Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Physiology - Paris xxx (2004) xxx–xxx
Journal of Physiology Paris
www.elsevier.com/locate/jphysparis

Pattern formation and cortical maps

Peter Dayan

Gatsby Computational Neuroscience Unit, University College London, Alexandra House, 17 Queen Square, London WCIN 3AR, UK

Abstract

The response selectivities of neurons in adult primary sensory cortices depend on intricate patterns of synaptic connections; these selectivities are arranged over cortex in equally rich fashion. Characterising these patterns, and particularly the activity-dependence (and independence) of their developmental trajectories, has been a major task for experimental and theoretical neuroscience. Here, we describe and analyse a paradigmatic algorithm for activity-dependent development of the refinement and generation of neuronal selectivities, and relate it to some of the wealth of suggestions in the literature.

1. Introduction

Systematic patterns in the connections received by, and the resulting arrangements of, cortical cells, abound in early visual [26,27], auditory [46], and somatosensory [65] neocortex, and many other structures (e.g. the thalamus, [52]). V1 cells, for instance, may be selective to at least least location on the retina, ocularity (favoring input from one eye over the other), and orientation, spatial frequency, direction and speed of movement of bars or gratings. Cells with differing selectivities are laid out in two-dimensional cortex in a labyrinthine manner [28]. These regularities, and ideas about the course of their development, have been the subject of a wealth of computational modelling (see [14,35,53] for some recent reviews).

Two selectivities that have been best characterised in this way are ocularity and orientation. Fig. 1 shows the result of an optical imaging experiment that investigated how they are arranged across a region of the primary visual cortex of a macaque monkey [13,41]. The thick lines show the boundaries of the ocular dominance stripes, showing a part of a characteristic fingerprint-like pattern. The thin lines show iso-orientation contours, i.e. locations where the preferred orientations are roughly constant. These indicate, by the regions they enclose, that whole neighborhoods of cells favor similar orientations, and show how these neighborhoods are arranged with respect to each other and with respect to the ocular dominance stripes. Large scale order is evident, as in the singularities in the orientation map (called pinwheels), where the patches for all orientations meet, and which tend to occur near the centers of ocular dominance stripes, and as in the linear zones, where the iso-orientation domains run parallel to each other, and which tend to occur at, and to run perpendicular to, the boundaries of the ocular dominance stripes. These relationships between orientation and ocular dominance concern local order; the maps are also known in some species such as macaques to have more global order, in the sense that the two-dimensional power spectra of the patterns of each across extended patches of cortex are elliptical, and the major axes of the ellipses for ocular dominance and orientation are orthogonal to each other [2,41].

Abstractly, both the selectivities of individual cells generated by their synaptic inputs, and the manner in which these cells are laid out over cortex can be characterised in terms of patterns. Cortical cells with similar selectivities tend to be nearby, and, conversely, nearby cortical cells have similar selectivities. However, making this absolutely true is impossible, since cells are arranged on an essentially two-dimensional cortical sheet, but are selective in many more dimensions. Actual cortical maps show regularities as to how these two general rules are violated, and it is these regularities that models of the selectivity maps must capture. Some of the regularities arise from basic mathematical facts. For instance, it is impossible to have a continuous map between spaces

E-mail address: dayan@gatsby.ucl.ac.uk (P. Dayan).

0928-4257/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.jphysparis.2004.01.006

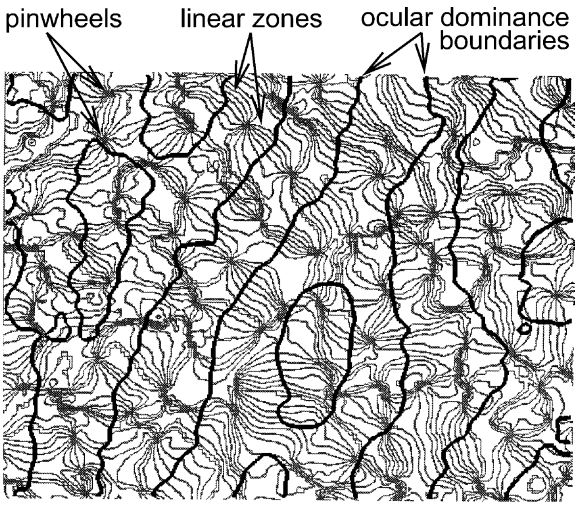

Fig. 1. Orientation domains and ocular dominance. Contour map showing iso-orientation contours (grey lines) and the boundaries of ocular dominance stripes (black lines) in a 1.7 mm patch of macaque primary visual cortex. Iso-orientation contours are drawn in intervals of 11.25°. Pinwheels are singularities in the orientation map where all the orientations meet; linear zones are extended patches over which the iso-orientation contours are parallel. From Obermayer and Blasdel [41], Erwin [13], Dayan and Abbott [9].

with different topological characteristics (called homotopy classes) without having singularities. This constrains the nature of the map of orientation (which lives in a space that is characterised as the unit circle) onto cortex (treated as a two-dimensional sheet). Given that pinwheels are the singularities, it is possible to draw further conclusions about the order in which orientation changes around each [53,55].

Here, we consider such arrangements from a slightly more general perspective of pattern formation, one of the more general concerns of mathematical biology [40]. Indeed, the similarity between the ocular dominance stripes and other biological patterns, such as fingerprints or the stripes on a zebra, has long been noted, and various of the mathematical models employ very similar mechanisms to reaction-diffusion equations, which were first suggested by Turing [57] as a general mechanism for biological pattern formation.

Many of the questions and models focus on the ways that these patterns come about, and the degree to which they can be perturbed by external manipulations. There is an intricate developmental interaction between innate specification and environmental influence, and the development of the maps is determined by both activity-independent and activity-dependent means. A conventional view has been that activity-independent mechanisms control the initial targeting of axons, choose an appropriate layer for them to make connections, and establish a coarse topographic order in projections. Then, other activity-independent and activity-dependent mechanisms refine this order and help to create or boost and preserve the regular receptive fields of individual cells, and the patterns of arrangement of multiple cells, such as ocular dominance stripes (see [66]). However, although activity-dependent development has been a particularly seductive target for modelling because it fits so well with the extensive study of synaptic learning rules, it is likely that a large majority of neural developmental processes are unaffected by activity, and there is an active experimental debate about the true extent of activity-dependence, even for such complex maps as the orientation map (e.g. [4,29,50]).

Note that there is no necessary equivalence between activity-dependence and *environmental* influence or visual experience—the patterns of activity that drive adaptation (perhaps, for instance, waves of activity moving slowly across the retinas of ferrets during early development, [64]) can be created by internal mechanisms and reflect any external milieu only indirectly. Further, very different biophysical mechanisms, even ones that do not involve activity-dependent synaptic plasticity at all, can be characterised mathematically as forming patterns in rather similar ways [61].

Finally, experimental data on the formation of these maps are currently in considerable flux, significantly outpacing most of the models. For instance, such critical factors as the extent to which the development of the pattern of orientation selectivity precedes the development of the pattern of ocular dominance (see Erwin and Miller, 1998), and the relative degrees of innervation of contra-lateral and ipsi-lateral projections during the formation of ocular dominance stripes [5] are only now becoming clear.

One problem with the field of self-organising pattern formation is that there is a wealth of closely related, though not quite identical, models. Here, we consider a variant of a simple and fairly abstract competitive Hebbian model [9,44] for the activity-dependent refinement of topography and development of ocularity. Since this model combines aspects of popular existing suggestions, and yet is analytically tractable (at least in one spatial dimension, which is all we study), it helps make clear some of the critical aspects of pattern formation in such systems. We also consider the relationship between our weight-based model with one of the standard feature-based (i.e. low-dimensional) accounts [30,31].

2. The model

We consider the pattern forming capacity of a reasonably abstract competitive Hebbian model [9,44] to suggest at least some of the critical aspects of many related developmental models. The model specifies how synaptic connections change from an initial, essentially undifferentiated, synaptic state, on the basis of neural input. The model is intended to capture those aspects of

the patterns that might possibly develop on the basis of neural activity, putatively the refinement of a coarse topography (the initial form of which is specified in a quantity called an arbor function) and the creation of monocular cells and ocular dominance stripes. We use conventional analysis techniques (e.g. [35]).

2.1. Architecture

Fig. 2 shows the model. Two input layers (each containing N units), laid out in a single spatial dimension, connect to an output layer (with N units too), also laid out in a single spatial dimension. The input layers represent a one-dimensional version of the eye-specific layers of the lateral geniculate nucleus (labelled 'L' and 'R'); the output layer represents a one-dimensional version of layer IV in the cortex. Labelling the cells as if there is a continuum of them, $\mathbf{W}^{L}_{(\vec{a},\vec{b})}$ and $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$ represent the weight of each connection from the neurons at position $\vec{b}$ in the left and right input layer to the neuron at position $\vec{a}$ in the cortex; $\mathbf{A}_{(\vec{a},\vec{b})}$ represents the multiplicity of each such connection, and is also known as an *arbor function* [35]. The net connection strengths from $\vec{b}$ to $\vec{a}$ are the products $\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}\mathbf{A}_{(\vec{a},\vec{b})}$ and $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}\mathbf{A}_{(\vec{a},\vec{b})}$. Fig. 3(A) shows an example arbor function; Fig. 3(B) shows the final weights $\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ and $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$ for one application of the model.

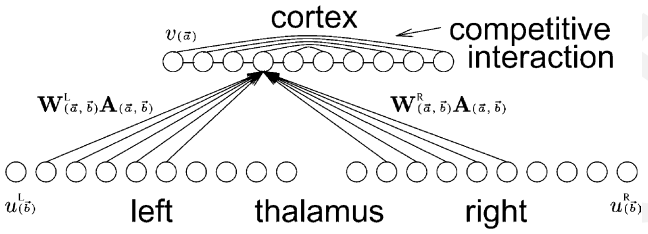

Fig. 2. Competitive ocular dominance model. Left (L) and right (R) input units (with activities $u^{\mathrm{L}}_{(\vec{b})}$ and $u^{\mathrm{R}}_{(\vec{b})}$ at the same location $\vec{b}$ in input space) project through weights ($\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ and $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$) and a restricted topography arbor function $\mathbf{A}_{(\vec{a},\vec{b})}$ to an output layer, which is subject to lateral competitive interactions.

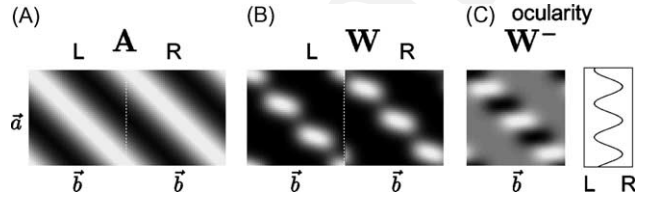

Fig. 3. Ocular dominance patterns from the competitive Hebbian model. (A) Gaussian arbor function $\mathbf{A}_{(\vec{a},\vec{b})}$ showing explicitly the connections from the L and R projections. Toroidal boundary conditions are used to avoid edge effects. (B) Stable weight patterns $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$ showing ocular dominance. (C) (Left) difference in the connections $\mathbf{W}^{-}_{(\vec{a},\vec{b})} = \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} - \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ from right and left eye; (right) sum difference across $\vec{b}$ showing the net ocularity for each $\vec{a}$. Here, $\sigma_A = 0.2$, $\sigma_I = 0.08$, $\sigma_U = 0.075$, $\beta = 10$, $\gamma = 0.95$ and multiplicative normalisation is employed with the weights adding to 3. There are $N = 100$ units in each input layer and the output layer. Circular (toroidal) boundary conditions are also used.

