18

Cortical field models for perception

The world is continuous. Humans walk along corridors and streets, move their arms, turn their head, and orient the direction of gaze. All of these movements and gestures can be described by continuous variables such as position, head direction, gaze orientation, etc. These continuous variables need to be represented in the brain. Field models are designed to encode such continuous variables.

Objects such as houses, trees, cars, pencils have a finite extension in three-dimensional space. Visual input arising from these and other objects is projected onto the retinal photoreceptors and gives rise to a two-dimensional image in the retina. This image is already preprocessed by nerve cells in the retina and undergoes some further processing stages before it arrives in the cortex. A large fraction of the primary visual cortex is devoted to processing of information from the retinal area around the fovea. As a consequence, the activation pattern on the cortical surface resembles a coarse, deformed and distorted image of the object (Fig. 18.1). Topology is largely preserved so that neighboring neurons in the cortex process neighboring points of retinal space. In other words, neighboring neurons have similar receptive fields, which give rise to cortical maps; see Chapter 12.

In this chapter we describe the activity of local groups of cortical neurons. We exploit the fact that neighboring neurons have similar receptive fields so as to arrive at a continuum model of cortical activity. Neural continuum models are often called field models. Neurons in a field model typically receive input on a forward path, from sensory modalities such as vision, audition, or touch, but they also interact with each other. Field models of sensory areas in the cortex are suitable to explain some, but not all, aspects of perception. For example, the lateral interaction of neurons in the visual cortex (and also in the retina!) gives rise to visual phenomena such as Mach bands or grid illusions (Fig. 18.2).

Field models are also used for some forms of working memory. As an example, let us focus on the sense of orientation. Normally, when reading this book for example, you are able to point spontaneously in the direction of the door of your room. Let us imagine that you spin around with closed eyes until you get dizzy and lose your sense of orientation. When you open your eyes your sense of orientation will re-establish itself based on the visual cues in your surroundings. Thus, visual input strongly influences self-orientation. Nevertheless, the sense of orientation also works in complete darkness. If somebody turns the lights off, you still remember where the door is. Thus, humans (and animals) have a

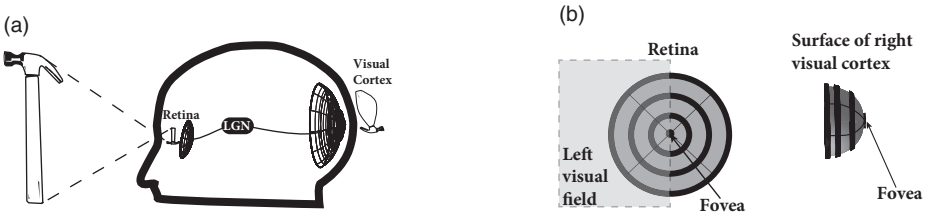

Fig. 18.1 Transformation from visual space to the cortex. (a) The image of an object is projected onto the retina and from there transmitted, via the optic tract and lateral geniculate nucleus (LGN), to the visual cortex. The pattern of activation on the medial surface of visual cortex represents a "distorted" image. (b) Concentric rings on the retina give rise to vertical stripes of activation along the surface of the cortical tissue. A large fraction of neurons in V1 process information from a small region around the fovea in the retina. From Kandel *et al*. (2000).

rather stable "internal compass" which can be used, even in complete darkness, to walk a few steps in the right direction.

While the memory effect of the "internal compass" can be related to the regime of bump attractors in a continuum model, perceptual phenomena can be described by the same class of continuum models, but in the input-driven regime. Both regimes will be discussed in this chapter. We start with a general introduction to continuum models in Section 18.1. We then turn to the input-driven regime in Section 18.2 in order to account for some perceptual and cortical phenomena. Finally we discuss the regime of bump attractors in the context of the sense of orientation (Section 18.3).

18.1 Spatial continuum model

We focus on networks that have a spatial structure. In doing so we emphasize two characteristic features of the cerebral cortex, namely the high density of neurons and its virtually two-dimensional architecture.

Each cubic millimeter of cortical tissue contains more than 10^4 neurons. This impressive number suggests that a description of neuronal dynamics in terms of an averaged *population activity* is more appropriate than a description on the single-neuron level. Furthermore, the cerebral cortex is huge. More precisely, the unfolded cerebral cortex of humans covers a surface of 2200–2400 cm^2, but its thickness amounts on average to only 2.5–3.0 mm. If we do not look too closely, the cerebral cortex can hence be treated as a continuous two-dimensional sheet of neurons. Neurons will no longer be labeled by discrete indices but by continuous variables that give their spatial position on the sheet. The coupling of two neurons i and j is replaced by the *average* coupling strength between neurons at position x and those at position y, or, even more radically simplified, by the average coupling strength of two neurons being separated by the distance $|x-y|$. Similarly to the notion of an average coupling strength we will also introduce the *average activity* of neurons located at position x and describe the dynamics of the network in terms of these averaged quantities only. The details of how these average quantities are defined have been discussed in Chapter 12.

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

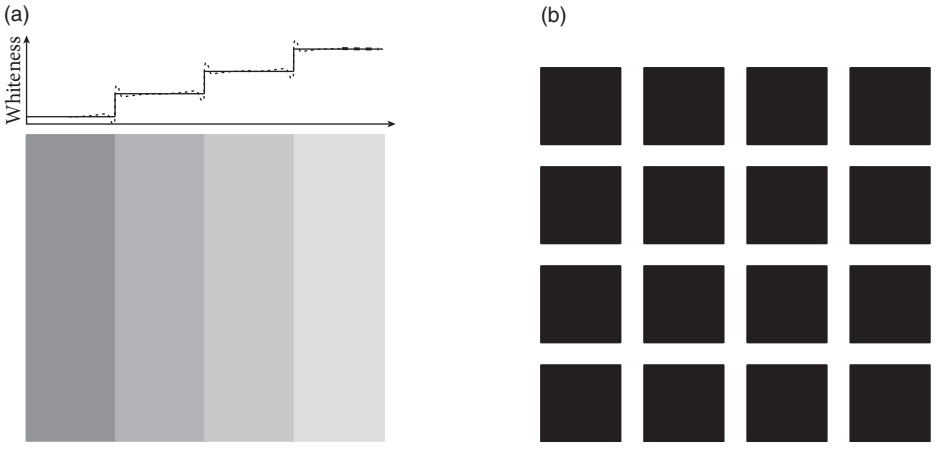

Fig. 18.2 Visual illusions. (a) Mach bands. Gray stripes touching each other are perceived as non-uniform, despite the fact that inside each stripe the gray value is constant. Top: Actual (solid line) and perceived (dashed line) whiteness (vertical axis) as a function of position along the horizontal axis of the image. (b) Helmholtz grid illusion. A grid of black squares on a white background gives rise to scintillating bright or gray dots at the intersections of the white bands.

We now introduce – without a formal justification – field equations for the spatial activity $A(x,t)$ in a spatially extended, but otherwise homogeneous, population of neurons.

We work with the population rate equations that we introduced in Chapter 15. As we have seen, these equations neglect rapid transients and oscillations that can show up in simulations of spiking neurons. On the other hand, in the limit of high noise and short refractoriness, the approximation of population dynamics by differential equations is good; see Chapter 15.

Consider a single sheet of densely packed neurons. We assume that all neurons have similar properties and that the connectivity is homogeneous and isotropic, i.e., that the coupling strength of two neurons is a function of their distance only. We loosely define a quantity $h(x,t)$ as the average input potential at time t of the group of neurons located at position x. We have seen in Chapter 12 that in the stationary state the "activity" of a population of neurons is strictly given by the single-neuron gain function $A_0(x) = F[h(x)]$; see Fig. 18.3. If we *assume* that changes of the input potential are slow enough so that the population always remains in a state of incoherent firing, then we can set

$$A(x,t) = F[h(x,t)]\,, \tag{18.1}$$

even for time-dependent situations. According to Eq. (18.1), the activity $A(x,t)$ of the population around location x is a function of the input potential at that location.

The synaptic input current to a given neuron depends on the level of activity of its presynaptic neurons and on the strength of the synaptic couplings. We assume that the amplitude of the input current is simply the presynaptic activity scaled by the average coupling strength of these neurons. The total input current $I^{\rm syn}(x,t)$ to a neuron at position x is

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

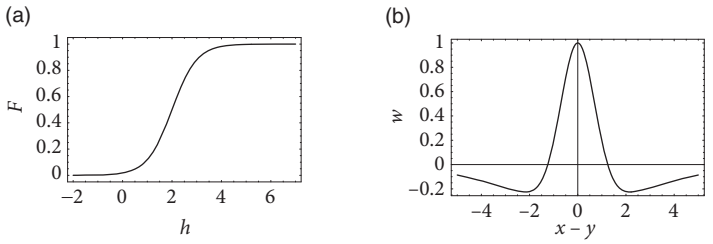

Fig. 18.3 (a) Generic form of the sigmoidal gain function F of graded response neurons that describes the relation between the potential h and the "activity" of the neural population. (b) Typical "Mexican-hat"-shaped function for the coupling w of two neurons as a function of their distance x.

therefore

$$I^{\text{syn}}(x,t) = \int \mathrm{d}y\, w(|x-y|)\, A(y,t)\,. \tag{18.2}$$

Here, w is the average coupling strength of two neurons as a function of their distance.

