# MODELING ACTIVITY-DEPENDENT DEVELOPMENT IN THE RETINOGENICULATE PROJECTION

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PSYCHOLOGY
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Gary Lawrence Haith
June 1998

© Copyright 1998 by Gary Lawrence Haith
All Rights Reserved

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Dr. David Heeger
(Principal Adviser)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Dr. Brian Wandell

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Dr. Kenneth Miller

Approved for the University Committee on Graduate Studies:

# Abstract

In higher mammals, the primary visual pathway starts with the ("retinogeniculate") projection from the retina to the dorsal lateral geniculate nucleus (dLGN) of the thalamus, which in turn projects to visual cortex. Although the retinal axons initially innervate the dLGN in a relatively disorganized manner, they are precisely arranged by maturity. Some dominant features of this organization emerge only under the influence of activity, yet these features are established before eye-opening or photoreceptor function. The crucial activity is supplied by spontaneous bursts of action potentials that propagate in waves across the immature retinal ganglion cell layer that projects to the dLGN. Under the influence of retinal activity, the retinal axons segregate into eye-specific layers, on/off sublayers, and precise retinotopic maps.

This dissertation describes a formal computational framework for modeling and exploring the activity-dependent development of the retinogeniculate projection. The model is the first to support the development of layers, sublayers, and retinotopy in a unified framework. The model is constructed so as to be directly biologically interpretable and predictive. It refines based on realistic patterns of wave activity, retinal axon arbor change, and Hebbian synaptic weight change. In addition, the model is relatively tractable to formal analysis. This tractability makes the model relatively undemanding to simulate computationally and provides analytic insight into the dynamics of the model refinement. Several experimental predictions that follow directly from the model are described.

# Acknowledgments

It's been a long road, and I've made it with more than a little help from my colleagues and friends. Good advice has helped a lot. Thank you to David Heeger for keeping me sane and on track. David's excellent advice, unflagging confidence in the project, intellectual rigor, and clarity of thought have all greatly contributed to my graduate career and my progress on this project. Thank you to Ken Miller for generously agreeing to be on my orals and reading committee despite being overburdened and being in a different academic community. Many parts of this project would not have been possible without Ken's expertise, guidance and patience. Thank you to Brian Wandell for being extremely generous with resources and time over the last 3 years. This project would be much more limited without a computer and without Brian's willingness to be on my orals and reading committee. Thank you to Bill Newsome for making the timely completion of my orals possible. Thank you to Ellen Markman for introducing me to science and scientific writing and for looking out for me, I appreciate it. Thank you to David Rumelhart for intellectual inspiration, being on my orals committee and for darts. Thank you to Roger Shepard for providing me with the intellectual seed of this project.

Those people on the data collection side of things have also been generous with their time. I would especially like to thank Rachel Wong for going beyond the call of duty and making my life much easier. I would also like to thank Anna Penn, Marla Feller, Dan Butts, and Carla Shatz for their time and answers.

The grad students and post-docs in the vicinity of my office know all too well how much they've helped me. Thanks to Sean for frisbee, countless questions and answers, great conversations and turning me on to some neat ideas. Thanks to Rehan for explaining the simplest concepts to me; Geoff for helping me sort out and understand MarkerSize and repmat among other things; Alex Huk for setting me up with good music, helping me think things through and high hilarity; Jon Demb for listening for me talk to myself and him; and Jennifer Smith for some crucial question answering and technical expertise.

I have also realized just how much I depend on friends and family during this project. Thank you to my friends for patience, support, and helping me keep my perspective. Thank you to Kirsten for the million or so big and little things that she did to support me, including proof-reading. Thank you to my father for planting so many fruitful seeds in my life. Thank you to my mother for saying "just finish it" and other good advice.

I would also like to thank Stephen Eglen for pointing me to his dissertation work. His work has been a valuable resource and has greatly influenced this project.

# Contents

**Abstract** **iv**

**Acknowledgments** **v**

**1 Introduction** **1**

1.1 Adaptation and Exploitation . . . . . 1

1.2 Why The Retinogeniculate Projection? . . . . . 2

1.3 The Model Domain . . . . . 3

1.4 Biological Realism . . . . . 3

1.4.1 Realistic Inputs . . . . . 3

1.4.2 Sprouting and Lateral Interactions . . . . . 4

1.4.3 Details Matter . . . . . 4

1.5 Thesis Organization . . . . . 5

**2 Biological Background: Maturity** **6**

2.1 The Eye . . . . . 6

2.2 Afferent Projections . . . . . 7

2.3 The dLGN . . . . . 7

2.3.1 Laminar Structure . . . . . 8

2.3.2 Retinotopic Maps . . . . . 9

2.3.3 dLGN relay cells: Expansion of the X and Y pathways . . . . . 9

2.3.4 Extra-retinal and Intra-dLGN Influences . . . . . 10

2.3.5 Physiology: GABA subtypes . . . . . 12

2.3.6 dLGN Response Modes: Tonic, Bursting and Spindle Waves . . . . . 13

**3 Biological Background: Development** **15**
3.1 Problems to be Solved During Development . . . . . 15
3.1.1 Possible Mechanisms of Retinotopy . . . . . 17
3.1.2 Possible Mechanisms of Layer Segregation . . . . . 19
3.2 Developmental Trends in the Eye and Retina . . . . . 19
3.2.1 Increasing Acuity: Anatomical Issues . . . . . 20
3.2.2 Processing and Relay Circuitry before Transduction . . . . . 20
3.2.3 Center to Periphery . . . . . 21
3.2.4 X, Y and W Pathways . . . . . 22
3.3 Retinogeniculate Projection . . . . . 22
3.3.1 Anatomical Background: Milestones and dLGN Tissues . . . . . 23
3.3.2 Arrival of RGC axons at the dLGN . . . . . 23
3.3.3 Axonal and Dendritic Arborization in the dLGN . . . . . 23
3.3.4 Time-Course of X and Y Arborization . . . . . 24
3.3.5 Retinotopy . . . . . 25
3.3.6 Laminar Organization . . . . . 26
3.3.7 Activity Dependence of Eye-Specific Layer Segregation . . . . . 28
3.3.8 Activity Dependence of On/Off Sublayer Segregation . . . . . 32
3.3.9 Maps and Layers: Shifting . . . . . 32
3.4 Physiological Development . . . . . 33
3.4.1 Immature dLGN Physiology . . . . . 33
3.4.2 LTP . . . . . 35
3.4.3 RGC Spontaneous Activity . . . . . 35
3.5 Summary of Relevant Biological Background . . . . . 38

**4 The Model** **39**
4.1 Overview . . . . . 39
4.2 Brief Summary . . . . . 42
4.3 Detailed Description . . . . . 43
4.3.1 Activity in the Model . . . . . 43
4.3.2 Inputs . . . . . 44

4.3.3 Initialization . . . . . 49
4.3.4 dLGN Update . . . . . 50
4.3.5 Hebbian Weight Update . . . . . 53
4.3.6 Synaptic Sprouting . . . . . 53
4.3.7 Weight Normalization . . . . . 54

**5 Model Results** **57**
5.1 Format of Results . . . . . 59
5.1.1 1-Dimensional Result Display . . . . . 59
5.1.2 2-Dimensional Result Display . . . . . 62
5.1.3 Speed Issues . . . . . 62
5.1.4 Scale Issues . . . . . 64
5.2 Organization of Results . . . . . 65
5.3 Canonical Model . . . . . 65
5.4 Longer Time-course . . . . . 66
5.5 Multiple Domains . . . . . 67
5.6 Blurring and Sprouting in the dLGN . . . . . 69
5.6.1 Blurring . . . . . 69
5.6.2 Sprouting . . . . . 70
5.7 No On/Off Activity . . . . . 70
5.8 Biases . . . . . 70
5.8.1 Alternate Types of Bias . . . . . 72
5.9 Asymmetries in Sprouting and Lateral Interactions . . . . . 72
5.9.1 No Asymmetries . . . . . 72
5.9.2 Static Asymmetries . . . . . 72
5.9.3 No Inter-Layer/Sublayer Interactions . . . . . 73
5.10 Pre- and Post-Synaptic Normalization . . . . . 75
5.10.1 No Pre-Synaptic Normalization . . . . . 75
5.10.2 No Post-Synaptic Normalization . . . . . 75
5.10.3 Capped Normalization and Monocular Deprivation . . . . . 76
5.11 Summary . . . . . 78

**6 Formal Analysis** **79**
6.1 Analytic Difficulties . . . 79
6.2 Alternate Formulation . . . 80
6.2.1 Simulation Results . . . 82
6.3 Linear Analysis . . . 83
6.3.1 Weight Matrix Eigenmodes . . . 83
6.4 Summary . . . 85

**7 Previous Models** **87**
7.1 General . . . 87
7.1.1 Willshaw and von der Malsburg . . . 87
7.1.2 Kohonen . . . 88
7.2 LGN Specific Models . . . 89
7.2.1 Keesing and Stork . . . 89
7.2.2 Eglen . . . 90
7.2.3 Lee and Wong . . . 96

**8 Discussion** **98**
8.1 Biological Value of the Model . . . 98
8.2 Activity-Dependent Developmental Strategies . . . 99
8.2.1 Constrained Experience . . . 99
8.2.2 Self-Induced Experience . . . 100
8.3 Unaddressed Issues . . . 101
8.3.1 Feedback Circuits . . . 101
8.3.2 Activity-Dependent Intra-dLGN Development . . . 101
8.3.3 Hebbian Synaptic Change and LTP . . . 101
8.3.4 Activity-Independent Factors . . . 102
8.3.5 Visual Experience . . . 102
8.3.6 Scale Issues . . . 102
8.4 Proposed experiments . . . 103
8.4.1 Point Spread . . . 103
8.4.2 Relay Cell Physiology . . . 103

8.4.3 Differentiation Between Off-Cells and On-Cells . . . . . 103
8.5 Conclusion . . . . . 104

**A LGN Update and Regularization** **105**

**B LGN Update and the Membrane Equation** **107**
B.1 Derivation . . . . . 109

**C Eigenmode Analysis** **111**
C.1 Translation into Matrix Notation . . . . . 111
C.2 Correlation Formalism . . . . . 113
C.2.1 Normalized Eigenmodes . . . . . 113

**D The Covariance-Rule and Sparse Inputs** **115**

**Bibliography** **117**

# List of Tables

3.1 Developmental Milestones . . . . . 16

4.1 Wave Parameters . . . . . 44
4.2 dLGN Update Symbol Table . . . . . 50

5.1 Canonical Model Parameters . . . . . 58
5.2 Results Table . . . . . 59

B.1 Symbol Table . . . . . 108

C.1 Symbol Table . . . . . 112

D.1 Example: Silence and Negative Correlations . . . . . 116

# List of Figures


4.1 Model Schematic . . . . . 40
4.2 Model System . . . . . 41
4.3 Wave Illustration . . . . . 45
4.4 Correlation Index Plots . . . . . 46
4.5 On and Off-RGC Correlations: D56-63 . . . . . 48
4.6 dLGN Response over Development . . . . . 51
4.7 Inter-layer and Sublayer Interactions . . . . . 52

5.1 Matrix Tableau Explanation . . . . . 60
5.2 1-Dimensional Result Display Explanation . . . . . 61
5.3 2-Dimensional Result Display Explanation . . . . . 63
5.4 Result: 2-D Canonical Model . . . . . 65
5.5 Result: 1-D Canonical Model . . . . . 66
5.6 Result: Realistic Time-Course . . . . . 66
5.7 Input Correlation Structure, Multiple Domains . . . . . 68
5.8 Result: Multiple Domains . . . . . 68
5.9 Result: No Sprouting . . . . . 69
5.10 Result: No Intra-dLGN Interactions . . . . . 70
5.11 Result: No Noisy Off-RGC Firing . . . . . 71
5.12 Result: No Biases . . . . . 71
5.13 Result: No Asymmetries . . . . . 73
5.14 Result: Static Asymmetries . . . . . 73
5.15 Result: Static Asymmetries, Inter-Layer $\rightarrow 0$ . . . . . 74
5.16 Result: No Inter-Layer/Sublayer Interactions . . . . . 74

5.17 Result: No Pre-Synaptic Normalization . . . . . 75
5.18 Result: No Post-Synaptic Normalization . . . . . 76
5.19 Results: Capped Normalization . . . . . 77
5.20 Result: Capped Normalization and Monocular Deprivation . . . . . 78

6.1 Result: Subtractive Model . . . . . 82
6.2 Eigenmodes During Eye-specific Layer Segregation . . . . . 84
6.3 Eigenmodes: During On/Off Sublayer Segregation . . . . . 85

7.1 The Covariance Rule and Post-Synaptic Normalization . . . . . 94
7.2 The Covariance Rule and Pre-Synaptic Normalization . . . . . 95

# Chapter 1

# Introduction

It is known that neural activity can change brain structure. The mechanisms involved have been explored at several levels of analysis and scale: on the scale of the entire organism, it has been found that rats reared in complex environments have more extensive synaptic connectivity in some brain areas than rats reared in sparse environments (Greenough et al., 1987); on the scale of individual neurons, synaptic connectivity has been found to change in response to specific programs of stimulation (LTP and LTD, Madison et al., 1991); and on the scale of neural systems, there is evidence that projections between brain areas can be shaped by patterns of neural activity (Goodman and Shatz, 1993).

This thesis attempts to formalize the dynamics of activity-dependent neural change in a well-studied and understood system, the retinogeniculate projection. The current model is the first to capture the activity-dependent development of the main retinogeniculate structures (eye-specific layers, on/off sublayers, and retinotopic refinement) in a unified framework. The model is structured so as to be biologically realistic and analytically tractable so as to be maximally interpretable and predictive. I've chosen to model the retinogeniculate projection because it provides a wonderful window into some complex developmental strategies that organisms have evolved to take advantage of neural activity.

## 1.1 Adaptation and Exploitation

Mechanisms of activity-dependent neural change are valuable to an organism. Most obviously, they provide a means for adaptation or learning. Through experience, the particulars of the environment can generate patterns of neural activity that shape the processing structure of the organism, adapting the organism's behavior patterns to environmental pressures. Sensitivity to experience also has a less obvious effect. It allows developmental processes to exploit the structure inherent in experience in order to "cheaply" create useful neural structures. Whereas adaptation allows neural structures to take on novel configurations in response to novel experience, exploitation allows

neural structures to take on stereotypical configurations in response to stereotypical aspects of the environment.[1]

One would expect the exploitation of experience to occur if a useful neural structure robustly develops from the interaction between mechanisms of activity-dependent neural change and experiences caused by reliable features of the environment. In this case, it is wasteful to spend resources on building a useful structure instead of exploiting the environment, assuming that the organism already has mechanisms that support activity-dependent neural change. What if a specific set of experiences (in combination with mechanisms of activity-dependent neural change) can give rise to a useful structure, but the needed experiences aren't reliably provided by the environment? A counter-intuitive option is that the organism can "simulate" the experience by intrinsically generating the needed activity — a viable option if generating the activity is relatively "cheap" (in terms of energy and genetic information required) compared to generating the resulting structure. Such a situation exists in the development of visual circuits and probably in many other neural structures.

## 1.2 Why The Retinogeniculate Projection?

The crucial question at the onset of this thesis is: "Why study the development of the retinogeniculate projection of ferrets and cats?"

My primary answer is that the system is surprising. It generates its own activity, activity that is sensible in the framework of vision (e.g. spatio-temporally correlated). As a result, the system is able to undergo significant activity-dependent refinement *before* visually evoked retinal activity is present. As with visual experience, the structural development adaptively responds to alterations in the normal course of this intrinsically generated activity. For example, if this activity is blocked in one eye, the projection from the blocked eye will decay and the spared eye's projection will take over more territory.

This pre-visual period of retinogeniculate development thus puts a strange and significant twist on the relation between intrinsic and extrinsic influences ("nature and nurture"). In this system, intrinsic factors express themselves via a putatively extrinsic pathway by inducing activity at the sensory surface. Despite the flexibility that is inherent in this external experience-dependent pathway, a stereotyped path of development results. Importantly, this stereotyped structure is not entirely determined by intrinsic constraints on the final form of the projection. Rather, the form of the intrinsically generated activity plays a crucial role in shaping the projection's development and mature form. In this system, the elegant mechanisms that allow the organism to learn and adapt to its environment seem to have been co-opted for the development of a stereotyped system by providing a stereotyped activity environment. The retinogeniculate system thus provides an opportunity to see a complex interaction between intrinsic and extrinsic factors in shaping a complex neural circuit.

[1] Adaptation and exploitation correspond roughly to "experience-dependent" and "experience-expectant" structure in Greenough et al.'s (1987) terminology.

Fortunately, this system also affords concrete investigations into the balance between the intrinsic and extrinsic factors because it is a well-studied and described system with a striking developmental trajectory.

## 1.3 The Model Domain

The model system is the retinogeniculate projection of ferrets and cats. These animals have a similar developmental time-course and adult retinogeniculate projection, but ferrets are born $\sim$3 weeks earlier than cats ($\sim$1 month before eye-opening). Because they are born so immature, ferrets are ideal experimental animals for investigating very early ("pre-natal") development.

Their retinogeniculate projection consists of the axons of retinal ganglion cells (RGCs) that travel through the optic nerve and the optic tract into the thalamus. They arborize in the dorsal lateral geniculate nucleus (dLGN) of the thalamus and establish synapses with thalamo-cortical/relay cells as well as interneurons there. In the adult ferret, the axons from the two eyes and the different cell types arborize in separate layers in the dLGN. In addition, neighboring RGCs establish strong synaptic connectivity with neighboring dLGN relay cells, maintaining a precise retinotopic map in the dLGN. This mature state arises from a much more disordered initial state. At birth, each RGC axon connects to dLGN cells across layers. The axons segregate into different layers before eye-opening. This segregation is activity-dependent, and is supported by spontaneous retinal activity. During the period of refinement, the RGCs fire spontaneous bursts of action potentials with a spatial propagation pattern that resembles waves of bursting activity spreading across the retina.

## 1.4 Biological Realism

Because the developing retinogeniculate projection is well-studied, the model described in this thesis can be relatively grounded in biological detail. Hopefully, the biological realism will make the model maximally interpretable and predictive. In order to make the approach of the model more concrete, some features of the model are sketched below.

### *1.4.1 Realistic Inputs*

Because the rich spatio-temporal structure of the retinal activity may play a role in the development of the retinogeniculate projection, the simulated inputs are closely matched to in vitro measurements of the retinal activity during this period. The spontaneous retinal waves are simulated as if each "unit" in the model corresponds to a neuron. Measurements of wave speed, width, height (i.e. peak firing rates), and refractory periods are directly translated into the simulated wave inputs (see Table4.1). Incorporating these details helps to constrain parameter choices.

Because of this explicit simulation of the wave speed and refractory period length, there is an

intrinsic time in the model that corresponds directly to biological time. Roughly the same number of waves will occur in 1 week of intrinsic model time as would occur in 1 week of biological time (see Sec. 5.1.3 for discussion of this point in relation to simulation details).

### *1.4.2 Sprouting and Lateral Interactions*

The model also attempts to make realistic assumptions about the mechanisms of activity spread and synaptic weight change in the dLGN. Traditional models of activity-dependent development make two assumptions that are inaccurate in the retinogeniculate system. The first is that there is an initial period of exuberant connectivity, in which each input cell (RGC) connects to all the output cells. Refinement consists of the inappropriate synapses dying off and the appropriate synapses being preserved. In fact, RGCs maintain a relatively constant and restricted axonal arbor width in the dLGN over the course of development. This fact is somewhat problematic because there is reason to believe that the arbors must shift over development (see Sec. 3.3.5). The current model resolves this contradiction by assuming that the axonal arbor width is in a state of dynamic equilibrium during this period, with sprouting processes at work concurrently with culling processes (see Sec 7.2 for other models that use sprouting).

In order to ensure global organization, models traditionally make a second assumption that lateral interactions in the output tissue (the dLGN in this case) are excitatory at short range and inhibitory at long range. The excitatory portion is usually taken to start off large — shrinking gradually over the course of development. A literal interpretation of this assumption is problematic in the dLGN because inhibitory pathways are essentially non-functional in the developing dLGN, and there is no evidence for long-range excitatory connections at any point in development. The current model makes the weaker assumption of nearest neighbor excitatory interactions, and incorporates observed biophysical changes in dLGN relay cell responsivity in order to generate effective long-range interactions in the dLGN that shrink over development.

### *1.4.3 Details Matter*

Given that spontaneous activity plays a role in the development of the retinogeniculate projection, a natural question is: “How?”. This question is sometimes glossed over in experimental neurobiology. The assumption being that intuition and demonstrated principles of self-organization are sufficient to understand the role of activity in a given system.

In fact, it is non-trivial to make a self-organizing model capture the observed structural refinement in a specific system, given the anatomical and physiological constraints in that system. Presumably, in a biological system, there is a hierarchy of mechanisms that are often redundant or balanced against each other. This entire hierarchy conspires to make the refinement robust to some insults, thus insuring the proper development of key structures, while remaining sensitive to other influences so as to make development adaptable to unexpected circumstances. Understanding and modeling such a hierarchy is complex.

The original models of self-organization suggest that with appropriate spatio-temporal correlations in the inputs and appropriate lateral interactions in the outputs, a system should support self-organization. The catch is that the details matter. There are plenty of intuitively plausible models of a given system that can not support development of the observed structures. In fact, after experimentation with models, one often ends up wondering if the given mechanisms are capable of developing the given structures at all. What is obvious in principle can be extremely difficult to rigorously demonstrate in a particular instance. For example, why do layers segregate nearly perfectly? Why doesn't the retinogeniculate projection end up with fractured or intermingled layers? In fact, many models can end up with this type of end state, including the present model in some parameter domains. The balance between the different aspects of the model is crucial to generate robust and appropriate development.

On the flip side, there are many intuitively plausible mechanisms that do support appropriate development, but are simply not present in the actual system. For example, long-range inhibitory interactions in the developing dLGN would help support the development of retinotopy and layer segregation, but they are probably too weak to be useful at the relevant points in development. Likewise, negative correlations between the eyes or the on/off-center RGCs would probably help support layer and sublayer segregation in the dLGN, but these negative correlations are probably not present. The goal of this work is to outline a model that works and is in accordance with the biological observations. The hope is to use the model to gain insight into the crucial features of the system, and their inter-relations.

## 1.5 Thesis Organization

This thesis will first present a detailed biological background of the early visual system at maturity and through development (Chs. 2-3). These chapters cover many biological details that are peripheral to the model, but which are included because they could potentially influence retinogeniculate development. A brief summary of the directly relevant biological background is presented in Sec. 3.5 at the end of Ch. 3 for those readers who wish to skip the detailed background. The thesis then goes on to describe the model in depth (Ch. 4). Simulations that explore the convergence behavior of the model as well as the importance and function of the individual model mechanisms are presented in Ch. 5. A formal linear analysis of the model dynamics is presented in Ch. 6. Ch. 7 is a brief review of previous models which concentrates on their relations to the current model. Finally, Ch. 8 is a general discussion with suggested experiments.

# Chapter 2

# Biological Background: Maturity

There are a few key structures in the early visual system that are central to this discussion. A picture of the mature state of these structures will provide a useful framework for the outstanding questions about their development.

This description is ordered from peripheral to more central structures in the early visual system. Thus it will proceed from the sensory surface, through the retinal outputs — retinal ganglion cells (RGCs), and eventually to their primary target —the dorsal lateral geniculate nucleus (dLGN) of the thalamus. The overriding focus will be the organization of the cells, afferents and intrinsic circuits in the dLGN.

## 2.1 The Eye

The mature retina has a laminar structure. The receptors (rods and cones) are farthest from the back of the eye ("inner"), while the retinal ganglion cells and their axons are closest to the back of the eye ("outer"). The bipolar cell bodies form a layer between the receptors and the RGCs. The bipolar cells have their dendrites in the outer plexiform layer and their axons in the inner plexiform layer. In each of these layers are feed-forward ribbon synapses. In addition, there are laterally modulatory "surround" circuits formed by the horizontal cells in the outer plexiform layer and amacrine cells in the inner plexiform layer. These lateral circuits are based on conventional synapses (Daw, 1995; Chalupa and White, 1990; Sterling, 1990).

The output cells of the retina, retinal ganglion cells (RGCs), are traditionally classified morphologically as $\alpha$, $\beta$, or $\gamma$ cells (corresponding to the Y, X, or W physiological classifications respectively). $\alpha$, $\beta$ and $\gamma$ cells account for 5%, 80%, and 15% of the RGCs respectively (Sterling, 1990). The $\alpha$ cells have a large dendritic tree, soma and axon diameter, and thus are fast with low spatial acuity. Because the $\beta$ cells have a smaller dendritic tree, soma, and axon diameter, they have better spatial acuity and a slower time-course. The $\gamma$ cells (projecting primarily to the superior

colliculus (SC)) have smaller somas and axons and have a fine and wide dendritic tree, and they have a 'sluggish' time-course and low spatial acuity (see Sterling, 1990 for review).

There are also on and off subtypes of both $\alpha$ and $\beta$ cells. Cells in these subtypes respond either to a onset ('on-center') or offset ('off-center') of a stimulus in the center of their receptive field. The on and off-center RGCs can be differentiated based on the location of their dendritic arbors within the inner-plexiform layer (IPL) of the retina. The dendrites of the on-center cells stratify in the outermost portion of the IPL while the dendrites of the off-center cells stratify in the innermost portion of the IPL. Ignoring the W- cells, the mature cat retina is effectively tiled 4 times, once each by the dendrites of the on and off $\alpha$ and the on and off $\beta$ cells (Chalupa and White, 1990). These subtypes make up parallel pathways that remain largely separate — even up to primary visual cortex.

## 2.2 Afferent Projections

The adult retinas in the cat and ferret have two main sub-cortical visual projections: the superior colliculus (SC, termed the 'tectum' in lower animals) and the dorsal lateral geniculate nucleus (dLGN). The SC projects to many areas including the dLGN and higher visual areas in the cortex. The dLGN provides most of the input to the primary visual cortex. The primary visual cortex has reciprocal (descending) connections to both the SC and the dLGN. The projection from the retina to the dLGN (the retinogeniculate system) is part of the primary visual system in higher animals and will be the focus of this paper.

As in other mammals, there is partial decussation of RGC axons in the ferret (Jeffery, 1990; Cucchiaro and Guillery, 1984) and cat (Shatz, 1983). The contralateral projection is much larger ($\sim$8 times in ferret, Zahs and Stryker, 1985), and develops earlier than the ipsilateral projection. Albino animals have an amplified pattern of decussation with almost all of the RGC axons crossing at the chiasm (Cucchiaro and Guillery, 1984).

## 2.3 The dLGN

The dLGN is the primary target of the retinal projection in the ferret and cat (and other higher mammals). It serves as a "gating station" for the retinal signals bound for primary visual cortex and the higher visual areas. Though the dLGN has long been assumed to passively relay the retinal signal to cortical targets, the predominance of non-retinal synapses and complex synaptic structures involving interneurons suggest that it may perform some important functions in modulating the retinal signal.

### 2.3.1 Laminar Structure

The mature cat (Sherman, 1985) and ferret (Cucchiaro and Guillery, 1984) have a typical ”carnivore” geniculate organization. The dLGN is organized into 3 main layers: A (farthest from the optic tract, dorsal), A1 and the C group (closest to the optic tract, ventral). The C group of laminae is further subdivided into C, C1, C2, and C3. The afferents from a given eye innervate roughly every other layer, (Shatz, 1983) with almost no overlap at full maturity. Thus, the RGC axons from the temporal portion of the ipsilateral retina innervate A1 and C1 while the nasal portion of the contralateral eye innervates A, C, and C2. C3 is innervated by the midbrain (Sherman and Koch, 1990).[1]

The shape and laminar structure of the cat and ferret dLGN are somewhat different. Whereas all the layers in the cat dLGN span the entire length of the nucleus (see figure 8.3 in (Sherman and Koch, 1990)), only layer A spans all of the nucleus in the ferret (see Fig. 4.1, and see Figs. 3 and 4 in (Linden et al., 1981)). This laminar structure coincides with a more L-shaped dLGN in the ferret. The upper portion of the L (more rostromedial) is innervated only by the contralateral retina and the lower portion of the L (more caudolateral) is innervated in layers by both retinas (Linden et al., 1981). Same-eye “Layers” in the ferret also appear less distinct than in the cat, i.e. layers A, C and C2 in the ferret are connected by a ’bridge’ at the most medial portion of the nucleus (compare figures 3 & 4 in (Linden et al., 1981) to figure 2 in (Sherman and Koch, 1990)).

In both cats and ferrets, the axons of different functional subtypes of RGCs (i.e. X, Y, and W cells or on and off cells) also tend to focally arborize in some layers of the dLGN while only traversing or simply avoiding other layers. X and Y axons innervate the A laminae. In cats, Y axons innervate the most dorsal portion of the C lamina, while the W cells innervate the the ventral portion of C, and all of C1-C3 (see fig. 8.3 (Sherman and Koch, 1990)). In ferrets, the A laminae are further divided into on and off sublayers. dLGN cells with on-center receptive fields and cells with off-center receptive fields (RFs) are primary located in the inner and outer portion of the A laminae respectively (Linden et al., 1981; Shatz, 1983; Casagrande and Condo, 1988). These sub-laminae mainly reflect differential innervation by the X RGCs, which have relatively homogeneous and restricted arborization patterns relative to the Y RGCs (Roe et al., 1989).

In the ferret, the segregation of axons is two-fold with eye-specific and functional-class specific layers forming. The most robust, early and marked segregation is between the afferents from the two eyes. That is, the RGC axons from the each eye segregate over the course of development such that in the adult animal they arborize in roughly alternating layers. Axons cross the layers innervated by the other eye (if necessary) with a minimal axonal shaft, rarely retaining any side branches, terminal boutons, or synapses in the inappropriate layers (Shatz, 1983; Sretavan and Shatz, 1986; Sretavan

[1] Due to considerations of space and complexity, I am choosing to ignore the medial interlaminar nucleus (MIN). The MIN lies just medial to the dLGN, is innervated completely by Y-cells and has a laminar structure exactly parallel to dLGN, making the two appear nearly continuous (at least in schematic form — see figure 8.3 (Sherman and Koch, 1990)). In my preliminary investigations the anatomical separation of the MIN belies the lack of idiosyncratic information about its structure and function. I thus consider it here to be largely a physical extension of the dLGN.

and Shatz, 1987). The result of this selective arborization of the different functional subclasses of RGCs is that almost every relay cell in the dLGN receives input from only one kind of cell (X/Y and on/off-center) from one of the two retinas. Only 10% of cells receive mixed X and Y input. Very rarely do cells receive binocular input (and these are all located at laminar borders or in interlaminar zones), and no cells receive input from both on and off cells (Archer et al., 1982, and see (Mastronarde, 1992) for details of cells with mixed input).

### *2.3.2 Retinotopic Maps*

The dLGN, SC & primary visual cortex all receive visual projections in a "retinotopic map". That is, nearby cells in the tissues have receptive fields centered on nearby points in space. Furthermore, the mature retinal projection to the dLGN is "sharp" in that each cell in the recipient tissue receives input from one or a small number of RGCs (Mastronarde, 1992; Sherman and Koch, 1990). A measure of the degree of retinotopy in the maps is the degree of scatter in the receptive fields of dLGN cells as one moves laterally across the tissue (i.e. using an electrode track technique). The scatter in the position of the receptive fields is only $.5^\circ$ of visual angle. Comparable measures find 10-$30^\circ$ of scatter in SC and 20-$30^\circ$ of scatter in the optic tract (Voigt et al., 1983; Sanderson, 1971).

The dLGN in the ferret receives input from the central and nasal part of the contralateral retina and the most temporal part of the ipsilateral retina. Each of these projections form a continuous topographic map of their portion of the retina. These adult maps, although in different layers, are in register such that a vertical track through the dLGN represents a single point in binocular visual space. Because there is limited binocular overlap in ferrets (due to the lateral placement of the eyes), the alignment of RFs in the maps means that there is only a limited extent to which the maps in the different layers overlap — namely in the most most temporal portion of the ipsilateral retina, which receives input from the same portion of the visual field as the central portion of the contralateral retina (see figure 1 in Jeffery, 1990).

The density of RGCs in the central retina (area centralis) is approximately 80 times greater than the density in the peripheral retina (Chalupa and White, 1990). This concentration of resources in the center of the visual field is preserved in the projections to the dLGN, SC and primary visual cortex where the area centralis has a largely exaggerated representation, i.e. commands many cells, compared to an equal area of retinal tissue in the periphery.

### *2.3.3 dLGN relay cells: Expansion of the X and Y pathways*

dLGN relay cells are the output cells of the dLGN, eventually projecting to cortical targets. dLGN relay cell classification is usually performed physiologically (thus their designation as X and Y cells versus $\alpha$ and $\beta$ cells), but they are differentiated morphologically as well. Their morphology mirrors the morphology of their inputs in the retina, that is X cells are smaller and have thinner dendrites and axons. Additionally, the X- cells are oriented along projection lines (i.e. perpendicular to laminar boundaries) and have many dendritic appendages (sites of synaptic glomeruli) whereas

the Y- cells are more circular and have smoother dendrites (Sherman and Koch, 1990). These characteristics reflect the patterns of physiology and connectivity in these two functional pathways with X-cells responding more slowly and synapsing in glomeruli.

dLGN relay cells eventually receive input from one or a few RGCs (almost always from the same functional class), and most can be considered to inherit their physiological properties directly from their inputs (Mastronarde, 1992). On the other hand, each RGC connects to many dLGN cells — thus amplifying the number of neurons in the pathway by $\sim 4\times$. This fan-out is far more pronounced in the Y pathway than in the X pathway. While X RGCs outnumber Y RGCs by more than 10:1, the percentages of X and Y relay cells in the dLGN are approximately equal (i.e. 45% each, W cells being the remaining 10%, but see Sherman and Koch, 1990 for 2:1 X:Y figure). This relative growth of the Y pathway is due to the extensive fan out from each Y RGC cell to 30-50 dLGN relay cells — as compared to 4-5 for each X RGC. This fan out is reflected in their relative arbor widths and bouton numbers: X arbors are 100 $\mu m$ wide with 500-1000 boutons each, while Y arbors are 300 $\mu m$ wide with 800-2000 boutons (Sur, 1988). At maturity the Y pathway is much more amplified than the X pathway.

In the X subclass, RGCs connect primarily to cells with soma within their afferent arbors, rather than to all cells with dendrites overlapping their afferents (Sherman and Koch, 1990). Further, RGCs connect to only a few of the many functionally appropriate dLGN cells within their arbor. Therefore, the selective connectivity observed from RGCs to dLGN cells cannot derive from purely mechanical processes, rather it must also rely on mechanisms of selective synapse formation and stabilization (Hamos et al., 1987).

### *2.3.4 Extra-retinal and Intra-dLGN Influences*

In addition to the feed-forward connections from the retinal afferents there are also lateral and descending connections to the dLGN that modulate the incoming retinal signal. In fact, approximately 90% of the synapses to dLGN relay cells are non-retinal in origin. 25% of these connections are associated with inhibitory (GABAergic) interneurons in the dLGN or the midbrain. Up until the last year it was thought that the cortex supplies the dominant input to both interneurons (37%) and dLGN relay cells (58%) in the dLGN (Montero, 1991), but recent evidence indicates that as many as half of the synapses that were thought to be cortical in origin may actually originate in the brainstem (Erisir et al., 1997).

*Local Inhibition.* The retinal synapses on interneurons tend to occur at glomeruli, the axonal appendages interwoven with dendritic appendages (95%, Montero, 1991). Glomeruli are the site of many of the classic triadic synaptic structures where an afferent fiber synapses on a relay neuron and an interneuron, and the interneuron synapses on the same relay neuron. Because interneurons are very electrotonically spread (i.e. parts of the dendritic tree are electrically isolated from other parts of the arbor and the soma), they are not good candidates for wide ranging inhibition. More likely,

interneurons, and glomeruli in particular, provide a substrate for local computations (Sherman and Koch, 1990).

*Medium Range Inhibition.* There is evidence that the inhibitory surround present in RGC cell responses is accentuated by surround inhibition in the dLGN (Eysel et al., 1987; Eysel and Pape, 1987; Eysel et al., 1986; Sillito and Kemp, 1983; Singer et al., 1972). This heightened surround inhibition is thought to be responsible for the improved spatial resolving power of dLGN cells over and above the RGC inputs. This added surround inhibition could stem from the convergence of several neighboring RGCs onto individual interneurons (Norton and Godwin, 1992; Sillito and Kemp, 1983). These post-synaptic interneurons can then serve to inhibit dLGN relay cells via either dendro-dendritic or more traditional axonal connections (Bloomfield and Sherman, 1989; Pape and McCormick, 1995; Montero, 1991). The interneuron's axon terminates within its dendritic arbor (Uhlrich and Cucchiaro, 1992) and synapses on relay neurons, and possibly other interneurons (see below) (Pape and McCormick, 1995; Williams et al., 1996). This convergence could serve to mediate the inhibition of relay cell firing due to the firing of RGCs neighboring its primary input RGC. In addition, there is some convergence of interneurons onto dLGN cells (Uhlrich and Cucchiaro, 1992; Soltesz and Crunelli, 1992). This convergence could mediate the heightened surround inhibition of relay cell firing due to firing of neighboring interneurons, which presumably receive input primarily from neighboring RGCs.

The interactions between interneurons are relatively unknown. There are GABAergic terminals on the dendrites and soma of interneurons in dLGN, and, because the reticular nucleus thalami/peri-geniculate nucleus (RNT/PGN)[2] seems to selectively innervate relay neurons (Pape and McCormick, 1995), it has been suggested that these inhibitory terminals likely arise from other interneurons.

Interneurons are primarily involved in the X pathway, thus Y RGCs synapse mainly on dLGN relay cells (Sherman and Koch, 1990; Friedlander and Tootle, 1990; Uhlrich and Cucchiaro, 1992, but see Mastronarde, 1992 for evidence of Y interneurons). The frequency of these complex circuits in the X pathway suggests that it is more modulated by inhibitory circuitry than the Y pathway. Because each interneuron receives input from 4-5 RGCs, there is more RGC convergence on interneurons than on relay cells (Sillito and Kemp, 1983; Norton and Godwin, 1992).

*Long-Range Inhibition.* Typically, dLGN cells in the adult cat receive excitatory input from only one optic nerve, but inhibitory connections from both nerves. dLGN cells have purely inhibitory receptive fields for the non-dominant eye which overlap the excitatory receptive fields from the dominant eye. This inhibitory input is di or tri-synaptic, while the excitatory input is monosynaptic (Shatz and Kirkwood, 1984; Lindstrom, 1982; Pape and Eysel, 1986). It follows that the excitatory inputs to dLGN relay cells stem directly from retinal afferents, while the inhibitory inputs

[2]These structures are neighbors and appear to be quite similar in projections, structure and role (Soltesz and Crunelli, 1992), thus they are treated as a single complex for the purposes of this paper.

are communicated via dLGN interneurons or other nuclei in the midbrain, e.g. the reticular nucleus of the thalamus (Pape and Eysel, 1986).

Because of the large electrotonic size of interneurons, and the local arborization of their axons, any wide-ranging lateral inhibition in the dLGN would likely be mediated by external structures that provide input to the dLGN. The likely candidates for this role is the RNT/PGN. RNT/PGN cells receive inputs from the cortex and dLGN relay cell collaterals, and they send afferents to dLGN relay cells, each other, and dLGN interneurons (Soltesz and Crunelli, 1992). The specific role of the RNT/PGN afferents in mediating long range lateral inhibition is unknown. dLGN relay neurons can receive input from the very same cells that they synapse on in the RNT/PGN (Lo and Sherman, 1994). As RNT/PGN neurons have fairly large RF's ($\sim 20°$, Montero, 1991) this finding suggests that the inhibitory feedback from the RNT/PGN could modulate activity in a limited area around a dLGN neuron (Lo and Sherman, 1994; Eysel et al., 1986; Eysel and Pape, 1987). Others have taken the large RF size to suggest that the RNT/PGN probably has more influence on global responsivity of the dLGN than on shaping local RFs (Ahlsen et al., 1985). Because RNT/PGN afferents cells establish 90% of their synapses on dLGN relay cells in regions that also receive input from cortical axons, they probably are largely involved in modulating the cortical influence on the dLGN (Ahlsen et al., 1985; Cucchiaro et al., 1991).

