applications, is the most interesting part of the theory. Divergences are usually very slow and may halt completely due to reasonable boundary conditions, such that an activity-based correlational learning scheme may organize anisotropies in the connections, which stabilizes the anisotropic activities as assumed in Bressloff (2005) and exploited in Schierwagen and Werner (1996). A theoretical account of the interaction of activity dynamics and learning was studied in Dong and Hopfield (1992) in relation to the activity effect on feature maps (cf. Mayer et al., 2002).

2 The Neural Field Equation

The neural field model describes the activations of a layer of neurons when the geometry of interactions rather than the specific connectivity among the neurons is relevant. We assume positions $\mathbf{r} \in \mathbb{R}^2$ for neurons with continuous-valued activations $u(\mathbf{r}, t)$. The synaptic weights between neurons at the positions $\mathbf{r}$ and $\mathbf{r}'$ are expressed by isotropic interaction kernel $w(\mathbf{r}, \mathbf{r}') = w(|\mathbf{r} - \mathbf{r}'|)$ of Mexican hat shape. Neurons are activated if their total input is greater than zero. We will study only equilibrium solutions without external input, and we neglect slow learning effects, so the synaptic weights are constant over time.

The activation at a position results from a weighted integration over the inputs from all other active locations in the field and a natural decay toward a resting potential denoted by h. The dynamics of the neural field is thus determined by the equation

$$\tau \frac{\partial u(\mathbf{r}, t)}{\partial t} = -u(\mathbf{r}, t) + \int_{R[u]} w(|\mathbf{r} - \mathbf{r}'|) \, d\mathbf{r}' + h, \qquad (2.1)$$

where $R[u] = \{x \mid u(\mathbf{r}) > 0\}$ is the excited region, that is, a neuron receives input only from neurons within R. The boundary of $R[u]$ is assumed to be smooth. Rescaling time, τ can be set to unity without loss of generality. Equilibrium solutions are defined by

$$u(\mathbf{r}, t) = \int_{R[u]} w(|\mathbf{r} - \mathbf{r}'|)\,d\mathbf{r}' + h \qquad (2.2)$$

and depend on the value of the threshold parameter h and the particular form of w. Here we use a smooth kernel which is more general than the quasi-constant kernel function in Herrmann et al. (2004). The kernel is constructed as a difference of gaussian functions and is defined by four parameters: K, k, M, m:

$$w = K \exp\left(-k\|\mathbf{r} - \mathbf{r}'\|^2\right) - M \exp\left(-m\|\mathbf{r} - \mathbf{r}'\|^2\right). \qquad (2.3)$$

If nonrotationally symmetric solutions are excluded from the beginning from the consideration of equation 2.2, the situation simplifies dramatically, and it can be shown (Taylor, 1999) that one-bump solutions $u(\|\mathbf{r}\|)$ of certain radii are stationary states of the dynamics 2.1. Analogously, a ring-shaped solution (Werner & Richter, 2001) or a stripe-like solution (i.e., a degenerate ring of infinite radius) can occur. However, when an arbitrarily small perturbation of the solution is being considered, the symmetry might be broken and new phenomena can appear.

3 Stability

It has previously been shown that equation 2.1 admits rotationally invariant stationary solutions with disc-shaped activated region (one-bump solutions). Generically these arise in the course of a "blue sky bifurcation" (Strogatz, 1994): no solution is present in the subcritical parameter region, whereas two solution branches bifurcate as the control parameter exceeds the critical value. The two solutions are rotationally invariant one-bumps. Stability analysis of these states with respect to rotationally invariant perturbation is essentially equivalent to stability analysis of the one-bump solution of the one-dimensional model. It reveals that the unstable branch generically corresponds to the bump of smaller radius. Upon destabilization, the region of activation expands as the solution approaches the stable branch, corresponding to the circular bump of larger radius (see Figure 1).

Seeking a stationary solution $\overline{u}$ of the two-dimensional model, equation 2.1, assuming rotational invariance of the field (i.e., $\overline{u}(\mathbf{r}) \equiv \overline{u}(r)$ with $r = \|\mathbf{r}\|$), leads to a problem that is essentially equivalent to that of finding
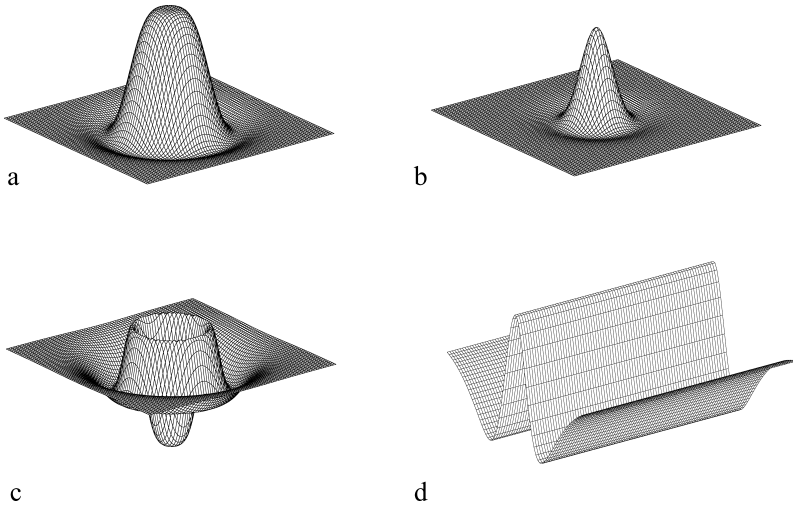
Figure 1: (a) Stable one-bump solution of the neural field equation. Parameters are $K = 2.5$, $k = 5$, $M = 0.5$, $m = 0.5$, and $h = -0.281$. (b) Unstable one-bump solution for the parameter values $K = 2.5$, $k = 5$, $M = 0.5$, $m = 0.5$, and $h = -0.294$. (c) Annular solution with the parameters $K = 2.5$, $k = 5$, $M = 0.5$, $m = 0.5$, and $h = -0.115$. (d) Part of a stable stripe-like solution. Parameters as in c. An unstable solution also exists at the same parameters.

stationary states in the one-dimensional model (see the appendix). Two types of solutions besides the circular one-bumps are readily constructed: solutions with annular activated regions and solutions with a stripe-shaped region of activation (see Figure 1). The possibility of the existence of the former has been pointed out previously, while the latter can be seen as degenerate annuli of infinite inner radius.