To complete the definition of the model, we need to specify a relation between the input current and the resulting membrane potential. To keep things simple we treat each neuron as a leaky integrator. The input potential is thus given by a differential equation of the form

$$\tau \frac{\partial h}{\partial t} = -h + RI^{\text{syn}} + RI^{\text{ext}}, \tag{18.3}$$

with τ being the time constant of the integrator and I^{ext} an additional external input. In what follows we work with unit-free variables and set $R = 1$. If we put things together we obtain the field equation

$$\tau \frac{\partial h(x,t)}{\partial t} = -h(x,t) + \int \mathrm{d}y\, w(|x-y|)\, F[h(y,t)] + I^{\text{ext}}(x,t)\,; \tag{18.4}$$

see Wilson and Cowan (1973); Feldman and Cowan (1975); Amari (1977). This is a non-linear integro-differential equation for the average membrane potential $h(x,t)$.

18.1.1 Mexican-hat coupling

In order to be more specific, we now specify the connection strength w as a function of interneuron distance. In what follows we consider a connectivity pattern that is excitatory for proximal neurons and predominantly inhibitory for distal neurons. Figure 18.3b shows the typical "Mexican-hat" shape of the corresponding coupling function.

There are a few simplifying assumptions in the choice of the Mexican-hat function which are worth mentioning. First, Eq. (18.2) assumes that synaptic interaction is instantaneous. In a more detailed model we could include the axonal transmission delay and synaptic time constants. In that case, $A(y,t)$ on the right-hand side of Eq. (18.2) should be replaced by $\int \alpha(s) A(y,t-s)\,\mathrm{d}s$ where $\alpha(s)$ is the temporal interaction kernel.

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

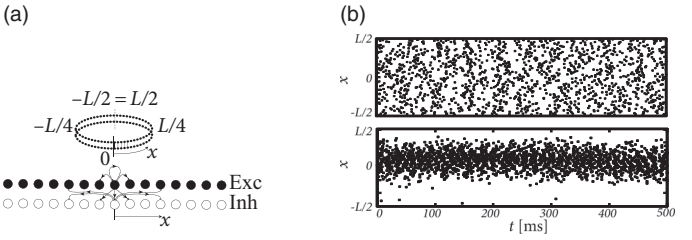

Fig. 18.4 Ring model with spiking neurons. (a) Neurons are organized on a ring, i.e., neurons at position $x = L/2$ are neighbors of neurons at position $x = -L/2$ (top). The dashed vertical line indicates the position where the ring was cut for the unfolded figure at the bottom. At each position x excitatory and inhibitory neurons interact with each other. Inhibitory interactions have a longer range than the excitatory ones. Both neuron types are integrate-and-fire models with identical model parameters and interaction patterns. (b) Spike patterns of sample neurons. In the input-driven regime, activity is spatially homogeneous (top) while for stronger lateral interactions an activation bump forms spontaneously (bump attractor, bottom). Vertical axis shows different neurons, plotted according to their position on the ring. In both cases input is spatially uniform. From Shriki *et al*. (2003) by permission of MIT Press Journals.

Second, in Eq. (18.2) presynaptic neurons at location x can give rise to both excitatory and inhibitory input whereas cortical neurons are either excitatory or inhibitory. The reason is that our population model is meant to be an effective model. There are two alternative ways to arrive at such an effective model. Either we work out the effective coupling model using the mathematical steps of Chapter 16 starting from a common pool of inhibitory neurons shared by all excitatory cells. Or we assume that excitatory and inhibitory neurons at each location receive the same mean input and have roughly the same neuronal parameters (Shriki *et al*., 2003), but have slightly different output pathways (Fig. 18.4a). In both cases, inhibition needs to be of longer range than excitation to arrive at the effective Mexican-hat coupling. A third, and biologically more plausible alternative is discussed in Section 18.2.4.

With Mexican-hat coupling, two different regimes emerge (Fig. 18.4):

(i) In the *input-driven regime* the activity pattern is, in the absence of input, spatially uniform. In other words, any spatial structure in the neuronal activity pattern is causally linked to the input.

(ii) In the *regime of bump attractors*, spatially localized activity patterns develop even in the absence of input (or with spatially uniform input).

The two regimes are discussed in Sections 18.2 and 18.3, respectively.

Example: Ring model with spiking neurons

Excitatory and inhibitory integrate-and-fire models are coupled with a ring-like topology (Fig. 18.4a). Excitatory neurons make connections to their excitatory and inhibitory

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

neighbors. Inhibitory neurons make long-range connections. Let us suppose that the ring model is driven by a spatially uniform input. Depending on the interaction strength (Shriki *et al*., 2003), either a spatially homogeneous spike pattern or a spatially localized spike pattern can emerge (Fig. 18.4b). The first case corresponds to the input-driven regime, the second case to the regime of bump attractors.

18.2 Input-driven regime and sensory cortex models

In this section we study the field equation (18.4) in the input-driven regime. Thus, if the input is spatially uniform, the activity pattern is also spatially uniform. From a mathematical perspective, the spatially uniform activity pattern is the homogeneous solution of the field equation (Section 18.2.1). The stability of the homogeneous solution is discussed in Section 18.2.2.

A non-trivial spatial structure in the input gives rise to deviations from the homogeneous solution. Thus the input drives the formation of spatial activity patterns. This regime can account for perceptual phenomena such as contrast enhancement as shown in Section 18.2.3. Finally, we discuss how the effective Mexican-hat interaction, necessary for contrast enhancement, could be implemented in the cortex with *local* inhibition (Section 18.2.4).

18.2.1 Homogeneous solutions

Although we have kept the above model as simple as possible, the field equation (18.4) is complicated enough to prevent comprehensive analytical treatment. We therefore start our investigation by looking for a special type of solution, i.e., a solution that is uniform over space, but not necessarily constant over time. We call this the homogeneous solution and write $h(x,t) \equiv h(t)$. We expect that a homogeneous solution exists if the external input is homogeneous as well, i.e., if $I^{\text{ext}}(x,t) \equiv I^{\text{ext}}(t)$.

Substitution of the ansatz $h(x,t) \equiv h(t)$ in Eq. (18.4) yields

$$\tau \frac{\mathrm{d}h(t)}{\mathrm{d}t} = -h(t) + \bar{w} F[h(t)] + I^{\text{ext}}(t)\,, \tag{18.5}$$

with $\bar{w} = \int \mathrm{d}y\, w(|y|)$. This is a nonlinear ordinary differential equation for the average input potential $h(t)$. We note that the equation for the homogeneous solution is identical to that of a single population without spatial structure; see Chapter 15.

The fixed points of the above equation with $I^{\text{ext}} = 0$ are of particular interest because they correspond to a resting state of the network. More generally, we search for stationary solutions for a given constant external input $I^{\text{ext}}(x,t) \equiv I^{\text{ext}}$. The fixed points of Eq. (18.5) are solutions of

$$F(h) = \frac{h - I^{\text{ext}}}{\bar{w}}\,, \tag{18.6}$$

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

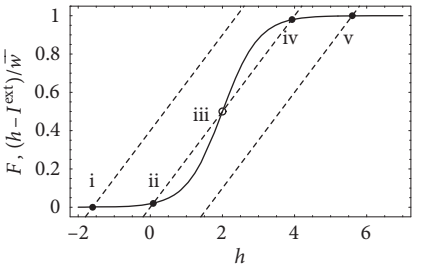

Fig. 18.5 Graphical representation of the fixed-point equation (18.6). The solid line corresponds to the neuronal gain function $F(h)$ and the dashed lines to $(h-I^{\rm ext})/\bar{w}$ for different amounts of external stimulation $I^{\rm ext}$. Depending on the amount of $I^{\rm ext}$ there is either a stable fixed point at low activity (i), a stable fixed point at high activity (v), or a bistable situation with stable fixed points (ii–iv) separated by an unstable fixed point at intermediate level of activity (iii).

which is represented graphically in Fig. 18.5.

Depending on the strength of the external input three qualitatively different situations can be observed. For low external stimulation there is a single fixed point at a very low level of neuronal activity. This corresponds to a quiescent state where the activity of the whole network has ceased. Large stimulation results in a fixed point at an almost saturated level of activity which corresponds to a state where all neurons are firing at their maximum rate. Intermediate values of external stimulation, however, may result in a situation with more than one fixed point. Depending on the shape of the output function and the mean synaptic coupling strength $\bar{w}$, three fixed points may appear. Two of them correspond to the quiescent and the highly activated state, which are separated by the third fixed point at an intermediate level of activity.