This complex and organized wiring pattern in the early visual system serves to maintain parallel signal pathways which can be modulated, amplified or combined selectively. Further, via topographic projections, the wiring preserves useful information intrinsic to the spatial relations between RGCs.

### *2.3.5 Physiology: GABA subtypes*

Investigations into the function and distribution of the two GABA receptor subtypes, GABAa and GABAb, have revealed that there are two distinct inhibitory mechanisms in the dLGN. The GABAa subtype is fast acting (30-40 ms duration), works via $Cl^-$ conductance and has a reversal potential around the resting potential of dLGN relay cells (-65 mV, suggesting a "shunting" role). GABAa seems to be the dominant inhibitory pathway. GABAb is slow acting (20-30 ms to onset, 200-300 ms duration), works via $K^+$ conductance and has a hyper-polarizing reversal potential (-80 mV). GABAb probably mediates global responsivity in the dLGN (Soltesz and Crunelli, 1992; Crunelli and Leresche, 1991). The GABAb receptor is also strongly voltage dependent, with its efficacy sharply decreasing when the cell is depolarized from its resting potential. The GABA subtypes also have a distinct distributions and action in the dLGN. The GABAa receptor mediated pathway acts strongly on dLGN relay cells, but is not found in inhibitory interneurons (Soltesz and Crunelli, 1992). GABAb acts fairly weakly on dLGN relay cells. Its main role is probably to regulate relay cell responsivity, rather than to shape relay cell RFs (Soltesz and Crunelli, 1992). GABAb also has little effect on interneurons (Williams et al., 1996).

The spatial distribution of GABA receptors can also help illuminate the roles of inhibition.

The placement of GABAa receptors at locations like the base of glomeruli and close to the soma (Sherman and Koch, 1990; Uhlrich and Cucchiaro, 1992) could help support the shunting role of GABAa receptors because the receptors have maximal influence on the dendritic signals to the soma in this position. Interneuron terminals also occur frequently in glomerular structures, where they are thought to regulate the incoming retinal signals.

RNT/PGN terminals are most frequently found on the dendrites of relay cells in close proximity to excitatory cortical terminals, suggesting that the RNT/PGN inhibition probably modulates the cortical influence on the dLGN relay cells. These RNT/PGN afferents generate IPSPs in dLGN interneurons via GABAa receptors, and IPSPs in dLGN relay cells via GABAa and GABAb receptors (Soltesz et al., 1989; Soltesz and Crunelli, 1992; Crunelli and Leresche, 1991). The primary influence of RNT/PGN afferents is mediated by GABAa, as it elicits IPSPs at very low levels of RNT/PGN stimulation (which are blocked by a GABAa antagonist). GABAb mediated IPSPs only appear after much stronger RNT/PGN stimulation (Sanchez-Vives and McCormick, 1997).

### *2.3.6 dLGN Response Modes: Tonic, Bursting and Spindle Waves*

GABAb is thought to mediate the shift between tonic firing mode — in which the dLGN relay cells respond to retinal inputs in a linear manner, and the bursting mode — in which the dLGN relay cells respond to retinal inputs non-linearly with bursts of action potentials. The tonic mode is associated with alert responsivity to retinal inputs. The bursting mode possibly underlies both supersensitivity to inputs (i.e. firing a burst of action potentials in response to a weak input effectively amplifies the input), and non-responsivity to retinal inputs during spindle waves (common during REM sleep). In general GABAb is assumed to reduce the dLGN relay cell sensitivity to sensory inputs.

On the other hand, GABAb could push the cells toward bursting by evoking low-threshold $Ca^{2+}$ potentials which "prime" the dLGN cells to fire bursts of action potentials (Soltesz and Crunelli, 1992; Crunelli and Leresche, 1991; Emri et al., 1996). When in bursting mode, spindle waves can occur in response to barrages of IPSPs from the RNT/PGN. These waves arise because the membrane, hyper-polarized by the IPSPs, rebounds and overshoots its resting potential and generates a low-threshold $CA^{2+}$ potential that can elicit bursts of action potentials in the dLGN relay cells. These action potentials in the dLGN excite the RNT/PGN neurons via thalamocortical axon collaterals, and a feedback loop is created that supports the oscillation of activity between the RNT/PGN and the dLGN. The resulting spindle wave activity is often seen during REM sleep (Crunelli and Leresche, 1991).

In the adult, GABAb IPSPs might be able to mimic the role of the barrages of IPSPs from the RNT/PGN by directly induce bursting responses in dLGN relay cells via hyper-polarization and rebound spiking. Despite this evidence for its involvement in the bursting behavior of dLGN relay cells, GABAb doesn't seem to be necessary for the spindle waves because blocking GABAb receptors does not abolish them (von Krosigk et al., 1993). Indeed, some research suggests that

GABAa is primarily responsible for spindle waves in the dLGN, with GABAb playing a significant role only in pathological oscillatory behavior in the dLGN (Sanchez-Vives and McCormick, 1997). Interneurons seem not to have a bursting mode of responsivity (Williams et al., 1996; Crunelli and Leresche, 1991).

# Chapter 3

# Biological Background: Development

This chapter will discuss the anatomical development of the early visual system in the neo-natal ferret. It will focus on activity-dependent development of map-like projections and laminated structures. It is an attempt to establish the parameters relevant to a formal model of the development of the ordered retinogeniculate projection.

Much is known about the fetal development of the ferret visual system because ferrets are born, and are thus available for study, in an extremely immature state. Furthermore, the developmental time-course and mature structure of the ferret's visual system are closely parallel to that of the well studied cat (see Cucchiaro and Guillery, 1984 for review and see Table 3.1 for a listing of developmental milestones discussed below).[1] Findings from the cat will be presented alongside those from the ferret to help construct a complete picture. For purposes of comparison, the time from conception to eye opening is comparable in the two animals (72-74 days in the ferret (Wong et al., 1993), and 75 days in the cat (Shatz, 1983)), but ferrets are born 3 weeks earlier than cats (43 days and 65 days after conception respectively, "D43" and "D65"). The early birth of ferrets essentially makes possible "fetal" studies of the developing visual system without fetal surgery.

## 3.1 Problems to be Solved During Development

The emergence of structure in the retinogeniculate projection is one of the most investigated questions in developmental neuroscience, but the exact hierarchy of mechanisms involved is still not known in detail.

The developing brain has some practical engineering problems to solve. RGC's eventually must connect to targets that have a precise and predictable 'horizontal/lateral' location (i.e. their appropriate topographic position) and 'vertical' location (i.e. in the appropriate layer) in the dLGN. Furthermore, the retinotopic representations in the different layers must end up in register, such

[1] This common neural organization is sometimes referred to as the "carnivore" pattern.

| Days | Event in Cats | Days | Event in Ferrets |
|---|---|---|---|
| 19 | RGC's in Optic Stalk [1] | 20-30 | LGN and Subplate Neurogenesis [3] |
| 22-54 | W (small) RGC Neurogenesis [1] | 20-36 | Cortical Layer 6 Neurogenesis [3] |
| 22-37 | SC Neurogenesis [1] | 20-43 | RGC birth [3] |
| 22-32 | dLGN Neurogenesis [1] | 24 | Axons leave LGN & V1 [3] |
| 26 | Axons leave LGN [3] | 24 | RGC's in optic stalk/chiasm [3] |
| 29-32 | X (medium) RGC Neurogenesis [1] | 27 | dLGN $\leftrightarrow$ & Cortical axons arrive [3] |
| 30 | Retinal IPL formed [1] | 28 | RGC axons $\rightarrow$ contra-dLGN [8] |
| 32-34 | Y (big) RGC Neurogenesis [1] | 31-41 | Cortical Layer 4 Neurogenesis [3] |
| 32-35 | RGC axons $\rightarrow$ dLGN [8] | 32 | RGC axons $\rightarrow$ dLGN [8] |
| 32 | RGC axons $\rightarrow$ contra-LGN [8] | $\leq$40 | Coarse SC and dLGN retinotopy[9] |
| 35 | RGC axons arrive at ipsi-LGN [8] | 41-43 | **BIRTH** |
| 36 | dLGN axons $\rightarrow$ Striate Cortex subplate [3] | 41-69 | Waves present [2] |
| 39 | Functional Retinogeniculate synapses [6] | 41-43 | RGC axons start layer segregation [8,2] |
| 40 | Peak #'s of RGCs [1] | 43-97 | LGN dendrites grow [11] |
| 45 | Horizontal & Amacrine synapses[1] | 41-69 | LGN Layer Segregation [2] |
| 45 | Peak in Eye Overlap in dLGN [6] | 45 | Conventional Retinal Synapses appear [2] |
| 50 | Retinal OPL formed [1] | $\leq$50 | Eye-specific axon segregation done [12] |
| 52 | RGC X/Y Morph. Differentiation [1] | 51 | dLGN axons invade cortex [3] |
| 60 | dLGN axons invade cortex [1] | 51-58 | A/A1 dLGN cell lamination [5] |
| 60 | A/A1 dLGN cell lamination [1] & | 55-64 | Rods Mature, Retinal photosensitivity... |
| | RGC lamination in retina complete | | Off-cell noise, & On/Off-sublayer formation [2] |
| 60 | RGC axon segregation is complete [6] | 59 | Ribbon Synapses appear in Retina [2] |
| 65 | Mature RGC density in area centralis [1] | 62 | On/off RGC axon segregation complete [13] |
| 65 | **BIRTH** | | |
| 68-79 | A1/C3 dLGN Cell Lamination [1] | $\leq$69 | Waves Disappear [2] |
| 71-73 | Receptors Mature | 71 | Laminar segregation complete [5] |
| 71-74 | Cell Lamination complete [6] | 73 | Orientation Columns in Striate Cortex [4] |
| 75 | **EYE OPENING** | 73 | **EYE OPENING** |
| 75 | Orientation Columns in V1 [4] | | |
| 75-85 | Highest Synaptogenesis in Retina [10] | | |
| 79 | LGN dendritic growth [1] | | |
| 86 | Direct LGN to Layer 4 [4] | | |
| 86 | Can correlate RGCs Morph. & Phys. [7] | | |
| 95 | Lateral synapses mature [10] | | |
| 115 | Ribbon synapses mature [10] | | |

Table 3.1: Time-line of events in the development of the visual systems of cats and ferrets. All dates are number of days after conception. Note that many of the main milestones are very similar in the two animals, but ferrets are born approximately 3 weeks earlier. Sources: 1) Chalupa and White, 1990, 2) Wong et al., 1993, 3) Johnson and Casagrande, 1993, 4) Daw, 1995, 5) Linden et al., 1981, 6) Shatz, 1983, 7) Garraghty and Sur, 1993, 8) Cucchiaro and Guillery, 1984, 9) Jeffery, 1990, 10) Friedlander and Tootle, 1990, 11) Sutton and Brunso-Bechtold, 1991, 12) Penn et al., 1998, 13) Wong and Oakley, 1996. "TCC" is an acronym for thalamo-cortical cell, or dLGN relay cell.

that a vertical track through the nucleus represents a single RF in visual space. Remarkably, there appears to be very little error in the overall pattern of these connections. For example most dLGN relay cells receive input from only 1-4 retinotopically appropriate RGCs and there is virtually no overlap between afferents from the two eyes or different functional classes in the dLGN.

The final organization of the projection emerges early in development (before eye opening) and is quite predictable (e.g. the orientation of the retinotopic maps and ordering of the layers is similar across animals). These factors suggest a genetic mechanism for generating organization. On the other hand, the degree of precision in the final arrangement seems unlikely to be directly determined genetically. It simply is too much of a task to specify the exact wiring of hundreds of thousands of intermingled RGC's projecting in such an elaborate pattern to a target located 2-3 cm away.

There are a variety of mechanisms aside from direct genetic control that might play a role in the establishment of horizontal and vertical specificity in connectivity. Three of the main types of mechanisms that have been investigated are mechanical, chemical and activity-dependent. This section will explore the possible mechanisms involved, and will posit plausible mixtures of mechanisms that are at work. The relevant empirical work will be laid out in the remainder of this chapter.

### *3.1.1 Possible Mechanisms of Retinotopy*

The formation of a retinotopic projection to the dLGN essentially means that neighboring dLGN neurons must come to receive inputs from either the same or neighboring RGC's.

*Mechanical Mechanisms.* A mechanical explanation seems to be an obvious candidate for preserving these spatial relations. That is, neighboring cells could simply remain next to each other throughout their journey through the optic nerve, at the optic chiasm, through the optic tract, and eventually innervate neighboring tissues, thus preserving retinotopic order. Indeed, much of the information about spatial gradients of maturation in the retina points to a likely role for mechanical processes in maintaining neighbor relations in the retinogeniculate projection, but this role is limited. Stated simply, the fibers become disordered in transition to the dLGN where they eventually become highly ordered. The fiber organization in the optic tract is far more disordered than that in the dLGN ($30^\circ$ versus $.5^\circ$ of scatter, see Sec. 2.3.2 and Voigt et al., 1983). In light of this secondary precision, there must be intervening processes which serve to organize the retinogeniculate projection.

*Trophic Factors and Cell Death.* Another mechanism, elucidated by Sperry (Sperry, 1963), depends on tropic factors. The afferents could preferentially be attracted and/or stabilized by mediating chemical factors that vary over the pre-synaptic and post-synaptic cells. Such a chemical foundation could be specified in at least two ways. There could be a particular chemical or combination of chemicals specific to each afferent and each target cell — akin to having an individual

lock and key for each cell. This arrangement is highly unlikely simply because the number of individual chemicals needed is prohibitive. A more plausible scheme uses chemical gradients instead of different chemicals to specify location, two perpendicular gradients are sufficient for a flat sheet. For example, an incoming afferent fiber would have a preferential attraction, determined by the location of its soma in the source tissue, to specific preferred concentrations of the two chemicals — essentially matching up coordinates in the source and target tissues.

There could also be activity-related trophic factors that could guide axons selectively toward co-active (thus appropriate) locations without shifting the entire arbor. In fact, some individual axons (after D57) have been observed to make gross trajectory corrections in the cat (Sretavan and Shatz, 1987). These rare axons make sharp turns along the mediolateral axis and arborize at a new lateral location. The majority of axons seem to grow straight into the nucleus, perpendicular to the layers, and monotonically grow a more complex arbor in the appropriate spot. Activity might also affect patterns of cell death so as to eliminate projection errors, thus refining retinotopy. Cell death has been shown in some systems to be an active mechanism in the refinement of retinotopy (Sretavan and Shatz, 1987).

*Activity-Dependent Mechanisms.* Finally, activity-dependent modification of synapses, e.g. Hebbian learning (Hebb, 1949), could utilize patterns of activity to sculpt initially disordered projections into precisely organized topographic projections. This mechanism essentially uses correlations between the spontaneous (the retinogeniculate projection becomes organized before eye opening) activity of nearby cells (e.g. due to lateral spread of excitation) in the retina in conjunction with that induced in the dLGN to selectively strengthen the synapses from spatially contiguous RGCs to spatially contiguous dLGN relay cells. This selective strengthening encourages neighboring cells in the target tissue to have neighboring receptive fields. If this principle is successfully enforced across all the cells in the target tissue, the result is a topographically correct projection (Willshaw and von der Malsburg, 1976).

Given the precision of the adult map and the intervening disorder in the projection, it is likely that activity-dependent mechanisms are needed to support precise refinement. Most models of activity-dependent topographic refinement rely on the pruning of initially wide-ranging "blurry" connections to form a focused arbor in the appropriate location. In contrast, RGC axons do not have wide ranging arbors at any time during development. As a result the RGC axonal arbors must be capable of shifting over time in order to be capable of correcting initially poor (non-retinotopic) arbor placement. In lower animals entire arbors have been observed to shift (e.g. in the retinotectal projection of *Xenopus*, Udin and Fawcett, 1988).

Of course, these mechanisms might also work in concert to attain the high degree of precision in the mature projection. In fact it appears that a combination of all three of these mechanisms helps shape the formation of a retinotopic projection. Mechanical factors appear to limit the degree of disorder in the initial in-growth. Chemical factors help the projection reach fairly high levels of precision even very early in development. Finally, activity-driven mechanisms most likely help

eliminate erroneous projections and make fine adjustments in the horizontal position of afferent arbors and their target cells.

### *3.1.2 Possible Mechanisms of Layer Segregation*

A similar story applies to the development of laminar structure (vertical specificity), but there seems to be a much heavier reliance on activity-dependent refinement.

Because afferents that eventually arborize in different layers are intermingled in the retina the mechanical situation in layer formation is unlike that in retinotopy. Instead of preserving spatial relations present in the retina, any mechanical processes would have to preserve spatial groupings stemming from differential timing of axonal migration and innervation into the dLGN. By capitalizing on the differences in timing between contralateral and ipsilateral afferents and X and Y afferents, mechanical factors could achieve some measure of segregation.

Similarly to retinotopic organization, chemical factors that are differentially distributed in the different layers of the dLGN could be involved in differentially attracting and stabilizing RGC arborizations in the appropriate layer.

Activity-dependent processes could serve to cluster afferents that tend to be co-activated in clusters/layers in the dLGN. For this process to work, afferents that eventually segregate would have to be less correlated in their firing than afferents that are eventually clustered into the same layer. This means that afferents originating from the same eye or functional subclass would have to be more correlated than those originating from different eyes or functional subclasses.

It appears that mechanistic processes and chemical processes help bias the ordering of layers toward the stereotypical pattern, but they establish only very rough segregation and ordering with a large degree of overlap. Activity-dependent processes play a much larger role in the establishment and refinement of laminar segregation than they do in establishing retinotopic projections.

A large body of information and research comes to bear on questions about the mechanisms underlying the development of elaborate structure in the retinogeniculate system. Findings will be presented in the same general framework used above in the outline of mature visual system organization. The discussion will again proceed from the eye through RGC's and the optic nerve, chiasm and tract and into finally into the dLGN. Crucial aspects of retinal and dLGN physiology will also be presented in hopes of specifying the potential role of activity-dependent processes.

## 3.2 Developmental Trends in the Eye and Retina

Understanding the development of the eye and retina can help guide reasoning and assumptions about the development of its projections to more central targets. Of clear relevance are features of development concerning the spatial, temporal or functional ordering of neurogenesis, maturation

and axonal outgrowth. These orderings might be the source of mechanical mechanisms of generating and preserving organization in the developing RGC projection. In addition, information about the development of optics and synapses can help generate more subtle insights as to the state of early circuits and apparent priorities in their maturation.

### 3.2.1 Increasing Acuity: Anatomical Issues

The increase, over development, in behaviorally measured spatial acuity has often been thought to be a function of increasing acuity of the neural visual circuitry (Daw, 1995). This reasoning supported the assumption that connectivity in the neonatal visual system is "blurry", meaning imprecise and likely exuberant with transient synapses. The existence of this sort of blurry initial projection would support the role of activity-driven topographic refinement, for it indicates that mechanical and chemical processes support only coarse order. Yet, there are several features of the immature eye that account for the lack of early spatial acuity without making any assumptions about the blurriness of the neural circuitry — thus challenging the accuracy of these assumptions about neonatal visual circuitry.

First, the eye doubles in size from about 10 to 20 $mm$ over maturation. The area of the retina illuminated by a spot of light roughly increases roughly by 4 times (A $1^\circ$ spot projects to a 129 $\mu m$ and 213 $\mu m$ diameter spot at birth and adulthood respectively). It can be assumed that if more retina processes a given portion of space, spatial acuity in that portion of space will increase (Friedlander and Tootle, 1990). Therefore, the mere smallness of the neo-natal eye can help account for its low spatial acuity.

Further, in the kitten there is a persistent pupillary membrane and vascular network supplying the developing lens. This network scatters light and dramatically limits the resolution of the image that reaches retina (Friedlander and Tootle, 1990). Prior to D86 in cats the lens significantly blurs the retinal image; it becomes close to mature around D89. Notably, ferrets (Daw, 1995) have a clear lens by the time of eye-opening (D74) and monkeys and humans have a clear lens at birth (Friedlander and Tootle, 1990).

### 3.2.2 Processing and Relay Circuitry before Transduction

Two other trends in retinal development, termed "inner to outer (centripetal)" and "lateral to feed-forward" here, might help clarify priorities in the developing retina. Though the retina is primarily a transduction organ, the evidence suggests that the retina generates and relays spontaneous excitation *before* developing the machinery for transduction. These trends potentiate activity-dependent mechanisms of refinement because they support the generation of spatiotemporally correlated patterns of activity in the immature retina (see below).

The inner to outer trend is manifest in the order of laminar development in the retina. Maturation progresses centripetally from the inner (RGC) to the outer (receptor) layers. The inner

plexiform layer (IPL) is first formed around D30 (in newborn ferrets (D43) the IPL is the only layer present (Wong et al., 1993)), while the outer plexiform layer (OPL) isn't formed until D50 — lamination being nearly complete around D65 (Chalupa and White, 1990). Exemplifying this trend, the receptors mature quite late in the cat (D71-73 (Friedlander and Tootle, 1990); in the ferret, structures crucial for transduction in the rods only appear around D57-64 (Wong et al., 1993)). Some aspects of the the relay and modulatory circuitry have a developmental lead on the transduction circuitry of almost 6 weeks.

There is also a somewhat surprising, yet robust, ordering in synaptogenesis. In the IPL, the conventional synapses that modulate activity laterally appear and mature before the ribbon synapses that pass information centripetally (i.e. feed-forward connections) (Chalupa and White, 1990). Speaking in terms of specific dates, in the IPL conventional synapses that modulate activity laterally appear by D45 (ferret) (Wong et al., 1993) and mature by D95 (cat) (Friedlander and Tootle, 1990), while ribbon synapses appear by D59 (ferrets) (Wong et al., 1993) and mature by D115 (cat) (Friedlander and Tootle, 1990). This ordering suggests a primacy in setting up the intrinsic circuitry of the retina, as opposed to the retina's ability to take in visual information (clearly its primary function) — perhaps to support the generation and propagation of useful patterns of spontaneous activity (described below).

### *3.2.3 Center to Periphery*

A mechanical mechanism for preserving retinotopy could be based on a lateral gradient in the maturation. If the cells that matured earliest migrated and innervated the dLGN first they would be the only cells that would need to be guided to their exact target, later maturing cells could just fill in the next available space. In this way afferents would simply be 'laid down' in the order of their maturational gradient, preserving spatial order.

This scheme could play a role in ordering the retinogeniculate projection because the maturation of the retina does in fact have pronounced spatial gradients. The retina of ferrets and cats matures laterally from the center outward. This trend can be seen in patterns of neurogenesis, cell death and lamination. In the cat, the first RGC's are born in the central retina on D21, while the last RGC's are born around D36 in the peripheral retina (Chalupa and White, 1990). RGC's are generated from D20-D45 in the ferret (Johnson and Casagrande, 1993).[2] Cell death appears to be as spatiotemporally programmatic as cell birth.

While these spatial trends could potentially play a role in the development of retinotopy, they do not provide the necessary groupings for layer formation in dLGN. Mechanical means of layer

[2]But these figures are based on tritiated thymadine studies where label is incorporated into cells during of cell division — allowing their birth-date to be determined by label concentration. Unfortunately, only the cells that survive until adulthood can be dated, thus the figures omit 'transient' cells that die before adulthood. This omission could be misleading if the transient cells play crucial pioneer roles in setting up projections. Indeed, RGC axons have been seen in the optic stalk as early as D19 (Chalupa and White, 1990).

formation would need the RGC's from each eye, and from each functional subclass, to be differentially grouped at some point in development. Different developmental time-courses in axonal outgrowth of the RGC's from each eye and the X, Y and W subclasses might play a role in the development of layers in the dLGN.

### *3.2.4 X, Y and W Pathways*

Unfortunately, details about the maturation of each cell class are somewhat tentative for methodological reasons.

Using morphological differences, some RGC's can be classified into cell classes by D52 (Chalupa and White, 1990), but due to methodological considerations, one can't correlate the morphology and physiology of these cells until D86-D100 weeks post-natal (Garraghty and Sur, 1993). These difficulties limit our knowledge of the early developmental time-courses of different cell classes. Fortunately, tritiated thymadine studies reveal an interesting order in the genesis of different sizes of RGC's.

These studies have determined that small ganglion cells (presumably W-cells) are generated throughout the period of RGC genesis (D22-D36), and small cells thought to be displaced amacrine cells continue to be born until D54. Medium RGC's (presumably X-cells) begin to be generated slightly later, the last are born D29 in the central retina and D32 in the peripheral retina. Last generated are the large cells (presumably Y-cells), which are generated in the central retina at D32 and in the peripheral retina D34 (Chalupa and White, 1990). These experiments suggest that X-cells have a developmental lead of 4 days on Y-cells. X-cells are born before Y-cells, they reach the optic tract first and most likely they grow into the dLGN before Y cells. Given that the X-cells come to dominate the deepest (most dorsal) dLGN in the absence of activity (see below), these facts suggest that, through mechanical mechanisms, dLGN structure may reflect the sequential order in maturation of RGC cell types (Garraghty et al., 1986; Garraghty and Sur, 1993).

## 3.3 Retinogeniculate Projection

Developmental details of the retinogeniculate projection of the ferret and cat are the centerpoint of this paper. The key issues are: the extent and character of disorder in the original projection, the time-course of laminar and retinotopic development, and the effects of various manipulations on the maturation of these projections.

Some background information, including milestones of development (see Table 3.1) and the nature of the target tissue, will help provide a framework for investigating these issues.

### 3.3.1 Anatomical Background: Milestones and dLGN Tissues

In both ferrets (Johnson and Casagrande, 1993) and cats (Chalupa and White, 1990), RGC axons are first seen in the optic stalk around D24 (sometimes as early as D19). Ferret RGC axons arrive between D28-32 while in the cat they arrive somewhat later, D32-35. Afferent segregation in the dLGN is complete by D60 and cellular segregation is complete by D71-74 (Shatz, 1983; Wong et al., 1993).

There are several subtypes of dLGN relay cell and interneurons as classified by Mastronarde (1992). dLGN neurogenesis of both dLGN relay cells and interneurons occurs from D22-32 in the cat (Weber et al., 1986; Chalupa and White, 1990), and D20-30 in the ferret (Johnson and Casagrande, 1993). In contrast to retinal neurogenesis, there are only weak spatial gradients (if any at all) in dLGN neurogenesis in the cat and ferret. X and Y cells in the dLGN can develop rough morphological differences without afferent input (Casagrande and Condo, 1988) — suggesting that they have some intrinsic cell fate. Yet, their physiological differences are thought to be primarily a result of the properties of their inputs.

Synaptic development in the dLGN is overwhelmingly postnatal in the cat. The period of most rapid increase is D73-85, around the time of eye opening. Later development in the Y pathway grows slowly and monotonically, with the number of Y synaptic boutons increasing by a factor of approximately 3, and with the synaptic bouton density increasing by a factor of 2 (Friedlander and Tootle, 1990). Presumably the number of boutons in the X pathway actually decrease or stay the same over this same period because the X arbors are shrinking (see below). To my knowledge this issue has not been directly explored.

### 3.3.2 Arrival of RGC axons at the dLGN

The ordered arrival and in-growth of RGC axons into the dLGN could support mechanical processes of afferent segregation. Axons first arrive in the dLGN from D25-32. The earliest arrivals are from the contralateral eye, with the ipsilateral eye arriving roughly 3 days later (Cucchiaro and Guillery, 1984). Superimposed on the contralateral before ipsilateral trend is the X- before Y- cell trend as mentioned above. Thus, X-axons from the contralateral eye are most likely be the pioneers into the dLGN.

### 3.3.3 Axonal and Dendritic Arborization in the dLGN

RGC's running along the optic tract tend to take approximately 90° turns into the dLGN both at initial in-growth and maturity. Initial in-growth is relatively simple; the fibers grow in parallel fashion from the optic tract (outer) across the putative layers and towards the middle of the thalamus (inner). There is no evidence of exuberance in the width of RGC arborizations in the cat, rather the axons appear to penetrate into the dLGN in a relatively focal and orderly manner (Sretavan and Shatz, 1987).

In the dLGN RGC axons initially form diffuse connections spanning several layers. The prototypical immature RGC axon has many short side branches along its trunk, making it resemble a straight 'hairy' fiber approximately 100 $\mu m$ wide. As development progresses, the dLGN differentiates into eye specific layers that have differential cytoarchitecture, afferents and efferents. Leading this process, the afferent fibers progressively eliminate short side branches in inappropriate layers and ramify their arbors focally in the appropriate layers (Shatz, 1994; Sterling, 1990). There is a pronounced trend in arbor maturation in the dLGN from medial to lateral — probably inherited from the gradient of RGC maturation (Sretavan and Shatz, 1987).

The dendritic trees of dLGN neurons in ferrets from D43 through D97 (peak in arbor size), decrease rapidly until D133 and then more gradually to adult levels (Sutton and Brunso-Bechtold, 1991; Sutton and Brunso-Bechtol, 1993). There seems to be a marked period over which dendritic appendages are formed and then eliminated. These transient appendages account for the bulk of dendritic growth at the onset of vision (D73-D99) (Sutton and Brunso-Bechtol, 1993). Most of these appendages are eliminated in later development.

### *3.3.4 Time-Course of X and Y Arborization*

As mentioned above, because it is impossible to correlate physiology and anatomy of RGC's before D86-D100, there is nothing known about the differential development of X and Y arbors until well after birth. However, the postnatal development of the arbors of the different cell classes is illuminating. The general picture that emerges is that the development of the Y pathway is more sensitive to activity patterns than the X pathway.

Garraghty, Sur (Garraghty et al., 1986; Garraghty and Sur, 1993; Sur, 1988) and their colleagues have extensively investigated and analyzed the development of these arbors in the cat. They found that both X and Y arbors extensively remodel in opposite directions during the period from D92 to adulthood. At D65 the retinogeniculate arbors (undifferentiable into X and Y classes) are uniformly narrow — 80-100 $\mu m$ in width (Sur, 1988). X arbors are exuberant in early postnatal development and then shrink down over the first 3 months, while Y arbors grow monotonically throughout development. From D86 to D149, the mean X arbor width shrinks from 180 $\mu m$ to the adult width of 115 $cm$. Over the same period, the mean Y arbor width grows from 193 $\mu m$ to 293 $\mu m$ (for review see Friedlander and Tootle, 1990).

These findings suggest that X arbors are culled back by expanding Y arbors, essentially reflecting a shift from X to Y dominance in vision. Further complementing the story are studies involving disrupted visual experience (for review see Garraghty and Sur, 1993). These interventions almost always cause Y arbors to be smaller in the A lamina (where they overlap with X axons), but to remain the same in the C lamina (where they do not overlap with the X axons). Conversely, in these same animals the X arbors retain their immature exuberance in the A lamina. This same contrast between X and Y arbor development is apparent in laminar development (see below).

### 3.3.5 Retinotopy

Preservation of retinotopy is one of the most striking features of the retinogeniculate projection. It is a highly specific, organized and precise connection scheme. The exact mechanisms behind the development of retinotopy are unknown. In addition, details about the development of precise retinotopy (Daw, 1995) and the effects of subtle manipulations on retinotopic development are unknown for methodological reasons. Development on a somewhat coarser scale is well investigated though.

*Direct Measures: dLGN.* Jeffery (1985,1989,1990) has done the main to date that directly investigates retinotopy in the retinogeniculate projection. He ablates a small portion of the retina and then anterogradely stains the projections to the dLGN and the SC. To the extent that the projection is retinotopic, there is a gap in the pattern of label in the target tissue. Furthermore, the sharpness of the borders of the gap indicate the sharpness of the projection.[3]

Jeffery has demonstrated that the very early retinogeniculate projection is retinotopic and that the degree of retinotopy, as reflected in border sharpness, remains unchanged from D40 through adulthood.[4] Furthermore, he finds that the degree of sharpness is the same in the SC and the dLGN (Jeffery, 1985). These findings are informative in several ways. First, they demonstrate that there is a notable degree of retinotopy even in the very early (D40) retinogeniculate and retinocollicular projections. Second, in conjunction with physiological work they establish upper bounds on the resolution of Jeffery's method. Retinotopy in the dLGN is far more precise, .5° of scatter, than that in the SC, 10-30° of scatter (for review see Voigt et al., 1983), but this difference is invisible using this method. It follows that Jeffery's method has a resolution of, at best, 10° of scatter.

Because of these limitations, the time-course of fine retinotopic organization in the retinogeniculate projection is unknown. It is possible that the retinotopy is fairly sharp from the earliest in-growth of RGC axons, or the projection could be relatively disordered initially (like the mature SC) and then refine gradually. If the latter is true then activity-dependent mechanisms may play an important role in refining the projection.

*Direct Measures: SC.* The development of retinotopy has been investigated far more extensively in the SC (and the tectum of lower animals) than in the dLGN. In hopes of shedding some light on general trends in the development of retinotopy, the findings are presented in brief form here.

[3] Sretavan and Shatz, 1987 performed this same experiment on several animals in which the retinas were accidentally damaged and, as expected, found retinotopic gaps in the dLGN. The level of resolution was poor enough for them to state that "the degree of topographic precision in the projection from the retina to the dLGN during fetal development requires much further direct study".

[4] Topographic organization is not demonstrable in the part of the dLGN closest to the optic tract (C laminae) until later ages (around D44). That is, the gap in the labeling does not extend the full mediolateral extent of the nucleus, rather a band of labeling remains between the gap and the optic tract (Jeffery, 1989).

Chalupa and his colleagues (1996,1990) have done elegant experiments to determine the degree of retinotopy in the developing retinocollicular projection in cats. Because these experiments require *in utero* surgery, the precise deposit of a retrograde tracer into the target tissue, and then keeping the animal alive for up to 2 days — they are difficult to perform. Because the tracer deposits are visually guided, these experiments are likely performed in the SC because that nucleus is not covered by the cortex and is thus visible during the relevant period of pre-natal development. There are no comparable studies in the dLGN (Chalupa, personal communication).

These studies have found less than 1% of RGC's made gross projection errors. At all ages (E37 to adult) there was a focused area of high density labeling in the retina. The ectopic cells were scattered thinly across the entire retina. These findings contrast with the finding (in the rat) that there is more disorder in the original projection than in the adult projection and that the ectopic cells stem mostly from immediately neighboring retinal regions (Simon and O'Leary, 1990). These differences could stem from methodological or species differences.

### *3.3.6 Laminar Organization*

There are two main substrates of laminar structure: the segregation of RGC axons ("afferent segregation"), and physical segregation of dLGN cell bodies corresponding to the formation of interlaminar cell-sparse spaces. Cellular lamination lags afferent segregation by about 2 weeks, starting around D60 and continuing until just before eye opening (D73) in the ferret (Linden et al., 1981).

The segregation of axons is two-fold with eye-specific layers and on/off sublayers forming. The most robust, early and marked segregation is between the afferents from the two eyes. That is, the RGC axons from the each eye segregate over the course of development such that in the adult animal they arborize in roughly alternating layers. Axons cross the layers innervated by the other eye (if necessary) with a minimal axonal shaft, rarely retaining any side branches, terminal boutons, or synapses in the inappropriate layers (Shatz, 1983; Sretavan and Shatz, 1984; Sretavan and Shatz, 1986).

The axons of different functional subtypes of RGC's (i.e. X, Y, and W cells or on and off cells) also tend to focally arborize in some layers of the dLGN while only traversing or simply avoiding inappropriate layers. Axons from on and off center RGC's do not segregate in the cat, but they do separate to form separate sublayers in the A laminae of the ferret. The on/off layers form after the eye-specific layers (Wong and Oakley, 1996).

*Segregation of Afferents.* The axons of different functional subclasses of RGC's (e.g. on and off and X- and Y- ) are grouped in the optic nerve (Sretavan and Shatz, 1987). However, when axons grow into the dLGN, the arbors from the two eyes and from different functional classes overlap extensively. This tendency toward initial overlap of the afferents is weaker in cats than in ferrets or monkeys. According to some reports monkeys have 100% anatomical overlap in the immature

dLGN (Rakic, 1977 p. 206 as quoted in (Shatz, 1983)), while ferrets have nearly complete overlap in the binocular portion of the nucleus (Linden et al., 1981, but see Penn et al., 1998 for illustrations of incomplete overlap at D41).

In both ferrets and cats, the imperfect mixing of afferents reflects the early contralateral dominance of the dLGN. Contralateral afferents arrive first and penetrate all the way to the medial border (the A laminae), ahead of the ipsilateral axons. It is thought (Shatz, 1983; Garraghty and Sur, 1993) that this gives the contralateral axons a competitive advantage in occupying the deepest A laminae (i.e. the portion of the dLGN farthest from the optic tract). In fact, the deepest portion of the cat's A laminae is never innervated significantly by the ipsilateral afferents (see figure 13 (Shatz, 1983)). In ferrets, the ipsilateral component fills the entire depth of the nucleus early in development (Cucchiaro and Guillery, 1984), although it never innervates the monocular segment of the nucleus (see Fig.4.1 and see Sec.3.3.7).

In cats and ferrets alike, afferent segregation starts between D43-47 and is complete around D60, about 2 weeks before eye opening (Cucchiaro and Guillery, 1984). In cats, the extent of overlap peaks around at D45, at this point 90% of dLGN cells can be driven binocularly and the percentage of dLGN that is highly innervated by both eyes is 40%. At birth (D65) around 50% of the cells can be driven binocularly and the afferents are highly segregated so that less than 10% of the dLGN is innervated significantly by both eyes. In addition, there must be significant postnatal development of the connections, because binocularly driven cells are quite rare in the adult cat dLGN (Shatz and Kirkwood, 1984; Shatz, 1983).

The order of lamination proceeds roughly from the deepest portions of the dLGN to those closest to the optic tract. At D47 in cats the afferents from the contralateral eye overlaps A1, except in the posterior pole, where lamination has presumably started. At D54 afferents to A1 and A are segregated but the C lamina divisions are difficult to discern. By D60 in cats the segregation between layer A1 and layer C is visible (Shatz, 1983). Adjacent layers innervated by the same eye, e.g. on and off leaflets, are segregated after those innervated by different eyes (Casagrande and Condo, 1988).

At D43 (birth) the ferret dLGN is unlaminated. The eye-specific lamination of RGC arbors is nearly complete by D48-D51 (Penn et al., 1998; Wong and Oakley, 1996). From D43-D46 the ipsilateral fibers progressively concentrate in the caudal and medial region, and the contralateral fibers retreat to the lateral area and the rostral medial area. Between D47-D51 the ipsilateral representation expand, densely innervates A1 and lightly innervates C1. The contralateral fibers overlap the borders of A1 and all of C at this point. The afferents are completely segregated into eye-specific layers after D58 (Linden et al., 1981). In ferrets, the segregation of on and off afferents in dLGN occurs after the eye-specific axon segregation — from D55-64 (Wong and Oakley, 1996). Mature dLGN cells in the ferret and cat are monocularly driven and are functional class specific physiologically (see Mastronarde, 1992 for counter-examples).

*Development of Cellular Lamination.* In cats, cellular lamination lags afferent lamination by about 2 weeks. Cytoarchitectural segregation of A/A1 starts around D60 (Shatz, 1983). The A1/C interlaminar zone emerges around D68 reaching its mature cell-sparse character by D79 (Chalupa and White, 1990).

In ferrets, lamination between layer A/A1 occurs between D51-58. Cellular lamination is apparent when the RGC arbors are still overlapping the interlaminar boundaries (Linden et al., 1981). The laminar maturation of the nucleus is complete by D71 (Wong et al., 1993; Linden et al., 1981). Although the C laminae segregate after those of the A laminae, many of the cells in the C laminae are actually more morphologically mature than those in the A laminae. This paradoxical trend reflects of the relatively early maturation of the W-cells, the primary inputs to the C laminae (Friedlander and Tootle, 1990).

The cell sparse interlaminar regions receive inputs from extra-retinal sources after the process of lamination is well under way. It is thought that this in-growth (perhaps just following the path of least resistance) may add mechanical assistance to the presumably chemical (trophic) influence of activity on laminar development (Casagrande and Condo, 1988). Interestingly, the cell free space in the dLGN representing the optic disc forms at the same time that the interlaminar spaces are forming. This suggests a link between refinement of retinotopy and laminar segregation (Casagrande and Condo, 1988).

### *3.3.7 Activity Dependence of Eye-Specific Layer Segregation*

Lid suture, dark rearing, pharmacological block of activity in the eye or dLGN, and enucleation (removal of an eye) all eliminate different aspects of activity during development.

