

# Development of columnar structures in visual cortex

Miguel Á. Carreira-Perpiñán & Geoffrey J. Goodhill  
Department of Neuroscience  
Georgetown University Medical Center  
3900 Reservoir Road NW, Washington, DC 20007, US  
{miguel,geoff}@giccs.georgetown.edu

April 30, 2001

## Abstract

Many features of visual scenes are represented in the visual cortex in the form of maps. The best studied of these are the maps of features such as ocular dominance and orientation in primary visual cortex (V1). The beautifully regular structure of these maps and their dependence on patterns of neural activity have inspired several different computational models. In this chapter we focus on what can be explained by models based on the idea of optimising a trade-off between coverage and continuity, in particular the elastic net.

## 1 Introduction

Visual cortical map development is a well-studied area of computational neuroanatomy, characterised by an abundance of both experimental data and successful models. Starting with the classic experiments of Hubel and Wiesel in the 1970s, the characterisation of both the adult structure of the primary visual cortex (V1) and how that structure arises developmentally have become continuously more precise and detailed as experimental techniques have improved. Current methodologies include single- and multielectrode physiology, 2-deoxyglucose and cytochrome oxidase staining, optical imaging based on both intrinsic signals and voltage-sensitive dyes, and anatomical tract tracing techniques. In parallel several different types of theoretical models have developed, some of which are very well understood analytically. To comprehensively review the full range of both experimental data and theoretical models relevant to map formation in visual cortex would require far more space than is available in this chapter. We therefore restrict ourselves to only briefly reviewing recent experimental data regarding map structure and map development, and then focussing on one particular class of low-dimensional models which attempt to optimise a trade-off between coverage and continuity. We further restrict ourselves to issues regarding maps overall, rather than details of the receptive fields of the individual neurons from which these maps are made. Broader reviews of models can be found in Swindale [1, 2].

## 2 Structure of adult maps

Receptive fields of V1 neurons are highly selective along a number of feature dimensions of the stimulus. These feature dimensions include position in the visual field, eye of origin (ocular dominance), orientation, direction of movement, spatial frequency of a grating and disparity. Neurons lying along a line or column orthogonal to the surface of V1 respond in approximately the same way to visual stimuli. However, responses vary in an organised way in the tangential direction, parallel to the surface. As is common in this field we will therefore discuss only the two-dimensional structure of the visual cortex. Such two-dimensional organisation of preferred responses to a particular stimulus feature is termed a *map* (e.g. ocular dominance map, orientation map and so on), and several such maps coexist on the same neural substrate. The map of preferred location in the visual field is topographic on a large scale (i.e., moving systematically across the visual field roughly corresponds to moving systematically across the cortex), though more convoluted on a fine scale [3]. The ocular dominance map consists of alternating stripes or blobs with a regular periodicity, with neurons in each stripe/blob responding preferentially to stimuli in one eye [4]. The orientation map is also striped with an overall periodicity [5], but is characterised by point singularities or pinwheels, around which a circular path meets once all orientations from  $0^\circ$  to  $180^\circ$  [6]. The

structure of the spatial frequency map is somewhat controversial [7], with competing claims for both a binary [8, 9] and a more continuous representation [10].

Different maps are not independent from each other: the stripes of ocular dominance and orientation tend to run locally orthogonal to each other and orientation singularities tend to lie in the centre of ocular dominance stripes [11, 12, 9]. Besides this local structure some global structure is also apparent: in monkeys ocular dominance columns tend to run parallel to the shorter axis of V1, orthogonal to V1 boundaries, and are more irregular in the foveal region [13, 14]. Maps of two individual animals of the same species are qualitatively similar, but maps of two individual animals of different species differ in the amount of columnar segregation and its type (stripes, blobs), periodicity, pinwheel density and other structural characteristics. A problem which several theoretical researchers have become engaged in is to find effective ways to quantitatively characterise map structure (e.g. [15, 16, 17, 18]).

### 3 Map development: role of activity

The view of development of V1 (layer IV) that has been universally accepted until recently can be summarised as follows (see e.g. [19, 20, 21]). (1) During early stages of circuit development, genetically specified molecular signals guide axonal outgrowth and targetting. These early connections are typically diffuse and imprecise. (2) These connections are then refined and some are eliminated in response to visual activity, giving rise to the adult pattern of connectivity.

In particular, ocular dominance column formation requires a prolonged activity-dependent segregation process (lasting several weeks in cats and ferrets). This starts from an initial state in which lateral geniculate nucleus afferents representing both eyes overlap extensively and ends in a mature state where eye-specific afferents occupy stripes. The hypothesis that column formation is activity-dependent is based on a large number of experiments dating back to the 1960s. When TTX (tetrodotoxin) is injected into both eyes of the cat during the critical period (the period when visual experience affects column formation), thus blocking all retinal activity, ocular dominance columns are not seen either anatomically or physiologically [22]. When one eye is occluded or sewn shut during the critical period, thus causing an imbalance in the amount of activity from the two eyes, it is found that (a) a higher proportion of cortical cells than normal are completely monocular, (b) substantially more of the cells in layer IV can be driven by the normal eye as compared to the deprived eye, and (c) ocular dominance stripes are now of different thicknesses for the two eyes. The stripes receiving input from the normal eye expand at the expense of the stripes from the deprived eye [19, 23, 24]. When an occluded eye competes with an eye injected with TTX there is a shift towards the occluded eye [25]. The cortical imbalance in the representation of the two eyes is greater when an occluded eye competes with a normal eye than when a TTX-injected eye competes with a normal eye [26]. When strabismus is induced during the critical period, thus preventing image registration in the two retinæ and decreasing the strength of between-eye correlations, all cells become entirely monocular, and the pattern of stripes becomes correspondingly sharper [27]. It was originally thought that overall column periodicity was left unchanged; recent data has suggested otherwise in the cat [28] though this is controversial [29].

The data regarding the effect of activity on orientation map development is more complex and controversial; see section 6 and Swindale [2] for review.

### 4 Theoretical models: coverage and continuity

The role of activity in shaping cortical maps has usually been modelled via Hebbian learning rules. Such rules can often be interpreted as implementing gradient ascent/descent in some objective function, so that the effect of the developmental process is to optimise (at least to some extent) that function. A particularly useful class of objective functions implements a trade-off between two competing tendencies, **coverage uniformity** (or completeness) and **continuity** (or similarity). However, even though several mathematical definitions of coverage uniformity and continuity have been given, the principles of coverage uniformity and continuity remain conceptually vague. For example, coverage uniformity and completeness are strictly different: the former means that each combination of stimuli values (e.g. any orientation in any visual field location of either eye) has equal representation in the cortex, while the latter means that any combination of stimuli values is represented somewhere in cortex. Thus, coverage uniformity implies completeness (disregarding the trivial case of a cortex uniformly nonresponsive to stimuli), but not vice versa, since it is possible to have over- and underrepresented stimuli values. Besides, it is not possible to represent all values of a continuous, higher-dimensional stimulus space with a two-dimensional

cortex<sup>1</sup>. A practically useful middle ground is to consider that the set of stimulus values represented by the cortex be roughly uniformly scattered in stimulus space—whatever that set is. Continuity is even less well defined than coverage uniformity. Loosely, we can say continuity means that neurons which are physically close in cortex tend to have similar stimulus preferences, or (non-equivalently) that similar features are represented nearby in the cortex. This can be motivated in terms of economy of cortical wiring [30].

The striped structure of several of the maps can be understood to represent a compromise between coverage and continuity. An early idea based on these principles is the *icecube* model of Hubel and Wiesel [4], where stripes of ocular dominance run orthogonally to stripes of orientation and all combinations of eye and orientation preference are represented within a cortical region smaller than a cortical point image (the collection of neurons whose receptive fields contain a given visual field location).

These general optimisation principles of coverage and continuity do not in themselves support any specific development rule, since there may be different ways in which they can be optimised. However, heuristic rules that obtain local optima are rather more plausible than a global search where the best of all possible configurations is found. It is appealing from both biological and computational perspectives to consider that visual cortical structure is the result of small developmental changes driven by neural activity. It is then plausible to think that such principles—abstractions based on physical and biological constraints and on adaptation to the environment—have a strong influence on the cortical structure, but with the following caveats: that they probably are not the only principles at work; and that they are only partly optimised in real organisms.

