Chapter 3

Previous Models

3.1 Introduction

In the previous chapter we reviewed a range of experimental data, and also outlined the main organizing principles that have been proposed to account for that data. In this chapter, we first discuss informal models for how these organizing principles are realised, and we then review a number of computational models. Each of the latter specifies a mechanism or a combination of mechanisms for map formation in a form precise enough for implementation as a computer program, and thus helps to make clear the capabilities and limitations of each mechanism.

Some important questions to be borne in mind when considering these models are as follows:

- How are presynaptic locations specified?
- How is information about presynaptic positions communicated to the postsynaptic sheet?
- How are synaptic strengths updated?
- What cooperative and competitive interactions take place to ensure a smooth map across the postsynaptic sheet?

Although there are several models, they are all based on quite a small number of principal mechanisms, usually some form of the Hebb rule. The defining characteristics of a particular model are therefore the precise mathematical forms adopted for each of these mechanisms. One such characteristic that will be important later in this thesis is the type of normalization used to limit synaptic strengths.

We judge each model according to the criteria outlined in the introduction: their ability to develop topography and ocular dominance from the same mechanisms, their biological plausibility, their success in the presence of positive correlations between the eyes, the properties of the ocular dominance stripes produced, and the predictions they make regarding stripe width.

3.2 Informal models for topography

Over the past several decades there has been much discussion and controversy over the mechanisms of retinotectal map formation in amphibians and fish. We take as our starting point the hypothesis of chemical matching between pre- and postsynaptic cells, discussed first by Langley (e.g. [Langley 1895], see also [Brown et al 1991]) and then in much greater detail by Sperry.

3.2.1 Chemospecificity

To account for the experimental facts that retinal axons can find their normal sites of termination in the tectum even when optic nerve fibres are artificially scrambled, Sperry postulated a scheme of chemical matching between retina and tectum (see [Sperry 1963] for review). He describes it thus:

> The establishment and maintenance of synaptic associations [is] conceived to be regulated by highly specific cytochemical affinities that arise systematically among the different types of neurons involved via self-differentiation, induction through terminal contacts, and embryonic gradient effects [Sperry 1963].

This "chemospecificity" hypothesis has been extremely fruitful in both providing an informal account of various experimental results, and inspiring theoretical models. A distinction can be made between the three related chemical mechanisms mentioned in the above quotation. (1) The idea of self-differentiation is that some process takes place whereby each retinal and tectal cell assumes a unique chemical identity, and there is a process of matching between appropriately located retinal and tectal cells. One possibility is that there might be no relation between the labels associated with neighbouring retinal or tectal cells: this is sometimes referred to as a "mosaic"of labels. (2) An alternative to mosaics is explicitely mentioned by Sperry: that there are *gradients* of chemical differences in both retina and tectum, and it is the gradients rather than individual cells that are matched. At least two gradients are needed to match two-dimensional structures. (3) The notion of induction through terminal contacts is that the chemical identities of retinal cells are passed via retinal axons into the tectum. It is not clear from the above whether Sperry intended this to be simply a way for the chemical identities of retinal and tectal cells to be brought together for matching purposes, or a map-forming mechanism in itself. We will discuss computational models of map formation based on each of these three mechanisms later.

Besides the unscrambling of abnormal projections in Sperry's own experiments, the chemospecificity hypothesis fits well with the eye rotation results of [Hunt & Jacobson 1972] and others (for a fuller review of data supporting this hypothesis see [Gaze & Keating 1972]). However, subsequent experimental data began to cause problems for the hypothesis. Firstly, the expansion and compression results of [Yoon 1971] and others (see for example [Gaze & Keating 1972, Udin & Fawcett 1988]) showed that the matching scheme cannot be "rigid": even though there are a mismatched number of termination sites, a smooth map still forms. An interpretation of these results favoured by Yoon was that "respecification" of retinal or tectal cells occurs after surgical intervention, so that each cell develops a new label appropriate for the new context. However, as pointed out by [Gaze & Keating 1972] this is a post-hoc explanation, and has no predictive power (see also [Prestige & Willshaw 1975]). In addition, the respecification hypothesis requires that the pre- or postsynaptic sheet has direct knowledge about how much the other sheet was altered by the surgery. The expansion and compression results mentioned above have been referred to by [Gaze & Keating 1972] as "systems-matching". They put forward the important hypothesis that this could occur by a competitive mechanism: retinal fibres compete for space in the tectum, and since each fibre has equal competitive strength they spread out to fill the entire space available, be that smaller or larger than the normal case.

We have presented these issues regarding the role of chemical specificity in map formation here to aid understanding of the computational models reviewed in this chapter that are based on these ideas. More recently debate has continued, fueled by the results of ever more exotic surgical manipulations (for reviews see [Trisler 1982, Meyer 1982(a), Schmidt 1985, Udin & Fawcett 1988, Fraser & Perkel 1990]), and also direct evidence for the existence of gradients of chemical markers in retina and tectum (reviewed in [Gottlieb & Glaser 1981]). We do not discuss this further here, since our primary concern in the following chapters is the effects of correlated retinal activity on map formation.

3.2.2 Correlated activity

The idea that correlations in presynaptic activity could drive the formation of an ordered map was discussed by several experimentalists before being formulated as a computational model in [Willshaw & von der Malsburg 1976]. The informal notion behind the model is as follows.

Consider a postsynaptic cell $\mathtt{p}$ receiving connections from two presynaptic cells, $\mathtt{a}$ and $\mathtt{b}$. The following assumptions are made:

- The activity of $\mathtt{p}$ is related (perhaps by a threshold function) to the sum of the activities of $\mathtt{a}$ and $\mathtt{b}$ weighted by their connection strengths.
- Connection strengths are updated according to a Hebbian rule such that at each step the strength of each connection is increased in proportion to the product of pre- and postsynaptic activity.
- The activities of $\mathtt{a}$ and $\mathtt{b}$ are highly correlated only if they are close together in the presynaptic sheet.