We now turn to the analysis of the stability properties of the above-mentioned stationary states. Consider the one-bump solution with a disc-shaped activated region (i.e., $u(r) > 0$ iff $r < R$). Consider the dynamics of a small perturbation $\epsilon\eta$,

$$u(\mathbf{r}, t) = \overline{u}(\|\mathbf{r}\|) + \epsilon\eta(r,t). \tag{3.1}$$

Inserting into equation 2.1 and keeping the terms of up to the first order in ϵ, one finds that to linear-order the dynamics of the perturbation obeys

$$\frac{\partial \eta\,(\mathbf{r}, t)}{\partial t} = -\eta(\mathbf{r}, t) + \int_{\mathbb{R}^2} w(\|\mathbf{r} - \mathbf{r}'\|)\, \delta(\overline{u}(\mathbf{r}'))\, \eta(\mathbf{r}', t)\, d\mathbf{r}', \tag{3.2}$$

where $\delta(\overline{u}(\mathbf{r}'))$ is a Dirac delta function (see the appendix). Substituting an ansatz of the form $\eta(\mathbf{r}, t) = \exp(\lambda t)\xi(\mathbf{r})$, one arrives at an eigenvalue problem that in polar coordinates becomes (see the appendix)

$$\lambda\xi(r, \theta) = -\xi(r, \theta) + \Gamma \int_0^{2\pi} g(\theta - \theta')\,\xi(r, \theta')\,d\theta'. \tag{3.3}$$

Here, g is a 2π periodic function depending on kernel $w(\|\mathbf{r} - \mathbf{r}'\|)$, and Γ is a constant given by

$$\Gamma = R \left| \frac{\partial \overline{u}(r)}{\partial r} \right|_{r=R}^{-1}, \tag{3.4}$$

that is, the ratio of the radius of the activated region R to the absolute value of the slope of the radial profile of the stationary solution at $r = R$. Clearly, explicit evaluation of Γ requires calculating the stationary solution, which is implicitly given by equation 2.2 in terms of a double integral. Equation 3.3 is known as Fredholm's integral equation of the second kind. The integral operator on the right-hand side of equation. 3.3 is compact, bounded, and self-adjoint, implying that every spectral value is an eigenvalue, all eigenvalues are real, each eigenspace is finite-dimensional, and zero is the only possible accumulation point of eigenvalues (Kreyszig, 1978). Solving equation 3.3, we obtain eigenvalues and eigenfunctions that in polar coordinates read (see the appendix)

$$\lambda_n = -1 + \Gamma \int_0^{2\pi} g(\theta) \cos(n\theta)\,d\theta$$

$$\xi_n = \int_0^{2\pi} w(r, \theta, R, \theta') \cos(n\theta')\,d\theta', \tag{3.5}$$

where R is the radius of the circular activated region. The nth eigenfunction is $2\pi n$-periodic in θ, implying that it is D_n-symmetric (symmetry with respect to rotation by $2\pi/n$ around the origin and with respect to reflections on the n respective symmetry planes; the square is, e.g., D_4-symmetric, whereas a regular hexagon is D_6-symmetric) and corresponds to a multiperiodic deformation of a circle. Eigenvalue spectra for one-bump solutions are given in Figure 2. Certain features of these are readily interpretable. For example, we see that for the bump with a smaller radius, the eigenvalue λ_0, corresponding to rotationally invariant deformation is positive, implying instability with respect to perturbations in the radius of the bump. The spectrum, corresponding to the larger bump, however, is nonpositive, implying stability with respect to arbitrary perturbation in agreement with earlier results. Also, the eigenvalue λ_1, which corresponds to a 2π-periodic
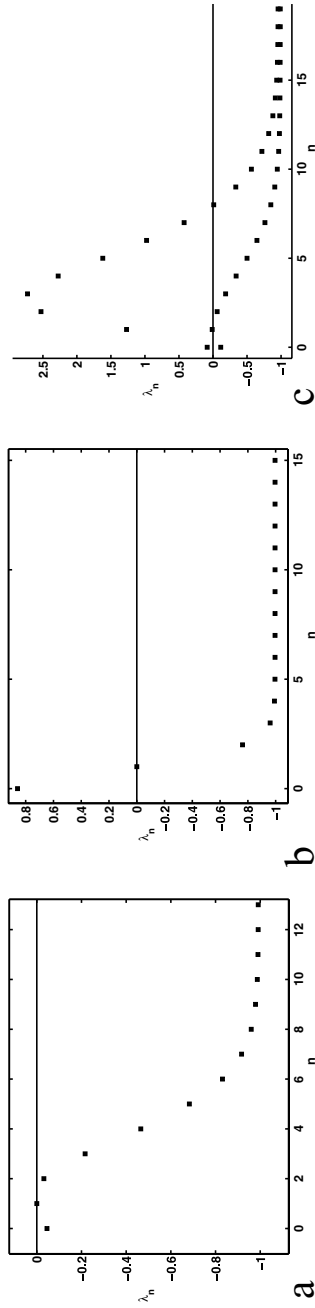
Figure 2: Spectra corresponding to the solutions in Figures 1a–1c, resp. The positivity of the zeroth mode in *b* indicates the instability of this solution with respect to circular perturbations. The solution (a) is stable, with a marginal instability with respect to a lateral shift as visible from the vanishing first eigenvalue. The instability of the annulus solution in *c* shows a characteristic scale that causes the solution to split into three single bumps.
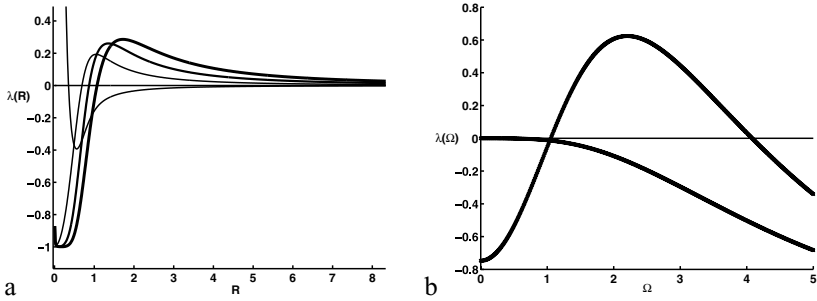
Figure 3: (a) Four eigenvalues versus bump radius R. The curve that represents λ_0 describes the stability with respect to perturbations in bump radius. The others are (from left to right) λ_2 (reflection-invariant deformation), λ_3 (D_3-invariant eigenmode), and λ_4 (D_4-invariant eigenmode). λ_1 is identically zero, reflecting the metastability of the solutions with regard to lateral shifts. Parameters are $K = 1.5$, $k = 5$, $M = 0.5$, and $m = 1.5$. (b) The spectrum determining the stability of the stripe-shaped solutions shown in Figure 1d. Only the information for the stable solution if given here. This solution loses stability for a band of values of Ω (see section A.3).