Particularly successful examples of such heuristic rules are the elastic net (EN) [31, 30, 32] and self-organising map (SOM) [33, 34]. In these models, the competition can be explained in a dimension reduction framework, where a two-dimensional cortical sheet twists in a higher-dimensional stimulus space to cover it as uniformly as possible while minimising some measure of discontinuity. These models differ in their explicit mathematical definitions, but produce maps which are similar and display a quantitatively good match to the observed phenomenology of cortical maps [35, 1]. This includes: (1) the striped structure of ocular dominance and orientation columns with appropriately related periodicities and, for orientation, the existence of singularities (pinwheels); (2) the interrelations between different maps, such as the tendency of orientation and ocular dominance stripes to be locally orthogonal and of the pinwheels to lie on the centre of ocular dominance stripes; and (3) the effect on the maps of various abnormal conditions during development, such as strabismus or monocular deprivation. Shortly we discuss specific details of the elastic net and its application to cortical mapping problems. First however we discuss coverage and continuity from a more general mathematical perspective, which helps to illustrate how the elastic net fits into the broader picture.

## 4.1 Mathematical formulation of coverage uniformity and continuity

Given a representation  $\mathbf{M}$  of a cortical map, a mathematically convenient way of writing the tradeoff between the goals of attaining uniform coverage and respecting the constraints of cortical wiring is to assume that cortical maps maximise an objective function

$$\mathcal{F}(\mathbf{M}) \stackrel{\text{def}}{=} \mathcal{C}(\mathbf{M}) + \lambda \mathcal{R}(\mathbf{M}) \quad (1)$$

where  $\mathcal{C}$  is a measure of the uniformity of coverage,  $\mathcal{R}$  is a measure of the continuity and  $\lambda > 0$  specifies the relative weight of  $\mathcal{R}$  with respect to  $\mathcal{C}$ . This formulation is formally akin to regularisation theory [36]. We assume that maximising either  $\mathcal{C}$  or  $\mathcal{R}$  separately does not lead to a maximum of  $\mathcal{F}$  and therefore that maxima of  $\mathcal{F}$  imply compromise values of  $\mathcal{C}$  and  $\mathcal{R}$ . By quantitatively defining  $\mathcal{C}$  and  $\mathcal{R}$  in terms of the map representation  $\mathbf{M}$  it is in principle possible to perform a numerical optimisation of  $\mathcal{F}$  to generate a map. We examine several possibilities next.

### 4.1.1 Model-based formulation

For a model, the map representation  $\mathbf{M}$  is the set of model parameters, such as synapse strengths, tuning widths, mapping parameters, receptive field centres, etc. Such parameters uniquely determine the values of orientation, etc. at any point in the model cortical sheet. For the elastic net (which is discussed in more detail in section 5), the parameters are the locations in stimulus space of the reference vectors  $\{\mathbf{y}_i\}_{i=1}^N$  and the width  $k$ . From eq. (9) we can define

$$\mathcal{C}(\{\mathbf{y}_i\}_{i=1}^N, k) \stackrel{\text{def}}{=} k \sum_i \log \sum_j \Phi(\|\mathbf{x}_i - \mathbf{y}_j\|, k) \quad \mathcal{R}(\{\mathbf{y}_i\}_{i=1}^N, k) \stackrel{\text{def}}{=} - \sum_j \|\mathbf{y}_{j+1} - \mathbf{y}_j\|^2 \quad (2)$$

---

<sup>1</sup>We should say *practically* not possible, since from set theory we know that the cardinal of  $\mathbb{R}^D$  is equal to the cardinal of  $\mathbb{R}$   $\forall D \geq 1$ : there exists a continuous, one-to-one mapping from  $\mathbb{R}^D$  to  $\mathbb{R}$ .

where  $\{\mathbf{x}_i\}_{i=1}^N$  is the sample of stimulus values,  $\Phi$  is defined in eq. (8) and  $\lambda \stackrel{\text{def}}{=} \frac{\beta}{2\alpha}$ . Note that the standard elastic net algorithm minimises  $-\alpha\mathcal{F}$  instead of maximising  $\mathcal{F}$  and also anneals  $k$  rather than optimising over it, but the EM (expectation-maximisation) algorithm version does optimise over  $k$  and the reference vectors jointly. In the probabilistic interpretation of the elastic net, the  $\mathcal{C}$  function is simply the log-likelihood of the parameters  $\{\mathbf{y}_i\}_{i=1}^N$  and  $k$  for the data sample  $\{\mathbf{x}_i\}_{i=1}^N$ , i.e., the probability density that an elastic net density model with those parameter values generated the sample. The  $\mathcal{R}$  function is a negative “length” of the net (it would be exactly the Euclidean length of the elastic net from  $\mathbf{y}_1$  to  $\mathbf{y}_N$  if the terms were not squared). It attains its maximal value at 0 for a point-like net where  $\mathbf{y}_1 = \dots = \mathbf{y}_N$ : the more stretched the net is, the less “continuous” it is. This has been motivated in terms of economy of cortical wiring (the cost of setting up a given neural connectivity pattern; [30]).

Unlike the elastic net, the self-organising map learning rule cannot be integrated to give an objective function [37]. However, slight variations of its rule leading to basically the same behaviour can be integrated and give an objective function very similar to that of the elastic net. For example, in the probabilistic variation of self-organising maps of Utsugi [38]:

$$\mathcal{C}(\{\mathbf{y}_i\}_{i=1}^N, \sigma) \stackrel{\text{def}}{=} \sum_i \log \sum_j e^{-\frac{1}{2} \left\| \frac{\mathbf{x}_i - \mathbf{y}_j}{\sigma} \right\|^2} \quad \mathcal{R}(\{\mathbf{y}_i\}_{i=1}^N, \sigma) \stackrel{\text{def}}{=} -\frac{1}{2} \sum_j \|(\mathbf{D}\mathbf{y})_j\|^2 \quad \mathcal{F} \stackrel{\text{def}}{=} \mathcal{C} + \lambda\mathcal{R} \quad (3)$$

where  $\mathbf{D}$  is a discretised differential operator, e.g. first order gives  $(\mathbf{D}\mathbf{y})_j = \mathbf{y}_{j+1} - \mathbf{y}_j$ , second order  $\mathbf{y}_{j+1} - 2\mathbf{y}_j + \mathbf{y}_{j-1}$  and so on. Similarly generalised regularisation terms could also be used in the elastic net [39].

#### 4.1.2 Model-free formulation

A model-free, resolution-dependent representation of a map can be defined as a two-dimensional array (not necessarily rectangular) of vector values of the stimulus variables of interest. Each position  $(i, j)$  in the array represents an ideal cortical cell. Let us call  $\mathcal{C}$  the set of all such *cortical positions* (e.g.  $\mathcal{C} = \{1, \dots, M\} \times \{1, \dots, N\}$  represents an  $M \times N$  rectangular array). There is a vector of *stimulus values*  $\boldsymbol{\mu}_{ij}$  associated to each cortical position  $(i, j)$ . Stimulus variables of interest include the retinotopic position (or receptive field centre in the visual field)  $(x, y)$  in degrees, the preferred orientation  $\theta \in [0^\circ, 180^\circ)$ , the ocular dominance  $n$  ( $-1$ : left eye,  $+1$ : right eye) and the spatial frequency  $m \in \{-1, 1\}$ . Therefore  $\boldsymbol{\mu}_{ij} \stackrel{\text{def}}{=} (n_{ij}, m_{ij}, \theta_{ij}, x_{ij}, y_{ij})$  for  $(i, j) \in \mathcal{C}$  can be considered a *generalised receptive field centre*; a *receptive field* could then be defined by a function sitting on the receptive field centre and monotonically decreasing away from it. The collection  $\mathbf{M} \stackrel{\text{def}}{=} \{\boldsymbol{\mu}_{ij}\}_{(i,j) \in \mathcal{C}}$  of such receptive field centres, together with the two-dimensional ordering of cortical positions in  $\mathcal{C}$ , defines the cortical map. This representation is applicable to maps measured empirically, for example with optical recording techniques.

Swindale [40] (see also [41]) introduced the following mathematical definition of coverage, which is applicable to this representation independently of any model. Given an arbitrary stimulus  $\mathbf{v}$ , the total amount of cortical activity that it produces is defined as

$$A(\mathbf{v}) \stackrel{\text{def}}{=} \sum_{(i,j) \in \mathcal{C}} f(\mathbf{v} - \boldsymbol{\mu}_{ij}) \quad (4)$$

where  $f$  is the (generalised) receptive field of cortical location  $(i, j)$ , assumed translationally invariant (so it depends only on the difference of stimulus and generalised receptive field centre values);  $f$  is taken as a product of functions: Gaussian for orientation and retinotopic position (with widths derived from biological estimates of tuning curves) and delta for ocular dominance and spatial frequency.  $A$  is calculated for a regular grid in stimulus space, which is assumed to be a representative set of stimulus values. The measure of coverage uniformity is finally obtained as

$$c' \stackrel{\text{def}}{=} \frac{\text{stdev}\{A\}}{\text{mean}\{A\}} \quad (5)$$

that is, the magnitude of the (normalised) dispersion of the total activity  $A$  in the stimulus space. Intuitively,  $c'$  will be large when  $A$  takes different values for different stimuli and zero if  $A$  has a positive value independent of the stimulus. Thus, it is a measure of lack of coverage uniformity, and we could define  $\mathcal{C} \stackrel{\text{def}}{=} -c'$ .