There are now two possibilities for the development of the two connection strengths in this system. If $\mathtt{a}$ is near to $\mathtt{b}$, they will tend to be active at the same time, and thus be likely to make $\mathtt{p}$ active at this time. This will lead to both connection strengths being increased. In other words, $\mathtt{a}$ and $\mathtt{b}$ *cooperate*. However, if $\mathtt{a}$ and $\mathtt{b}$ are widely separated then their activities are not correlated, and the two connection strengths evolve independently. If a normalization rule is introduced to limit connection strengths (e.g. the sum of connection strengths for cell $\mathtt{p}$ is maintained at a constant), then in the latter case $\mathtt{a}$ and $\mathtt{b}$ actively *compete*. Thus we see how locally correlated activity in the presynaptic sheet has a tendency to cause neighbouring presynaptic cells to be connected to the same postsynaptic cell.

However, two other components now need to be added to explain map formation. So far there is no pressure for neighbouring postsynaptic cells to be connected to (1) different but (2) neighbouring cells in the presynaptic sheet. The first can be achieved by limiting the total synaptic strength available to each postsynaptic cell (this will become clearer when we discuss [Prestige & Willshaw 1975] below). The second can be achieved by introducing lateral interactions in the postsynaptic sheet. The principal requirement is that neighbouring postsynaptic cells excite one another: this is equivalent to specifying that their activities be correlated. Then, by a similar argument to that for correlations between presynaptic cells, neighbouring postsynaptic cells will have a tendency to receive connections from the same presynaptic cell, or more generally, presynaptic cells with correlated activities, i.e. those that are close together. In this way an activity-based learning mechanism can lead to the formation of a locally-ordered map.

By itself, such a mechanism can generate only local order. One additional assumption that assists the establishment of global order is that there exists some other mechanism to specify an overall bias or polarity for the map. The experimental evidence supports the notion that map formation is activity-dependent but that there is also some other mechanism that establishes a crude map in the absence of activity [Meyer 1983]. Many theoreticians have postulated a two-stage process whereby a crude initial map is formed by matching of chemical gradients, followed by refinement of this map due to activity-based mechanisms (first proposed in [Willshaw & von der Malsburg 1976]). The mathematical effects of this initial symmetry breaking are discussed in [Häussler & von der Malsburg 1983].

3.2.3 Summary

Cytochemical matching, normalization, and a combination of correlated presynaptic activity and lateral excitation of neighbouring postsynaptic cells, are the main mechanisms that have been put forward to explain topographic map formation in the retinotectal and retinocortical systems. We will see shortly various ways in which these qualitative notions have been formalized into computational models, and how different implementations of these ideas can lead to different results.

3.3 Informal models for ocular dominance

Two issues involved in the development of ocular dominance are: (1) Why should tectal or cortical cells become monocular despite being initially binocularly innervated, and (2) Given that half the cells are dominant for one eye and half for the other, why should these cells be laid out in a periodic pattern of stripes?

The first issue was discussed by [Hubel & Wiesel 1965] (see also [Stent 1973]). In their "artificial strabismus" experiments they made kittens unable to focus both eyes on the same point, in effect severely reducing the normal correlations between cells at corresponding positions in the two eyes. This led to an increased tendency for cortical cells to be monocular. [Hubel & Wiesel 1965] suggested that this could be explained by similar arguments to those made above for the topography case: uncorrelated afferents converging on the same postsynaptic cell compete. They refer to a "lack of synergy between the two afferent paths": following from our discussion above it is clear that Hubel and Wiesel were assuming some form of competition (e.g. normalization) between afferents. However, this argument does not account for how monocularity arises in normal cats and monkeys, where there is a high degree of correlation between the two eyes. Also, we note that both [Hubel & Wiesel 1965] and [Stent 1973] refer to "asynchronous" rather than uncorrelated activities in the two afferent pathways. These are not equivalent: asynchronous means "occuring at different times" (Collins English Dictionary, third edition), which implies anticorrelation between the two pathways. Uncorrelated afferents may still sometimes be synchronous. If the two eyes move effectively independently, it seems more reasonable to suppose that activity is uncorrelated than anticorrelated. We emphasise this point here because the precise nature of the correlations between the two eyes will be important in later chapters.

Moving now to the second issue, an idea which we will refer to as the "balancing" model was first introduced in [LeVay et al 1975]:

> How is the columnar pattern arrived at during development? The segregation of the two sets of afferents into alternating bands might be understood as the consequence of two conflicting tendencies, the specification of the two sets to occupy the same cortical space according to a single topographic map, and an opposing tendency for grouping within each set, as if caused by mutual repulsion between members of the two sets. The first principle would tend to increase the degree of interpenetration of the two sets, that is, to reduce the maximum distance from any cortical point to the nearest column border. If this distance....is kept constant, then a repulsion between the two sets would tend to produce an arrangement with the smallest total length of interface. Alternating

bands give a shorter interface than any other arrangement such as a checkerboard or circular islands in a matrix [LeVay et al 1975].

The idea is that the final pattern is a balance between two contradictory constraints: that corresponding cells in the two eyes are mapped close together, and that afferents from neighbouring positions in each eye are kept together. This notion was also proposed more specifically to explain ocular dominance stripe formation in the retinotectal system in [Constantine-Paton 1983]. She suggests that the first tendency is caused by chemospecific matching, and states "Only the striped pattern optimizes each retina's ability to contact all regions of the tectal lobe while simultaneously minimizing the borders where ganglion cells must synapse adjacent to the non-neighbouring cells of the second retina".

The notion of minimal boundary length was also mentioned by von der Malsburg in [von der Malsburg & Willshaw 1976], and explored more quantitatively in [von der Malsburg 1979]. He imagined a postsynaptic sheet with periodic boundary conditions (i.e. a torus), and considered the topologically different ways in which this can be divided into two regions by closed lines. He showed that the topologies with the shortest boundary lengths are circles or stripes, and went on to calculate the minimal boundary lengths for both cases as a function of α, the fraction of the plane taken up by one ocularity. For $1/\pi < \alpha < 1 - 1/\pi$, stripes have shorter boundary length than circles: the normal case of $\alpha = \frac{1}{2}$ is in this range. [von der Malsburg 1979] also predicted that the determiner of stripe width in the balancing model is the relative strength of these two tendencies.