Across these manipulations, X arbors seem to have their development specified relatively independent of activity, and tend to ramify in the appropriate layer in almost all cases. Y axons, on the other hand, seem to rely on activity to enable them to competitively grow their arbors after birth. Without normal patterned binocular input the Y arbors tend to be smaller than normal in layers that receive X input, and the Y arbors will invade deafferentiated layers, even if they are normally inappropriate targets (Garraghty and Sur, 1993). In summary, the laminar location of X arbors is relatively activity-independent while that of Y arbors is relatively activity-dependent.

*Dark Rearing, Lid Suture, and Strabismus.* The methods of lid suture, dark rearing, and strabismus [5] are only possible after birth (D43 in ferrets, D65 in cats) and are only useful after the onset of photoreceptor function (after D57 in the ferret (Wong et al., 1993) and after D71 in the cat — 5 days prior to eye opening (Friedlander and Tootle, 1990)). In the cat these manipulations can influence the maintenance of afferent segregation (the initial segregation is largely completed

[5] Strabismus is misaligning the two eyes through surgery. This procedure alters visual input such that points in the opposing retinas that would normally have overlapping receptive fields no longer do.

by D65 in cat), the postnatal intralaminar development of axonal arbors (in the cat peaking from D65-D93 and continuing until D140 (Garraghty and Sur, 1993)), and the development of cell free inter-laminar zones (starts D60 in the cat and ferret).

All of these these manipulations eliminate the fine binocular pattern information that would normally be present in visual stimuli. There is an interesting contrast between the results of binocular lid suture and dark rearing. While lid suture selectively prevents visual pattern information from reaching the retina[6], dark rearing entirely eliminates the influence extrinsic activity. Interestingly, the mere presence of relatively unpatterned light seems to affect the development of RGC arbors in the dLGN.

In dark reared animals, postnatal development is slowed but there are not significant changes in lamination or X/Y axon morphology in the dLGN. The truly surprising result is that development is normal in darkness. If the Y axons are taken to expand under the influence of externally provided input, then the Y arbors in the dark reared animals would be expected to be small to the extent that there is no activity in the eyes. Clearly there is no activity that is correlated between the two eyes in a normal fashion. Intrinsic activity (analogous to the retinal waves present before eye-opening) might provides enough information to support normal development (Garraghty and Sur, 1993). It is known for example that RGC's in the adult ferret retina are spontaneously active during sleep, though not in the highly spatio-temporally patterned way they are in fetal development (Wong et al., 1993). Perhaps extrinsic sources of light act as a trigger to end the period of spontaneous patterned activity in the retina and this signal is bypassed in dark-reared animals.

In contrast, binocular lid suture and strabismus both prevent RGC axons from developing normally. Specifically, both manipulations cause the the Y arbors in the A laminae to be smaller than normal, but do not effect the Y arbors. This pattern suggests that the Y arbors are disadvantaged in the A lamina where they compete for space with X axons, but are normally sized in the layers where they don't have to compete (Garraghty and Sur, 1993). This disadvantage stems from both lack of fine pattern information and misaligned binocular information.

One possible mechanism supporting the abnormal development after strabismus is the asynchronization of firing in nearby dLGN cells (Garraghty and Sur, 1993). The normal alignment of the retinotopic maps in adjoining layers might tend to support the synchronous firing of nearby cells in different layers because those cells have receptive fields in nearby parts visual space. It is possible that the misalignment induced by strabismus causes a asynchronicity in the firing of nearby dLGN cells in different layers. This asynchronicity could undermine mechanisms of "cooperative growth" between Y cells in adjacent layers (e.g. locally released trophic factors) that might normally be present.

*Enucleation.* In case of binocular enucleation there are, of course, no afferents to be segregated in the dLGN. It has been noted though, that interlaminar spaces fail to form in cases of

[6]See (Garraghty and Sur, 1993) for discussion about visual information that does get through the eye-lids.

binocular enucleation, suggesting that this process is dependent on the presence of retinal afferents (Casagrande and Condo, 1988; Guillery et al., 1985). Overall dendritic arbors tend to grow slightly less in binocularly enucleated ferrets. The incidence of long unbranched dendrites seem more common in enucleates than in normals (Sutton and Brunso-Bechtol, 1993). The development of some aspects of dendritic morphology also seem to be slowed by binocular enucleation. For example, club like appendages (terminal clusters/glomeruli) don't show up until after D64 in enucleated animals, but show up before D64 in normals (Sutton and Brunso-Bechtol, 1993).

Monocular enucleation (and the application of TTX) after the formation of eye-specific layers (i.e. at birth in cats, or 3 weeks after birth in ferrets) prevents the normal postnatal development of the X and Y arbors. The X arbors tend to retain the exuberant arbor size that they achieve between D86-93. The Y arbors tend to be smaller than average in the A laminae, and they also tend to sprout into inappropriate deprived layers and arborize there. They seem to follow the course of least resistance, ramifying their arbors in areas with minimal competition (Garraghty and Sur, 1993). Under these manipulations, both X arbors and Y arbors are more diffuse than in normal development.

Monocular enucleation in combination with lid suture of the remaining eye results in more clearly defined lamination than the monocular enucleation alone. If one interprets laminar structure as a result of even competition between the eyes (and functional cell classes) then this finding is not surprising. Essentially the suture attenuates the competitive advantage that is given to the non-enucleated eye — providing a more balanced activity (Garraghty et al., 1986).

Monocular enucleation before the period of layer segregation (pre-natally in cats, or at birth in ferrets) leads to radically abnormal laminar development in the dLGN. Roughly two layers form, a magnocellular-like layer likely consisting of collapsed A, A1 and dorsal C, and a parvocellular-like layer likely consisting of the more ventral C layers. These layers are separated by a cell free zone. The X and Y arbors have different etiologies in these animals. X axons terminate in regions that would be the appropriate layers for their eye of origin in a laminated dLGN — in the inner or outer portion of the magno-like layer in contralateral and ipsilateral dLGN respectively. Y axons, on the other hand, arborize throughout the magno-like layer, presumably due to lack of competition (Sur, 1988).

*Phamacological Blockade.* Application of TTX before the period of laminar segregation blocks eye-specific layer development (Shatz and Stryker, 1988). Some evidence suggests that, even under the influence of TTX, the ipsilateral afferents do not innervate all of layer A (i.e. they don't reach the innermost border of the nucleus). The dLGN does grow to normal size during TTX application and arbors develop, though abnormally, so TTX does not simply freeze development (Sretavan et al., 1988).

TTX actually encourages exuberant growth in the deprived RGC arbors. Under the influence of TTX, the retinogeniculate arbors from each eye arborize in the entire nucleus, sometimes crossing its entire mediolateral extent. The silenced RGC's have much wider, more extensive, and less

selectively ramifying arbors (Sretavan et al., 1988). This expanded arbor size may account for the elongated, diffuse and sometimes patchy receptive fields of deprived dLGN relay cells when compared to non-deprived cells (Archer et al., 1982). Following this tendency toward exuberance in the absence of activity, the number of dendritic spines in cats with TTX applied to the optic nerve and thalamus is 300% the number in normals (Dalva et al., 1994). These dendritic spines are capable of forming functional synapses (Dubin et al., 1986) even though they are frozen at the point of impulse blockade, retaining immature characteristics (Kalil et al., 1988).

Application of TTX to neonatal kittens prevents the dLGN cells from receiving input from only one RGC type. Normal dLGN cells receive inputs only from one of each of these subtypes of RGC's. In contrast, dLGN cells after neo-natal TTX application receive inputs from both X and Y RGC's and from both on and off-center RGC's. TTX seems to allow RGC's to form functional synapses on cells that are not their matched partners. These findings are consistent with the abnormal growth seen in RGC arbors under the influence of TTX. These effects are attenuated by allowing a recovery period after the deprivation or by delaying the initial deprivation (the measure is number of mixed on/off cells (Archer et al., 1982)). The attenuation of the effect of deprivation demonstrates that dLGN connectivity loses some plasticity after early post-natal development. The ability of older dLGN (D100-D114) cells to partially recover from deprivation demonstrates that the dLGN retains some plasticity long after segregation of the afferents (D58), and cellular segregation are complete (D80) (Archer et al., 1982; Dubin et al., 1986).

*Intra-ocular Blockade.* Recently researchers have managed to block activity at the eye using intra-ocular injections (Penn et al., 1998). Intra-ocular blockade has the advantage that it can be used to investigate the role of activity-dependent competition between synapses. Complete blockade of activity at the dLGN cannot provide information about the role of competition, because activity at all synapses are blocked simultaneously.

Using injection of epibatidine (a nAchR cholinergic receptor blocker, time-released via impregnated latex micro-spheres), the researchers completely abolished wave activity and action potentials in the RGCs of injected eyes. Repeated injections during the period of eye-specific layer segregation (D41-D50) blocked the laminar segregation of retinal afferents. In fact, in the ipsi-lateral eye, the projection from the treated eye decreased substantially (even disappearing in some animals). The treated eye's projection to the contra-lateral retinal remained almost normal in the monocular portion of the nucleus, but was much smaller in the binocular portion of the nucleus (i.e. layer A is much smaller). The remaining projection to the binocular portion sometimes remained intermingled with the ipsi-lateral projection from the untreated eye. In the case of binocular injection the monocular regions of the dLGN still received inputs from only the contralateral eye. In the binocular region the axons from both eyes are intermingled, with little sign of eye-specific layer formation. It appears that layer formation is dependent on the activity-dependent competitive interactions between RGC axons.

The researchers also point out that their experiment suggests that retinotopy is preserved without the influence of activity, because the monocular portion of the retina appears to receive inputs from the appropriate retinal sources without the influence of activity. This degree of retinotopy is roughly equivalent to that demonstrated in Jeffery's work discussed above in Section 3.3.5.

### *3.3.8 Activity Dependence of On/Off Sublayer Segregation*

The formation of on/off sublayers in the ferret is activity-dependent. Hahm and colleagues (1991) blocked N-methyl-D-asparate (NMDA) receptor mediated activity by administering D-APV (an NMDA receptor blocker) to the thalamus during the period of on/off layer segregation (D55-D62). They found that the sublayers failed to segregate, and RGC axonal arbors often span the sublayers. Even when the arbors are relatively restricted, they are positioned without respect to the sublayer borders. Notably, APV does not abolish geniculate activity (unlike TTX) — it attenuates the post-synaptic activity. Their results highlight the specific role of NMDA receptor mediated activity in the formation of the sublayers.

### *3.3.9 Maps and Layers: Shifting*

Jeffery's work (1985,1989,1990) has some interesting implications beyond demonstrating the presence of retinotopy in the early retinogeniculate projection. His findings demonstrate that, in neo-natal ferrets, both the ipsilateral and contralateral projections to the dLGN innervate almost the entire nucleus. However, as was mentioned above, the ipsilateral projection is small and only comes from a small patch of the temporal part of the retina. The contralateral projection comes from the entire retina in the neonate and only the central and nasal portion of the retina in the adult. The most temporal portion of the contralateral projection dies sometime between D41-45. The portion of visual space represented by the neonatal ipsilateral projection is only in correspondence with the most caudal portion of the contralateral projection. This means that the neonatal projections from the projections from the two eyes are not in register as they are in the adult (compare Figs. 1 and 11 in Jeffery, 1990). The projections must shift in relation to each other, such that the ipsilateral projection is limited to the caudal portion of the nucleus is aligned with the corresponding portion of the contralateral projection.

Because these maps shift and begin to come into register as laminar segregation is occurring, the maps must "shift through each other" during this portion of development. This phenomenon demonstrates that axonal arbors in the dLGN are capable of adjusting relative to other arbors without requiring the cell to prune an initially exuberant arbor down to a final shape. Jeffery postulates that the two maps first come into register at a single point, the caudal pole of the nucleus, when the temporal portion of the contralateral projection dies out. This loss brings the most caudal portion of both the ipsilateral and the contralateral projection into registration — i.e. they then represent the same point in visual space. After this point Jeffery postulates that the maps seek registration at all points, shrinking the ipsilateral projection down to the only portion of the contralateral projection to

which it it corresponds and leaving the lateral portion of the ferret dLGN with purely contralateral innervation (compare Figs.1 and 11 in Jeffery, 1990).

To the extent that this shifting occurs after eye opening, activity could play a role in aligning the maps from the two eyes. Before that point activity can not help align the layers, because the retinal activity at corresponding points in the two retinas is not especially correlated.

## 3.4 Physiological Development

Activity-dependent processes are qualitatively different than mechanical or chemical mechanisms because they are founded in the transient electrical activity of the cells rather than anatomy. Key issues about activity-dependent mechanisms are: whether there is activity present at the right time and in the right form to effect developing projections, whether the activity that is present is propagated between tissues in a way that could support the development of useful structure, and whether the immature dLGN displays long-term potentiation (LTP) or other Hebbian processes.

### *3.4.1 Immature dLGN Physiology*

In order for activity-dependent mechanisms to be at work in the retinogeniculate projection, the immature dLGN must have two physiological properties: responsiveness to retinal input and LTP or a similar mechanism of synaptic change. Both of these properties have been found in the developing dLGN, as described below. Because the details of the dLGN's response to retinal input are central to the present model, they will be discussed at length.

*Retinogeniculate Transmission.* The earliest functional retinogeniculate connection is D39 (Shatz and Kirkwood, 1984), well before the period of laminar segregation. Recent studies have shown that bursting excitation during the first post-natal weeks is transmitted to the dLGN and may even have a wave-like character there (Penn et al., 1995; Mooney et al., 1995).

The influence of this retinal input is potentiated by several factors, including the lack of spindle waves in dLGN relay cells before ∼D65 (McCormick et al., 1995). The ability of dLGN relay cells to fire $CA^{2+}$ spikes also doesn't appear before ∼D62 (Ramoa and McCormick, 1994a). The dLGN relay cells are thus stuck in the "tonic" state during this period of development. The absence of these forms of intrinsic dLGN activity presumably supports the transfer and use of retinal excitation, and is likely linked to immature inhibitory circuitry.

*Attenuated GABA inhibition.* Anatomical and physiological evidence indicates that GABA inhibition in the dLGN is highly attenuated during the layer and sublayer segregation. dLGN tissue stained with antibodies for GABA and GABAergic enzymes (glutamic acid decarboxylase, GAD) does not show detectable staining in the A layers until the end of eye-specific layer segregation

(D59) in the cat (Shotwell et al., 1986). Even at D59 the distribution of staining is far weaker than at maturity — near mature levels of staining aren't seen until 2 months after birth. The presence of staining at glomerular terminals is only present after the first postnatal month. Likewise, at the post-synaptic membrane, the number of GABA receptors on glomeruli increases 10-20 fold from P0 to maturity (Friedlander and Tootle, 1990). Ultrastructural analysis of the glomerular structures supports the conclusion that glomerular structures are very immature at birth and only begin to take on their distinctive structure after after the third postnatal week (Mason, 1982).

These findings are in accordance with the physiological finding that GABA mediated functions in the dLGN mature largely after the focal period of refinement. In the cat, inhibition has a measurable modulatory effect on dLGN relay cell response only after D59 in the cat, at a time when segregation of afferents is nearly complete. The response of dLGN relay cells to GABAa stimulation emerges at a similar time, becoming measurable about 10 days before eye opening and slowly maturing after that point (Friedlander and Tootle, 1990; Shatz and Kirkwood, 1984; Ramoa and McCormick, 1994b).

Both GABAa and GABAb mediated IPSPs are absent at birth in the ferret. GABAa mediated IPSPs mature gradually from birth to ∼P21, while GABAb mediated IPSPs don't mature until after P21 (Ramoa and McCormick, 1994b; White and Sur, 1992). Interestingly, it is possible to generate recurrent IPSPs by stimulation of the RNT/PGN well before IPSPs can be evoked by feed-forward stimulation (Ramoa and McCormick, 1994b), thus GABA receptors seem to be operational before they are used by the feed-forward sensory circuits.

*Heightened NMDA excitation.* The dLGN relay cells during the period of segregation actually respond especially easily due to more efficacious NMDA receptor function, inhibited burst firing, attenuated resting potential (Ramoa and McCormick, 1994a), longer action potential durations (4-5 times longer), and longer EPSPs (∼2 orders of magnitude longer, Ramoa and McCormick, 1994b). These effects are due partly to the lengthened membrane time constant and NMDA generated EPSC duration in immature dLGN relay cells, but they are more dramatically effected by the initial lack of GABA mediated IPSPs (Ramoa and McCormick, 1994b).

Because the EPSPs are untruncated by GABA inhibition, a 25Hz excitatory input (the firing rate of an RGC at the peak of a retinal wave) will cumulatively depolarize an immature dLGN relay cell, but not a mature cell. Further, in an immature cell a weak depolarization is more likely to generate an action potential because the membrane has a lower resting potential, and is closer to its firing threshold.

Interestingly, the NMDA component of the excitatory response is eventually overshadowed by other excitatory channels. Blocking the NMDA receptors eliminates this EPSP summation in the first 2-3 weeks of development, but after D65 NMDA blockade has little effect on the responsivity of dLGN relay cells (Ramoa and McCormick, 1994b).

*Receptive Field Development.* dLGN relay cells have weak inhibitory surrounds and large receptive fields prior to ∼D65 in cats (i.e. the surround strength is only about 50% of its adult magnitude by D62, Friedlander and Tootle, 1990). There is evidence suggesting that the lack of inhibitory surround is due to immaturity in the inhibitory pathways in the dLGN. Berardi and Morrone (1984) tested the effects of application of GABA and a GABA agonist (bicuculline) on the responses of X-type dLGN relay cell responses to sine gratings. Although GABA application inhibited the cell's response, the application of the bicuculline had no measurable effect on the cell's response properties before D95 in cats. Bicuculline has easily measurable effects on the response properties of mature dLGN dLGN relay cells. This finding suggests that even after the inhibitory pathways are largely in place, they are not functioning to shape dLGN relay cell response until the second post-natal month.

Other researchers have found that spatial resolving power and surround inhibition in the dLGN change dramatically around D80. At this point the accentuation of the inhibitory surround seen in the dLGN starts to reach mature levels (Ikeda and Tremain, 1978; Daniels et al., 1978).

### *3.4.2 LTP*

The immature dLGN exhibits LTP (Mooney et al., 1993), perhaps even heightened LTP due to heightened NMDA channel efficacy compared to the mature dLGN (Ramoa and McCormick, 1994b). This suggests that during the period of laminar segregation, the dLGN has heightened responsivity to retinal signals and the ability to use those signals to guide refinement of the retinogeniculate projection.

### *3.4.3 RGC Spontaneous Activity*

It has been found that the developing retina is active in a surprisingly organized fashion. This activity seems well-suited to support activity-dependent refinement of the retinogeniculate projection. Because the activity is highly systematic and correlated within each eye, and probably not correlated between the eyes, these patterns of activity can support eye-specific segregation of afferents in the dGLN. Rough calculations also suggest that the waves could support activity-driven refinement of the retinotopic projection to adult levels of precision.

Using arrays of micro-electrodes (Meister et al., 1991; Meister et al., 1994; Wong et al., 1993), and voltage-sensitive dyes (Wong et al., 1995; Feller et al., 1996), researchers have been able to record activity in the immature ferret retina. They have found this activity to be quite regular and systematic — essentially, waves of high firing rates, or bursting, travel across the retina. These waves propagate through the ganglion (and amacrine) cell layer during the period from D43 (at latest) up to D64. This activity is independent of light, in fact the photoreceptors are not even functional during this period. The waves disappear ∼D73 just before eye opening (D75). The adult retina is also spontaneously active, but only in the dark, and not in such an organized manner (Wong et al., 1993). Importantly, the timing of this ordered activity coincides with the time that eye-specific

layers of the dLGN are differentiating and perhaps when retinotopic maps are being refined in the dLGN (Wong et al., 1993).

The waves are bands of ganglion cells firing in sustained bursts at the crest, with neural firing rates falling off with distance from the crest. Waves generally occur one at a time, staggered by quiet periods of 30 to 60 seconds (Wong et al., 1993). Waves become broader and more frequent over development, but they move at the same speed regardless of age (Wong et al., 1993). These waves only include transient subdomains of the retina (e.g. a large one might cover a third of the retina), they move quite slowly (around 1% to 10% the speed of normal synaptic transmission, 100-300 $\frac{\mu m}{sec}$ with bursts of speed up to 500 $\frac{\mu m}{sec}$), and they start from random points and propagate in random directions (Wong et al., 1993; Feller et al., 1996). Investigations into the anatomical basis for these waves suggests that they propagate through synaptic transmission, and that cholinergic amacrine cells are involved in wave propagation (Feller et al., 1996).

*Retinotopy.* (Wong et al., 1993) made calculations showing that these waves have the right temporal properties to support long-term potentiation (LTP) in the dLGN. Studies in the hippocampus indicate that spikes arriving within .05s of each other ("coincident spikes"), can potentiate LTP. During the time of wave activity, nearby cells have up to a $100\times$ the number of coincident spikes than would be expected at random (P0, P21: $20\times$, P30: $4\times$, adult: $1\times$). The number of coincident spikes declines exponentially in proportion to distance at all ages. Roughly, at birth (D43) ganglion cells separated by 200 $\mu m$ have spikes which are less than half as strongly correlated as cells that are neighbors. There is still a relatively high correlation between many of the cells separated by up to 600 $\mu m$ (average of 5 times as correlated as random cells from D43-D65, but the authors postulate that correlations drop to random around 1 mm).

Because the waves propagate in different directions, and can be oriented in any direction, they can serve to cluster the axons of RGCs at relatively distant points (i.e. the length of the wavefront) in the retina. The extent to which the waves may aid in the refinement of retinotopy requires the converse calculation — i.e. what is the minimum distance at which RGC's are differentiated by their activity patterns. For example, a very wide wave would do little to differentially activate neighboring RGC's and could not support the segregation of their afferents.

In concrete terms, the overall width of waves is about 200-400 $\mu m$ (Meister et al., 1994); this corresponds to a width of approximately 5-15 RGC's (calculated from RGC densities reported in (Sanderson, 1971). (Wong et al., 1993) suggests that the distance over which the number of coincident firings drops by 50% is a plausible lower bound on this minimally differentiable distance. This bound roughly doubles from 170 to 300 $\mu m$ from D43-D48 and then stays constant through D73. In conjunction, the retina expands during this time (D48-D73) reducing the density of cells in the mid-peripheral retina by 50%. The number of cells that are within the crucial distance for supporting LTP thus goes down by 50%, effectively sharpening the extent of the waves in terms of RGC's even though the waves are physically wider (Wong et al., 1993). Determining whether

this degree of precision is sufficient to sharpen the projection to adult levels of retinotopy requires further calculations.

$1^\circ$ of visual angle is approximately 130 $\mu m$ in a D79 cat retina (Friedlander and Tootle, 1990), yielding a lower bound on sharpening is approximately $2.5^\circ$ ($300\mu m/130\frac{\mu m}{\text{deg}}$). Thus, wave activity can differentially activate RGC's that are $2.5^\circ$ apart. The RGC axonal arbors in the dLGN overlap those of other RGC's from $2^\circ$ (area centralis) to $11^\circ$ distant (periphery). Similarly, the scatter in receptive fields of cells that have RF centers at the same point in visual space have an overall scatter of 2-$10^\circ$ (area centralis to periphery), with the standard deviation of the scatter being .5-$2^\circ$ (Sanderson, 1971).[7] Given that the recordings of retinas were done between the blind spot and the periphery (Meister et al., 1991; Wong et al., 1993), they likely are between 5-$20^\circ$ eccentricity in the retina. Therefore, in the adult the minimal amount of visual distance one can shift and be sure of shifting laterally in terms of dLGN cells is approximately 3-$5^\circ$. By these calculations, the waves could serve to refine retinotopy to adult levels in the dLGN.

*Laminar Segregation.* Wave activity generates coincident spikes within a retina but not between the two retinas. Because the waves are passed by synaptic activity, wave activity in the two eyes is probably independent. In addition, the short duration of a given wave (4-5 s) and the relatively long quiet periods (40-60 s) between waves insure that most of the time when there is a wave in one retina the other retina will be silent. These factors will strongly support the emergence of eye-specific clustering of afferents in the retinogeniculate projection. Further, there is some correlation information in the wave activity that could serve to support function class specific segregation.

During the period of on/off sublayer segregation in the dLGN are segregating (D55-67), the on and off-RGC firing patterns differentiate. Before this period, the on and off afferents act homogeneously. The resulting correlations between cell firing are independent of on/off-cell characterizations. During the period in which the afferents are segregating the on and off inputs become less correlated in their activity patterns. The shift is rooted in the fact that the off cells tend to burst much more frequently than the on cells (i.e. 3-4$\times$ as much). Up to 75% of the time when the off cells are firing the on cells in the same patch of retina are silent. The off cells also become somewhat de-correlated with each other at this point, suggesting that their firing becomes more "noisy" during this period (Wong and Oakley, 1996). The firing of on cells remains highly correlated, and they will still have a higher proportion of their spikes be coincident with other on cells than with off cells.

[7] These figures are consistent with some figures about RGC arborization and the degree of visual magnification in the dLGN. The width of RGC arbors is 100-200 $\mu m$ around this time (Sretavan and Shatz, 1986; Sretavan and Shatz, 1984), and one degree of visual angle corresponds to 600-50 $\mu m$ (area centralis to the periphery). Consequently, two RGC's with barely touching arbors would be separated by 200-400 $\mu m$, .3 to $8^\circ$ . These separations are approximately the same magnitude as the actual distances between cells that have overlapping arbors.

## 3.5 Summary of Relevant Biological Background

Ferrets are born ∼D41 (cats, ∼D65), and they open their eyes 4 weeks later (∼ D73). The RGC axons invade the dLGN before birth (Sec. 3.3.2). The majority of the RGC axons cross at the optic chiasm and innervate the contralateral dLGN (Sec. 2.2). RGC axons innervate the entire contralateral dLGN, and the binocular portion of the ipsilateral dLGN (Sec. 2.3.1). Individual arbors are restricted in width throughout development (Sec. 3.3.3), but they initially innervate and form synapses throughout the depth of the dLGN (Sec. 3.3.3 and Fig. 4.1). The position of the RGC axons is known to be very coarsely retinotopic, but the initial extent and subsequent refinement of precise retinotopy is unknown (Sec. 3.3.5). In the mature animal the receptive fields of the dLGN neurons are arranged in a very precise retinotopic map (Sec. 2.3.2).

In the 3 weeks after birth, the RGC axons segregate into 4 sublayers based, at least partially, on activity-dependent mechanisms. During the first two post-natal weeks (D41-D55) the axons from the contralateral and ipsi-lateral eyes segregate into layer A and A1 respectively (Sec. 3.3.6). During the third post-natal week the axons from the on- and off-center RGCs segregate into separate sublayers within layers A and A1 (Sec. 3.3.6). By eye-opening, each sublayer (and each dLGN relay cell) receives inputs from either the on or off-center cells from one of the eyes.

During the period from birth to eye-opening there is spontaneous activity in the retinas that propagates spatially in wave-like patterns (Sec. 3.4.3). During the first two weeks, this activity is correlated within each eye, but independent between the eyes. Around the end of the second post-natal week the on- and off-center cells within each eye begin to have more independent patterns of activity as well. Pharmacological blockade and enucleation studies suggest that this spontaneous retinal activity is crucial to proper development of layers (Sec. 3.3.7) and sublayers (Sec. 3.3.8) in the dLGN.

During this period of refinement, GABAergic pathways in the dLGN are immature and largely ineffective (Sec. 3.4.1). This lack of inhibition is thought to heighten the dLGN response to retinal inputs during this period (Sec. 3.4.1).

# Chapter 4

# The Model

The previous section presented a detailed account of the development of the retinogeniculate projection in cats and ferrets. The model presented in this section necessarily simplifies this picture of development. It outlines a set of mechanisms that, in conjunction with simulated wave activity, is capable of supporting three of the most striking aspects of retinogeniculate refinement: eye-specific layer segregation, on/off sublayer segregation and the formation of precise retinotopy. This model is the first to support the development of all three of these structures in a unified framework.

The model is proposed as a coherent set of mechanisms that are capable of supporting some crucial features of the observed development. Undoubtedly, this model is inaccurate in some respects. However, it takes steps toward a formal and biologically detailed model of this system. The processes in the model are intended to be biologically interpretable and plausible, and the model dynamics are largely tractable to formal analysis.

This chapter is organized into three main sections. While each section covers the whole model framework, the level of detail increases significantly with each section.

## 4.1 Overview

Because of computational limitations, the model is formulated on a much smaller scale than the entire retinal surface and the entire dLGN. It simulates the refinement of the axons projecting from two patches of retina (one from each eye) to a single "column" of the dLGN. The scale in both cases is on the order of a millimeter (see Fig. 4.1 and Fig. 4.2). The model is also limited in its time-course. Development is simulated from "birth" (D41, in ferrets) through the first 3 postnatal weeks (i.e. 1 week before eye-opening).

Although the developmental processes occur concurrently, we iterate through a four step algorithm for computational convenience:

1. Compute dLGN activity in response to a "snapshot" presentation of a retinal wave

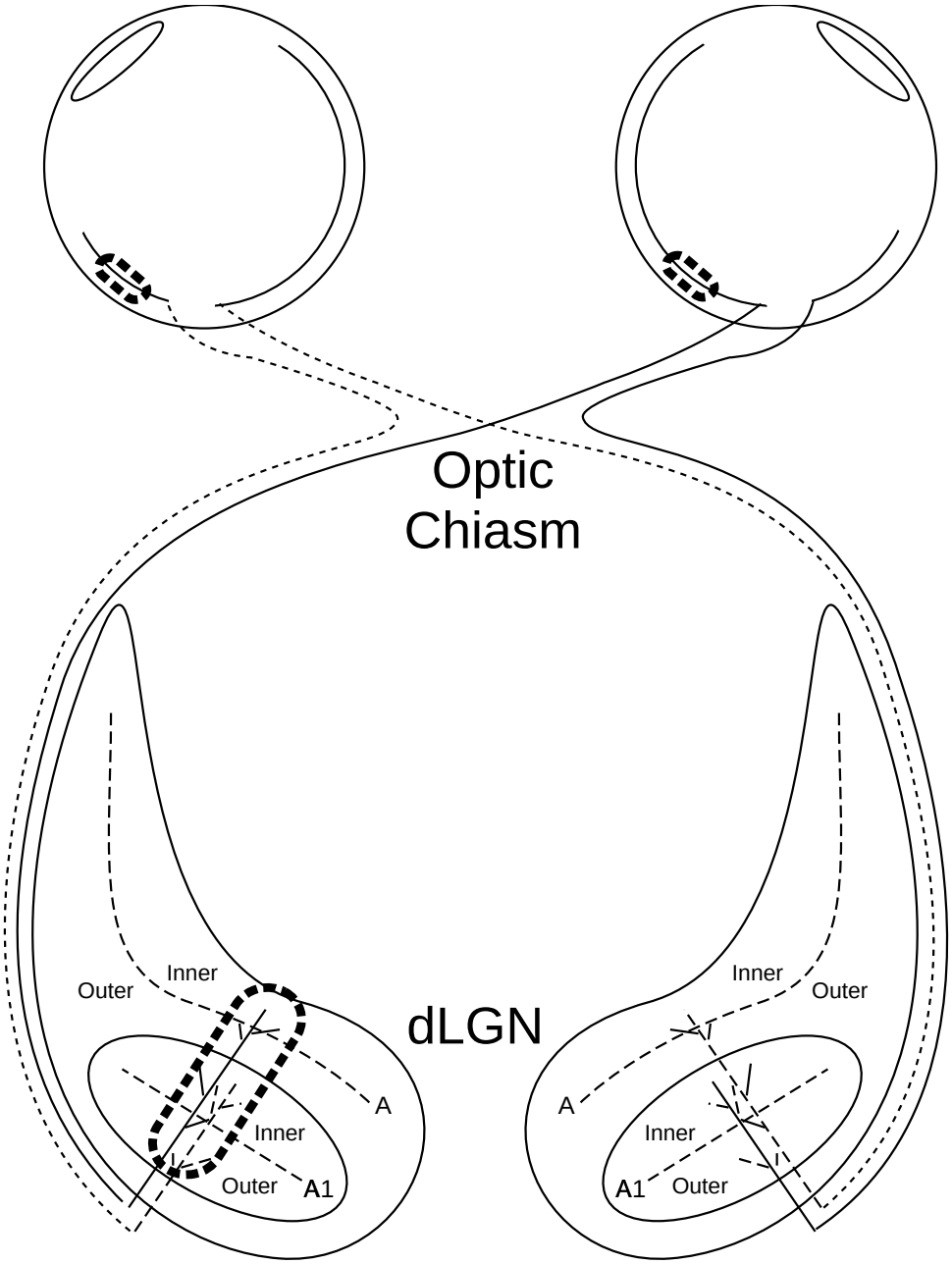


Figure 4.1: A schematic drawing of the retinogeniculate system of the ferret. The heavy dotted lines surround the areas of interest. Two small patches of retina are modeled, with two distinct cell types (on and off-center) from each eye. The patches project to the same "column" of dLGN tissue. That column is located in the binocular region of the dLGN, such that all 4 sublayers are manifest in the column at maturity.

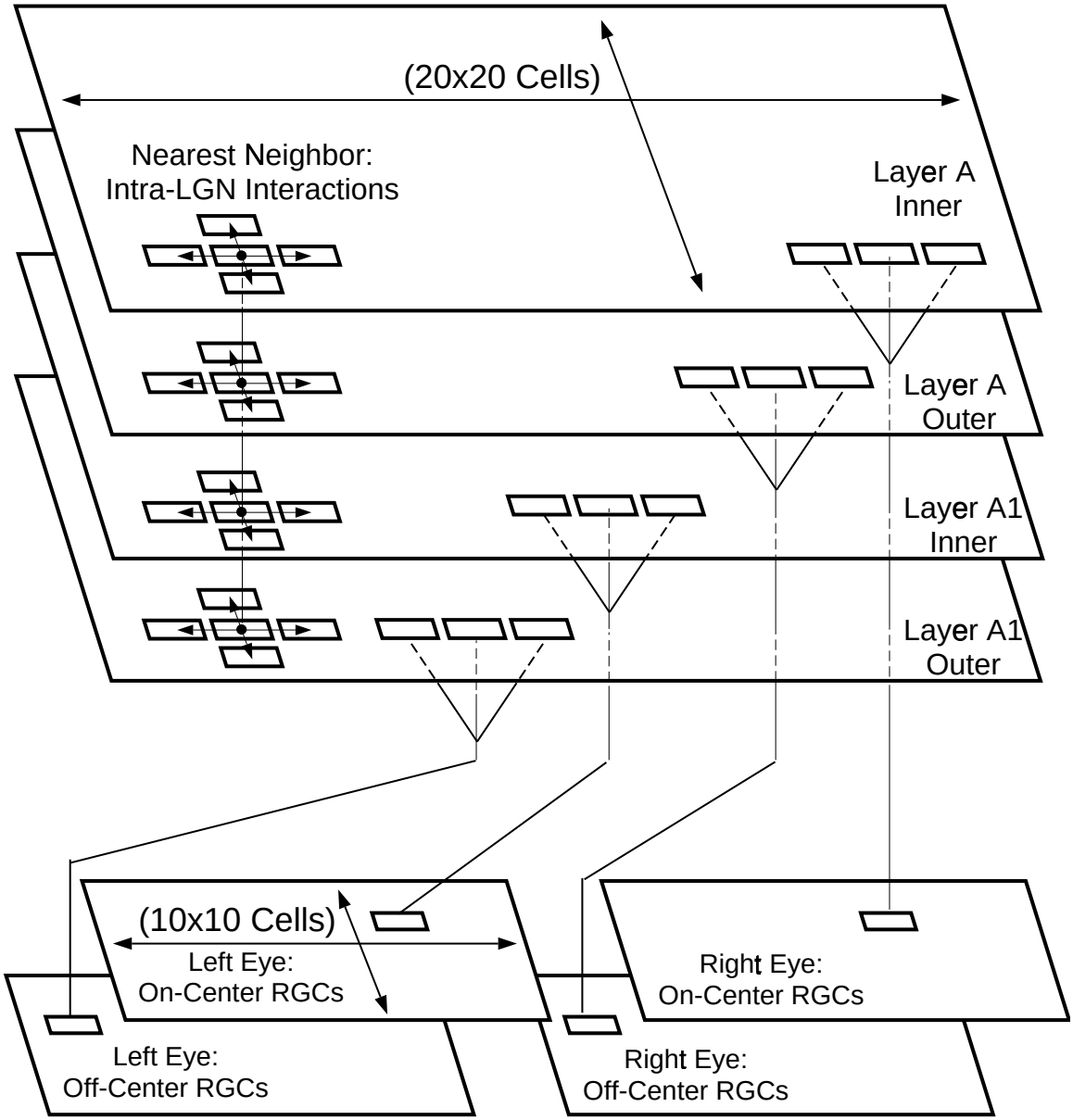


Figure 4.2: A schematic drawing of the model architecture after segregation into layers and sublayers. Four patches of RGCs project to four sublayers of dLGN cells. In all of the models discussed in the results section, the dLGN is 20 cells square, and the retinal patches are 10 cells square. The one dimensional models are like a slice through the two dimensional model, such that the retinal strips are 1x10 cells and the dLGN strips are 1x20 cells. Each non-border dLGN cell connects to its 4 nearest intra-laminar neighbors, and with less strength, to its 2 nearest inter-laminar neighbors. Boundaries do not wrap around, but are treated specially during the computation of steady-state dLGN activities (see Sec. 4.3.4).

2. Hebbian weight update
3. Synaptic sprouting
4. Weight normalization and synaptic retraction

This algorithm is repeated for tens of thousands of iterations, gradually refining the projection of RGC's to the dLGN relay cells.

## 4.2 Brief Summary

*Initialization.* The model is initialized with each RGC arbor innervating all future layers of the dLGN. In contrast to many self-organizing models, the arbors have roughly the same width throughout development, consonant with the biological observations (see Sec. 3.3.3). The model is initialized with some bias toward a particular retinotopic alignment and a particular layer configuration. The bias makes the model converge more robustly and supports a stereotypical final state as seen in the biological system.

*Inputs and dLGN Update.* The retinal waves in step 1 are idealizations of the retinal waves seen in pre-visual ferret and cat retinas (see Sec. 3.4.3). Because they move slowly relative to synaptic transmission speeds, the model assumes that they can be considered as static ("snapshots") until the dLGN activity reaches steady state. The dLGN reaches a steady state based on its synaptically weighted retinal inputs and the activity of its neighbors. The feedback of activity between the dLGN neurons effectively blurs the retinal input. The propagation of activity in the dLGN is formalized so that the steady state dLGN activities for a given retinal input can be calculated in one step.

*Weight Update.* Given the dLGN activity and the retinal activity, computing the weight change is straightforward. The change in each synaptic strength is proportional to the product of the pre-synaptic (RGC) activity and the post-synaptic (dLGN cell) activity. The weight change is scaled by the learning rate and by the synapse's previous strength/weight. The update is a slightly modified version of the Hebb rule, where weights between co-active cells tend to grow, but that growth isn't explicitly scaled by the previous weight.

*Sprouting.* The width of RGC axonal arbors remains relatively stable over the course of development, yet the arbors shift between layers and within maps over development (see Sec. 3.3.9). In this model, sprouting and retraction processes reconcile these two observations. The assumption is that the arbors continuously sprout new branches and synapses, and simultaneously cull other branches and synapses. These mechanisms maintain the arbor width in a state of dynamic stability, while allowing the arbor to shift appreciably over time. Computationally, sprouting in the model

is accomplished by spatially blurring the weight matrix corresponding to each ganglion cell arbor. While this approach doesn't capture directed sprouting (e.g. via trophic factors), it does allow directional movement over time, as evidenced by the convergence behavior of the model.

*Normalization.* In the Hebbian update step of this algorithm, all the weight changes are positive. In order to prevent infinite weight growth, some weights need to be decreased. A simplistic view would be that synaptic changes that are below the mean positive change are decreased, while above average changes remain as increases — a process known as "synaptic competition".

In this model, the resources of both RGC and dLGN cells are assumed to be approximately fixed. In other words, the synaptic weights are constrained such that the total synaptic weight supported by any given cell in the model remains roughly constant during a simulation. This constraint is enforced multiplicatively, which has important ramifications for the model behavior (for discussion and an alternate "subtractive" model see Sec. 6.2).

Given the updated, sprouted and normalized synaptic weights, the whole process is repeated with the next snapshot of a wave as it moves across the retina. The details of the algorithm are crucial to its behavior, so they are presented at length in the following section.