Note that the function  $A$  can be seen as a kernel density estimate [42] for the sample  $\{\boldsymbol{\mu}_{ij}\}_{(i,j) \in \mathcal{C}}$  with smoothing parameter given by the width of the kernel function  $f$ . In fact, its spirit is the same as that of the probabilistic interpretation of the elastic net and self-organising map. This is because the latter are vector quantisation methods, where the dimension reduction mapping is implicitly defined by the reference vectors in stimulus space. If the mapping was defined explicitly via parameters, as e.g. in the generative topographic mapping (GTM) model [43], the resulting  $\mathcal{C}$  and  $\mathcal{R}$  functions would be quite different.

More difficult is to define  $\mathcal{R}$  because the cortical wiring constraints are largely unknown and possibly result from the combined effect of several factors. At an abstract level, we can define  $\mathcal{R}$  as a similarity measure where the preferred stimuli of nearby cortical neurons are similar, as in the expression [44]:

$$\sum_{(i,j),(k,l) \in \mathcal{C}} F((i,j), (k,l)) G(\boldsymbol{\mu}_{ij}, \boldsymbol{\mu}_{kl}).$$

In fact, many models of cortical development implicitly or explicitly implement such a definition of continuity for specific choices of the similarity measures  $F$  and  $G$ , like the elastic net, often with a flavour of wire length or local similarity of neural responses (e.g. [45]; see Goodhill [46] for discussion). However, this implies introducing model assumptions. At present continuity is too vaguely defined to be quantitatively characterised for a model-free map representation.

## 5 The elastic net algorithm

The elastic net algorithm [31] was originally developed as an approximate method for the Travelling Salesman Problem (TSP), a well-known NP-complete combinatorial optimisation problem. Here the objective is to find the shortest distance a salesman can travel to visit a set of  $N$  cities in a plane and return to where they started. The key idea is that this problem is analogous to the problem of forming topographic maps in the nervous system, where cities represent input or feature points and the tour represents the ordering of these points onto the target structure—though generally in the nervous system there is no “return to where you started” constraint. The TSP problem has been extensively discussed in the combinatorial optimisation literature, being both easy to state and hard to solve. For  $N$  cities there are  $\frac{(N-1)!}{2}$  possible routes: for large  $N$  it is impossible to search them all to find the optimal tour. Therefore many heuristic algorithms have been investigated, which aim to provide good solutions in reasonable time (for review see Lawler et al. [47]). The set of valid tours for a TSP of size  $N$  can be represented as the vertices of an  $N$ -dimensional hypercube, and most techniques aim to provide good ways of stepping from one vertex to another to gradually improve the quality of the solution. However, an alternative method used by the elastic net is for the search to proceed through the continuous space inside the body of the hypercube, only converging to a valid solution in the final state.

The dynamics of the elastic net algorithm are closely related to those of the Kohonen algorithm. Both algorithms grew out of earlier models of retinotectal map formation (Kohonen: Willshaw and von der Malsburg [48]; elastic net: Willshaw and von der Malsburg [49]). The basic framework is an array of cortical cells which receive weighted connections from points in an input space. In the so-called low-dimensional version of both algorithms, input points represent features—such as a line segment of a particular orientation located at a particular point in the image—rather than individual image pixels. Generally, each feature is represented by an orthogonal dimension: one for  $x$  position, one for  $y$  position, one for ocularity (degree of left or right eye dominance), one for orientation, one for direction, and so on. The weight vectors of cortical cells can be represented as points in the input space. The input space is densely populated with such points so that all appropriate feature combinations are represented. In the Kohonen algorithm, inputs are presented in turn. The initial activity of each cortical cell is the sum of pixel values times weights. This activation rule is equivalent to calculating the distance between the input vector and each weight vector (see e.g. Kohonen [34] for mathematical details). In Willshaw and von der Malsburg [48] there was an explicit pattern of lateral cortical connections, consisting of short range excitation and longer range inhibition, and the activity of cortical cells was iterated until the cortical activity pattern stabilised. The most strongly responding cortical cells then had their weights updated according to a Hebbian learning rule (see Willshaw and von der Malsburg [48] for mathematical details). Kohonen’s insight was that the end result of this process is usually (though not always) a blob of activity centred on the cortical cell that initially received the largest input. He therefore proposed an algorithmic shortcut, whereby it is assumed from the outset that the only cortical cells that should have their weights updated are those close to the unit that initially responded most strongly, which is the unit whose weight vector was closest to the input vector. The elastic net works in a similar way to this, with two important differences. Firstly, whereas Kohonen uses hard competition between cortical units, the elastic net uses soft competition. This means that in the elastic net, *all* cortical units are updated in proportion to how strongly they respond to each input pattern, rather than in the Kohonen algorithm where just the most strongly responding are updated. Secondly, the elastic net algorithm operates in *batch* mode: all input points are considered simultaneously for updating cortical cells, rather than presenting them one at a time as in Kohonen’s algorithm. The biologically appealing interpretation that input points are being seen one at a time and that the cortex is responding to each in turn is now lost; however this modification makes little difference from a mathematical perspective, and allows a useful statistical interpretation as described below.

Refer to the positions of points in the input (feature) space as  $\mathbf{x}_i$ , and the positions of the weight vectors of cortical cells in the input space as  $\mathbf{y}_j$ . The change in the position  $\Delta\mathbf{y}_j$  of each cortical unit at each time step is given by

$$\Delta\mathbf{y}_j = \alpha \sum_i w_{ij}(\mathbf{x}_i - \mathbf{y}_j) + \beta k \sum_{j' \in \mathcal{N}(j)} (\mathbf{y}_{j'} - \mathbf{y}_j). \quad (6)$$

The first term is a matching term that represents the ‘‘pull’’ of feature points for cortical units, which is traded off with ratio  $\alpha/\beta k$  against a regularisation term representing a ‘‘tension’’ in the cortical sheet, i.e., a desire for neighbouring cortical cells to represent neighbouring points in the feature space (short-range excitation between cortical cells). The sum over all input points  $i$  indicates that the algorithm operates in batch mode.  $\mathcal{N}(j)$  refers to the set of cells in the cortical sheet that are neighbouring to  $j$ . The  $w_{ij}$ ’s (rather confusingly termed ‘‘weights’’ by Durbin and Willshaw [31]) say how much cortical cell  $j$  is activated by input  $i$  as a function of the difference between  $\mathbf{x}_i$  and  $\mathbf{y}_j$ :

$$w_{ij} = \frac{\Phi(\|\mathbf{x}_i - \mathbf{y}_j\|, k)}{\sum_p \Phi(\|\mathbf{x}_i - \mathbf{y}_p\|, k)} \quad (7)$$

where

$$\Phi(\|\mathbf{x}_i - \mathbf{y}_j\|, k) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{y}_j\|^2}{2k^2}\right). \quad (8)$$

The sum over all cortical cells in the denominator of equation 7 is a normalisation term that says that each input produces the same overall amount of activity in the cortical sheet: the  $w_{ij}$ ’s give the distribution of that activity.  $k$  is a scale parameter that determines the spread of this overall activity: if  $k$  is large then many cortical cells are roughly equally activated by an input, whereas if  $k$  is small then only those cortical cells whose weight vectors are closest to the input vector are significantly updated. Over the course of a simulation, the scale parameter  $k$  is gradually reduced, so that the matching term comes to dominate the regularisation term. Equation (6) implements Hebbian learning in the sense that the degree to which a cortical cell is updated by a particular input depends on the degree of similarity between the input pattern and the weight vector of that cortical cell, i.e., the extent to which the activity of a cortical cell is correlated with that pattern of input activity.