Neither [LeVay et al 1975, Constantine-Paton 1983, von der Malsburg 1979] put forward any convincing justification for why the system should attempt to minimize the length of stripe boundaries. [LeVay et al 1975] refer to a repulsive force between opposite-eye fibres, but without saying how this force arises. [Constantine-Paton 1983] implies that there is a "cost" incurred if neighbouring tectal cells represent different eyes. The isogenic-eye experiments [Ide et al 1983] seem to rule out a global chemical difference between the two eyes as being a factor in stripe formation. Whether or not a mechanism based on correlated activity could implement a suitable cost function depends on the precise form of correlations between the two eyes and the nature of the competition in the tectum/cortex.

3.4 Computational models - topography

The qualitative models of topography described above were systematically developed as computational models in [Prestige & Willshaw 1975, Willshaw & von der Malsburg 1976, von der Malsburg & Willshaw 1977, Willshaw & von der Malsburg 1979]. We describe these in some detail, and then discuss other models that have been presented.

Some of the models are phrased in terms of mappings in lower vertebrates from retina to tectum, some in terms of mappings in higher vertebrates from LGN to cortex, and some in terms of a general mapping between a pre- and postsynaptic sheet. However, most of the models described here do not rely on the details of either the retino-tectal or geniculo-cortical system beyond the general idea of a mapping between a pre- and postsynaptic sheet. The difference lies in which set of data the model attempts to explain. If the emphasis is on higher vertebrates, the data usually considered is that regarding ocular dominance and deprivation results, whereas for lower vertebrates it is expansion, compression, translocation and rotation of maps. For completeness, in the latter case we briefly indicate which set of data each model can account for, though these results are not part of the criteria for model evaluation listed in the introduction. At the end of the chapter we summarize in tabular form the important properites of each model.

3.4.1 Prestige and Willshaw

[Prestige & Willshaw 1975] formalized notions of chemospecific matching suggested by Sperry's hypothesis. They drew a clear distinction between two forms of chemical matching, which they termed "group I" and "group II" (often subsequently referred to as type I and type II). In group I (which they also refer to as *direct*) matching, each presynaptic cell j has an affinity for just a small neighbourhood of postsynaptic cells, with peak affinity for the cell topographically matching to j in the postsynaptic sheet: a mosaic model. While such a rigid scheme can form a map under normal conditions, it cannot account for certain experimental data without invoking the unsatisfactory notion of respecification (see earlier). In group II matching, "...*all* axons have maximum affinity for making and retaining contacts at one end of the postsynaptic sheet of cells, and progressively less for the cells at greater distances from that end. Similarly, all postsynaptic cells have maximum affinity for axons from one end of the postsynaptic set, and axons remote from this end have correspondingly less likelihood of retaining *any* contact. We may thus talk of *graded affinity* within both pre- and postsynaptic sets..." [Prestige & Willshaw 1975]. Simulations showed that this mechanism is capable of forming an ordered map if competition is introduced, in the form of normalization. They specified that each presynaptic cell could make only a fixed number of contacts N_{pre} among the postsynaptic cells, and similarly each postsynaptic cell could only support a fixed number of contacts N_{post} from presynaptic cells. This then ensures an even spread of connections: without competition for postsynaptic sites every presynaptic cell would connect only to the highest affinity end of the postsynaptic set, while without competition for presynaptic sites every postsynaptic cell would receive connections only from cells at the highest affinity end of the presynaptic set. However, to explain compression and expansion of retinotectal maps, they found that it was necessary to make the additional assumption ("regulation") that N_{pre} and N_{post} were also altered. For instance, if half the postsynaptic set is removed then, unless N_{post} is increased, connections will only be made from the highest-affinity end of the presynaptic set.

3.4.2 The Neural Activity model

[Willshaw & von der Malsburg 1976] proposed a theory in which topographic map formation is driven by correlated activity in the presynaptic sheet. A two-dimensional sheet of presynaptic cells is totally connected, with synapses of variable strengths, to a two-dimensional sheet of postsynaptic cells. Initial connectivity is random, with a small initial bias to model the effect of an initial crude polarity-matching mechanism: the initial weights between a few topographically matching pairs of pre- and postsynaptic cells were artificially strengthened. There also exist fixed lateral connections in the postsynaptic sheet, excitatory at short range and inhibitory at long range (similar to [von der Malsburg 1973]). At each time step, a pattern of "dipole" activity is imposed on the presynaptic sheet. This pattern is generated by choosing a cell at random and giving it and one of its neighbours unit activity, while all other cells in the presynaptic sheet have zero activity. Presynaptic cells are thus anticorrelated with all other presynaptic cells apart from their immediate neighbours. The activity of cells in the postsynaptic sheet is calculated by adding together the feedforward and lateral interactions and iterating until a stable state is reached. All postsynaptic cells with activity above a certain threshold then have their connection strengths updated according to a Hebbian rule. This is a form of competition between units, similar to [von der Malsburg 1973]. In order to prevent the weights (which are always positive) growing without bounds, the total sum of connection strengths for each postsynaptic unit is normalized to a constant value by dividing each connection strength by the total sum. They showed that such a mechanism is capable of generating topographic maps, given a small degree of initial order.

Although providing an elegant account of how activity can drive map formation under normal conditions, the Neural Activity model fails to account for translocated maps in graft transplantation experiments, and also some of the compound eye experiments (for review see [Willshaw & von der Malsburg 1979]).

