Modelling the development of the retinogeniculate pathway

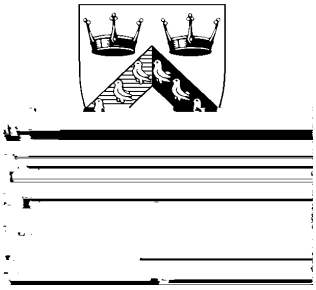
Stephen Eglen

CSRP 467

July, 1997

ISSN 1350–3162

UNIVERSITY OF



Cognitive Science
Research Papers

Modelling the development of the retinogeniculate pathway

Stephen Eglen

Summary

How does the visual system develop before the onset of visually-driven activity? By the time

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The problem of visual system development

How do mammalian visual pathways develop in the absence of patterned vision? By the time cats are born, the retinogeniculate pathway, the pathway between the retina and the lateral geniculate nucleus (LGN), has already developed into a near adult form. It is possible that each retinal cell is told exactly which LGN cells it should connect to by some genetic plan (for example, the chemospecificity hypothesis (Sperry, 1963)). However, given the large number of retinal and geniculate cells, it is more likely that some other more general guiding principles are at work (von der Malsburg & Singer, 1988). One such principle is that neural activity generated by visual stimulation drives development (Blakemore & Cooper, 1970; Hirsch & Spinelli, 1970). At this stage however, the photoreceptors in the retina are still de

layers:

1. Retinotopic mappings — neighbouring retinal cells tend to connect to neighbouring geniculate cells. Initially this mapping is quite coarse, and refines during development.

2. Ocular segregation — LGN cells initially receive inputs from both eyes before becoming selective to only one eye. Cells responding to the same eye are grouped into eye-specific layers within the LGN.

1.2 Why consider the retinogeniculate pathway?

There are many visual pathways in the mammalian brain. There are several reasons why this thesis focuses on just the retinogeniculate pathway:

- Although the retinogeniculate pathway is one of the simplest visual pathways, it is highly likely that there are general principles of development at work. By studying these principles in a simple pathway, it is hoped that they might be applicable to other more complicated pathways. For example, it is suggested that the retinal waves of activity generate action potentials in LGN cells which could in turn drive geniculocortical development (Penn, Gallego, Mooney, & Shatz, 1995).

-

1.3.1 Testing hypotheses

1.4 Aims of the thesis

This thesis aims to cover the following issues:

- How can topography and ocular segregation develop in the LGN? Can it develop by way of just the one mechanism, in the same way as was shown for ocular dominance and topography in the cortex (Goodhill, 1992)?
- How are retinotopic maps affected by the dimensionality of the output structure? Most models only consider a two-dimensional output structure, whereas this thesis is concerned with the three-dimensional structure of the LGN.
- What are the within-eye and between-eye correlations that the model requires for development and are they similar to the correlations that exist in the developing retina?
- Can the segregation of on- and off-centre cells be explained by the same mechanisms as those needed for topography and ocular segregation, assuming certain correlations between on- and off-centre cells?

1.5 Outline of the thesis

Chapter 2 provides a review of the relevant biological literature for this thesis. It contains a brief introduction to the early stages of the mammalian visual pathway, concentrating on the properties of cells in the retina and the LGN. The factors influencing neural development of visual pathways, especially the retinogeniculate pathway, are discussed. The chapter concludes by introducing the central hypothesis of the thesis: spontaneous retinal waves drive development of the retinogeniculate pathway.

Chapter 3 reviews the previous computer models of retinotopic map formation and ocular dominance. Although there are many models which superficially seem quite different, they share many common mechanisms. Despite the large number of previous models, only one model has explicitly examined the role of spontaneous activity in retinogeniculate development (Keesing et al., 1992). A crucial feature missing from all models however is the lack of information on how retinotopic maps can develop in the

specific waves can cause postsynaptic units to become responsive to either on- or off-centre presynaptic units, but the input correlations required by the model are at odds with the current biological data.

Chapter 6 investigates the forms of ocular dominance patterns that can be produced by cortical models. Most cortical models generate ocular dominance stripes similar to those found in area 17 of the cortex, but can these models also be applied to the pa

Chapter 2

Development of the retinogeniculate pathway

This chapter summarises the relevant biological data about the development of the retinogeniculate pathway. It is composed of two main sections. The first section introduces the retinogeniculate pathway, describing its main components and how they function. Since cross-species comparisons are difficult, most of the data is taken either from the cat or f

ing. For example, around 10% of cells in area IT of the monkey respond maximally to images

receive antagonistic input from horizontal cells, which normally receive synapses from photoreceptors over a wider area.

This combination of direct input from a narrow range of photoreceptors and antagonistic indirect input from a wider range of photoreceptors produces a centre-surround type receptive field profile for the bipolar cells. A bipolar cell with inhibitory input from central photoreceptors is called an on-centre cell since it depolarises in response to stimulation of the central photoreceptors and hyperpolarises in response to stimulation of more p

for causing the temporal retinal axons to stay on the same side of the brain, whereas the nasal retinal axons cross over to the other side (Wingate & Thompson, 1995).) The axons leave the optic chiasm to form an optic tract on each side of the brain to innervate the appropriate LGN. In this way, each LGN receives inputs from areas of the two retinae which correspond to the same part of the visual field.

For most animals, the size of the inputs to the LGN from the two retinae are different: the contralateral retinal input (that coming from the retina on the opposite side of the brain to the LGN) is larger than the ipsilateral (same side) input. The di

Y-like, depending on whether they show a null-response to sinusoidal gratings, as an analogue of the retinal X and Y classes (Bowling & Wieniawa-Nakiewicz, 1986).

2.3.2 Receptive fields of LGN relay cells

The spatial receptive field profiles of relay cells are similar to RGCs, although the inhibitory surround region tends to be stronger (Hubel & Wiesel, 1961), probably due to the effect of the inhibitory geniculate interneurons. This similarity of receptive fields is due to the low convergence of RGCs onto relay cells: typically a geniculate cell receives its inputs from between 1–5 retinal cells which tend to be from the same eye and of the same polarity (on- or off-centre) (Mastronarde, 1987a).

This similarity of retinal and geniculate spatial receptive fields has led to the notion of the geniculate as a "relay station", transferring retinal information without significant processing to the cortex. For example, geniculate cells show a slight orientation selectivity, but this property originates in the retina, and is formed by the oriented dendritic fields of retinal cells rather than being generated in the retinogeniculate pathway (Levick & Thibos, 1982; Leventhal & Schall, 1983). Similarly, the direction-selectivity found in some geniculate cells (Thompson, Zhou, & Leventhal, 1994) is most likely a reflection of the sensitivity of their retinal inputs (Shou, Leventhal, Thompson, & Zhou, 1995).

In the temporal domain however, geniculate cell responses differ quite strongly from retinal cells. Around two-thirds of geniculate cells exhibit a lagged response to retinal stimuli (Mastronarde, 1987a, 1987b). This temporal distinction has led to geniculate cells being further categorised as either lagged or non-lagged (Humphrey & Weller, 1988). Although the function of this lagged response is not clear, it could be used to generate various cortical responses such as direction selectivity (Saul & Humphrey, 1990).

Finally, X and Y relay cells can fire in one of two modes: tonic or bursting (Guido, Lu,

Lin, 1979). However, the projection of retinal space into the geniculate is not just a conventional topographic mapping. Most maps project an input space into an output space of either the same or lower dimensionality. In the LGN however, the dimensionality of the spaces is reversed: although the retina is regarded as a two-dimensional sheet of cells, the LGN is a three-dimensional block of cells. The mapping is organised such that a point of visual space in the retina projects to a column of cells within the LGN. This group of LGN cells is termed a 'projection column', which is normally defined as a column of LGN cells which contains 90% of all cells with receptive fields (RFs) responding to the same region of visual space (Sanderson, 1971b). Typically, these columns are oriented perpendicular to the LGN layers.

Although cells within a projection column receive input from the same region of visual space, they can respond in different ways to the same stimulus. For example, X relay cells at different depths within a column respond to the same stimuli with different timing latencies. These latencies could be useful for creating certain selectivities of cortical cells, such as direction selectivity (Bowling, 1989a).

The mapping of retinal space into the cat LGN is shown in Figure 2.3. The layout of retinal space in the ferret follows similar retinotopic principles, although the ferret LGN is positioned differently (Zahs & Stryker, 1985). Just as the retina has more cells representing the central visual field, so does the LGN, as can be seen from the amount of geniculate covering the central $\pm 2^{\circ}$ in comparison with the amount covering 10–20°. However, equal numbers of RGCs project to equal volumes in the LGN, and so the gradient of visual field representation in the geniculate is a reflection of RGC density gradient (Sanderson, 1971b).

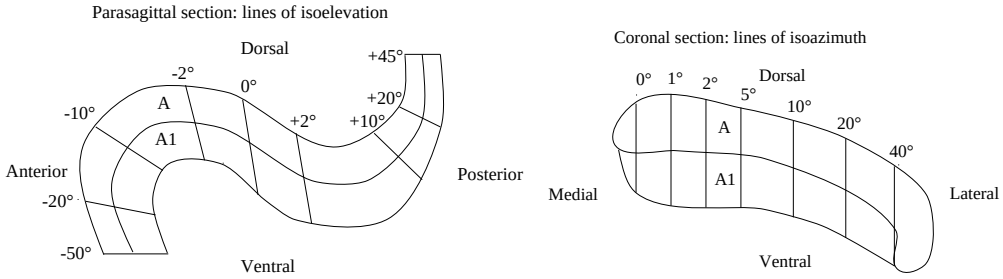
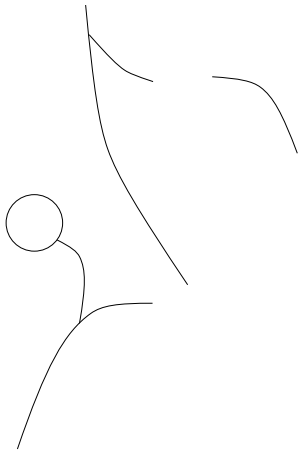


Figure 2.3: The topography and layering of the cat LGN — parasaggital and coronal views. Parasagittal view adapted from Figure 9 of (Sanderson, 1971a). Coronal view adapted from Figure 14 of (Sanderson, 1971a). The C layers have been omitted here for clarity, but lie just below



2.4.2 Other theories of LGN function

The main theories regarding the role of the LGN in visual processing are dominated by the existence of the corticogeniculate pathway and have been outlined above. In this section we briefly consider some of the other theories.

One role of the LGN may be to organise the retinal inputs in a way that can help cortical processing by grouping together certain cells so that they are close together. For example, the segregation of inputs in the LGN is maintained in the projection to the cortex. In the cat, central regions of the A laminae project to layer IVb of cortical area 17, whereas border regions project to layer IVa (Bowling, 1989b).

Alternatively, since there are around four times as many geniculate relay cells than RGCs (Sanderson, 1971b), there is considerable divergence in the retinogeniculate pathway. This divergence could give the cortex a fairer representation of the X and Y cell pathways. In the retina, there are around five times as many X RGCs as Y RGCs. In the geniculate, this ratio is reduced to around 2:1, since Y-RGCs innervate many more LGN relay cells than X RGCs (Sur et al., 1987; Sherman & Koch, 1990). This increase in the proportion of Y-like cells in the geniculate reduces the dominance of the X pathway that exists in the retina.

Another role of the LGN may be captured in the way that geniculate cells respond over time. Lagged and non-lagged geniculate cells can transform retinal responses in the temporal domain. It has been hypothesised that just as the retina removes spatial correlations using centre-surround receptive fields (Atick & Redlich, 1992), the LGN removes temporal correlations using the temporal response characteristics of lagged and non-lagged cells (Dan, Atick, & Reid, 1996). Finally, it has also been suggested that geniculate cells adjust the variance of retinal signals on their way to the cortex, transforming the non-linear variance of retinal cell firing rates to mean firing rates into linear variances. (Levine, Cleland, Mukherjee, & Kaplan, 1996).