Equation (6) can be integrated to produce an *energy* function  $E$  which is such that  $\Delta\mathbf{y}_j = -k \frac{\partial E}{\partial \mathbf{y}_j}$ :

$$E = -\alpha k \sum_i \log \sum_j \Phi(\|\mathbf{x}_i - \mathbf{y}_j\|, k) + \frac{\beta}{2} \sum_j \|\mathbf{y}_{j+1} - \mathbf{y}_j\|^2. \quad (9)$$

As explained in section 4.1.1, this function realises the competition between coverage uniformity and continuity. The first term (coverage uniformity) simply tries to produce a one-to-one match between input features and cortical cells. The second term (continuity) says that only the similarity between neighbouring cortical cells is considered, and that dissimilarity in the input space is given by squared distance (the distances are squared for computational convenience). Carefully chosen versions can be interpreted as corresponding to particular patterns of lateral connections in the cortex [39]. This rather abstract version of Hebbian learning has the advantage of an elegant statistical interpretation [50]. Consider each unit on the elastic sheet as a Gaussian generator of data in the feature space. Each Gaussian has the same variance, determined by  $k$ . At each value of  $k$ , the optimisation of the first term in the energy function corresponds to finding the positions of points on the sheet where they are most likely to have generated the data: a maximum likelihood model. The second term acts as a prior on the model that favours solutions with short distances between Gaussian centres: alternative forms for this term therefore correspond to different priors. The algorithm implements a form of graduated non-convexity or deterministic annealing [51]. At large values of  $k$  the energy function has a unique minimum. As  $k$  is reduced, the energy function bifurcates to produce several local minima, and the algorithm tracks one of these. This process continues as  $k$  is further reduced. In common with all heuristic optimisation methods, the algorithm is not guaranteed to find the global minimum. Durbin et al. [50] analyzed the application of this algorithm to the TSP, where *cortical* units form a one dimensional loop of points on the tour and *feature* points are cities in a plane. By calculating the first and second derivatives of the energy function with respect to the positions of cortical points, they showed that the centre of gravity of the feature points (i.e., the configuration where all cortical receptive fields are coincident at this point) is always an extremum of the energy function. They calculated the value of  $k$  for which the centre ceases to be a stable minimum, i.e., the energy function bifurcates, to be  $k \approx \sqrt{\lambda}$ , where  $\lambda$  is the principal eigenvalue of the covariance matrix of the feature points. At this stage the cortical cells form a line (or sheet) along the principal axis of the covariance matrix of the feature points. By varying the structure of the covariance matrix of the feature points, it is therefore possible to vary the order in which different maps (e.g. orientation, ocular dominance) develop. We return to this issue later.

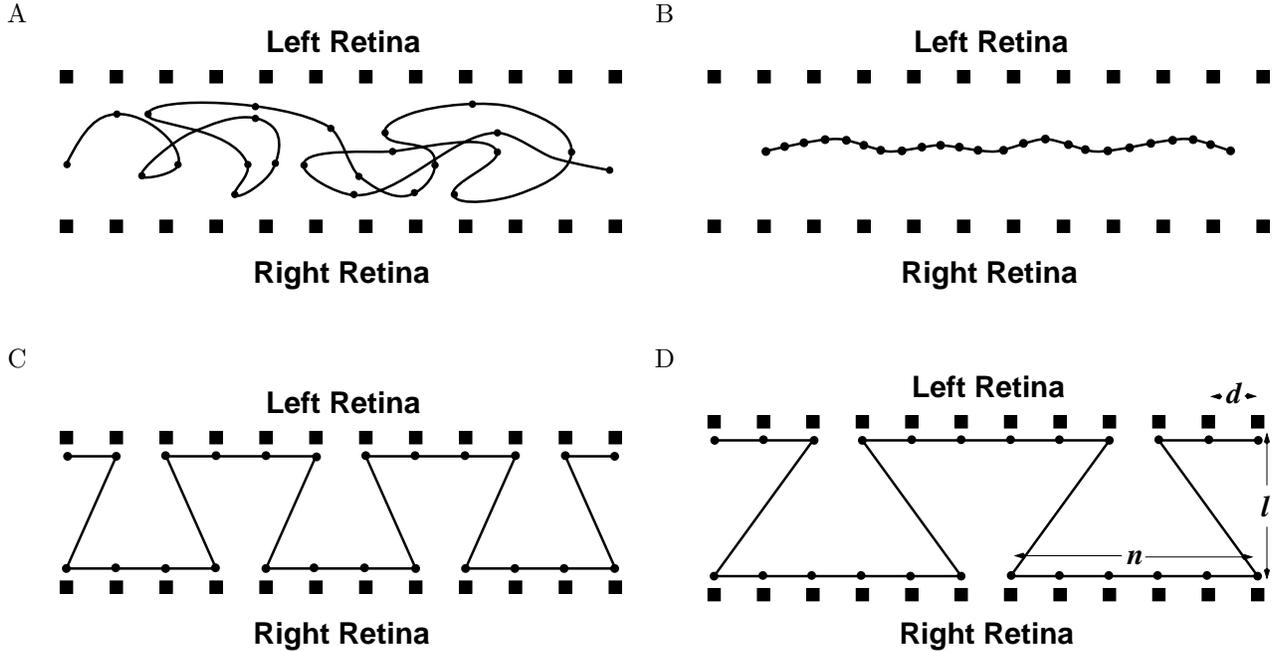


Figure 1: Stages of development of the elastic net for the one-dimensional ocular dominance problem [32]. Each retina is represented by a row of points in a feature space where distances are taken to represent correlations. The mapping to the cortex is represented by an elastic rope, where points on the rope represent cells in the cortex. The position of each cortical cell in this abstract space is updated iteratively so as to simultaneously match cortical cells to retinal locations and keep the distance between neighbouring cortical cells as small as possible (i.e., maximise the degree to which neighbouring cortical cells receive highly correlated inputs). **A.** Cortical cells are initially positioned randomly in the feature space, except for a crude initial topographic bias. **B.** As development proceeds, the cortical mapping assumes an ordered topography. Cortical cells lie roughly equidistant between the two retinæ, signifying a completely binocular mapping. **C.** Eventually cortical cells become committed to particular retinal locations, and a periodic pattern emerges (at this stage cortical points lie on top of retinal points: for clarity in the picture their positions have been slightly offset vertically). **D.** Moving the two rows of retinal points further apart (increasing  $l$  in the terminology of the diagram) causes wider columns to be formed. For a striped solution it is straightforward to calculate that the optimal width is  $n = 2l/d$ . For simplicity, the same number of cortical cells as retinal points have been drawn.

**Ocular dominance** The behaviour of the elastic net for the ocular dominance column mapping in one dimension is shown in figure 1 and equivalent results in two dimensions are illustrated in figure 2 [32, 52]. In both figures, it can be seen that column periodicity depends on the degree of correlation between the eyes. This can be understood theoretically as follows. The elastic net tries to find a mapping that maximises the degree to which neighbouring cells in the cortex receive inputs that are highly correlated. In the abstract representation shown in figure 1, this is equivalent to minimising the length of the path that joins all input points (in fact, for reasons of computational efficiency, the elastic net minimises the sum of squares of distances). Refer to the distance (in correlation space) between neighbouring points in the same eye as  $d$ , and the distance between corresponding points in the two eyes as  $l$  (see figure 1d). It is easy to show that the optimal width  $n$  of a striped solution<sup>2</sup> is  $n = 2l/d$  [32, 52]. As  $l$  increases the correlation between the two eyes decreases [54], and the stripe width increases. An experimental prediction following from this result is that kittens raised with divergent strabismus should have wider ocular dominance columns than normal kittens [52, 55]. Evidence in favour of this prediction was found by Löwel [28] (for further discussion see Goodhill and Löwel [56]), though this is controversial [29]. The effect of reduced activity in one eye [23, 24] can be modelled with the elastic net by reducing the effective

<sup>2</sup>For  $l > d$ , the solution that traverses the complete length of the left eye followed by the complete length of the right eye is in fact more optimal than any striped solution. In this case  $n = 2l/d$  is a local rather than a global minimum. However, in practice the elastic net always finds a striped solution, since the dynamics of the algorithm always establish initial topography with the two ends of the cortex at opposite ends of the two retinæ for sufficiently large values of the annealing parameter. Biologically, a biased initial topography is established by activity-independent molecular cues: see e.g. Goodhill and Richards [53].