In our one-dimensional model, the monocularity of cortical cells is represented by having $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} - \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ being either all positive (the right projection dominating) or all negative (the left projection dominating) over a single receptive field, i.e. $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} - \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ should have the same sign over $\vec{b}$ for each $\vec{a}$ (though different signs for different $\vec{a}$). Stripes of ocular dominance translate into a pattern of alternating preferences across the cortex as in Fig. 3(B).

Different patterns or modes of synaptic connections grow at different rates, in a way that depends on the parameters and form of the model, and can be approximately characterised near the undifferentiated state using a linear difference (or differential) equation. In many cases, those modes that grow fastest tend to dominate even once non-linear aspects of the models, such as saturation bounds on the strengths of synapses, become important, and so we determine a likely outcome of the complex developmental process from the simple linear difference equation. We can then analyse how the components of the model such as the degree of competition and the restricted topography control stripe width and its dependence on input correlation. There are two different competitive aspects to ocular dominance: competition at the level of single cells for one eye to dominate, i.e. to create monocular cells at all, and second, competition at the level of the maps between input from the two eyes.

2.2. Components

The model has four key characteristics. First is the *arbor* function $\mathbf{A}_{(\vec{a},\vec{b})}$ (Fig. 3(A)) which specifies the basic topography of the map at the time that the pattern of synaptic growth is being established. The arbor function is typically considered to be the product of activity-independent axonal targetting mechanisms, the molecular basis of which is under intense experimental and theoretical investigation [15,18,21,56]. In our model, we consider

$$\mathbf{A}_{(\vec{a},\vec{b})} \propto \mathrm{e}^{-(\vec{a}-\vec{b})^2/2\sigma_A^2} \qquad (1)$$

where σ_A is a parameter specifying the width of the arbor. Two revealing special cases of the arborisation are $\mathbf{A}_{(\vec{a},\vec{b})}$ being constant ($\sigma_A = \infty$), and $\mathbf{A}_{(\vec{a},\vec{b})} \propto \delta(\vec{a} - \vec{b})$ ($\sigma_A = 0$, the case of rigidly specified topography). For some algorithms which explicitly or implicitly model the growth and retraction of axons, the arbor function can change over time in consort with the synaptic weights, for instance if branches of an axonal arbor are stabilised if they support any non-zero synaptic weights. We study the initial formation of patterns, assuming that such

changes do not affect the gross characteristics of the patterns that form.

The second component of the competitive Hebbian model is the nature of the input activity during development. There is evidence about waves of activity moving across the retinæ over development [64] and even about aspects of the activity in the visual thalamus (e.g. [34,58]; however this is far from a full characterisation. For simplicity, we ignore most details (for example, the known differences between ON-center and OFF-center inputs) and consider highly spatially simplified input activities at location $\vec{b}$ in the left ($u^{\mathrm{L}}_{(\vec{b})}$) and right ($u^{\mathrm{R}}_{(\vec{b})}$) reflecting just a single, randomly located, Gaussian bump (of width σ_U) which is stronger to the tune of γ in (a randomly chosen) one of the input projections than the other

$$u^{\mathrm{L}}_{(\vec{b})} = 0.5(1+z\gamma)\mathrm{e}^{-(\vec{b}-\xi)^2/2\sigma_U^2}$$
$$u^{\mathrm{R}}_{(\vec{b})} = 0.5(1-z\gamma)\mathrm{e}^{-(\vec{b}-\xi)^2/2\sigma_U^2} \quad (2)$$

where $\xi \in [0,1)$ is the randomly chosen input location, z is set randomly to -1 or 1 (with probability 0.5 each), and determines whether the input is more from the right or left projection. Parameter $0 \leqslant \gamma \leqslant 1$ governs the weakness of correlations between the projections. Of course, activity patterns must really be substantially more complicated than just a single Gaussian bump.

The third component is the way that input activities and the weights conspire to form output activities. This happens in three steps, the first a linear combination of the inputs through the weights and the arbor

$$\text{linear} \quad v(\vec{a}) = \int \mathrm{d}\vec{b}\,\mathbf{A}_{(\vec{a},\vec{b})}\left(\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}u^{\mathrm{L}}_{(\vec{b})} + \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}u^{\mathrm{R}}_{(\vec{b})}\right) \quad (3)$$

the second involving competition between the output units in which their activities are raised to a power β to sharpen them, and are then normalised

$$\text{competitive} \quad v^{\mathrm{c}}_{(\vec{a})} = (v_{(\vec{a})})^{\beta} \Big/ \int \mathrm{d}\vec{a}'(v(\vec{a}'))^{\beta} \quad (4)$$

and the third involving cortical interaction, in which active units can excite their neighbors

$$\text{interactive} \quad v^{\mathrm{i}}_{(\vec{a})} = \int \mathrm{d}\vec{a}'\,\mathbf{I}_{(\vec{a},\vec{a}')}v^{\mathrm{c}}_{(\vec{a}')} \quad (5)$$

In Eq. (4), $\beta \geqslant 1$ is a parameter governing the strength of competition between the cortical cells. As $\beta \to \infty$, the activation process becomes more strongly competitive, ultimately having a winner-takes-all effect. This is the same sort of idealisation of the equilibrium patterns of activity in the neural activity model [60] that was adopted to good effect in the definition of the self-organising map [30,31]. We will see that the case of $\beta = 1$ is quite closely related to a standard non-competitive model [35]. The separation between the competition and cooperation between the output units in Eqs. (4) and (5) is somewhat artificial, since the same cortical connections presumably instantiate both. However, it qualitatively captures the outcome of some more faithful activation rules. We consider a Gaussian interaction function

$$\mathbf{I}_{(\vec{a},\vec{a}')} = \mathrm{e}^{-(\vec{a}-\vec{a}')^2/2\sigma_I^2} \quad (6)$$

which would generally not produce ocular dominance stripes of a width smaller than the arbor function in non-competitive models [35].

The fourth component is the learning rule governing the change in the weights. This depends on the Hebbian correlation between input and output activities (with contributions such as $\langle v^{\mathrm{i}}_{(\vec{a})}u^{\mathrm{L}}_{(\vec{b})}\rangle_{\xi z}$, averaging over input patterns ξz). We ignore temporal aspects of Hebbian plasticity (Markram et al., 1997; see [9]). Hebbian learning is generally unstable, leading to weights of infinite magnitude, and so normalisation and/or saturation constraints are usually required. We constrain the weights $\mathbf{W}_{(\vec{a},\vec{b})} \in [0,1]$, and normalise the sum of the weights of each postsynaptic cell to be constant

$$\int \mathrm{d}\vec{b}\,\mathbf{A}_{(\vec{a},\vec{b})}\left(\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} + \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}\right) = \Omega \quad (7)$$

in a multiplicative manner

$$\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} \to \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} + \epsilon\left(\langle v^{\mathrm{i}}_{(\vec{a})}u^{\mathrm{L}}_{(\vec{b})}\rangle_{\xi z} - \lambda_{(\vec{a})}\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}\right) \quad (8)$$

where $\lambda_{(\vec{a})} = \lambda_{(\vec{a})}(\mathbf{W}^{\mathrm{L}}, \mathbf{W}^{\mathrm{R}})$ is chosen to enforce equality in Eq. (7). The update equations for $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$ follow similarly. There is an interaction between the upper saturating value for the weights (taken here to be 1) and the value of Ω. For instance, if Ω is too large, then it can become impossible for ocular dominance to develop. Note that only the bottom-up weights are subject to Hebbian plasticity, the intracortical weights defining β and $\mathbf{I}$ are either fixed, or are made to change in systematic ways ($\beta \to \infty$, $\sigma_I \to 0$) over the course of adaptation. Note that in models without activity competition, subtractive normalisation is often considered, because it offers another way to induce the sort of competition that leads to the preferential formulation of ocularity [37]. For this model, in many parameter regimes, it does not make a substantial difference.

The initial conditions for the weights are taken to have the Gaussian form

$$\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} = \omega\mathrm{e}^{-(\vec{a}-\vec{b})^2/2\sigma_W^2} + \eta\delta\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$$
$$\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} = \omega\mathrm{e}^{-(\vec{a}-\vec{b})^2/2\sigma_W^2} + \eta\delta\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} \quad (9)$$

where ω is chosen to satisfy the normalisation constraints, η is small, and $\delta\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ and $\delta\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$ are Gaussian or uniformly distributed random perturbations constrained to satisfy

$$\int \mathrm{d}\vec{b}\,\mathbf{A}_{(\vec{a},\vec{b})}\left(\delta\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} + \delta\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}\right) = 0 \quad (10)$$

for all postsynaptic cells. Given the existence of the arbor function, it is most natural to use $\sigma_W^2 = \infty$ for the initial conditions for the weights, so that each synapse is roughly equally efficacious, although one could also conceive that weights in different regions of an axonal arbor could have different characteristic values. In any case, we will see that it is important to consider values of $\sigma_W^2 < \infty$, since they emerge as equilibrium values of the weights if there is competition ($\beta > 1$) or a restricted arbor ($\sigma_A^2 < \infty$).

It turns out that, for flat topography (and initial conditions, $\sigma_A = \infty$ and $\sigma_W = \infty$) and $\beta = 1$, this model behaves quite like a simple version of a non-competitive Hebbian model [35], with the exception that ocular dominance stripes of a finite width can develop even if $\mathbf{I}$ is purely Gaussian. The key questions are the differences as the model is more competitive (as β gets larger), and as the arbor becomes more peaked (as σ_A gets smaller).

3. Analysis

3.1. Linearisation

The essential analysis technique for this pattern formation problem is to linearise Eq. (8) around the equilibrium values for the weights, and then analyse the effects of perturbations in these values. Since left and right inputs are symmetrical, we consider sum and difference modes

$$\mathbf{W}^{+}_{(\vec{a},\vec{b})} = \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} + \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} \quad \mathbf{W}^{-}_{(\vec{a},\vec{b})} = \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} - \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} \tag{11}$$

and, equivalently,

$$\delta\mathbf{W}^{+}_{(\vec{a},\vec{b})} = \delta\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} + \delta\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$$
$$\delta\mathbf{W}^{-}_{(\vec{a},\vec{b})} = \delta\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} - \delta\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} \tag{12}$$

3.1.1. The sum mode

We first analyse the behavior of the sum mode $\mathbf{W}^{+}_{(\vec{a},\vec{b})}$, because it generally governs the degree of refinement of the topography of the weights, through its equilibrium values. These can then exert a strong influence over the behavior of the difference mode.