## 4.3 Detailed Description

### *4.3.1 Activity in the Model*

Activity in model neurons generally corresponds to instantaneous firing rates. In real neurons, the instantaneous firing rate can be estimated by averaging the response (spike train) over many measurements, assuming that spike generation is a random process. Under the assumption of independent spiking, simulated spike trains can be generated from the instantaneous firing rates by a Poisson process.

In order to link the dynamics of the model dLGN directly to biophysical properties of the dLGN relay cells, the dLGN activities can also be interpreted as membrane potentials (Sec. 4.3.4). The presumed equivalence of membrane potentials and firing rates assumes that the firing rate is linearly proportional to the cell's thresholded membrane potential, where the threshold is the resting potential.

Of course, these assumptions ignore many details of spike generation and are surely inaccurate accounts of the relations between membrane potentials, spikes, and firing rates. Nevertheless, they are commonly used approximations (Movshon et al., 1978) and are probably sufficient for the level of analysis in this model.

| | |
|---|---|
| RGC Separation | $\sim 30\mu m$ |
| Speed | 200 $\mu m$/sec |
| Width | Gaussian ($\sigma = 75\mu m$, thus $\sim 300\mu m$ across ) |
| Height | Peak Firing Rate = 25 spikes/sec |
| Refractory Period | $80 \pm 16 sec$ |

Table 4.1: Some basic parameters of the simulated waves, matched to the in vitro measurements of retinal waves (Wong et al., 1993)

### *4.3.2 Inputs*

Because the model is formulated at the level of individual cells, matching input parameters to in vitro anatomical and physiological measurements is relatively straightforward. A wave is modeled as a bar-like section of a spreading 2-d ring with Gaussian cross-section. The center of the ring is located at a random point on a large circle surrounding the retinal patch (radius = 100 RGCs). The random orientation of the wave-front ensures any two cells in a 1-d retinal patch can be co-active. Consequently, the correlation between distant points within an eye are non-negative (in contrast see Sec. 7.2.2 and Eglen, 1997).[1] The width, frequency, height, and speed of the waves matches the reported retinal measurements (Wong et al., 1993, see Table 4.1). Two representative waves are illustrated in Fig. 4.3.

Both on and off RGCs are indistinguishable based on activity for the first "2 weeks" of model development, but bursting noise is added to the off cell population during "week 3". This change brings the bursting rates and correlational structure of the input into approximate agreement with in vitro measurements of retinal activity during on/off sublayer segregation (Wong and Oakley, 1996 and see Section 3.4.3).

*D41-D55: Eye-Segregation.* The realism of the inputs was checked by calculating the same statistics for the simulated waves that were calculated for the in vitro waves. Wong et. al. (1993) constructed a measure of the correlation in firing between cells, and calculated its relation to the separation between those cells. Their correlation index quantifies the extent to which two cells fire action potentials within .05s of each other, relative to the baseline co-occurrence of action potentials among independent cells.

I calculated this index for the inputs in my model during the period corresponding to D41-D55 (i.e. birth to the onset of differentiated on/off activity). Because the index is calculated from spike trains, I used a Poisson model of spike generation to generate spike trains from the instantaneous

[1] Given that co-occurrence of waves in the same retina has been observed (Feller, personal communication) and, to my knowledge, there is no evidence to suggest that distant cells within a retina do not fire independently. In the current model, distant RGCs in the same retina fire independently.

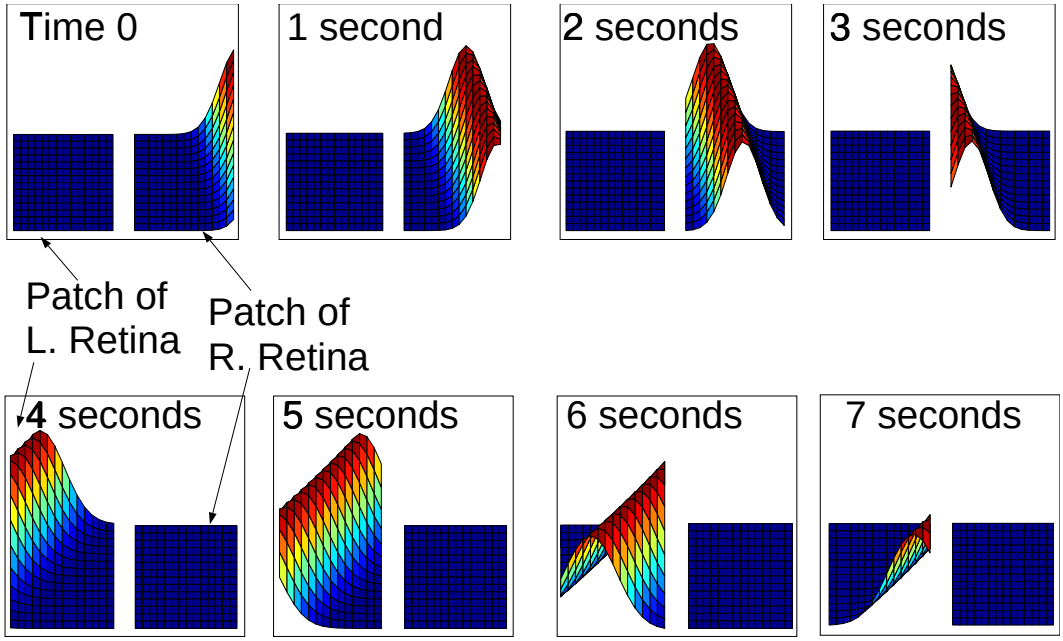


Figure 4.3: 8 snapshots of simulated wave activity at 1 second intervals are shown. Waves do not normally follow one another immediately, because the retinas are in a silent refractory periods for the majority of the time. Note that the waves originate and propagate in random directions, and that they tend not to occur in the two (right/left retina) patches simultaneously.

firing rates used in my simulations. I followed the Wong et. al. procedure for calculating the correlation index from the spike train. The correlation index is defined as:

$$\frac{N_{AB(-0.05s,+0.05s)}T}{N_A N_B(0.1s)} \tag{4.1}$$

where $T$ is the total duration of the measurement, $N_{AB(-0.05s,+0.05s)}$ is the count of spikes co-occurring in cell $A$ and cell $B$ within $0.05s$ of each other from time 0 to time $T$, $N_A$ is the total number of spikes in cell $A$, and $N_B$ is the total number of spikes in cell $B$.

This index provides a measure of how often a pair of cells fires together relative to chance. For example, a pair of cells with a correlation index of 10, fire together 10 times more often than a pair of cells that only fire together by chance (correlation index = 1).

The correlational structure of the simulated waves and observed waves is compared in Fig. 4.4. Following Wong et. al. (1993, figures 8 and 9), the correlation indices calculated from the observed ferret retinas are plotted as a function of distance between the cells. The same data from the simulated waves is also plotted. The plots demonstrate that the simulated waves have roughly the same correlational structure (in addition to specific parameters) as the observed waves. Two discrepancies between the observed and simulated waves are notable: the simulated waves have higher correlation

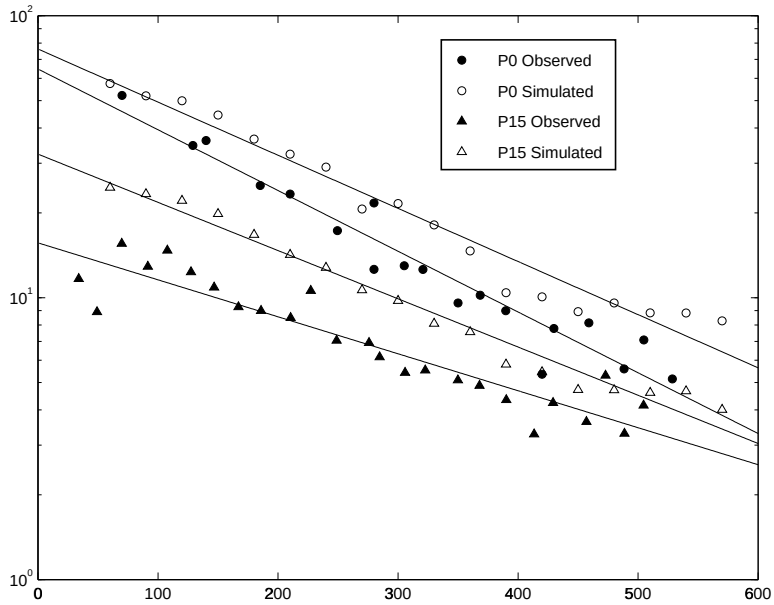


Figure 4.4: The correlation index between pairs of cells plotted as a function of distance between the cells. Exponential fits to each set of data are plotted as well. The correlation indices calculated in Wong et al., 1993 (closed symbols) are shown in addition to correlation indices calculated for the simulated waves (open symbols). Simulated and observed data are plotted at birth (P0) and at the onset of on/off activity (P15). The indices indicate how often a pair of cells will fire together relative to chance (see text). The correlation index seems to decrease exponentially over distance.

indices than the measured waves (by a factor of $\sim$3 in neighboring RGCs at P15), and the observed waves change over development.

The retinal waves change over the course of development in ways that are not captured in the current model. For example, the waves are thought to broaden and degenerate gradually over development until they eventually dissolve into uniform background noise around birth. In contrast, the simulated waves are generated from static parameters that are modeled on the P0 observations. The only change in the simulated waves is the addition of off-cell noisy bursting at the end of the second week. As is evidenced in the plots of the P15 correlation indices in Fig. 4.4, this change suffices to significantly lower the correlation indices between nearby cells, but does not change the slope of the spatial decrease (as in the observed waves). The correlations are lowered because the off-cells and on-cells are interspersed, and the noise in the off-cell firing lowers the average correlations between pairs of cells at all distances.

The higher correlation indices in the simulated waves are probably related to at least three factors: the static wave parameters mentioned above, the equivalence of model retinal patches and single domains, and the unrealistically “clean” simulated waves. As discussed in Sec. 3.4.3 the propagation of the in vitro waves is limited by the boundaries of “domains” of the retina, which transiently shift over time. Because each wave includes all of the cells in the relevant retinal patch,

the implicit assumption is that a retinal patch is never on the border between domains. The assumption that each patch is always in the same transient domain tends to bias the activity toward higher correlations than are seen in the actual retina. The unrealistically cohesive quality of the simulated waves also would tend to amplify the correlation indices. In the retina some RGCs fail to fire, even if a wave passes directly over them (see Figs. 2 and 3 in Wong et al., 1993). All other wave parameters being equal, this "noisy" aspect of the waves would tend to lower correlations between pairs of cells at all distances. It is possible that the amplified correlations provide unrealistically strong support to model convergence, but Sec. 5.5 suggests that the exact strength of the correlations is not crucial as long as the spatial correlational structure is kept intact.

*D56-D62: On/Off Sublayer Segregation.* Noise is added to the off-RGC populations during the period of on/off sublayer segregation so as to increase the off-RGC firing rates to the levels observed in vitro and to decrease the pairwise correlations between off-RGCs and both on and off-RGCs. Wong and Oakley (1996) calculated the correlations for the waves in the ferret retina by converting their $\sim$1Hz measurements of $CA^{+2}$ levels in the RGCs into dichotomous bursting "events" by thresholding the derivatives of the levels, and then calculating the correlations between trains of these events. I compare these measurements with the the raw firing rates from the simulated waves, as there was no straight-forward translation from their methodology to the simulations. The correlations between on-on RGCs, off-off RGCs and on-off RGCS are plotted following Wong and Oakley (Figure 7B), in Fig. 4.5.

There are two basic differences between the measured correlations and the correlations in the simulated waves. The first is that the on/on correlations decrease slightly in the in vitro measurements, while the simulated correlations remain at the same level. The discrepancy is due to the the general lowering of correlations over development in the ferret, compared the simulated wave structure which remains fixed over development. This increased correlation may bias the model slightly toward grouping the on- cells in one sublayer.

The second discrepancy, the lower off/off correlations found in the model relative to in vitro measurements, would provide unrealistically weak support to the grouping of off-cell arbors in a single sublayer. While the cause of this difference is not immediately obvious, there are some likely explanations. In the model, independent noise was added to the off-RGCs until their bursting rate was in line with in vitro measurements ($\sim$4 times the bursting rate of the on- cells). Perhaps the non-wave bursting in the ferret off-RGCs is not independent. If the in-vitro "noise" is spatially correlated, such that bursting in an off-RGC is more likely if a nearby off-RGC is bursting, the off/off correlations would be higher than those used in the model.

*Wave Subdomains.* In most of the simulations each wave traverses the entire patch in that eye. Because each in vitro wave only involves transient subdomains of the retina (see Sec. 3.4.3), the patch is implicitly assumed to consistently be included in the same transient subdomain. Even if the assumption held true, the results of the model are not necessarily applicable to the clustering of

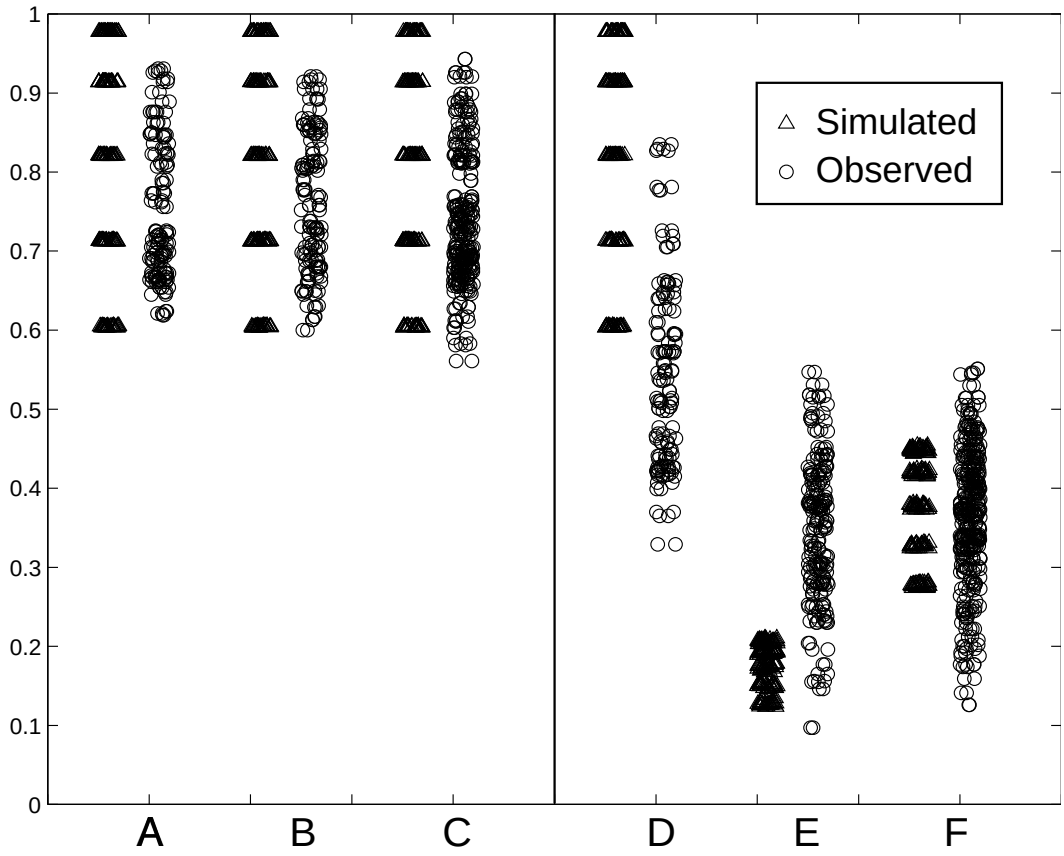


Figure 4.5: The correlations between in the RGC cell groups during the first two weeks of development (A-C), and the third week of development (D-E). (A,D) on-on RGCs, (B,E) off-off RGCs (C,F) on-off RGCs.

the axons of neighboring retinal patches into the same sublayer, because these neighboring patches would only sometimes be involved in the same wave. Indeed, because distant RGCs would *never* be involved in the same wave, this assumption could strongly limit interpretation of the results of the model. Fortunately, the characteristics of the waves suggest a computational method for testing the generality of the model.

In the ferret, it is likely that the transient aspect of the subdomains plays an important role in supporting global segregation. Because the domains are transient, neighbors will be correlated as a function of their distance, rather than being strongly influenced by their sub-domain membership. The development of global segregation is akin to development of global retinotopy, it results from the chaining of local influences. Because each RGC's arbor tends to segregate into the same layer as co-active cells (e.g. neighbors), the configuration that optimally satisfies these local tendencies is global segregation.

To test the strength of this chaining effect, simulations were performed in which each retinal patch was taken to be chronically on the shifting border of two subdomains. Each wave included

some, but not all, of the patch, and the most distant RGCs in the patch were never involved in the same wave (i.e. their activities are negatively correlated). Nonetheless, the patch segregated normally, presumably because of these chaining processes (see Sec. 5.5 and Fig. 5.7).

### *4.3.3 Initialization*

The model is initialized to mimic the projection at birth from a small retinal patch from each eye to a column of dLGN tissue. A retinal patch from a given eye actually represents two intermingled populations of RGCs, on- and off-center cells; there are 4 groups of distinct RGCs total. These RGCs are assumed to project to a 4 layer deep column of the dLGN (each simulated layer is only 1 neuron deep). In rough concordance with the magnification factor in the ferret and cat dLGN, there are 4 dLGN cells for every RGC in the 2-d version of the model (see Fig. 4.2).

*Retinotopy* Each RGC arbor is initialized as a Gaussian (s.d. $\approx 38\mu m$) that is centered on a retinotopically biased position in the dLGN. The location of the Gaussian is initially the same in all layers. Noise is added to this initial projection in two ways. First, the arbor centers are randomized by computing a weighted average of the retinotopically correct position and a random position in dLGN (typical weightings were .2 and .8 respectively, "80% noise"). Second, each synaptic weight is multiplied by a random number (typically [.2 1]). Multiplication ensured that arbors remained relatively compact while breaking symmetry between layers and adding randomness.

*Layer and sublayer Segregation* Biases are added to this original innervation in order to support the development of a stereotypical mature layer configuration. While the model can develop eye-specific layers, retinotopy, and on/off sublayers without these biases (see Sec. 5.8) — the consistency in the mature configuration of the biological system suggests that there are significant biases at work. A mechanism that may provide a bias toward the stereotypical eye-specific segregation pattern is the early ingrowth of contra-lateral RGC axons in the dLGN (see Sec. 3.3.6). Early ingrowth gives these afferents a competitive advantage in the deepest (i.e. A) layers of the dLGN, perhaps even preventing the a full initial innervation into these layers by the ipsi-lateral afferents. This ingrowth bias is simulated in the initial projection by weakening the ipsi-lateral projection to the deepest layers by reducing the strength of ipsilateral synapses in the A layers by $\sim 20\%$.

There is no evidence for similar staggered ingrowth of the on/off afferents into the inner and outer sublayers, but the stereotypical pattern of segregation in the adult suggests that there is some bias at work. The canonical form of this model assumes that this bias is instantiated as a very small strengthening of the weights from the on-center cells to the inner sublayer's at each time-step.[2] The intention is to simulate the action of trophic factors that preferentially support the formation of

[2]The bias is much too small to support afferent segregation without appropriately structured retinal activity, see Sec. 5.7.

| | |
|---|---|
| $l_{[x]}$ | Activity of dLGN cell $x$ |
| $r_{[y]}$ | Activity of RGC $y$ |
| $w_{[xy]}$ | Synaptic weight from RGC $y$ to dLGN cell $x$ |
| $\tilde{w}_{[xy]}$ | Updated synaptic weight from RGC $y$ to dLGN cell $x$ before normalization |
| $h$ | Learning rate |
| $C$ | Membrane Capacitance |
| $g_l$ | Leak conductance |
| $g_e$ | Excitatory synaptic conductance |
| $g_i$ | Inhibitory synaptic conductance |
| $E_e$ | Excitatory potential |
| $E_i$ | Inhibitory potential |

Table 4.2: The symbol table for the terms used in this section.

synapses by one of the two afferent types. In fact, the specific form of the bias is not important (see Sec. 5.8), but in order to develop a stereotypical configuration there must be biases present.

### *4.3.4 dLGN Update*

Geniculate potentials are computed as a (weighted) sum of their inputs from the retina and their neighbors in the dLGN. We assume that the dLGN cells reach steady state for a given retinal input because the speed of wave propagation is much slower than synaptic transmission and membrane depolarization. The steady state is calculated by iterating the following update equation until the geniculate activities stabilize:

$$\frac{\delta l_{[x]}}{\delta t} = \frac{\lambda}{1+2\lambda}\left(\frac{1}{\sqrt{\lambda}}\sum_{y} w_{[xy]} r_{[y]} + \sum_{x\in\eta} l_{[x]}\right) - l_{[x]} \tag{4.2}$$

where $l_{[x]}$ is the activity of dLGN cell $x$, $\lambda$ is a constant, $w_{[xy]}$ is the synaptic weight from RGC $y$ to dLGN cell $x$, $r_{[y]}$ is the firing rate of RGC $y$, and $\eta$ includes the dLGN cell's 2 nearest intra-layer neurons. If the neuron is at the boundary it will have only 1 intra-sublayer neighbor (see Appendix A for details on activity update in boundary neurons).[3]. The dLGN cell activity is calculated from the scaled product of its synaptic inputs and the RGC firing rates in addition to the activities of the neighboring dLGN cells and self-inhibition.

In signal processing terms, the intra-dLGN influence acts as a feedback linear system that regularizes (blurs) the weighted retinal input. Regularization smoothes a partial or noisy input,

[3]These numbers assume a 1-d model, there are 4 nearest neighbors in the 2-d model. Note also that this equation ignores the influence of the inter-layer and sublayer influence, see App. A for the generalization of this update.

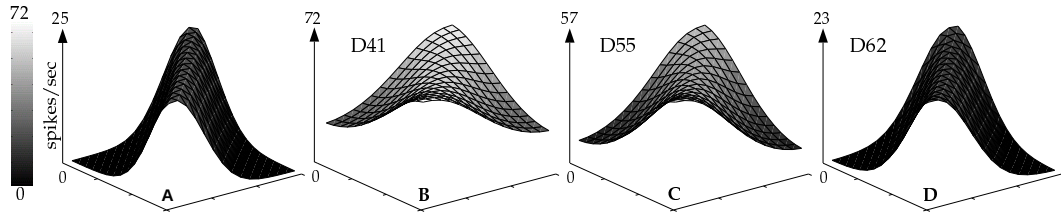


Figure 4.6: (A) Retinal activity during a wave. (B-D) dLGN response at different stages in development. The retinal signal evokes a large, blurred response in the neonatal dLGN, a less blurred response later in development, and a virtual copy of the retinal input at eye opening. For the purpose of this figure, the retinogeniculate projection was assumed to be 2-dimensional and perfectly retinotopic in all three cases.

effectively de-noising the input and interpolating missing values under the assumption that the initial input is smooth (see Appendix A for details on the relation between the dLGN update equation and regularization).

The amount of blurring/smoothing and the amplification of the dLGN response are both scaled by $\lambda$, which decreases over development (see Fig. 4.6). The range of $\lambda$ (20 at birth and 1 at eye opening) as well as the scale factor ($\frac{1}{\sqrt{\lambda}}$) on the effective inputs to the dLGN cells was chosen to match the roughly three-fold decrease in geniculate responsivity observed over the course of development (see AppendixA, Sec. 3.4.1 and Ramoa and McCormick, 1994b).[4] Because it supports global order early in development and local precision later in development, this decrease in blurring plays an analogous role to more traditional neighborhood functions that shrink over development (Kohonen, 1993).

*Intra-dLGN Asymmetries.* The inter-layer and inter-sublayer weights are much weaker then the intra-layer weights. This asymmetry is crucial to the robust development of layer segregation (see Sec. 5.9 and Fig. 5.13). In the canonical model the inter-layer and inter-sublayer weights decrease linearly over development. The time-course of decrease is shown in Fig. 4.7. The inter-layer weights start off quite weak (.1 versus intra-sublayer weights of 1) and are decreased linearly during the first 2 weeks of postnatal development, so that during the third week there is no excitation between the two eye-specific layers (i.e. the inter-layer weights go to 0). The inter-sublayer weights start off at 1, decrease to the initialization value of the eye-specific weights (.1) over the first two weeks and then decrease linearly to zero during the third week of postnatal development.[5]

Layers and sublayers will segregate in the model without this decrease, but the segregation is slightly imperfect. Intact inter-layer connections blur the boundaries between layers and sub-layers even after segregation. Further, the existence of excitatory inter-layer interactions supports the "clustering" of off-center RGC axons from the two eyes. This clustering disappears when the

[4]Standard regularization preserves the sum under the blurred image, thus the peaks of the image tend to drop as blurring increases. Adding the scale factor of $\sqrt{\lambda}$ to the inputs counter-acts this tendency, and makes the dLGN response increase with $\lambda$, as seen in the developing dLGN, Sec. 3.4.1.

[5]There is no variance or noise in the intra-dLGN synaptic weights.

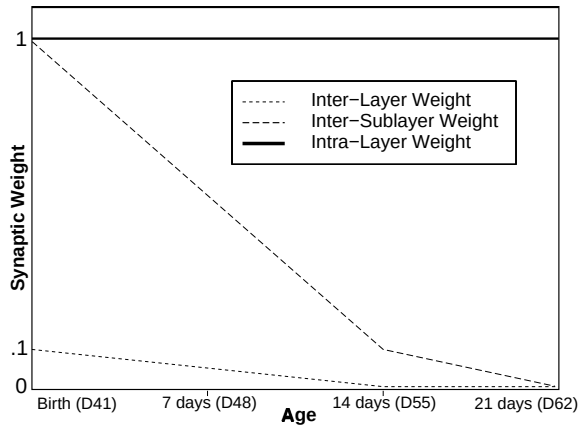


Figure 4.7: In the canonical model, the inter-layer weights decrease to zero during eye-layer segregation (D41-D55). The inter-sublayer weights start off equal to the intra-layer weights, decrease to .1 during the period of eye-layer segregation, and decreasing to 0 during the period of on/off sublayer segregation (D55-D62). The intra-layer weights remain constant over the course of development.

excitatory weights are removed (see Sec. 5.9.2 and Figs. 5.14 and 5.15). A linear decrease in the inter-layer weights was adopted because it is simple and straightforward. The eventual disappearance of inter-laminar weights is not a particularly strong assumption given that the layers are eventually separated physically by the cell sparse interlaminar zones. Indeed, in the adult there are even inhibitory connections between the eye-specific layers (see Sec. 2.3.4). To my knowledge, there is no data on the development or existence of interactions between the sublayers.

*Biophysical Interpretation.* A biophysical interpretation of Eq. 4.2 can be stated in terms of the membrane equation:

$$C\frac{\delta l_{[x]}}{\delta t} = -(g_{l[x]}l_{[x]} + g_{e[x]}(l_{[x]} - E_e) + g_{i[x]}(l_{[x]} - E_i)) \tag{4.3}$$

which describes a geniculate cell's membrane potential ($l_{[x]}$) as a function of its excitatory ($g_{e[x]}$), inhibitory ($g_{i[x]}$), and leak ($g_{l[x]}$) conductances, their respective reversal potentials ($E_e, E_i,$ and 0), and the capacitance ($C$) of the membrane.

Eq. 4.3 becomes algebraically equivalent to Eq. 4.2 by eliminating the scale factor $\frac{1}{\sqrt{\lambda}}$ from Eq. 4.2 and equating the following terms (see Appendix B for derivation details):

$$\begin{aligned} C &= \frac{1+2\lambda}{\lambda} \\ g_{e[x]} &= \frac{\sum_y w_{[xy]}r_{[y]} + \sum_{x\in\eta} l_{[x]}}{E_e} \\ g_{i[x]} &= -\frac{(g_{l[x]} + g_{e[x]} - \frac{1+2\lambda}{\lambda})l_{[x]}}{l_{[x]} - E_i} \end{aligned}$$

Simulations of the dLGN relay cell activity using Eq. 4.3 and Eq. 4.2 display parallel changes over development. The decrease in $\lambda$ over the course of development (see above), indicates that the capacitance $C$ will increase by a factor of $1.5$, and the inhibitory conductances $g_{i[x]}$ increase by a factor dependent on the magnitude of leak conductance $g_{l[x]}$ assumed. These dynamics are consistent with physiological measurements of developing dLGN relay cells (Ramoa and McCormick, 1994a; Ramoa and McCormick, 1994b). Under these conditions the inhibitory and excitatory conductance are in a push-pull relationship, meaning that as $g_{e[x]}$ increases, $g_{i[x]}$ decreases.

Note that although algebraically identical, the two equations behave slightly differently because the inhibitory conductances are rectified such that they cannot go below 0, and because the input scaling factor $\frac{1}{\sqrt{\lambda}}$ has been eliminated. Under the assumption that $E_i = 0$ (i.e. the inhibitory reversal potential is equal to the leak reversal potential)[6] the elimination of $\frac{1}{\sqrt{\lambda}}$ and the rectification of the inhibitory conductance tend to compensate for each other. The reason is that if $g_{i[x]}$ is allowed to be negative it can amplify the responsivity of the cell, but if it is rectified the dLGN cell activity will saturate at a level dependent on $g_l$ and $E_e$. As a result, rectifying $g_{i[x]}$ effectively establishes a saturation level for dLGN cell firing rates. The elimination of the scale factor $\frac{1}{\sqrt{\lambda}}$ on the inputs has the opposite effect, because its removal amplifies the input signal — a change which amplifies the dLGN response. The approximate balance between the rectification and the removal of the input scale factor has been verified in simulations.

### 4.3.5 Hebbian Weight Update

Second, the retina-to-dLGN weights are updated using a modified Hebbian rule:

$$\delta w_{[xy]} = h\, w_{[xy]}\, l_{[x]}\, r_{[y]} \tag{4.4}$$

The weight change ($\delta w_{[xy]}$) is the product of the learning rate ($h$), the current weight ($w_{[xy]}$), the steady-state geniculate activity ($l_{[x]}$), and the retinal activity ($r_{[y]}$).

Scaling the Hebbian update by the previous weight generates a feed-back cycle in which strong weights tend to become stronger. This tendency allows for the development of relatively precise connectivity, even in the case of multiplicative normalization (see Eq. 4.5 below) which would otherwise tend to favor graded synaptic connectivity (Miller and MacKay, 1994).

### 4.3.6 Synaptic Sprouting

Third, the RGC axonal arbors are blurred (i.e. the weights are convolved with a Gaussian) to mimic a branching factor. The degree of blurring for each RGC axonal arbor is scaled by that RGC's activity. The arbors of those RGC's that are currently firing bursts of action potentials (i.e.

[6] Following the biological data, this assumes that inhibitory conductances in question correspond to GABAa shunting inhibition, see Sec. 2.3.5 and Sec. 3.4.1.

are at the peak of a wave) will be blurred strongly whereas the the arbors of silent RGCs will not be blurred at all. Because Hebbian update tends to sharpen the connectivity, this scaling helps maintain equilibrium in the model. The scaling is especially important during the period of on/off activity in the retina, for in this period there are many retinal "snapshots" with only a few spiking RGCs. Without the scaling, this sparse activity would blur silent RGC arbors for many time-steps without sharpening them, and thus the RGC arbors will tend to diffuse.

Consistent with observations of the developing dLGN (Sretavan and Shatz, 1986), the use of sprouting yields RGC axonal arbors that can shift their geniculate position appreciably, yet change their width only slightly over development. Because it is strongest at small distances, sprouting encourages neighboring neurons to receive synaptic inputs from the same afferents and, as a result, neighboring dLGN cells will tend to have correlated firing patterns. Therefore, sprouting plays a similar role to the direct excitation provided by neighborhood functions. Indeed, a model with only sprouting can topologically organize, and in some cases segregate into layers, in the absence of lateral interactions in the target tissue (Keesing et al., 1992; Eglen, 1997 and see Sec. 5.6.2).

Parallel to the intra-dLGN weights, sprouting is also asymmetric. That is, the weights are blurred more within each layer than between layers. The degree of asymmetry and its increase over the course of training is identical to that of the intra-dLGN weights (see Sec. 4.3.4 and Fig. 4.7).

*Biological Interpretation.* Axonal sprouting and retraction is know to occur in the tecta of some fish and amphibians (Fawcett, 1993). It is thought by some to be mediated by trophic factors originating in active sites. A candidate factor is nitric oxide (NO). NO is thought to act as a retrograde messenger that changes pre-synaptic efficacy. In at least one model (Krekelberg and Taylor, 1997), NO uptake at the pre-synaptic site is proportional to the pre-synaptic activity — parallel with the instantiation of sprouting in this model.

*Computational Details* Intra-layer blurring is simulated by convolving the axonal weight matrix with a Gaussian. Inter-layer and inter-sublayer sprouting is simulated by adding the blurred weights in the adjoining layer(s), scaled by the inter-layer or inter-sublayer sprouting strength. The blurred version of the synaptic weights is then scaled by the learning rate, the pre-synaptic activity, and the sprouting rate and is added to the unblurred weight matrix. The learning rate is shared with the Hebbian update, but the sprouting rate is an independent scale factor on the sprouting.

### *4.3.7 Weight Normalization*

Finally, the weights are normalized divisively such that the total connection strength of each RGC and dLGN neuron stays roughly constant.

$$w_{[xy]} = \frac{2\tilde{w}_{[xy]}}{\sum_x \tilde{w}_{[xy]} + \sum_y \tilde{w}_{[xy]}} \tag{4.5}$$

where $\tilde{w}_{[xy]}$ denotes the updated weight before normalization, $\sum_x \tilde{w}_{[xy]}$ is the total synaptic weight provided by an RGC axonal arbor, and $\sum_y \tilde{w}_{[xy]}$ is the total synaptic weight in a dLGN cell's dendritic tree. The factor of 2 is included in the numerator so that for the limiting case of equal numbers of RGCs and dLGN cells, the sum of each cell's synaptic weights is 1. In fact, the sum of each cell's synaptic weights is a function of the ratio between the number of pre and post-synaptic cells.

$$\begin{aligned}\sum_x w_{[xy]} &\approx \frac{2 \times \#\,\text{Post-synaptic Cells}}{\text{Total}\,\#\,\text{of Cells}} \\ \sum_y w_{[xy]} &\approx \frac{2 \times \#\,\text{Pre-synaptic Cells}}{\text{Total}\,\#\,\text{of Cells}}\end{aligned}$$

In the canonical model, with a 2:1 dLGN cell to RGC ratio, each RGC supports a total synaptic weight of 1.33 and each dLGN cell supports a total synaptic weight of .66.

Because all of the weight changes are initially positive, this step is crucial to prevent the weights from increasing without limit. The assumptions here are that that each neuron can only support a limited number of synapses and neurons will compete for synaptic sites. Presumably, this process is accomplished biologically through synaptic competition. Given a finite amount of resources, the biological system has to take resources from some synapses (i.e. decrease their strength) in order to grow others. While this is surely not true in the exact sense assumed by canonical form of the model, competition for resources does occur in neural systems (see Miller, 1996 for review).

Notably, the rigidity of this normalization procedure precludes the model's ability to usefully model deprivation results. The problem is that even silent inputs are constrained to maintain a fixed total synaptic strength. Thus, the observation that monocular silencing leads to a decreased projection from the deprived eye (see Sec. 3.3.7) cannot be modeled with this type of normalization.

*Capped Normalization* With a variant of normalization, derived by analogy to Eglen's work (section 5.3.1 in Eglen, 1997, and Sec. 7.2.2), it is possible to partially model monocular deprivation results in the current framework. This variant on normalization limits ("caps") the total allowed weight by applying normalization to cells that exceed the allowed total synaptic weight, and ignoring cells that are under the limit. With capped normalization inactive cells can lose their synaptic connections.

However, taken alone, this approach is not sufficient to bring about decay of retinal synapses, because the only way for synapses to decay in this model is through normalization driven competition (in contrast to a covariance weight change rule, see Sec. 7.2.2). Because of the canonical limits imposed on total synaptic strength, a retina can only compete for half of the dLGN synapses available. Consequently, the other retina will keep the other half though simple lack of competition.

A partial solution is to raise the cap on total RGC synaptic strength to near 200% of the normal amount, such the retinal patches taken together are capable of supporting $\sim 2\times$ more synaptic strength than the entire dLGN. This change allows a subset of the RGCs to actively compete for the entire retinogeniculate projection.

The capped normalization procedure for the canonical model is:

$$w_{[xy]} = \frac{2\tilde{w}_{[xy]}}{\tau + \sum_y \tilde{w}_{[xy]}} \tag{4.6}$$

$$\tau = \begin{cases} 1.33 \times \frac{\sum_y \tilde{w}_{[xy]}}{\theta} & \text{if } \sum_x \tilde{w}_{[xy]} > \theta \\ 1.33 & \text{otherwise} \end{cases}$$

where $\tau$ is the "effective" (for purposes of normalization) synaptic weight provided by an RGC axonal arbor, and $\theta$ is the cap on the total synaptic weight provided by an RGC axonal arbor. This method of normalization treats RGCs with total synaptic strength below $\theta$ as if they are already normalized, while decreasing synapses from RGCs with a total synaptic strength above $\theta$.

Preliminary explorations suggest that with this capped version of normalization the model is able to capture some of the observed effects of deprivation (see Sec. 5.10.3).

# Chapter 5

# Model Results

There are regions of the parameter space of this model that I haven't explored, simply because the space is overwhelmingly large. My approach has been to fix as many of the parameters in the model as possible using the available biological data. With these parameters fixed, I have searched for regions of the free parameter space that robustly support stereotypical development with the canonical set of model mechanisms. I then explored the behavior of the model in these regions. I did not always rigorously characterize these explorations, because the inter-relations between parameters have often proven too complex to be understood through simulations. There are probably other coherent unexplored areas of the parameter space, and the results that I present below may not hold completely in these unexplored areas.

That said, I have explored the behavior of this model in depth, and I believe I have solid intuitions about the dynamics and crucial features of the model. I also feel that one specific model that supports the development of structures seen in the retinogeniculate projection provides a useful and testable framework for exploring the activity-dependent development of the system.

It is important to note that although some model configurations presented appear to categorically fail, they can sometimes be made to work with appropriate adjustments of the model parameters or with different initialization conditions and simulated waves (both are stochastic features of the simulations).[1] The parameter regime was chosen to robustly support the convergence of the canonical model, and therefore the robustness of the other results are somewhat variable. One should be careful in drawing categorical conclusions from the presented results, rather they should be considered as indicating general tendencies in the model dynamics. Table 5.2 attempts to summarize the quantitative influence of the model mechanisms, in which the numbers given are related to the results shown as well as to accumulated experience running many simulations that are not presented.

[1]The results shown were initialized with the same random seed unless stated otherwise. This procedure ensures that the starting conditions and the wave activity were the same in all of the results, and eliminates spurious differences that result from different random waves and starting configurations.

| | |
|---|---|
| Retinal Patch Dimensions | $1 \times 10$ cells |
| dLGN Layer Dimensions | $1 \times 20$ cells |
| Duration in biological time | D41-D62 † |
| Onset of On/Off activity | D55 |
| Time-step ($\delta t$) | .1 seconds |
| Learning Rate (h) | .005 |
| Total dLGN Cell Synaptic Weight ($\sum_y w_{[xy]}$) | .66 |
| Total RGC Synaptic Weight ($\sum_x w_{[xy]}$) | 1.33 |
| Retinotopic Bias at Initialization | 20% |
| Eye-Specific Bias at Initialization | 20% |
| Iterative On/Off Bias | .005% |
| RGC Axonal Arbor Width at Initialization | $\pm 2$ sd's $= 150\mu m$ |
| dLGN Blurring ($\lambda$) | $20 \rightarrow 1$ over development |
| Intra-Layer Interaction Strength | 1 ‡ |
| Inter-Layer Interaction Strength | $.1 \rightarrow 0$ (D41-D55) |
| Inter-Sublayer Interaction Strength | $1 \rightarrow .1$ (D41-D55), $.1 \rightarrow 0$ (D55-D62) |
| Sprouting Rate (see Sec. 4.3.6) | 10 |
| Intra-Layer Sprouting Strength | 1 †† |
| Inter-Layer Sprouting Strength | $.1 \rightarrow 0$ (D41-D55) |
| Inter-Sublayer Sprouting Strength | $1 \rightarrow .1$ (D41-D55), $.1 \rightarrow 0$ (D55-D62) |

Table 5.1: The canonical model parameters. Unless stated otherwise, all the simulations in this section assume these parameter settings. Arrows indicate linear change with time.