Any potential physical relevance of fixed points clearly depends on their stability. Stability under the dynamics defined by the ordinary differential equation (18.5) is readily checked using standard analysis. Stability requires that at the intersection

$$F'(h) < \bar{w}^{-1}\,. \tag{18.7}$$

Thus all fixed points corresponding to quiescent or highly activated states are stable whereas the middle fixed point in the case of multiple solutions is unstable; see Fig. 18.5. This, however, is only half the truth because Eq. (18.5) describes only homogeneous solutions. Therefore, it may well be that the solutions are stable with respect to Eq. (18.5), but unstable with respect to *inhomogeneous* perturbations, i.e., to perturbations that do not have the same amplitude everywhere in the net.

18.2.2 Stability of homogeneous states (*)

In what follows we will perform a linear stability analysis of the homogeneous solutions found in the previous section. Readers not interested in the mathematical details can jump directly to Section 18.2.3.

We study the field equation (18.4) and consider small perturbations about the homogeneous solution. A linearization of the field equation will lead to a linear differential equa-

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

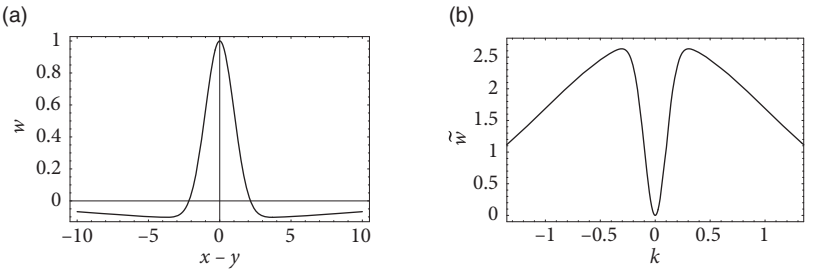

Fig. 18.6 (a) Synaptic coupling function with zero mean as in Eq. (18.14) with $\sigma_1 = 1$ and $\sigma_2 = 10$. (b) Fourier transform of the coupling function shown in (a); see Eq. (18.16).

tion for the amplitude of the perturbation. The homogeneous solution is said to be stable if the amplitude of every small perturbation is *decreasing* whatever its shape.

Suppose $h(x,t) \equiv h_0$ is a homogeneous solution of Eq. (18.4), i.e.,

$$0 = -h_0 + \int \mathrm{d}y\, w(|x-y|)\, F[h_0] + I^{\mathrm{ext}}. \tag{18.8}$$

Consider a small perturbation $\delta h(x,t)$ with initial amplitude $|\delta h(x,0)| \ll 1$. We substitute $h(x,t) = h_0 + \delta h(x,t)$ in Eq. (18.4) and linearize with respect to δh,

$$\tau \frac{\partial}{\partial t}\delta h(x,t) = -h_0 - \delta h(x,t) + \int \mathrm{d}y\, w(|x-y|)\,[F(h_0) + F'(h_0)\,\delta h(y,t)] + I^{\mathrm{ext}}(x,t) + \mathcal{O}(\delta h^2). \tag{18.9}$$

Here, a prime denotes the derivative with respect to the argument. Zero-order terms cancel each other because of Eq. (18.8). If we collect all terms linear in δh we obtain

$$\tau \frac{\partial}{\partial t}\delta h(x,t) = -\delta h(x,t) + F'(h_0) \int \mathrm{d}y\, w(|x-y|)\,\delta h(y,t). \tag{18.10}$$

We make two important observations. First, Eq. (18.10) is *linear* in the perturbations δh – simply because we have neglected terms of order $(\delta h)^n$ with $n \geq 2$. Second, the coupling between neurons at locations x and y is mediated by the coupling kernel $w(|x-y|)$ that depends only on the distance $|x-y|$. If we apply a Fourier transform over the spatial coordinates, the convolution integral turns into a simple multiplication. It suffices therefore to discuss a single (spatial) Fourier component of $\delta h(x,t)$. Any specific initial form of $\delta h(x,0)$ can be created from its Fourier components by virtue of the superposition principle. We can therefore proceed without loss of generality by considering a single Fourier compo-

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

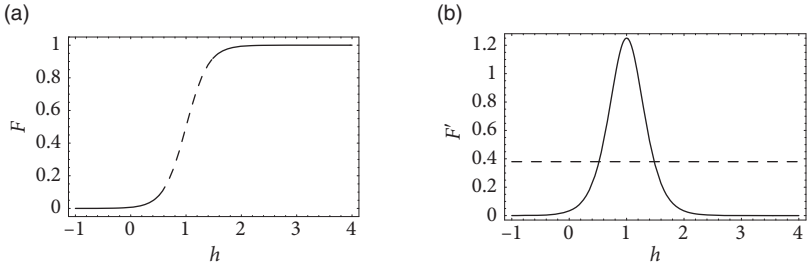

Fig. 18.7 (a) Gain function $F(h) = \{1 + \exp[\beta(x-\theta)]\}^{-1}$ with $\beta = 5$ and $\theta = 1$. The dashed line indicates that part of the graph where the slope exceeds the critical slope s^*. (b) Derivative of the gain function shown in (a) (solid line) and critical slope s^* (dashed line).

nent, namely, $\delta h(x,t) = c(t)\,\mathrm{e}^{ikx}$. If we substitute this ansatz in Eq. (18.10) we obtain

$$\begin{aligned}\tau\, c'(t) &= -c(t)\left[1 - F'(h_0)\int \mathrm{d}y\, w(|x-y|)\,\mathrm{e}^{ik(y-x)}\right] \\ &= -c(t)\left[1 - F'(h_0)\int \mathrm{d}z\, w(|z|)\,\mathrm{e}^{ikz}\right], \end{aligned} \tag{18.11}$$

which is a *linear* differential equation for the amplitude c of a perturbation with wave number k. This equation is solved by

$$c(t) = c(0)\,\mathrm{e}^{-\kappa(k)t}, \tag{18.12}$$

with

$$\kappa(k) = 1 - F'(h_0)\int \mathrm{d}z\, w(|z|)\,\mathrm{e}^{ikz}. \tag{18.13}$$

Stability of the solution h_0 with respect to a perturbation with wave number k depends on the sign of the real part of $\kappa(k)$. Note that – quite intuitively – only two quantities enter this expression, namely the slope of the activation function evaluated at h_0 and the Fourier transform of the coupling function w evaluated at k. If the real part of the Fourier transform of w stays below $1/F'(h_0)$, then h_0 is stable. Note that Eqs. (18.12) and (18.13) are valid for an arbitrary coupling function $w(|x-y|)$. In the following, we illustrate the typical behavior for a specific choice of the lateral coupling.

Example: "Mexican-hat" coupling with zero mean

We describe Mexican-hat coupling by a combination of two bell-shaped functions with different width. For the sake of simplicity we will again consider a one-dimensional sheet of neurons. For the lateral coupling we take

$$w(x) = \frac{\sigma_2\,\mathrm{e}^{-x^2/(2\sigma_1^2)} - \sigma_1\,\mathrm{e}^{-x^2/(2\sigma_2^2)}}{\sigma_2 - \sigma_1}, \tag{18.14}$$

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

with $\sigma_1 < \sigma_2$. The normalization of the coupling function has been chosen so that $w(0) = 1$ and $\int \mathrm{d}x\, w(x) = \bar{w} = 0$; cf Fig. 18.6a.

As a first step we search for a homogeneous solution. If we substitute $h(x,t) = h(t)$ in Eq. (18.4) we obtain

$$\tau \frac{\mathrm{d}h(t)}{\mathrm{d}t} = -h(t) + I^{\mathrm{ext}}. \tag{18.15}$$

The term containing the integral drops out because $\bar{w} = 0$. This differential equation has a single stable fixed point at $h_0 = I^{\mathrm{ext}}$. This situation corresponds to the graphical solution of Fig. 18.5 with the dashed lines replaced by vertical lines ("infinite slope").