The trickiest aspect of the analysis of the sum mode is that the unperturbed initial values of the weights (putting $\eta = 0$ in Eq. (9)) may not be equilibrium points for the full dynamics. If they are not, then the dynamics of changes to the weights will typically exhibit two timescales, a fast one in which the weights change towards the equilibrium values, and a slower one in which more subtle weight changes such ocular dominance occur.

Analysis of the equilibrium values is easy for multiplicative normalisation of the weights. In this case, the equilibrium values of the weights can be found by solving

$$\left\langle v^{\mathrm{i}}_{(\vec{a})} u^{\mathrm{L}}_{(\vec{b})} \right\rangle = \lambda_+ \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} \quad \left\langle v^{\mathrm{i}}_{(\vec{a})} u^{\mathrm{R}}_{(\vec{b})} \right\rangle = \lambda_+ \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} \tag{13}$$

for the λ_+ determined such that the normalisation constraint $\int \mathrm{d}\vec{b}\, \mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} + \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})} = \Omega$ is satisfied for all $\vec{a}$. Although $v_{(\vec{a})}$ is a non-linear function of the weights, with the given input distribution, this implies that the equilibrium values of $\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ and $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$ are the same, determined by

$$\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})} = \omega \mathrm{e}^{-(\vec{a}-\vec{b})^2/2\sigma_W^2} \tag{14}$$

for a particular width σ_W that depends on σ_I, σ_A, σ_U and β according to a simple quadratic equation and a value of ω that depends on the normalisation constraint. We assume that $\omega < 1$, so the weights do not reach their upper saturating limit.

Fig. 4 shows how this equilibrium value of σ_W depends on β, σ_A and σ_I. The solid lines are based on the same parameter values as in Fig. 3 apart from the parameter on the abscissa. Fig. 4(A) shows that the width rapidly asymptotes as β grows, and it only gets large as the arbor function gets large for β near 1. Fig. 4(B) shows this in another way. For $\beta = 1$ (the dashed line), which closely parallels the non-competitive case of Hebbian learning, σ_W grows roughly like the square root of σ_A as the arborisation gets flatter. However, for any $\beta > 1$, one equilibrium value of σ_W has a finite asymptote with σ_A. For absolutely flat topography and $\beta > 1$, there are actually two equilibrium values for σ_W, one with $\sigma_W = \infty$, i.e. flat weights; the other with σ_W taking values such as the asymptotic values for the dotted and solid lines in Fig. 4(B). If the flat equilibrium point is unstable, and the peaked equilibrium point is stable, then topography will be refined over the course of development. For other values of σ_A or β, there is only one equilibrium solution, and it is stable. The stable equilibrium value of σ_W governs the degree of refinement of the final topography.

To assess the stability of the equilibrium solutions, we linearise the solution about each equilibrium point and calculate the resulting eigenvalues and eigenfunctions. Any eigenfunction that grows (in the face of the

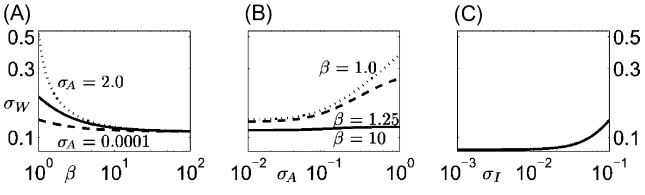

Fig. 4. Log–log plots of the equilibrium values of σ_W in the case of multiplicative normalisation. Solid lines based on parameters as in Fig. 3 ($\sigma_A = 0.2$, $\sigma_I = 0.08$, $\sigma_U = 0.075$, $\beta = 10$). (A) σ_W as a function of β for $\sigma_A = 0.2$ (solid), $\sigma_A = 2.0$ (dotted) and $\sigma_A = 0.0001$ (dashed). (B) σ_W as a function of σ_A for $\beta = 10$ (solid), $\beta = 1.25$ (dashed) and $\beta = 1.0$ (dotted). (C) σ_W as a function of σ_I. Other parameters as for the solid lines.

multiplicative constraint) perturbs the equilibrium. The update equation for perturbations to the sum mode that satisfy the normalisation condition in Eq. (10) is

$$\delta \mathbf{W}^{+}_{(\vec{a},\vec{b})} \rightarrow (1-\epsilon\lambda_{+})\delta \mathbf{W}^{+}_{(\vec{a},\vec{b})} + \epsilon\frac{\beta}{2}\int\int \mathrm{d}\vec{a}_1\,\mathrm{d}\vec{b}_1\,\mathsf{O}_{(\vec{a},\vec{b},\vec{a}_1,\vec{b}_1)}\delta \mathbf{W}^{+}_{(\vec{a}_1,\vec{b}_1)} - \epsilon\lambda'_{(\vec{a})}\mathbf{W}^{+}_{(\vec{a},\vec{b})} \quad (15)$$

where the operator $\mathsf{O} = \mathsf{O}^1 - \mathsf{O}^2$ is defined by

$$\mathsf{O}^1_{(\vec{a},\vec{b},\vec{a}_1,\vec{b}_1)} = \frac{1}{N}\int\int \mathrm{d}\xi\,\mathrm{d}\vec{a}_2\,\mathbf{I}_{(\vec{a},\vec{a}_2)}v^{\mathrm{c}}_{(\vec{a}_2)}\frac{\delta(\vec{a}_1-\vec{a}_2)}{v_{(\vec{a}_1)}}\mathbf{A}_{(\vec{a}_1,\vec{b}_1)}u^{\mathrm{R}}_{(\vec{b}_1)}u^{\mathrm{R}}_{(\vec{b})} \quad (16)$$

$$\mathsf{O}^2_{(\vec{a},\vec{b},\vec{a}_1,\vec{b}_1)} = \frac{1}{N}\int\int \mathrm{d}\xi\,\mathrm{d}\vec{a}_2\,\mathbf{I}_{(\vec{a},\vec{a}_2)}v^{\mathrm{c}}_{(\vec{a}_2)}\frac{v^{\mathrm{c}}_{(\vec{a}_1)}}{v_{(\vec{a}_1)}}\mathbf{A}_{(\vec{a}_1,\vec{b}_1)}u^{\mathrm{R}}_{(\vec{b}_1)}u^{\mathrm{R}}_{(\vec{b})} \quad (17)$$

where the values of $v_{(\vec{a})}$ and $v^{\mathrm{c}}_{(\vec{a})}$ are determined by the values of ξ and $z = 1$ and the unperturbed initial values of the weights ($\eta = 0$), and the values of λ_+ and

$$\lambda'_{(\vec{a})} = \frac{\beta}{2\Omega}\int\int\int \mathrm{d}\vec{b}\,\mathrm{d}\vec{a}_1\,\mathrm{d}\vec{b}_1\,\mathbf{A}_{(\vec{a},\vec{b})}\mathsf{O}_{(\vec{a},\vec{b},\vec{a}_1,\vec{b}_1)}\delta \mathbf{W}^{+}_{(\vec{a}_1,\vec{b}_1)} \quad (18)$$

come from the normalisation condition. Here, λ_+ is determined by $\mathbf{W}^{+}_{(\vec{a},\vec{b})}$ and not by $\delta\mathbf{W}^{+}_{(\vec{a}_1,\vec{b}_1)}$. Except in the special case that $\sigma_A = \infty$, which we discuss below, the term $\epsilon\lambda'_{(\vec{a})}\mathbf{W}^{+}_{(\vec{a},\vec{b})}$ generally keeps stable the equilibrium solution.

We consider the full eigenfunctions of $\mathsf{O}_{(\vec{a},\vec{b},\vec{a}_1,\vec{b}_1)}$ below. However, for flat topography, with both $\mathbf{A}_{(\vec{a},\vec{b})} = 1$ and the unperturbed values of the initial weights $\mathbf{W}_{(\vec{a},\vec{b})} = \omega$ being constant, the operator simplifies to

$$\frac{1}{\sqrt{2}\omega N^2}(\mathrm{e}^{-(\vec{a}-\vec{a}_1)^2/2\sigma_I^2} - \bar{\mathbf{I}})\mathrm{e}^{-(\vec{b}-\vec{b}_1)^2/4\sigma_U^2} \quad (19)$$

where $\bar{\mathbf{I}}$ is the average value of $\mathbf{I}_{(\vec{a},\vec{a}')}$ across $\vec{a}'$ for each $\vec{a}$. The two parts of this operator decouple, and its eigenfunctions are just sines and cosines, i.e. products of real and imaginary parts of

$$\mathbf{W}_{\boldsymbol{k},\boldsymbol{l}}(\vec{a},\vec{b}) = \mathrm{e}^{2\pi \mathrm{i}k\vec{a}} \times \mathrm{e}^{2\pi \mathrm{i}l\vec{b}} \quad (20)$$

where k and l govern the frequencies of variation in the *projective* field of a single input unit and the *receptive* field of a single output unit, respectively. Note that stripes of a finite (albeit overly large) width can form even if $\mathbf{I}$ is purely excitatory (cf. [35]), whatever the strength of the competition (i.e. whatever the value of β).

Remember that the assumption of a continuously sampled system is only an approximation. In the simulations generating figures such as Fig. 3, we treat a finite number of neurons N (so $\vec{a} \in \{1/N, 2/N \ldots 1\}$) and circular boundary conditions (e.g. the neuron with $\vec{a} = 1/N$ is considered to be adjacent to the neurons with $\vec{a} = 2/N$ *and* $\vec{a} = 1$). This means that the continuum of frequencies such as k and l should be replaced by a discrete, quantised set ($k, l \in \{0, 1, 2, 3, \ldots\}$). Furthermore, although the circular boundary conditions make the system translation invariant (so each unit plays exactly the same role), it is not a perfect model for an infinitely large system without circular boundary conditions. In particular, the finite system cannot faithfully represent broad inputs or broad connectivity (i.e. large σ_U, σ_A or σ_W), and this can have significant effects. We typically use a mixed notation with integrals instead of sums, but showing the scaling with the number of units N in an explicit manner.