3.4.3 The Tea Trade model

The one remaining idea of chemospecific matching referred to by Sperry that had not been investigated theoretically was that of map formation being somehow dependent on induction of chemicals from the retina into the tectum. [von der Malsburg & Willshaw 1977, Willshaw & von der Malsburg 1979] proposed a model based on this mechanism (for a summary of the model see [von der Malsburg & Willshaw 1981(a)]; [von der Malsburg & Willshaw 1981(b)] presents a more formal statement of the equations, and [von der Malsburg 1989] is a recent review). The model is expressed in terms of chemical markers, but markers intrinsic only to the presynaptic sheet. There are no pre-existing markers in the postsynaptic sheet: it is assumed that presynaptic markers are transported to the postsynaptic sheet via induction along retinal axons (evidence for such a mechanism in the natural system comes from the experiments of [Schmidt 1978, Schmidt et al 1978]). Several markers exist in the presynaptic sheet, the sources of which are spaced out in the presynaptic sheet at fairly regular intervals. An analogy presented to explain the working of the model was the import of tea from plantations in India to British towns [von der Malsburg & Willshaw 1977]: hence the name "Tea Trade model".

Initially, markers diffuse in the presynaptic sheet until a stable distribution is set up. It is assumed that each presynaptic axon then induces the vector of markers existing at that point in the presynaptic sheet into the postsynaptic sheet, where the markers diffuse according to the same rule as in the presynaptic sheet. The rate of induction at each synapse is proportional to the strength of that synapse. Synaptic strengths are updated periodically according to the degree of similarity between the vector of markers each fibre carries and the vector of markers already existing at those points it contacts in the postsynaptic sheet. An interesting aspect of this model is that synaptic updating occurs with a molecular analogue of the Hebb rule: the strength of connection between presynaptic cell i and postsynaptic cell j is increased in proportion to the similarity of their vectors of chemical markers, rather than on the basis of the correlations in their activities as Hebb originally formulated the rule. A slight overall bias in the connection strengths is specified initially in order to provide a global orientation for the map. Synaptic strengths are normalized by division so that each *presynaptic* cell can only support a certain total strength of connections. This ensures that every presynaptic axon makes contacts in the postsynaptic sheet. Willshaw and von der Malsburg showed that such a process reaches a stable state and can form appropriate maps under normal conditions, and also correctly predict the outcome of a range of results under abnormal conditions. In particular, some regeneration experiments (e.g. [Schmidt 1978]) can be accounted for under this scheme by assuming that when regeneration occurs, the postsynaptic sheet holds a "memory" of the previous pattern of innervation in terms of the previous stable distribution of markers.

An algorithm similar to the Tea Trade model has been analysed mathematically for the one-dimensional case in [Häussler & von der Malsburg 1983]. The technique of analysis used is synergetics: "Linear analysis near the initial state, classification of the linear modes into principal and ancillary, adiabatic elimination of the latter from the original system and discussion of the resulting equations". Their analysis makes the assumption that the two modes corresponding to an ordered map (one in each direction) are the only modes with positive eigenvalue. A linear analysis shows that eventually these modes dominate. By considering higher-order terms, they were also able to show that asymmetries in the initial configuration can select between the two modes (cf [Turing 1952]).

3.4.4 Cowan's model

The three models described so far each address one of the basic mechanisms that could underlie map formation, and each can account for a different subset of the experimental data. In order to explain a more complete range of data, some experimentalists (e.g. [Meyer 1982(a), Fraser & Perkel 1990]) have argued that it is important to consider the simultaneous operation of several mechanisms. These are hypothesised to act in concert in normal map formation, but under abnormal conditions different mechanisms can dominate to produce varying outcomes.

One such model is that of [Whitelaw & Cowan 1981]. This was an attempt to integrate both marker- and activity-based mechanisms in map formation, by combining a gradient of adhesive specificity (i.e. a group II mechanism) with synaptic updating as in the Neural Activity model [Willshaw & von der Malsburg 1976]. They employed an activity regime and updating rule similar to the Neural Activity model, but with changes in synaptic strengths multiplied by the degree of adhesion between the corresponding pre- and postsynaptic cells (it is unclear how such a mechanism could be implemented in the natural system). Both pre- and postsynaptic normalization are employed. They predict a range of the experimental literature (including expansion and compression, mismatch, rotation and compound eye experiments), and also draw attention to experimental evidence contradictory to the induction hypothesis of the Tea Trade model.

Recent additions to this model [Cowan & Friedman 1990(a), Cowan & Friedman 1991] have increased the range of data it can account for. An extra term in the updating equations has been added to represent ".... a postulated random depolarization which occurs at synapses due to the quantal release of neurotransmitter...." [Cowan & Friedman 1990(a)]. This means that a crude map will now form in the absence of electrical activity, in agreement with the experimental data. In addition, motivated by experimental evidence that fibre-fibre interactions play a part in map formation (see e.g. [Fraser & Perkel 1990]), a tendency for fibres to stick to their retinal neighbours has been added. With these and other modifications the model reproduces the results of [Schmidt et al 1978], among others.

3.4.5 Fraser's model

Another example of the consideration of multiple influences on map formation is the model of [Fraser 1980, Fraser 1985, Fraser & Perkel 1990]. Fraser introduced the notion that the state of the system could be described by an "adhesive free energy" G, which depends on how successfully a number of constraints are satisfied. In his model the form of the final mapping is found from minimising G by simulated annealing [Kirkpatrick et al 1983]. In [Fraser & Perkel 1990], the constraints employed (in order of decreasing weighting in G) are (1) a position independent adhesion between retinal and tectal cells, (2) a general competition among retinal axons for tectal space, (3) a tendency for neighbouring axon terminations in the tectum to stabilize if those axons come from neighbouring positions in the retina (implemented perhaps by correlated activity), (4) a dorsoventral gradient of adhesive specificity in retina and tectum, and (5) an anteroposterior gradient in retina and tectum. In [Fraser & Perkel 1990] an impressive array of results is presented that qualitatively accounts for a large proportion of the experimental literature. However, Fraser does not provide a developmental mechanism which could perform such a minimization. Also, the way he proposes constraints (4) and (5) to be implemented is somewhat unclear. In fact a group I type of specificity is employed, with the suggestion that this could be implemented by gradients of different molecules running in opposite directions [Fraser & Perkel 1990]. However, he does not properly explain how the addition of two linear gradients results in a local minimum of G for each retinal cell.