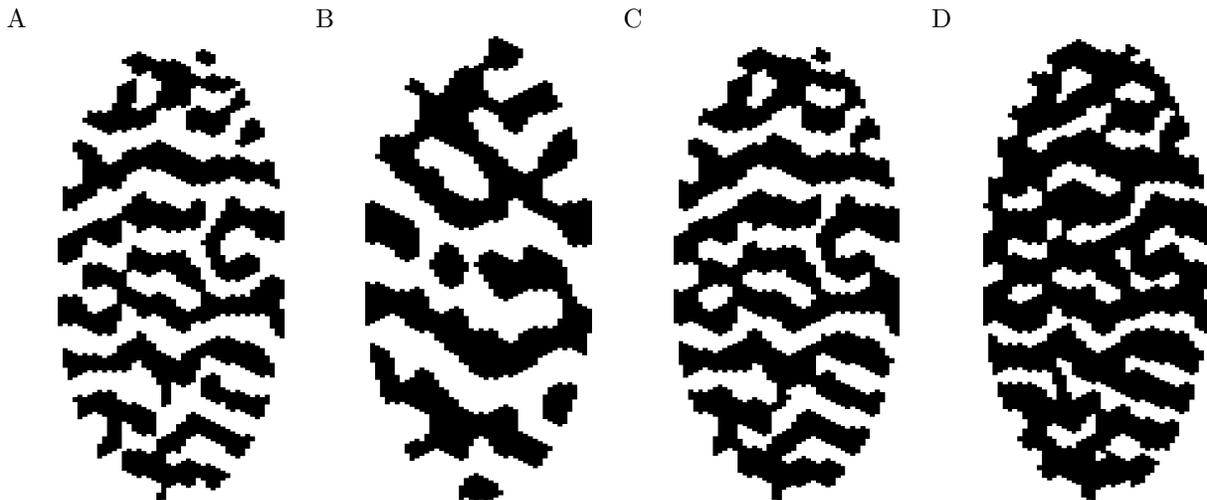


Figure 2: Ocular dominance maps produced by the elastic net in two dimensions. Cortical points are coloured white or black depending on the eye to which they are committed. The two retinal sheets were hexagonal grids with a circular boundary, and the cortical sheet was a hexagonal grid with an elliptical boundary. **A.** Normal development. **B.** Strabismic development (reduced correlation between the eyes). **C, D.** Effects of monocular deprivation. **C.** Activity in the deprived eye reduced by 25%. **D.** Activity in the deprived eye reduced by 50%. The same initial conditions were used in each case.

pull that points in one eye exert on cortical cells relative to points in the other eye [57]. This corresponds to reducing the parameter  $\alpha$ . Illustrative results are shown in figure 2c,d.

In the macaque, besides local structure ocular dominance stripes are also characterised by an overall orientation that varies with position in V1: ocular dominance stripes are less parallel in the foveal region, tend to be orthogonal to the borders of the neighbouring visual cortical area and decrease in width from the fovea to the periphery. Goodhill et al. [58] attempted to model this global structure using the elastic net. They identified three factors with an influence on the global structure of cortical maps and implemented them with the elastic net as summarised:

Factor	Reason proposed	Elastic net implementation	Elastic net result
Spatially nonuniform correlational structure of activity in the retina	The increase in retinal ganglion cell density proceeding from peripheral to central retina	Foveal region of increased density of retinal points (smaller spacing), representing stronger correlations between neighbouring retinal ganglion cells in the fovea	Columns are wider and more disordered in the cortical representation of the fovea
Spatially anisotropic correlational structure of activity in the retina	The asymmetric way in which the retina develops.	By squashing the two retinal sheets (in the training set) so that the spacing between points is less in one direction than the other	Columns tend to line up orthogonal to the direction of stronger correlations
Elongated cortical shape		Circular retina and elliptical cortex	Columns line up parallel to the short axis of the cortical sheet.

They thus concluded that (a) the widening of the columns in the foveal region results from stronger correlations (decreased spacing of retinal points); and (b) the increased disorder in the foveal representation results from two competing effects: an elliptical cortex that causes column alignment parallel to the short axis and anisotropic input correlations that cause column alignment parallel to the long axis. This model predicts that in a strongly anisotropic visual environment (e.g. kittens raised with cylindrical lenses), columns should tend to line up parallel to the direction of the weaker correlations.

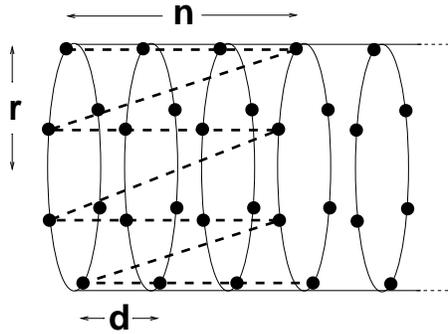


Figure 3: One possible type of mapping between a cylindrical feature space representing orientation and one spatial dimension, and a one-dimensional elastic net (solid circles: feature points; dashed line: elastic net; solid lines to aid visualisation). Here the net forms isoorientation domains of smoothly varying orientation, each of which traverses  $n$  units in the spatial dimension.

**Orientation** The elastic net was first applied to the formation of orientation maps by Durbin and Mitchison [30]. A standard trick in these types of models to capture the intrinsic periodicity of the orientation dimension is to wrap it around into a cylinder, so that orientation is now represented by a circular manifold in two dimensions rather than as just a line in one dimension. An equivalent periodicity analysis to that described earlier for ocular dominance can be performed, to yield an expected periodicity of orientation columns of  $2\pi r/d$  (see figure 3) which is in good agreement with simulation results [59]. A quantitative comparison of several different models of column formation [35] found that the elastic net and Kohonen algorithms actually produced the best match to real orientation and joint orientation/ocular dominance maps.

Two aspects of the elastic net applied to the joint formation of ocular dominance and orientation columns were examined by Goodhill and Cimponeriu [59]: which model parameters control the order of development of the two sets of columns, and how this ordering affects the final patterns of columns produced. Applying the bifurcation analysis of Durbin et al. [50] discussed earlier, Goodhill and Cimponeriu [59] derived expressions for the critical values of  $k$  for expansion along each of the visual space, orientation and ocular dominance dimensions in terms of the parameters  $d$ ,  $l$  and  $r$  of the feature space (see also [60, 61]). Combining these expressions with those quoted earlier for the expected ocular dominance and orientation periodicities, it is possible to relate the ratio of critical  $k$  values to the ratio of final periodicities without any reference to the parameters  $d$ ,  $l$  and  $r$ . Approximately, if the orientation wavelength is greater than the ocular dominance wavelength then orientation columns developed first, otherwise ocular dominance columns developed first. The model then predicts that in normal macaque monkeys ocular dominance develops first, in normal cats orientation develops first, but in strabismic cats ocular dominance develops first. Whether this is true has not yet been experimentally determined.

An intriguing property of a particular version of the elastic net was discovered by Wolf and Geisel [62]. They carried out a symmetry-based theoretical analysis of the dynamics of pinwheels during visual development, assuming only a very general class of activity-dependent learning rules. They showed that these symmetry properties together with some general assumptions of Gaussian statistics and homogeneous correlations imply that, if starting from a random unselective initial state, a minimal expected spatial density of pinwheels  $\rho$  emerges when orientation selectivity is first established:  $\rho \geq \frac{\pi}{\Lambda^2}$ , where  $\Lambda$  is the typical spacing of the orientation stripes. Therefore, if pinwheels are found in adult animals at a density  $\rho < \frac{\pi}{\Lambda^2}$ , then there must have been motion and annihilation of pinwheel pairs of opposite sign (as is also observed in the physics of defects). This result is robust to variations in the particular details of the developmental rule employed, and depends only on its intrinsic symmetries. The elastic net falls within this general class and displays this tendency for pinwheels to move and annihilate during development, if it is run in the *non-annealed* regime. In this version,  $k$  is kept fixed during a simulation at a value below all relevant critical values, so that expansion occurs along all dimensions simultaneously. Pinwheel annihilation leads to growing regions of stripe-like isoorientation domains which locally resemble a plane wave. Using different sets of stimuli tuned to attain a variable degree of ocular dominance, the presence of ocular dominance columns slows down or even stops pinwheel annihilation, in proportion to the degree of ocular dominance segregation; thus, pronounced ocular dominance is associated to high scaled density of pinwheels. As yet no theoretical justification has been given for this. In the annealed version the orientation map eventually stabilises; whether annihilation occurs to any significant extent depends on how slow the rate of annealing is.

## 6 Discussion

As we have seen above, the elastic net (and related models such as the self-organising map) can accurately reproduce a large amount of the observed phenomenology of ocular dominance and orientation maps. In addition, these models make several interesting and often surprising predictions about the outcomes of certain experiments. These include predictions of an increase in overall ocular dominance column periodicity with strabismus, a difference in the order of development between normal and strabismic cats, and pinwheel annihilation during development. While these claims are experimentally somewhat controversial, the generation of such clear and testable predictions has helped ferment an active and productive dialog between theoretical and experimental researchers. However, it is also appropriate to focus on some issues raised by recent data that these models do not account for.