deformation (or, equivalently, a translation of the bump), vanishes, reflecting the translational invariance of equation 2.1. Exploiting this observation, it appears possible to reexpress Γ in equation 3.4 more explicitly as $\Gamma = 1/\int_0^{2\pi} g(\theta)\cos(\theta)\,d\theta$, whereby the expressions for the other eigenvalues simplify to

$$\lambda_n = -1 + \frac{\int_0^{2\pi} g(\theta)\cos(n\theta)\,d\theta}{\int_0^{2\pi} g(\theta)\cos(\theta)\,d\theta}. \tag{3.6}$$

Note that contrary to equation 3.5, equation 3.6 does not explicitly contain the stationary solution $\bar{u}$, which allows us to calculate the nth eigenvalue as a function of the radius of the activated region without evaluating double integrals in the implicit expression for $\bar{u}$. Only integrals over one-dimensional manifolds appear in equation 3.6, greatly simplifying the calculation of the spectrum. Apart from theoretical considerations, this is of importance for technical applications since the knowledge of stability properties of solutions of equation 2.1 could affect their use for representing probability distributions, for example, in implementations of autonomous robot memory.

Previous work (Werner & Richter, 2001) conjectured that the bifurcation branch corresponding to the stable one-bump remained stable for all values of the control parameter. Using equation 3.6, this is readily proved false: higher and higher frequency eigenmodes progressively turn unstable as the radius of the stationary solution is increased (see Figure 3).

Figure 4: (Top) Time evolution of an unstable bump undergoing stability loss at a D_2-eigenmode. Parameters are $K = 1.5$, $k = 5$, $M = 0.5$, $m = 1.5$, and $h = -1.46 \cdot 10^{-2}$. (Bottom) Time evolution of an unstable bump undergoing stability loss at a D_3 invariant eigenmode. Parameters: $K = 1.5$, $k = 5$, $M = 0.5$, $m = 1.5$, and $h = -4.43 \cdot 10^{-3}$.

Strictly speaking, the assertion that the linear stability analysis as outlined above correctly determines the stability properties of stationary solutions relies on additional assumptions on operators appearing on the right-hand side of equation 2.1 (Golubitsky, Schaeffer, & Stewart, 1985). For infinitely dimensional nonsmooth fields, these generically need not hold. In order to check whether stability analysis is adequate, correctly determining the stability properties of the stationary solutions, we performed a number of numerical simulations. Figure 4 depicts a simulation of unstable one-bump whose corresponding stability spectrum reveals that the maximal (positive) eigenvalue is that of the D_2-symmetric eigenmode. As time elapses, the initially circular activated region keeps deforming, forming bloblike protrusions. These subsequently bud off from the middle bump. The newly formed activated domains keep splitting, progressively tiling the plane in a hexagonal pattern. We stress that the pattern that is formed immediately after destabilization of the stationary state is D_2-symmetric, as would be expected from the properties of the eigenvalue spectrum. That is, stability properties of the stationary solutions, as well as qualitative aspects of the pattern, forming upon destabilization of the steady state, are correctly predicted from the linear stability analysis.

Symmetry breaking accompanying destabilization of a stationary bump was examined for a broad range of parameters (e.g., parameters that yielded solutions with a maximal eigenvalue of the respective spectrum corresponding to D_3-, D_4-, and D_8-symmetric perturbations; see Figure 4). In all of the cases, the course of the symmetry breaking was appropriately determined by linear stability.

Stability analysis of annular solutions proceeds along the same lines as that of one-bumps (see the appendix). The essential difference is that instead of a single equation, 3.3, a system of two equations results, meaning that to

every nonnegative whole number corresponds a pair of real eigenvalues. (In general, to every boundary of an activated domain, there corresponds an equation in the corresponding eigenvalue problem.) We find that in the case of the annular solution shown in Figure 1c, the largest eigenvalue corresponds to a D_3-symmetric perturbation. Simulations demonstrate that initially a rotationally invariant annulus splits into three adjacent blobs that gradually drift apart (not shown). The symmetry of the resulting state is the same as that of the largest eigenvalue. Again, the stability of the stationary solution, as well as qualitative aspects of the emerging pattern, are readily predicted from the analytically computed spectrum.