3.4.6 The arrow model

Shortly after [Prestige & Willshaw 1975], an alternative way of looking at map formation was proposed [Hope et al 1976], the so-called "arrow" model. The assumptions of this model are that (a) retinal fibres that terminate next to each other in the tectum are able to compare their relative positions of origin, and (b) retinal fibres can identify both rostral-caudal and medial-lateral axes of the tectum (this could be implemented in terms of a polarity gradient of marker along each of these axes). The procedure is basically the bubblesort algorithm, though they do not give a biological implementation for this. Starting from initially random connections where each retinal fibre contacts one tectal cell, two retinal fibres are chosen at random that terminate next to each other in the tectum. Their retinal positions are compared, and if appropriate the sites of termination of the fibres are exchanged. This simple algorithm is capable of forming topographic maps under normal conditions, and also rotated maps when a piece of the tectal array is rotated. However, it fails to account for the translocated map experiments. This algorithm has since been extended in a number of ways [Overton & Arbib 1982]. The tectum is now considered to be continuous, retinal fibres are assumed to branch and make several contacts in the tectum, and "....the movement of a retinal fibre is determined by the overall trend of the movements of its individual branches" [Overton & Arbib 1982]. They also explicitly introduced group II position markers for each axis in retina and tectum. This "Extended Branch Arrow model" is able to account for more of the experimental data than the original arrow model.

3.4.7 The work of Amari

[Amari 1980, Takeuchi & Amari 1979] performed a mathematical analysis of a one-dimensional version of the Neural Activity model, but assuming that the postsynaptic sheet is a continuous neural field, and that signals are chosen at random in a continuous rather than discrete presynaptic sheet. In [Amari 1980] inputs were assumed to be single points, but in [Takeuchi & Amari 1979] the case of inputs with finite spatial extent was analysed. This latter case produces more interesting results, since now there is an interaction between the two length scales involved: the width of the input stimuli L_{in}, and the width of lateral interactions in the postsynaptic sheet L_{lat}. It was shown that if $L_{in} < L_{lat}$, a smooth topographic map forms as normal, but if $L_{in} > L_{lat}$ the map forms a series of *microstructures*: although topographic on a large scale, on a fine scale the map is broken up into small blocks, which Amari identified with columns in visual cortex [Hubel & Wiesel 1977]. Later [Amari 1983, Amari 1988] further extended the analysis to show that the area in the postsynaptic sheet devoted to the representation of a particular part of the presynaptic sheet is proportional to the probability of occurrence of stimuli in that part of the presynaptic sheet. That is, more frequently occurring stimuli are represented in more detail in the postsynaptic sheet (in agreement with experimental data on somatosensory cortex - see for example [Sur et al 1980]). This analysis has recently been extended to the case of a dual population of postsynaptic cells, some inhibitory and the rest excitatory in accordance with Dale's law, and similar results are obtained [Da Silva Filho 1991].

3.4.8 Kohonen's algorithm

The models of topographic map formation so far discussed have generally been addressed towards the retino-tectal or retinocortical systems. We now mention a *general* map formation technique that sacrifices biological plausibility for algorithmic expedience, since it will be important later in the thesis. This is the algorithm of [Kohonen 1982(a)], which is basically the Neural Activity model [Willshaw & von der Malsburg 1976], abstracted away from biological details and simplified to make it much faster computationally [Kohonen 1982(a), von der Malsburg 1989]. Kohonen's insight was to see that the time-consuming process of iterating activity in the postsynaptic sheet to reach a stable state in response to input both from the presynaptic sheet and from the lateral excitatory and inhibitory connections in the postsynaptic sheet could be replaced by a much simpler process with approximately the same behaviour. In Kohonen's algorithm, it is *assumed* that the result of iteration in the postsynaptic sheet is that a small "bubble" of activity forms, centred on the postsynaptic cell with the largest activity in response to purely feedforward input, such that these cells are the only postsynaptic cells with non-zero activity: they are assumed to inhibit all other cells. Only the synapses for this small bubble of cells are updated in response to that pattern. The process of iteration is thus now simplified to the step of just finding the postsynaptic cell with the largest activity (or equivalently the largest input, if the activation function of cells is assumed to be monotonic) for the input pattern presented. This winning cell and its geometric neighbours then have their connection strengths updated. This implements a form of competition between cells, and is similar to competitive learning [Rumelhart & Zipser 1986, Carpenter & Grossberg 1988] with a topology on the output units. It has been shown analytically that this process converges to an ordered map in the case of a discrete one-dimensional presynaptic sheet mapping to a discrete one-dimensional postsynaptic sheet [Kohonen 1982(b)]. Additionally simulations show that this algorithm leads to maps with interesting geometric structure in a number of more complex cases, including mapping between spaces of different dimensionalities [Kohonen 1982(a), Kohonen 1988]. The final results seem to be reasonably insensitive to the precise form of the neighbourhood function, although different functions can lead to quite different convergence times [Erwin et al 1991]. One important aspect of Kohonen's simulations is that usually the width of the neighbourhood function is changed over the course of a simulation. Best results are obtained if this width is initially large and then gradually reduced: a process often referred to in this context as "annealing", though it is rather different from the technique of simulated annealing [Kirkpatrick et al 1983]. The most obvious correlate for such a process in the biological system is the critical period: however, there is no evidence for such changes in lateral connectivity during natural development.

[Cottrell & Fort 1986] applied Kohonen's algorithm specifically to the case of retino-cortical mappings, although still in an algorithmic rather than biological form. They considered a discrete array of retinal cells making contact with a continuous cortical sheet. A small part of the retina was stimulated at each time, and the relative positions of the centres of mass of retinal weights in the cortex was used to assess the topography of the map. Unusually, the initial bias was in the form of fixing the boundaries of the map. They showed theoretically that this variant of Kohonen's algorithm leads to a topographic map in the one dimensional case (where as stated above topography is measured in the retina-cortex direction), and presented simulation results suggesting that the algorithm also produces a topographic map in the two dimensional case.