We still have to check the stability of the homogeneous solution $h(x,t) = h_0$ with respect to inhomogeneous perturbations. In the present case, the Fourier transform of w,

$$\int \mathrm{d}x\, w(x)\, \mathrm{e}^{ikx} = \frac{\sqrt{2\pi}\,\sigma_1\,\sigma_2}{\sigma_2 - \sigma_1} \left(\mathrm{e}^{-k^2\sigma_1^2/2} - \mathrm{e}^{-k^2\sigma_2^2/2} \right), \tag{18.16}$$

vanishes at $k = 0$ and has its maximum at

$$k_m = \pm \left[\frac{2\ln(\sigma_2^2/\sigma_1^2)}{\sigma_2^2 - \sigma_1^2} \right]^{1/2}. \tag{18.17}$$

At the maximum, the amplitude of the Fourier transform has a value of

$$\hat{w}_m = \max_k \int \mathrm{d}x\, w(x)\, \mathrm{e}^{ikx} = \frac{\sqrt{2\pi}\,\sigma_1\,\sigma_2}{\sigma_2 - \sigma_1} \left[\left(\frac{\sigma_1^2}{\sigma_2^2} \right)^{\frac{\sigma_1^2}{\sigma_2^2 - \sigma_1^2}} - \left(\frac{\sigma_1^2}{\sigma_2^2} \right)^{\frac{\sigma_2^2}{\sigma_2^2 - \sigma_1^2}} \right], \tag{18.18}$$

see Fig. 18.6b. We use this result in Eqs. (18.12) and (18.13) and conclude that stable homogeneous solutions can only be found for those parts of the graph of the output function $F(h)$ where the slope $s = F'(h)$ does not exceed the critical value $s^* = 1/\hat{w}_m$,

$$s^* = \frac{\sigma_2 - \sigma_1}{\sqrt{2\pi}\,\sigma_1\,\sigma_2} \left[\left(\frac{\sigma_1^2}{\sigma_2^2} \right)^{\frac{\sigma_1^2}{\sigma_2^2 - \sigma_1^2}} - \left(\frac{\sigma_1^2}{\sigma_2^2} \right)^{\frac{\sigma_2^2}{\sigma_2^2 - \sigma_1^2}} \right]^{-1}. \tag{18.19}$$

Figures 18.6 and 18.7 show that, depending on the choice of coupling w and gain functions F, a certain interval for the external input exists without a corresponding stable homogeneous solution. In this parameter domain a phenomenon called *pattern formation* can be observed: small fluctuations around the homogeneous state grow exponentially until a characteristic pattern of regions with low and high activity has developed; see Fig. 18.8.

18.2.3 Contrast enhancement

More than a hundred years ago Mach described the psychophysical phenomenon of edge enhancement or contrast enhancement (Mach, 1906): the sharp transition between two

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

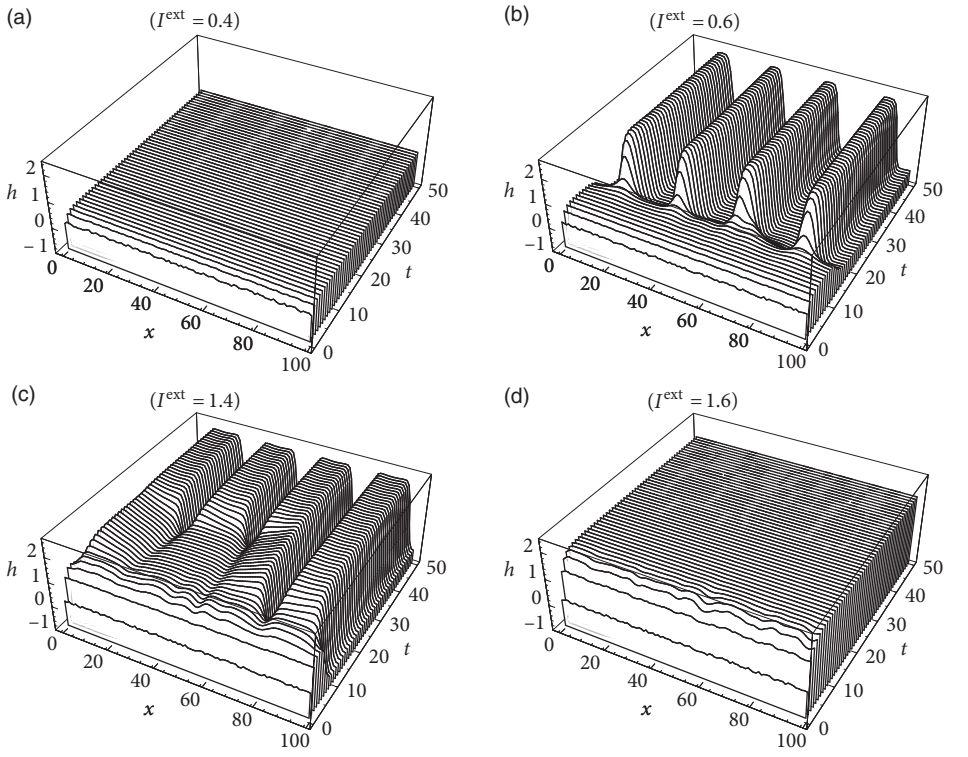

Fig. 18.8 Spontaneous pattern formation in a one-dimensional sheet of neurons with "Mexican-hat" type of interaction and homogeneous external stimulation. The parameters for the coupling function and the output function are the same as in Figs. 18.6–18.7. The graphs show the evolution in time of the spatial distribution of the average membrane potential $h(x,t)$. (a) For $I^{\text{ext}} = 0.4$ the homogeneous low-activity state is stable, but it loses stability at $I^{\text{ext}} = 0.6$. (b) Here, small initial fluctuations in the membrane potential grow exponentially and result in a global pattern of regions with high and low activity. (c) Similar situation to that in (b), but with $I^{\text{ext}} = 1.4$. (d) Finally, at $I^{\text{ext}} = 1.6$ the homogeneous high-activity mode is stable.

regions of different intensities generates perceptual bands along the borders that enhance the perceived intensity difference (Fig. 18.2a). Edge enhancement is already initiated in the retina (Mach, 1865), but likely to have cortical components as well.

Field models with a Mexican-hat interaction kernel generically generate contrast enhancement in the input-driven regime (Fig. 18.9a). Because of the nonlinear lateral interactions, an incoming spatial input pattern is transformed (Wilson and Cowan, 1973; Grossberg, 1973). For example, a spatial input with rectangular profile boosts activity at the borders, while a smooth input with sinusoidal modulation across space boosts activity at the maximum (Fig. 18.9b). A spatial input with a staircase intensity profile generates activity patterns that resemble the perceptual phenomenon of Mach bands.

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

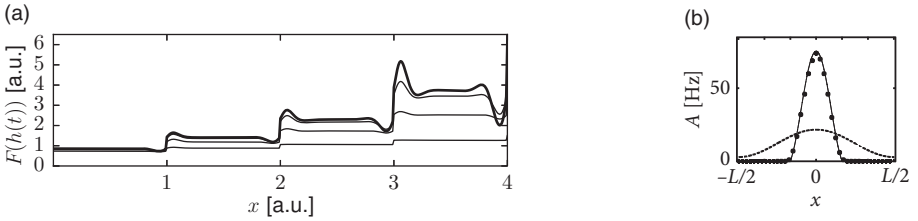

Fig. 18.9 (a) Mach bands in a field model with Mexican-hat coupling. Reflecting Fig. 18.2a, the external current $I^{\text{ext}}(x,t) = I^{\text{ext}}(x)$ forms a staircase as a function of distance. The resulting activity $F(h(t))$ is shown for four different times. The equilibrium solution is indicated by a thick line. (b) An implementation of a field model of excitatory and inhibitory spiking neurons, stimulated with a sinusoidal spatial profile (dashed line) generates a peak at the maximum. From Shriki *et al*. (2003) by permission of MIT Press Journals.

Example: An application to orientation selectivity in V1

Continuum models can represent not only spatial position profiles, but also more abstract variables. For example, ring models have been used to describe orientation selectivity of neurons in the visual cortex (Ben-Yishai *et al*., 1995; Hansel and Sompolinsky, 1998; Shriki *et al*., 2003).

As discussed in Chapter 12, cells in the primary visual cortex (V1) respond preferentially to lines or bars that have a certain orientation within the visual field. There are neurons that "prefer" vertical bars; others respond maximally to bars with a different orientation (Hubel, 1988). Up to now it is still a matter of debate where this orientation selectivity comes from. It may be the result of the wiring of the input to the visual cortex, i.e., the wiring of the projections from the LGN to V1, or it may result from intracortical connections, i.e., from the wiring of the neurons within V1, or both. Here we will investigate the extent to which intracortical projections can contribute to orientation selectivity.

We consider a network of neurons forming a so-called hyper column. These are neurons with receptive fields which correspond to roughly the same zone in the visual field but with different preferred orientations. The orientation of a bar at a given position within the visual field can thus be coded faithfully by the population activity of the neurons from the corresponding hyper column.