2.5 Factors influencing neural development

The first half of this chapter has focused on the properties of the visual pathway from the retina, through the LGN, up to the cortex, and back down to the LGN. The second half of this chapter considers the central question of how this pathway comes into existence during the early stages of the animal's development. In general, we are concerned with how one set of cells, the presynaptic cells, connects to another group of cells, the postsynaptic

after the optic nerve had been severed. Their results showed

2.5.2 The role of experience in development

The role that visual experience plays in the development of visual pathways was first assessed by a series of experiments looking at the development of ocular dominance in cat visual cortex (Wiesel & Hubel, 1963a, 1963b; Hubel & Wiesel, 1963). In a newborn kitten, most cortical cells in the primary visual cortex (area 17) respond to stimulation of either eye, although there is a bias favouring stimulation of one of the eyes. After around six weeks of normal visual experience, most cells have adapted to respond to stimulation of only one eye, ignoring stimulation from the other eye. Overall, both eyes normally innervate an equal number of cortical cells. However, if one of the eyes is covered from birth for six weeks so that it does not receive any visual input, most (if not all) of the cortical cells do not respond to stimulation of the eye that was deprived of vision. Wiesel and Hubel (1963a) repeated these experiments with kittens of different ages. They found that the older the kitten was at the onset of deprivation, the weaker the effects of deprivation. Furthermore, depriving one eye of vision during adulthood did not affect the pattern of ocular dominance at all. These results indicated a critical period of development, during which visual deprivation radically

found. Hence the ability to generate barrels is not due to some intrinsic property of somatosensory

These experiments confirmed the existence of spontaneous activity, correlated across neighbouring retinal cells, in the developing retina.

Since then, two developments in recording techniques have revealed more details about the nature of this spontaneous activity. First, the development of the multielectrode array has allowed the activity of around one hundred neighbouring retinal ganglion cells to be recorded simultaneously (Meister et al., 1991). These recordings show that the cells

Smetters, Hahm, and Sur (1994) rightly concluded that there could be other activity-dependent mechanisms at work, or that the NMDA receptors are used to detect correlated activity for other purposes, such as topographic map refinement (Cline & Constantine-Paton, 1989).

using activity-dependent processes which eliminate inappropriate synapses and strengthens topographically correct synapses. It is likely that this combination of early activity-independent processes followed by later activity-dependent processes apply to other pathways, both visual and non-visual (Goodman & Shatz, 1993).

2.9 Summary

This chapter has introduced the main elements of the early visual pathway in mammals from the retina to the geniculate, and from the visual cortex back down to the geniculate. The central hypothesis of this thesis has been introduced, stating that the LGN develops to a near-adult state using a combination of activity-independent and activity-dependent processes:

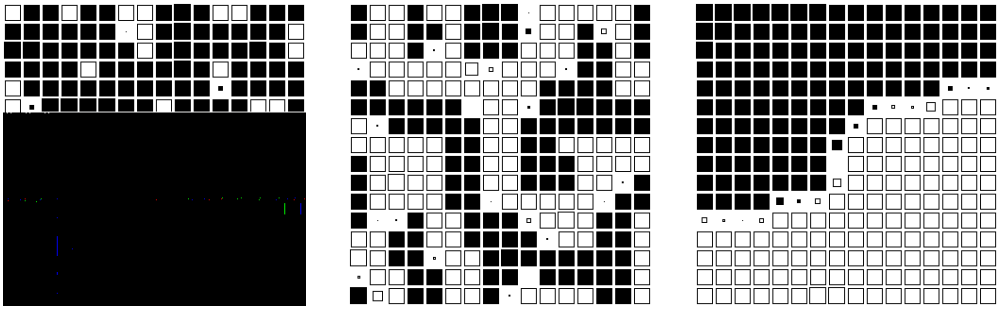
- Retinal axons from both eyes meet at the optic chiasm and project into the appropriate LGN using a combination of activity-independent markers (Hankin & Lund, 1991; Goodman & Shatz, 1993). The initial retinotopic map is quite coarse and most geniculate cells are bi534(n)-4.10691(d)(n)-4.10914(10914(,)-211026(p)-.10914(m)-230.235(b)-11(m)-230.235(b)-116.9)-4.11

Chapter 3

- patchy maps — Figure 3.1(b). These maps show local smooth organisation, but no global

- stripes — Figure 3.2(b). In comparison to patches, when the inputs from both eyes are of the same size, stripes tend to arise. The best-known example of stripes is the pattern of ocular dominance in area 17 of visual cortex. (Hubel & Wiesel, 1972; LeVay, Hubel, & Wiesel, 1975).

- layers — Figure 3.2(c). The output map can organise so that all of the inputs from one eye are gathered together into one large region. This is closest to the form found in the LGN (Sanderson, 1971a).



(a) Patches (b) Stripes (c) Layers

Figure 3.2: Common patterns of ocular dominance. Each postsynaptic unit is coloured black if it

Term	Meaning
	Index for presynaptic unit
	Index for postsynaptic unit
	Feedforward weight connecting presynaptic unit to postsynaptic unit
	Lateral weight connecting postsynaptic unit to postsynaptic unit
pre	Number of presynaptic units
post	Number of postsynaptic units
	Activity of presynaptic unit
	Activity of postsynaptic unit
σ	Size of neighbourhood
ε	Rate of weight-update

Table 3.1: Terms used for describing model elements.

Trdt-4.11137(nd)-4.11.48 -4.151879((a)]TJ 259.)-4.110914(d)-4..521
presynaptic axon and the postsynaptic dendrite. Additionally, some 14(s)-5.52048(e.9307]TJ/R11-6.93.103(s)-5

- Correlated output.

  In a topographic map, neighbouring postsynaptic units have

salient features from retinal activity. However, this means that a network using low-dimensional inputs cannot discover new features — for example, if the feature vector codes only for centre

3.6 Normalisation methods and resource limitations

Neural systems work within several constraints, including restrictions on firing rates and synaptic strengths which must be both positive and bounded at some upper limit. These constraints are often introduced into models for two reasons. First, many Hebbian-based modification rules are unstable, adapting weights without bounds. Normalisation constraints ensure that weights remain within limits. Second, the normalisation constraints introduce competition between weights: as one connection strength increases, others must decrease to keep the normalisation sum constant. It can also be argued that constraints are introduced simply because they are present in the natural system: synaptic strengths cannot increase without bounds, and cells have upper limits on firing rates. However, it is often better to keep the model simple and introduce extra mechanisms only

contrast, subtractive normalisation changes both the direction and magnitude of the weight vector. This difference is demonstrated for a two-dimensional vector in Figure 3.4. Divisive normalisation maintains the direction of the weight vector, whereas subtractive normalisation will tend to push elements of the weight vectors to extreme values. When these normalisation schemes are used in conjunction with modification rules, they can drastically affect the development of a postsynaptic unit's properties such as ocular dominance (Miller & Mackay, 1994; Goodhill & Barrow, 1994).

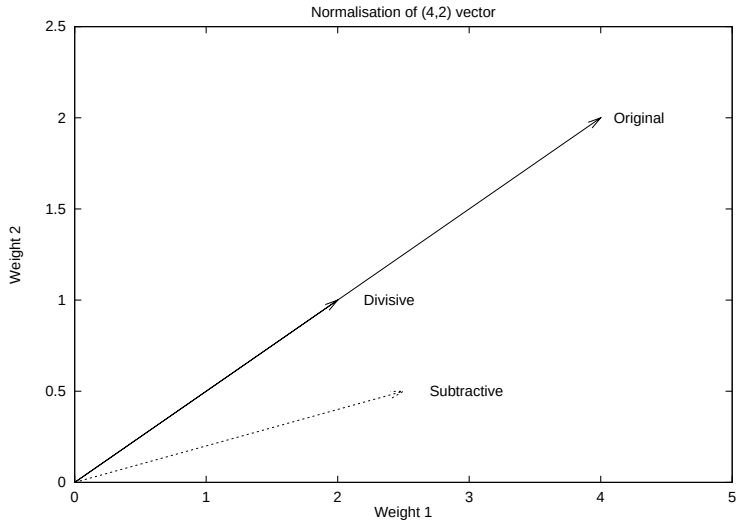


Figure 3.4: A simple comparison of divisive and subtractive normalisation. The original weight vector $(4\ 2)$ is to be normalised such that $_1 + \ _2 = 3$. Under divisive normalisation, the vector becomes $(2\ 1)$, which lies in the same direction as the original vector. Under subtractive normalisation however, the vector becomes $(2.5\ 0.5)$, which moves the vector closer to the $_1$ axis.

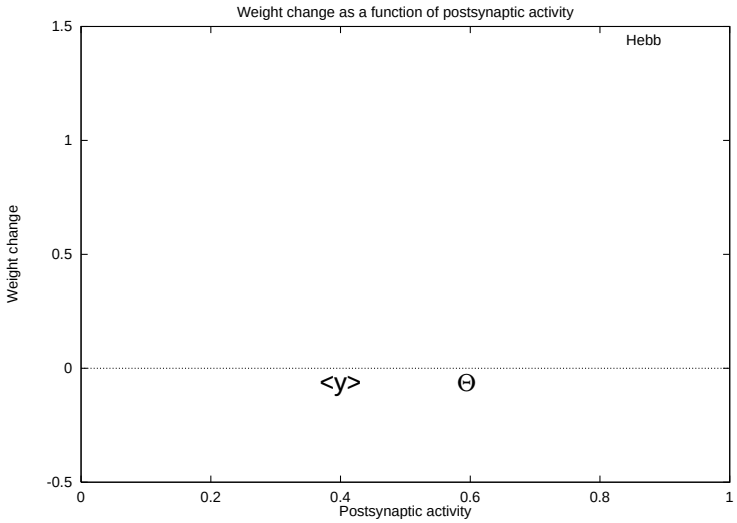
contrast, other experiments have shown that there are limits on the number of contacts each postsynaptic cell can make. First, the addition of an extra eye to innervate the optic tecta of tadpoles did not affect the number and size of connections to tectal cells when compared with normal two-eyed tadpoles (Norden & Constantine-Paton, 1994). Second, the size of retinal axonal arbors varied in accordance with changes in the number of retinal or tectal cells (Xiong, Pallas, Lim, & Finlay, 1994). For details of other related experiments, see (Hayes & Meyer, 1988b; Sabel & Schneider, 1988; Pallas & Finlay, 1991). These experiments show that there are limits on the number of contacts that a cell (either presynaptic or postsynaptic) can make, although whether constraints simultaneously exist at both sites within one system is still unknown.

3.7 Weight modification rules

Modification rules specify how the weights should be updated in response to the state of elements of the network. Such elements include the activation level of units and the value of other weights. To keep the rules fairly plausible with respect to biological principles, the rules should only make reference to quantities that are simple to calculate and available locally to the weights being changed. Many such rules have been created for modelling neural development. In this section, the two main types of modification rule used for modelling the development of topography and ocular dominance are described.

3.7.1 Correlational rules

Most modification rules capture the basic principle of the Hebbian synapse (Hebb, 1949). This states that if a presynaptic cell and postsynaptic cell are b


Weight change as a function of postsynaptic activity
Hebb
Weight change
Postsynaptic activity
1.5
1
0.5
0
-0.5
0
0.2
0.4
0.6
0.8
1
<y>
Θ

postsynaptic activity was beneath the low threshold, neither LTD nor LTP occurred. This threshold for LTD has been incorporated into a BCM-like rule to create a rule with two thresholds, known

units (or both) are therefore required to specify weight decrements which prevent weights increasing without bounds. Both the Kohonen and Obermayer/Goodhill rules are “hard-competitive” rules because at each iteration, only the weights of a subset of postsynaptic units are updated.