Finally, let us consider the stripe-shaped solutions. Their stability is governed by an eigenvalue problem, similar to that governing the stability of the annuli. However, the corresponding linear operator is no longer compact (this is a consequence of activated region being unbounded), and the spectrum no longer needs to remain discrete. Actually, in this case, the spectrum is continuous: to every real corresponds a pair of (real) eigenvalues. Figure 3b shows results of the linear stability analysis of the stripe solution, depicted in Figure 1d. In the corresponding simulation, the stripe is seen to split into a row of separate bumps—a "chain-of-pearls" configuration. The interbump separation is the same as the wavelength of the eigenmode, corresponding to the largest eigenvalue. Again, stability and semiquantitative properties of patterns resulting from the stationary state destabilization are readily predicted from the respective eigenvalue spectrum.

In summary, the preceding section describes all of the stationary non-homogeneous solutions of the two-dimensional Amari equation known to date and exhaustively examines their stability properties. Quite strikingly, stability analysis of the nonhomogeneous steady state is possible since the eigenvalue problem, equation 3.5, is effectively one-dimensional, although a two-dimensional system is being considered.

4 Modified Equation

A long-standing question regarding the Amari model is the existence of nonrotationally invariant stationary solutions with bounded and connected regions of activation. Such states are likely to bifurcate from circular solutions on destabilization at a nonrotationally invariant eigenvalue. In fact, the stability loss of one branch is always accompanied by the emergence of another in its vicinity, provided that the mappings defining the dynamical system under consideration are sufficiently smooth (Crandall & Rabinowitz, 1971). However, in the simulations, exclusively periodic patterns resulted on destabilization of circular solutions.

In order to address the existence of the nonrotationally invariant solutions of equation 2.1 with a bounded and connected activated region, we modify the original Amari model, equation 2.1, so as to obtain a related (modified) equation fulfilling the following three conditions:

1. Stationary solutions of the modified equations should be solutions of the unmodified equation, 2.2.
2. The modified equation should not admit solutions with an unbounded activated region.
3. The stability of a solution to equation 2.2 should remain unaltered by the modification.

According to condition 3, rotationally invariant stationary solutions behave like those of the unmodified Amari model, admitting symmetry breaking at a nonrotationally invariant eigenvalue. However, the symmetry breaking cannot result in a spatially extended periodic pattern according to condition 2. Consequently, a nonrotationally invariant localized state is likely to emerge. It will be a solution of the original (unmodified) Amari model with desired properties, provided that its region of activation remains connected in the course of destabilization. Note, however, that conditions 1 through 3 do not suffice to ensure that the activated domain will not start splitting into a separate disconnected region upon symmetry breaking.

We now turn to the construction of a modification of equation 2.1 satisfying conditions 1 to 3. Consider a circular one-bump solution $\overline{u}_h$ of equation 2.1 with resting potential h and the area of activated region $A[\overline{u}_h]$. Let us modify the equation according to

$$\partial_t u = -u(\mathbf{r}, t) + \int_{R[u]} [w(|\mathbf{r} - \mathbf{r}'|) + q] \, d\mathbf{r}' - q \, A[\overline{u}_h] + h', \qquad (4.1)$$

where q is any real number. Suppose that $\overline{u}$ is a stationary solution of equation 4.1 for a certain q. Substituting $\overline{u}$ into equation 4.1, one obtains with the area of $R[\overline{u}]$ being denoted by $A[\overline{u}]$:

$$0 = -\overline{u}(\mathbf{r}, t) + \int_{R[\overline{u}]} [w(|\mathbf{r} - \mathbf{r}'|) + q] \, d\mathbf{r}' - q \, A[\overline{u}_h] + h' =$$

$$= -\overline{u}(\mathbf{r}, t) + \int_{R[\overline{u}]} w(|\mathbf{r} - \mathbf{r}'|) \, d\mathbf{r}' + q \, A[\overline{u}] - q \, A[\overline{u}_h] + h', \qquad (4.2)$$

implying that $\overline{u}$ is a solution of equation 2.2 with the resting potential h replaced by $h' - q \, A[\overline{u}_h] + q \, A[\overline{u}]$. Consequently, condition 1 is satisfied. Note that the circular one-bump solution $\overline{u}_h$ of the original problem, equation 2.2, which was used when constructing equation 4.1, solves the modified problem 4.1 with $h' = h$ and any q.

Equation 4.1 does not admit stationary solutions with an unbounded region of activation. Indeed, assuming that such a solution $\hat{u}$ exists, substituting $\hat{u}$ into the equation, we would conclude that the integral term of the right-hand side of equation 4.1 is infinite, which forms a contradiction. Therefore, condition 2 above is satisfied.

Finally, we show that modification 4.1 preserves stability properties of the stationary solution $\overline{u}_h$ (which is a stationary solution of both the modified and the unmodified problems, 2.1 and 4.1, respectively, by construction). Recall that in deriving equation 3.5, we did not make use of any particular assumptions on the form of the integral kernel $w(|\mathbf{r} - \mathbf{r}'|)$. Consequently, this expression for the eigenvalue spectrum is equally valid for the modified problem as well as for the unmodified one. Note, however, that when deriving a stability spectrum in the case of the modified problem, equation 4.1, we shall exchange $g(\theta - \theta')$ by $g(\theta - \theta') + q A[\overline{u}_h]$ (see the derivations in the appendix). Using $\int_0^{2\pi} q A[\overline{u}_h] \cos(n\theta)\, d\theta = q A[\overline{u}_h] \int_0^{2\pi} \cos(n\theta)\, d\theta = q A[\overline{u}_h] 2\pi \delta_{n0}$, it now follows from equation 3.5 that all eigenvalues of the stability spectrum of $\overline{u}_h$ (except for λ_0 corresponding to perturbations of the radius of the bump) remain unaltered by the modification. Consequently, if $\overline{u}_h$ is unstable with respect to some nonrotationally invariant perturbation in the original problem, 2.1, it is unstable with respect to such a perturbation in the modified problem, 4.1, whereby condition 3 holds.