Instead of using spatial coordinates to identify a neuron in the cortex, we label the neurons in this section by their preferred orientation θ which may vary from $-\pi/2$ to $+\pi/2$. In doing so we assume that the preferred orientation is indeed a good "name tag" for each neuron so that the synaptic coupling strength can be given in terms of the preferred orientations of presynaptic and postsynaptic neuron. Following the formalism developed in the previous sections, we assume that the synaptic coupling strength w of neurons with preferred orientation θ and θ' is a symmetric function of the difference

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

$\theta - \theta'$, i.e., $w = w(|\theta - \theta'|)$. Since we are dealing with angles from $[-\pi/2, +\pi/2]$ it is natural to assume that all functions are π-periodic so that we can use Fourier series to characterize them. Non-trivial results are obtained even if we retain only the first two Fourier components of the coupling function,

$$w(\theta - \theta') = w_0 + w_2 \cos[2(\theta - \theta')]\,. \tag{18.20}$$

Similarly to the intracortical projections we take the (stationary) external input from the LGN as a function of the difference of the preferred orientation θ and the orientation of the stimulus θ_0,

$$I^{\text{ext}}(\theta) = c_0 + c_2 \cos[2(\theta - \theta_0)]\,. \tag{18.21}$$

Here, c_0 is the mean of the input and c_2 describes the modulation of the input that arises from anisotropies in the projections from the LGN to V1.

Analogously to Eq. (18.4) the field equation for the present setup thus has the form

$$\tau \frac{\partial h(\theta,t)}{\partial t} = -h(\theta,t) + \int_{-\pi/2}^{+\pi/2} \frac{\mathrm{d}\theta'}{\pi}\, w(|\theta - \theta'|)\, F[h(\theta',t)] + I^{\text{ext}}(\theta)\,. \tag{18.22}$$

We are interested in the distribution of the neuronal activity within the hyper column as it arises from a stationary external stimulus with orientation θ_0. This will allow us to study the role of intracortical projections in sharpening orientation selectivity.

In order to obtain conclusive results we have to specify the form of the gain function F. A particularly simple case is the piecewise linear function,

$$F(h) = [h]_+ \equiv \begin{cases} h, & h \geq 0, \\ 0, & h < 0, \end{cases} \tag{18.23}$$

so that neuronal firing increases linearly monotonously once the input potential exceeds a certain threshold.

If we assume that the average input potential $h(\theta,t)$ is always above threshold, then we can replace the gain function F in Eq. (18.22) by the identity function. We are thus left with the following *linear* equation for the stationary distribution of the average membrane potential,

$$h(\theta) = \int_{-\pi/2}^{+\pi/2} \frac{\mathrm{d}\theta'}{\pi}\, w(|\theta - \theta'|)\, h(\theta') + I^{\text{ext}}(\theta)\,. \tag{18.24}$$

This equation is solved by

$$h(\theta) = h_0 + h_2 \cos[2(\theta - \theta_0)]\,, \tag{18.25}$$

with

$$h_0 = \frac{c_0}{1 - w_0} \quad \text{and} \quad h_2 = \frac{2\,c_2}{2 - w_2}\,. \tag{18.26}$$

As a result of the intracortical projections, the modulation h_2 of the response of the

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

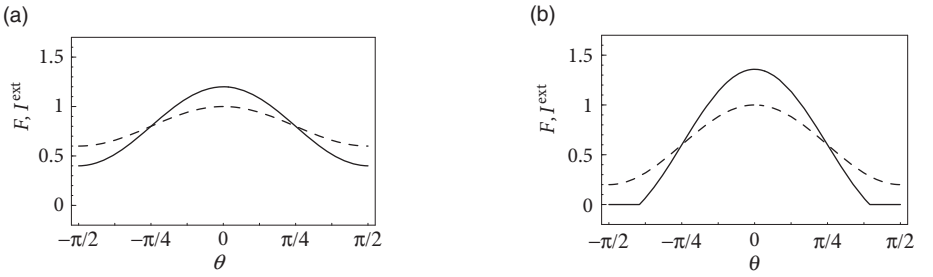

Fig. 18.10 Activity profiles (solid line) that result from stationary external stimulation (dashed line) in a model of orientation selectivity. (a) Weak modulation ($c_0 = 0.8$, $c_2 = 0.2$) of the external input results in a broad activity profile; see Eq (18.24). (b) Strong modulation ($c_0 = 0.6$, $c_2 = 0.4$) produces a narrow profile; see Eq. (18.27). Other parameters are $\omega_0 = 0$, $\omega_2 = 1$, $\theta_0 = 0$.

neurons from the hyper column is thus amplified by a factor $2/(2-w_2)$ compared to the modulation of the input c_2.

In deriving Eq. (18.24) we have assumed that h always stays above threshold so that we have an additional condition, namely, $h_0 - |h_2| > 0$, in order to obtain a self-consistent solution. This condition may be violated depending on the stimulus. In that case the above solution is no longer valid and we have to take the nonlinearity of the gain function into account (Ben-Yishai *et al*., 1995), i.e., we have to replace Eq. (18.24) by

$$h(\theta) = \int_{\theta_0-\theta_c}^{\theta_0+\theta_c} \frac{\mathrm{d}\theta'}{\pi}\, w(|\theta-\theta'|)\, h(\theta') + I^{\mathrm{ext}}(\theta)\,. \qquad (18.27)$$

Here, $\theta_0 \pm \theta_c$ are the cut-off angles that define the interval where $h(\theta)$ is positive. If we use (18.25) in the above equation, we obtain together with $h(\theta_0 \pm \theta_c) = 0$ a set of equations that can be solved for h_0, h_2, and θ_c. Figure 18.10 shows two examples of the resulting activity profiles $F[h(\theta)]$ for different modulation depths of the input.

Throughout this example we have described neuronal populations in terms of an averaged input potential and the corresponding firing rate. At least for stationary input and a high level of noise this is indeed a good approximation of the dynamics of spiking neurons. Figure 18.11 shows two examples of a simulation based on SRM_0 neurons with escape noise and a network architecture that is equivalent to what we have used above. The stationary activity profiles shown in Fig. 18.11 for a network of spiking neurons are qualitatively similar to those of Fig. 18.10 derived for a rate-based model. For low levels of noise, however, the description of spiking networks in terms of a firing rate is no longer valid, because the state of asynchronous firing becomes unstable (see Section 14.2.3) and neurons tend to synchronize (Laing and Chow, 2001).

18.2.4 Inhibition, surround suppression, and cortex models

There are several concerns when writing down a standard field model such as Eq. (18.4) with Mexican-hat interaction. In this section, we aim to move field models closer to biology and consider three of these concerns.

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

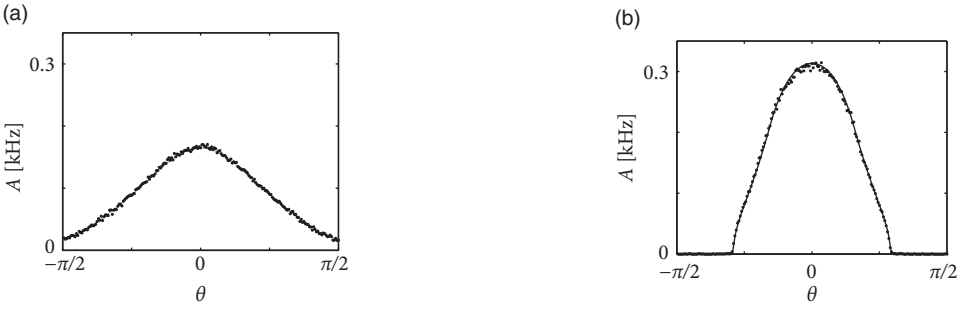

Fig. 18.11 Activity profiles in a model of orientation selectivity obtained by simulations based on SRM_0 neurons (dots) compared to the theoretical prediction (solid line) during stimulation with a low-contrast orientation input at $\theta = 0$. (a) If lateral coupling is not distance-dependent [$\omega_2 = 0$; see Eq. (18.20)] the activity profile reflects the weak modulation of the input pattern. (b) Excitatory coupling between cells of the same orientation and long-range inhibition ($\omega_2 = 10$) generates a sharp activity profile centered at $\theta = 0$. From Spiridon and Gerstner (2001) with permission of Informa Healthcare.

A. Does Mexican-hat connectivity exist in the cortex? The Mexican-hat interaction pattern has a long tradition in theoretical neuroscience (Wilson and Cowan, 1973; Grossberg, 1973; Kohonen, 1984), but, from a biological perspective, it has two major shortcomings. First, in field models with Mexican-hat interaction, the same presynaptic population gives rise to both excitation and inhibition whereas in the cortex excitation and inhibition require separate groups of neurons (Dale's law). Second, inhibition in Mexican-hat connectivity is of longer range than excitation whereas biological data suggests the opposite. In fact, inhibitory neurons are sometimes called local interneurons because they make only local interactions. Pyramidal cells, however, make long-range connections within and beyond cortical areas.