The elastic net (Durbin & Willshaw, 1987) combines both a regularisation and a matching term into one rule which is similar to the Kohonen rule. In the elastic net, each postsynaptic unit i has a receptive field of Gaussian shape with width σ centred on the input space at $\mathbf{w}_i$. For each input vector $\mathbf{x}$

other relevant models which investigate the development of

visual pathway (such as tectal grafts and mismatch experiments). Some of the relevant biological manipulations will be mentioned, although for more details refer back to the original papers.

Table 3.3 describes the components of the major models for the development of topographic mappings. As can be seen from this table, the various model components can be combined in different ways to investigate retinotopic map development. All models produce a refinement of topography at the single-unit level. For those models not using chemical markers to code for proximity, extra mechanisms are needed to choose between alternative global layouts of the maps,

development, there are no markers in the postsynaptic sheet, and so the global orientation of the map is dominated by the initial connection strengths. In regeneration experiments however, postsynaptic units already have markers as a result of the initial development before regeneration. These regeneration experiments produced novel map formations if the markers preferred a different global map orientation to the orientation specified in the regenerated weights.

The tea trade model was abstracted into a set of equations suitable for theoretical analysis by Häussler and von der Malsburg (1983). An eigenvector analysis showed that by suitable selection of model parameters, the eigenvalues of all non-diagonal eigenvectors were negative, leaving just the two diagonal eigenvectors with positive eigenvalues. The diagonal eigenvectors therefore dominate development to produce an ordered retinotopic mapping.

Whitelaw and Cowan (1981) took a different approach by using both neural activity and chemical markers to encode proximity of presynaptic units. In the model, group II labels were interpreted as adhesion coefficients to represent a tendency for presynaptic units to bind to certain postsynaptic units. Weight update was Hebbian based modulated by the adhesion coefficients. This model was later updated to account for more biological data (Cowan & Friedman, 1991). First, it introduced a tendency for neighbouring presynaptic units to stick together to account for the polarity mismatch experiments (Meyer, 1979). Second, it introduced random depolarisation of synapses so that in the absence of any retinal activity, a rough retinotopic map would still form.

was not surprising given the results from earlier models using anticorrelated between-eye inputs (Miller et al., 1989).

Elliot, Howarth, and Shadbolt (1996) adopted a rather different approach to ocular dominance by using a network which allowed constant sprouting and retraction of connections. At each time step, sprouting or retraction of presynaptic axons was performed probabilistically depending on how the sprouting or retraction affected the energy of the system. The energy function was constructed to directly introduce competition between inputs so that no extra normalisation terms were required. In the original model, the two eyes were never correlated, although later work showed that segregation of inputs from different eyes can occur in this model even in the presence of strong between-eye correlations (Elliot & Shadbolt, 1996).

The last entry in Table 3.4 describes a mathematical analysis and simulation of ocular dominance within a competitive network (Bauer, Brockmann, & Geisel, 1997). This was not the first model to use a competitive network for ocular dominance, but the other models also investigated the development of topography, and so are discussed in the next section. Analysis of the Kohonen rule showed that ocular dominance can develop in the presence of between-eye correlations, and that the larger the between-eye correlations, the narrower the width of the ocular dominance stripes. Numerical simulation of the model showed a close fit to the analytical results, verifying the analysis. Additionally, the simulation results were unaffected by the choice of postsynaptic normalisation (divisive or subtractive), in comparison to earlier work by Goodhill (1992).

Models of ocular dominance development

Model	Encoding of ocularity	Correlated output	Modification rule	Constraints	Stripe width
Malsburg (76)	Retinal inputs: correlated within-eye and anticorrelated between-eyes.	Fixed Mexican-hat lateral interactions.	Correlational rule.	Postsynaptic normalisation of weights.	Range of inhibitory lateral connections.
Malsburg (79)	Fixed ocularity markers unique to each eye in presynaptic units. Topographic markers fixed for all units.	Lateral diffusion of ocularity markers in postsynaptic sheet.	Similarity between presynaptic and postsynaptic markers.	Normalisation of markers at each postsynaptic unit. Divisive presynaptic normalisation of weights.	Compromise between topographic and ocularity markers.
Swindale (80)	Synaptic density functions with Mexican-hat profile for synapses from same-eye and inverse Mexican-hat profile for synapses from different eyes.	Modification rule sums over neighbouring synapses using synaptic density functions.	Weight change proportional to convolution of neighbouring synapses and synaptic density functions within and between eyes.	Implicit postsynaptic normalisation of weights. Upper bound on each weight.	Extent of within-eye and between-eye synaptic density functions.
Bienenstock (82)	Retinal inputs: uncorrelated between eyes.	Not considered.	BCM rule.	Time varying threshold.	Not considered.
Miller (89)	Correlation matrix: Gaussian within-eye correlations and zero (or inverse Gaussian) between-eye correlations.	Fixed lateral interactions.	Correlational rule.	Postsynaptic normalisation.	Extent of within-eye correlations and lateral interactions.
Montague (91)	Retinal inputs: anticorrelated waves.	Lateral diffusion of postsynaptic molecule.	Covariance-based rule.	Bounds on each weight. Other normalisations not mentioned.	Not considered.
Elliot (96)	Retinal inputs: anticorrelated between eyes.	Energy function sums over nearest postsynaptic neighbours.	Probabilistic sprouting/retraction according to energy of system.	None (sprouting and retraction are independent).	Not considered.
Bauer (97)	Retinal inputs: correlated within and between eyes.	Neighbourhood weight update.	Kohonen rule.	None.	Between-eye correlations and width of neighbourhood.

Table 3.4: Summary of the main models of ocular dominance development. Items not investigated by the models are marked “Not considered”. Full references for each model can be found in the text.

Models of the joint development of ocular dominance and topography

Model	Encoding of ocularity and topography	Correlated output	Modification rule	Constraints	Stripe width
Obermayer (91)	Feature vector $(x, y, \phi, ,)$.	Neighbourhood weight update.	Kohonen rule.	NonenyNo	

3.12 Discussion

next chapter, we therefore consider the form of the retinotopic map produced across the whole postsynaptic sheet of the Keesing model.

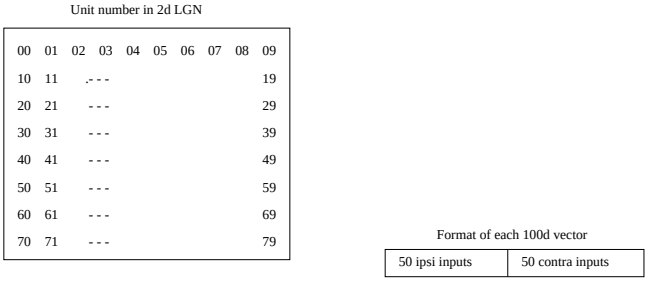
Most of the previous modelling work has demonstrated that development of monocular units depends on the correlations between eyes, the type of weight modification rule, and the type of normalisation schemes used. First of all, when the two eyes are anticorrelated, all of the models can produce monocular postsynaptic units. This is expected since presynaptic units from different eyes are never simultaneously active. For the correlational rule, an eigenvector analysis of a correlation matrix with between-eye anticorrelations predicts that monocular receptive fields will dominate development (Miller & MacKay, 1992; Miller & Mackay, 1994). Most models additionally require some form of normalisation to keep weights within bounds and to introduce competition. Out of these models, only the tea trade model used presynaptic, rather than postsynaptic, weight normalisation, but even this model also used postsynaptic normalisation of ocular markers (von der Malsburg, 1979). Presynaptic normalisation, when used, is only required to keep presynaptic units connected to the same postsynaptic sheet.

4.4 Results

In the following sections, we present the results of using the Keesing model under various different conditions. To provide a fair comparison between the different simulations, unless stated otherwise, all parameter values were kept the same. A list of the model parameters, along with their meaning and typical values, are given in Table 4.1.

Parameter	Meaning	Typical value
	Probability of a new wave starting	0.02
σ	Standard deviation of the wave	0.85 – 1.00
R	Refractory period between waves	1
T	Number of time steps wave is present on retina	50 ($_{\text{pre}}$)
$_{\text{pre}}$	Width of each retina	50
$_{\text{pre}}$	Number of presynaptic units	100 ($_{\text{pre}} \times$ 2 eyes)
$_{\text{post}}$	Width of the LGN	10
$_{\text{post}}$	Height of the LGN	8
$_{\text{post}}$	Number of postsynaptic units	80 ($_{\text{post}} \times$ $_{\text{post}}$)
$_{\text{pre}}$	Target sum for weights from each presynaptic unit	1.0
$_{\text{post}}$	Target sum for weights to each postsynaptic unit	1.25 †
	Enforcement rate of subtractive normalisation	1.0
ε	Rate of weight update in covariance rule	0.001
α	Presynaptic threshold in covariance rule	0.1
β	Postsynaptic threshold in covariance rule	0.0125
γ	Constant for growth term	0.1
	Radius of growth rule	[2,1,0]
	Probability of using growth rule	0.01
	Number of epochs between radius decreasing by one	200

networks, in practice the range of weights in most experiments is roughly constant. (The two notable exceptions to this are shown in Figure 4.12, in the absence of any normalisation, and Figure 5.4, when the probability of waves being generated is very small.)



To reduce the effect of the non-dominant eye on the centre of mass measurements, we take the centre of mass and standard deviation for the subset of the weight vector from the eye providing the dominant input to the unit (given by). The centre of mass, , and the standard deviation, , for each postsynaptic unit are given by:

$$= \left(\frac{1}{} \sum_{=1}^{\text{pre}} ((-
$$

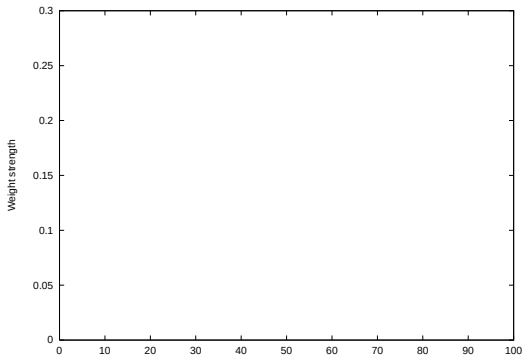
In addition to these plots, which show the state of the network at one particular time during development, two types of plot are used to summarise how the network develops over the time course of the simulation:

- Refinement of receptive field size.

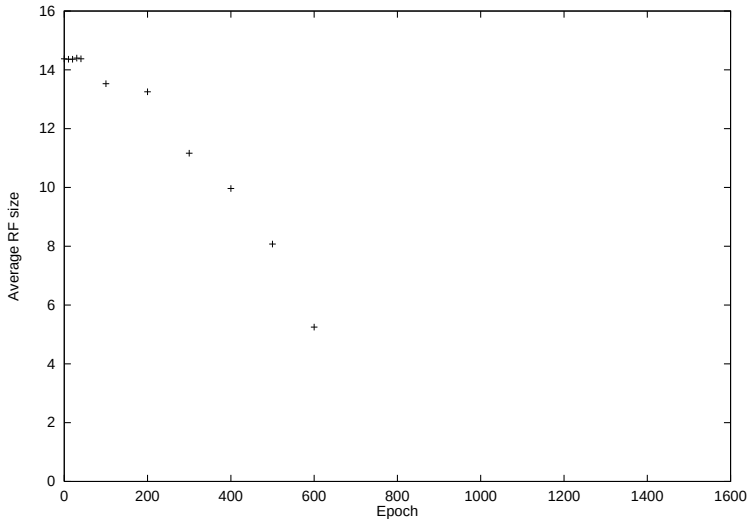
  The standard deviation of the centre of mass for a postsynaptic unit's weight vector can be used as a rough indicator of the width of the unit's receptive field. These widths can be averaged over all postsynaptic units to produce a "mean receptive field width", which is plotted at various stages of development. Error bars for each point indicate ± 1.0 standard deviation of the receptive field width. An example is given in Figure 4.5(a).