### 6.1 Retinotopy distortions

The retinotopic map of V1, given by the receptive field centres in visual space of every cortical neuron, is roughly uniform and has received far less attention than the ocular dominance and orientation maps. However, it contains both local and global distortions. Global distortions result from the fact that the magnification factor and receptive field size change monotonically from the fovea to the periphery. The retinotopic map has been approximately described by a complex logarithm [63]. It is not hard to produce global distortions with the elastic net and self-organising map, simply by systematically changing the density of feature points across the input space (e.g. [64]). More troubling are the local distortions, different from the scatter<sup>3</sup> of receptive fields, which result from discontinuities of the retinotopic map and are matched with the discontinuities (or pinwheels) of the orientation map [3]. That is, along a tangential penetration the receptive field centre varies relatively smoothly from neuron to neuron and neighbouring receptive fields overlap considerably; but receptive fields of cells lying on opposite sides of a pinwheel do not overlap. Das and Gilbert [3] also found that the rates of variation of receptive field centre (normalised by receptive field size) and of orientation preference were positively correlated. These two findings pose problems, so far unsolved, for cortical development models. On the one hand, some types of models, such as correlational models [45], have relied so far on a perfect retinotopy and have not addressed topography distortions. On the other hand, dimension reduction models, such as the elastic net and self-organising map, do not assume a perfect retinotopy, but simulations result precisely in anticorrelations rather than positive correlations: when one stimulus variable varies, the others tend to remain constant [30]. Why this is so is unknown: it is possible to have a rich behaviour in the joint variation of stimulus variables (e.g. all variables varying at the same cortical location) while still satisfying coverage and continuity. Thus, the likely reason must lie in the specific formulation of the objective function or the training algorithm, that tends to single out solutions with anticorrelations. One factor could be the discrete character of the elastic net and self-organising map, both being vector quantisation methods, which limits the different gradient values that may arise in a local optimum.

The results obtained by Das and Gilbert [3] used a coarse sample, since neighbouring cells were recorded at approximately intervals of 50  $\mu\text{m}$  for tangential penetrations and of 400  $\mu\text{m}$  for a two-dimensional grid. Ideally, the correlations and discontinuities should be obtained from a two-dimensional sample on a wide area of the cortex at a much finer cortical separation in order to obtain meaningful gradients as a function of the cortical location. This would also show the distribution over cortex of the effects of discontinuity matching and positive correlation, which may not be uniform in view of the global distortions mentioned above. Unfortunately such area-wide measurements, that have advanced our knowledge of the orientation and ocular dominance maps considerably, are not currently possible for the retinotopic map. Das and Gilbert [3] claim that their results are robust against scatter because the latter is very small. However, the scatter could affect their results in a different way: while the retinotopic map obtained from measurements of a collection of cells at roughly the same cortical depth is noisy due to the inherent scatter of cell receptive fields, a more homogeneous map might be obtained as the aggregate receptive field resulting from cells in the same column. However, our inability to make spatial and volumetric measurements of this kind currently prevents experimental investigation of this issue.

### 6.2 Activity-independent mechanisms in column development

Earlier we briefly reviewed the large body of evidence for a role for neural activity (both visually evoked and spontaneous) in visual cortical map formation and plasticity. However, in the past few years a number of pieces

---

<sup>3</sup>Quantitative estimates of this scatter vary, but recent results place its standard deviation between 0.1 and 0.5 times the receptive field size (e.g. [3, 65]).

of experimental data have appeared which challenge to varying extents the hypothesis that initial column development is activity dependent. One reason for the rise of interest in activity-independent explanations for map formation in the cortex is the dramatic increase in our understanding of activity-independent mechanisms of axonal targetting. Since 1994 several large families of molecules, many previously unknown, have been identified to play crucial roles in the development of neuronal connections. These include the netrins, semaphorins, slits and ephrins (reviewed in [66, 67]). Molecules which may be particularly relevant to understanding map development in visual cortex are the ephrins, signalling through receptors of the Eph family. Low anterior to high posterior gradients of ephrins exist in the optic tectum and its mammalian homologue the superior colliculus, while low nasal to high temporal gradients of Eph receptors exist in the retina ([68, 69]; reviewed in [53]). Extensive evidence suggests that these gradients play a crucial role in guiding retinal ganglion cell axons to their targets (e.g. [70, 71]; other data reviewed in [72]). However, although ephrins have also been found in the lateral geniculate nucleus [73] and in somatosensory cortex [74], as yet there is no direct evidence that they play a role in guiding axons to appropriate targets in primary visual cortex. In particular, in order to control column formation one would expect to find ephrins in an initially patchy distribution in visual cortex, and this has not so far been observed.

**Ocular dominance maps** To re-examine the issue of whether the retinae or retinal activity are required for the establishment of ocular dominance columns, Crowley and Katz [75] enucleated ferrets<sup>4</sup> very early in life and let them develop. To visualise patterns of lateral geniculate nucleus (LGN) axons in maturity to determine whether ocular dominance segregation occurs they used anterograde LGN injections (i.e., injection of a tracer into eye-specific cells in the LGN) and retrograde cortical injections. They found that, with or without information derived from the retina, geniculocortical axons organised into discrete ocular dominance stripes. Crowley and Katz [76] further found that columns are not present at birth but appear as early as 16 days later (equivalent to a week before birth in cats). Therefore, Crowley and Katz [75] removed the eyes *before* columns form. In addition, when they removed just one eye at an age when LGN axons have innervated the cortex but before columns have formed, normal-looking ocular dominance columns still resulted (with the same periodicity: neither shrunk nor expanded). These results suggest that the establishment of ocular dominance columns and the plasticity of ocular dominance columns are two temporally different phases of visual cortex development.

**Orientation maps** Using optical imaging of intrinsic signals and single-unit microelectrode penetrations in both eyes at P15 (postnatal day 15) or younger, Crair et al. [77] found that the orientation map forms before P14 in both normal and binocularly deprived (BD; by bilateral lid suture) cats. The similarity between the orientation maps of each eye for normal and BD cats varied with age. From P0 to around P21: the similarity increases monotonically in the same way for both normal and BD cats. From around P21: maps remain identical for normal cats, but progressively dissimilar for BD cats. They thus concluded that patterned visual experience is required for the maintenance of orientation selectivity rather than for the initial development of the orientation map and the deterioration of maps coincides with the critical period. Orientation columns emerge independently of patterned visual experience during the second postnatal week and patches of ipsilateral eye responses appear early in the third week. Experience then makes responses to become stronger, more selective and nearly equal for both eyes by the beginning of the fourth week; with continued BD, ipsilateral eye responses never become very strong or selective.

In another experiment, Gödecke and Bonhoeffer [78] investigated the influence of activity on the fact that orientation maps are precisely matched for both eyes (at least for binocularly-driven neurons)—which is essential for disparity detection and so for stereoscopic vision. They raised kittens with reverse suturing, so that both eyes were never able to see at the same time. They found that the orientation maps in area 18 for both eyes were identical, as in normal kittens (except for minor differences such as a slight shift of some pinwheels that they imputed to technical limitations). They argued that this effect cannot be due to spontaneous retinal activity, since this would have to be synchronised between both eyes. They thus deduced that correlated visual input is not required for the alignment of orientation maps.

However, Wolf et al. [79] showed that it is possible to replicate the results of Gödecke and Bonhoeffer [78] in simulations with self-organising maps by using a cortex with a specific shape: a narrow, elongated cortex results in matched orientation maps while a square cortex results in unmatched orientation maps. Wolf et al. [79] explain this in terms of symmetry-breaking and pattern formation in physical systems, where the qualitative behaviour depends generally on the ratio between system size and characteristic wavelength of the emerging structure.

---

<sup>4</sup>Ferrets are ideal for experiments for two reasons: they have robust ocular dominance columns and well-defined critical periods (like cats); and their nervous systems at birth are not yet developed (e.g. 3 weeks less developed than those in cats, in which ocular dominance columns form before birth). This allows the detection of earlier developmental events.

In another experiment, Weliky and Katz [80] examined the effect of perturbed patterns of neural activity on orientation maps in ferret V1. They implanted a stimulating cuff around the optic nerve from one eye (the other eye was removed), and stimulated it for about 2 seconds every 20 seconds from around P16 to P42. While they found effects on the receptive field structure of individual neurons, the overall structure of the orientation map looked apparently normal. Although this could be interpreted as evidence that overall orientation map structure is not activity dependent, it is important to note that the aberrant stimulation occurred for only a small proportion of the total time (see Goodhill [81] for further discussion).

These studies and others suggest that the role of activity in columnar development is less determining than originally thought. However, it is important to note that even though the eyes are deprived or removed, spontaneous activity is still likely to be present in both lateral geniculate nucleus and cortex. In the LGN spontaneous activity emerges from multiple mechanisms, including endogenous network oscillations and feedback connections [82]. Thus, the LGN does not simply relay patterns of retinal activity to the cortex, but rather this activity is reshaped and transformed by corticothalamic interactions.