†For the purposes of comparison to the biological data, these dates assume that model development spans the equivalent of 3 weeks of biological time. In fact, for computational efficiency, the canonical model only develops during ∼3 hours worth of wave inputs, and has a relatively large learning rate to speed development (see Secs. 5.1.3 and 5.4 for discussion and validation).

‡The coefficient on each term in the sum $\sum_{x \in \eta} l_{[x]}$ in Eq. 4.2.

††1 is an arbitrary unit used to compare the strength of intra-layer sprouting with that of inter-layer sprouting. The intra-layer sprouting is simulated by convolving the weights (i.e. the model RGC axonal arbor) with a Gaussian with standard deviation = 1.25 dLGN cells ($\sim 38\mu m$), truncated to a 5 cell width.

| Model Variant | Retinotopy | Eye Layers | On/Off Sublayers |
|---|---|---|---|
| Canonical Model | 5 | 5 | 5 |
| Long Time-Course | 4 | 5 | 4 |
| Multiple Domains | 4 | 4 | 4 |
| dLGN Blurring Only | 4 | 5 | 4 |
| Sprouting Only | 4 | 5 | 4 |
| No Noisy Off Firing | 5 | 5 | 0 |
| Biases | | | |
| • None | 4(0) | 5(0) | 3(0) |
| • All Initial | 5 | 5 | 0 |
| • All Iterative | 5 | 5 | 5 |
| Asymmetries | | | |
| • Layer/sublayer interactions = 1 | 2 | 0 | 0 |
| • Static | 4 | 5 | 4(0) |
| • Static, inter-layer $\rightarrow$ 0 | 4 | 5 | 5 |
| • Layer/sublayer interactions = 0 | 1 | 5 | 1 |
| Normalization | | | |
| • Post-synaptic Only | 0 | 0 | 0 |
| • Pre-synaptic Only | 0 | 4 | 2 |
| • Capped | 3 | 4 | 0 |
| • Capped & Monocular Deprivation | 3 | 0 | 1 |

Table 5.2: Qualitative judgments of the robustness of convergence for the model variants discussed in this section, where 5 = very robust and 0 = no tendency to develop. Numbers in parentheses indicate the tendency to develop the stereotypical configuration, and are only included where the development of a non-stereotypical configuration is seen. The organization of the table reflects the organization of the results presented.

## 5.1 Format of Results

Because this model is structured with a fair amount of realism, and because it encompasses eye-specific layers, on/off sublayers, and retinotopy, the format used to display the results is probably non-obvious. The majority of the results presented below are simply plots of the synaptic weight matrices used in the simulations. Because the model simultaneously simulates the projections from 4 patches of retina to 4 layers of LGN, each weight matrix is a collection of 16 sub-matrices.

### *5.1.1 1-Dimensional Result Display*

The format of the weight matrix is linked to the format of RGC activity and LGN potential vectors. As illustrated in Fig. 5.1 these vectors are simply concatenations of the activities in each patch and sublayer. Using this format, calculating the weighted synaptic input to the LGN (not the

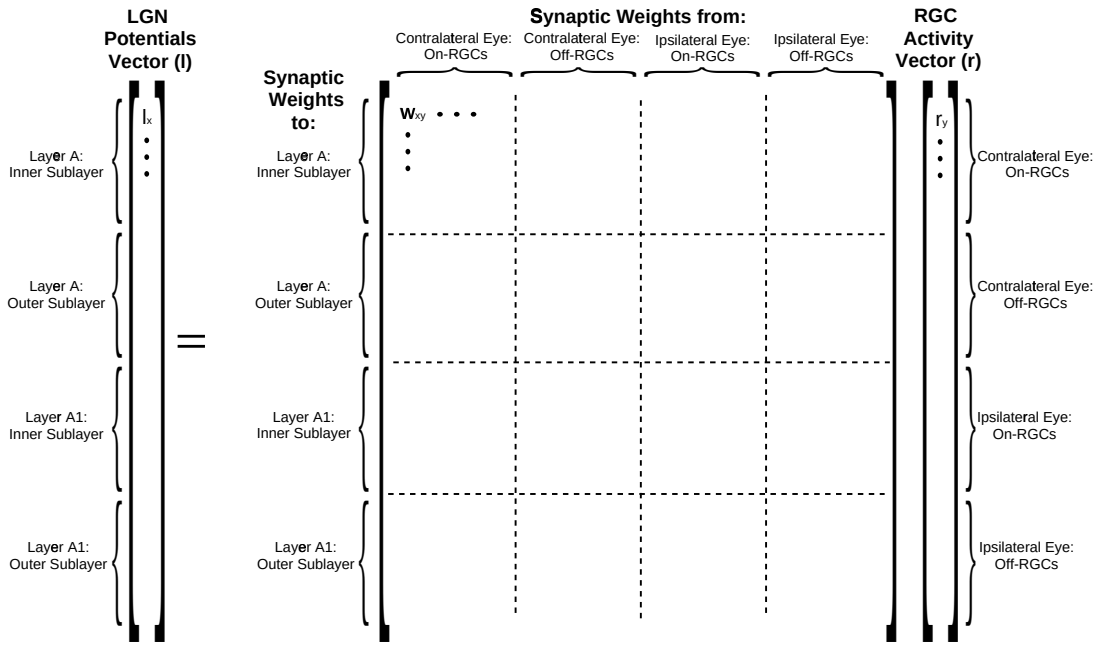


Figure 5.1: The model can be expressed in matrix notation by concatenating the activities of the 4 retinal patches into a single vector and collecting the synaptic weights from each retinal patch to each LGN sublayer into a single matrix. Multiplying these together yields an LGN potentials vector that is a concatenation of the potentials of LGN cells in each of the 4 sublayers ($\mathbf{l} = \mathbf{W}\,\mathbf{r}$). The specific organization of the vectors and the weight matrix are explained in the figure. Within each of the 4 sub-vectors of RGC and LGN activities, neighboring entries correspond to spatial neighbors. Thus, the spatial pattern of activity across the retina and LGN can be imaged by simply plotting the vector of activations for a patch.

stable state, see Appendix C for discussion) is a simple matter of multiplying the retinal activity vector by the matrix of synaptic weights.

In each plot of a weight matrix, the darkness of a dot indicates the strength of the synaptic weight from an RGC (corresponding to the column) to an LGN relay cell (corresponding to the row), normalized by the largest synaptic weight in the entire simulation. Fig. 5.2 illustrates the meaning of the 1-d results plots.

Note that only 5 points in the development of each 1-d simulation are shown. The 5 time-points were chosen to best illustrate initialization, the earliest completion of eye-specific segregation, the earliest precise retinotopy, the point just after the onset of on/off segregation, and the final projection state. Because of these constraints, the results displays are not directly comparable, but should be maximally informative. Note also that different durations of development pass between the onset

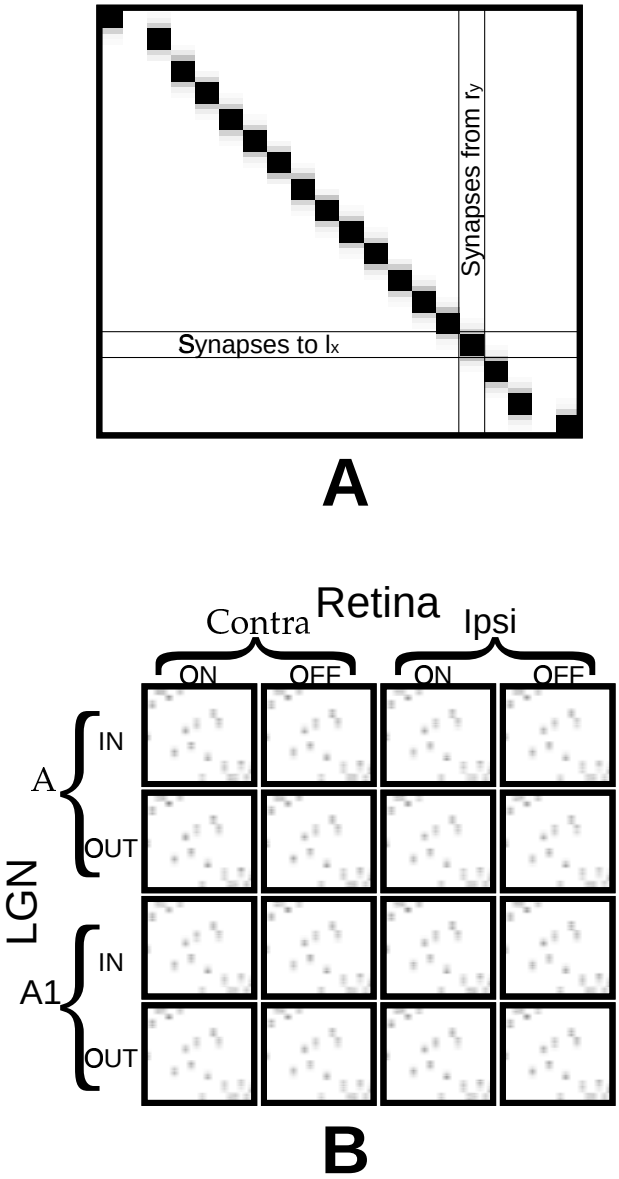


Figure 5.2: Plots of weight matrices are used to display the results of the 1-d simulations. A) The weights from a single retinal patch to a single sublayer are shown. The darkness of an entry indicates the strength of the synapse. From this plot we can conclude that the projection is almost perfectly retinotopic, because neighboring points in the retinal patches project most strongly to neighboring points in the LGN. Perfect retinotopy would correspond to the identity matrix in the case of equal numbers of LGN cells and RGCs (in fact the ratio of LGN cells to RGCs is 2:1 in the 1-d simulations and 4:1 in the 2-d case). B) The structure of the complete weight matrix is shown. In this case the layers are not yet segregated, evidenced by the presence of synaptic connections from all retinal patches to all sublayers, and retinotopy is still quite coarse, evidenced by the dissimilarity between the weight matrix and the identity matrix. The matrices are collected such that a perfect final configuration will look like one large identity matrix.

of on/off segregation and the next saved time-point, so the speed of on/off segregation can be a bit misleading. As a rule on/off segregation occurs very quickly ($\sim 1/2$ day in biological time).

### 5.1.2 2-Dimensional Result Display

The weight matrices for 2-d simulations are not easily interpreted because their spatial layout is arbitrarily mapped into the 1-d structure of vectors. The 2-d results are displayed in a more abstract form (see Fig. 5.3). The centers of the receptive fields (RFs) for each dLGN relay cell are plotted and connected by lines to the RF centers of their nearest dLGN relay cell neighbors. Each of the 16 sub-quadrants maps the RFs of the dLGN relay cells in a particular sublayer (corresponding to the row), in a particular retinal patch (corresponding to the column).

The location of a dot within one of the sub-quadrants of a plot corresponds to the position of that RF in visual/retinal space. Note that there are 4 dLGN cells for each RGC, so a perfectly sharp projection in which each dLGN cell receives synaptic inputs from only 1 RGC would be a 10x10 grid with 4 overlapping RFs centers plotted at each vertex (i.e. there would appear to be only 100 dLGN RF centers plotted, instead of 400). The size of a dot indicates the strength of the synaptic input from that retinal patch (column) to the given dLGN relay cell. Very small dots indicate weak projections, and no dot indicates that the total projection strength from that retinal patch to that dLGN sublayer is less than $1/1000$th the strength of the strongest patch to sublayer projection. A perfect grid within each sub-quadrant indicates perfect retinotopy, while blank sub-quadrants everywhere but the main diagonal indicate perfect layer segregation.

### 5.1.3 Speed Issues

The model waves initiate at a frequency and propagate at a speed that is directly linked to in vitro measurements of the waves (see Table 4.1). Because of these parallels, there are roughly the same number of waves in a week of model "biological time" as in a week of in vivo time, and the intrinsic time in the model can be directly related to the time-scale of biological development. However, most of the simulations shown below have a much shorter developmental time-course than the biological system. Because of computational constraints they only develop for $\sim$3 hours in biological terms. The days that are presented in the captions of most of the results presented below are normalized, assuming that the course of the simulation corresponds to a period of $\sim$3 weeks (D41-D62).

My assumption in this model is that a long running simulation with a relatively small learning rate will yield roughly equivalent results to a short simulation with a relatively large learning rate. To the extent that this approach alters the behavior of the model, I would expect that it is biased against appropriate development. My reasoning, and that of other modelers, is that the system can develop appropriately because it is able to smooth out idiosyncrasies that occur on a short time-scale by averaging activity and synaptic weight change over a long span of time. An unrealistically fast time-course could push a model toward local minima, preventing it from converging to global

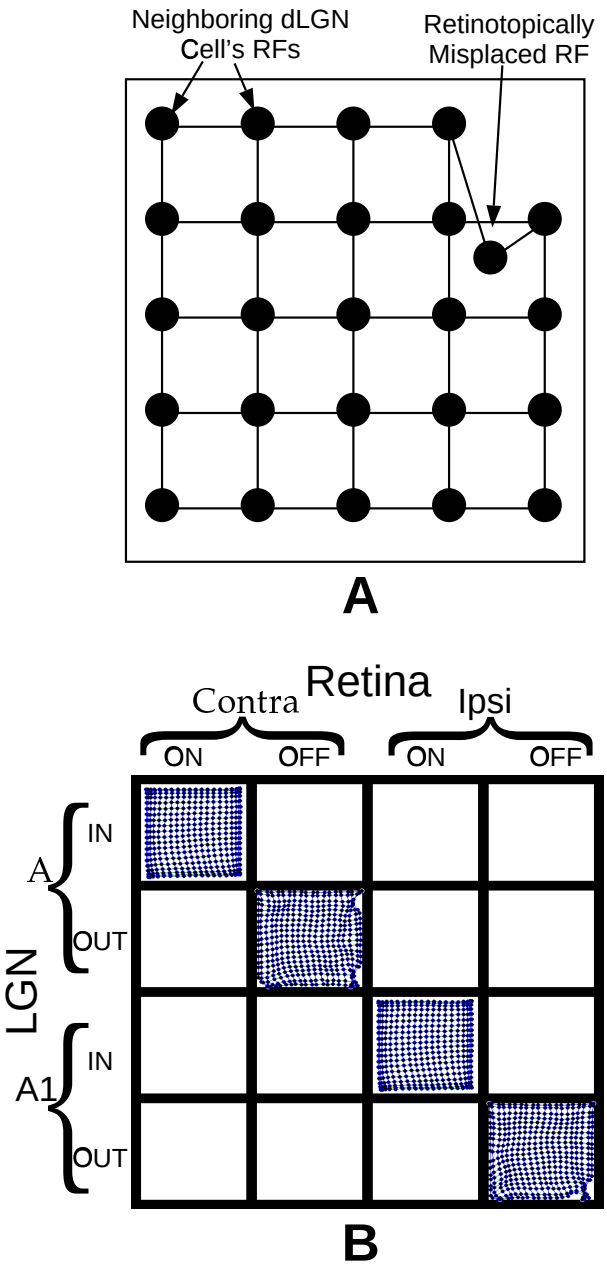


Figure 5.3: Plots of the receptive field (RF) centers are used to display the 2-d simulation results. (A) The weights from a single retinal patch to a single sublayer are shown. The receptive field centers of neighboring dLGN cells are linked by lines. A regular grid indicates perfect retinotopy. The size of each dot indicates the strength of the projection from the retinal patch to the corresponding dLGN cell. (B) The structure of the complete weight matrix is shown. In this case, the layers are perfectly segregated in the stereotypical configuration with nearly perfect retinotopy in each sublayer. The blank sub-quadrants indicate that there is only a very weak projection from the given retinal patch to the given dLGN sublayer ($< .1\%$ of the strength of the strongest patch to sublayer projection).

order. In this view, if the model achieves global order with a fast time-course it is a testament to the robustness of the mechanisms involved. In order to validate these assumptions and ensure that the results are not spuriously supported by the acceleration in the time-course of development, a simulation with a more realistic time-course was run (see Sec. 5.4) and the results were in line with the other simulations.

### *5.1.4 Scale Issues*

Most of the simulations presented below are 1-d simulations with 1x10 RGCs in each patch and 1x20 dLGN relay cells in each sublayer. There is also a 2-d simulation with 10x10 RGCs in each patch and 20x20 dLGN relay cells in each sublayer. The 1-d simulations are used to explore the model behavior, because they are much less computationally demanding then the 2-d simulations [2].

The 2-d model helps to address several potential problems in the 1-d simulation. Each simulated wave in the 1-d simulation is modeled as a 1-d section through a 2-d wave that propagates from a random point located on a circle (radius = 100 RGCs) around the retinal patch. As a result, some wave-fronts are oriented parallel to the orientation of the retinal strip. Effectively this causes a very wide, very quickly moving wave to excite all of the RGCs in a 1-d patch simultaneously (this is not true in the multiple domain 1-d simulation, Sec. 4.3.2). In the 1-d simulations these parallel waves could unrealistically support the grouping of distant RGCs in the same retina. In the 2-d model there can never be a wave that simultaneously activates all of the RGCs (because the waves limited in their width) so the convergence of the 2-d model indicates that this unrealistic aspect of the 1-d model is not necessary for appropriate model development.

The 2-d model probably also presents a more realistic picture of the lateral interactions in the dLGN than the 1-d model because each dLGN cell has more nearest neighbors. In the 1-d model there are 2 nearest intra-sublayer neighbors (except at the edges), while in the 2-d model there are 4. Because the number of inter-sublayer connections is the same in both the 1-d and the 2-d simulation (2 or 1 at a boundary), the asymmetries in intra-dLGN influence are accentuated in the 2-d model. This increase in asymmetry seems to support faster and more robust layer segregation in the 2-d relative to the 1-d simulations, suggesting that judgments of robustness based on the 1-d results are probably slightly conservative.

Because the 2-d model incorporates a 4:1 dLGN relay cell to RGC ratio, it also more accurately models the neural amplification factor found in cats, and ferrets (see Sec. 2.3.3). In my experience the degree of neural amplification (i.e. “dimension expansion”) has little effect on the convergence behavior of the model.

[2] In the full 2-d simulation the synaptic weights matrix is 1600x400, and the matrix that mediates the lateral interactions in the dLGN is 1600x1600. At this scale the model takes more than a week to simulate.

NOT DONE

Figure 5.4: The 2-d simulation of the canonical model can support the development of precise retinotopy, eye-specific layer segregation and on/off specific sublayer segregation. RF centers of dLGN cells in the 4 sublayers are plotted in each of the 4 retinal patches (Fig. 5.3). The developmental time-course is shown at ages (A) D41, (B) D42, (C) D44, (D) D47, (E) D50, (F) D54, (G) D56, (H) 62.

## 5.2 Organization of Results

The main mechanisms in the model are explored and discussed below. Because these mechanisms do not fall into a hierarchy, there is not a natural organization for the explorations. In order to provide some structure, Table 5.2 provides an outline of the organization and order of the results.

## 5.3 Canonical Model

The mechanisms in the model can support the development of eye-specific layers, on/off sublayers and precise retinotopy from an initially disordered projection. The time-course of refinement is qualitatively similar to refinement in the actual projection: eye-specific layers segregate early in development, on/off sublayers segregate during differential on/off activity in the retina, and the refinement of retinotopy is relatively prolonged. Over the course of development the connectivity sharpens, but the arbors stay within a relatively constrained range of widths (as seen in the ferret and cat, see Sec 3.3.3).

A large 2-dimensional simulation is shown in Fig. 5.4. The initial projection is shown in Fig. 5.4A. The small amount of bias toward retinotopy, and eye-specific layer segregation is the same in this simulation as in all of the others (unless stated otherwise), but the display method makes the degree of initial disorder more clear. Notice that the layers segregate quickly relative to retinotopic refinement, and that the on/off sublayers segregate quickly after the onset of on/off differentiated RGC activity. The extent of segregation is less evident in the 2-d model than in the 1-d models because the threshold for ignoring residual weights is very low in the 2-d plots. Weak synaptic weights in the 1-d plots disappear at a higher threshold because they are on a continuous scale with the strongest synaptic weights over the entire course of development.

For comparison to the other results shown below, a 1-d simulation is shown in Fig. 5.5. The sharpness of the off-center RGC arbors relative to the on-center RGCs at the end of development is due to the differences in their activity patterns during on/off sublayer segregation. Because of noisy firing during this period, the off-center RGCs often fire in relative isolation from other RGCs (see Sec. 4.3.2). This isolated firing allows synapses to grow maximally because there is no competition for synaptic resources from other cells. The on-center RGC's continue to fire only in waves when

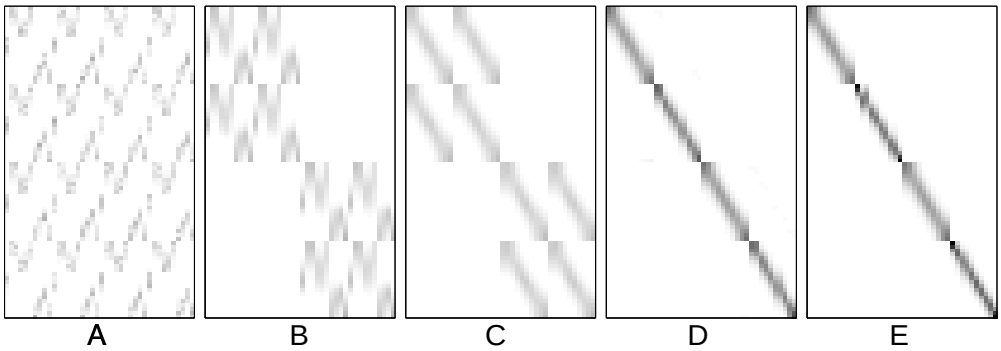


Figure 5.5: A 1-d simulation of the canonical model can support the development of precise retinotopy, eye specific layer segregation and on/off sublayer segregation. Weight matrix is shown at ages (A) D41, (B) D42, (C) D46, (D) D56, (E) D62.

many of their neighbors are simultaneously actively competing for resources so their arbors remain relatively broad.

## 5.4 Longer Time-course

As discussed above in Sec. 5.1.3 most of the simulations of the current model have an unrealistically short time-course and fast learning rate in order to make the simulations run faster. In order to validate this approach, a simulation with a realistic time-course was run. Fig. 5.6 illustrates the development of a simulation that took 21 days in model biological time.

The convergence of this realistically timed simulation demonstrates that the model dynamics do not depend on a speeded time-course. The model also demonstrates that with this extended time-course, the parameters can be chosen such that the time-course of layer segregation accurately

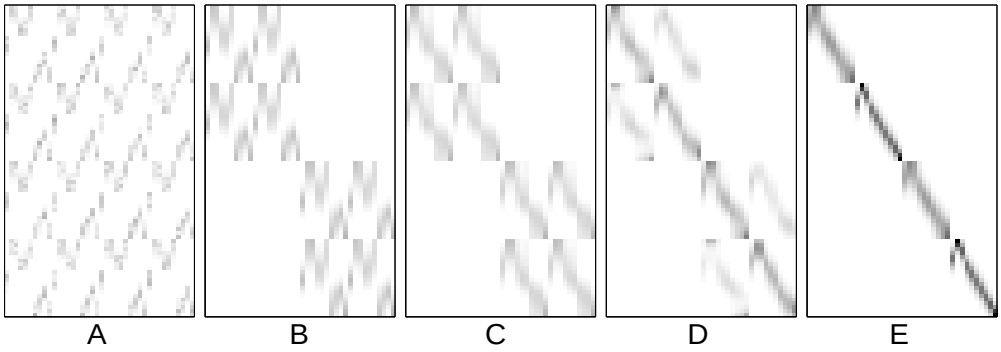


Figure 5.6: A simulation with a realistic time-course develops similarly to the speeded standard simulations. Weight matrix with a biological time-course of 3 weeks is shown at ages (A) D41, (B) D46, (C) D53, (D) D57, (E) D62.

follows the observed time-course of segregation.[3] In this simulation, the learning rate was decreased in inverse proportion the change in simulation time and further by a factor of 10 in order to bring the time-course of layer formation into accordance with biological observations (from $5 \times 10^{-3}$ to $5 \times 10^{-6}$). Due to computational constraints, the time-step used in this model was .5s instead of .1s. All other parameters were the same as in the standard simulations.

## 5.5 Multiple Domains

As discussed above in Sec. 4.3.2 the model assumes that each retinal patch is always in the same transient subdomain, and so always has a wave pass completely over it. In order for the model to be generally applicable to the development of clustering or to the afferents from each entire retina, the model must be applicable to multiple sub-domains. In order to model multiple subdomains, and to bias the simulation against unrealistic mechanisms of layer segregation, a simulation was run using waves that passed over only $1/5$—$4/5$ of a retinal patch, leaving the rest of the RGCs silent. This activity pattern corresponds to the retinal patch being chronically located on the shifting border between two transient subdomains. The active and silent RGCs were at either end of the retina, such that RGCs at the opposite ends of the retinal sheet were never co-active. Because the RGCs within the same retinal patch are anti-correlated in this simulation (Fig. 5.7), these afferents might be expected to cluster less than they do in the biological system, where they fire independently (Wong et al., 1993 and Daniel Butts, personal communication).

In spite of this anti-correlation, the model still supports layer and sublayer segregation (Fig. 5.8). The factor working in favor of eye-specific segregation is the "chaining" of the tendency of neighbors to cluster their afferents together. It follows that the transient/shifting nature of the subdomain boundaries is crucial to the global segregation of afferents, because this shifting prevents subdomain boundaries from chronically supporting sharp drops in correlations between the activity of neighbors on opposite sides of the subdomain boundaries. The length of time needed to achieve layer segregation is more prolonged in this model than in the models with only one subdomain. This observation can help explain why the segregation of layers is unrealistically fast in most of the simulations ($\sim$1 day versus $\sim$7-9 days in the biological system), because the biological system has to rely on the chaining effect over a relatively large distance in order to support the global clustering of afferents from each eye. Intuition suggests that as the number of subdomains increases (i.e as the retinal size increases, or the wave duration decreases) the period of activity-dependent layer segregation should become longer.

[3] Along with more rigorous matching of the simulated waves to the in vitro measurements and further data on the mechanisms of synaptic change, the model biological time-course of segregation could serve as a useful constraint on the model parameters. For example, the learning rate ($h$) that supported an appropriate time-course of development could provide a rough estimate of the strength of synaptic change mechanisms.

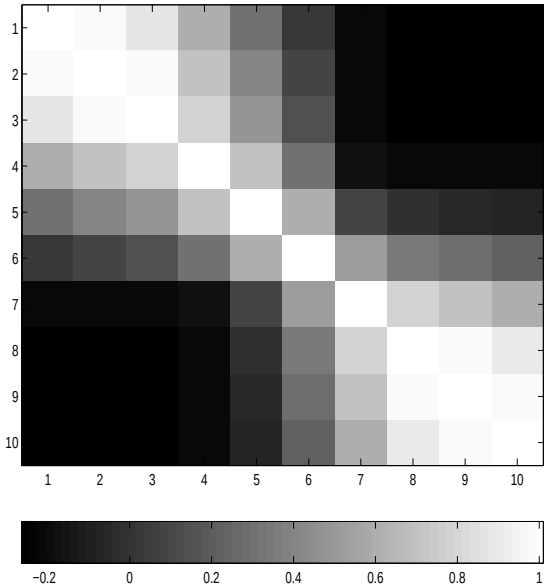


Figure 5.7: The correlation matrix of the RGCs in one patch during non-overlapping waves. The color-bar below indicates the correlation value corresponding to the plot above. Note that the correlations are negative between RGCs that are separated by large distances. See Sec. 4.3.2 for details of the activity patterns and their importance.

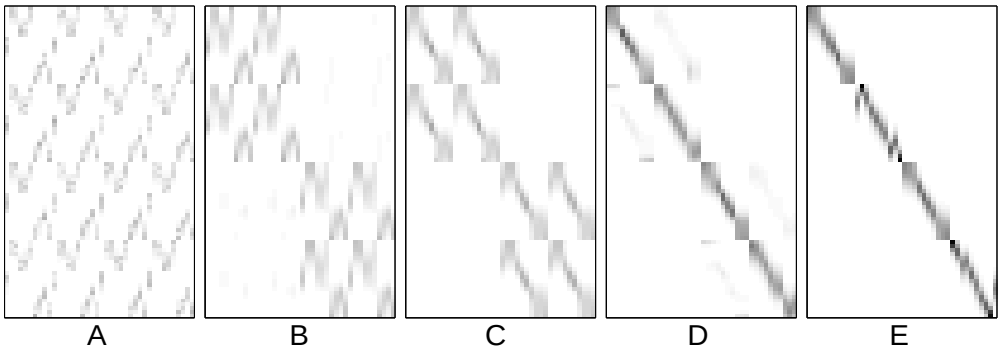


Figure 5.8: Multiple sub-domains of retinal activity can support layer and sublayer segregation. Weight matrix with non-overlapping waves (see text for details, Sec. 5.5) is shown at ages (A) D41, (B) D43, (C) D47, (D) D56, (E) D62.

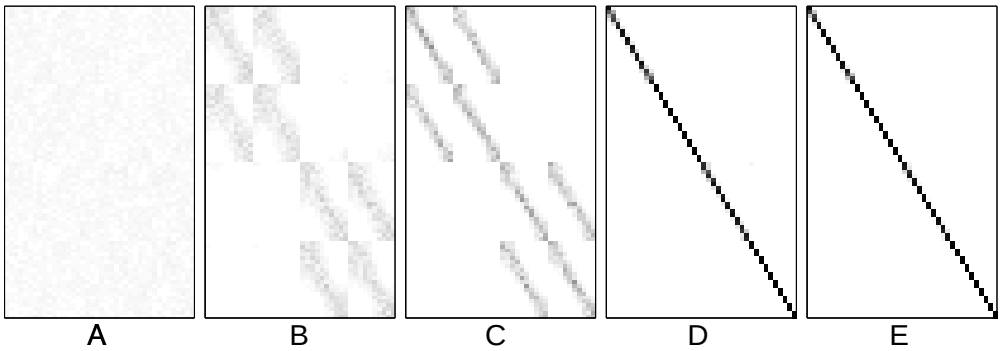


Figure 5.9: Assuming exuberant initial connectivity lateral interactions can support the development of retinotopy, eye-specific layer segregation, and on/off sublayer segregation in the absence of sprouting. Weight matrix with exuberant initial conditions and no sprouting is shown at ages (A) D41, (B) D42, (C) D45, (D) D56, (E) D62.

## 5.6 Blurring and Sprouting in the dLGN

The effects of sprouting and those of lateral influence in the dLGN are largely redundant. Both support retinotopic refinement, and the asymmetric nature of both supports layer and sublayer segregation.

### *5.6.1 Blurring*

Lateral interactions alone can support the development of retinotopy, eye-specific layers and on/off sublayers, Fig. 5.9. Because arbors cannot shift their positions appreciably without sprouting processes, it was necessary to assume exuberant initial connectivity (i.e. every RGC synapses on every dLGN relay cell) to ensure that convergence was possible. This condition was satisfied by increasing the width of the initial RGC axonal arbors to $10\times$ the canonical width (s.d. $= 380\mu m$).

In previous models it has been found that maps converge more robustly when the lateral influence in the dLGN starts off wide-ranging and then slowly decreases over development (a "shrinking neighborhood function"). This shrinking influence is instantiated in the current model as a shrinking in the parameter $\lambda$ (Fig 4.6, corresponding to an increase in the inhibitory conductances, Sec. 4.3.4). In the current parameter regime, this shrinking is not necessary for proper convergence. Simulations with $\lambda = 1$ or $\lambda = 20$ throughout development had results very similar to those shown in Fig. 5.9. There are some inappropriately placed synapses in the case of $\lambda = 1$, but convergence is nearly perfect. In my experience with other parameter regimes, static large or small lateral interactions do make convergence to layer segregation and retinotopy less robust, but not markedly so.

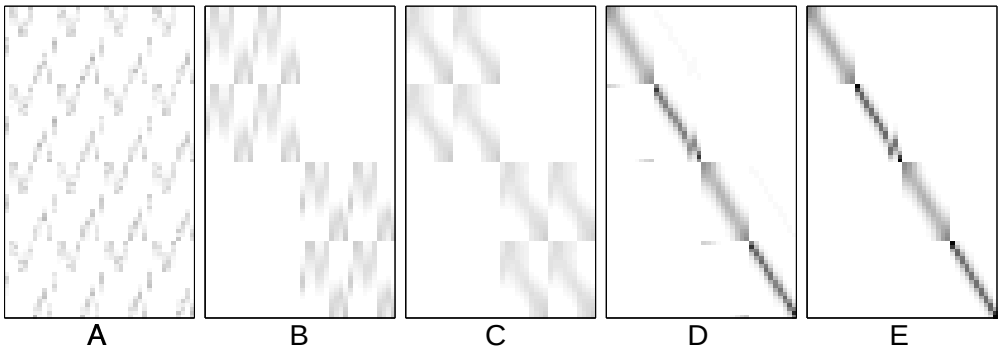


Figure 5.10: Sprouting can support the development of retinotopy, eye-specific layer segregation in the absence of lateral excitatory interactions in the dLGN. Weight matrix with no lateral interactions is shown at ages (A) D41, (B) D43, (C) D47, (D) D56, (E) D62.

### *5.6.2 Sprouting*

Sprouting alone can support the development of retinotopy, eye-specific layers and on/off sublayers, Fig. 5.10. While previous models with only sprouting have demonstrated that sprouting can support the emergence of retinotopy and eye-specific layers (Keesing et al., 1992; Eglen, 1997), attempts to support on/off specific segregation with only sprouting have been unsuccessful (see Sec. 7.2.2 and Eglen, 1997). The ability of this model to support on/off sublayer segregation with only sprouting is probably due to the presence of asymmetries in sprouting and biases toward sublayer segregation.

While sprouting and lateral interactions redundantly support the clustering of correlated afferents, sprouting also allows arbors to shift gradually over development. Thus, it is possible that lateral interactions are not necessary for activity-dependent retinogeniculate refinement, but some form of sprouting is probably necessary to ensure that arbors can shift their positions over development.

## 5.7 No On/Off Activity

Because the iterative on/off bias and the asymmetries in the model do not independently support sublayer segregation, sublayers fail to segregate in the absence of differentiated on/off RGC activity, Fig. 5.11.

## 5.8 Biases

In order to illustrate the role of the initial biases (Sec. 4.3.3) in terms of the model behavior, a simulation was run in which there was no bias toward retinotopy, eye-specific layer segregation, or on/off sublayer segregation. As illustrated in Fig. 5.12 the simulation can develop continuous

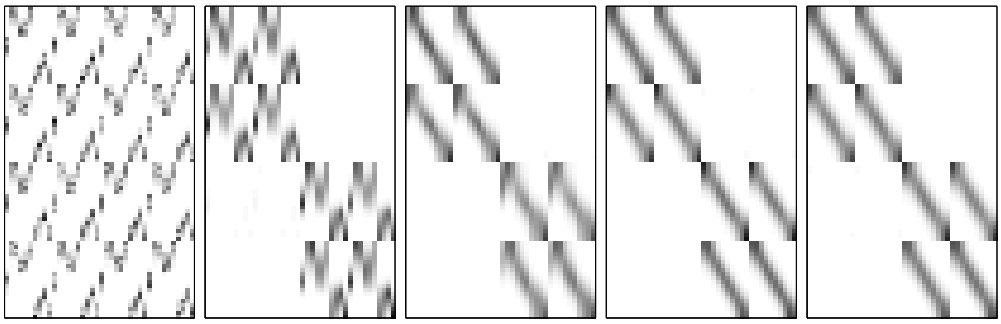


Figure 5.11: Differentiated on/off RGC activity is crucial to the development of on/off dLGN sublayers. Weight matrix with no noisy off-cell firing is shown at ages (A) D41, (B) D42, (C) D45, (D) D56, (E) D62. The weights are darker during the initial parts of development because they are larger relative to the blurry and unsegregated final state of the projection than to a normal mature projection.

retinotopy, eye-specific layers, and on/off sublayers without any biases. Note that, as one would expect, the mature configuration is not necessarily stereotypical without biases.

More surprising is the fact that, without biases, convergence toward any segregated configuration is more fragile, especially convergence to segregated on/off sublayers. This suggests that biases not only support a stereotypical final configuration, but they also help to support the layer and sublayer segregation in general. On/off sublayer segregation is probably especially dependent on biases because the correlational structure in the retinal activity differentiates the on/off RGCs less than it differentiates the RGCs from the two eyes.

Simulations suggest that it is possible to compensate for the influence of the missing biases by increasing the influence of asymmetries on layers segregation. Lowering the the inter-layer and inter-sublayer interactions (i.e. from .1 to .001) makes layer and sublayer segregation fairly robust.

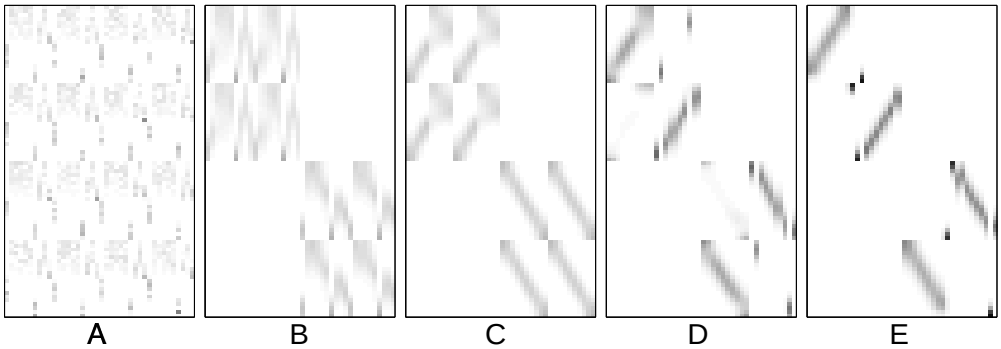


Figure 5.12: A simulation without any biases and with the inter-layer and inter-sublayer interactions = .001 can develop retinotopy, eye-specific layer segregation, and on/off sublayer segregation. Weight matrix is shown at ages (A) D41, (B) D42, (C) D54, (D) D56, (E) D62.

### 5.8.1 Alternate Types of Bias

The actual form of the biases is not crucial to proper convergence. The canonical model uses an ingrowth bias (one step, at initialization) for the eye-specific layers, and a trophic bias (iteratively applied every step) for the on/off sublayers. Using trophic biases or an ingrowth bias for both layers and sublayers also supports convergence to the stereotypical mature state. However, the ingrowth bias for the on/off sublayers must be enforced at the onset of on/off sublayer segregation or it attenuates before the period of sublayer segregation.

## 5.9 Asymmetries in Sprouting and Lateral Interactions

### 5.9.1 No Asymmetries

Asymmetries in the intra-dLGN interactions are crucial to the proper convergence of the model (see Fig. 5.13). In the canonical model the inter-sublayer influence is 10% of the strength of the intra-sublayer influence between nearest neighbors. Likewise, RGC axonal arbor sprouting is stronger within a layer than between layers, by the same factor. To my knowledge this use of asymmetrical lateral interactions to support layer formation is unique to this model, and has not been tested biologically.[4]

Without these asymmetries rough retinotopy emerges, but each RGC arbor innervates all sublayers. In fact, each RGC arbor duplicates itself in all of the sublayers. This duplication between layers arises in the absence of asymmetries because structures in one layer can cooperatively reinforce similar structures in adjacent layers. When this tendency is strong, the structures in all the sublayers are identical, as seen in Fig. 5.13.

### 5.9.2 Static Asymmetries

Although asymmetries are assumed to increase over the course of development in the canonical model (see Fig. 4.7), this assumption is not necessary for appropriate convergence. A model in which the inter-layer and inter-sublayer weights remain .1 over development robustly develops layer and sublayer segregation. However, even with the bias toward a stereotypical layer configuration, the sublayer configuration robustly ends up switched such that the on-center cells cluster together at the border between the layers (see Fig 5.14). That is, the on-center RGCs from the two eyes end up projecting to the outer sublayer of layer A and the inner sublayer of layer A1 (see Fig. 5.1).

[4]Previous analytic work (Miller, 1996) suggests that the importance of asymmetries to the formation of layers and sublayers in the dLGN is a specific instance of the tendency of lateral interactions to determine the layout of topographic maps. Likewise, the development of precise receptive fields (eye and on/off specific) is a specific instance of the tendency of correlations in the inputs to determine receptive field structure.