These arguments imply that a rotationally invariant solution $\overline{u}_h$ of equation 2.1 that is unstable at a noncircular eigenmode also solves the modified equation, 4.1, and is unstable with respect to the same eigenmode of the dynamics of equation 4.1. Furthermore, contrary to the case of equation 2.1, the dynamics of equation 4.1 can never result in a periodic pattern with unbounded activated region if $q < 0$.

As stated above, conditions 1 to 3 do not suffice to guarantee that the activated region will remain connected as $\overline{u}_h$ follows the dynamics of equation 4.1. Nevertheless, by tuning the parameter q in equation 4.1, one is able to trap the dynamics in the vicinity of instability in a state with connected activated region.

Assume that for $q = 0$, the dynamics in the vicinity of the bifurcation tends to increase the area of the activated region. Note that q could be understood as a Lagrange multiplier, which ensures the constancy of the area for one special value, which we certainly exclude. Suppose now that q is chosen to be negative. The additional term in the integrand of equation 4.1 will tend to counterbalance the area increase. We can now choose q such that two effects counterbalance, and a nonrotationally invariant steady state with bounded and connected region of activation will result. The calculation of effective values of q requires the consideration of higher orders of the dynamics. Here, we instead resort to numerical simulations. These confirm the emergence of nonrotationally invariant steady states with bounded and connected regions of activation. For example, Figure 5 shows the destabilization of a circular one-bump that is unstable at a D_2-symmetric eigenmode, developing into a nonrotationally invariant steady state with an ellipse-shaped activated region. Only one-quarter of the domain was simulated (the field in the other three quadrants is determined by that on the simulated quadrant due to the Euclidean symmetry of
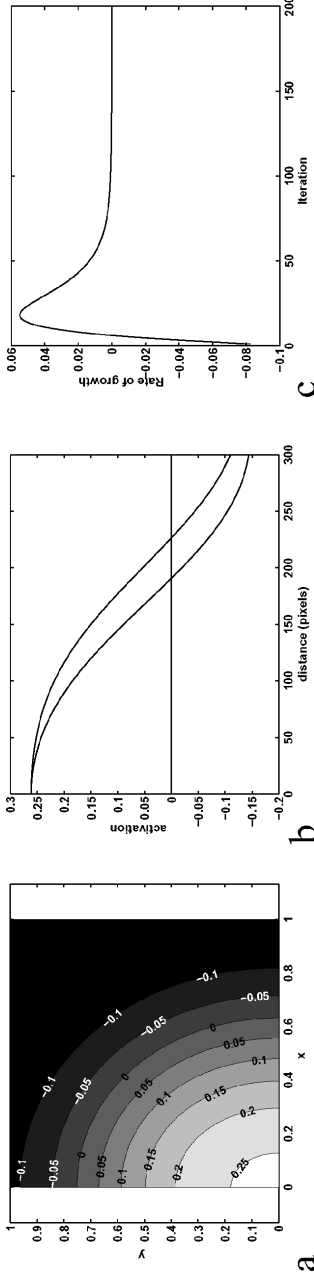
Figure 5: (a) Contour plot of the D_2-symmetric solution superimposed with its level curves. The zeroes' level curve (corresponding to $u = 0$) delimits the activated region. The symmetries of the equation and initial condition allow restricting the simulation to one-quarter of the domain. The resulting blob is roughly ellipse shaped, with the vertical semi-axis being longer than the horizontal one. (b) Cross-section through the profile of the stationary solution along the coordinate axes. (c) Rate of growth of the solution versus the iteration step calculated as a maximal deviation between two subsequent time steps. Parameters are $K = 1.5$, $k = 5$, $M = 0.5$, $m = 1.5$, $h = 1.50$, and $q = 0.02850$.

dynamic equations and the D_2-symmetry of initial conditions) on a grid of 300×300 pixels in order to increase accuracy. Symmetries corresponding to higher eigenvalues are irrelevant because the eigenvalues λ_n are stable for $n \geq 3$ for the given parameters. The length difference of the half-axes of the activated domain of the resulting D_2-symmetric stationary state was much larger than the spatial discretization step (some 30 pixels), allowing the conclusion that the stationary solution obtained is not a discretization artifact.

In conclusion, we stress that the particular choice of modification according to equation 4.1 is very restrictive. In fact, many other equations whose stationary states satisfy equation 2.1 and do not admit solutions with unbounded regions of activation are readily constructed along the same lines. For instance, $\partial_t u = -u + \int_{R[u]} w(|\mathbf{r} - \mathbf{r}'|)\, d\mathbf{r}' - \int_{R[u]} |\mathbf{r} - \mathbf{r}'|\, d\mathbf{r}'\, d\mathbf{r} + h$ can be shown to have these properties, arguing essentially as when proving that conditions 1 and 2 are satisfied by stationary states of equation 4.1. We believe that further investigation of such modifications will provide insight into the properties of unstable solutions of the Amari equation.

5 Discussion

Bifurcation theory describes the time course to critical behavior in low-dimensional systems. Under certain conditions, the parametric destabilization of an activity distribution does not lead to a nearby stable state but initiates a cascade of symmetry-breaking events that eventually approaches a new distant stable configuration. A motivation for this study is to demonstrate the complex evolution of the state of the field after the loss of stability. It is these configurations that bear the greatest computational potentials.

For the Amari equation with a simple kernel, only the existence of rotationally invariant bumps has previously been proven. This work demonstrates the existence of two other nonperiodic stationary states: stripe shaped and annular solutions. Stability analysis of these would be expected to be very involved. Surprisingly, the special form of the Amari equation makes this stability problem amenable to analytical treatment. The total synaptic input to a given neuron is dependent only on the shape of the boundary of the activated domain, making the eigenvalue problem effectively one-dimensional, thereby allowing fairly straightforward calculation of the spectrum for each of these cases.