B. Are there electrophysiological correlates of contrast enhancement? Simple and complex cells in the visual cortex respond best if they are stimulated by a slowly moving grating with optimal orientation and of a size that is matched to the cells' receptive field; see Chapter 12. If the grating is optimally oriented but larger than the receptive field, the response is reduced compared to that of a smaller grating (Fig. 18.12). At first sight, this finding is consistent with contrast enhancement through Mexican-hat interaction: a uniform large stimulus evokes a smaller response because it generates inhibition from neurons which are further apart. Paradoxically, however, neurons receive *less* inhibition (Fig. 18.13) with the larger stimulus than with the smaller one (Ozeki *et al*., 2009).

C. How can we interpret the "position" variable in field models? In the previous sections we varied the interpretation of the "space" variable from physical position in the cortex to an abstract variable representing the preferred orientation of cells in the primary visual cortex. Indeed, in the visual cortex several variables need to be encoded in parallel: the location of a neuron's receptive field *and* its preferred orientation *and* potentially its preferred color *and* potentially the relative importance of input from left and right eye – while

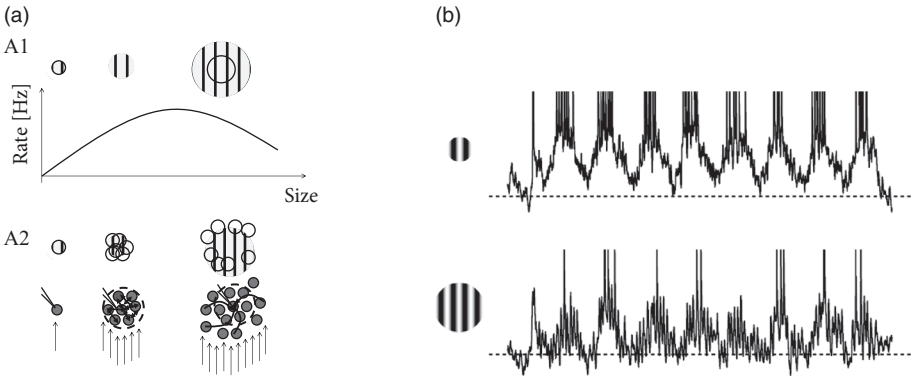

Fig. 18.12 Surrounding suppression. (a) Schematic. A1. Firing rate of a V1 cell as a function of the size of a moving grid stimulus. The grid has optimal orientation and optimal line spacing. Larger grids cause weaker responses than smaller ones. A2. Heuristic interpretation of surrounding suppression. The feedforward pathway from LGN to a cell (arrow, bottom row) gives rise to a small receptive field (RF size and location indicated above cell). Neighboring neurons with overlapping receptive fields excite each other and can be grouped into a local population (dashed circle). If the size of the stimulus is slightly larger, the response of the recorded neuron (middle) is enhanced because of excitatory input from neighboring cells. Right: Distal neurons inhibit the central neuron. Therefore an even larger stimulus suppresses the firing rate of the recorded neuron. (b) Experimental data. A moving grating causes a modulation of the membrane potential and spike firing. The number of spikes and the membrane potential are larger for a small grating than for a bigger one. Dashed horizontal line: mean membrane potential in the absence of stimulation. From Ozeki *et al*. (2009) with permission from Elsevier.

each neuron also has a physical location in the cortex. Therefore a distance-dependent connectivity pattern needs to be distance dependent for several dimensions in parallel while respecting the physical properties of a nearly two-dimensional cortical sheet.

In what follows, we present a model by Ozeki *et al*. (2009) that addresses concerns A and B and enables us to comment on point C.

We group neurons with overlapping receptive fields of similar orientation preference (Fig. 18.12a) into a single population. Inside the population neurons excite each other. We imagine that we record from a neuron in the center of the population. Neurons with receptive fields far away from the recorded neuron inhibit its activity.

Inhibition is implemented indirectly as indicated in Fig. 18.13c. The excitatory neurons in the central population project onto a group of local inhibitory interneurons, but also onto populations of other inhibitory neurons further apart. Each population of inhibitory neurons makes only *local* connections to the excitatory population in their neighborhood. Input to the central group of excitatory neurons therefore induces indirect inhibition of excitatory neurons further apart. Such a network architecture therefore addresses concern A.

To address concern B, the network parameters are set such that the network is in the inhibition-stabilized regime. A network is said to be inhibition-stabilized if the positive feedback through recurrent connections within an excitatory population is strong enough

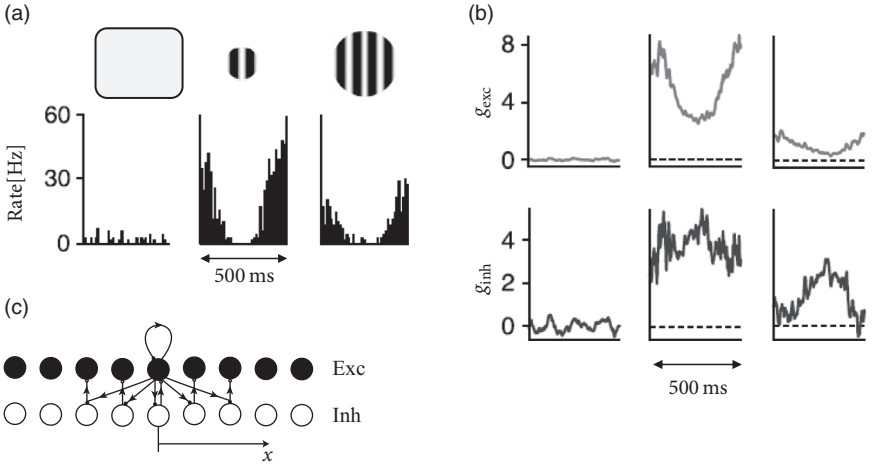

Fig. 18.13 Network stabilized by local inhibition. The schematic model could potentially explain why larger gratings lead not only to less excitatory input g_{exc}, but also to less inhibitory input g_{inh}. (a) The firing rate as a function of the phase of the moving grating for the three stimulus conditions (blank screen, small grating, and large grating). (b) Top: Excitatory input into the cell. Bottom: Inhibitory input into the same cell. As in (a), left, middle, and right correspond to a blank screen, a small grating, and a large grating. Note that the larger grating leads to a reduction of both excitation and inhibition. Adapted from Ozeki *et al.* (2009). (c) Network model with long-range excitation and local inhibition. Excitatory neurons within a local population excite themselves (feedback arrow), and also send excitatory input to inhibitory cells (downward arrows). Inhibitory neurons project to local excitatory neurons.

to cause run-away activity in the absence of inhibition. To counterbalance the positive excitatory feedback, inhibition needs to be even stronger (Tsodyks *et al.*, 1997). As a result, an inhibition-stabilized network responds to a positive external stimulation of inhibitory neurons with a decrease of both excitatory and inhibitory activity (see Exercises).

If the coupling from excitatory populations to neighboring inhibitory populations is stronger than that to neighboring excitatory populations, an inhibition-stabilized network can explain the phenomenon of surrounding suppression *and* at the same time account for the fact that during surrounding suppression both inhibitory and excitatory drive are reduced (Fig. 18.13b). Such a network architecture therefore addresses concern B (Ozeki *et al.*, 2009).

In the above simplified model we focused on populations of neurons with the same preferred orientation, say vertical. However, in the same region of the cortex, there are also neurons with other preferred orientations, such as diagonal or horizontal. The surrounding suppression effect is much weaker if the stimulus in the surroundings has a different orientation than in the central region. We therefore conclude that the cortical connectivity pattern does not simply depend on the physical distance between two neurons, but also on

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

the difference in preferred orientation as well on the neuron type, layer, etc. Therefore, for generalized field models of the primary visual cortex the coupling from a neuron j with receptive field center x_j to a neuron i with receptive field center x_i could be written as

$$w_{ij} = w(x_i, x_j, \theta_i, \theta_j, type_i, type_j, layer_i, layer_j), \tag{18.28}$$

where *type* refers to the type of neuron (e.g., pyramidal, fast-spiking interneuron, non-fast spiking interneuron) and *layer* to the vertical position of the neurons in the cortical sheet. Other variables should be added to account for color preference, binocular preference, etc.

18.3 Bump attractors and spontaneous pattern formation

In this section we study a continuum model with strong recurrent connections such that a spatial activity profile emerges even in cases when the input is homogeneous.

18.3.1 "Blobs" of activity: inhomogeneous states

From a computational point of view bistable systems are of particular interest because they can be used as "memory units." For example, a homogeneous population of neurons with all-to-all connections can exhibit a bistable behavior where either all neurons are quiescent or all neurons are firing at their maximum rate. By switching between the inactive and the active state, the neuronal population is able to represent, store, or retrieve one bit of information. The exciting question that arises now is whether a neuronal net with distance-dependent coupling $w(|x-y|)$ can store more than just a single bit of information, but *spatial patterns* of activity. Sensory input, such as visual stimulation, could switch part of the network to its excited state whereas the unstimulated part would remain in its resting state. Owing to bistability this pattern of activity could be preserved even if the stimulation is turned off again and thus provide a neuronal correlate of working memory.