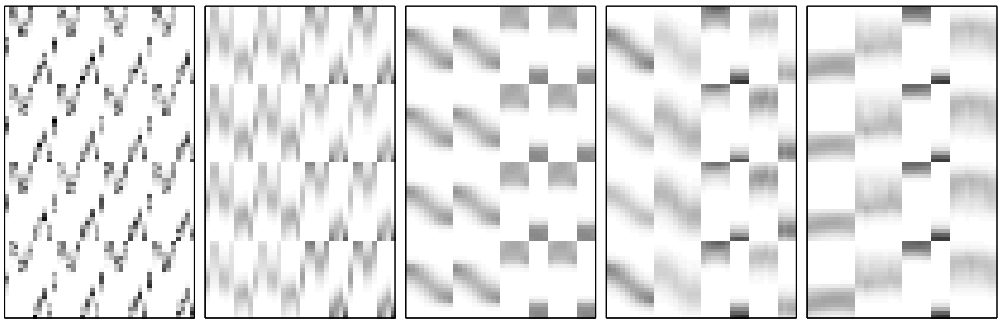


Figure 5.13: Asymmetries in intra-dLGN interactions and RGC sprouting are crucial to model convergence. Weight matrix with no asymmetries in either sprouting or intra-dLGN interactions is shown at ages (A) D41, (B) D42, (C) D48, (D) D56, (E) D62.

Unlike the canonical asymmetries, static asymmetries provide excitatory interactions between the layers during the period of on/off sublayer segregation. Indeed, the stereotypical configuration re-establishes itself if this inter-layer excitatory interaction disappears (i.e. inter-layer interactions = 0) during the period of on-off segregation (see Fig. 5.15). The clustering probably arises because the correlations between the on-RGCs from the two eyes are higher than the correlations between the on-RGCs from one eye and the noisily firing off-RGCs from the other eye. Because dLGN cells that have an excitatory influence on each other tend to have more correlated inputs, the inter-layer excitation supports the grouping of the on-center afferents from the two eyes.

### *5.9.3 No Inter-Layer/Sublayer Interactions*

An absence of inter-layer and inter-sublayer interactions also disrupts appropriate model convergence (Fig. 5.16). It seems that the retinotopic maps in the sublayers need mutual reinforcement,

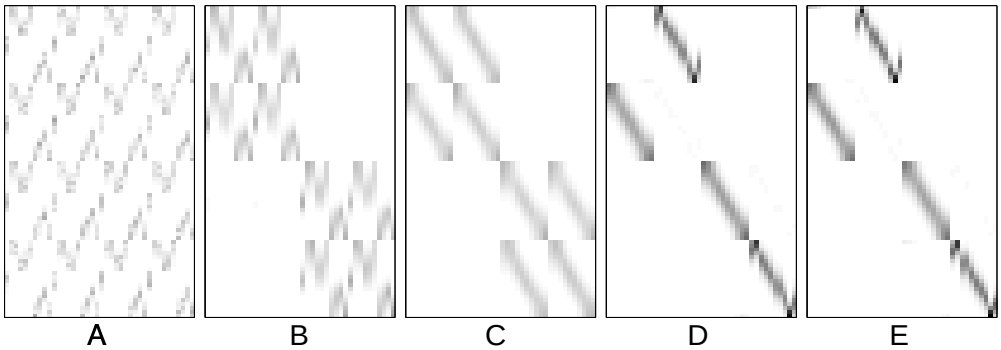


Figure 5.14: Static asymmetries in intra-dLGN interactions and RGC sprouting support the segregation of layers and sublayers, but not in the stereotypical configuration. Weight matrix with inter-layer and inter-sublayer interactions fixed at .1 is shown at ages (A) D41, (B) D42, (C) D47, (D) D56, (E) D62.

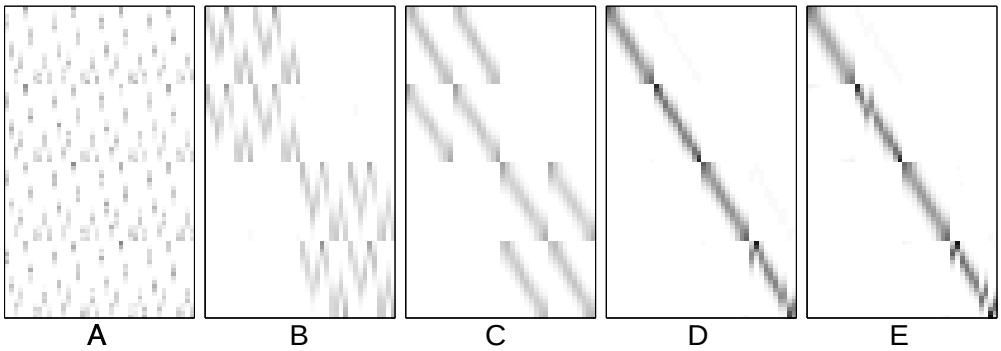


Figure 5.15: Static asymmetries in intra-dLGN interactions and RGC sprouting, coupled with the disappearance of inter-layer interactions during on/off sublayer segregation, supports proper model convergence. Weight matrix with inter-layer and inter-sublayer interactions = .1 from D41-D55, and inter-layer interactions = 0 from D55-D62 is shown at ages (A) D41, (B) D42, (C) D47, (D) D56, (E) D62.

or they will diverge and eventually fracture. Although each RGC arbor only innervates one sublayer, RGCs from the same patch do not segregate into the same sublayer.

In summary, inter-layer and sublayer interactions must be present and much weaker than intra-sublayer interactions to support layer and sublayer segregation as well as retinotopic refinement. In addition, the inter-layer interactions must decrease to 0 (at least) around the onset of on/off segregation in order to prevent on-center afferents from the two eyes from clustering in adjacent sublayers.

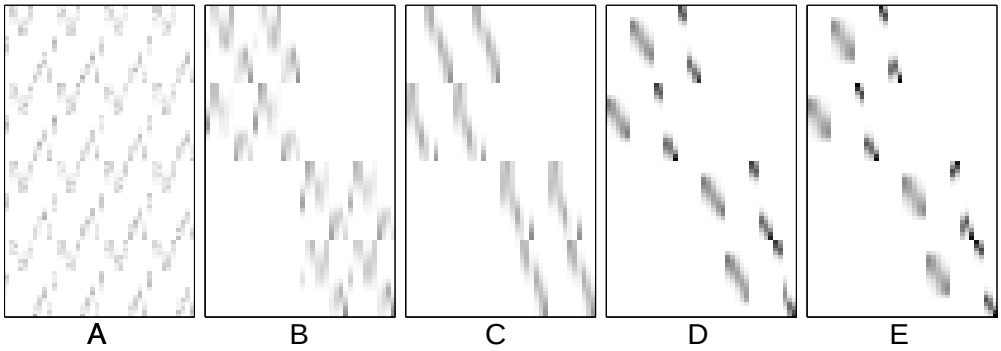


Figure 5.16: Layer segregation can occur in the absence of interactions between dLGN layers and sublayers, but retinotopy and sublayer segregation require some interactions. Weight matrix with inter-layer and inter-sublayer interactions = 0 is shown at ages (A) D41, (B) D42, (C) D47, (D) D56, (E) D62.

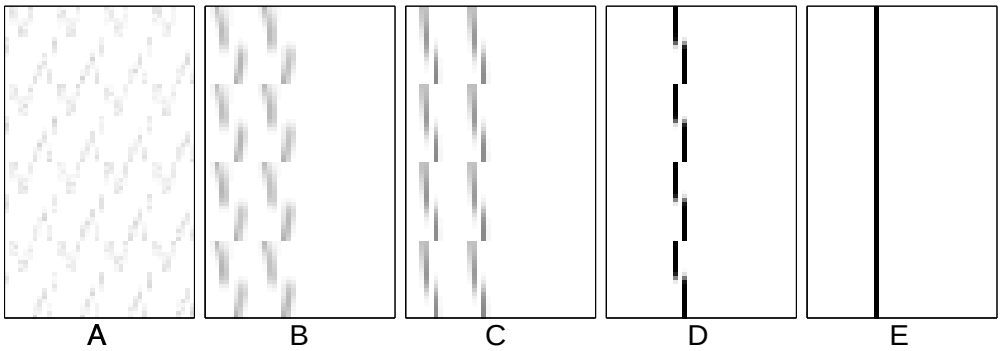


Figure 5.17: In the absence of pre-synaptic normalization all of the synaptic weight concentrates in one or a few RGCs, leaving the other RGCs without synapses. Weight matrix with no pre-synaptic normalization is shown at ages (A) D41, (B) D42, (C) D45, (D) D56, (E) D62.

## 5.10 Pre- and Post-Synaptic Normalization

Both pre and post-synaptic normalization are crucial to the model's stability. As shown in Fig. 5.18, normalizing only the pre-synaptic weights results in a handful of dLGN cells receiving all of the retinal synapses. Likewise, Fig. 5.17 results in one RGC dominating the entire retinal projection. This sharpening of the retinogeniculate projection is supported by the combination between scaled synaptic weight update and multiplicative normalization.

### *5.10.1 No Pre-Synaptic Normalization*

In the absence of pre-synaptic normalization, the total synaptic strength of an RGC is only constrained by the number of dLGN cells and their total synaptic strength. Under these conditions one or a few RGCs eventually dominate the entire projection (see Fig. 5.17). This concentration of synaptic weight is possible because activity spreads in the dLGN. A strong synapse will also tend to reinforce the growth of synapses from the pre-synaptic RGC to nearby dLGN cells (sprouting has a similar effect). Eventually, this process tends to amplify the projection of an RGC with strong synapses until it dominates all post-synaptic cells. During periods that the dominant RGC is silent, other RGCs will be able to compete for resources, but when the dominant RGC is active it will more than recoup its losses. Once discrepancies arise between the total synaptic strength of RGCs, a dominant RGC will tend to increase the gap between it and the other neurons over development.

### *5.10.2 No Post-Synaptic Normalization*

In the absence of post-synaptic normalization, the total synaptic strength of a dLGN cell is only constrained by the number of RGCs and their total synaptic strength. Under these conditions one or a few dLGN cells eventually dominate the entire projection (see Fig. 5.18). One mechanism that

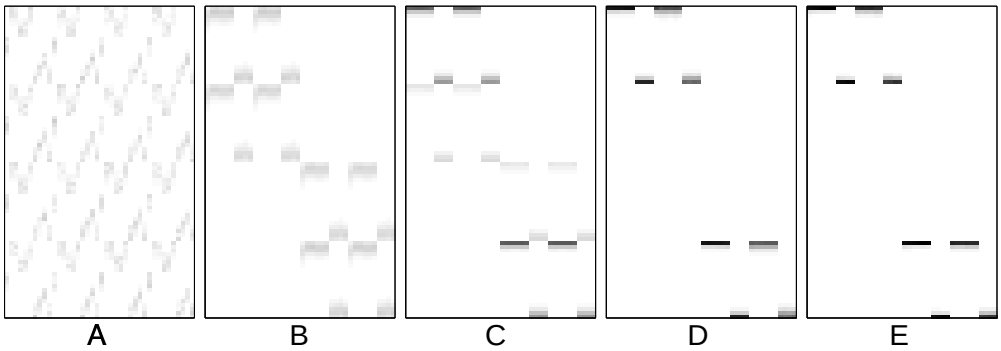


Figure 5.18: In the absence of post-synaptic normalization all of the synaptic weight concentrates in one or a few dLGN cells, leaving the other dLGN cells without synapses. Weight matrix with no post-synaptic normalization is shown at ages (A) D41, (B) D42, (C) D45, (D) D56, (E) D62.

supports this concentration of synaptic weight is that a post-synaptic neuron with a large incoming synaptic strength will tend to be active more than a neuron with less incoming synaptic strength. Due to this fact, any synapses from RGCs onto a strongly connected dLGN cell will tend to be strengthened more at all activity levels (save silence) than the synapses from the same RGC to other dLGN cells, because those other dLGN cells are, on average, less active and the synaptic change is scaled by the post-synaptic activity. The result is that the connections from all RGCs to a strongly connected dLGN cell will tend to grow faster than synapses to less strongly connected dLGN cells.

The above description will only be accurate to the extent that all RGCs connect to all dLGN cells. Sprouting pushes synaptic connectivity toward this state, although not completely and especially not between layers (due to asymmetries). The result is that the dominant dLGN cell can only capture the retinal inputs to its sublayer, rather than all the retinal inputs. After development this results in one cell in each sublayer that captures all the retinal synapses to that layer (see Fig. 5.18).

### *5.10.3 Capped Normalization and Monocular Deprivation*

Monocular deprivation results in the deprived eye losing much of its projection to the binocular region of the dLGN, and the spared eye expanding its axons into the layers that it would normally not innervate. The canonical form of the current model cannot account for these results, because the total synaptic weight supported by the pre-synaptic cells is tightly constrained (normalized) and independent of activity. However using a variant of normalization ("capped normalization", see Sec. 4.3.7), derived by analogy to Eglen's work (Eglen, 1997, and Sec. 7.2.2), it is possible to model some of the monocular deprivation results in the current framework. Capped normalization allows RGC axonal arbors to sustain a total synaptic strength that is small or 0, but limits their growth to a maximal value ($\theta$, see Sec. 4.3.7).

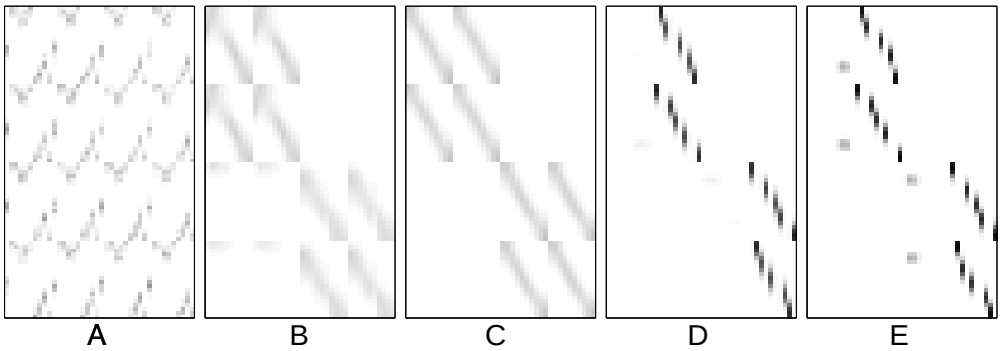


Figure 5.19: Capped normalization supports the development of rough retinotopy and eye-specific layers, but fails to support the development of on/off sublayers. Weight matrix with capped normalization ($\theta = 2.33$) is shown at ages (A) D41, (B) D45, (C) D48, (D) D56, (E) D62.

*Canonical Inputs* Capped normalization supports the development of retinotopy, and eye-specific layer segregation, but not on/off sublayer segregation. The results of simulating the model development with this form of normalization (with $\theta = 2.33$) is shown in Fig. 5.19. The specific value of $\theta$ was chosen to best support normal development with canonical inputs and appropriate development with monocular inputs. Lower values of $\theta$ tend to form eye-specific layers in both cases, and higher values tend to support the domination of the projection by a single eye in both cases.

Note that, because they have a higher activity level, the off-RGCs come to dominate the projection during the period of on/off sublayer segregation. This finding highlights a potential dilemma in modeling monocular deprivation and on/off layer segregation simultaneously. The deprivation results suggest that RGCs with low activity levels will tend to lose their synaptic connections. Yet, there are no obvious asymmetries between the normal projections from the on and off-RGCs — even though the on-RGCs are only 25% as active as the off-cells during the period of on/off segregation. Resolution of this dilemma could be provided by: undiscovered asymmetries in the on and off-RGC projections to the dLGN; added complexities in the relation between pre-synaptic activity and total synaptic strength; or added complexities in the dynamics of weight change.

*Monocular Deprivation* When deprivation conditions are simulated by silencing one retina, capped normalization yields results in accordance with deprivation experiments (see Sec. 3.3.7). In this simulation the deprived eye loses its projection to the dLGN (the binocular region that is the focus of the current model), while the spared eye takes over the entire dLGN (Fig. 5.20).

In this case, the off-RGCs are prevented from dominating the monocular projection because $\theta$ will only allow them to support $\sim 2\times$ their normal synaptic strength, where their domination of the entire projection would require them to support $4\times$ their normal synaptic weight. The on-RGCs retain a portion of the synaptic weight because there is no competition for their resources, analogous to the silent RGCs in the canonical model with monocular deprivation.

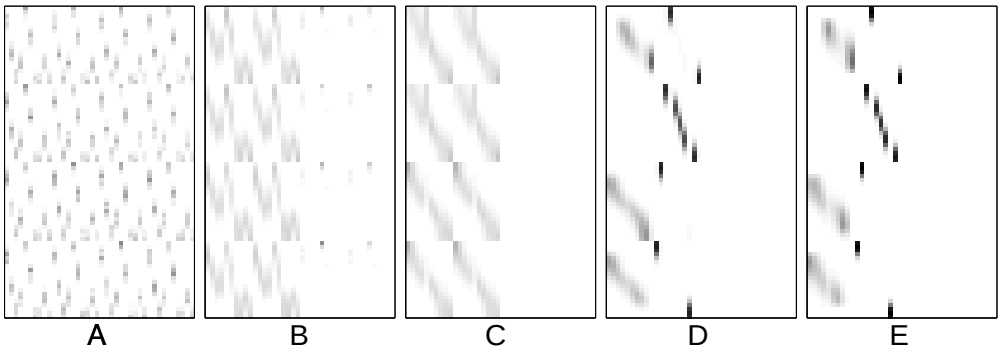


Figure 5.20: Capped normalization with monocular retinal deprivation (silencing) results in the domination of the retinogeniculate projection by the spared eye. In this monocular projection there is rough retinotopy, and some on/off layer segregation. Weight matrix with capped normalization ($\theta = 2.33$) and monocular deprivation is shown at ages (A) D41, (B) D42, (C) D44, (D) D57, (E) D62.

## 5.11 Summary

It was demonstrated that the canonical version of the model, as well as versions with only sprouting or lateral excitatory interactions in the dLGN can support the development of precise retinotopy, eye-specific layer segregation and on/off sublayer segregation. It was further shown that the model shows robust convergence at different temporal and spatial scales. The effects of some of the parameters and mechanisms in the model were discussed and illustrated.

# Chapter 6

# Formal Analysis

Understanding the behavior of the model through simulations is difficult because there are simply too many parameters to explore simultaneously. A more powerful approach is to understand the dynamics of the equations and the effects of the parameters on those dynamics by addressing the mathematics of the model at an analytic level. This type of approach can point out the crucial aspects of the model and even, in some cases, predict its dynamics without resorting to simulations.

## 6.1 Analytic Difficulties

In its main formulation, the model is quite difficult to approach analytically. There are three problems: the model depends on normalizing the weights over both the pre- and post-synaptic cells; the model utilizes limited connectivity and sprouting; and the weight dynamics in the model are non-linear because the weight update is dependent on the previous weights.

First, the use of both pre and post- synaptic normalization generates two conflicting constraints that prevent the exact satisfaction of either constraint. The resulting dynamics do not conform to any constrained objective function. Therefore, the approximate normalization used in the model prevents exact interpretation of the model in terms of previous analytic work (Wiskott and Sejnowski, 1997; Wiskott and Sejnowski, 1998; Miller and MacKay, 1994), and other models that enforce either pre or post-synaptic normalization, but not both (e.g. Miller et al., 1989).

Second, the model assumption of limited connectivity coupled with sprouting, although biologically realistic, prevents direct comparison with previous models that use all-to-all connectivity for analytic and computational simplicity. This difficulty is mitigated by the fact that sprouting, at least in the case of one model, has been shown to be formally equivalent to more traditional fully connected models (Miller, 1997b). The equivalence between the models depends on the *possible* positions of the arbors being equivalent in the two models (i.e. that their "arbor functions" are equivalent), and on both models having subtractive normalization.

Third, explicitly scaling the weight update by the previous weight makes the weight update

intractable to linear analysis. Fortunately, guided by previous analytic work (Wiskott and Sejnowski, 1998; Miller and MacKay, 1994), it was possible to construct an alternative formulation of the model that makes the normalization subtractive and simultaneously addresses the third problem by making the synaptic weight change independent of the previous weight (except as it influences dLGN activity).

## 6.2 Alternate Formulation

The recent work of Wiskott and Sejnowski (Wiskott and Sejnowski, 1998; Wiskott and Sejnowski, 1997) has created a mathematical foundation for translating a model in which the synaptic weight change is scaled by the existing weight into a more traditional, linear and tractable model in which the synaptic weight change is not explicitly effected by the previous weight. This change is coupled with a change from multiplicative normalization to subtractive normalization. Wiskott and Sejnowski's analysis, as well as previous analytic work (Miller and MacKay, 1994), suggests that the changes in the weight growth rule and the normalization rule compensate for each other's influence.

Wiskott and Sejnowski's analysis applies in a straightforward fashion to models with either pre- or post-synaptic normalization, but not both. Difficulties arise when both forms of normalization are enforced simultaneously (as in the current model), because the constraints interfere with each other — resulting in inexact satisfaction of both constraints. In order to construct a subtractive version of the current model, I followed Wiskott and Sejnowski's discussion of a canonical case with only pre- or post- synaptic normalization and generated an analogous subtractive version for the current model. The resulting model is thus motivated by their analysis, but not rigorously derived from it.[1] The resulting model was then validated through simulations, as discussed in the next section.

Wiskott and Sejnowski outline a model with scaled Hebbian learning and a multiplicative post-synaptic normalization rule that is described by the equations:[2]

[1]Formally, Wiskott and Sejnowski translate from one set of weight update and normalization equations to another by conceptualizing the model as a set of dynamics (driven by weight update) that minimize an objective function subject to constraints (the normalization procedure). The model dynamics are framed as an optimization problem, where the model eventually settles into a minimum of the objective function. The benefit of this approach is that the same objective function can be minimized in more than one way. Two sets of equations that minimize the same objective function will result in similar end states even if their paths (their dynamics) are different. If the end state of one of the sets of equations is predictable analytically, the end point of the other set is known as well. Alternate sets of equations are derived by transforming the objective function via coordinate transformations (i.e. changes of basis) that preserve the minima in the objective function, but change its error surface. This changed error surface changes the dynamics of the minimization and the constraints. Because the set of transformed equations minimizes the same objective function (under a coordinate transformation) the two sets of equations can be taken to be roughly equivalent in terms of the mature structures that they support.

[2]The notation and approach in Wiskott and Sejnowski (1998) is quite different from that used in this thesis. For purposes of clarity and consistency, I use terminology that is consistent with the current discussion. This departure from Wiskott and Sejnowski's framework results in some loss of fidelity to the details of Wiskott and Sejnowski's approach.

$$\frac{\delta w_{[xy]}}{\delta t} = w_{[xy]}\, l_{[x]}\, r_{[y]} \tag{6.1}$$

$$w_{[xy]} = \frac{\tilde{w}_{[xy]}}{\sum_y \tilde{w}_{[xy]}} \tag{6.2}$$

They show that these equations minimize the same objective function (under a coordinate transformation) as the following equations that enforce unscaled Hebbian learning and a subtractive normalization rule:

$$\frac{\delta w_{[xy]}}{\delta t} = l_{[x]}\, r_{[y]} \tag{6.3}$$

$$\begin{aligned} w_{[xy]} &= \tilde{w}_{[xy]} + \frac{1}{R}(1 - \sum_y \tilde{w}_{[xy]}) \\ &= \tilde{w}_{[xy]} - \frac{\sum_y \tilde{w}_{[xy]}}{R} + \frac{1}{R} \end{aligned}$$

$\tilde{w}_{[xy]}$ denotes the updated and unnormalized synaptic weight, and if $\tilde{w}_{[xy]} < 0,$ then $w_{[xy]} = 0.$

By analogy I constructed a normalization rule that applies to the simultaneous pre and post-synaptic normalization used in the current model:[3]

$$\begin{aligned} \frac{\delta w_{[xy]}}{\delta t} &= l_{[x]}\, r_{[y]} \\ w_{[xy]} &= \tilde{w}_{[xy]} - \frac{1}{2}(\frac{\sum_x \tilde{w}_{[xy]}}{L} + \frac{\sum_y \tilde{w}_{[xy]}}{R}) + \frac{2}{L+R} \end{aligned}$$

constrained such that:

$$\begin{aligned} &\text{if } \tilde{w}_{[xy]} < 0, w_{[xy]} = 0 \\ &\text{if } \tilde{w}_{[xy]} > 1, w_{[xy]} = 1 \end{aligned}$$

The factors of 2 are added to counteract the summing operations over both pre and post-synaptic indices, and the second constraint is added because this normalization is approximate, so that even after normalization some cells can support more synaptic weight than maximally allowed.

[3] It is possible that there are other unscaled subtractive formulations that have the desired convergence behavior, but constructing one is not trivial. Before learning of Wiskott and Sejnowski's work, I made several unsuccessful attempts at constructing unscaled subtractive versions of the model. David Heeger's help was crucial to the formulation of this normalization rule.

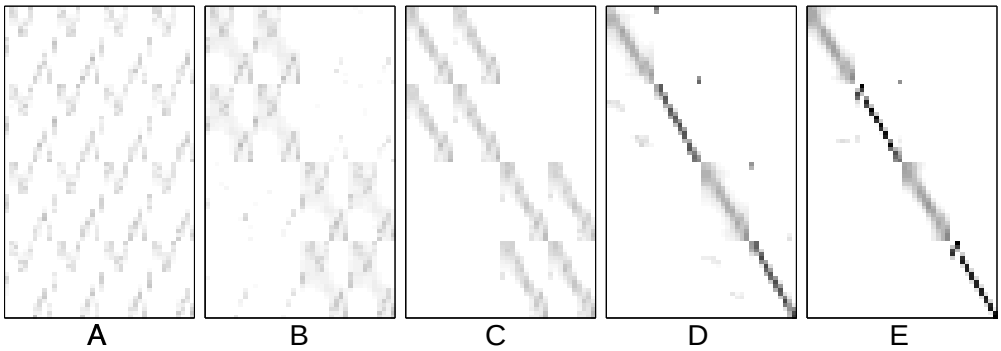


Figure 6.1: The subtractive version of the model can support the development of retinotopy, eye-specific layer segregation and on/off sublayer segregation. Weight matrix with $h = 5 \times 10^{-5}$ is shown at ages D41 (A), D43 (B), D51 (C), D58 (D), D62 (E).

In combination with the unscaled Hebbian learning rule (Eq. 6.3), this subtractive normalization rule was predicted to generate similar convergence behavior to the canonical model.[4] In addition, the model using these equations is tractable to linear analysis because the weight change does not explicitly depend on the previous weight.

### *6.2.1 Simulation Results*

I ran some simulations using this "unscaled subtractive" set of equations in order to check the validity of the construction. With the learning rate lowered by a factor of 100 (to $5 \times 10^{-5}$), the simulations yielded results similar to the former set of equations (see Fig. 6.1), as predicted. The latter set of equations describes an analytically tractable version of the model that generates retinotopy, eye-specific layers and on/off sublayers.

Above this reduced learning rate the subtractive simulations did not robustly display convergence. The fragile nature of the subtractive version is possibly due to hard limits imposed on the weights after subtractive normalization to prevent them from increasing beyond 1 or decreasing below 0. Large jumps past these limits might introduce unstable dynamics in the weight change. Alternatively, the tendency of the subtractive dynamics to strongly sharpen the projection could cause the weight configuration to become locked into local minima relatively easily.

[4]There may be some empirical reasons to favor the one formulation over the other. For example, some research has shown that at least some types of LTP and LTD seem to act multiplicatively (Turrigiano et al., 1996), but the dynamics of synaptic change are still relatively unknown.

## 6.3 Linear Analysis

The dynamics of this unscaled subtractive model are tractable to linear analysis. In linear analysis, the dynamics of the weights are related to the possible configurations of weights (the arbor function), the lateral influences in the target tissue (the neighborhood function), and the correlational structure of the input patterns. The trick is to determine the fastest growing patterns in the weights (via eigenvector analysis), and then extrapolate the ultimate weight pattern (Haussler and von der Malsburg, 1983; Miller, 1990; Miller and MacKay, 1994).

### *6.3.1 Weight Matrix Eigenmodes*

My analytic approach, guided by previous analytic work[5], was to determine the eigenmodes of the weight matrix assuming canonical retinal activity patterns and intra-LGN interactions (see Appendix C for an explanation and derivation of the eigenmodes). The eigenmodes can be interpreted as the fastest growing components of any weight matrix in the unscaled subtractive model. The particular modes that dominate the eventual projection will depend on relative growth rates as well as biases and noise in the initial projection and stochasticity in the retinal activity.

The 25 largest eigenmodes (and growth rates/eigenvalues) of the specific dLGN interactions and retinal activity patterns used in the unscaled subtractive (and canonical) version of the current model are shown in Figs. 6.2 and 6.3. For purposes of indicating individual eigenmodes, I will refer to them by number (1-25) in decreasing order of eigenvalue (i.e. left to right, top to bottom). Note that the influence of sprouting is ignored. More accurately, it is effectively assumed that each RGC arbor initially projects to all LGN neurons.

The actions of some of the eigenmodes are discussed below. It should be kept in mind that the eigenmodes act in parallel during development. For example, the combined action of horizontal and vertical gradients (e.g. Fig. 6.2, 6 and 7) could amplify a retinotopic projection within each of the 16 sub-projections.

*D41: Eye-Specific Layer Segregation* Fig. 6.2 shows the eigenmodes during the early part of eye-specific layer segregation, when there is no noisy off-RGC firing, the inter-sublayer connections are small (.1), and the inter-sublayer connections are large (1).

During the period of eye-specific layer segregation, the two fastest growing eigenmodes (Fig. 6.2, 1 and 2) strongly differentiate the inputs from the two eyes. Eigenmode 2 directly supports segregated projections from each eye to a separate layer, and Eigenmode 5 supports differential growth of weights to the two layers. Eigenmodes 13, 15, 19 and 21 will amplify retinotopic projections, because they selectively amplify weights along the diagonals of each of the 16 sub-quadrants.

[5] I could not have done this analysis without Ken Miller's expertise and generous guidance.

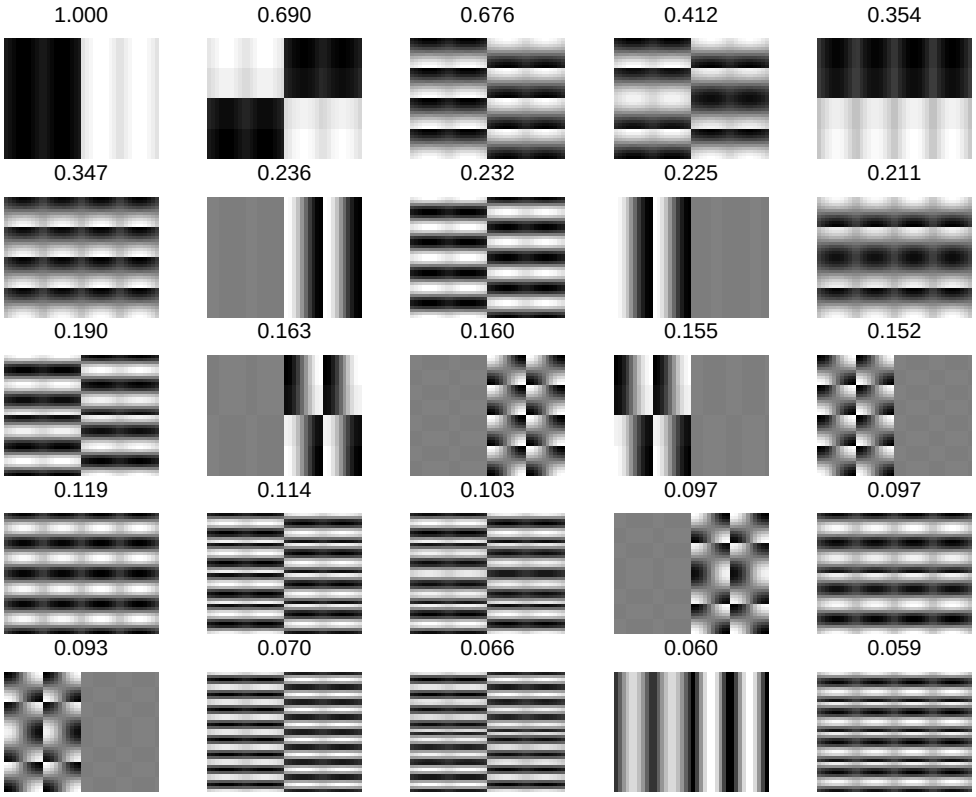


Figure 6.2: The 25 fastest growing weight patterns (eigenmodes) during eye-specific layer segregation (D41-D55) in the canonical model. Above each eigenmode is its rate of growth, normalized to make the fastest rate of growth = 1. Actual fastest growth rate is 10000.

*D55: On/Off Sublayer Segregation* Fig. 6.3 shows the eigenmodes during the early part of on/off sublayer segregation when the off cells are dominated by noisy firing, the inter-layer connections are zero, and the inter-sublayer connections are small (.1).

Comparing Fig. 6.2 to Fig. 6.3, it is apparent that many more of the eigenmodes during on/off sublayer segregation selectively influence the projection to one of the two layers. This tendency toward independent growth is probably due to the lack of inter-layer connections during this period. There is also a tendency for the eigenmodes to differentiate the inputs from the on and off-RGCs (4 vertical stripes, quite pronounced in 2,16 and 17) due to the addition of noisy off-cell firing. Like the eigenmodes during the period of eye-specific segregation, three of the dominant eigenmodes (1,2 and 5) generate eye-specific and layer-specific projections. The principle eigenmode (1) directly supports eye-specific layer segregation. In contrast to the eigenmodes during the period of eye-specific segregation, there are several eigenmodes that directly support differential projections to the

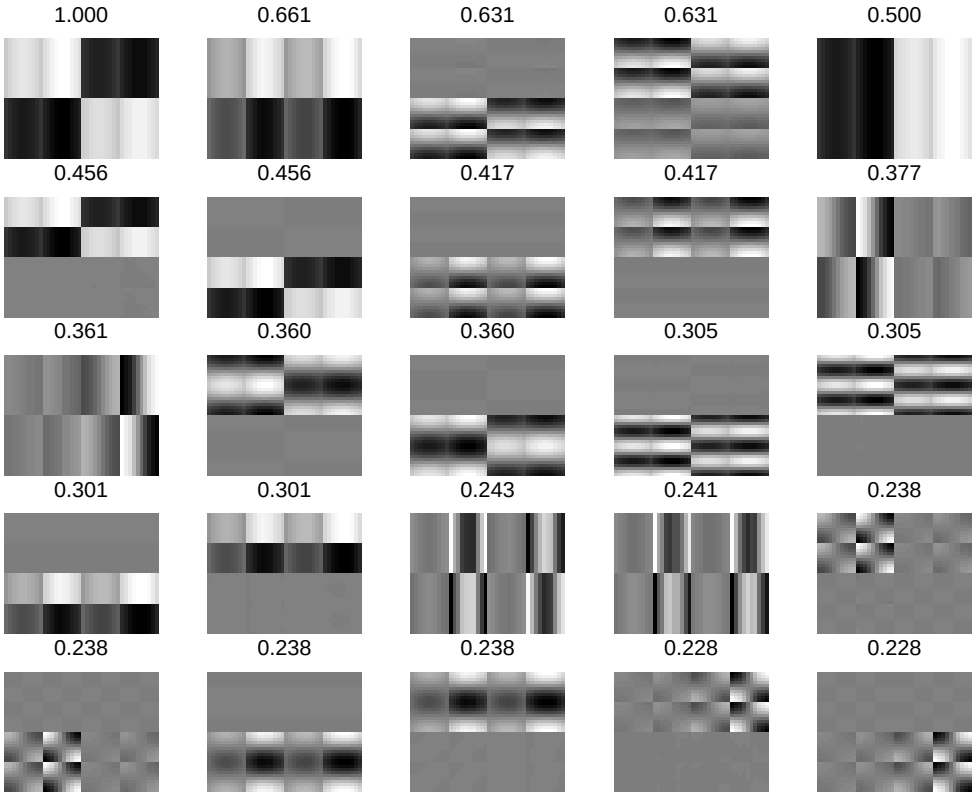


Figure 6.3: The 25 fastest growing weight patterns (eigenmodes) during on/off sublayer segregation (D55-D62) in the canonical model. Above each eigenmode is its rate of growth, normalized to make the fastest rate of growth = 1. Actual fastest growth rate is 500.

inner and outer sublayers from the two eyes (6 and 7) and the on/off cells (16 and 17). Presumably the differential growth between the sublayers arises because the inter-sublayer weights are weak during this period. The eigenmodes that support retinotopy during this period are (20, 21, 24, and 25).

## 6.4 Summary

An unscaled subtractive version of the model is presented. This version follows loosely from previous analytic work, converges appropriately in simulations, and is tractable to the type of linear analysis that has previously been applied to similar models. Linear analysis demonstrates that some of the fastest growing weight patterns from D42-D55 support eye specific layer segregation

and retinotopy. From D55-D62 some of the fastest growing patterns enforce the eye-specific layer segregation, the segregation of inputs to the sublayers, and the differential growth of inputs from the on- and off-RGCs.

This analysis of the model dynamics is useful because its helps formalize intuitive understandings of the model mechanisms. All of the major structures that are supported by the current model (eye-specific layers, on-off sublayers, and retinotopy) arise because of the specific mechanisms and assumptions of the model. These analyses indicate that the specific assumptions about intra-LGN interactions, Hebbian learning, and the realistic patterns of retinal activity will tend to amplify these patterns. The other implementational details are probably best viewed as providing biological realism and an environment that prevents spurious patterns ("local minima") from becoming entrenched, allowing the dominant eigenmodes to shape the global order of the projection.

# Chapter 7

# Previous Models

Self-organizing maps, and their relevance to neural circuit formation, have been explored in a variety of models since the 1970s. They have been applied to diverse systems including the development of retinotopy in the tectum, the organization of dermal topographic maps, the segregation of groups of sensory inputs, and the coordination between overlaid maps (e.g. ocular dominance, retinotopy, and orientation selectivity). Because there have been several extensive reviews of these models published recently (Swindale, 1996; Erwin et al., 1995; Eglen, 1997) this review will cover only a handful of the most relevant previous models. Notably this review will not cover models that concentrate on the role of trophic factors in activity dependent refinement (Whitelaw and Cowan, 1981; von der Malsburg and Willshaw, 1977), because trophic factors are only incorporated into the current framework as simple biases toward specific final configurations.

The details of the models differ substantially, but they all highlight the basic principle that local interactions can support the development of global order.

## 7.1 General

### *7.1.1 Willshaw and von der Malsburg*

The original starting point for the present model was the model developed by Willshaw and von der Malsburg in 1976 . They developed a simple model with center-surround lateral interactions that develops retinotopy based on spatially correlated inputs. Their model differs in several crucial respects from the current model: it does not encompass layer segregation, it uses abstract inputs, it uses a non-linear model of LGN receptive fields, and it assumes that inhibitory lateral interactions are present in the dLGN.

This model pioneered explorations into self-organizing maps. The point of its development was to formalize understandings of the development of topography, a problem made salient by work on growth, regeneration, and re-organization in the retinotectal projection in frogs, goldfish, and chicks

(see Fawcett, 1993 for review). Their work provided a non-trivial and powerful demonstration that local connections can bring about global order. While they were not concerned with segregation of inputs, the model principles can be interpreted to include segregation of inputs as a special case. That is, segregation of inputs is essentially the topographic mapping appropriate for inputs with dominant groups in the correlational structure.

Their model simulated visual inputs as clusters of co-activated RGCs. Because they were randomly placed, the inputs generated spatially graded correlations between model retinal cells. While these inputs lack many of the details of retinal waves, their correlational structure is similar to the correlational structure of the waves in a single patch of RGCs in the current model.

Their model instantiates activity in the tectum as a membrane potential that, above a given threshold, supports the generation of action potentials. The tectal potentials are iteratively updated with the lateral feedback from the other tectal cells until the activity settles to a steady state, analogous to dLGN update in the current model. The explicit inclusion of action potentials makes their model LGN activity more directly interpretable than the current model, unfortunately it also introduces non-linear behavior. This makes the computation of LGN activity both more demanding and harder to address analytically.