The eigenvalue of an eigenfunction in Eq. (20) is

$$e_{k,l} = \frac{2\pi\sigma_I\sigma_U}{\omega}(\mathrm{e}^{-\frac{1}{2}\sigma_I^2k^2 4\pi^2} - \delta(k))\mathrm{e}^{-\sigma_U^2 l^2 4\pi^2} \quad (21)$$

If normalisation is imposed, Eq. (10) implies a restriction on the perturbations such that eigenmodes with $l = 0$ are not excited. Their growth would in any case be arrested by the $\epsilon\lambda'_{(\vec{a})}\mathbf{W}^{+}_{(\vec{a},\vec{b})}$ component of Eq. (15). The remaining mode with the largest eigenvalue has $k = 1$, $l = 1$. This mode will grow if

$$\frac{\beta}{2}e_{2\pi,2\pi} > \lambda_+ \quad (22)$$

It turns out for this case that $\lambda_+\omega = \pi\sigma_I\sigma_U$, and so the mode will grow if

$$\beta > \mathrm{e}^{4\pi^2(\sigma_I^2+2\sigma_U^2)/2} \quad (23)$$

If this condition is *not* satisfied, then the flat mode is stable, and topography will not be refined. If this condition *is* satisfied, then the flat mode is unstable. Although the pattern of weights that grows the fastest has $k = 1$, $l = 1$ (see Fig. 5(B)), the terminal pattern of weights (provided that ocular dominance does not form, see the next section) at the peaked, equilibrium, value of σ_W often has more sharply refined topography (Fig. 5(C)). This, more refined, topographical solution is, in general, stable when the flat solution is unstable. Inequality 23 couples the degree of competition and the spatial characteristics of the intracortical connections and the input.

To summarise, for $\sigma_A < \infty$, the only equilibrium solution for the weights has a refined topography, and this is stable. This width depends on the parameters in a way shown in Fig. 4, in particular, reaching a non-zero asymptote even as β gets very large. For a flat arbor ($\sigma_A = \infty$), and $\beta = 1$, the only equilibrium solution has flat weights, and is stable. For $\beta > 1$, there are two equilibrium solutions, one flat, and one with refined topography (the asymptotic values of the curves in Fig. 4(D)). For multiplicative normalisation, for sufficient intracortical competition (β sufficiently large, as judged by Eq. (23)), the flat solution is unstable, and the topography can refine. Altogether, there can therefore

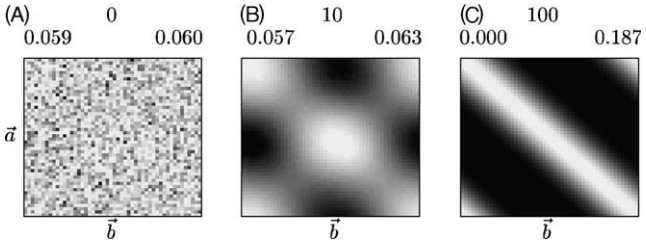

Fig. 5. Topographical refinement given a flat arbor ($\sigma_A = \infty$). The figures show the weights using a grayscale, with the minimum (black) and maximum (white) values above each. (A) Initial weights—there is a very slight bias in favor of diagonal to orient the ultimate solution. (B) Weights after 10 iterations, showing that the $k = 1$, $l = 1$ mode (in the form of $\cos(2\pi\vec{a})\cos(2\pi\vec{b})$) is dominating the unstable growth. (C) Equilibrium weights showing very sharp refinement. The same parameters are used as in Fig. 3, except with $\beta = 5$ and $\gamma = 0.1$ (to prevent the formation of ocular dominance). Also, $\epsilon = 0.1$.

be a qualitative difference between an assumption that the initial arborisation is flat and one that it is even fairly coarsely topographic, at least if there are competitive cortical interactions.

3.1.2. The difference mode

The sum mode mostly controls the refinement of topography, whereas the difference mode controls the development and nature of ocular dominance. Given the simple form of inputs adopted, the development of $\delta\mathbf{W}^-_{(\vec{a},\vec{b})}$ follows almost exactly the same equations as that of the sum mode. The development of ocular dominance requires that a mode of $\delta\mathbf{W}^-_{(\vec{a},\vec{b})} \neq 0$ grows, for which each output cell has weights of only one sign (either positive or negative). The stripe width is determined by changes in this sign across the output layer. The example shown in Fig. 3 is typical for multiplicative normalisation. Fig. 3(C) shows the final value of $\mathbf{W}^-_{(\vec{a},\vec{b})}$ explicitly together with the net ocularity of each output unit.

The main differences between the development of $\delta\mathbf{W}^-_{(\vec{a},\vec{b})}$ and $\delta\mathbf{W}^+_{(\vec{a},\vec{b})}$ given multiplicative normalisation are, first, that the equilibrium value of $\mathbf{W}^-_{(\vec{a},\vec{b})}$ is always 0, independent of the other parameters, since the projections are assumed equivalent. Therefore, we need only consider the linearised dynamics about $\mathbf{W}^-_{(\vec{a},\vec{b})} = 0$. The linearised difference equation is

$$\delta\mathbf{W}^-_{(\vec{a},\vec{b})} \rightarrow (1-\epsilon\lambda_+)\delta\mathbf{W}^-_{(\vec{a},\vec{b})} + \epsilon\frac{\beta\gamma^2}{2}\int\int \mathrm{d}\vec{a}_1\,\mathrm{d}\vec{b}_1\,\mathsf{O}_{(\vec{a},\vec{b},\vec{a}_1,\vec{b}_1)}\delta\mathbf{W}^-_{(\vec{a}_1,\vec{b}_1)} \qquad (24)$$

which is almost the same as Eq. (15) (with the same operator O), except that the multiplier for the integral is $\beta\gamma^2/2$ rather than $\beta/2$. Since $\gamma < 1$, the eigenvalues for the difference mode are therefore all less than those for the sum mode, and by the same fraction. The multiplicative decay term $\epsilon\lambda_+\delta\mathbf{W}^-_{(\vec{a},\vec{b})}$ uses the same λ_+ as Eq. (15), whose value is determined exclusively by properties of $\mathbf{W}^+_{(\vec{a},\vec{b})}$; but the non-multiplicative term $\epsilon\lambda'_{(\vec{a})}\mathbf{W}^+_{(\vec{a},\vec{b})}$ is absent. Note that the equilibrium values of the weights (controlled by σ_W) affect the operator O, and hence its eigenfunctions and eigenvalues.

The second difference is the initial conditions for the perturbations $\delta\mathbf{W}^-_{(\vec{a},\vec{b})}$. Whereas for multiplicative normalisation for the sum mode, the initial perturbations $\delta\mathbf{W}^+_{(\vec{a},\vec{b})}$ must satisfy Eq. (10), there is no such constraint on $\delta\mathbf{W}^-_{(\vec{a},\vec{b})}$.

The simple form of the inputs allows the eigenfunctions of operators O^1 and O^2 to be calculated. Provided that the arbor and the initial values of the weights are not both flat ($\sigma_A \neq \infty$ or $\sigma_W \neq \infty$), the principal eigenfunctions can be shown to have the general form

$$\mathbf{W}^-_{(\vec{a},\vec{b})} = \mathrm{e}^{2\pi ik\vec{a}}\mathrm{e}^{-d^2(\vec{b}-\vec{a})^2+2\pi ilk(\vec{b}-\vec{a})}p_n(\vec{b}-\vec{a},k) \qquad (25)$$

where $p_n(r,k)$ is a polynomial of degree n in r whose coefficients depend on k and d and l are constants that depend on the various parameters of the model. Here k controls the periodicity in the projective field of each input cell $\vec{b}$ to the output cells, and ultimately the periodicity of any ocular dominance stripes that might form. The remaining terms control the receptive fields of the output cells since they depend on $\vec{b}-\vec{a}$ which governs position relative to the center of the receptive field.

Operator O^2 has zero eigenvalues for the polynomials of degree $n > 0$. Unfortunately, the expressions for the non-zero eigenvalues of O^1 and O^2, and also for the coefficients of the polynomials, are too complicated to give here. However, we can use them to predict the outcome of development. Remember that the use of continuous labelling for the neurons is only an approximation and the continuum of frequencies should be replaced by a discrete, quantised set (0, 1, 2,...).

Fig. 6 shows an example of this analysis. The left 5×3 block shows eigenfunctions and eigenvalues of O^1 for $k = 0\ldots5$ and $n = 0, 1, 2$; the middle 5×3 block, the equivalent eigenfunctions and eigenvalues of O^2. As mentioned above, the eigenvalues of O^2 for $n > 0$ are actually 0, and so the operator has a very large null subspace. The eigenfunctions for $n = 0$ look just like conventional Töplitz eigenfunctions, except confined by the equilibrium values of the weights σ_W and the arbor to just the central region. The eigenfunctions for $n = 1$ and $n = 2$ are a little different. When plotted out in the single dimension of $\vec{b}-\vec{a}$, the eigenfunctions look like Gabor functions, whose frequency is set by n, and this is what results in the apparently complicated shapes.

The numbers on top of the eigenfunctions are the eigenvalues. For $n = 0$, they actually have the form of a Gaussian in k (i.e. strictly, just like the discrete Fourier transform of a sampled Gaussian). The Gaussian for O^2 is narrower (though scaled). This is particularly easy to see in Eq. (21), for which the eigenvalues of O^2 have the shape of a delta function in k. To a crude first approx-

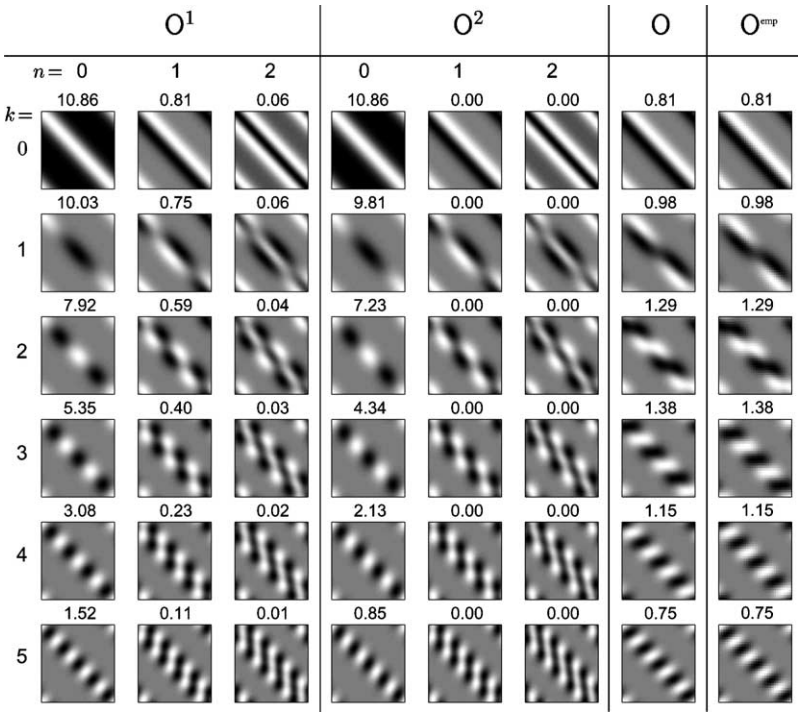

Fig. 6. Eigenfunctions and eigenvalues of O^1 (left block), O^2 (center block), and and the theoretical and empirical approximations to O (right columns). Here, as in Eq. (25), k is the frequency of alternation of ocularity across the output; n is the order of the Hermite polynomial. The numbers on top of each eigenfunction is the associated eigenvalue. Parameters are $\sigma_A = 0.2$, $\sigma_I = 0.08$, $\sigma_U = 0.075$, $\beta = 10$, $\gamma = 1$, $N = 100$.

imation, therefore, the eigenvalues of O resemble the difference of two Gaussians in k, and so have a peak at a non-zero value of k, i.e. a finite ocular dominance periodicity.