3.5 Computational models - ocular dominance

Attempts have been made to extend a number of the models of topography discussed in the previous section to account for ocular dominance stripes with the same mechanisms. We first review these, and then some other models which have been proposed specifically for ocular dominance.

3.5.1 The Neural Activity model

A model similar to the Neural Activity model was applied to the formation of ocular dominance stripes by von der Malsburg in [von der Malsburg & Willshaw 1976]. He assumed that activity in the two presynaptic sheets is anticorrelated between the two eyes: the activity regime in the model consists of activity in one eye *or* the other, but never both eyes simultaneously. He showed that this model is capable of producing ocular dominance stripes, and he argued that the anticorrelation between the two eyes could be produced by inhibition within the LGN. However, there is no physiological evidence such inhibition. This model is somewhat similar to that explored in chapter 6 of this thesis, except that in the model of chapter 6 ocular dominance segregation can occur with distributed patterns of activity presented in both eyes simultaneously.

3.5.2 The Tea Trade model

[von der Malsburg 1979] also proposed a version of the Tea Trade model which produces ocular dominance segregation. This assumes the same chemical markers in the two sheets, plus one extra unique marker for each sheet. These markers become active once the regular markers have set up a smooth map, and segregation occurs. Stripe width is determined by the balance between two factors: a tendency for cells to map to their correct retinotopic locations as in the monocular model, and a tendency for cells from one eye to map close to one another due to the attractive force of the chemical unique to that eye. However, the existence of such a global chemical difference between the eyes appears unlikely in light of the experiments where stripes are formed between fibres from compound eyes [Fawcett & Willshaw 1982], and the isogenic eye experiments [Ide et al 1983]. In addition, there is no evidence for the proposed time course of marker activation.

3.5.3 The models of Cowan and Fraser

The most recent versions of Cowan's model also generate ocular dominance stripes [Cowan & Friedman 1990(b)]. However, segregation only occurs if the correlation between the two eyes is less than or equal to zero (J. Cowan, personal communication).

Ocular dominance stripes can be formed in Fraser's model if it is assumed that constraint (3), the tendency for neighbouring axon terminations in the tectum to stabilize if those axons come from neighbouring positions in the retina, is stronger for terminals from the same eye than for terminals from different eyes [Fraser 1985, Fraser & Perkel 1990]. This could be so if activity within an eye is generally more correlated than activity between the two eyes. While this is reasonable to a degree (though fibres from the two eyes can still be highly correlated if they are from corresponding positions), since this model does not propose a learning rule that can identify such differences in correlation, it is simply equivalent to the assumption of a global label marking one eye as different from the other. Also, there is no consideration of the fact that in the natural system cells are initially binocularly innervated. [Fraser & Perkel 1990] state that stripe width is determined purely by the width of terminal arbors (a fixed parameter of the model), rather than on the balance of the constraints.

3.5.4 Kohonen's algorithm

An application of Kohonen's algorithm to the problem of the simultaneous formation of topography, ocular dominance, and also orientation selectivity, has recently been presented in [Obermayer et al 1990, Ritter et al 1990, Obermayer, Blasdel & Schulten 1991, Obermayer, Ritter & Schulten 1991]. This can be looked at in two ways: either as a biological model based on the presentation of a sequence of images to the retina and the subsequent evolution of synaptic strengths between the retinae and the cortex, or as an abstract optimization model where the aim is to find a "good" mapping between an appropriate feature space and a cortical sheet, without placing emphasis on the biological interpretation of the mechanism used to find the mapping. The first point of view (stressed in [Obermayer et al 1990]) suggests analogies with the model presented in chapter 6 of this thesis, while the second point of view (pervading more recent articles) suggests analogies with the elastic models of chapter 5. We will discuss these analogies in the appropriate chapters: here we describe the basic framework.

The model consists of an array of cortical units receiving input from an array of retinal units onto which patterns of activity are imposed. These patterns are localised, oriented patches, where only one patch in the retina is active at a time. In the binocular case, activity patterns in the two eyes are generally of different strengths but otherwise identical. The cortical units compete for each pattern and are updated according to a neighbourhood function, as in Kohonen's algorithm. The degree to which the stimulus is oriented and ocularly dominant is determined by two parameters, q and z. $q = 0$ corresponds to round stimuli, and as q increases the pattern becomes more elongated (i.e. more sharply tuned to a particular orientation). $z = 0$ corresponds to the pattern being of equal strengths in the two eyes: $z > 0$ represents the pattern stronger in one eye, and $z < 0$ represents the pattern stronger in the other eye. In the simulations presented in [Obermayer, Ritter & Schulten 1991] the inputs were chosen at random positions, but with fixed orientation tuning $q = q_{pat}$ and fixed dominance $|z| = z_{pat}$.

One of the most interesting aspects of this model is that Obermayer et al have been able to calculate critical values for their parameters, which determine when orientation and ocular dominance stripes will be formed. This is related to the issue of *automatic selection of feature dimensions*, which was introduced by Kohonen (see e.g.

[Kohonen 1988]) and has been analysed for Kohonen's algorithm in [Ritter & Schulten 1988]. The basic question is to predict, when mapping from a high-dimensional feature space to (for instance) a two-dimensional sheet, which variables from the feature space will be "primary" in the sense of being mapped smoothly, and which variables will be secondary: that is, changing rapidly on a local scale. For instance, in the case of mapping two eyes to the cortex, if space is the primary variable then a pattern of ocular dominance stripes will be formed. However, if space is secondary to ocularity, projections from the two eyes completely segregate such that each eye is mapped over one half of the cortex. In the Obermayer et al model, there are 3 possibilities for ocular dominance which depend on their parameter z_{pat}. If z_{pat} is small, cells do not become dominant (i.e. ocularity is not represented), and if z_{pat} is large the projections become completely segregated into two regions. It is only between these two critical values that stripes form [Obermayer, Ritter & Schulten 1991]. Similar results hold in the case where the dominance of the patterns is chosen uniformly in the range $z \in [-z_{pat}, z_{pat}]$ [Obermayer, Blasdel & Schulten 1991][1]. They also show that stripe width in their model is determined by the width of the cortical neighbourhood function [Obermayer, Blasdel & Schulten 1991].