Strictly speaking, spectral properties of the linearized operator do not guarantee the stability of the stationary state unless additional assumptions are satisfied. However, our extensive numerical investigation shows that stability is indeed correctly predicted by eigenvalue analysis. Furthermore, the spectrum allows predicting certain semiquantitative features of solutions approached after the onset of instability.

Our numerical experiments showed that exclusively spatially extended solutions (those with unbounded activated domain) appeared in the unstable parameter region. Yet a slight modification of the interaction kernel, introducing long-range interactions, circumvents this, yielding nonrotationally invariant stationary states with bounded and connected regions of activation that at the same time are stationary states of the original unmodified equation. This settles a long-standing question regarding the existence of solutions of this type. Note, however, that stability analysis of these is nontrivial, requiring the knowledge of analytical parameterization of the boundary of the activated domain. The existence of such states is also relevant for biological systems. Assuming a certain degree of shift-twist symmetry, the existence of elongated blobs suggests mechanisms for the emergence of orientation selectivity in neurons of the primary visual cortex (Tanaka, Ribot, Imamura, & Tani, 2006). Although the degree of asymmetry of nonrotationally invariant solutions was moderate, these effects could in principle be enhanced by (Hebbian) learning, which we disregarded in our treatment.

Appendix: Linear Stability of Rotationally Invariant Solutions

A.1 Circular One-bump Solution. Consider the development of a small perturbation $\varepsilon\eta$ of a stationary rotationally invariant one-bump $\bar{u}$. Substituting into the Amari equation and linearizing in ε, we have

$$\frac{\partial \eta(\mathbf{x}, t)}{\partial t} = -\eta(\mathbf{x}, t) + \int_{\mathbb{R}^2} w(|\mathbf{x} - \mathbf{x}'|)\, \delta(\bar{u}(\mathbf{x}'))\, \eta(\mathbf{x}', t)\, d\mathbf{x}'. \tag{A.1}$$

In polar coordinates, we write (somewhat informally) $w(|\mathbf{x} - \mathbf{x}'|) = w(r, \theta, r', \theta')$ and use the rotational invariance of $\bar{u}$, such that

$$\frac{\partial \eta(r, \theta, t)}{\partial t} = -\eta(r, \theta) + \int_0^{2\pi} \int_0^\infty r'\, w(r, \theta, r', \theta')\, \delta(\bar{u}(r'))\, \eta(r', \theta')\, dr'd\theta'. \tag{A.2}$$

Recall that

$$\delta(f(x)) = \sum_{x_i} \frac{\delta(x - x_i)}{\left| \frac{df(x_i)}{dx} \right|}, \tag{A.3}$$

where the sum is over the roots of f, provided that f is differentiable at the corresponding points. Using equation A.3, equation 2.2 simplifies to

$$\frac{\partial \eta(r, \theta)}{\partial t} = -\eta(r, t) + \int_0^{2\pi} Rw(r, \theta, R, \theta') \frac{1}{\left| \frac{\partial \bar{u}(R)}{\partial r} \right|} \eta(R, \theta')\, d\theta', \tag{A.4}$$

where R is the radius of the (circular) region of activation of $\bar{u}$. Substituting $\eta = e^{\lambda t} \xi(r, \theta)$, we arrive at the following eigenvalue problem,

$$\lambda \xi(r, \theta) = -\xi(r, \theta) + \Gamma \int_0^{2\pi} w(r, \theta, R, \theta') \xi(R, \theta') \, d\theta', \qquad (A.5)$$

where $\Gamma \equiv R/|(\partial \bar{u}(R)/\partial r)|$. By restriction of both sides to $r = R$, the eigenvalue problem A.5 is solved by

$$\xi_n(R, \theta) = \cos n\theta$$

$$\lambda_n = -1 + \Gamma \int_0^{2\pi} w(R, \theta, R, 0) \cos(n\theta) \, d\theta, \qquad (A.6)$$

where n are nonnegative integers. The last equation can be verified by noting that $w(R, \theta, R, \theta')$ is a function of $(\theta - \theta')$ alone and Fourier-expanding the integrand of equation A.5.

The r-dependence of the eigenfunctions can be derived from equation A.6 by exploiting the special form of the eigenvalue problem, equation A.5:

$$\xi_n(r, \theta) = \int_0^{2\pi} w(r, \theta, R, \theta') \cos(n\theta') \, d\theta'. \qquad (A.7)$$

Shift invariance allows us to conclude that $\lambda_1 = 0$. Thus, Γ is obtained more explicitly from equation A.6:

$$\Gamma = \frac{1}{\int_0^{2\pi} w(R, \theta, R, 0) \cos(\theta) \, d\theta}. \qquad (A.8)$$

Now the eigenvalues can be calculated without evaluating the slope of the stationary solution $\bar{u}$:

$$\lambda_n = -1 + \frac{\int_0^{2\pi} w(R, \theta, R, 0) \cos(n\theta) \, d\theta}{\int_0^{2\pi} w(R, \theta, R, 0) \cos(\theta) \, d\theta}. \qquad (A.9)$$

Equation A.9 is particularly convenient for examining the stability properties of circular one-bump solutions.