- Development of monocularity.

  This plot shows the average value of the monocularity index $,$ for all postsynaptic units at various points during development for both the left ($, > 0$) and right ($, < 0$) eyes. Error bars for each point indicate ± 1.0 standard deviation of the $,$ values. An example is given in Figure 4.5(b).

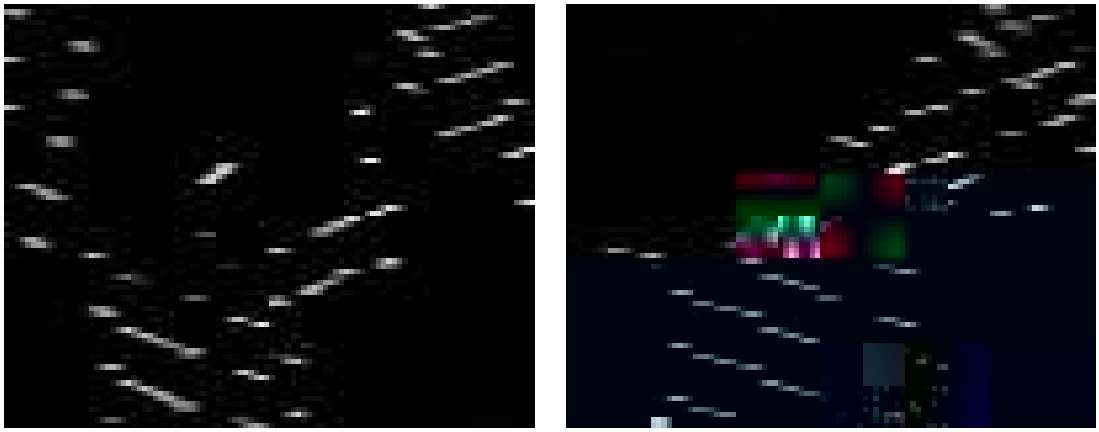

Weight strength
0.3
0.25
0.2
0.15
0.1
0.05
0
0
10
20
30
40
50
60
70
80
90
100

ferret retinal inputs cluster into patches rather than layers. This evidence can be interpreted as suggesting that extra information must be provided for the formation of layers, and it is likely that this information is specified by activity independent mechanisms (Sur, 1995; Angelucci, Clasca,


Average RF size
Epoch
0
2
4
6
8
10
12
14
16
0
200
400
600
800
1000
1200
1400
1600

Pre+Post normalisation

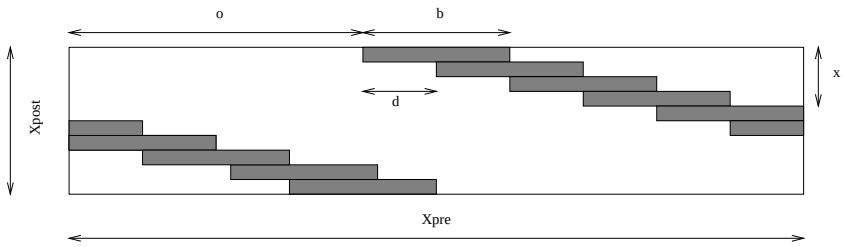
Pre normalisation only



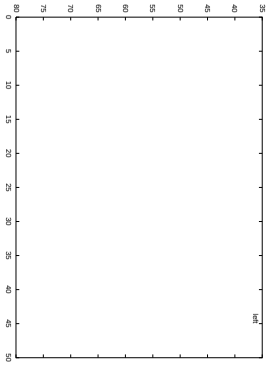
(a) Raw weight matrix

Pre+Post normalisation

Pre normalisation only


o
b
x
d
Xpost
Xpre

bias, the amount of topographic bias was reduced to = 10. The initial weights, along with


80
75
70
65
60
55
50
45
40
35
0
5
10
15
20
25
30
35
40
45
50
RH

4.6 The importance of normalisation

As mentioned in Chapter 3, most previous models of ocular dominance require postsynaptic normalisation of weights to allow ocular dominance to develop. However, the results from Figure 4.6 and Figure 4.7 show that ocular dominance develops within this model either with or without any postsynaptic normalisation. These experiments only investigated network development with ocularity bias in the weights. To address the issue of normalisation more thoroughly, a set of experiments were therefore run to examine the effects of all possible combinations of normalisation techniques.

Normalisation can be applied either divisively or subtractively to both the presynaptic and postsynaptic units. In addition, there is the option to ignore normalisation of either pre- or postsynaptic units. This gives us nine combinations of normalisation techniques. The same network with weights initially biased for ocularity and topography (using the same initial weights as shown in Figure 4.10) was run nine times using each combination of normalisation technique. The results from these nine experiments are summarised in Figures 4.12, 4.13, 4.14 and 4.15.

From these figures, several points can be made concerning the role of the different normalisation techniques upon network development.

4.6.1 The effect of normalisation upon topography

First of all, for the normal pattern of topography to develop, the presynaptic normalisation must be divisive. As long as the presynaptic normalisation is divisive, the form of the postsynaptic normalisation is mostly redundant, as shown by Figures 4.12

tioo theelah ws lhb-4.10914(h)-4.64422(l)-6.9307(l)-4.106939ddlffarancsl Fiaslf thee poj-6.93181(t)-6.93404(e)5.6

and therefore the presynaptic units have disconnected from the postsynaptic sheet. (Any activity in these presynaptic units is therefore not propagated to the postsynaptic sheet.) If the growth rule is removed from the model, all presynaptic units can remain connected to the postsynaptic units in the absence of presynaptic normalisation (Figure 4.18(d)), although this prevents the nor-

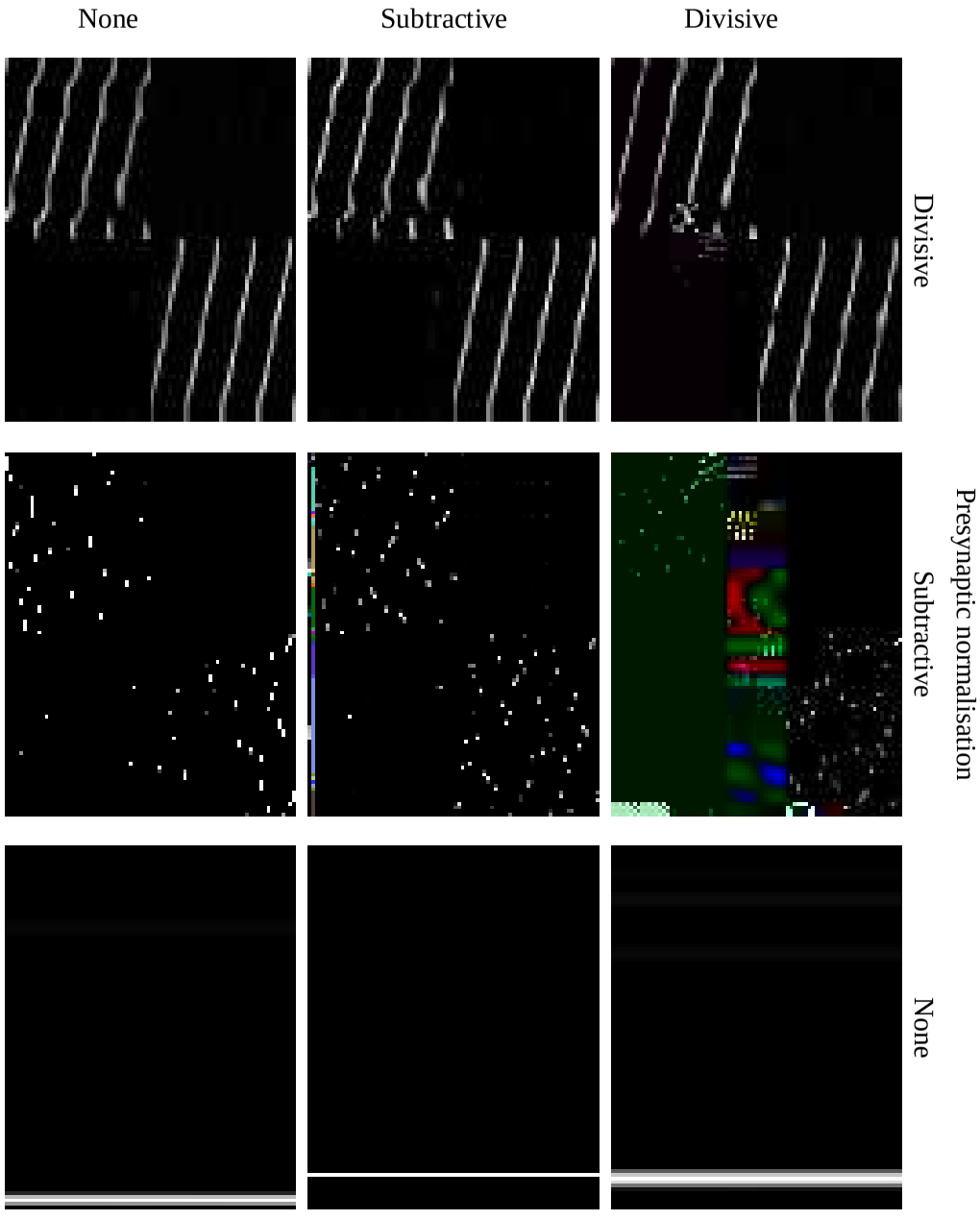


Figure 4.12: The effect of different normalisation methods – 1: weight matrices.

Presynaptic normalisation

Divisive

Subtractive

None

Postsynaptic normalisation

Divisive

Presynaptic normalisation

Divisive | Subtractive | None

Divisive

81

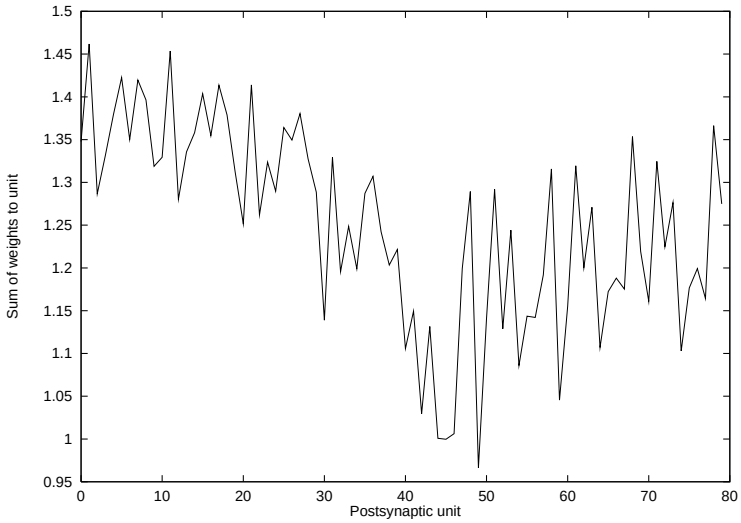


Figure 4.16: Sum of weights to each postsynaptic unit in a network with divisive presynaptic normalisation and no postsynaptic normalisation. For each postsynaptic unit , the sum $\sum^{pre}$ is shown.