**Models incorporating neuronal activity, molecular guidance cues and gene expression** Models such as the elastic net and the self-organising map have been very successful at accounting for much of the observed phenomenology of visual cortical maps. Although they are conventionally thought of as implementing activity-dependent learning rules, it is important to remember that the elastic net was originally derived from a mapping model based entirely on activity-independent molecular mechanisms [49]. Nevertheless, it seems reasonable to conclude that the early stages of cortical map formation are driven by a combination of molecular guidance cues and gene expression in addition to neural activity, and that a different type of model explicitly including all three factors may be necessary to account for map formation. One such kind of models are *gene networks* [83]. A gene network is a cluster of genes in which (1) the expression of the genes in the cluster is affected by specific stimuli, such as exposure to a hormone or neurotransmitter, and (2) the protein products of some members of the cluster act as transcription factors that regulate, positively or negatively, the expression of other members. The total array of genetic regulation in neurons and other cells is a gene network with a large number (of the order of thousands) of interactions. Current models of gene networks describe the rates of change of the concentrations of gene products (mRNAs and proteins) with ordinary differential equations, as a function of the levels of transcription factors or other effector molecules and can incorporate stochastic fluctuations in molecule numbers. Logical networks, where the expression of each gene in the network is assumed to be either ON or OFF, also have been proposed. These models give rise to a rich variety of qualitative nonlinear behaviours, including multistability and oscillations. However, gene network models present some difficulties. Modelling differential equations requires short time steps and so a high computational cost, which may make them impractical for large gene networks or lengthy processes such as development of tissues or organisms. Models of specific gene networks need to be based, insofar as possible, on values of biochemical parameters measured in vivo, which is difficult. Hence, new methods for gathering detailed data are necessary. Finally, so far they mostly do not take into account the spatial organisation of gene expression, which is essential to explain biological pattern formation and in particular cortical map formation. Models for the *Drosophila* segmentation problem such as those of Sharp and Reinitz [84] and von Dassow et al. [85] are a promising step in this direction.

## References

- [1] Swindale, N. V. (1996) The development of topography in the visual cortex: A review of models. *Network* **7**, 161–247.
- [2] Swindale, N. V. (2001) Development of ocular dominance stripes, orientation selectivity and orientation columns. In *Modelling Neural Development* (van Ooyen, A., ed.), MIT Press, Cambridge, MA, chapter 12.
- [3] Das, A. and Gilbert, C. D. (1997) Distortions of visuotopic map match orientation singularities in primary visual cortex. *Nature* **387**, 594–598.
- [4] Hubel, D. H. and Wiesel, T. N. (1977) Functional architecture of the macaque monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* **198**, 1–59.
- [5] Blasdel, G. G. and Salama, G. (1986) Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* **321**, 579–585.

- [6] Bonhoeffer, T. and Grinvald, A. (1991) Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature* **353**, 429–431.
- [7] Issa, N. P., Trepel, C., and Stryker, M. P. (2000) Spatial frequency maps in cat visual cortex. *J. Neurosci.* **20**, 8504–8514.
- [8] Shoham, D., Hübener, M., Schulze, S., Grinvald, A., and Bonhoeffer, T. (1997) Spatio-temporal frequency domains and their relation to cytochrome oxidase staining in cat visual cortex. *Nature* **385**, 529–533.
- [9] Hübener, M., Shoham, D., Grinvald, A., and Bonhoeffer, T. (1997) Spatial relationships among three columnar systems in cat area 17. *J. Neurosci.* **17**, 9270–9284.
- [10] Everson, R. M., Prashanth, A. K., Gabbay, M., Knight, B. W., Sirovich, L., and Kaplan, E. (1998) Representation of spatial frequency and orientation in the visual cortex. *Proc. Natl. Acad. Sci. USA* **95**, 8334–8338.
- [11] Bartfeld, E. and Grinvald, A. (1992) Relationships between orientation-preference pinwheels, cytochrome oxidase blobs, and ocular-dominance columns in primate striate cortex. *Proc. Natl. Acad. Sci. USA* **89**, 11905–11909.
- [12] Obermayer, K. and Blasdel, G. G. (1993) Geometry of orientation and ocular dominance columns in monkey striate cortex. *J. Neurosci.* **13**, 4114–4129.
- [13] LeVay, S., Connolly, M., Houde, J., and Van Essen, D. C. (1985) The complete pattern of ocular dominance stripes in the striate cortex and visual field of the macaque monkey. *J. Neurosci.* **5**, 486–501.
- [14] Horton, J. C. and Hocking, D. R. (1996) Intrinsic variability of ocular dominance column periodicity in normal macaque monkeys. *J. Neurosci.* **16**, 7228–7339.
- [15] Obermayer, K. and Blasdel, G. G. (1997) Singularities in primate orientation maps. *Neural Comput.* **9**, 555–575.
- [16] Löwel, S., Schmidt, K. E., Kim, D.-S., Wolf, F., Hoffsummer, F., Singer, W., and Bonhoeffer, T. (1998) The layout of orientation and ocular dominance domains in area 17 of strabismic cats. *Eur. J. Neurosci.* **10**, 2629–2643.
- [17] Müller, T., Stetter, M., Hübener, M., Sengpiel, F., Bonhoeffer, T., Gödecke, I., Chapman, B., Löwel, S., and Obermayer, K. (2000) An analysis of orientation and ocular dominance patterns in the visual cortex of cats and ferrets. *Neural Comput.* **12**.
- [18] Kaschube, M., Wolf, F., Geisel, T., and Löwel, S. (2000) Quantifying the variability of patterns of orientation domains in the visual cortex of cats. *Neurocomputing* **32–33**, 415–423.
- [19] Hubel, D. H., Wiesel, T. N., and LeVay, S. (1977) Plasticity of ocular dominance columns in monkey striate cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **278**, 377–409.
- [20] Shatz, C. J. (1990) Impulse activity and the patterning of connections during CNS development. *Neuron* **5**, 745–756.
- [21] Katz, L. C. and Shatz, C. J. (1996) Synaptic activity and the construction of cortical circuits. *Science* **274**, 1133–1138.
- [22] Stryker, M. P. and Harris, W. A. (1986) Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *J. Neurosci.* **6**, 2117–2133.
- [23] Shatz, C. J. and Stryker, M. P. (1978) Ocular dominance in layer IV of the cat’s visual cortex and the effects of monocular deprivation. *J. Physiol.* **281**, 267–283.
- [24] LeVay, S., Wiesel, T. N., and Hubel, D. H. (1980) The development of ocular dominance columns in normal and visually deprived monkeys. *J. Comp. Neurol.* **191**, 1–51.
- [25] Chapman, B., Jacobson, M. D., Reiter, H. O., and Stryker, M. P. (1986) Ocular dominance shift in kitten visual cortex caused by imbalance in retinal electrical activity. *Nature* **324**, 154–156.

- [26] Rittenhouse, C. D., Shouval, H. Z., Paradiso, M. A., and Bear, M. F. (1999) Monocular deprivation induces homosynaptic long-term depression in visual cortex. *Nature* **397**, 347–350.
- [27] Hubel, D. H. and Wiesel, T. N. (1965) Binocular interaction in striate cortex of kittens reared with artificial squint. *J. Neurophysiol.* **28**, 1041–1059.
- [28] Löwel, S. (1994) Ocular dominance column development: Strabismus changes the spacing of adjacent columns in cat visual cortex. *J. Neurosci.* **14**, 7451–7468.
- [29] Sengpiel, F., Gödecke, I., Stawinski, P., Hübener, M., Löwel, S., and Bonhoeffer, T. (1998) Intrinsic and environmental factors in the development of functional maps in cat visual cortex. *Neuropharmacology* **37**, 607–621.
- [30] Durbin, R. and Mitchison, G. (1990) A dimension reduction framework for understanding cortical maps. *Nature* **343**, 644–647.
- [31] Durbin, R. and Willshaw, D. (1987) An analog approach to the traveling salesman problem using an elastic net method. *Nature* **326**, 689–691.
- [32] Goodhill, G. J. and Willshaw, D. J. (1990) Application of the elastic net algorithm to the formation of ocular dominance stripes. *Network* **1**, 41–59.
- [33] Kohonen, T. K. (1982) Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43**, 59–59.
- [34] Kohonen, T. K. (1995) *Self-Organizing Maps*. Springer-Verlag, Berlin.
- [35] Erwin, E., Obermayer, K., and Schulten, K. (1995) Models of orientation and ocular dominance columns in the visual cortex: A critical comparison. *Neural Comput.* **7**, 425–468.
- [36] Tikhonov, A. N. and Arsenin, V. Y. (1977) *Solutions of Ill-Posed Problems*. Scripta Series in Mathematics. John Wiley & Sons, New York, London, Sydney. Translation editor: Fritz John.
- [37] Erwin, E., Obermayer, K., and Schulten, K. (1992) Self-organizing maps: Ordering, convergence properties and energy functions. *Biol. Cybern.* **67**, 47–55.
- [38] Utsugi, A. (1997) Hyperparameter selection for self-organizing maps. *Neural Comput.* **9**, 623–635.
- [39] Dayan, P. (1993) Arbitrary elastic topologies and ocular dominance. *Neural Comput.* **5**, 392–401.
- [40] Swindale, N. V. (1991) Coverage and the design of striate cortex. *Biol. Cybern.* **65**, 415–424.
- [41] Swindale, N. V., Shoham, D., Grinvald, A., Bonhoeffer, T., and Hübener, M. (2000) Visual cortex maps are optimised for uniform coverage. *Nat. Neurosci.* **3**, 822–826.
- [42] Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London, New York.
- [43] Bishop, C. M., Svensén, M., and Williams, C. K. I. (1998) GTM: The generative topographic mapping. *Neural Comput.* **10**, 215–234.
- [44] Goodhill, G. J. and Sejnowski, T. J. (1997) A unifying objective function for topographic mappings. *Neural Comput.* **9**, 1291–1303.
- [45] Miller, K. D., Keller, J. B., and Stryker, M. P. (1989) Ocular dominance column development: Analysis and simulation. *Science* **245**, 605–615.
- [46] Goodhill, G. J. (1998) The influence of neural activity and intracortical connectivity on the periodicity of ocular dominance stripes. *Network* **9**, 419–432.
- [47] Lawler, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G., and Shmoys, D. B. (1986) *The Travelling Salesman Problem*. John Wiley & Sons, Chichester, England.
- [48] Willshaw, D. J. and von der Malsburg, C. (1976) How patterned neural connections can be set up by self-organization. *Proc. R. Soc. Lond. B Biol. Sci.* **194**, 431–445.