However, this first approximation is too crude. Although the eigenfunctions of O^1 and O^2 shown in Fig. 6 *look* almost identical, they are, in fact, subtly different. Indeed, the operators do not commute, making it hard to infer the eigenfunctions and eigenvalues of O from those of O^1 and O^2. Nevertheless, the similarity between the eigenfunctions makes it possible to approximate the eigenfunctions of O very closely by expanding those of O^2 in terms of O^1 (or vice-versa). This only requires knowing the overlap between the eigenfunctions, which can be calculated analytically from their form in Eq. (25). Expanding for $n \leqslant 2$ leads to the approximate eigenfunctions and eigenvalues for O shown in the penultimate column on the right of Fig. 6. The difference, for instance, between the eigenfunction of O for $k = 3$ and those for O^1 and O^2 is striking, considering the similarity between the latter two. In the special cases of flat and rigid topography, O^1 and O^2 *do* commute, and so the eigenfunctions and eigenvalues can be calculated exactly. Just for comparison, the farthest right column shows empirically calculated eigenfunctions and eigenvalues of O (using a 50×50 grid). These are clearly very close (note that there is no reason for the eigenfunctions to have the same spatial phase).

Putting $\delta\mathbf{W}^-$ back in terms of ocular dominance, we require that eigenmodes O resembling the modes with $n = 0$ should grow more strongly than the normalisation makes them shrink; and then the value of k associated with the largest eigenvalue will be the stripe frequency that should be expected to dominate. For the parameters of Fig. 3, the case with $k = 3$ has the largest eigenvalue—and indeed, note how close the outcome of development of $\mathbf{W}^-$ in Fig. 3(C) is to this analytically calculated eigenfunction.

We are now in a position to make qualitative predictions about the outcome of development for any set of parameters, in the face of multiplicative normalisation. First, the analysis of the behavior of the sum mode (including, if necessary, the point about multiple equilibria for flat initial topography) allows a prediction of the equilibrium value of σ_W, which indicates the degree of topographic refinement. Second, this value of σ_W can be used to calculate the value of the normalisation parameter λ_+ that affects the growth of $\delta\mathbf{W}^+$ and $\delta\mathbf{W}^-$. There is then a barrier of $2\lambda_+/\beta\gamma^2$ that the eigenvalues of O must surmount for a solution that is not completely binocular to develop. Third, if the peak eigenvalue of O is indeed sufficiently large that ocular dominance develops, then the favored periodicity is set by the value of k associated with this eigenvalue. Of course, if many eigenfunctions have similarly large eigenvalues, then

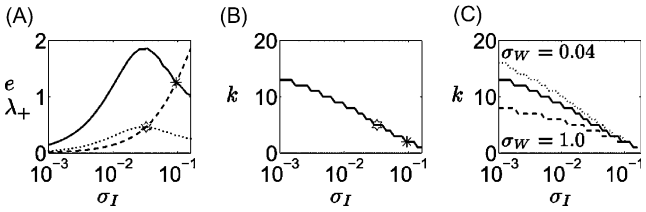

Fig. 7. (A) The constraint term $\lambda_+(\Omega/N)$ (dotted line) and the ocular dominance eigenvalues $e(k)(\Omega/N)$ (solid line $\gamma = 1$; dotted line $\gamma = 0.5$) of $\beta\gamma^2 \mathsf{O}/2$ as a function of σ_I, where k is the stripe frequency associated with the maximum eigenvalue. For σ_I too large, the ocular dominance eigenfunction no longer dominates. The star and hexagon show the maximum values of σ_I such that ocular dominance can form in each case. The scale in (A) is essentially arbitrary. (B) Stripe frequency k associated with the largest eigenvalue as a function of σ_I. The star and hexagon are the same as in (A), showing that the critical preferred stripe frequency is greater for higher correlations between the inputs (lower γ). Only integer values are considered, hence the apparent aliasing. (C) Preferred stripe frequency when σ_W is fixed to $\sigma_W = 0.04$ (dotted line) or $\sigma_W = 1.0$ (dashed line) rather than being determined from the equilibrium state of the sum mode. The solid line is the same as in (B) for comparison.

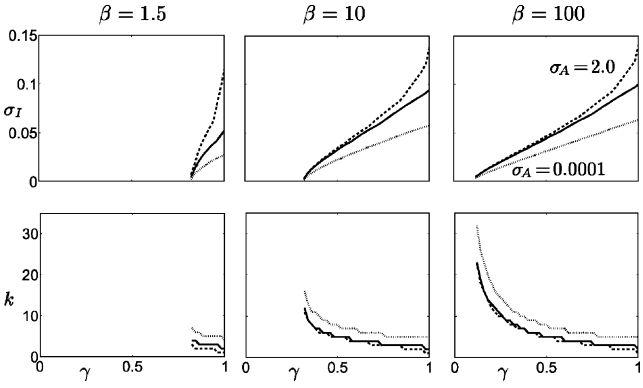

Fig. 8. (Upper row) maximal values of σ_I for which ocular dominance will develop as a function of γ. All other parameters as in Fig. 3, except that $\sigma_A = 0.2$ (solid), $\sigma_A = 2.0$ (dashed); $\sigma_A = 0.0001$ (dotted). (Lower row) value of stripe frequency k associated with the maximal eigenvalue for parameters as in the upper row at the critical value of σ_I.

slightly different stripe periodicities may be observed depending on the initial conditions.

The solid line in Fig. 7(A) shows the largest eigenvalue of $\beta\gamma^2 \mathsf{O}/2$ as a function of the width of the cortical interactions σ_I, for $\gamma = 1$, the value of σ_W specified through the analysis of the sum mode, and values of the other parameters as in Fig. 3. The dashed line shows λ_+, which comes from the normalisation. The largest value of σ_I for which ocular dominance still forms is indicated by the star. For $\gamma = 0.5$, the eigenvalues are reduced by a factor of $\gamma^2 = 0.25$, and so the critical value of σ_I (shown by the hexagram) is reduced. Fig. 7(B) shows the frequency of the stripes associated with the largest eigenvalue. The smaller σ_I, the greater the frequency of the stripes. This and other lines are jagged because only integers are acceptable as stripe frequencies given circular boundary conditions.

Fig. 8 shows the consequences of such relationships in a slightly different way. Some models consider the possibility that σ_I might not be fixed during development, but could change from a large to a small value. If the frequency of the stripes is most strongly determined by the frequency that grows fastest when σ_I is first sufficiently small that stripes grow, we can therefore analyse plots such as those in Fig. 7 to determine the outcome of development. The figures in the top row show the largest values of σ_I for which ocular dominance can develop; the bottom plots show the stripe frequencies associated with these critical values of σ_I (like the stars and hexagons in Fig. 7), in both cases as a function of γ. The columns are for successively larger values of β; within each plot there are three lines, for $\sigma_A = 0.0001$ (dotted); $\sigma_A = 0.2$ (solid), and $\sigma_A = 2.0$ (dashed). Where no value of σ_I permits ocular dominance to form, no line is shown. From the plots, we can see that the more similar the inputs, (the smaller γ) or the less the competition (the smaller β), the harder it is for ocular dominance to form. However, if ocular dominance does form, then the width of the stripes depends only weakly on the degree of competition, and slightly more strongly on the width of the arbors. The narrower the arbor, the larger the frequency of the stripes. For rigid topography, as $\sigma_A \to 0$, the critical value of σ_I depends roughly linearly on γ. We analyse this case in more detail below. Note that the stripe width predicted by the linear analysis does not depend on the correlation between the input projections *unless* other parameters (such as σ_I) change, although ocular dominance might not develop for some values of the parameters.

The last aspect of the solutions that bears comment is the effect of the existence of ocular dominance on the *topography* of the solution. Fig. 9 shows the general pattern of results, in this case using a different set of parameters from those in Fig. 3. Fig. 9(A) shows the final weights from both projections in the same format as Fig. 3(A). Fig. 9(C) shows the net ocular preference in favor of the right eye for these weights. Each output cell is then characterised by the weighted mean location in the left projection (solid lines in Fig. 9(B) and (D)), the right projection (dashed lines) and both projections (dotted lines). Fig. 9(B) shows these mean topographic locations directly (allowing for the wrap-around); Fig. 9(D) shows the difference between these mean locations and a purely linear progression from one end of the output to the other. It is apparent that the topography within a single projection changes most *slowly* when that projection is dominant, and most *quickly* when that projection is suppressed. The topographic locations within the two projections are equal at the most *monocular* regions.

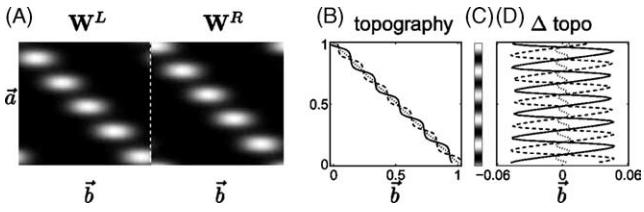

Fig. 9. The effect of ocular dominance on topography. (A) Weight matrix from both input projections ($\sigma_I = 0.04, \sigma_A = 0.5, \sigma_U = 0.075, \beta = 25, \gamma = 0.9$). (B, D) Mean topographic location (B) or difference between the mean topographic location and a purely linear progression (D) for the left input (solid line); right input (dashed line); and both inputs (dotted line). (C) Net ocular bias in favor of the right projection (white is right dominant; black is left dominant). In (B), wrap-around topography is accommodated by identifying $\vec{b} < 0$ with $1 + \vec{b}$ and $\vec{b} > 1$ with $\vec{b} - 1$.

We can now continue the analysis of the two special cases we considered above. First, for flat topography, i.e. $\sigma_A = \infty$, we saw that there are two equilibrium solutions for σ_W. For β that satisfies the inequality of Eq. (23), the equilibrium with flat weights $\sigma_W = \infty$ is unstable, and development proceeds as in the general case, based on the equilibrium solution with finite σ_W. However, for the sum mode, the eigenfunction associated with $k = 1$, $l = 0$ in Eq. (20) is not excited in the initial conditions (because of normalisation), and is, in any case, prevented from growing by the term $\lambda'_{(\vec{a})}\mathbf{W}^+_{(\vec{a},\vec{b})}$ which afflicts the update for the sum mode Eq. (15) but not the difference mode (Eq. (24)). Neither of these is true for the difference mode, and therefore if β satisfies

$$\frac{1}{\gamma^2}\mathrm{e}^{4\pi^2\sigma_I^2/2} < \beta < \mathrm{e}^{4\pi^2(\sigma_I^2+2\sigma_U^2)/2} \tag{26}$$

then ocular dominance will form, with a frequency of $k = 1$, but with completely flat receptive fields, i.e. no topographic refinement.