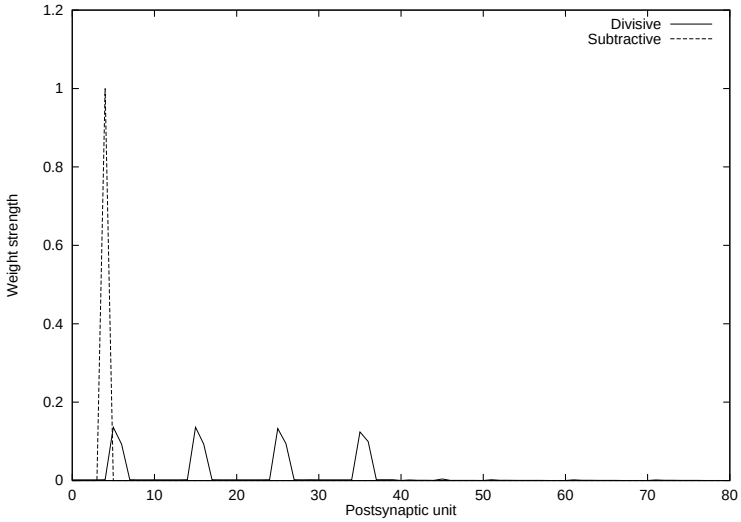


Figure 4.17: Projective field for one presynaptic unit from the right eye using either divisive (solid line) or subtractive (dashed line) presynaptic normalisation. These weights were taken from presynaptic unit 75 in the experiments with no postsynaptic normalisation.

4.6.3 The importance of the ordering of the normalisation techniques

The previous section has shown that presynaptic normalisation plays a crucial role in development, unlike postsynaptic normalisation. In these experiments, the presynaptic normalisation was always applied before the postsynaptic normalisation. Postsynaptic normalisation may be redundant simply because it is always performed after the presynaptic normalisation. To test for this possibility, a set of experiments were run varying the probability of whether presynaptic or postsynaptic normalisation was applied first.

A new parameter, , was therefore introduced. This controlled the probability of postsynaptic normalisation occurring before presynaptic normalisation. For all of the experiments presented so far in this chapter, this parameter was implicitly set to 0.0 so that divisive presynaptic normalisation was applied first and then subtractive postsynaptic normalisation afterwards. Using the same set of initial conditions as for the previous normalisation experiments, network development was monitored using a range of values for . The results of these experiments are shown in Figure 4.19.

For all values of , the network developed the usual two eye-specific layers, although high values of produced a small number of binocular units and monocular units in the wrong layer. The strongest effect of the parameter was that it affected receptive field size: high values of produced sharper and narrower postsynaptic receptive fields. This sharpening effect of the receptive fields is due to the subtractive normalisation which is sometimes applied first when > 0.0. As a consequence of the sharper receptive fields, the topography within each row of the LGN is not so smooth as the value of

normalisation applied second will have a very low error.) Th

4.6.5 Capping subtractive presynaptic normalisation

The previous experiments have shown the importance of divisive presynaptic normalisation for the development of topography in the LGN. Networks using subtractive presynaptic normalisation are unable to replicate the topographic projections due to the way that the normalisation pushes individual weights to extreme values. Since there is no maximum weight value imposed on weights in the network, subtractive normalisation forces all of the synaptic weight strength into one of the elements of the weight vector, with all other elements pushed to the minimum weight value of zero, as shown in Figure 4.17. If, however, there is a maximum value imposed on each weight (which will be called $_{\max}$), more than one weight for a unit is forced to take a non-zero value.

weights have strength 0 or $_{max}$. In contrast, simulations using divisive presynaptic normalisation produce weights distributed over a wide range of values.

Projective field

Weight matrix

Ocular dominance plot

4.7 Summary

This chapter has introduced the Keesing et al. (1992) model of retinogeniculate development. In the first half of this chapter, we have replicated the results of the Keesing model, showing in particular how the form of the initial weight bias influences network development. This replication stage was necessary since the only published information on the model omitted a number of crucial details of the model, including the form of the inputs and the nature of the initial weights (Keesing et al., 1992). We have been able to replicate these initial results, and also expand on the nature of the topographic map across all rows of the LGN, another aspect of the model that was not described in the original publication.

The second half of the chapter has analysed the role of the normalisation mechanisms in network development. In contrast to the original presentation of the model (Keesing et al., 1992), we have shown here that the form of the postsynaptic normalisation is redundant, as long as the presynaptic normalisation is implemented divisively, rather than subtractively. This is a new result, in contrast to previous models of ocular dominance which have used some form of postsynaptic normalisation (reviewed in Chapter 3) to ensure postsynaptic units become monocular. The only other model that relies on presynaptic normalisation of weights is the model by (von der Malsburg, 1979), although it also used a normalisation of the ocular marker concentrations induced into each postsynaptic unit. Postsynaptic normalisation (of either weights or markers) is very likely to produce monocular units, since any increase in weight strengths for some units from one eye is accompanied by a uniform decrease of all weights. Using subtractive enforcement of postsynaptic normalisation increases the tendency for monocular units due to the way that subtractive normalisation forces individual weights to extreme values. We have also shown that in some cases, the presynaptic normalisation can be implemented subtractively, as long as the normalisation is applied slowly and individual weights are constrained to

Chapter 5

5.2.1 The relationship between the rate of wave generation and the probability of activity in both eyes

When the waves are independently generated in each eye, thre

		0	1	2
0.002	0.091	0.826	0.165	0.008
0.003	0.130	0.756	0.227	0.017
0.005	0.200	0.640	0.320	0.040
0.020	0.500	0.250	0.500	0.250
0.200	0.909	0.008	0.165	0.826
0.500	0.962	0.001	0.074	0.925
0.800	0.976	0.001	0.047	0.952

Table 5.1: Theoretical probability of eye activity as a function of . Measurements of the corresponding probabilities from the simulations of eye activity were always within four percent of the predicted values.

1. $_0 > \max(\ _1,\ _2)$ $(\quad \leq 0.01)$. Here it is most likely that both eyes are quiet.

2. $_1 > \max(\ _0,\ _2)$ $(0.01 < \quad < 0.04)$. In this range, the most likely situation is that one eye is active while the other eye is quiet.

3. $_2 > \max(\ _0,\ _1)$ $(\quad \geq 0.04)$. In this last range (the widest of the three ranges), both eyes are likely to be jointly active most of the time.

the normal patterns of ocular dominance and topography. A common feature of both rules however is that the average receptive field width increases inversely with (see Table 5.2). These results must be interpreted with some caution however, due to the rather high standard deviation of receptive field widths in comparison to their mean values.

	Mean $\pm$ standard deviation	
	Covariance	Active-cov.
0.002	N/A	11.384 ± 4.718
0.003	N/A	5.793 ± 5.770
0.005	N/A	3.762 ± 4.965
0.020	4.305 ± 4.956	1.591 ± 2.780
0.200	2.199 ± 0.869	0.889 ± 0.046
0.500	2.184 ± 1.001	0.887 ± 0.054
0.800	2.109 ± 0.679	0.896 ± 0.111

Table 5.2: Mean receptive field width for different values of using either the covariance or active-covariance rule. For the entries listed N/A, all postsynaptic units were binocular, which precluded taking any receptive field measurements.

	Frequency of use / %				Weight change per epoch			
	1	2	3	4	1	2	3	4
0.002	0.495	0.011	15.468	84.027	95.5920	-1.1990	-666.186	5830.838
0.003	0.763	0.015	23.032	76.191	164.7320	-1.4550	-1062.795	5279.894
0.005	1.103	0.018	32.381	66.498	257.7480	-1.7380	-1571.903	4600.636
0.020	1.062	1.732	23.810	73.396	1344.4520	-393.7110	-7218.146	4541.466
0.200	1.324	3.427	23.559	71.690	3108.4650	-819.3380	-13704.282	4365.681
0.500	1.459	3.540	25.339	69.661	3339.5660	-852.4340	-14409.363	4250.804
0.800	1.609	3.453	27.526	67.412	3448.7770	-846.1370	-14552.396	4185.087

(a) Covariance rule

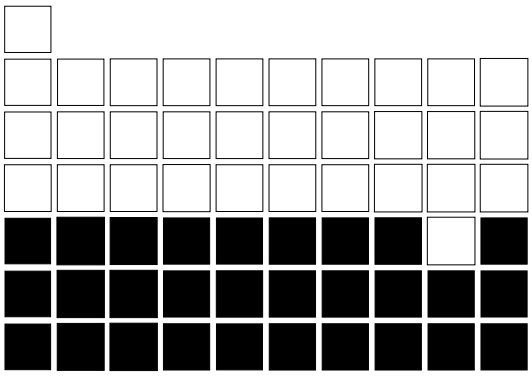
	Frequency of use / %				Weight change per epoch			
	1	2	3	4	1	2	3	4
0.002	0.382	0.124	11.815	87.680	112.7830	-24.6050	-806.365	6017.041
0.003	0.409	0.368	12.171	87.052	242.7040	-86.3470	-1652.527	5920.608
0.005	0.403	0.718	11.412	87.467	432.6690	-183.7100	-2872.521	5948.982
0.020	0.671	2.123	14.894	82.312	1513.0623	-562.6000	-8218.576	5538.130
0.200	1.063	3.689	18.945	76.303	3235.3990	-957.5940	-14278.994	4969.223
0.500	1.137	3.863	19.737	75.264	3473.8570	-999.0490	-15003.074	4875.800
0.800	1.249	3.813	21.323	73.615	3577.4920	-987.0899	-15108.254	4773.361

(b) Active-covariance rule

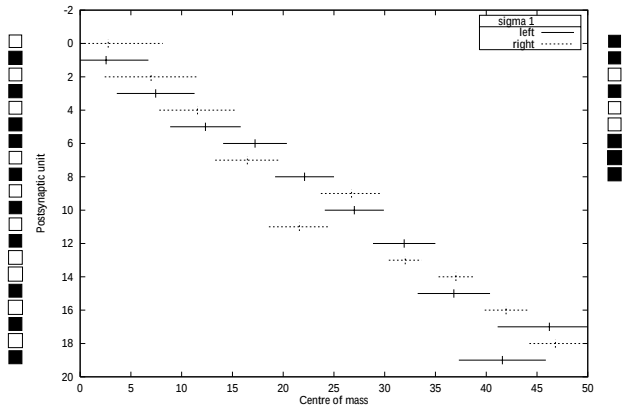
Table 5.3: The frequency of use and the amount of weight change per epoch for each case of the covariance and active-covariance rule. Cases refer to the cases of the covariance rule as shown in Table 3.2. For the active-covariance rule, although values of weight change for case four are given here, these weight changes were ignored.

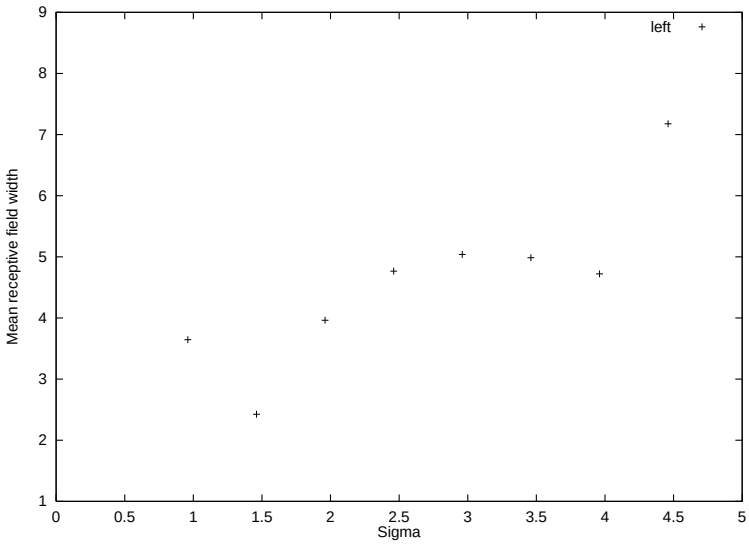
5.3 Development of the LGN under conditions of monocular deprivation

Multiplicative and divisive normalisation



Divisive normalisation


sigma 1
left
right
Postsynaptic unit
Centre of mass
-2
0
2
4
6
8
10
12
14
16
18
20
0
5
10
15
20
25
30
35
40
45
50


left
Mean receptive field width
Sigma
9
8
7
6
5
4
3
2
1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

unit i, the polarity dominance is denoted f_i, and is defined as:

$$f_i = f_i^{\text{on}}$$

postsynaptic units.