3.5.5 The elastic net algorithm

The elastic net is a developmental algorithm that has the virtue that there exists an energy function which the algorithm optimizes. It was originally formulated as a simplification of the Tea Trade model specifically for application to combinatorial optimization algorithms which have a geometric interpretation: in particular the Travelling Salesman problem [Durbin & Willshaw 1987]. In [Goodhill 1988, Goodhill & Willshaw 1990] it was applied to the topography and ocular dominance problem and shown to be capable of producing topography and ocular dominance from the same mechanisms. The retinae are represented as two parallel sheets of units lying in an abstract "feature" space. The directions parallel to the two planes represent the spatial dimensions: the third dimension is ocularity. The cortex is represented by a sheet of points in this feature space. Each cortical point is attracted to each retinal unit, and also to its neighbouring cortical points: this latter force attempts to minimize the distances between neighbouring cortical points. The trade-off between these two components serves to establish a map that matches cortical points to retinal units so that the distances between neighbouring cortical points as measured in the feature space are kept small. We discuss the elastic net in much greater detail in chapter 5.

We now turn to models designed exclusively for ocularity.

3.5.6 Swindale's model

One of the first mathematical theories specifically addressing the formation of ocular dominance stripes was presented by Swindale [Swindale 1979, Swindale 1980]. His model is expressed in terms of stimulating and inhibiting interactions between neighbouring synapses in the cortex, for which he suggests interpretations both in terms of correlated activity and diffusing chemical substances. In the second case, the model bears similarities with that of [Turing 1952], and could also account for pattern formation in other systems such as zebra or mackerel skin.

The model consists of a cortical array filled with synapses from both eyes that are initially of random strength and uniformly intermingled. A fixed topography is assumed, and only the distribution of synapses rather than individual cells is considered. There are four interaction functions between synapses, relating to the four possible left-eye, right-eye combinations, although in practice Swindale usually reduced this to just "same-eye" and "opposite-eye". Both of these interaction functions are radially symmetric with two regions: a narrow inner one and a broader outer one. For synapses from the same eye, the inner region is excitatory and the outer region inhibitory, and vice-versa for synapses from opposite eyes. It is assumed therefore that it is possible to distinguish fibres from the two eyes (for instance, on the basis of correlated activity). At each time step, the effects on each synapse are summed and the strength of that synapse increased or decreased according to whether the sum is positive or negative. Synaptic strengths are bounded between upper and lower limits. Various forms for the interaction function were investigated, and (providing they were of centre-surround type) the results were found to be reasonably insensitive to this.

[1] It has been pointed out by Peter Dayan (personal communication) that there is an inconsistency in the way this model is formulated. Under the abstract optimization interpretation, all feature dimensions (e.g. space and ocular dominance) should be regarded equally. However, when patterns are presented, corresponding positions within the two eyes are allowed to be active at the same time, but not widely separated positions in the two eyes.

This mechanism produces results that share a number of the morphological features of ocular dominance stripes as described in chapter 2: narrowing, branching, and perpendicular intersection with the boundaries. Although the stripes twist and turn more than in the natural system, a more regular global orientation can be produced by assuming growth of the cortex during development [Swindale 1980]. Narrowing of stripes in response to monocular deprivation can be produced by the model if it is assumed that deprivation corresponds to a reduction in the strength of the interaction functions for one eye. Qualitative arguments are presented based on the form of the interaction functions which suggest that the stable patterns should indeed be stripes. A similar model has also been formulated to account for the pattern of orientation columns seen in V1 [Swindale 1982], and more recently for the interaction between orientation columns and ocular dominance columns [Swindale 1992].

3.5.7 The model of Bienenstock

An influential model of cortical receptive field development is that of [Bienenstock et al 1982, Bienenstock 1980] (for a more recent discussion, including a possible physiological implementation of this algorithm, see [Bear & Cooper 1990]). Although primarily a theory for the development of orientation selectivity of isolated cortical cells, it is relevant here since the development of binocularly innervated cells is also considered. Weights to a cortical cell evolve according to a Hebbian rule, driven by input patterns randomly chosen from an "environment" of patterns. The unusual feature of the particular rule they use is the way that decreases in the weights are specified, which is a different method from that of normalization. Firstly, the learning rule says that in the absence of input, weights decay exponentially to zero. More importantly, they introduce a threshold for the output of the cell such that below this threshold all weights decrease. This threshold is chosen to be a nonlinear function of the response of the cell. Bienenstock at al showed that for certain environments of oriented patterns, weights converge to a stable state that is orientationally selective. With respect to ocular dominance, they showed the following results. (1) If inputs to the two eyes are perfectly correlated (which they referred to as "normal rearing"), cells are binocular. (2) Monocular cells develop under conditions of "binocular deprivation" (constant vector presented to both eyes, but with random variation in relative strength between the eyes), "monocular deprivation", and "uncorrelated rearing" (inputs in the two eyes completely uncorrelated). Again we have the problem of explaining why cells become monocular even when the two sets of inputs are positively correlated.