Let us suppose we prepare the network in a state where neurons in one spatial domain are active and all remaining neurons are quiescent. Will the network stay in that configuration? In other words, we are looking for an "interesting" stationary solution $h(x)$ of the field equation (18.4). The borderline where quiescent and active domains of the network meet is obviously most critical to the function of the network as a memory device. To start with the simplest case with a single borderline, we consider a one-dimensional spatial pattern where the activity changes at $x = 0$ from the low-activity to the high-activity state. This pattern could be the result of inhomogeneous stimulation in the past, but since we are interested in a memory state we now assume that the external input is simply constant, i.e., $I^{\text{ext}}(x,t) = I^{\text{ext}}$. Substitution of $h(x)$ for $h(x,t)$ in the field equation yields

$$h(x) - I^{\text{ext}} = \int \mathrm{d}y \, w(|x-y|) \, F[h(y)] \, . \tag{18.29}$$

This is a nonlinear integral equation for the unknown function $h(x)$.

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

We can find a particular solution of Eq. (18.29) if we replace the output function by a simple step function, for example

$$F(h) = \begin{cases} 0\,, & h < \vartheta, \\ 1\,, & h \geq \vartheta\,. \end{cases} \tag{18.30}$$

In this case $F[h(x)]$ is either zero or 1 and we can exploit translation invariance to define $F[h(x)] = 1$ for $x > 0$ and $F[h(x)] = 0$ for $x < 0$ without loss of generality. The right-hand side of Eq. (18.29) now no longer depends on h and we find

$$h(x) = I^{\text{ext}} + \int_{-\infty}^{x} \mathrm{d}z\, w(|z|)\,, \tag{18.31}$$

and in particular

$$h(0) = I^{\text{ext}} + \frac{1}{2}\,\bar{w}, \tag{18.32}$$

with $\bar{w} = \int \mathrm{d}y\, w(|y|)$. We have calculated this solution under the assumption that $F[h(x)] = 1$ for $x > 0$ and $F[h(x)] = 0$ for $x < 0$. This assumption imposes a self-consistency condition on the solution, namely that the membrane potential reaches the threshold ϑ at $x = 0$. A solution in the form of a stationary border between quiescent and active neurons can therefore only be found if

$$I^{\text{ext}} = \vartheta - \frac{1}{2}\,\bar{w}\,. \tag{18.33}$$

If the external stimulation is either smaller or greater than this critical value, then the border will propagate to the right or to the left.

Following the same line of reasoning, we can also look for a localized "blob" of activity. Assuming that $F[h(x)] = 1$ for $x \in [x_1, x_2]$ and $F[h(x)] = 0$ outside this interval leads to a self-consistency condition of the form

$$I^{\text{ext}} = \vartheta - \int_{0}^{\Delta} \mathrm{d}x\, w(x)\,, \tag{18.34}$$

with $\Delta = x_2 - x_1$. The mathematical arguments are qualitatively the same if we replace the step function by a more realistic smooth gain function.

Figure 18.14 shows that solutions in the form of sharply localized excitations exist for a broad range of external stimulations. A simple argument also shows that the width Δ of the blob is stable if $w(\Delta) < 0$ (Amari, 1977). In this case blobs of activity can be induced without the need for fine tuning the parameters in order to fulfill the self-consistency condition, because the width of the blob will adjust itself until stationarity is reached and Eq. (18.34) holds; see Fig. 18.14a.

18.3.2 Sense of orientation and head direction cells

Head direction cells in the rodent entorhinal cortex are thought to be a neural correlate of the internal compass of rodents. A head direction cell responds maximally if the head of

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

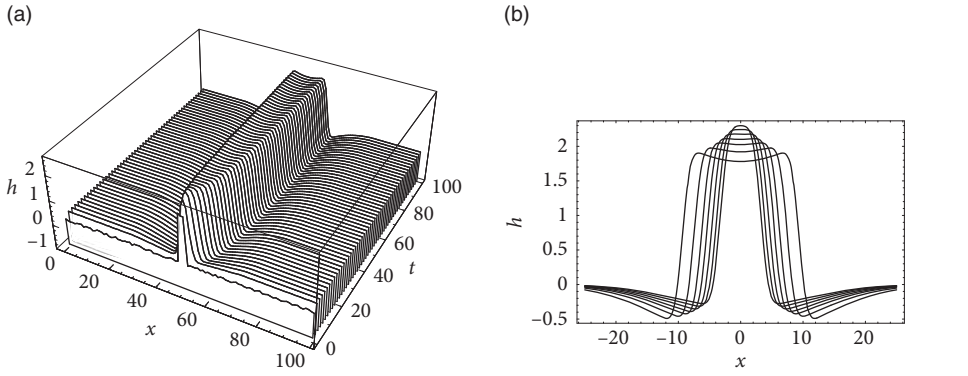

Fig. 18.14 Localized "blobs" of activity. (a) A small initial perturbation develops into a stable blob of activity. (b) Stationary profile of a localized excitation for various amounts of external stimulation ($I^{\text{ext}} = 0, 0.5, \ldots, 0.3$ in order of increasing width). Note that, for strong stimuli, neurons in the center of the activity blob are less active than those close to the edge of the blob.

the animal points in the direction preferred by this cell (Taube and Muller, 1998). Different head direction cells have different receptive fields. The preferred direction of head direction cells depends not on absolute north and south, but rather on visual cues. In a laboratory setting, one specific head direction cell will, for example, always fire when a salient cue card is 60° to the left of the body axis, another one when it is 40° to the right (Fig. 18.15a).

In an experiment during recordings from the entorhinal cortex (Zugaro *et al.*, 2003), a rat was trained to sit still while eating food with its head pointing in the preferred direction of one of the cells in the entorhinal cortex. Recordings from this cell show high activity (Fig. 18.15b). The cell remains active after the light has been switched off, indicating that the rat has memorized its momentary orientation. We emphasize that the orientation of the head can be described by a continuous variable. It is therefore this continuous value which needs to be kept in memory!

During the second phase of the experiment, while lights are switched off, the cues are rotated. Finally, in the last phase of the experiment, lights are switched on again. After a few milliseconds, the internal compass aligns with the new position of the cues. The activity of the recorded cell drops (Fig. 18.15b), because, in the new frame of reference, the head direction is now outside its receptive field (Zugaro *et al.*, 2003). Two aspects are important from a modeling perspective. First, memory content depends on the external input. Second, the memory must be capable of storing a continuous variable – in contrast to the memory models of Chapter 17 where discrete memory items are stored.

In order to represent the continuous head direction variable, ring models have been used (Redish *et al.*, 1996; Zhang, 1996). The location of a neuron on the ring represents its preferred head direction (Fig. 18.16). Cells with similar preferred head direction excite each other whereas cells with opposite preferred head direction inhibit each other. Thus,

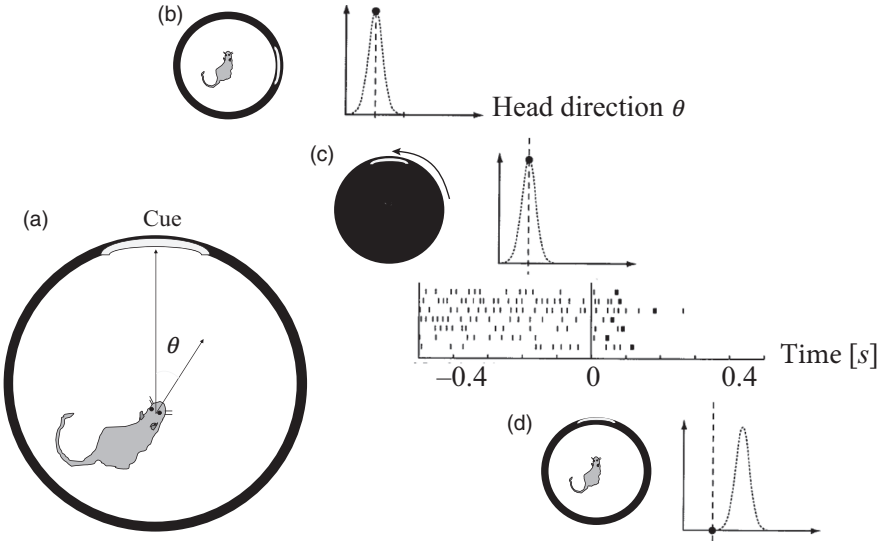

Fig. 18.15 Head direction cells. (a) The head direction θ is defined with respect to salient cues in the environment. (b) Firing rate as a function of θ. The cell responds maximally when the head points in the cell's preferred direction. (c) The cell remains active, even if the lights are switched off. Spike raster shows action potentials of a single head direction cell. For $t < 0$ the room is dark. During the lights-off phase, the cue is manually rotated. At $t = 0$ the light is switched on again, with the cue in the new position. (d) Compared to the new cue position, the head of the rat no longer points in the cell's preferred direction and, after a few milliseconds, the cell becomes silent. Adapted from Zugaro *et al*. (2003).

the ring model has a Mexican-hat connectivity. Parameters are chosen such that the model is in the regime of bump attractors, so that the present value of presumed head direction is kept in memory, even if stimulation stops.