To my knowledge, Willshaw and von der Malsburg were the first self-organizing map modelers to assume a center-surround (short-range excitatory, long-range inhibitory) neighborhood function in the tectum. This neighborhood function, in conjunction with the threshold model of activity, tends to generate a "bubble" of activity around the area of the map with the strongest synaptic connections from the retinal region within the activity cluster. The formation of only one bubble per input pattern arises because cells that reach a firing threshold first (i.e. because of strong synaptic inputs) inhibit the activation of distant cells in the map through their long-range inhibitory connections. This process plays a central role in the robust development of global order, insuring that only one part of the tectum "represents" a given area of retina and that tectal neighbors (tectal neurons in the bubble) connect to co-active (i.e.neighboring) areas of the retina.

Because it supports robust convergence, long-range inhibition has been a standard assumption (either explicit, as in neighborhood functions, or implicit, as in activity bubbles) in most self-organizing models (see Swindale, 1996 for review). One of the major goals of the current model was to explore models that do not assume the presence of long-range inhibition during map refinement, because this inhibition is largely absent during early development in the retinogeniculate system (see Sec. 3.4.1).

### *7.1.2 Kohonen*

Because it is computationally efficient, converges robustly, and captures some of the essential dynamics of self-organizing processes, Kohonen's self-organizing feature map (SOFM) algorithm has been widely used in both neural models and engineering applications (Kohonen, 1995). It is the most extensively used and analyzed of the self-organizing map algorithms.

The SOFM approach (Kohonen, 1993) can be understood as an abstraction of Willshaw and von der Mahlburg's model (1976). The SOFM algorithm explicitly calculates the output unit that receives the strongest connections from the given input, generates an activity bubble surrounding that output unit (as an approximation to the bubble that arises from the dynamics of lateral feedback), and then adjusts the synaptic weights so as to increase the strength of connections from the neurons in the activity bubble to the active inputs. The SOFM algorithm assumes a non-local winner-take-all mechanism, usually taken to arise from center-surround lateral interactions in the target tissue. This approach short-cuts the calculation of the activity bubble and low-level Hebbian learning, and concentrates on the effective dynamics in the Willshaw and von der Malsburg model.

Kohonen's model is more general than Willshaw and von der Malsburg's model in the sense that it is extensible to applications where a high-dimensional input space is collapsed onto a lower dimensional representational space — e.g. the representation of orientation selectivity, ocular dominance, and 2-d retinotopy on the 2-d sheet of primary visual cortex. This feature links the model closely to vector quantization and related engineering and statistical applications (see Haykin, 1994; Kohonen, 1995 for review).

The convergence of the algorithm to a globally ordered map is greatly increased by decreasing the learning rate over development (often by several orders of magnitude) and by using a neighborhood function that starts off very large (often covering the entire target map) and shrinks over development. Sharpness in the eventual map is ensured by shrinking the neighborhood function (i.e. the activity bubble) down to only one excitatory neuron in the final stages. The shrinking neighborhood function can result from the growth of lateral inhibition, which sharpens the activity bubble. Because lateral inhibition in the LGN increases over the time-course of afferent segregation (Sec. 3.4.1), this type of inhibition driven sharpening could complement the mechanisms that support the sharpening of lateral interactions in the present model. Sharpening in the current model is supported by physiological changes in the membrane response of LGN relay cells (see Fig. 4.6).

## 7.2 LGN Specific Models

Three models of development in the retinogeniculate system have been suggested (Eglen, 1997; Keesing et al., 1992; Lee and Wong, 1996). These models are all concerned with details specific to the retinogeniculate system, and each highlights different aspects of the system.

### *7.2.1 Keesing and Stork*

Keesing et. al. (1992) proposed the first specific model of retinogeniculate refinement. Their model was developed soon after the first reports of the retinal waves were published, and their model uses simulated waves as inputs. In their model, two retina strips (each 1x10 neurons) project to a sheet of LGN tissue several neuron tall/deep (8x50). The model assumes no lateral interactions

in the LGN — the authors note the lack of lateral inhibition during biological development. Self-organization is supported by axonal sprouting and retraction, similar to the current model.

Given an initial bias toward retinotopy and a particular layer configuration (i.e. the contralateral eye growing in first and penetrating deeper), the model develops both retinotopy and eye-specific layer segregation. With no initial biases, the model forms alternating domains, akin to ocular dominance columns, rather than layers. Their model is an elegant demonstration of the power of sprouting and retraction processes in supporting the development of topology.

Though the model is not analyzed in detail, it provides an important demonstration of the ability of a model with minimal assumptions and wave-like inputs to form two of the most striking features of the retinogeniculate projection: retinotopy and eye-specific layer segregation. The model does not deal with on/off sublayer formation, nor does it incorporate intra-dLGN interactions.

### *7.2.2 Eglen*

Eglen's work (1997) provides some in depth exploration of Keesing et. al.'s model framework as well as an elegant application of the SOFM algorithm to the retinogeniculate projection.

*Model I.* Eglen's first model is the same as that of Keesing and Stork's model. He models the projection of two 1-d strips of RGCs to 8 1-d LGN strips stacked on top of one another to create layers/columns. The model projects two 1-dimensional retinas to a 2-dimensional LGN. The shape of the retinal arbor along the depth of the LGN is important in the model because it is interpreted as a columnar structure. For example, Eglen finds that axonal sprouting can support the development of repeated and aligned retinotopic maps in each layer of the LGN. This arrangement nicely models the biological observation that RGC axons arborize in columns in the LGN, or similarly that a recording track perpendicular to the layers will tend to remain centered on the same point in the visual field.

As in Keesing and Stork's model, Eglen assumes that there are no direct interactions between dLGN cells during development, rather the clustering of like inputs is accomplished solely through sprouting and retraction processes. The extent of sprouting decreases over the course of development, eventually disappearing $\sim 2/3$ of the way through each simulation.

**Inputs.** One notable difference between Eglen's model and the current model is that Eglen generates negative within-eye correlations in his model because his simulated waves are narrow relative to the strip size (1:50), and have wave-fronts that are oriented perpendicular to the retinal strips. This arrangement ensures that RGCs at opposite ends of a retinal strip are never co-active, and as a result are negatively correlated (see Secs. 4.3.2 and 5.5). This feature of the retinal activity should make it unrealistically hard for Eglen's model to develop layer segregation. The successful segregation of layers in spite of these negative correlations testifies to the strength of the mechanisms that support layer segregation in the model. This situation is similar to the simulations of multiple retinal domains in my model (see Sec. 5.5).

**Initialization.** Eglen's model assumes a diffuse initial projection, where mean receptive field size is very large early in development and then shrinks down to a localized spot. There are some versions that use more restricted retinal arbors (namely to impose an ingrowth bias toward retinotopy and layer segregation), but the neurons are almost fully connected at initialization. This assumption is not justified in the retinogeniculate projection (see Sec. 3.3.3), but is probably not necessary for convergence in Eglen's model because the sprouting mechanisms that support constant arbor width in the current model are present in Eglen's model as well.

**Weight Update and Normalization** The model uses pre- and post-synaptic normalization and a covariance updating rule where the weight update is proportional to the activity levels minus a constant. In the terminology of the current model, Eglen's synaptic weight update equation is:

$$\delta w_{[xy]} \propto (l_{[x]} - \alpha)\,(r_{[y]} - \beta) \tag{7.1}$$

where the synaptic weight change ($\delta w_{[xy]}$) is the product of the geniculate activity minus a constant ($l_{[x]} - \alpha$) and the retinal activity minus a constant ($r_{[y]} - \beta$). In contrast, the Hebbian learning rule used in the current model uses the simple product of the pre- and post-synaptic normalization (Eq. 4.4) to update the synaptic weights. Eglen explores the effects of different types of normalization in depth. He investigates the different behavior of the subtractive and multiplicative normalization rules, finding that the multiplicative normalization rule supports both retinotopy and eye-segregation most effectively. He also finds that normalizing only over the pre-synaptic cells results in more robust development of retinotopy and layer segregation, a result that is not true in the current model (see Sec. 5.10.2).

In the case of mostly silent inputs, Eglen finds that his simulation tends towards a homogeneous projection where there is little sharpening by input type or receptive field. This tendency arises because, during silent periods, the covariance rule increases all the weights non-specifically. Eglen counteracts this tendency by enforcing weight change only when the pre-synaptic activities are above a certain value ("active covariance"). Using this weight update rule the between eye correlations are *effectively* negative, because the times when both retinas are silent are ignored (see Appendix D).

Eglen also explores models of monocular enucleation in which the treated eye's projection shrinks and the non-enucleated eye's projection expands. Using only pre-synaptic normalization, he is able to achieve compression of the spared eye's projection to a small number of dLGN relay cells, with corresponding expansion of the spared eye's projection. With a relaxed pre-synaptic normalization rule (similar to the capped normalization rule used in my model, Sec. 4.3.7), Eglen's model is able to simulate complete dominance of all dLGN relay cells by the spared eye's afferents.

**On/Off Segregation** Eglen explores the ability of his model to simulate the segregation of on/off sublayers as well. He simulates the differentiated on/off activity during the third postnatal week by

interpreting his retinal strips as alternating on and off-center RGCs (an assumption that is closely analogous to the use of two patches in my model). He demonstrates that the model is capable of supporting the segregation of inputs to individual dLGN relay cells but not to entire sublayers. The failure of on/off sublayer segregation is especially striking because the form of simulated on/off wave activity provides an unrealistic bias toward on/off afferent segregation. In Eglen's model the off-cell firing rates are assumed to increase through the introduction of off-RGC specific waves instead of off-RGC specific noise, as in the current model. Consequently, the off-off RGC correlations in the model stay unrealistically high. This high correlation should provide strong support to the clustering of off-RGC afferents and sublayer formation relative to the noisy off-RGC inputs used in the current model.

In order for his model to support the emergence of on/off specific cells, the on/off inputs must be negatively correlated (which, as Eglen notes, is not true of the observed waves, Wong and Oakley, 1996). With less negatively correlated on/off activity (in fact, the on/off cells are positively correlated, Fig. 4.5) the model does not support the emergence of on/off specific cells. Noting that realistic on/off activity does support the emergence of on/off specific cells in a previous model (see Sec. 7.2.3), Eglen suggests that this failure might be specifically caused by the use of the covariance rule, but he does not provide a more detailed analysis.

Eglen identifies several potential reasons that on/off segregation is not robustly supported in his model. A likely influence is that he does not bias the on or off-cells to project to a certain sublayer. Biases are found to be necessary to layer segregation in his model and are found to help support on/off sublayer segregation in my model (see Sec. 5.8). Eglen also notes that there is no significant development of on/off specificity in individual cells until sprouting has ceased ($\sim$D55). This fact makes sublayer segregation difficult if not impossible to achieve in his model, because there is no means by which post-synaptic cells of like type (i.e. on or off-center responsive) would tend to cluster in layers. Eglen suggests that sprouting may undermine dLGN cell specificity because it tends to reinforce the pre-synaptic cell's synaptic weights to post-synaptic cells irrespective of their on or off-center responsivity. Another possible factor is that Eglen uses unscaled weight update and multiplicative normalization. Under these conditions his model projection will not tend not to sharpen (Miller and MacKay, 1994). This tendency might help explain why the post-synaptic cells in his model are only able to sharpen their on/off specificity after the blurring influence of sprouting is removed. However, both non-specific sprouting and the tendency toward diffuse projections also apply to eye-specific segregation which is robustly supported by the model (even without the use of biases). So there are probably other influences that haven't been considered.

**Relevance to Current Model.** Although there are some unresolved discrepancies between the two models, Eglen's simulation results can provide some insight into the important aspects of the current model. Eglen's model, like the original Keesing and Stork model, highlights the fact that sprouting and retraction processes support the development of retinotopy and eye-specific layer segregation independent of other mechanisms. These results add weight to the conclusion that the intra-dLGN spread of excitation in my model is at least partially redundant with sprouting (Sec. 5.6.2).

There are several discrepancies between the behavior of the two models that help to highlight the different mechanisms and assumptions in the two models. Whereas Eglen finds that post-synaptic normalization is redundant with pre-synaptic normalization, in the current model both pre- and post-synaptic normalization are necessary (Sec. 5.10). This difference between the behavior of the two models is due to the difference in their learning rules, and in their inputs. The covariance rule can independently provide some post-synaptic normalization, especially in the presence of mexican hat spatial correlations among the inputs, like those present in Eglen's model (Fig. 5.8, Eglen, 1997).

The post-synaptic normalization tendency of the covariance rule even with uncorrelated inputs is illustrated in Figs. 7.1 and 7.2. Assuming that the projections start off with a maximally concentrated "pathological" projection, these examples explore the ability of the projection to develop toward a more distributed/normalized state. Fig. 7.1 demonstrates that a projection in which the synaptic weights are concentrated in a single pre-synaptic cell will remain concentrated in that cell in the absence of pre-synaptic normalization. On the other hand, Fig. 7.2 shows that a projection in which all the synaptic weights are concentrated in a single post-synaptic cell will tend to spread synaptic weight to the other post-synaptic cells, even in the absence of post-synaptic normalization. The covariance rule implicitly provides a normalization mechanism that is partially redundant with post-synaptic normalization (see Miller, 1996 for discussion of the normalization effects of the covariance rule). This feature of the covariance rule explains why Eglen was able to eliminate post-synaptic normalization in his model with no ill-effects. In contrast, the product rule used in the current model has no mechanism for decreasing a synaptic weight aside from normalization induced synaptic competition. As a result the product rule is not robust to the absence of either pre or post-synaptic normalization (see Sec. 5.10 and Sec. 4.3.7).

Eglen also finds that subtractive normalization does not support the development of smooth retinotopy, but it can in the current model (see Sec. 6.2.1 and Fig. 6.1). Mechanisms in the current model like the asymmetric lateral weights and sprouting and perhaps the details of the waves could contribute to these discrepancies in our results.

The failure of Eglen's model to support on/off cell and sublayer segregation also could have multiple causes. Given that asymmetries in sprouting and lateral intra-dLGN influence are crucial to sublayer segregation in the current model (see Sec. 5.9 and Fig. 5.13), the lack of similar asymmetries in Eglen's model can help explain the failure of his model to support on/off sublayer segregation. As suggested by Eglen, the lack of a bias toward sublayer segregation can also help explain the model's failure to form sublayers. Results from the current model suggest that biases do help support sublayer segregation (Sec. 5.8).

The ability of Eglen's model to form eye-specific layers without asymmetric sprouting contrasts with the failure of my model to develop eye-specific layers under the same conditions (see Sec. 5.9 and Fig. 5.13). The source of this discrepancy is not clear to me. It is possible that different instantiations of sprouting and layers in the two models is responsible for the discrepancy in the results. Sprouting (the "growth rule") is instantiated in Eglen's model by increasing a given weight

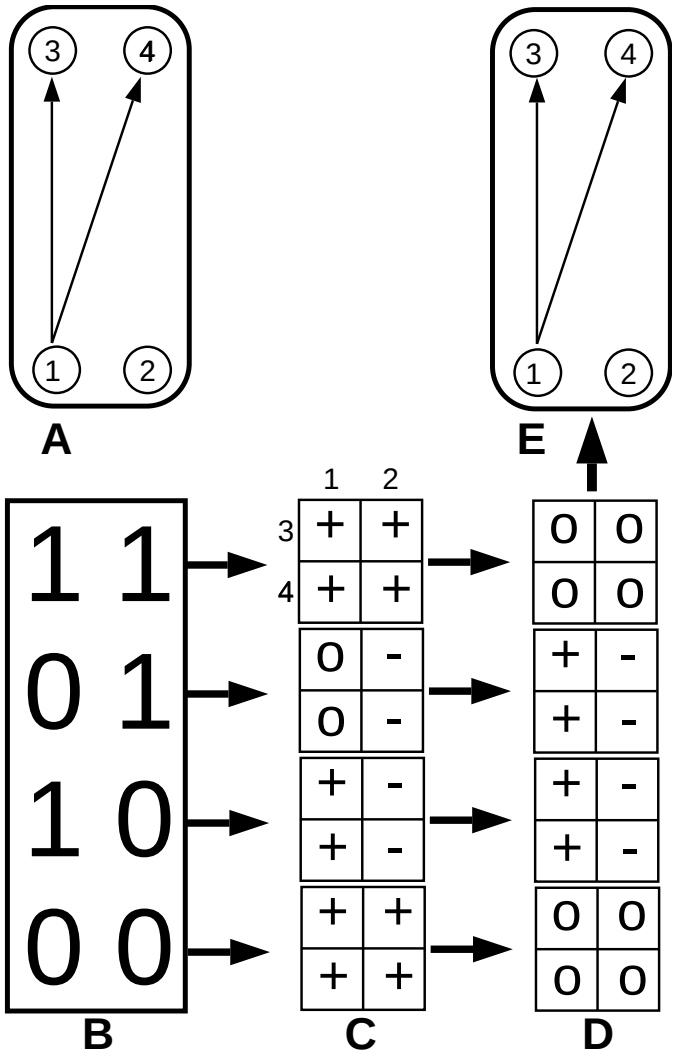


Figure 7.1: In conjunction with the covariance rule, post-synaptic normalization cannot ensure that a projection will not concentrate all its strength in one input cell. The 2x2 grids (C-D) are weight matrices with an analogous format used for the results in this paper (see Fig. 5.2A). The entries describe the change in that synaptic weight, where "+" denotes increase, "−" denotes decrease, and "0" denotes no change. The calculation of these changes assumes that a post-synaptic cell will be active if it receives a connection from an active pre-synaptic cell, and inactive otherwise. (A) An initial projection with the synaptic weights concentrated in one input cell, (B) The four possible (binary) input activities where 1 means the input cell is active and 0 means it is inactive, (C) The weight change caused by each input pattern under the covariance rule, (D) The effective weight change after post-synaptic normalization is enforced, (E) The updated projection in which all of the synaptic weight remains concentrated in one input cell. Note that (E) assumes that the initial weights are pinned at their extreme limits, so that a synaptic weight change in the direction of the limit has no effect, while a synaptic weight change away from the limit does have an effect.

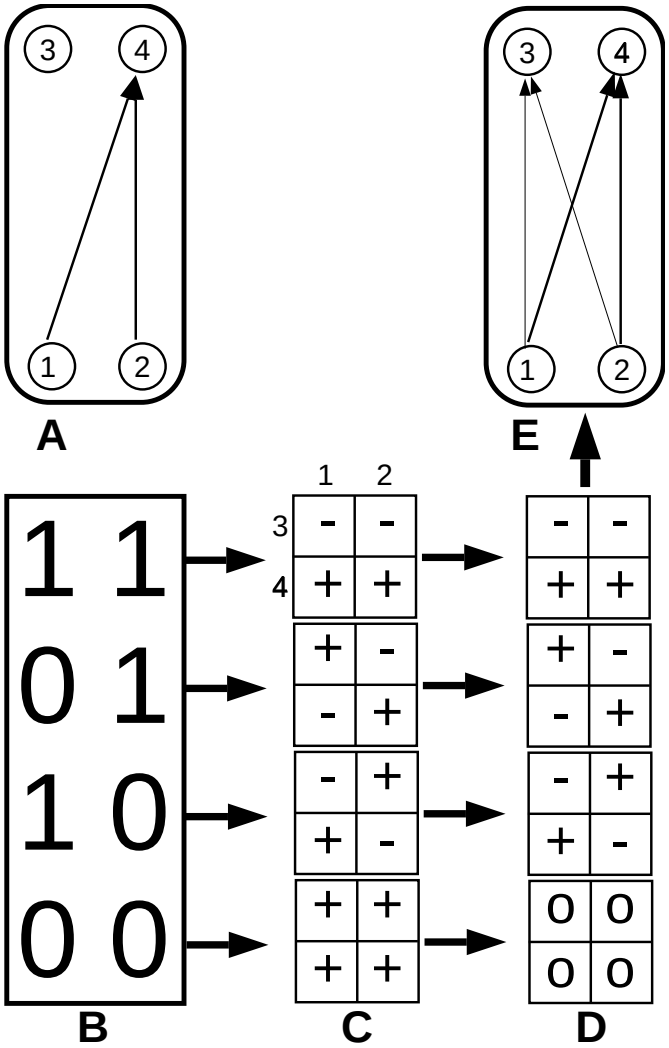


Figure 7.2: In conjunction with the covariance rule, pre-synaptic normalization can ensure that a projection will not concentrate all its strength in one output cell. The 2x2 grids (C-D) are weight matrices with an analogous format used for the results in this paper (see Fig. 5.2A). The entries describe the change in that synaptic weight, where "+" denotes increase, "−" denotes decrease, and "0" denotes no change. The calculation of these changes assumes that a post-synaptic cell will be active if it receives a connection from an active pre-synaptic cell, and inactive otherwise. (A) An initial projection with the synaptic weights concentrated in one output cell, (B) The four possible (binary) input activities where 1 means the input cell is active and 0 means it is inactive, (C) The weight change caused by each input pattern under the covariance rule, (D) The effective weight change after pre-synaptic normalization is enforced, (E) The updated projection in the initially isolated post-synaptic cell receives some inputs. Note that (E) assumes that the initial weights are pinned at their extreme limits, so that a synaptic weight change in the direction of the limit has no effect, while a synaptic weight change away from the limit does have an effect. Under this assumption, the two cases where one input is active and the other is not will tend to increase the zero weights, but will not increase the pre-existing weights.

by the scaled and summed weight of 4-15 of its nearest neighbors. This process is executed probabilistically on 1% of the iterations. The probabilistic nature of the sprouting process as well as its limited extent may discourage the replication of RGC axonal arbors in adjacent sublayers as seen in my model (Fig. 5.13).

*Model II: SOFM.* Eglen developed a second model that is an elegant demonstration of the possible importance of the shape of the dLGN on layer formation. This model uses a SOFM model that maps ocularity and retinotopy onto a 3 dimensional block of dLGN relay cells. The approach is to model the response of dLGN relay cells to abstract input features (i.e. ocularity and RGC position) rather than to actual inputs (i.e. vectors of activations across populations of RGCs). Eglen finds that the organization of the projection to the block is determined by the dimensions of the block in conjunction with the variance of the ocularity and retinotopy features. The lowest variance feature (i.e. ocularity) will tend to organize along the shortest axis, while high variance features (i.e. retinotopy) will tend to organized along the longest axis. In some regions of the parameter space, this model produces results that closely match features of the observed structures.

This approach is clever and parsimonious, as it can form layers and retinotopic maps with no initial bias, and uses the shape of the LGN itself to help shape the retinogeniculate projection. The major disadvantage of this approach is that the model is harder to map onto the biological system than less abstract models (like his first model and the current model).

### *7.2.3 Lee and Wong*

Lee and Wong (1996) have developed a simple model that demonstrates that, under some conditions, the activity present during on/off sublayer segregation can support the development of on or off- selectivity in a single post-synaptic cell. The model is initialized with a single post-synaptic cell receiving inputs from a group of on and off-center RGCs. The RGC activity patterns are modeled on the in vitro recordings (and the computed spatial correlations) of the spontaneous activity in the retina during on/off segregation. The synapses are then updated with a variant of the covariance rule, in which the raw post-synaptic activity is used ($\delta w_y \propto l(r_y - \theta_y)$). They find that the post-synaptic cell can develop either on or off selectivity, depending on the specific parameter ($\theta_y$) used in the weight update equation (Miller, 1997a). This model was an important demonstration that the correlations between RGCs during on/off sublayer segregation are capable of supporting specificity of connections to individual post-synaptic cells.

*Comparison of Lee and Wong's Rule to a Covariance Rule.* This result is an interesting contrast to Eglen's finding that, in the realistic case of positive correlations between on and off-center cells, his covariance rule did not robustly support specificity of connectivity to individual post-synaptic neurons. The differences between the Lee and Wong's rule and the covariance rule are probably the cause of these discrepancies.

The covariance rule and Lee and Wong's rule are most different in the way they treat periods when the post-synaptic cell is silent. In Lee and Wong's rule there is no synaptic change in this case, in the covariance rule there is often strong positive synaptic change (i.e. when the pre-synaptic cell is silent as well). Because the RGCs are silent for the vast majority of the time during this period, refinement under the covariance rule tends to be swamped by non-specific weight growth — preventing specificity of connections. In contrast, Lee and Wong's rule essentially ignore silent periods and thus refine under the influence of inputs that are effectively anti-correlated (see Appendix D). The importance of discounting the silent periods is highlighted by Eglen's finding that an "active covariance" rule (which ignores all instances where the input activity is below a given threshold) improves his model's convergence to specificity in dLGN cell connectivity.

# Chapter 8

# Discussion

## 8.1 Biological Value of the Model

The retinogeniculate projection is an ideal system for studying activity-dependent neural development. As the primary visual pathway to the cortex, it is well-studied, and it develops complex stereotypical structures under the influence of striking spontaneous activity. Because our knowledge of the system is highly detailed, a model can be highly constrained and can make steps toward exploring the hierarchy of candidate mechanisms for supporting activity-dependent refinement.

This is the first model of the retinogeniculate system to capture the activity-dependent development of eye-specific layers, on/off sublayers, and retinotopic refinement in a unified framework. The model is both biologically interpretable and analytically tractable. It uses realistic retinal activity and biologically motivated and detailed assumptions about the mechanisms that might be responsible. In addition, the model details are closely matched to anatomical, physiological, and biophysical observations of the retinogeniculate system and its development. This fidelity to biological detail, in combination with the phrasing of the model in terms of concrete spatial and temporal scales, make the model biologically interpretable and predictive. The analytic work presented in this thesis adds power to the model as an exploratory tool. For example, the dLGN update can be understood in terms of image processing and membrane biophysics, and using eigenmode analysis the developmental dynamics can be partially explored without resorting to computationally intensive simulations. Hopefully, the model will provide a framework for developing and exploring intuitions about activity-dependent refinement in this system, as well as making predictions about future experiments.

Of course, activity-dependent refinement is not limited to the retinogeniculate projection. However, because it is relatively well constrained, the retinogeniculate projection can provide a means to understand the interactions between activity-dependent mechanisms, and in combination with detailed models, can help establish the appropriate scale of some of the generic mechanisms supporting activity-dependent refinement.

## 8.2 Activity-Dependent Developmental Strategies

Retinogeniculate development can also provide some insight into general developmental strategies, helping to broaden our understanding of learning and its role in development. Stated in psychological terms, most learning presumably happens during experience with the world, and is learning "about" the world through the sensory systems. The activity-dependent change in early retinogeniculate development is also through sensory channels, and likely takes advantage of some of the same biophysical mechanisms of synaptic change used in experiential learning (e.g. LTP).

The intriguing aspect of this system is that the "sensory" input is generated by the organism itself, rather than the environment. Because the amount of genetically coded information needed to generate the retinal waves is probably less than that needed to specify the exact synaptic wiring of the retinogeniculate projection, the organism can more efficiently generate precise neural circuits with relatively economical spontaneous activity and some stock learning mechanisms. By generating activity, the organism only has to explicitly specify coarse features of the projection.

This developmental strategy appears to be quite general. The generation of spontaneous retinal activity in early development is common to vertebrates (Sernagor and Grzywacz, 1996). In addition, the findings of spatiotemporally patterned spontaneous activity in the auditory system (Lippe, 1994) and the somatosensory system (O'Donovan, 1994) suggest that this strategy may be used in different sensory systems (Cook and Becker, 1990).

### *8.2.1 Constrained Experience*

Perhaps our understanding of learning can also help broaden our understanding of activity-dependent development. Developmental theorists have postulated that the sequence of experience is crucial to the successful learning of complex domains. Surprisingly, limitations can play a crucial role in providing appropriate sequences, meaning that a period of highly limited immature functioning in early life can be necessary to achieving competence in an adult. The importance of limitations has been called the "Less is More" (Newport, 1990) and the "Starting Small" (Elman, 1993) principle. The validity of this principle has been demonstrated computationally in the context of language learning (Newport, 1990; Goldowsky and Newport, 1993; Elman, 1993).

The trick is that, in some situations, limitations can simplify a domain while still preserving some useful elements of the full complex domain — essentially performing task decomposition. This decomposition can make a complex task tractable because each subtask is relatively easy, and once mastered, contributes to learning and mastering the complete task. This principle also applies to a sequence of gradually lessening limitations. Like a good teacher, a sequence of gradually lessening limitations can tailor the complexity of the task to the competence of the learner. For example, Turkewitz and Kenny (1982) suggest that the programmatic and staggered onset of sensory function (cutaneous, vestibular, olfactory, auditory, then visual) could serve to support successful sensory integration. Their idea is that the sensations give a gradually increasing level of acuity, such

that experience is coarse grained at first and finely detailed only later. In terms of the cat's visual development, they suggest that delayed and gradual eye-opening, the poor acuity, and the lack of the ability to focus at different distances seen in the newborn could serve to highlight outlines, shapes and canonical object sizes. As vision develops, more detailed information about form and internal details are slowly integrated into the simple initial experiential framework.

An interesting thought experiment is to push this sequence even earlier in development when, presumably, even more simplified experience is optimal for learning. The newborn kitten's photoreceptors are functional by D70, but the kitten doesn't open its eyes until ∼D75. Given that light penetrates the lids of the eyes (see Sec. 3.3.7), some pattern information reaches the retina during this period. Of course, the pattern might be no more than one side of the retina receiving more light than the other side (i.e. when a body partially occludes a light source).[1] Even a pattern this simple carries information about future visual experience because cells in the same part of the two retinas will tend to be co-active when exposed to such patterns, as they will in mature visual experience. In addition, motion in the world will cause shifting shadows to be cast on both retinas in manner parallel with visual experience after eye opening. Perhaps this highly filtered early experience builds a set of circuits/competencies that are useful for parsing and using more detailed visual experience later in life.

### *8.2.2 Self-Induced Experience*

Following the "Starting Small" principle, activity that is even more simple than filtered experience could play a useful role. In some cases such simplified activity can be produced spontaneously. For example, even more basic than the "lesson" of co-activity between the eyes, is the fact that nearby cells in a given retina will tend to be co-active during visual experience. While experience with the world does not produce activity this simple (in isolation from co-activity in the other eye), the spontaneous wave activity discussed in this thesis does expose the organism to this spatiotemporally correlated characteristic of visual experience. The spontaneous activity can be seen as extending the programmatic sequence of experience back into pre-experiential development. In this view, intrinsically generated activity captures some extremely basic features of visual experience and essentially primes the visual system for the experiences to come.

Viewed from this perspective, the waves may have influence well beyond the segregation and topographic organization of retinal afferents. Heightened attention to motion has been documented in human infants (Nelson and Horowitz, 1987; Burnham, 1987; Haith, 1966) and probably serves to focus the newborn on animate objects and social interactions. Such a bias towards movement could be supported by pre-visual experience with "moving" wave stimuli. To the extent that such an attentional bias toward moving objects was supported by wave activity, it would presumably be tuned to the size and speed of the waves. If this reasoning is correct, then a specific and predictable

[1] The reader can simply pass their hands past their closed eyes while facing a bright scene or a light source to get a rough sense of visual input through the eyelids.

attentional bias toward certain types of motion would be present in ferrets and cats at the time of eye-opening. To my knowledge, an attentional bias toward motion has not been investigated in neonatal ferrets or cats.

## 8.3 Unaddressed Issues

The domain of this model is highly restricted, consequently it ignores some potential influences on retinogeniculate development. The following section discusses some important features of biological development that are not explicitly modeled, and which should be taken into account when interpreting the model. This discussion can also be read as an outline of some useful future directions in which to expand the model.

### *8.3.1 Feedback Circuits*

The current model does not incorporate any feedback connections to the dLGN. Only 10% of the synapses in the dLGN are retinal in origin, suggesting that activity in the dLGN is highly modulated by feedback connections and extra-retinal inputs (see Sec. 2.3.4). For example, cortical inputs and RNT/PGN inputs form feed-back loops with dLGN relay cells that strongly shape mature geniculate responsivity. There are reasons to believe that these feedback loops have a greatly attenuated influence during early development (Sec. 3.4.1), but they probably do play some role. Further, the feedback loops themselves might undergo significant activity-dependent development during this period. The model also neglects other sub-cortical areas that send afferents to and receive afferents from the dLGN. These interactions may be importantly involved in early geniculate functioning.

### *8.3.2 Activity-Dependent Intra-dLGN Development*

It is possible, and even likely that intra-dLGN connectivity is influenced by activity, but this dependency is outside the scope of this model. Other models have investigated the simultaneous development of feed-forward connectivity and lateral influence (Sirosh and Miikkulainen, 1993; Sirosh and Miikkulainen, 1994; Sirosh and Miikkulainen, 1996), though not explicitly in the dLGN. These models have found that co-development of the lateral connections can bring about stereotypical lateral interaction functions, and can simulate some details of lateral connectivity in the cortex.

### *8.3.3 Hebbian Synaptic Change and LTP*

There is little reference to the linkages between Hebbian learning and LTP in this thesis because there is a mismatch between the investigations of LTP that are possible at this point and the type of data that could directly inform model details. Presumably, LTP in the developing dLGN is better described as either scaled or unscaled by the previous weight, and knowing which version of

weight update was more accurate would add a valuable constraint to the model. Unfortunately, the data on LTP is not well suited to answering this type of question (Dan Madison, personal communication). In order to definitively differentiate between scaled and unscaled Hebbian learning, data is needed on temporal scales of weeks and spatial scales of hundreds or thousands of neurons. The experiments to date that investigate LTP concentrate on temporal scales of minutes to days and spatial scales of a few neurons. At these scales, LTP is dominated by the saturation of synaptic potentiation (presumably because there is neither the time nor the resources to activate "dormant" synapses and build new synapses). Hopefully, as the data on LTP becomes more clear and more relevant to the issues in this model, empirical observations will help to directly constrain our assumptions about Hebbian synaptic change.

### *8.3.4 Activity-Independent Factors*

The model assumes only the most simple types of trophic and mechanical influences. These activity-independent factors might be quite influential in the initial setup and refinement of the retinogeniculate projection, but they are down-played in the current model. While the current model does demonstrate how powerful activity-dependent mechanisms can be, it does not address how the two types of factors might dynamically interact over the course of development.

### *8.3.5 Visual Experience*

The development of the retinogeniculate projection continues well after the early post-natal period that is the focus of the model. For purposes of simplicity, this model assumes that layer alignment and dendritic and axonal development have developed to a stable state before eye opening. However, these assumptions are not entirely accurate. For example, X- and Y-RGCs undergo significant post-visual refinement (Sec. 3.3.4), and the alignment of the topographic maps between layers and eyes (Sec. 3.3.9) is probably at least partially dependent on patterns of visual activity. The model could be extended to post-visual development by adding a more extended and detailed time-course of the retinal activity and perhaps more detailed and variegated RGC types and activity patterns.

### *8.3.6 Scale Issues*

Although it was demonstrated that the mechanisms in the model can potentially deal with multiple retinal domains (Sec. 5.5), and the model was simulated at two scales, the scale of the model is $\sim$4 orders of magnitude smaller than the actual retinogeniculate projection. This reduced scale prevents the model from capturing the dynamics of a large retinal surface with many transient subdomains. A larger scale model could use more of the biological observations available to further constrain the characteristics of the retinal activity, and perhaps the other model parameters as well.

## 8.4 Proposed experiments

There are some direct predictions of the model that are, to my knowledge, untested. Other predictions could be made from the model behavior, but these are the predictions that are most robust across parameter domains.

### *8.4.1 Point Spread*

The model makes some strong predictions about asymmetry and the developmental time-course of point-spread in the dLGN. In order for layer segregation to develop properly in the model, the inter-layer physiological interactions must be much weaker than the intra-layer interactions in the dLGN (Sec. 5.9). The model predicts that the spread of excitation should be wider within a layer than between layers. To test this prediction one could excite a dLGN cell with an intracellular electrode while simultaneously recording from other dLGN cells at various distances within the same layer and across layers and sublayers. Failure to find this strongly asymmetric spread of excitation would suggest that additional mechanisms (e.g. a stronger bias in initial ingrowth and/or trophic factors) contribute more substantially to layer segregation than this model assumes. Note that these predictions are not necessarily anatomical, because only the effective physiology (activity spread) supported by the circuitry is important.

### *8.4.2 Relay Cell Physiology*

The biophysical interpretation of the model makes some interesting predictions about immature dLGN physiology. It predicts that inhibitory and excitatory conductances in the developing dLGN will be linked in a push-pull relationship. That is, as the excitatory conductances increase, the inhibitory conductances should decrease (Sec. 4.3.4). The model also predicts a 50% increase in dLGN whole cell capacitance over the course of development. These features of dLGN cell physiology could be measured via intra-cellular recording techniques (Ramoa and McCormick, 1994a; Ramoa and McCormick, 1994b) during the first few post-natal weeks in order to determine the dynamics of conductance change and the time-course of whole cell capacitance.

### *8.4.3 Differentiation Between Off-Cells and On-Cells*

The model also suggests that the off-center RGC axonal arbors will have some characteristics that differentiate them from the on-center RGC axonal arbors around the time of eye-opening. They will tend to be more focally arborized, and, assuming that normalization is somewhat flexible, they will tend to have more synaptic weight (Sec. 5.10.3) than the on-center RGCs. These differences result from the pronounced differences in the on and off-center RGC firing rates during the period of on/off sublayer segregation. These features of axonal arborization could be investigated through anterograde staining and tracing of the on- and off-center RGCs, and perhaps through the comparison

of previously extracted axonal arbors (for example, Mastronarde, 1992; Dalva et al., 1994; Sretavan and Shatz, 1984; Sretavan and Shatz, 1987).

## 8.5 Conclusion

Neural development is a complex, robust and powerful process. The neural circuitry develops with what seems to be a woefully inadequate amount of genetic information available to determine its structure. The retinogeniculate system provides a window into the features of neural development that make this feat possible, and this model helps to characterize the crucial mechanisms. Hopefully, this work will provide a useful framework organizing some of the broad range of investigations into the development of this system and neural development in general.

# Appendix A

# LGN Update and Regularization

This appendix explains the relations between the equation used for dLGN activity update in the current model (Eq. 4.2), and the image processing approach known as regularization (Poggio et al., 1985; Blake and Zisserman, 1987; Black and Rangarajan, 1996). Regularization essentially blurs an image so as to denoise it and interpolate missing input values. A common implementation of regularization, is expressed below (for the 1-d case).

$$l_{[x]}^{(i+1)} = \frac{1}{1+2\lambda}(u_{[x]} + \lambda l_{[x-1]}^{(i)} + \lambda l_{[x+1]}^{(i)}) \tag{A.1}$$

Where $l_{[x]}^{(i+1)}$ is the image value at position $x$ and time $i+1$, $u_{[x]}$ is the original image value (i.e. input) at position $x$, and $\lambda$ is a space constant that determines the amount of blurring performed by the regularization. This equation is iteratively applied until the image reaches a steady state.

The change in image value at point $x$ and time $i$ is,

$$\frac{\delta l_{[x]}}{\delta t} = l_{[x]}^{(i+1)} - l_{[x]}^{(i)} \tag{A.2}$$

Substituting Eq. A.1 for $l_{[x]}^{(i+1)}$ yields a version of the update equation that is independent of the specific time $i$.

$$\frac{\delta l_{[x]}}{\delta t} = \frac{1}{1+2\lambda}(u_{[x]} + \lambda l_{[x-1]} + \lambda l_{[x+1]}) - l_{[x]} \tag{A.3}$$

This equation can be restated in a form that assumes a constant influence of the nearest neighbors $l_{[x-1]}$ and $l_{[x+1]}$ (i.e. they are not scaled by $\lambda$):

$$\frac{\delta l_{[x]}}{\delta t} = \frac{\lambda}{1+2\lambda}(\frac{1}{\lambda}u_{[x]} + l_{[x-1]} + l_{[x+1]}) - l_{[x]} \tag{A.4}$$

In this formulation, a large $\lambda$ (e.g. $\lambda = 20$ versus $\lambda = 1$) tends to decrease the influence of the original image relative to the influence of neighbors, and tends to increase the strength of the positive blurring process relative to the negative stabilization term $l_{[x]}$.

Interpreting the dLGN activations as the image values, $u_{[x]}$ as the synaptically weighted retinal input to the dLGN and adding a further scaling factor of $\sqrt{\lambda}$ to this input (in order to generate the proper scaling of dLGN responsivity over development, Sec. 4.3.4), makes this equation equivalent to the dLGN update equation.

$$\frac{\delta l_{[x]}}{\delta t} = \frac{\lambda}{1+2\lambda}(\frac{1}{\sqrt{\lambda}}\sum_{y} w_{[xy]} r_{[y]} + \sum_{x \in \eta} l_{[x]}) - l_{[x]} \tag{A.5}$$

where $\eta$ includes the nearest neighbors of cell $x$.