The introduction of polarity-specific waves did not radically affect the nature of the topographic mapping, as shown by the topographic plots in Figure 5.12. The final projection columns for the network, shown in Figure 5.14, are well organised, with only a few units in the wrong position. The average wave width at the end of development was slightly bigger than for the previous simulations (mean = 3.795, s.d. = 0.695, compared with mean = 2.496, s.d. = 0.737 for L4R8 of Figure 4.11 and mean = 2.922, s.d. = 0.894 for L2R4 of Figure 4.11). This increase in receptive field width was expected due to the increase in the width of the retinal waves after 400 epochs.


1
0.8
0.6
0.4
0.2
0
-0.2

5.5.4 Weakening the anticorrelations between on- and off-centre inputs

The parameters $_{on}$ and $_{off}$ control the nature of the correlations between the on- and off-centre units. In the previous section, the values $_{on} = 0.5,\ _{off} = 1.0$ were used. This ensured that once the polarity-specific waves had been introduced, an on-centre unit and off-centre unit would never be jointly active, maximising the chance that units of different polarity would not jointly innervate the same postsynaptic unit. However, such anticorrelations have not been found in the developing retina. Instead, as mentioned before, when on-centre cells are active, neighbouring off-centre cells also tend to be active, but not vice-versa (Wong & Oakley, 1996). These weaker anticorrelations can be modelled by reducing the value of $_{off}$ below 1.0 to allow some retinal waves to have both on- and off-centre activity (using the default rule from Table 5.5).

Since off-centre cells are active more often than on-centre cells during the period of polarity-specific waves, a set of experiments were performed with $_{on}$ reduced below 0.5, and $_{off}$ set to $_{on} + 0.5$. This means that off-centre waves are present half of the time, with the rest of the time divided between on-centre waves and mixed on- and off-centre waves. As $_{on}$ is reduced to 0.0, only off-centre and mixed on- and off-centre waves are generated. This is the situation found in the developing retina (Wong & Oakley, 1996).

Figure 5.15 shows the results of development using $_{on} = 0.48, 0.46, 0.44, 0.42$ with $_{off} = _{on} + 0.5$ in each case. The plots of polarity segregation for each experiment clearly show that as $_{on}$ decreases, the overall degree of polarity segregation is greatly reduced. The other network features, ocular dominance and topography, were unaffected. The existence of mixed on- and

5.6 Comparison with biological data

light-evoked anticorrelations between on- and off-centre cells driving polarity segregation.

5.7 Discussion

The formulation of the waves used in this thesis is relativel

Chapter 6

preventing them from adapting in the same direction to the input stimuli. Miller's model is slightly more complicated: in the absence of inhibitory lateral connections, as long as there is presynaptic normalisation then stripes still develop (Miller et al., 1989).

6.2.2 Competition between topography and ocularity

For those models considering the development of both ocular dominance and topography, stripes arise through competition between these two features as wel

with highest variance, and represents them in the network. Other elements of the input vector with variance below a critical value do not get represented in the map (Ritter & Schulten, 1988). This is referred to as the "automatic selection of feature dimensions" (Kohonen, 1988). As is shown in more detail in the next section, the variance of the ocularity feature controls stripe formation: the higher the ocularity variance (compared with the other input variances), the wider the ocular dominance stripes.

6.3 Obermayer's model for stripe formation

All of the models mentioned in the last section provide different reasons for how stripe width varies under different conditions. Out of all of the models, the model by Obermayer et al. (1991) is arguably the simplest because it uses feature vectors to represent the neural activity distributed across two retinae. This model has therefore been chosen to investigate the nature of stripe formation in the cortical simulations. We also extend the postsynaptic sheet into a three-dimensional block so that the model can be applied to the problem of retinogeniculate development.

6.3.1 Implementation details of Obermayer experiments

The network consists of a presynaptic sheet with three units fully connected to a set of postsynaptic units. The postsynaptic units are arranged into a three-dimensional block for the purposes of neighbourhood weight updating. (For the two-dimensional simulations, the Z dimension of the postsynaptic block was set to one.)

The initial weights are set at random, with no topographic or ocular bias. One iteration of the model consists of generating and presenting a feature vector, calculating postsynaptic unit activations and updating the weights of the winning and neighbouring units. One epoch of the model corresponds to 100 iterations, after which various parameters, such as the weight-update rate and size of neighbourhood, are updated. Table 6.1 describes the equations and parameters used for these experiments.

6.3.2 Visualisation of maps formed in Obermayer experiments

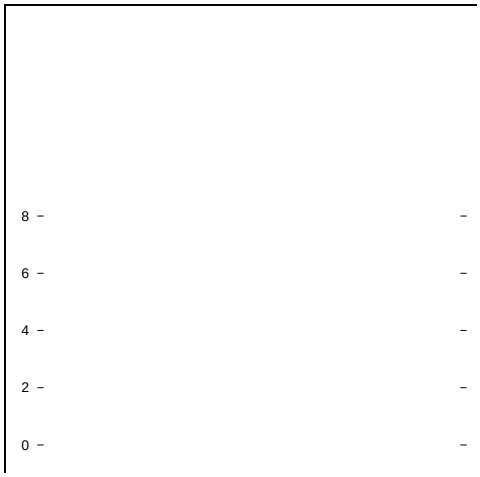
Each postsynaptic unit receives inputs via three weights. The first two weights code for the centre of mass and the third weight codes for the ocularity to which the unit is most responsive. Visualisation of each postsynaptic unit's preferred stimulus is straightforward. The centre of mass for each postsynaptic unit is drawn on a graph at the point (= , =). Lines are drawn between the points of neighbouring postsynaptic units to indicate the location of neighbouring postsynaptic units. The ocularity of each postsynaptic unit is encoded in = . The value of for all postsynaptic units is visualised using a Hinton diagram — the size of the square is proportional to the magnitude of (scaled to a maximum size of) and the colour of the square indicates the sign of (black for negative values representing dominant left-eye input and white for positive values representing dominant right-eye input).

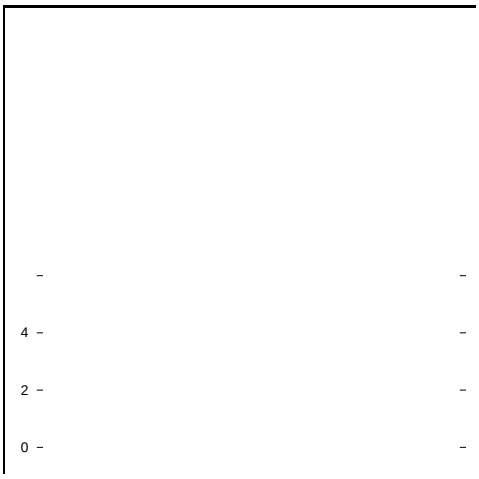
6.3.3 How the variance of ocularity affects stripe width

To illustrate the principle of feature selection within the standard cortical model, several experiments similar to those presented in (Obermayer et al., 1991) were replicated. Three-dimensional

Inputs	$x_1 = \text{rnd}(0,15)$, $x_2 = \text{rnd}(0,15)$, $x_3 = \pm$
Activation rule	$y_j = \sum_{i=1}^{3}(w_{ij} - x_i)^2$
Winner	Find j : $y_j \leq y_j \forall j$
Weight modification	$\Delta w_{ij} = \varepsilon(t)\, h(j, j^*)(x_i - w_{ij})$ (t is the epoch number)
Distance measure	$d(j, j^*) = (d_x(j) - d_x(j^*), d_y(j) - d_y(j^*), d_z(j) - d_z(j^*))$
Index function	$j = x(j) + (N_{post} \times y(j)) + ((N_{post} \times N_{post}) \times z(j))$†
Neighbourhood function	$h(x, y, t) = \exp($

input vectors (, ,) were used as input to a Kohonen network with a two-dimensional sheet of postsynaptic units of size $_{post} = 32$, $_{post} = 32$ (assuming $_{post} = 1$). Each feature vector repre-


8
6
4
2
0


4
2
0

6.4 A new model of retinotopic map formation in the LGN

In the previous experiments, the units in the postsynaptic sheet were arranged in a two-dimensional sheet. In the experiments presented in this section, the postsynaptic sheet is extended into a three-dimensional block to make the model similar to the LGN. The three-dimensional arrangement of the postsynaptic block is shown in Figure 6.3. All other elements of the model remained the same.

When the network is presented with the three-dimensional feature vector, postsynaptic units become monocular even with values of that were too low for ocularity to be represented in the two-dimensional network. Figure 6.4 shows the results of presenting the same retinal inputs as used in the two-dimensional experiments into the three-dimensional postsynaptic block of size ($_{\text{post}} = 32$, $_{\text{post}} = 32$, $_{\text{post}} = 4$), for $= 0.2$ and $= 2.0$.

To analyse the maps, the three-dimensional postsynaptic block has been divided into multiple two-dimensional planes (as illustrated in Figure 6.3). Z-plane corresponds to all of the units in the plane $=$ of the postsynaptic block. Figure 6.4 shows the results of network development with $= 0.2$. In this case, the top two Z-planes of the network responded to the right eye and the bottom two Z-planes of the network to the left eye. Within each Z-plane, the representation of visual space was complete and topographic. Figure 6.5 shows the topographic maps from the same network when the postsynaptic sheet is divided into multiple X-planes. (X-plane corresponds to all of the units in the plane $=$ of the postsynaptic block.) Each X-plane covers a small part of the visual space, but all of the X-planes taken together cover the entire visual space. Additionally, neighbouring X-planes cover neighbouring parts of the visual space. (X-planes 1 and 32 cover a slightly larger region of visual space than the other planes due to boundary effects.) The X-planes are similar to the projection columns found in the LGN (Sanderson, 1971a).

However, with a bigger value of , such as $= 2.0$, ocularity is the primary map feature and visual space a secondary feature (similar to the two-dimensional network shown in Figure 6.2). Each Z-plane of the network developed in the same manner, although the border between left- and right-eye regions varied systematically through the planes. In Figure 6.4, the left half of each Z-plane is responsive to the right eye, and the right half of each Z-plane is responsive to the left eye. In this case, the topographic plots in both the X- and Y-planes show no correspondence with the LGN projection columns. A sample of some topographic maps in the X- and Y-planes are shown in Figure 6.6. For the X-plane plots we find that the maps for the planes equidistant from the centre (X-planes and $33 -$ for all values of $= 1 \ldots 16$) cover almost the same region of the input space. For each pair of X-planes, X-plane contains units responsive to the right eye, and X-plane $33 -$ contains units responsive to the left eye. In contrast, each Y-plane topographic plot has folded over on top of itself, so that each part of visual space is covered by two units (one for each eye) in each Y-plane.

6.5 Correspondence with topography and ocularity in the LGN

The cat LGN has a very distinctive three-dimensional shape, as shown in Figure 2.3. It can be approximated however as a three-dimensional block of postsynaptic units, with roughly equal extent in the lateral-medial (X) and anterior-posterior (Y) dimensions, but a much smaller dorsal-ventral (Z) extent. The maps in Figure 6.4 can be compared with the retinotopic and ocularity

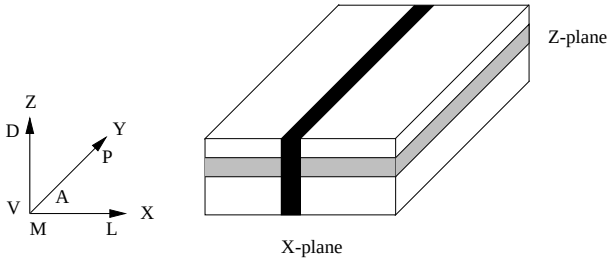


Figure 6.3: Three-dimensional arrangement of postsynaptic units in the Kohonen network. Postsynaptic units are positioned uniformly throughout the three-dimensional block. The X dimension of this block corresponds to the medial—lateral dimension of the LGN. Likewise, the Y dimension corresponds to the anterior—posterior dimension, and the Z dimension corresponds to the dorsal—ventral dimension. The figure shows two different two-dimensional slices through the postsynaptic block. The grey slice is a Z-plane which samples all postsynaptic units at the same dorso-ventral position. The black slice is a X-plane which samples all postsynaptic units at the same medio-lateral position. Key: A – anterior. P – posterior. D – dorsal. V – ventral. L – lateral. M – medial.

mappings within the LGN.