For rigid topography, as $\sigma_A \to 0$, the operators O^1 and O^2 come to commute, and we can calculate the eigenvalues exactly. Towards this limit, the eigenfunctions of $\beta\gamma^2\mathsf{O}/2$ become the real and complex parts of

$$\mathbf{W}^-_{(\vec{a},\vec{b})} = \mathrm{e}^{2\pi i k\vec{a}} \times \mathrm{e}^{-I\beta U(\vec{b}-\vec{a})^2/\kappa} \tag{27}$$

where

$$I = \frac{1}{2\sigma_I^2} \quad U = \frac{1}{2\sigma_U^2} \quad \kappa = I + \beta(I+U) \tag{28}$$

with eigenvalues (which scale with σ_A) of

$$N\sqrt{2\pi\sigma_A^2}\beta\gamma^2\frac{1}{\Omega}\sqrt{\frac{\pi\beta}{\kappa}}\mathrm{e}^{-(\beta+1)\pi^2k^2/\kappa}\left(1 - \mathrm{e}^{-\mu\pi^2k^2/\kappa}\right) \tag{29}$$

where $\mu = I(1+2\beta)/\beta U$. This peaks for a stripe frequency k that satisfies

$$\mathrm{e}^{-\mu\pi^2k^2/\kappa} = \frac{\beta+1}{\beta+1+\mu} \tag{30}$$

Further, the multiplicative normalisation term λ_+ becomes

$$\lambda_+ = N\sqrt{2\pi\sigma_A^2}\frac{1}{\Omega}\sqrt{\frac{\pi\beta}{\kappa}} \tag{31}$$

The sum mode is uninteresting in the limit of rigid topography, since there is no opportunity for topographic refinement. However, ocular dominance will form when the eigenvalue of Eq. (29) is larger than the normalisation term of Eq. (31).

One interesting limit for rigid topography is that as $\beta \to \infty$, i.e. infinitely sharp competition. In this limit, the largest σ_I such that ocular dominance will form satisfies

$$\sigma_I = \frac{\gamma}{\sqrt{eU}} = \gamma\sqrt{\frac{2}{e}}\sigma_U \tag{32}$$

which is linear in γ (as in the rightmost plot of Fig. 8), and the stripe frequency k that maximises the eigenvalue at the critical σ_I is

$$k = \frac{1}{\pi}\sqrt{I+U} = \frac{1}{2\pi}\sqrt{\left(\frac{e}{\gamma^2}+2\right)\frac{1}{\sigma_U^2}} \tag{33}$$

In this limit, the model behaves, at the equilibrium point, just like the self-organising map [30,31], which we consider below, and that has been extensively investigated in its own right. One complicating factor for sharp competition is that the analysis of pattern formation near the equilibrium solution may not accurately predict the final outcome. Consider what happens to the linear input $v_{(\vec{a})}$ in the artificial case that the weights reflect just a single eigenfunction

$$\begin{aligned}\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{a})} &= \omega + \mu\cos(2\pi k(\vec{a} - 0.5))\\ \mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{a})} &= \omega - \mu\cos(2\pi k(\vec{a} - 0.5))\end{aligned} \tag{34}$$

where $\omega = \Omega/2$, and, for definiteness, the input pattern is specified by Eq. (2) with $\xi = 0.5$ and $z = -1$. The linear aspect of the activation of the output units (Eq. (3)) is

$$v_{(\vec{a})} = \mathrm{e}^{-4\pi^2(\vec{a}-0.5)^2/2\sigma_U^2}(\omega - \mu\gamma\cos(2\pi k(\vec{a} - 0.5))) \tag{35}$$

If the magnitude of the mode satisfies

$$\mu > \frac{\omega}{\gamma(4\pi^2k^2\sigma_U^2 + 1)} \tag{36}$$

then $\vec{a} = 0.5$ is a local *minimum* of $v_{(\vec{a})}$ rather than a local maximum, and, in the fierce competition engendered by large β, $v^e_{(\vec{a})}$ will be large only for $\vec{a} \neq 0.5$, invalidating the analysis in the preceding sections. Further, it can be that the weights to some output units are never substantially altered, if neither they, nor their neighbors ever win the fierce competition. Artificial mechanisms are sometimes used to alleviate this effect, such as a fatigue mechanism that prevents units from winning too

frequently. With smaller values of β, this is less of a concern.

For such artificial, pure, eigenfunctions, the consequence of the lack of activation of units such as that with $\vec{a} = 0.5$ in the case above turns out to be that the fastest growing eigenmodes have a smaller stripe periodicity (i.e. smaller k) than expected from the equilibrium analysis. For smaller correlations between the input projections (i.e. smaller γ), this happens for smaller magnitudes of the modes, and favors smaller k. This effect depends strongly on the non-linear competition, and so is hard to analyse in the case that the weights reflect a sum of many different modes. However, broadly, it is a mechanism by which, the competitive model can generate wider ocular dominance stripes for smaller correlations, even if the width of cortical interactions is not changed. The same effect also happens for broader topography (i.e. $\sigma_A > 0$), but, empirically, depends on β being large. This effect is important [20] since it is one of the pieces of evidence adduced in support of activity-dependent effects.

4. Feature-based models

The form of the terminal weights in Figs. 3(B), (C) and 9 suggests that it is possible to abstract away the details of the synaptic weights for the output units, and instead represent each by two single numbers $\tilde{\mathbf{w}}_{(\vec{a})} = (x_{(\vec{a})}, z_{(\vec{a})})$—one, $x_{(\vec{a})}$, representing the net topographic location of the unit, the other $z_{(\vec{a})}$, representing the net difference in the strength of the connection from the two input projections. Feature-based models take exactly this step, also defining inputs and adaptation rules in terms of these simplified parameters. Here, we show the relationship between Kohonen's self-organising map [17,30,31] and the competitive Hebbian model; related analysis can be performed for another feature-based model called the elastic net [10,11,23]. Links to more abstract feature-based accounts [22,62] are somewhat more obscure.

The self-organising map is a feature-based version of the competitive Hebbian algorithm, copying almost all its characteristics. The inputs presented are characterised in the same two dimensions $\tilde{\mathbf{u}} = (\xi, z\gamma)$ as the weights, as an abstraction of the activities of the input units $u^{\mathrm{L}}_{(\vec{b})}$ and $u^{\mathrm{R}}_{(\vec{b})}$. As in Eq. (2), ξ is the topographic location of the pattern, $z \in \{-1, 1\}$, each with probability 0.5, indicates the ocular preference of the input, and γ is a constant representing the dissimilarity of the two projections. Unlike the feature-based model, the full model of the previous section could accommodate substantially more complex inputs.

The key idea underlying Eq. (3) in the competitive Hebbian model is that cortical interactions should select the output units according to how strongly activated they are by the input, which in turn depends on how closely aligned their weights $\mathbf{W}^{\mathrm{L}}_{(\vec{a},\vec{b})}$ and $\mathbf{W}^{\mathrm{R}}_{(\vec{a},\vec{b})}$ are with the input. In the feature-based model, the alignment is assessed by the discrepancy between the input pattern $\tilde{\mathbf{u}}$ and the weights $\tilde{\mathbf{w}}_{(\vec{a})}$

$$v_{\vec{a}} = \mathrm{e}^{-(\tilde{\mathbf{u}}-\tilde{\mathbf{w}}_{(\vec{a})})^2/2\sigma_U^2} \tag{37}$$

Then, $v^{\mathrm{c}}_{(\vec{a})}$ and $v^{\mathrm{i}}_{(\vec{a})}$ are calculated just as in Eqs. (4) and (5). This implements the competitive aspect of the model.

Similarly, the intuition behind the Hebbian portion of the weight-based learning rule in Eq. (8) is to pull the weights of the activated units towards the input pattern presented, making them respond more strongly to the same input on subsequent presentations. The remaining part of the rule implements multiplicative normalisation to prevent arbitrary weight growth. The feature-based model uses

$$\begin{aligned}\tilde{\mathbf{w}}_{(\vec{a})} &\rightarrow \tilde{\mathbf{w}}_{(\vec{a})} + \epsilon\left(\langle v^{\mathrm{i}}_{(\vec{a})}\tilde{\mathbf{u}}\rangle_{\xi z} - \lambda_{(\vec{a})}\tilde{\mathbf{w}}_{(\vec{a})}\right)\\ &= \tilde{\mathbf{w}}_{(\vec{a})} + \epsilon\langle v^{\mathrm{i}}_{(\vec{a})}(\tilde{\mathbf{u}} - \tilde{\mathbf{w}}_{(\vec{a})})\rangle_{\xi z}\end{aligned} \tag{38}$$

where, in the absence of any equivalent of the sum mode, the multiplicative weight control factor is given by

$$\lambda_{(\vec{a})} = \langle v^{\mathrm{i}}_{(\vec{a})}\rangle_{\xi z} \tag{39}$$

in both cases taking averages over the inputs ξ and z. As for the full model, the cortical interaction term σ_I can either be fixed, or annealed during adaptation. Note that for the case of rigid topography in the full model, the topographic locations in the input of the output units is fixed, since $\mathbf{A}_{(\vec{a},\vec{b})} = \delta_{(\vec{a}-\vec{b})}$ is fixed. For the self-organising map, the topographic location $x_{(\vec{a})}$ can change during development. However, such changes turn out to happen late in development, after the properties of ocular dominance are set.

Now [42,43], treating the N output units as a continuum, we can analyse the development of ocular dominance in this feature-based model by linearising about the equilibrium solution and assessing the growth of ocularity eigenmodes in the direction of $\delta z_{(\vec{a})}$. Since we employ wrap-around boundary conditions, the equilibrium solutions is $\tilde{\mathbf{w}}_{(\vec{a})} = (\vec{a}, 0)$. The similarity between the activation rules for the weight- and feature-based models means that the linearisations are almost the same

$$\delta z_{(\vec{a})} \rightarrow (1 - \epsilon\lambda_+)\delta z_{(\vec{a})} + \epsilon\frac{\beta}{\sigma_U^2}\gamma^2 \int\int \mathrm{d}\vec{a}_1\, \mathsf{O}_{(\vec{a},\vec{a}_1)}\delta z_{(\vec{a}_1)} \tag{40}$$

where

$$\lambda_+ = \sqrt{2\pi\sigma_I^2} \tag{41}$$

and the operator O is defined by a difference of two operators, similarly to Eqs. (16) and (17). The eigenfunctions of $\beta\gamma^2\mathsf{O}/\sigma_U^2$ are the real and complex parts of

$$\delta z_{(\vec{a})} = \mathrm{e}^{2\pi i k\vec{a}} \tag{42}$$

with eigenvalues

$$e_k = \frac{\beta}{\sigma_U^2}\gamma^2\sqrt{2\pi\sigma_I^2}\mathrm{e}^{-4\pi^2k^2\sigma_I^2/2}(1-\mathrm{e}^{-4\pi^2k^2\sigma_U^2/\beta}) \tag{43}$$

which can be directly compared with Eq. (29). The main difference comes because the inputs for the weight-based model are spatially extended. As above, ocular dominance only grows if the maximal $e_k > \lambda_+$, and then there is pressure for the stripe frequency that dominates to be the maximising k.