Example: Activation of head direction cells

Head direction cells rely on visual cues, but also on information arising from the acceleration sensors in the inner ears. One of the intriguing open questions is how the sensory information from the inner ear is transformed into updates of the head direction system (Redish *et al*., 1996; Zhang, 1996). A second question is how head direction information can be maintained in the absence of further visual or acceleration input. Here we focus on this second question.

Since the activity of head direction cells does not stop when the light is turned off, interactions in a ring model of head direction must be chosen such that activity bumps are stable in the absence of input. Initially, a strong visual cue fixates the bump of activity at one specific location on the ring. If the visual input is switched off, the bump remains at

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

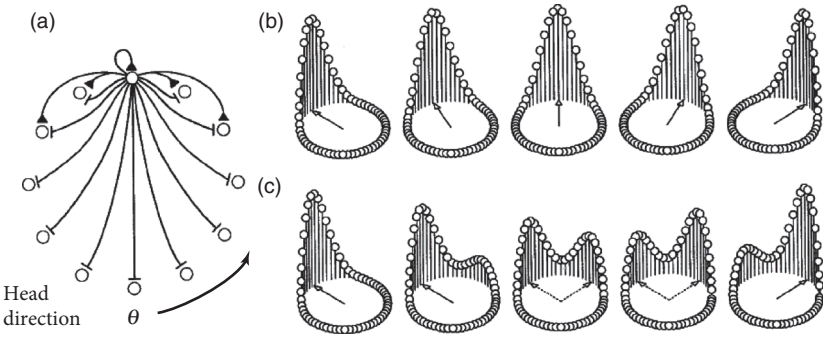

Fig. 18.16 Ring model of head direction cells. (a) Different neurons code for different head directions, visualized by their position on a ring. Neighboring neurons excited each other while neurons at larger distances along the ring inhibit each other. (b) If the input is switched to a new location the activity bump could slowly rotate to the new location particularly if old and new inputs are close to each other. (c) Alternatively, a switch in the input could cause the bump to reappear at the new location; see the experimental data from Fig. 18.15. From Zugaro *et al*. (2003).

its location so that the past head direction is memorized. When a new input with rotated cues is switched on, the activity bump could either rotate towards the new location or disappear at the old and reappear at the new location (Fig. 18.16). For large jumps in the input, the latter seems to be more likely in view of the existing experimental data (Zugaro *et al*., 2003).

18.4 Summary

The cortex is a large, but thin sheet of neurons. Field models, in their spatial interpretation, describe the population activity of neurons as a function of the location on the cortical sheet.

Field models are, however, also used in a more general setting. In sensory cortices, neuronal activity encodes continuous variables, such as position of an object, orientation of edges, direction of movement etc. Field models, in their more abstract interpretation, represent the distribution of activity along the axes representing one or several of these variables.

In field models, interactions between populations of neurons depend on the distance of neurons in the physical, or abstract, space. Classical field models assume a Mexican-hat interaction pattern where local excitation is combined with long-range inhibition. Field models with Mexican-hat interaction have two important regimes of parameter settings. In the input-driven regime, spatial activity patterns can only arise if the input has a non-trivial spatial structure. In the bump-attractor regime, however, localized blobs of activity emerge even in the absence of input.

Mexican-hat interaction combines excitation and inhibition from the same presynaptic

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

population and must therefore be considered as an *effective, mathematical* coupling scheme between neurons. It is, however, possible to construct similar field models with separate populations of excitatory neurons with long-range interactions, and inhibitory neurons with short-range interactions. These models bring the main results of field models a step closer to biological reality.

Exercises

1. **Bump formation**. *Consider a one-dimensional discrete recurrent network with population units* $1 \leq i \leq N$ *and update rule*

$$A_i(t+1) = F\left(\sum_k w_{ik} x_k(t) + \sum_j B_{ij} A_j(t)\right), \tag{18.35}$$

where w_{ik} *is the coupling strength to the input* x_k *and* B_{ij} *are the recurrent weights. Each neuron receives local excitation from its* d *neighbors on both sides,* $B_{ij} = 1$ *for* $|i-j| \leq d$*, and inhibition from all others,* $B_{ij} = -\beta \leq -1$ *for* $|i-j| > d$*. The gain function* $F(h)$ *is the Heaviside step function, i.e.,* $F(h) = 1$ *for* $h > 0$ *and* $F(h) = 0$ *for* $h \leq 0$*.*
(a) Imagine that one single unit is stimulated and therefore becomes active. This neuron will excite its neighbors. Show that in the steady state of the network the number N *of active neurons is larger than* $2d$*.*
Hint: Consider the balance of excitation and inhibition at the border of the blob.
(b) How does the value of β *influence the number of active neurons? What happens in the limit of* $\beta \to 1$*?*
(c) Assume that $d = 5$*,* $N = 1000$ *and the input to neuron* $i = 17$ *is 1. Compute the first three time steps of the network dynamics.*
2. **Stability of homogeneous solution with excitatory coupling.**
(a) Consider the purely excitatory coupling

$$w(x) = \frac{\bar{w}}{\sqrt{2\pi\,\sigma^2}}\, \mathrm{e}^{-x^2/(2\sigma^2)}, \tag{18.36}$$

with the mean strength $\int \mathrm{d}x\, w(x) = \bar{w}$ *and Fourier transform*

$$\int \mathrm{d}x\, w(x)\, \mathrm{e}^{ikx} = \bar{w}\, \mathrm{e}^{-k^2\sigma^2/2}. \tag{18.37}$$

Under what conditions is the homogeneous solution stable (assume $F'(h_0) > 0$*)?*
(b) Consider a general coupling function $w(x)$ *such that this function can be written as an auto-correlation*

$$w(x) = \bar{w} \int_{-\infty}^{\infty} f(x'-x) f(x') \mathrm{d}x', \tag{18.38}$$

for some real function $f(x)$*. Under what conditions is the homogeneous solution stable? Hint: The convolution theorem.*
3. **Phase plane analysis of inhibition-stabilized network**. *An excitatory population is coupled to an inhibitory population, controlled by the activity equations*

$$\begin{aligned} \tau_E \frac{\mathrm{d}A_E}{\mathrm{d}t} &= -A_E + F(w_{\mathrm{EE}} A_{\mathrm{E}} - w_{\mathrm{EI}} A_{\mathrm{I}} + I_{\mathrm{E}}), \\ \tau_I \frac{\mathrm{d}A_I}{\mathrm{d}t} &= -A_I + F(w_{\mathrm{IE}} A_{\mathrm{E}} - w_{\mathrm{II}} A_{\mathrm{I}} + I_{\mathrm{I}} - \vartheta). \end{aligned} \tag{18.39}$$

Assume that $F(h) = 0$ *for* $h < 0$*;* $F(h) = h$ *for* $0 \leq h \leq 1$ *and* $F(h) = 1$ *for* $h > 1$*.*

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press

(a) Draw the nullclines in the phase plane spanned by the variables A_E (x-axis) and A_I (y-axis) in the absence of input $I_E = I_I = 0$ and $\vartheta = 0.5$. Assume that $w_{EE} = w_{EI} = 2$, $w_{IE} = 1$ and $w_{II} = 0$.
(b) Assume that the inhibitory population receives positive input $I_I = 0.2$. Redraw the nullclines. Does the population activity of excitatory or inhibitory populations increase or decrease? Does this correspond to your intuition? Can you interpret the result?

4. **Surrounding inhibition**. *We study the model of the previous exercise in the linear region ($F(h) = h$). Two instantiations $i = 1, 2$ of the model are coupled via an additional connection from $A_{E,1}$ to $A_{i,2}$ and from $A_{E,2}$ to $A_{i,1}$ with lateral connections $w_{\text{lat}} > 0$.*
(a) Assume that the first excitatory population receives an input $I_{E,1} = 0.3$. Calculate the stationary population activity of all four populations.
(b) Assume that both excitatory populations receive an input $I_{E,1} = 0.3 = I_{E,2}$. Calculate the stationary population activity of $A_{E,1}$ and $A_{I,1}$. Do the population activities increase or decrease compared to the case considered in (a)?
(c) Can you interpret your result in the context of the surrounding suppression?

https://doi.org/10.1017/CBO9781107447615.022 Published online by Cambridge University Press