A.2 Annular Solutions. The existence of solutions with annular region of activation was suggested already by Amari (1977). Denoting the inner radius by R_1 and the outer radius by R_2, equation A.4 is immediately

rewritten as

$$\frac{\partial \eta (r, \theta)}{\partial t} = -\eta (r, \theta) + \Gamma_1 \int_0^{2\pi} w(r, \theta, R_1, \theta') \eta (R_1, \theta') d\theta'$$

$$+ \Gamma_2 \int_0^{2\pi} w (r, \theta, R_2, \theta') \eta (R_2, \theta') d\theta', \tag{A.10}$$

where $\Gamma_1 = R_1/|\partial \bar{u}(R_1)/\partial r|$, $\Gamma_2 = R_2/|\partial \bar{u}(R_2)/\partial r|$. Analogous to the derivation of equation A.5, we set $\eta = e^{\lambda t}\xi(r, \theta)$ and restrict both sides to R_1 and R_2:

$$\lambda \xi_1 = -\xi_1 + \Gamma_1 \int_0^{2\pi} w (R_1, \theta, R_1, \theta') \xi_1' \theta'$$

$$+ \Gamma_2 \int_0^{2\pi} w (R_1, \theta, R_2, \theta') \xi_2' d\theta'$$

$$\lambda \xi_2 = -\xi_2 + \Gamma_1 \int_0^{2\pi} w(R_2, \theta, R_1, \theta') \xi_1' d\theta'$$

$$+ \Gamma_2 \int_0^{2\pi} w (R_2, \theta, R_2, \theta') \xi_2' d\theta', \tag{A.11}$$

where $\xi_i = \xi(R_i, \theta)$ and $\xi_i' = \xi_i(\theta')$, $i \in \{1, 2\}$. It is natural to seek the eigenfunctions of the form $[\xi_1, \xi_2] = \mathbf{v} \cos(n\theta)$, where $\mathbf{v}$ is a two-dimensional vector. Using this substitution, it turns out that the eigenvalues of equation A.11 are the same as those of the matrices

$$\begin{bmatrix} -1 + \Gamma_1 \int_0^{2\pi} w_{11} \cos n\theta\, d\theta & \Gamma_2 \int_0^{2\pi} w_{12} \cos n\theta\, d\theta \\ \Gamma_1 \int_0^{2\pi} w_{12} \cos n\theta\, d\theta & -1 + \Gamma_2 \int_0^{2\pi} w_{22} \cos n\theta \end{bmatrix}, \tag{A.12}$$

such that equation A.12 allows us to construct the spectrum of equation A.11.

A.3 Stripe-Shaped Solutions. Stripe-shaped solutions can be constructed by assuming the region of activation of the form $R[\bar{u}] = \{(x, y) \mid 0 \leq x \leq L\}$ and can be seen as degenerate annuli with infinite inner radius, considered in the previous section. Interestingly, in this case, it appears possible to give an explicit expression for the stationary one-bump solution in terms of the model parameters:

$$\bar{u}(x, y) = \frac{\pi}{2km}(Km\,\mathrm{erf}\,(\sqrt{k}x) - kM\,\mathrm{erf}\,(\sqrt{m}x) -$$

$$- Km\,\mathrm{erf}\,(-\sqrt{k}L + \sqrt{k}x) + Mk\,\mathrm{erf}\,(-\sqrt{m}L + \sqrt{m}x)) - h. \tag{A.13}$$

The counterpart of equation A.10 now reads

$$\frac{\partial \eta}{\partial t} = -\eta + \int_{\mathbb{R}^2} w(|\mathbf{x} - \mathbf{x}'|)\, \delta\left(\bar{u}(\mathbf{x}')\right) \eta\, (d\mathbf{x}')$$

$$= -\eta + \int_{-\infty}^{\infty} w(x, y, 0, y') \frac{1}{\left| \frac{\partial \bar{u}(0)}{\partial x} \right|}\, \eta\, (0, y')\, dy' \tag{A.14}$$

$$+ \int_{-\infty}^{\infty} w(x, y, L, y') \frac{1}{\left| \frac{\partial \bar{u}(L)}{\partial x} \right|}\, \eta\, (L, y')\, dy', \tag{A.15}$$

where (x, y) and (x', y') are Cartesian coordinates of $\mathbf{x}$ and $\mathbf{x}'$, respectively. As in the former cases, substituting $\eta(x, y, t) = e^{\lambda t} \xi(x, y)$ and restricting to 0 or to L, we find

$$\lambda \xi_1 = -\xi_1 + \Gamma_1 \int_{-\infty}^{\infty} w(0, 0, 0, y')\, \xi_1'\, dy' + \Gamma_2 \int_{-\infty}^{\infty} w(0, 0, L, y')\, \xi_2'\, dy'$$

$$\lambda \xi_2 = -\xi_2 + \Gamma_1 \int_{-\infty}^{\infty} w(L, 0, 0, y')\, \xi_1'\, dy' + \Gamma_2 \int_{-\infty}^{\infty} w(0, 0, L, y')\, \xi_2'\, dy', \tag{A.16}$$

where ξ_i' now denotes $\xi_i(y')$. Subscripts 1 and 2 designate restrictions to $x = 0$ and $x = L$, respectively, according to $\Gamma_1 = 1/|\partial \bar{u}(0)/\partial x|$, $\Gamma_2 = 1/|\partial \bar{u}(L)/\partial x|$ and $\xi_1(y) = \xi(0, y)$, $\xi_2(y) = \xi(L, y)$. Arguing in exactly the same way as when considering annular solutions, we conclude that equation A.15 admits an uncountable infinity of eigenvalues that are the same as those of matrices

$$\begin{bmatrix} -1 + \Gamma_1 \hat{w}_0(\Omega) & \Gamma_2 \hat{w}_L(\Omega) \\ \Gamma_1 \hat{w}_L(\Omega) & -1 + \Gamma_2 \hat{w}_0(\Omega) \end{bmatrix}, \tag{A.17}$$

where $\hat{w}_0(\Omega) = \int_{-\infty}^{\infty} w(0, 0, 0, y') \cos(\Omega y')\, dy' = \int_{-\infty}^{\infty} w(L, 0, L, y') \cos(\Omega y')\, dy'$, $\hat{w}_L(\Omega) = \int_{-\infty}^{\infty} w(L, 0, 0, y') \cos(\Omega y')\, dy' = \int_{-\infty}^{\infty} w(0, 0, L, y') \cos(\Omega y')\, dy'$. To every nonnegative real number corresponds a pair of eigenvalues. The corresponding eigenfunctions can be evaluated from equation A.16.