In the limit as $\beta \to \infty$, which is the normal case for the self-organising map, the same pattern of results as Eqs. (32) and (33) holds. The largest σ_I such that ocular dominance will form satisfies

$$\sigma_I = \gamma\sqrt{\frac{2}{e}} \tag{44}$$

which parallels Eq. (32), and the resulting stripe frequency at the critical σ_I is

$$k = \frac{1}{2\pi}\frac{\sqrt{2}}{\sigma_I} = \frac{1}{2\pi}\sqrt{\frac{e}{\gamma^2}} \tag{45}$$

which parallels Eq. (33). Given slow annealing, this stripe frequency can be expected to dominate the final solution.

Therefore, as for the full model, the width of the stripes does not depend directly on the dissimilarity of the eyes, although they can be coupled if either σ_I is annealed from a sufficiently large value, or if non-linear effects such as those analysed in Eqs. (34)–(36) apply. In the latter case, the most strongly activated output unit can have $\vec{a} \neq \xi$, for input ξ, which perturbs the analysis.

Fig. 10 shows the consequences of applying this algorithm for three values of the dissimilarity parameter γ, and annealing the cortical interaction parameter σ_I to 0. As expected from Eq. (45), the favored stripe frequency is roughly linear in $1/\gamma$, and, as in area 17 of strabismic cats, is smaller, the less similar the projections. The ultimate stripe frequency is slightly larger than predicted from Eq. (45), perhaps because of the development that occurs for values of σ_I smaller than the critical value in Eq. (44). The fierceness of the competition accounts for the small absolute values of γ compared with Fig. 3. The annealing of σ_I makes the final values of $z_{(\vec{a})}$ close to -1 and 1 and creates the strong z-fold topography that the weight-based model lacked (see Fig. 9).

5. Discussion

We have discussed a simple and relatively abstract competitive Hebbian model for activity-dependent aspects of topographic refinement and ocularity development. The model is closely related to other weight-based (e.g. [17,35,47]) and feature-based [31] developmental models, and yet permits fairly extensive analysis, thus, hopefully, casting new light on their workings. We explicitly considered how the parameters of our high-dimensional, weight-based, model relate to those of

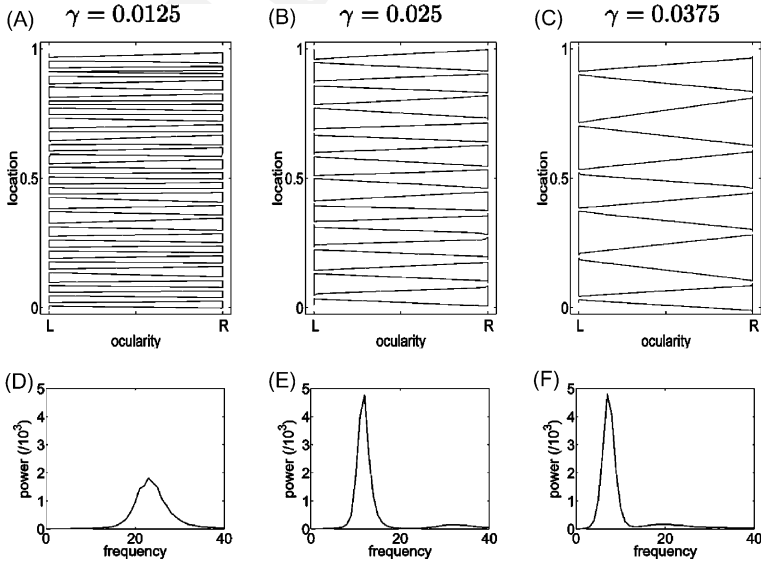

Fig. 10. Feature-based competitive model for ocular dominance and topography for $\gamma = 0.0125$ (A, D), $\gamma = 0.025$ (B, E), and $\gamma = 0.0375$ (C, F). The upper plots show sample maps (in the form of Fig. 3(C)); the lower plots show the average power spectra of the ocularity component. The cortical interactions σ_I are annealed, and the limit of fierce competition $\beta \to \infty$ is considered. The same accommodation for the wrap-around topography is used as in Fig. 9(B).

low-dimensional, feature-based models. We focused exclusively on the simplest properties of V1; ocular dominance stripes do not possess all the topological richness of orientation domains, and so the range of phenomena is more limited (there is no direct equivalent of pinwheel annhilation, for instance, [63]). However, it is to be hoped that such models might capture common principles of cortical plasticity.

The existence in the model of the competition parameter β (varying from essentially linear to maximising) and the topographic arbor width parameter σ_A (varying from flat to rigid) allows us to understand some of the essential pattern forming capacities of various models within the same framework.

Topographic refinement is determined by the properties of the sum mode $\mathbf{W}^+$. Under multiplicative normalisation, if the initial value of σ_W is infinite, refinement happens except for flat topography. Even if the topography is flat, refinement happens if the competition is sufficiently severe that Eq. (23) is satisfied. Note that this does not happen for $\beta = 1$. Empirically, the best studied example of topographic mapping is the retino-tectal system of lower vertebrates. In particular, growth and regrowth of connections are extremely robust to large manipulations, including ablating part of the retina or tectum at various points during development, and crushing, and allowing a regrowth of the optic nerve. In particular, some experiments indicate that the tectum has a form of molecular memory of a pre-existing map that does not depend on particular retino-tectal synapses (which wither following crushing of the optic nerve). One influential model of this (the tea trade model of [33,61], so called because of a rather complicated analogy between topographic maps and the importation, blending and distribution of varieties of tea) is based on molecular concentration gradients rather than synaptically mediated neural activity, but the mechanisms underlying it include ones that are quite closely related to those governing the development of the sum mode in our model.

In our model, the development of ocular dominance is controlled by properties of the difference mode $\mathbf{W}^-$. Under multiplicative normalisation, this depends on the same linear operator as for the sum mode near equilibrium, except that the growth rate is smaller, by a factor of γ^2, which is a measure of the dissimilarity between the input projections. Because of the normalisation of activity inherent in $v^{\mathrm{c}}_{(\vec{a})}$, the operator is a difference of two component linear operators. This difference leads the model to favor the development of periodic ocular dominance stripes, even if the cortical interaction function $\mathbf{I}$ is purely excitatory.

In general, the width of the ocular dominance stripes is determined by the initial instability in the direction of ocular dominance starting from the equilibrium solution of equal connections. This width does not directly depend on γ. However, if the width of cortical interactions is annealed, then the largest value of σ_I at which stripes can first form does depend on γ, and, through this dependence, the less similar the projections, the wider the stripes. A further non-equilibrium effect also tends to enforce the same relation. If the competition is very strong, i.e. β is very large, then as the strength of the stripes grows, there comes a point at which the different eigenfunctions change their relative growth rates, in a way that favors lower frequency modes. This is enhanced for weaker correlations. Therefore, it becomes necessary to consider the effective width of the connections at the time that ocular dominance is developing. Except for very short range cortical interactions, the preferred width of the stripes depends fairly weakly on this effective width. However, the less the correlation between the projections, the more substantial the development that occurs for greater effective widths of cortical interaction, and the wider the resulting stripes.

Many other aspects of pattern formation have also been modelled, and merit study in our competitive Hebbian learning account. Of particular interest is application [19,24] of the elastic net feature-based model to the global structure of the ocular dominance pattern [32], encompassing such phenomena as the perpendicularity of ocular dominance stripes to the boundaries of V1, and the alteration in stripe width in the fovea. Here, we only analysed the case of one spatial dimension. For the global structure, it is critical to extend our analysis to the two spatial dimensions of real cortex. The basic form of the pattern formation analysis in terms of symmetry breaking, linearisation, the effect of intra-cortical competition, *etc* remains the same. However, it is not possible to predict gross structure (for instance even that the ocular dominance stripes are elongated in one direction) merely from the preferred spatial frequency of the pattern, which is all that the linear analysis can be expected to provide. Non-linear analysis must be employed to predict which of the many possible structures with the same preferred frequency of alternation will dominate. This analysis would proceed along similar lines to the seminal recent work of Bressloff, Cowan and their collaborators on hallucinations (see [3]).

Further, although we treated the intracortical weights defining $\mathbf{I}$ as either constant or annealed in effective width, their ultimately patchy structure (e.g. [1,48]), reflecting the patchy selectivities of the associated neurons, suggest that they might also be plastic (e.g. [51]), and this might affect the developing map. Next, in our model and its congeners, the set of possible synaptic connections is fixed at the outset (via the arbor function). In fact, axons and dendrites are quite labile, and it would be interesting to understand the differences, if any, that arise if one takes into account growth, particularly in the three dimensions of cortex [12,39]. Last, but definitely not least, it is natural to assume that the point of activity-dependent plasticity is to allow cortical

selectivities to reflect and track significant statistical structure in the input. This is the focus of computational and statistical unsuperivsed learning algorithms [25]. It has proved difficult to link statistically sound models of the development of selectivity with models of the arrangement of these selectivities on cortex. Is the arrangement computationally epiphenomenal [45], though perhaps anatomically parsimonious [10,38]?

Recent evidence significantly complicates the picture of development of ocular dominance and its relationship with other aspects of the maps such as orientation [4,6,7,29,36,50]. The orientation map (which is also shown in Fig. 1, and has a complicated relationship with the underlying topographic map, [8]) is established extremely early in development and has been shown to remain stable even during the course of the drastic anatomical changes consequent on the development of ocular dominance, during which thalamocortical axonal arbors are undergoing very extensive remodelling. Important aspects of orientation selectivity are stable during monocular deprivation [16]. These results and others suggest the primacy of the orientation map over the map of ocular dominance. Further, there is evidence from kittens that there is an initial bias favoring connections from the contralateral eye over those from the ipsilateral eye that goes away during the time that the stripes are forming, provided that the animals are not deprived [5]. These factors and their ramifications have yet to be fully captured in models.

6. Uncited references

[49,54,59].

Acknowledgements

I am very grateful to Larry Abbott, without whom this paper would exist in a rather different form, Miguel Carreira-Perpiñán, Bard Ermentrout, Ed Erwin, Geoff Goodhill, John Hertz, Ken Miller, Klaus Obermayer, Read Montague, Nick Swindale, Peter Wiesing and David Willshaw for very helpful discussions, and two anonymous reviewers for their suggestions. I also thank Nicolas Brunel, Jean-Pierre Nadal Yves Fregnac and Claude Meunier for inviting me to and hosting Neurosciences et Computation. Funding came from the Gatsby Charitable Foundation.

References

[1] G.G. Blasdel, J.S. Lund, D. Fitzpatrick, Intrinsic connections of macaque striate cortex: axonal projections of cells outside lamina 4C, Journal of Neuroscience 5 (1985) 3350–3369.

[2] G. Blasdel, K. Obermayer, L. Kiorpes, Organization of ocular dominance and orientation columns in the striate cortex of neonatal macaque monkeys, Visual Neuroscience 12 (1995) 589–603.