First, when = 0.2, the plots show a similar retinotopic organisation to that found in the LGN: units at the same (,) position but at different depths ( ) receive input from the same part of visual space. Furthermore, the map has segregated into eye-specific layers (where a layer is used here to mean the same as a Z-plane) in the same way as the LGN. (In the LGN, the contralateral inputs always innervate the top of the LGN. In these experiments there is no preference for the contralateral inputs to go to the top of the LGN, and so the eye providing input to the top Z-plane varies from simulation to simulation. It should be relatively simple to ensure that the contralateral eye always dominates the top half of the postsynaptic block by placing a suitable bias on the initial weights.) The map has automatically oriented itself to represent the input features with the highest variance (,) along the largest dimensions of the postsynaptic block (,). The remaining input feature with the smallest variance, , is mapped onto the smallest dimension of the postsynaptic block ( ).

This segregation into eye-specific layers is dependent however on the number of Z-planes in the postsynaptic block. When = 4, the network can easily divide into two halves, so that Z-planes 1 and 2 can respond to one eye, and Z-planes 3 and 4 can respond to the other eye. When = 3 however, the network fails to completely segregate into separate layers, as shown in Figure 6.7. Monocular units for each eye are found in each Z-plane, although as shown in Table 6.2, there is a tendency for units from the left eye to settle in Z-planes 1 and 2 and for the right eye to settle in Z-planes 2 and 3. This non-complete segregation into eye-specific layers also affects the topography of units within each layer: units responding to different eyes respond to different parts of the visual space. This failure to segregate into eye-specific layers is most likely to be due to the odd number of Z-planes in the network, making it impossible for an equal number of Z-planes to be responsive to each eye.

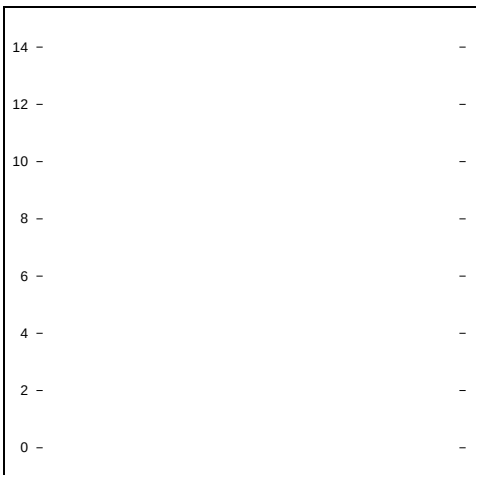
However, when = 2.0, the map develops a different retinotopic and ocular structure to that found in the LGN. For the map in Figure 6.4, the primary feature, ocularity, is mapped along the dimension of the postsynaptic block, rather than the dimension as before. The remaining

inputs, (x and y), map along the X and Y dimensions of the postsynaptic block. The structure of the map within each Z-plane is almost identical, except for a shift in boundary position between different ocularity values. This does not correspond to the topographic and ocularity maps found in the LGN, since both space and ocularity are primary features in the LGN.

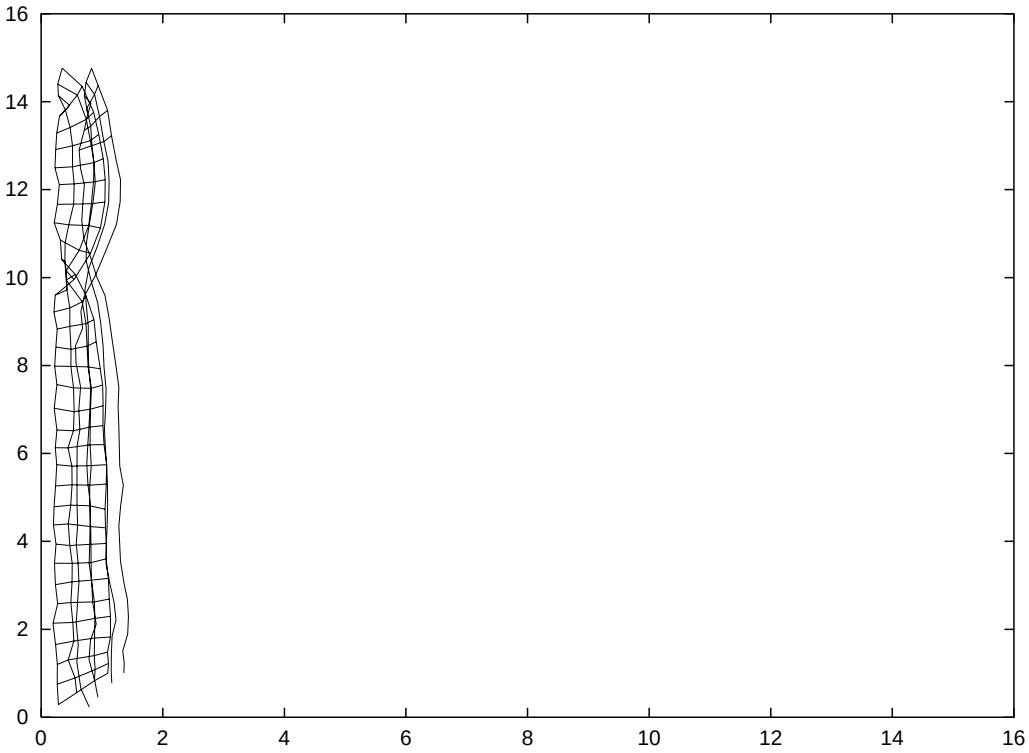
Figure		$_{post}$	Z-plane	>0 (right eye) n	mean $\pm$ s.d.	<0 (left eye) n	mean $\pm$ s.d.
6.1(a)	0.2	1	1	529	0.039 ± 0.029	495	-0.041 ± 0.030
6.1(b)	0.6	1	1	503	0.475 ± 0.162	521	-0.474 ± 0.154
6.1(c)	1.0	1	1	545	0.854 ± 0.301	479	-0.851 ± 0.305
6.2	2.0	1	1	509	1.964 ± 0.177	515	-1.962 ± 0.191
6.7	1.0	3	1	310	0.741 ± 0.306	714	-0.877 ± 0.230
			2	523	0.673 ± 0.347	501	-0.671 ± 0.346
			3	731	0.880 ± 0.227	293	-0.770 ± 0.305
6.4	0.2	4	1	1016	0.180 ± 0.028	8	-0.044 ± 0.026
			2	967	0.122 ± 0.034	57	-0.086 ± 0.041
			3	72	0.094 ± 0.035	952	-0.122 ± 0.032
			4	24	0.040 ± 0.037	1000	-0.178 ± 0.033
6.4	2.0	4	1	554	1.938 ± 0.284	470	-1.893 ± 0.383
			2	541	1.891 ± 0.351	483	-1.937 ± 0.216
			3	479	1.930 ± 0.244	545	-1.890 ± 0.355
			4	468	1.885 ± 0.390	556	-1.947 ± 0.242

Table

= 0.2 = 2.0



Z-plane 1


16
14
12
10
8
6
4
2
0
0
2
4
6
8
10
12
14
16

6.6 Discussion

In this chapter, we have presented three main arguments for the development of ocular dominance stripes. Although these arguments were developed with segr

block. So, even when the network does not need to perform dime

Chapter 7

Conclusions

The aim of this thesis has to been to investigate the hypothesis that spontaneous waves of activity

axis, in a similar fashion to the mapping of ocularity in the LGN. This self organisation can, in principle, account for the overall layout of visual space and ocularity in the LGN. In practice however, factors such as limited axonal branching may prevent such global reorganisation.

7.2 Future work

In this section we consider several directions in which this work could be extended.

7.2.1 Modelling of the retina

The model retina used here is very simplistic and could be imp

times during map development. In the case of retinogeniculate development, projection column maps should also be taken. Pictures of the initial map layout will indicate the amount of order in the system before activity-dependent processes begin to reorganise the pathway. Subsequent maps will then show the amount of reorganisation and refinement. These maps could be generated using retrograde labelling of neighbouring geniculate cells (I. Thompson, personal communication). Producing such maps of visual space in the LGN during the period of retinogeniculate development is complicated however due to the large increase in LGN volume at this time (Elgeti et al., 1976). This may mean that it is possible to create these maps only after development of the retinogeniculate pathway and the LGN has stabilised. Topographic maps of the mature LGN under altered conditions, such as activity blockade and monocular deprivation, would, however, be useful for comparison with both control maps and model predictions.

7.3.3 Development of on- and off-centre units

The segregation of LGN layers into polarity-specific sublaminae is believed to be activity-dependent (Cramer et al., 1996; Cramer & Sur, 1997). There are two outstanding questions related to this segregation. First, why do on-centre cells always go into the dorsal region of a lamina and off-centre cells to the ventral region? This could indicate some ll then (s)-5.52048(o3)-LT[(T)-2.64311(r)26.142(g)-4.iiul

er-itmo; alicbehuilm

Bibliography

Andrade, M. A., & Moràn, F. (1996). Structural study of the development of ocularity domains using a neural network model. , , 243–254.

Angelucci, A., Clasca, F., Bricolo, E., Cramer, K. S., & Sur, M. (1997). Experimentally induced retinal projections to the ferret auditory thalamus: development of clustered eye-specific patterns in a novel target. , , 2040–2055.

Archer, S. M., Dubin, M. W., & Stark, L. A. (1982). Abnormal development of kitten retinogeniculate connectivity in the absence of action potentials. , , 743–745.

Artola, A., Bröcher, S., & Singer, W. (1990). Different voltage-dependent thresholds for the induction of long-term depression and long-term potentiation in slices of the rat visual-cortex. , , 69–72.

Atick, J. J., & Redlich, A. N. (1992). What does the retina know about natural scenes? , , 196–210.

Barrow, H. G. (1987). Learning receptive fields. In , IV, pp. 115–121.

Barrow, H. G., & Bray, A. J. (1992). A model of adaptive development of complex cortical cells. In Aleksander, I., & Taylor, J. (Eds.), , . North-Holland.

Bauer, H.-U., Brockmann, D., & Geisel, T. (1997). Analysis of ocular dominance pattern formation,

Bodnarenko, S. R., Jeyarasasingam, G., & Chalupa, L. M. (1995). Development and regulation

Dale, H. H. (1935). Pharmacology and nerve endings. *ee f e s e f e e,* , 319–332.

Dan, Y., Atick, J. J., & Reid, R. C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *f e e e,* *16*, 3351–3362.

Dayan, P. S., & Goodhill, G. J. (1992). Perturbing Hebbian rules. In Moody, J. E., Hanson, S. J., & Lippmann, R. P. (Eds.), *e e f e s e* , Vol. 4, pp. 19–26. Morgan Kaufmann, San Mateo.

Dowling, J. E. (1987). *e e* . *e f e* . Harvard University Press.

Dubin, M. W., Stark, A., & Archer, S. M. (1986). A role for action-potential activity in the development of neuronal connections in the kitten retinogeniculate pathway. *f e e e,* *6*, 1021–1036.

Durbin, R., & Mitchison, G. (1990). A dimension reduction framework for understanding cortical maps. *e,* , 644–647.