Acknowledgments

References

Amari, S. I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, *27*, 77–87.

Bressloff, P. (2005). Spontaneous symmetry breaking in self-organizing neural fields. *Biological Cybernetics*, *93*, 256–274.

Camperi, M., & Wang, X. J. (1998). A model of visuospatial working memory in prefrontal cortex: Recurrent network and cellular bistability. *Journal of Computational Neuroscience*, *5*(4), 383–405.

Crandall, M. G., & Rabinowitz, P. H. (1971). Bifurcation from simple eigenvalues. *J. Funct. Anal. 8*, 321–340.

Dong, D. W., & Hopfield, J. J. (1992). Dynamic properties of neural networks with adapting synapses. *Network: Computation in Neural Systems*, *3*(3), 267–283.

Doubrovinski, K. (2005). *Dynamics, stability and bifurcation phenomena in a nonlocal model of cortical activity.* Available online at http://www.matj.uu.se/research/pub/Doubrovinski1.pdf.

Erlhagen, W., & Bicho, E. (2006). The dynamic neural field approach to cognitive robotics. *Journal of Neural Engineering*, *3*, R36–R54.

Erlhagen, W. , & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review*, *109*(3), 545–572.

Ermentrout, G. B. (1998). Neural networks as spatiotemporal pattern forming systems. *Rep. Prog. Phys.*, *61*(4), 353–430.

Ermentrout, G. B., & Cowan, J. D. (1979). A mathematical theory of visual hallucination patterns. *Biological Cybernetics*, *34*, 137–150.

Giese, M. A. (1998). *Dynamic neural field theory for motion perception*. Norwell, MA: Kluwer.

Golubitsky, M., Schaeffer, D., & Stewart, I. (1985). *Singularities and groups in bifurcation theory*. Berlin: Springer.

Gross, H.-M., Stephan, V., & Krabbes, M. (1998). *A neural field approach to topological reinforcement learning in continuous action spaces.* Paper presented at the IEEE World Congress on Computer Intelligence.

Herrmann, J. M., Pawelzik, K., & Geisel, T. (1999). Self-localization of autonomous robots by hidden representations. *Autonomous Robots*, *7*, 31–40.

Herrmann, J. M., Schrobsdorff, H., & Geisel, T. (2004). *Localized activations in a simple neural field model.* Paper presented at the annual Computational Neuroscience Meeting.

Iossifidis, I., & Steinhage, A. (2001). Controlling an 8 DOF manipulator by means of neural fields. In *Proceedings of the International Conference on Field and Service Robotics*.

Jirsa, V. K., Jantzen, K. J., Fuchs, A., & Kelso, J. A. S. (2002). Spatiotemporal forward solution of the EEG and MEG using network modeling. *IEEE Transactions on Medical Imaging*, *21*, 493–504.

Kishimoto, K., & Amari, S. I. (1979). Existence and stability of local excitations in homogeneous neural fields. *Mathematical Biology*, *7*, 303–318.

Kreyszig, E. (1978). *Introductory functional analysis with applications*. New York: Wiley.

Laing, C. R., & Chow, C. (2001). Stationary bumps in networks of spiking neurons. *Neural Computation*, *13*, 1473–1494.

Laing, C. R., & Troy, W. C. (2003). PDE methods for nonlocal problems. *SIAM Journal of Dynamical Systems*, *2*, 487–516.

Laing, C. R., Troy, W. C., Gutkin, B. S., & Ermentrout, G. B. (2002). Multiple bumps in a neuronal model of working memory. *SIAM Journal of Applied Mathematics*, *63*, 62–97.

Lieke, E. E., Frostig, R. D., Arieli, A., Ts'o, D. Y., Hildesheim, R., & Grinvald, A. (1989). Optical imaging of cortical activity: Real-time imaging using extrinsic dye-signals and high resolution imaging based on slow intrinsic-signals. *Annu. Rev. Physiol.*, *51*, 543–559.

Mayer, N., Herrmann, J. M., & Geisel, T. (2002). Curved feature metrics in models of visual cortex. *Neurocomputing*, *44–46*, 533–539.

Pismen, L. M. (1999). *Vortices in nonlinear fields: From liquid crystals to superfluids, from non-equilibrium patterns to cosmic strings.* Oxford: Clarendon Press.

Schierwagen, A., & Werner, H. (1996). Saccade control through the collicular motor map: Two-dimensional neural field model. *Lecture Notes in Computer Science*, *1112*, 439–444.

Schutte, A. R., Spencer, J. P., & Schöner, G. (2003). Testing the dynamic field theory: Working memory for locations becomes more spatially precise over development. *Child Development*, *74*, 1393–1417.

Steinhage, A. (2000). The dynamic approach to anthropomorphic robotics. In *Proceedings of the Fourth Portuguese Conference on Automatic Control, Controlo 2000*.

Strogatz, S. H. (1994). *Nonlinear dynamics and chaos: With applications in physics, biology, chemistry, and engineering*. New York: Perseus Books.

Suder, K., Wörgötter, F., & Wennekers, T. (2001). Neural field model of receptive field restructuring in primary visual cortex. *Neural Computation*, *13*, 139–159.

Takeuchi, A., & Amari, S. I. (1979). Formation of topographic maps and columnar microstructures. *Biological Cybernetics*, *35*, 63–72.

Tanaka, S., Ribot, J., Imamura, K., & Tani, T. (2006). Orientation-restricted continuous visual exposure induces marked reorganization of orientation maps in early life. *NeuroImage*, *30*, 462–477.

Taylor, J. G. (1999). Neural "bubble" dynamics in two dimensions: Foundations. *Biological Cybernetics*, *80*, 393–409.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, *24*, 1–84.

Werner, H., & Richter, T. (2001). Circular stationary solutions in two-dimensional neural fields. *Biological Cybernetics*, *85*, 211–217.

Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, *13*, 55–80.