- [49] Willshaw, D. J. and von der Malsburg, C. (1979) A marker induction mechanism for the establishment of ordered neural mappings: its application to the retinotectal problem. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **287**, 203–243.
- [50] Durbin, R., Szeliski, R., and Yuille, A. (1989) An analysis of the elastic net approach to the traveling salesman problem. *Neural Comput.* **1**, 348–358.
- [51] Rose, K. (1998) Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proc. of the IEEE* **86**, 2210–2239.
- [52] Goodhill, G. J. (1992) Correlations, competition, and optimality: Modelling the development of topography and ocular dominance. Cognitive Science Research Paper CSRP 226, Sussex University.
- [53] Goodhill, G. J. and Richards, L. J. (1999) Retinotectal maps: Molecules, models and misplaced data. *Trends Neurosci.* **22**, 529–534.
- [54] Yuille, A. L., Kolodny, J. A., and Lee, C. W. (1996) Dimension reduction, generalized deformable models and the development of ocularity and orientation. *Neural Networks* **9**, 309–319.
- [55] Goodhill, G. J. (1993) Topography and ocular dominance: A model exploring positive correlations. *Biol. Cybern.* **69**, 109–118.
- [56] Goodhill, G. J. and Löwel, S. (1995) Theory meets experiment: Correlated neural activity helps determine ocular dominance column periodicity. *Trends Neurosci.* **18**, 437–439.
- [57] Goodhill, G. J. and Willshaw, D. J. (1994) Elastic net model of ocular dominance: Overall stripe pattern and monocular deprivation. *Neural Comput.* **6**, 615–621.
- [58] Goodhill, G. J., Bates, K. R., and Montague, P. R. (1997) Influences on the global structure of cortical maps. *Proc. R. Soc. Lond. B Biol. Sci.* **264**, 649–655.
- [59] Goodhill, G. J. and Cimponeriu, A. (2000) Analysis of the elastic net applied to the formation of ocular dominance and orientation columns. *Network* **11**, 153–168.
- [60] Hoffsümmer, F., Wolf, F., Geisel, T., Löwel, S., and Schmidt, K. (1995) Sequential bifurcation of orientation- and ocular dominance maps. In *Proc. of the Fifth Int. Conf. on Artificial Neural Networks (ICANN95)* (Fogelman-Soulie, F. and Gallinari, R., eds.). EC2 & Cie, Paris, France, volume 1, pp. 535–540.
- [61] Hoffsümmer, F., Wolf, F., Geisel, T., Löwel, S., and Schmidt, K. (1996) Sequential bifurcation and dynamic rearrangement of columnar patterns during cortical development. In *Computational Neuroscience: Trends in Research 1995* (Bower, J. M., ed.). Academic Press, New York, pp. 197–202.
- [62] Wolf, F. and Geisel, T. (1998) Spontaneous pinwheel annihilation during visual development. *Nature* **395**, 73–78.
- [63] Schwartz, E. L. (1994) Computational studies of the spatial architecture of primate visual cortex: Columns, maps, and protomaps. In *Primary Visual Cortex in Primates* (Peters, A. and Rockland, K. S., eds.), Plenum Press, New York and London, volume 10 of *Cerebral Cortex*, chapter 9, pp. 359–411.
- [64] Wolf, F., Bauer, H.-U., and Geisel, T. (1994) Formation of field discontinuities and islands in visual cortical maps. *Biol. Cybern.* **70**, 525–531.
- [65] Hetherington, P. A. and Swindale, N. V. (1999) Receptive field and orientation scatter studied by tetrode recordings in cat area 17. *Vis. Neurosci.* **16**, 637–652.
- [66] Tessier-Lavigne, M. and Goodman, C. S. (1996) The molecular biology of axon guidance. *Science* **274**, 1123–1133.
- [67] Mueller, B. K. (1999) Growth cone guidance: First steps towards a deeper understanding. *Annu. Rev. Neurosci.* **22**, 351–388.
- [68] Cheng, H. J., Nakamoto, M., Bergemann, A. D., and Flanagan, J. G. (1995) Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. *Cell* **82**, 371–381.

- [69] Drescher, U., Kremoser, C., Handwerker, C., Loschinger, J., Noda, M., and Bonhoeffer, F. (1995) In vitro guidance of retinal ganglion cell axons by RAGS, a 25 KDa tectal protein related to ligands for Eph receptor tyrosine kinases. *Cell* **82**, 359–370.
- [70] Feldheim, D. A., Kim, Y.-I., Bergemann, A. D., Frisén, J., Barbacid, M., and Flanagan, J. G. (2000) Genetic analysis of Ephrin-A2 and Ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping. *Neuron* **25**, 563–574.
- [71] Goodhill, G. J. (2000) Dating behavior of the retinal ganglion cell. *Neuron* **25**, 501–503.
- [72] Flanagan, J. G. and Vanderhaeghen, P. (1998) The ephrins and Eph receptors in neural development. *Annu. Rev. Neurosci.* **21**, 309–345.
- [73] Feldheim, D. A., Vanderhaeghen, P., Hansen, M. J., Frisén, J., Lu, Q., Barbacid, M., and Flanagan, J. G. (1998) Topographic guidance labels in a sensory projection to the forebrain. *Neuron* **21**, 1303–1313.
- [74] Vanderhaeghen, P., Lu, Q., Prakash, N., Frisén, J., Walsh, C. A., Frostig, R. D., and Flanagan, J. G. (2000) A mapping label required for normal scale of body representation in the cortex. *Nat. Neurosci.* **3**, 358–365.
- [75] Crowley, J. C. and Katz, L. C. (1999) Development of ocular dominance columns in the absence of retinal input. *Nat. Neurosci.* **2**, 1125–1130.
- [76] Crowley, J. C. and Katz, L. C. (2000) Early development of ocular dominance columns. *Science* **290**, 1321–1324.
- [77] Crair, M. C., Gillespie, D. C., and Stryker, M. P. (1998) The role of visual experience in the development of columns in cat visual cortex. *Science* **279**, 566–570.
- [78] Gödecke, I. and Bonhoeffer, T. (1996) Development of identical orientation maps for two eyes without common visual experience. *Nature* **379**, 251–254.
- [79] Wolf, F., Bauer, H.-U., Pawelzik, K., and Geisel, T. (1996) Organization of the visual cortex. *Nature* **382**, 306–307.
- [80] Weliky, M. and Katz, L. C. (1997) Disruption of orientation tuning in visual cortex by artificially correlated neuronal activity. *Nature* **386**, 680–685.
- [81] Goodhill, G. J. (1997) Stimulating issues in cortical map development. *Trends Neurosci.* **20**, 375–376.
- [82] Weliky, M. and Katz, L. C. (1999) Correlational structure of spontaneous neuronal activity in the developing lateral geniculate nucleus in vivo. *Science* **285**, 599–604.
- [83] Smolen, P., Baxter, D. A., and Byrne, J. H. (2000) Mathematical modeling of gene networks. *Neuron* **26**, 567–580.
- [84] Sharp, D. H. and Reinitz, J. (1998) Prediction of mutant expression patterns using gene circuits. *Biosystems* **47**, 79–90.
- [85] von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. (2000) The segment polarity network is a robust developmental module. *Nature* **406**, 188–192.