Durbin, R., & Willshaw, D. (1987). An analogue approach to the travelling salesman problem using an elastic net method. *e,* *6*, 689–691.

Eglen, S. J. (1995). Modelling the development of the cat lateral geniculate nucleus with Hebbian learning. Tech. rep. CSRP 383, Cognitive and Computing Sciences, Sussex University.

Eglen, S. J. (1996). Modelling the prenatal development of the lateral geniculate nucleus. In Silva, F. L., Principe, J. C., & Almeida, L. B. (Eds.), *e e e e e e* , Vol. 37, pp. 33–41. IOS Press, Amsterdam.

Elgeti, H., Elgeti, R., & Fleischhauer, K. (1976). Postnatal growth of the dorsal lateral geniculate nucleus of the cat. *, , ,* *9*, 1–13.

Elliot, T., Howarth, C. I., & Shadbolt, N. R. (1996). Neural computation and statistical mechanics. *ee f e s e f e* , *6* , 601–606.

Elliot, T., & Shadbolt, N. R. (1996). A mathematical model of activity-dependent, anatomical segregation induced by competition for neurotrophic support. *, e e* , *5*, 463–470.

Erwin, E., Obermayer, K., & Schulten, K. (1995). Models of orientation and ocular dominance columns in the visual-cortex — a critical comparison. *e* , , 425–468.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *e e ex,* *1*, 1–47.

Feller, M. B., Wellis, D. P., Stellwagen, D., Weblin, F. S., & Shatz, C. J. (1996). Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. *s e e,* , 1182–1187.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical-cells. *f e s e f e* , , 2379–2394.

Földiák, P. (1991). Learning invariance from transformation sequences. *e* , , 194–200.

Frank, E. (1987). The influence of neuronal activity on patterns of synaptic connections. *e e e e,* *10*, 188–190.

Hankin, M., & Lund, R. (1991). How do retinal axons find their t

Kaas, J. H., Guillery, R. W., & Allman, J. M. (1972). Some prin

Linsker, R. (1986b). From basic network principles to neura

Métin, C., & Frost, D. O. (1989). Visual responses of neurons in somatosensory cortex of hamsters with experimentally induced retinal projections to somatosensory thalamus.

Meyer, R. L. (1979). Retinotectal projection in goldfish to an inappropriate region with a reversal in polarity.

Miller, K. D. (1994). A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between on- and off-center inputs.

Miller, K. D. (1996). Synaptic economics: competition and cooperation in synaptic plasticity.

Norden, J. J., & Constantine-Paton, M. (1994). Dynamics of retinotectal synaptogenesis in normal and 3-eyed frogs: evidence for the postsynaptic regulation of synapse number. *f* , , 461–479.

Obermayer, K., Blasdel, G. G., & Schulten, K. (1991). A neural network model for the formation and for the spatial structure of retinotopic maps, orientation- and ocular dominance columns. In Kohonen, T., Mäkisara, K., Simula, O., & Kangas, J. (Eds.), *f* Helsinki.

Obermayer, K., Ritter, H., & Schulten, K. (1990). A principle for the formation of the spatial structure of cortical feature maps. , , 8345–8349.

Obermayer, K., Ritter, H., & Schulten, K. (1991). Development and spatial structure of cortical feature maps: a model study. In Lippmann, R. P., Moody, J. E., & Touretzky, D. S. (Eds.), , Vol. 3, pp. 11–17. Morgan Kaufmann, San Mateo.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *f* , *15*, 267–273.

Overton, K. J., & Arbib, M. A. (1982). The extended branch-arrow model of the formation of retino-tectal connections. , *5*, 157–175.

Pallas, S. L., & Finlay, B. L. (1991). Compensation for population-size mismatches in the hamster retinotectal system: alterations in the organization of retinal projections. , *6*, 271–281.

Penn, A. A., Gallego, R., Mooney, R., & Shatz, C. J. (1995). Spontaneous retinal inputs drive postsynaptic action potentials in the LGN. In , Vol. 21, p. 591.5.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. , , 329–342.

Prestige, M. C., & Willshaw, D. J. (1975). On a role for competition in the formation of patterned neural connexions. , *190*, 77–98.

Reiter, H. O., & Stryker, M. P. (1988). Neural plasticity without postsynaptic action-potentials — less-active inputs become dominant when kitten visual cortical-cells are pharmacologically inhibited.

Sabel, B. A., & Schneider, G. E. (1988). The principle of conservation of total axonal arborizations — massive compensatory sprouting in the hamster subcortical visual-system after early tectal lesions. , , 505–518.

Sanderson, K. J. (1971a). The projection of the visual field to the lateral geniculate and medial interlaminar nuclei in the cat. , , 101–118.

Sanderson, K. J. (1971b). Visual field projection columns and magnification factors in the lateral geniculate nucleus of the cat. , , 159–177.

Saul, A. B., & Humphrey, A. L. (1990). Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate-nucleus. , , 206–224.

Schlaggar, B. L., & O'Leary, D. D. M. (1991). Potential of visual-cortex to develop an array of functional units unique to somatosensory cortex. , , 1556–1560.

Schmidt, J. T., & Buzzard, M. (1993). Activity-driven sharpening of the retinotectal projection in

Sherman, S. M., & Guillery, R. W. (1996). Functional-organization of thalamocortical relays. *f e , 6*, 1367–1395.

Sherman, S. M., & Koch, C. (1986). The control of retinogeniculate transmission in the mammalian lateral geniculate nucleus. *e e , 6* , 1–20.

Sherman, S. M., & Koch, C. (1990). Thalamus. In Shepherd, G. M. (Ed.), *f e* (3rd edition)., chap. 8, pp. 246–278. Oxford University Press, Oxford.

Shou, T. D., & Leventhal, A. G. (1989). Organized arrangement of orientation-sensitive relay cells in the cat's lateral geniculate nucleus. *f e e, 9*, 4287–4302.

Shou, T. D., Leventhal, A. G., Thompson, K. G., & Zhou, Y. F. (1995). Direction biases of X-type and Y-type retinal ganglion-cells in the cat.

Sretavan, D. W., Shatz, C. J., & Stryker, M. P. (1988). Modification of retinal ganglion cell axon morphology by prenatal infusion of tetrodotoxin. , , 468–471.

Stone, J. (1978). The number and distribution of ganglion cells in the cat's retina. , , 0, 753–771.

Stone, J. (1983). . Plenum Press, New York.

Stryker, M. P., & Harris, W. A. (1986). Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual-cortex. , , 2117–2133.

Stryker, M. P., & Zahs, K. R. (1983). On and off sublaminae in the lateral geniculate nucleus of the ferret.

von der Malsburg, C., & Singer, W. (1988). Principles of cort

Appendix A

Mathematical details

A.1 The derivation of $\frac{\mathbf{w}}{t} = \mathbf{C}\mathbf{w}$ from the covariance rule

Some models of visual system development assume that the covariance rule described by Sejnowski (1977) can be reduced to a rule of the form $\frac{\mathbf{w}}{t} = \mathbf{C}\mathbf{w}$ (Linsker, 1986a; Miller et al., 1989).

Assuming that the average value of each input is the same:

$$\langle x_i \rangle = \bar{x}, \quad \forall i \tag{A.4}$$

$$\langle \Delta w_i \rangle = \alpha \sum_j w_j \langle x_i x_j \rangle - \alpha_0 \bar{x} - \alpha x_0 \bar{x} \sum_j w_j + \alpha_0 x_0 \tag{A.5}$$

The covariance matrix of the inputs, **C**, is defined as:

$$\begin{aligned}
\mathbf{C}_{ij} &= \langle (x_i - \bar{x}_i)(x_j - \bar{x}_j) \rangle \\
&= \langle x_i x_j \rangle - \langle x_i \bar{x}_j \rangle - \langle \bar{x}_i x_j \rangle + \langle \bar{x}_i \bar{x}_j \rangle \\
&= \langle x_i x_j \rangle - \bar{x}_j \langle x_i \rangle - \bar{x}_i \langle x_j \rangle + \bar{x}_i \bar{x}_j \\
&= \langle x_i x_j \rangle - \bar{x}_i \bar{x}_j - \bar{x}_i \bar{x}_j + \bar{x}_i \bar{x}_j \\
\mathbf{C}_{ij} &= \langle x_i x_j \rangle - \bar{x}^2 \qquad \text{(using A.4)} \\
\langle x_i x_j \rangle &= \mathbf{C}_{ij} + \bar{x}^2
\end{aligned}$$

This value of $\langle x_i x_j \rangle$ can be substituted into equation A.5:

$$\begin{aligned}
\langle \Delta w_i \rangle &= \alpha \sum_j w_j (\mathbf{C}_{ij} + \bar{x}^2) - \alpha_0 \bar{x} - \alpha x_0 \bar{x} \sum_j w_j + \alpha_0 x_0 \\
&= \alpha \sum_j w_j \mathbf{C}_{ij} + \alpha \sum_j w_j \bar{x}^2 - \alpha_0 \bar{x} - \alpha x_0 \bar{x} \sum_j w_j + \alpha_0 x_0 \\
&= \alpha \sum_j w_j \mathbf{C}_{ij} + \sum_j w_j (\bar{x}^2 \alpha - \alpha x_0 \bar{x}) - \alpha_0 \bar{x} + \alpha_0 x_0 \\
&= \alpha \sum_j w_j \mathbf{C}_{ij} + \bar{x} \alpha \sum_j w_j (\bar{x} - x_0) - \alpha_0 (\bar{x} - x_0)
\end{aligned}$$

Assuming that the average input activity $\bar{x}$ is equal to x_0, $\langle \Delta w_i \rangle$ reduces to:

$$\langle \Delta w_i \rangle = \alpha \sum_j w_j \mathbf{C}_{ij}$$

$$\text{or } \frac{\mathbf{w}}{\tau} = \mathbf{C}\mathbf{w}, \mathbf{w} = (w_1, w_2, \ldots w_N)$$

Hence, given the two assumptions that weight changes occur on a slower timescale than presentation of inputs and that the average activation of all input units is x_0, the covariance rule reduces to $\frac{\mathbf{w}}{\tau} = \mathbf{C}\mathbf{w}$. This form of the covariance rule is much simpler because it relies only upon presynaptic, and not postsynaptic, activity levels.

A.2 Why eigenvectors dominate development in correlational-based modification rules

Modification rules of the form $\frac{\mathbf{w}}{\tau} = \mathbf{C}\mathbf{w}$ are often analysed by eigenvector analysis. Let us assume that the eigenvectors of **C** are $\mathbf{e}_i$ with eigenvalues λ_i. If C is real and symmetric, there are N real

eigenvectors. Writing **w** in terms of the eigenvectors of **C**:

$$\mathbf{w} = \sum \mathbf{e} \qquad \text{where} \qquad = \mathbf{e} \cdot \mathbf{w} \tag{A.6}$$

$$\underline{\mathbf{w}} = \mathbf{C}\mathbf{w} = \mathbf{C}\sum \mathbf{e} \tag{A.7}$$

$$= \mathbf{C}\mathbf{e}_1 \; _1 + \mathbf{C}\mathbf{e}_2 \; _2 + \cdots + \mathbf{C}\mathbf{e} \tag{A.8}$$

$$\underline{\mathbf{w}} = \lambda_1 \; _1\mathbf{e}_1 + \lambda_2 \; _2\mathbf{e}_2 + \cdots + \lambda \quad \mathbf{e} \tag{A.9}$$

Therefore the rate of growth of the weight vector in the direction of each eigenvector is determined by its eigenvalue. Any component of the weight vector in the direction of an eigenvector with negative eigenvalue is quickly removed. Components of the weight vector in the direction of an eigenvector with positive eigenvalue grows exponentially, with the eigenvector with highest eigenvalue quickly dominating development. Assuming that the maximum eigenvector dominates weight development before any constraint limits are met (su