3.5.8 Miller's model

[Miller et al 1989] proposed a model of ocular dominance stripe formation based on a "correlational" learning rule similar to that of [Linsker 1986], though the two models were developed independently (see also [Miller 1989, Miller 1990(a), Miller 1990(b), Miller & Stryker 1990]). Given certain assumptions about the input-output function of postsynaptic units and a simple mathematical implementation of the Hebb rule, Miller derived an expression for the rate of change of weights that depends only on the correlations between presynaptic units rather than between pre- and postsynaptic units. In particular, the behaviour of the system is determined largely by the eigenvalues and eigenvectors of the matrix of time-averaged correlations between presynaptic units. Miller assumed that the strength of these correlations drop off as a gaussian function of distance in each presynaptic sheet (representing the LGN), but that there are zero correlations between units in the two presynaptic sheets. Fixed lateral connections exist in the postsynaptic sheet (representing the cortex), generally of a short-range excitatory and long-range inhibitory form. At each time step, activity is iterated to a stable state in the postsynaptic sheet and then synaptic strengths are updated according to this pattern of activity and the presynaptic correlation matrix. Stripe width is determined largely by the width of the cortical interaction function. We discuss this model in much greater detail in chapter 4 of this thesis.

3.6 Conclusions

We have briefly reviewed the most important models of the development of topography, and ocular dominance, in the retinotectal and retinocortical systems. The properties of each model are summarized in table 3.1. It is clear that there are a number of different types of model, which depend on different types of mechanisms. The Tea Trade model, Cowan's model and Fraser's model are essentially in competition, since they tackle roughly

Model	Principal mechanisms	Topography		Ocular dominance		
		Refinement	Map	Monocularity	Stripes	Principal determiner of stripe width
Prestige and Willshaw	Group 1 and group 2 chemospecificity	-	Yes	-	-	-
Neural Activity	Activity	Yes	Yes	Yes if activity nonsimultaneous in two eyes		Width of cortical interaction function
Tea Trade	Induction of retinal markers	Yes	Yes	Yes if extra marker is introduced for one eye		Balance of chemical forces
Cowan	Group 2 chemospecificity and activity	Yes	Yes	Yes if no correlations between eyes		?
Fraser	Multiple	-	Yes	-	Yes	Width of terminal arbor
Amari	Activity	Yes	Yes	-	-	-
Swindale	Activity	Fixed		-	Yes	Width of cortical interaction function
Bienenstock	Activity	-	-	Yes	-	-
Miller	Activity	Fixed		See chapter 4		Width of cortical interaction function
Elastic net	See Chapter 5	(Yes)	Yes	(Yes)	Yes	Ratio of correlations within and between eyes
Obermayer et al	Activity	Yes	Yes	Yes under certain conditions		Width of cortical neighbourhood function

Table 3.1: Summary of models. We indicate for each of the models discussed (1) the main mechanisms underlying the model, (2) whether the model accounts for the formation of topography, and also provides a mechanism whereby an initially crude topographic map can refine, (3) whether the model explains how individual cells can become monocular, and how these monocular cells then come to form a pattern of stripes, (4) The main determiner of stripe width in the model. Dashes indicate that this is outside the scope of the model as originally presented. The entries for the elastic net should be read in the light of the discussion in chapter 5 of this thesis. There is a question mark for stripe width in Cowan's model since this has not been discussed by Cowan.

the same set of experimental data. The Tea Trade model is the most elegant, Cowan's is the most detailed, and Fraser's, being based on the direct optimization of an objective function, has the clearest computational level description. By contrast, [Prestige & Willshaw 1975] and the Neural Activity model are complementary: instead of attempting to account for a wide range of data, they each take a single mechanism and explore its capabilities. The same can be said of Bienenstock's and Swindale's models, except that the former is concerned only with the ocular dominance of single cells, and the latter only with the laying out of two projections in a pattern of stripes. Miller's is a more detailed model, that nonetheless has the virtue of yielding to extensive mathematical analysis. The models deriving from Kohonen are more algorithmic in nature, and are thus not directly competing with the more directly biological models such as the Neural Activity model and the Tea Trade model.

Strong analogies can be made between models based on chemical labelling and those based on activity labelling. Both specify that axons of presynaptic units are "similar" (in the sense of expressing similar chemical identities or having correlated activity) if they originate from nearby cells in the presynaptic sheet. The Tea Trade model is related to activity-based models in that they each specify some mechanism for information to be transmitted between postsynaptic units during map formation about which presynaptic axons connect to them. This is by diffusion of presynaptic markers in the postsynaptic sheet in the Tea Trade model, and lateral connections between neighbouring postsynaptic units in for instance the Neural Activity or Miller's model. Such models require at least some initial polarity to generate global order. Models such as Prestige and Willshaw, Cowan, and Fraser are somewhat different in that they assume fixed chemical gradients in the postsynaptic sheet, and communication between postsynaptic units of the connections they receive is not required. Such fixed gradients can produce global order.

To what extent do the models we have described fit into the criteria listed in the introduction? (We do not include Miller's and the elastic net model here, since they will be discussed in much greater detail later.) The first point is the development of topography and ocular dominance from the same mechanisms. This is difficult to judge for models such as Swindale's and Bienenstock's, since these were formulated exclusively for ocularity. It has not been shown that similar mechanisms could *not* also produce a topographic map: the issue has not been addressed. Regarding biological relevance, for extensions of models of topography we have indicated certain assumptions with relation to ocular dominance in each of the models which appear to be biologically implausible. Concerning the issue of correlations, the binocular versions of the Neural Activity model and Cowan's model require non-positive correlations between the eyes. The stripe morphology produced by the models is again hard to judge, since there are no generally accepted ways of assessing this.

It is apparent from our discussion that up to now, ocular dominance has often been regarded either as a separate phenomenon from the development of topography, or as an adjunct to theories of map formation. This has led to the development of models for ocular dominance that are strongly wedded to a pre-existing topography without showing how this could arise from the same mechanisms, and models for topography which consider ocular dominance only as an afterthought. Although a great amount of progress has been made by making clear distinctions between various phenomena to be explained, and between the mechanisms employed to explain them, we prefer a different perspective motivated by the experimental evidence reviewed in chapter 2. We propose that topography and ocular dominance should be regarded as equally important aspects of the functioning of the same basic mechanisms. This leads us to develop new models where from the outset topography and ocular dominance are considered equal partners, with a model considered inadequate unless it can naturally encompass both phenomena.