[3] P.C. Bressloff, J.D. Cowan, M. Golubitsky, P.J. Thomas, M.C. Wiener, Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex, Philosophical Transactions of the Royal Society, Series B 356 (2001) 299–330.

[4] B. Chapman, I. Gödecke, T. Bonhoeffer, Development of orientation preference in the mammalian visual cortex, Journal of Neurobiology 41 (1999) 18–24.

[5] M.C. Crair, D.C. Gillespie, M.P. Stryker, The role of visual experience in the development of columns in cat visual cortex, Science 279 (1998) 566–570.

[6] M.C. Crair, E.S. Ruthazer, D.C. Gillespie, M.P. Stryker, Relationship between the ocular dominance and orientation maps in visual cortex of monocularly deprived cats, Neuron 19 (1997) 307–318.

[7] M.C. Crair, E.S. Ruthazer, D.C. Gillespie, M.P. Stryker, Ocular dominance peaks at pinwheel center singularities of the orientation map in cat visual cortex, Journal of Neurophysiology 77 (1997) 3381–3385.

[8] A. Das, C.D. Gilbert, Distortions of visuotopic map match orientation singularities in primary visual cortex, Nature 387 (1997) 594–598.

[9] P. Dayan, L.F. Abbott, Theoretical Neuroscience, MIT Press, Cambridge, MA, 2001.

[10] R. Durbin, G. Mitchison, A dimension reduction framework for cortical maps, Nature 343 (1990) 644–647.

[11] R. Durbin, D.J. Willshaw, An analogue approach to the travelling salesman problem using an elastic net method, Nature 326 (1987) 689–691.

[12] T. Elliott, N.R. Shadbolt, A mathematical model of activity-dependent, anatomical segregation induced by competition for neurotrophic support, Biological Cybernetics 75 (1996) 463–470.

[13] Erwin, E., 1994. Structure and development in the Macaque primary visual pathway, Ph.D. thesis, University of Illinois, Urbana.

[14] E. Erwin, K. Obermayer, K. Schulten, Models of orientation and ocular dominance columns in the visual cortex: A critical comparison, Neural Computation 7 (1995) 425–468.

[15] J.G. Flanagan, P. Vanderhaeghen, The ephrins and Eph receptors in neural development, Annual Reviews of Neuroscience 21 (1998) 309–345.

[16] I. Gödecke, T. Bonhoeffer, Development of identical orientation maps for two eyes without common visual experience, Nature 379 (1996) 251–254.

[17] G.J. Goodhill, Topography and ocular dominance: a model exploring positive correlations, Biological Cybernetics 69 (1993) 109–118.

[18] G.J. Goodhill, Mathematical guidance for axons, Trends in Neuroscience 21 (1998) 226–231.

[19] G.J. Goodhill, K.R. Bates, P.R. Montague, Influences on the global structure of cortical maps, Proceedings of the Royal Society, Series B 264 (1997) 649–655.

[20] G.J. Goodhill, S. Löwel, Theory meets experiment: correlated neural activity helps determine ocular dominance column periodicity, Trends in Neurosciences 18 (1995) 437–439.

[21] G.J. Goodhill, L.J. Richards, Retinotectal maps: molecules, models and misplaced data, Trends in Neuroscience 22 (1999) 529–534.

[22] G.J. Goodhill, T.J. Sejnowski, A unifying objective function for topographic mappings, Neural Computation 9 (1996) 1291–1303.

[23] G.J. Goodhill, D.J. Willshaw, Application of the elastic net algorithm to the formation of ocular dominance stripes, Network: Computation in Neural Systems 1 (1990) 41–61.

[24] G.J. Goodhill, D.J. Willshaw, Elastic net model of ocular dominance: overall stripe pattern and monocular deprivation, Neural Computation 6 (1994) 615–621.
[25] G.E. Hinton, T.J. Sejnowski, Unsupervised Learning, MIT Press, Cambridge, MA, 1999.
[26] D.H. Hubel, Eye, Brain, and Vision, WH Freeman, New York, NY, 1988.
[27] D.H. Hubel, T.N. Wiesel, Functional architecture of macaque monkey visual cortex, Proceedings of the Royal Society of London, Series B: Biological Sciences B 198 (1977) 1–59.
[28] M. Hubener, D. Shoham, A. Grinvald, T. Bonhoeffer, Spatial relationships among three columnar systems in cat area 17, Journal of Neuroscience 17 (1997) 9270–9284.
[29] L.C. Katz, J.C. Crowley, Development of cortical circuits: lessons from ocular dominance columns, Nature Reviews Neuroscience 3 (2002) 34–42.
[30] T. Kohonen, Self-organised formation of topologically correct feature maps, Biological Cybernetics 43 (1982) 59–69.
[31] T. Kohonen, Self-Organizing Maps, Springer-Verlag, Berlin, New York, 1995.
[32] S. LeVay, M. Connolly, J. Houde, D.C. Van Essen, The complete pattern of ocular dominance stripes in the striate cortex and visual field of the macaque monkey, Journal of Neuroscience 5 (1985) 486–501.
[33] C. von der Malsburg, D.J. Willshaw, How to label nerve cells so that they can interconnect in an ordered fashion, Proceedings of the National Academy of Sciences of the United States of America 74 (1977) 5176–5178.
[34] D.N. Mastronarde, Correlated firing of cat retinal ganglion cells. I. Spontaneously active inputs to X- and Y-cells, Journal of Neurophysiology 49 (1983) 303–324.
[35] K.D. Miller, Receptive fields and maps in the visual cortex: models of ocular dominance and orientation columns, in: E. Domany, J.L. van Hemmen, K. Schulten (Eds.), Models of Neural Networks, III, Springer-Verlag, New York, 1996, pp. 55–78.
[36] K.D. Miller, E. Erwin, A. Kayser, Is the development of orientation selectivity instructed by activity?, Journal of Neurobiology 41 (1999) 44–57.
[37] K.D. Miller, D.J.C. MacKay, The role of constraints in Hebbian learning, Neural Computation 6 (1994) 100–126.
[38] G. Mitchison, A type of duality between self-organizing maps and minimal wiring, Neural Computation 7 (1995) 25–35.
[39] P.R. Montague, J.A. Gally, G.M. Edelman, Spatial signaling in the development and function of neural connections, Cerebral Cortex 1 (1991) 199–220.
[40] J.D. Murray, Mathematical Biology, Springer, New York, 1989, p. 548.
[41] K. Obermayer, G.G. Blasdel, Geometry of orientation and ocular dominance columns in monkey striate cortex, Journal of Neuroscience 13 (1993) 4114–4129.
[42] K. Obermayer, G.G. Blasdel, K. Schulten, Statistical–mechanical analysis of self-organization and pattern formation during the development of visual maps, Physical Review A 45 (1992) 7568–7589.
[43] K. Obermayer, H. Ritter, K. Schulten, A principle for the formation of the spatial structure of cortical feature maps, Proceedings of the National Academy of Sciences of the United States of America 87 (1990) 8345–8349.
[44] C. Piepenbrock, K. Obermayer, The role of lateral cortical competition in ocular dominance development, in: M.S. Kearns, S.A. Solla, D.A. Cohn (Eds.), Advances in Neural Information Processing Systems 11, MIT Press, Cambridge, MA, 1999.
[45] D. Purves, D.R. Riddle, A.S. LaMantia, Iterated patterns of brain circuitry (or how the cortex gets its spots), Trends in Neuroscience 15 (1992) 268–362.
[46] F. de Ribaupierre, Acoustical information processing in the auditory thalamus and cerebral cortex, in: G. Ehret, R. Romand (Eds.), The Central Auditory System, OUP, New York, NY, 1997, pp. 317–397.
[47] M. Riesenhuber, H.U. Bauer, T. Geisel, Analyzing phase transitions in high-dimensional self-organizing maps, Biological Cybernetics 75 (1996) 397–407.
[48] K.S. Rockland, J.S. Lund, Intrinsic laminar lattice connections in primate visual cortex, Journal of Comparative Neurology 216 (1983) 303–318.
[49] A.W. Roe, D.Y. Ts'o, Visual topography in primate V2: multiple representation across functional stripes, Journal of Neuroscience 15 (1995) 3689–3715.
[50] F. Sengpiel, P. Stawinski, T. Bonhoeffer, Influence of experience on orientation maps in cat visual cortex, Nature Neuroscience 2 (1999) 727–732.
[51] J. Sirosh, R. Miikkulainen, Cooperative self-organization of afferent and lateral connections in cortical maps, Biological Cybernetics 71 (1994) 65–78.
[52] M. Steriade, E.G. Jones, D.A. McCormick (Eds.), Thalamus, Elsevier, Amsterdam, Holland, 1997.
[53] N.V. Swindale, The development of topography in the visual cortex: a review of models, Network: Computation in Neural Systems 7 (1996) 161–247.
[54] N.V. Swindale, J.A. Matsubara, M.S. Cynader, Surface organization of orientation and direction selectivity in cat area 18, Journal of Neuroscience 7 (1987) 1414–1427.
[55] D. Tal, E.L. Schwartz, Topological singularities in cortical orientation maps: the sign theorem correctly predicts orientation column patterns in primate striate cortex, Network 8 (1997) 229–238.
[56] M. Tessier-Lavigne, C.S. Goodman, The molecular biology of axon guidance, Science 274 (1996) 1123–1133.
[57] A.M. Turing, The chemical basis of morphogenesis, Philosophical Transactions of the Royal Society of London B 237 (1952) 37–72.
[58] M. Weliky, L.C. Katz, Correlation structure of spontaneous neuronal activity in the developing lateral geniculate nucleus in vivo, Science 285 (1999) 599–604.
[59] T.N. Wiesel, Postnatal development of the visual cortex and the influence of environment, Nature 299 (1982) 583–591.
[60] D.J. Willshaw, C. von der Malsburg, How patterned neural connections can be set up by self-organization, Proceedings of the Royal Society of London, Series B: Biological Sciences 194 (1976) 431–445.
[61] D.J. Willshaw, C. von der Malsburg, A marker induction mechanism for the establishment of ordered neural mappings: Its application to the retinotectal problem, Philosophical Transactions of the Royal Society B 287 (1979) 203–243.
[62] L. Wiskott, T.J. Sejnowski, Constrained optimization for neural map formation: a unifying framework for weight growth, Neural Computation 10 (1998) 671–716.
[63] F. Wolf, T. Geisel, Spontaneous pinwheel annihilation during visual development, Nature 395 (1998) 73–78.
[64] R.O. Wong, A. Chernjavsky, S.J. Smith, C.J. Shatz, Early functional neural networks in the developing retina, Nature 374 (1995) 716–718.
[65] T.A. Woolsey, H. Van der Loos, The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units, Brain Research 17 (1970) 205–242.
[66] R. Yuste, M. Sur (Eds.), Development and plasticity of the Cerebral Cortex, Journal of Neurobiology 41 (1) (1999).