Fortunately, there is a natural way to handle boundaries in the framework of regularization. The term $\frac{\lambda}{1+2\lambda}$ can be generalized to $\frac{\lambda}{1+\text{ (\# of neighbors) }\lambda}$. The unequal contribution of inter-layer and sublayer neighbors necessitates the addition of a scale factor as well.

The general form of the dLGN update equation is

$$\frac{\delta l_{[x]}}{\delta t} = \frac{\lambda}{1+(\sum_{x \in \eta} S_x)\lambda}(\frac{1}{\sqrt{\lambda}}\sum_{y} w_{[xy]} r_{[y]} + \sum_{x \in \eta} S_x l_{[x]}) - l_{[x]} \tag{A.6}$$

where $S_x$ is the interaction strength from the neighboring neuron. $S_x = 1$ in the case of intra-layer neighbors and $S_x$ = the inter-layer or inter-sublayer interaction strength in the case of neighbors in different sublayers.

# Appendix B

# LGN Update and the Membrane Equation

This appendix describes the relations between the dLGN update equation and the membrane equation. The goal is to establish an algebraic relationship between the dLGN update and the membrane equation, and to derive expressions for the conductances in the membrane equation in terms of the retinal and geniculate activities and the blurring parameter $\lambda$.

The dLGN update equation (Eq. 4.2) is,

$$\frac{\delta l_{[x]}}{\delta t} = \frac{\lambda}{1+2\lambda}(\frac{1}{\sqrt{\lambda}}u_{[x]} + l_{\eta}) - l_{[x]} \tag{B.1}$$

dropping the scale factor $\frac{1}{\sqrt{\lambda}}$ amplifies the dLGN response, and yields the equation,

$$\frac{\delta l_{[x]}}{\delta t} = \frac{\lambda}{1+2\lambda}(u_{[x]} + l_{\eta}) - l_{[x]} \tag{B.2}$$

| | |
|---|---|
| $l_{[x]}$ | Activity of dLGN cell $x$ |
| $l_\eta$ | $\sum_{x\in\eta} l_{[x]}$, where $\eta$ includes the nearest neighbors |
| $r_{[y]}$ | Activity of RGC $y$ |
| $w_{[xy]}$ | Synaptic weight from RGC $y$ to dLGN cell $x$ |
| $u_{[x]}$ | $\sum_y w_{[xy]} r_{[y]}$, the effective retinal input to dLGN cell $x$ |
| $h$ | Learning rate |
| $C$ | Membrane Capacitance |
| $g_l$ | Leak conductance |
| $g_e$ | Excitatory synaptic conductance |
| $g_i$ | Inhibitory synaptic conductance |
| $E_e$ | Excitatory potential |
| $E_i$ | Inhibitory potential |

Table B.1: The symbol table for the terms used in this appendix.

The membrane equation (Eq. 4.3) is,

$$\begin{aligned} C\frac{\delta l_{[x]}}{\delta t} &= -(g_l l_{[x]} + g_e(l_{[x]} - E_e) + g_i(l_{[x]} - E_i)) \\ &= g_e E_e + g_i E_i - (g_l + g_e + g_i) l_{[x]} \\ \frac{\delta l_{[x]}}{\delta t} &= \frac{g_e E_e + g_i E_i - (g_l + g_e + g_i) l_{[x]}}{C} \end{aligned}$$

By assuming the following equalities:

$$\begin{aligned} C &= \frac{1+2\lambda}{\lambda} \\ g_e &= \frac{u_{[x]} + l_\eta}{E_e} \\ g_i &= \frac{(g_l + g_e - C) l_{[x]}}{E_i - l_{[x]}} \end{aligned}$$

The dLGN update equation (without $\frac{1}{\sqrt{\lambda}}$) and the membrane equation can be equated.

$$\frac{g_e E_e + g_i E_i - (g_l + g_e + g_i) l_{[x]}}{C} = \frac{1}{C}(u_{[x]} + l_\eta) - l_{[x]} \tag{B.3}$$

The equalities above suggests that $g_e$ and $g_i$ are in a push-pull relationship, where increases in $g_e$ are offset by corresponding decreases in $g_i$, and the exact dynamics are influenced by $g_l$ and the blurring factor $\lambda$.

## B.1 Derivation

Eq. B.3 can be restated as:

$$g_e E_e + g_i E_i - (g_l + g_e + g_i) l_{[x]} = (u_{[x]} + l_\eta) - C l_{[x]} \tag{B.4}$$

Collecting terms from

$$g_e E_e + g_i E_i = u_{[x]} + l_\eta + (g_l + g_e + g_i - C) l_{[x]} \tag{B.5}$$

By assuming that,

$$g_i E_i = (g_l + g_e + g_i - C) l_{[x]} \tag{B.6}$$

Eq. B.5 can be simplified to:

$$\begin{aligned} g_e E_e &= u_{[x]} + l_\eta \\ g_e &= \frac{u_{[x]} + l_\eta}{E_e} \end{aligned}$$

which yields an expression for the excitatory conductance that relies only on the input activities and the activities of the neighboring dLGN cells.

Eq. B.6 can be simplified to yield an expression for the inhibitory conductance.

$$\begin{aligned} g_i E_i - g_i l_{[x]} &= (g_l + g_e - C) l_{[x]} \\ g_i (E_i - l_{[x]}) &= (g_l + g_e - C) l_{[x]} \\ g_i &= \frac{(g_l + g_e - C) l_{[x]}}{E_i - l_{[x]}} \end{aligned}$$

Assuming $E_i = 0$ the equation simplifies to:

$$
\begin{aligned}
g_i &= \frac{(g_l + g_e - C)l_{[x]}}{-l_{[x]}} \\
g_i &= -(g_l + g_e - C)
\end{aligned}
$$

# Appendix C

# Eigenmode Analysis

## C.1 Translation into Matrix Notation

This section first expresses the weight update in terms of the input activities and the spread of activity in the LGN. In order to do so, it will prove useful to express the model in matrix notation.

The LGN update equation,

$$\frac{\delta l_{[x]}}{\delta t} = \frac{\lambda}{1+2\lambda}(\frac{1}{\sqrt{\lambda}}\sum_{y} w_{[xy]} r_{[y]} + \sum_{x \in \eta} l_{[x]}) - l_{[x]} \tag{C.1}$$

can be expressed more compactly and intuitively in matrix notation. By assuming that the constant scale factors (i.e. those factors that include $\lambda$) are included in $\mathbf{W}$ and $\mathbf{V}$, and likewise the diagonal of $\mathbf{V}$ incorporates the last term (i.e. the self-inhibition/leak term), this equation can be written simply in matrix notation.

$$\frac{\delta \mathbf{l}}{\delta t} = \mathbf{W}\,\mathbf{r} + \mathbf{V}\,\mathbf{l} \tag{C.2}$$

| | |
|---|---|
| $l_{[x]}$ | Activity of LGN cell $x$ |
| $r_{[y]}$ | Activity of RGC $y$ |
| $w_{[xy]}$ | Synaptic weight from RGC $y$ to LGN cell $x$ |
| $h$ | learning rate |
| $\mathbf{l}$ | Vector of dLGN activities |
| $\mathbf{r}$ | Vector of retinal activities |
| $\mathbf{W}$ | Matrix of synaptic weights, where $\mathbf{W}_{[\mathbf{xy}]} = w_{[xy]}$ |
| $\mathbf{V}$ | Matrix of intra-LGN interaction weights, where $\mathbf{V}_{[\mathbf{xx'}]}$ is the interaction strength between LGN cell $x'$ and LGN cell $x$, $\mathbf{V}_{[\mathbf{xx'}]} = 0 \,\text{if}\, x \notin \eta_{x'}$ |
| $\mathbf{C}$ | Expected cross-product of vector of retinal activities $\langle \mathbf{r}\,\mathbf{r'} \rangle$ where $\mathbf{C}_{[\mathbf{yy'}]} = \langle yy' \rangle$, the mean product of the activities of RGC $y'$ and RGC $y$. |
| $\mathbf{\Lambda}$ | A linear operator that incorporates the actions of $\mathbf{V}$, $\mathbf{C}$ and the subtractive normalization rule on $\mathbf{W}$. |

Table C.1: A symbol table for the terms used in this appendix.

The steady state LGN potentials can be found by setting their change equal to 0.

$$\begin{aligned} 0 &= \mathbf{W}\,\mathbf{r} + \mathbf{V}\,\mathbf{l} \\ -\mathbf{V}\,\mathbf{l} &= \mathbf{W}\,\mathbf{r} \\ \mathbf{l} &= (-\mathbf{V^{-1}})\,\mathbf{W}\,\mathbf{r} \end{aligned}$$

The weight update equation from the subtractive version of the model:

$$\delta w_{[xy]} = h\, l_{[x]}\, r_{[y]} \tag{C.3}$$

can similarly be expressed in matrix notation.

$$\delta \mathbf{W} = h\,\mathbf{l}\,\mathbf{r'} \tag{C.4}$$

where $\mathbf{l}\,\mathbf{r'}$ is the cross-product between the geniculate and retinal activity vectors. Substituting the steady state LGN potentials yields,

$$\delta \mathbf{W} = h\,(-\mathbf{V^{-1}})\,\mathbf{W}\,\mathbf{r}\,\mathbf{r'} \tag{C.5}$$

## C.2 Correlation Formalism

The first step toward the eigenmode analysis is to make the assumption that the synaptic weights change very slowly relative to the retinal inputs. Under this assumption we can move from the formulation that explicitly depends on the individual inputs, to a formulation that depends on the expected cross-product of the inputs $\langle \mathbf{r}\,\mathbf{r}' \rangle$ — the matrix $\mathbf{C}$.

Taking Eq. C.5, ignoring the step learning rate (because we are no longer dealing with individual time-steps), and substituting the expected product matrix ($\mathbf{C}$) in for the retinal activities ($\mathbf{r}\,\mathbf{r}'$) yields:

$$\delta \mathbf{W} = (-\mathbf{V^{-1}})\,\mathbf{W}\,\mathbf{C} \tag{C.6}$$

### *C.2.1 Normalized Eigenmodes*

The method for calculating the eigenmodes requires a bit of restructuring of terms in order to include the action of normalization. The approach taken here is stack the columns of $\mathbf{W}$ to construct the vector $\mathbf{w}$. Using this vector construction, the action of $-\mathbf{V^{-1}}$ and $\mathbf{C}$ is equivalent to that of the linear operator $\mathbf{\Lambda}$, such that (cf. Eq. C.6):

$$\mathbf{w}^{i+1} = \mathbf{\Lambda}\mathbf{w}^{i} + \frac{2}{L+R} \tag{C.7}$$

where $\mathbf{\Lambda}$ is constructed by setting:

$$\begin{aligned} \mathbf{\Lambda}^{g}_{[xy,x'y']} &= -\mathbf{V}^{-1}_{[x'y']}\,\mathbf{C}_{[xy]} \\ \mathbf{\Lambda}^{n}_{[xy,x'y']} &= \frac{1}{2}\left(\frac{\delta_{[xy]}}{L} + \frac{\delta_{[x'y']}}{R}\right) \end{aligned}$$

$$\delta_{[xy]} = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{w}^{i+1}$ and $\mathbf{w}^{i}$ refer to the weight vector at time $i+1$ and $i$ respectively, $\mathbf{\Lambda}^{g}$ (the weight update/growth operator) and $\mathbf{\Lambda}^{n}$ (the subtractive normalization operator) are both $LR \times LR$. This construction relates the entries in the $-\mathbf{V^{-1}}$ and $\mathbf{C}$ matrices to the stacked form of $\mathbf{w}$. Likewise, the double indices $(xy, x'y')$ correspond to the stacked position of entries in $\mathbf{\Lambda}$ such that $[xy, x'y']$ is equivalent to $[x + L(y-1), x' + R(y'-1)]$.

The growth and the normalization operators are combined in order to yield a linear operator

that produces an updated and normalized weight vector ($w^{i+1}$) from a given weight vector ($w^i$). First, we define the unnormalized weight change ($\delta\tilde{\mathbf{w}}$) as,

$$\delta\tilde{\mathbf{w}} = \mathbf{\Lambda}^g \, \mathbf{w}^i \tag{C.8}$$

The updated but unnormalized weights are:

$$\begin{aligned} \tilde{\mathbf{w}} &= \mathbf{w}^i + \mathbf{\Lambda}^g \, \mathbf{w}^i \\ &= \left(\mathbf{I} + \mathbf{\Lambda}^g\right) \mathbf{w}^i \end{aligned}$$

where **I** is the identity matrix. The normalized weights are:

$$\begin{aligned} \mathbf{w}^{i+1} &= \tilde{\mathbf{w}} - \mathbf{\Lambda}^n \, \tilde{\mathbf{w}} \\ &= \left(\mathbf{I} - \mathbf{\Lambda}^n\right) \tilde{\mathbf{w}} \\ &= \left(\mathbf{I} - \mathbf{\Lambda}^n\right) \left(\mathbf{I} + \mathbf{\Lambda}^g\right) \mathbf{w}^i \end{aligned}$$

this yields the linear operator of interest:

$$\mathbf{\Lambda} = \left(\mathbf{I} - \mathbf{\Lambda}^n\right) \left(\mathbf{I} + \mathbf{\Lambda}^g\right) \tag{C.9}$$

The eigenvectors of this operator describe the dominant growth modes of **w** under unscaled weight update and subtractive normalization (ignoring the constant factor $\frac{2}{L+R}$ from Eq. C.7). The eigenvalues describe how quickly the eigenmodes will grow relative to each other. For display purposes, the eigenvectors of $\Lambda$ are reshaped into an $L \times R$ matrix so as to be directly interpretable as modes of weight matrix growth. If the normalization was exact, all of the column and row sums in each eigenmode be 0, but the simultaneous pre and post-synaptic normalization used here is only approximate. As a result, some eigenmodes support growth that is relatively unnormalized.

# Appendix D

# The Covariance-Rule and Sparse Inputs

Under some common assumptions about activity-dependent change in synaptic circuits, sparse inputs are effectively more anti-correlated than they appear. To my knowledge, this fact and its role in activity-dependent development has not been discussed previously.

A commonly used formalism for synaptic change is the covariance rule, which incorporates the Hebbian process of strengthening connections between co-active cells, and is robust to changes in the baseline level of pre- and post-synaptic activity. Considering a system with several pre-synaptic cells and a single post-synaptic cell, the covariance rule is,

$$\frac{\delta w_i}{\delta t} \propto (y - \langle y \rangle)\,(x_i - \langle x_i \rangle) \tag{D.1}$$

where the change in the synaptic weight from pre-synaptic neuron $i$ ($\delta w_i$) is proportional to the covariance between the post-synaptic activity ($y$) and the pre-synaptic activity ($x_i$), where angle brackets denote averages.

If $w_i$ changes little in response to each input (i.e. the learning rate is small), then one can average the effect of all presentations of the inputs and phrase the synaptic changes in terms of the covariance between the input activities,

$$\frac{\delta \mathbf{w}}{\delta t} = \mathbf{C}\,\mathbf{w} \tag{D.2}$$

where **w** is the vector of synaptic weights and **C** is the covariance matrix for the inputs. The characteristics of the covariance matrix determine how the synaptic weights grow over development (Miller, 1990; Miller and MacKay, 1994; Miller, 1996; Eglen, 1997).

In terms of a single synapse, the covariance rule can be characterized by 3 situations: when both cells are above their mean firing rates, then the synapse is potentiated (long-term potentiation, LTP); when one is above its mean firing rate and the other is below its mean firing rate then the synapse is weakened (long-term depression, LTD); and when both cells are below their mean firing

| | $x_i$ | $x_j$ | $(x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle)$ | $x'_i$ | $x'_j$ | $(x'_i - \langle x'_i \rangle)(x'_j - \langle x'_j \rangle)$ |
|---|---|---|---|---|---|---|
| | 1 | 1 | $1/4$ | 1 | 1 | $1/9$ |
| | 1 | 0 | $-1/4$ | 1 | 0 | $-2/9$ |
| | 0 | 1 | $-1/4$ | 0 | 1 | $-2/9$ |
| | 0 | 0 | $1/4$ | | | |
| Cov $\langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$ | | | 0 | $-1/9$ | | |
| Corr $(\text{Cov}/(\sigma_{x_i}\,\sigma_{x_j}))$ | | | 0 | $-1/9 \div 2/3 = -1/6$ | | |

Table D.1: Demonstration that eliminating zeros in non-correlated inputs results in negative correlations. "Cov" denotes covariance, "Corr" denotes correlation, and angle brackets denote averages.

rate the synapse is potentiated. Because there is no known biophysical means of potentiating "quiet" synapses (where both pre and post-synaptic cells are silent), this third situation is often explicitly disregarded. Quiet situations can be ignored by assuming that the pre-synaptic cell's activity must exceed a threshold in order for the synaptic weight to change (Willshaw and von der Malsburg, 1976; Eglen, 1997; Linsker, 1986). Likewise, non-linear synaptic change functions can simply go to zero when the pre and post-synaptic input are silent (Bienenstock et al., 1982).

While the omission of silent inputs would seem to be a relatively trivial decision, in fact it has important ramifications for development because it decreases the covariance and correlation between inputs. Consider the numerical case of two inputs given in Table 1. This example illustrates that uncorrelated inputs become anti-correlated when the silent periods are ignored. The more silent periods originally present (i.e. the more "sparse" the inputs), the more pronounced the decrease in correlation.

It is often assumed that negative correlations must be present in **C** in order to achieve competition/segregation between groups of inputs under a covariance based synaptic growth rule (Miller, 1996; Eglen, 1997; Linsker, 1986). The present analysis suggests that in lieu of anti-correlated inputs, the necessary correlational structure can arise from a learning rule that ignores silent periods combined with sparse uncorrelated input activity.

# Bibliography


Ahlsen, G., Lindstrom, S., and Lo, F. (1985). Interaction between inhibitory pathways to principal cells in the lateral geniculate nucleus of the cat. *Experimental Brain Research*, 58(1):134–43.

Archer, S. M., Dubin, M. W., and Stark, L. A. (1982). Abnormal development of kitten retino-geniculate connectivity in the absence of action potentials. *Science*, 217:743–745.

Berardi, N. and Morrone, M. (1984). Development of gamma-aminobutyric acid mediated inhibition of x cells of the cat lateral geniculate nucleus. *Journal of Physiology*, 357:525–37.

Bienenstock, E., Cooper, L., and Munro, P. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2(1):32–48.

Black, M. and Rangarajan, A. (1996). On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19:57–92.

Blake, A. and Zisserman, A. (1987). *Visual Reconstruction.* MIT Press, Cambridge, MA.

Bloomfield, S. A. and Sherman, S. M. (1989). Dendritic current flow in relay cells and interneurons of the cat's lateral geniculate nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 86(10):3911–4.

Burnham, D. K. (1987). The role of movement in object perception by infants. In McKenzie, B. E. and Day, R. H., editors, *Perceptual Development in Early Infancy: Problems and Issues.* Erlbaum, Hillsdale, NJ.

Casagrande, V. A. and Condo, G. J. (1988). Is binocular competition essential for layer formation in the lateral geniculate nucleus? *Brain Behavior and Evolution*, 21:198–208.

Chalupa, L. M., Snieder, C. J., and Kirby, M. A. (1996). Topographic organization in the retinocollicular pathways of the fetal cat demonstrated by retrograde labeling of ganglion cells. *Journal of Comparative Neurology*, 368:295–303.

Chalupa, L. M. and White, C. A. (1990). Prenatal development of visual system structures. In Coleman, J. R., editor, *Development of Sensory Systems in Mammals*, Wiley Series in Neurobiology, chapter 1, pages 3–60. John Wiley and Sons, New York.

Cook, J. E. and Becker, D. L. (1990). Spontaneous activity as a determinant of axonal connections. *European Journal of Neuroscience*, 2:162–169.

Crunelli, V. and Leresche, N. (1991). A role for gabab receptors in excitation and inhibition of thalamocortical cells. *Trends in Neurosciences*, 14(1):16–21.

Cucchiaro, J. and Guillery, R. W. (1984). The development of the retinogeniculate pathways in normal and albino ferrets. *Proceedings of the Royal Society of London*, B223:141–164.

Cucchiaro, J., Uhlrich, D., and Sherman, S. (1991). Electron-microscopic analysis of synaptic input from the perigeniculate nucleus to the a-laminae of the lateral geniculate nucleus in cats. *Journal of Comparative Neurology*, 310(3):316–36.

Dalva, M. B., Ghosh, A., and Shatz, C. J. (1994). Independent control of dendritic and axonal from in the developing lateral geniculate nucleus. *Journal of Neuroscience*, 14(6):3588–3602.

Daniels, J., Pettigrew, J., and Norman, J. (1978). Development of single-neuron responses in kitten's lateral geniculate nucleus. *Journal of Neurophysiology*, 41(6):1373–93.

Daw, N. W. (1995). *Visual Development*. Perspectives in Vision Research. Plenum Press, New York.

Dubin, M. W., Stark, L. A., and Archer, S. M. (1986). A role for action potential activity in the development of neuronal connections in the kitten retinogeniculate pathway. *Journal of Neuroscience*, 6(4):1021–1036.

Eglen, S. (1997). *Modelling the Development of the Retinogeniculate Pathway*. PhD thesis, University of Sussex.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48:71–99.

Emri, Z., Turner, J., and Crunelli, V. (1996). Tonic activation of presynaptic gaba(b) receptors on thalamic sensory afferents. *Neuroscience*, 72(3):689–98.

Erisir, A., Horn, S. V., and Sherman, M. (1997). Relative numbers of cortical and brainstem inputs to the lateral geniculate nucleus. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 94(4):1517–1520.

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

Eysel, U. and Pape, H. (1987). Lateral excitation in the cat lateral geniculate nucleus. *Experimental Brain Research*, 67(2):291–8.

Eysel, U., Pape, H., and Schayck, R. V. (1986). Excitatory and differential disinhibitory actions of acetylcholine in the lateral geniculate nucleus of the cat. *Journal of Physiology*, 370:233–54.

Eysel, U., Pape, H., and Schayck, R. V. (1987). Contributions of inhibitory mechanisms to the shift responses of x and y cells in the cat lateral geniculate nucleus. *Journal of Physiology*, 388:199–212.

Fawcett, J. W. (1993). Refinement of topographic projections in the rodent, avian, amphibian and fish visual systems. In Sharma, S. C. and Fawcett, J. W., editors, *Formation and Regeneration of Nerve Connections*, pages 91–101. Birkhauser, Boston.

Feller, M. B., Wellis, D. P., Stellwagen, D., Werblin, F. S., and Shatz, C. J. (1996). Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. *Science*, 272:1182–1187.

Friedlander, M. J. and Tootle, J. S. (1990). Postnatal anatomical and physiological development of the visual system. In Coleman, J. R., editor, *Development of Sensory Systems in Mammals*, Wiley Series in Neurobiology, chapter 2, pages 61–124. John Wiley & Sons, New York.

Garraghty, P. E. and Sur, M. (1993). Competitive interactions influencing the development of retinal axonal arbors in cat lateral geniculate nucleus. *Physiological Reviews*, 73(3):529–543.

Garraghty, P. E., Sur, M., and Sherman, S. M. (1986). Role of competitive interactions in the postnatal development of x and y retinogeniculate axons. *Journal of Comparative Neurology*, 251(2):216–39.

Goldowsky, B. N. and Newport, E. L. (1993). Modeling the effects of processing limitations on the acquisition of morphology: The less is more hypothesis. In Clark, E. V., editor, *The Proceedings of the Twenty-fourth Annual Child Language Research Forum*, pages 125–138, Stanford, CA. CSLI and University of Chicago Press.

Goodman, C. S. and Shatz, C. J. (1993). Developmental mechanisms that generate precise patterns of neuronal connectivity. *Cell*, 72:77–98.

Greenough, W. T., Black, J. E., and Wallace, C. S. (1987). Experience and brain development. *Child Development*, 58:539–559.

Guillery, R. W., Lamantia, A. S., Robson, J. A., and Huang, K. (1985). The influence of retinal afferents upon the development of layers in the dorsal lateral geniculate nucleus of mustelids. In Eglen, 1997.

Hahm, J.-O., Langdon, R. B., and Sur, M. (1991). Disruption of retinogeniculate afferent segregation by antagonists to nmda receptors. *Nature*, 351:568–570.

Haith, M. M. (1966). The response of the human newborn to visual movement. *Journal of Experimental Child Psychology*, 3:235–243.

Hamos, J. E., Horn, S. C. V., Raczkowski, D., and Sherman, S. M. (1987). Synaptic circuits involving an individual retinogeniculate axon in the cat. *Journal of Comparative Neurology*, 259:165–192.

Haussler, A. and von der Malsburg, C. (1983). Development of retinotopic projections: An analytical treatment. *Journal of Theoretical Neurobiology*, 2:47–73.

Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation.* MacMillan Publishing Company, New York.

Hebb, D. O. (1949). *Organization of Behavior*. John Wiley & Sons, Inc., New York.

Ikeda, H. and Tremain, K. (1978). The development of spatial resolving power of lateral geniculate neurones in kittens. *Experimental Brain Research*, 31(2):193–206.

Jeffery, G. (1985). Retinotopic order appears before ocular separation in developing visual pathways. *Nature*, 313:575–576.

Jeffery, G. (1989). Shifting retinal maps in the development of the lateral geniculate nucleus. *Developmental Brain Research*, 46:187–196.

Jeffery, G. (1990). The topographic relationship between shifting binocular maps in the developing dorsal lateral geniculate nucleus. *Experimental Brain Research*, 82:408–416.

Johnson, J. K. and Casagrande, V. A. (1993). Prenatal development of axon outgrowth and connectivity in the ferret visual system. *Visual Neuroscience*, 10:117–130.

Kalil, R. E., Dubin, M. W., Scott, G., and Stark, L. A. (1988). Elimination of action potentials blocks the structural development of retinogeniculate synapses. *Nature*, 323:156–158.

Keesing, R., Stork, D. G., and Shatz, C. J. (1992). Retinogeniculate development: The role of competition and correlated retinal activity. In *Advances in Neural Information Processing Systems 4*, pages 91–97, San Mateo, CA. Morgan Kaufmann.

Kohonen, T. (1993). Physiological interpretation of the self-organizing map algorithm. *Neural Networks*, 6:895–905.

Kohonen, T. (1995). *Self-organizing maps*. Springer.

Krekelberg, B. and Taylor, J. (1997). Nitric oxide: What can it compute? *Network: Computation in Neural Systems*, 8:1–16.

Lee, C. W. and Wong, R. O. L. (1996). Developmental patterns of on-off retinal ganglion cell activity lead to segregation of their afferents under a hebbian synaptic rule. In *Society For Neuroscience Abstracts*, volume 22, page 1202.

Linden, D. C., Guillery, R. W., and Cucchiaro, J. (1981). The dorsal lateral geniculate nucleus of the normal ferret and its postnatal development. *Journal of Comparative Neurology*, 203(2):189–211.

Lindstrom, S. (1982). Synaptic organization of inhibitory pathways to principal cells in the lateral geniculate nucleus of the cat. *Brain Research*, 234(2):447–53.

Linsker, R. (1986). From basic network principles to neural architecture: Emergence of spatial-opponent cells. *Proceedings of the National Academy of Science of the United States of America*, 83:7508–7512.

Lippe, W. R. (1994). Rhythmic spontaneous activity in the developing avian auditory system. *Journal of Neuroscience*, 14(2):1486–1495.

Lo, F. and Sherman, S. (1994). Feedback inhibition in the cat's lateral geniculate nucleus. *Experimental Brain Research*, 100(2):365–8.

Madison, D. V., Malenka, R. C., and Nicoll, R. A. (1991). Mechanisms underlying long-term potentiation of synaptic transmission. *Annual Review of Neuroscience*, 14:379–97.

Mason, C. (1982). Development of terminal arbors of retino-geniculate axons in the kitten–ii. electron microscopical observations. *Neuroscience*, 7(3):561–82.

Mastronarde, D. N. (1992). Nonlagged relay cells and interneurons in the cat lateral geniculate nucleus: Receptive-field properties and retinal input. *Visual Neuroscience*, 8(5):407–441.

McCormick, D., Trent, F., and Ramoa, A. (1995). Postnatal development of synchronized network oscillations in the ferret dorsal lateral geniculate and perigeniculate nuclei. *Journal of Neuroscience*, 15(8):5739–52.

Meister, M., Pine, J., and Baylor, D. (1994). Multi-neuronal signals from the retina: Acquisition and analysis. *Journal of Neuroscience Methods*, 51:95–106.

Meister, M., Wong, R. O. L., Baylor, D. A., and Shatz, C. (1991). Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science*, 252(5008):939–943.

Miller, K. D. (1990). Correlation-based models of neural development. In Gluck, M. A. and Rumelhart, D. E., editors, *Neuroscience and Connectionist Theory*, pages 267–253. Lawrence Erlbaum Associates, Hillsdale, NJ.

Miller, K. D. (1996). Synaptic economics: Competition and cooperation in synaptic plasticity. *Neuron*, 17:371–374.

Miller, K. D. (1997a). Analysis of lee and wong model of wong and oakley's results. $ftp://ftp.keck.ucsf.edu/pub/ken/leewong_anal.ps$.

Miller, K. D. (1997b). Equivalence of a sprouting-and-retraction model and correlation-based plasticity models of neural development and correlation-based rules with subtractive constraints. *Neural Computation*, 10:528–547.

Miller, K. D., Keller, J. B., and Stryker, M. (1989). Ocular dominance column development: Analysis and simulation. *Science*, 245(4918):605–615.

Miller, K. D. and MacKay, D. J. C. (1994). The role of constraints in hebbian learning. *Neural Computation*, 6:100–126.

Montero, V. M. (1991). A quantitative study of synaptic contacts on interneurons and relay cells of the cat lateral geniculate nucleus. *Experimental Brain Research*, 86:257–270.

Mooney, R., Madison, D. V., and Shatz, C. J. (1993). Enhancement of transmission at the developing retinogeniculate synapse. *Neuron*, 10(5):815–25.

Mooney, R., Penn, A. A., and Shatz, C. J. (1995). Periodic synaptic currents in the neonatal lgn are generated by retinal activity. In *Society for Neuroscience Abstracts*, volume 21, page 1504.

Movshon, J., Thompson, I., and Tolhurst, D. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology*, 183:53–77.

Nelson, C. A. and Horowitz, F. D. (1987). Visual motion perception in infancy: A review and synthesis. In Salapatek, P. and Cohen, L., editors, *Handbook of Infant Perception: From Perception to Cognition*, volume 2. Academic Press, New York.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, 14(1):11–28.

Norton, T. and Godwin, D. (1992). Inhibitory gabaergic control of visual signals at the lateral geniculate nucleus. *Progress in Brain Research*, 90:193–217.

O'Donovan, M. (1994). Calcium imaging of rhythmic network activity in the developing spinal cord of the chick embryo. *Journal Of Neuroscience*, 14(1):6354–6369.

Pape, H. and Eysel, U. (1986). Binocular interactions in the lateral geniculate nucleus of the cat: Gabaergic inhibition reduced by dominant afferent activity. *Experimental Brain Research*, 61(2):265–71.

Pape, H. and McCormick, D. (1995). Electrophysiological and pharmacological properties of interneurons in the cat dorsal lateral geniculate nucleus. *Neuroscience*, 68(4):1105–1125.

Penn, A. A., Gallego, R., Mooney, R., and Shatz, C. J. (1995). Spontaneous retinal inputs drive postsynaptic action potentials in the lgn. In *Society for Neuroscience Abstracts*, volume 21, page 1504.

Penn, A. A., Riquelme, P. A., Feller, M. B., and Shatz, C. J. (1998). Competition in retinogeniculate patterning driven by spontaneous activity. *Science*, 279:2108–2112.

Poggio, T., Torre, V., and Koch, C. (1985). Computational vision and regularization theory. *Nature*, 317(6035):314–319.

Ramoa, A. S. and McCormick, D. A. (1994a). Developmental changes in electrophysiological properties of lgnd neurons during reorganization of retinogeniculate connections. *Journal of Neuroscience*, 14(4):2089–2097.

Ramoa, A. S. and McCormick, D. A. (1994b). Enhanced activation of nmda receptor responses at the immature retinogeniculate synapse. *Journal of Neuroscience*, 14(4):2098–2105.

Roe, A. W., Garraghty, P. E., and Sur, M. (1989). Terminal arbors of single on-center and off-center x and y retinal ganglion cell axons within the ferret's lateral geniculate nucleus. *Journal of Comparative Neurology*, 288:208–242.

Sanchez-Vives, M. and McCormick, D. (1997). Functional properties of perigeniculate inhibition of dorsal lateral geniculate nucleus thalamocortical neurons in vitro. *Journal Of Neuroscience*, 17(22):8880–8893.

Sanderson, K. J. (1971). Visual field projection columns and magnification factors in the lateral geniculate nucleus of the cat. *Experimental Brain Research*, 13:159177.

Sernagor, E. and Grzywacz, N. (1996). Influence of spontaneous activity and visual experience on developing retinal receptive fields. *Current Biology*, 6(11):1503–1508.

Shatz, C. J. (1983). The prenatal development of the cat's retinogeniculate pathway. *Journal of Neuroscience*, 3(3):482–499.

Shatz, C. J. (1994). Role of electrical activity in nervous system development. *Molecular Biology of the Cell*, 5(Supplemental):2A.

Shatz, C. J. and Kirkwood, P. A. (1984). Prenatal development of functional connections in the cat's retinogeniculate pathway. *Journal of Neuroscience*, 4(5):1378–1397.

Shatz, C. J. and Stryker, M. P. (1988). Prenatal tetrodotoxin infusion blocks segregation of retinogeniculate afferents. *Science*, 242(4875):87–89.

Sherman, M. and Koch, C. (1990). Thalamus. In Shepherd, G. M., editor, *The Synaptic Organization of the Brain*, pages 246–278. Oxford University Press, New York, 3rd edition.

Sherman, S. M. (1985). Development of retinal projections to the lateral geniculate nucleus. *TINS*, pages 350–355.

Shotwell, S. L., Carla, C. J., and Luskin, M. B. (1986). Development of glutamic acid decarboxylase immunoreactivity in the cat's lateral geniculate nucleus. *Journal of Neuroscience*, 6(5):1410–23.

Sillito, A. and Kemp, J. (1983). The influence of gabaergic inhibitory processes on the receptive field structure of x and y cells in cat dorsal lateral geniculate nucleus(dlgn). *Brain Research*, 277(1):63–77.

Simon, D. K. and O'Leary, D. D. M. (1990). Limited topographic specificity in the targeting and branching of mammalian retinal axons. *Developmental Biology*, 137:125–134.

Singer, W., Poppel, E., and Creutzfeldt, O. (1972). Inhibitiory interaction in the cat's lateral geniculate nucleus. *Experimental Brain Research*, 14:210–226.

Sirosh, J. and Miikkulainen, R. (1993). How lateral interaction develops in a self-organizing feature map. In *IEEE International Conference on Neural Networks*, San Francisco, CA.

Sirosh, J. and Miikkulainen, R. (1994). Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. Technical Report AI94-225, The University of Texas at Austin, Department of Computer Sciences, Austin, Tx.

Sirosh, J. and Miikkulainen, R. (1996). Self-orgainization and functional role of lateral connections and multisize receptive fields in the primary visual cortex. *Neural Processing Letters*.

Soltesz, I. and Crunelli, V. (1992). Gabaa and pre- and post-synaptic gabab receptor-mediated responses in the lateral geniculate nucleus. *Progress in Brain Research*, 90:151–169.

Soltesz, I., Lightowler, S., Leresche, N., and Crunelli, V. (1989). On the properties and origin of the gabab inhibitory postsynaptic potential recorded in morphologically identified projection cells of the cat dorsal lateral geniculate nucleus. *Neuroscience*, 33(1):23–33.

Sperry, R. W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proceedings of the National Academy of Sciences of the United States of America*, 50:703–710.

Sretavan, D. and Shatz, C. J. (1984). Prenatal development of individual retinogeniculate axons during the period of segregation. *Nature*, 308.

Sretavan, D. W. and Shatz, C. J. (1986). Prenatal development of retinal ganglion cell axons: Segregation into eye-specific layers within the cat's lateral geniculate nucleus. *Journal of Neuroscience*, 6(1):234–251.

Sretavan, D. W. and Shatz, C. J. (1987). Axon trajectories and pattern of terminal arborization during the prenatal development of the cat's retinogeniculate pathway. *Journal of Comparative Neurology*, 255:386–400.

Sretavan, D. W., Shatz, C. J., and Stryker, M. P. (1988). Modification of retinal ganglion cell axon morphology by prenatal infusion of tetrodotoxin. *Nature*, 336:468–471.

Sterling, P. (1990). Retina. In Shepherd, G. M., editor, *The Synaptic Organization of the Brain*, pages 3–32. Oxford University Press, New York, 3rd edition.

Sur, M. (1988). Development and plasticity of retinal x and y axon terminations in the cat's lateral geniculate nucleus. *Brain, Behavior and Evolution*, 31:243–251.

Sutton, J. K. and Brunso-Bechtol, J. K. (1993). Dendritic development in the dorsal lateral geniculate nucleus of ferrets in the postnatal absence of retinal input: A golgi study. *Journal of Neurobiology*, 24(3):317–334.

Sutton, J. K. and Brunso-Bechtold, J. K. (1991). A golgi study of dendritic development in the dorsal lateral geniculate nucleus of normal ferrets. *Journal of Comparative Neurology*, 309:71–85.

Swindale, N. V. (1996). The development of topography in visual cortex: A review of models. *Network: Computation in Neural Systems*, 7:161–247.

Turkewitz, G. and Kenny, P. A. (1982). Limitations on input as a basis for neural organization and perceptual development: a preliminary theoretical statement. *Developmental Psychobiology*, 15:357–368.

Turrigiano, G., Leslie, K., Desai, N., and Nelson, S. (1996). Activity-dependent regulation of quantal amplitude in visual cortical cultures. In *Society for Neuroscience Abstracts*, number 22, page 275.

Udin, S. B. and Fawcett, J. W. (1988). Formation of topographic maps. *Annual Review of Neuroscience*, 11:289–327.

Uhlrich, D. and Cucchiaro, J. (1992). Gabaergic circuits in the lateral geniculate nucleus of the cat. progress in brain research. *Progress in Brain Research*, 90:171–92.

Voigt, T., Naito, J., and Wassle, H. (1983). Retinotopic scatter of optic tract fibres in the cat. *Experimental Brain Research*, 52:25–33.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the stiata cortex. *Kybernetik*, 14:85–100.

von der Malsburg, C. and Willshaw, D. J. (1977). How to label nerve cells so that they can interconnect in an ordered fashion. *Proceedings of the National Academy of Sciences of the United States of America*, 74(5176-5178).

von Krosigk, M., Bal, T., and McCormick, D. (1993). Cellular mechanisms of a synchronized oscillation in the thalamus. *Science*, 261(5119):361–4.

Weber, A. J., Kalil, R. E., and Hickey, T. L. (1986). Genesis of interneurons in the dorsal lateral geniculate nucleus of the cat. *Journal of Comparative Neurology*, 252:285–391.

White, C. and Sur, M. (1992). Membrane and synaptic properties of developing lateral geniculate nucleus neurons during retinogeniculate axon segregation. *Proceedings of the National Academy of Sciences of the United States of America*, 89(20):9850–4.

Whitelaw, V. A. and Cowan, J. D. (1981). Specificity and plasticity of retinotectal connections: A computational model. *The Journal of Neuroscience*, 1(12):1369–1387.

Williams, S., Turner, J., Anderson, C., and Crunelli, V. (1996). Electrophysiological and morphological properties of interneurones in the rat dorsal lateral geniculate nucleus in vitro. *Journal of Physiology*, 490(1):129–47.

Willshaw, D. J. and von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London*, 194:431–445.

Wiskott, L. and Sejnowski, T. (1997). Objective functions for neural map formation. In *4th Joint Symposium on Neural Computation, Los Angeles, CA*.

Wiskott, L. and Sejnowski, T. J. (1998). Constrained optimization for neural map formation: A unifying framework for weight growth and normalization. *Neural Computation*, 10(3):671–716.

Wong, R. O. and Oakley, D. M. (1996). Changing patterns of spontaneous bursting activity of on and off retinal ganglion cells during development. *Neuron*, 16:1087–1095.

Wong, R. O. L., Cherjavsky, A., Smith, S. J., and Shatz, C. J. (1995). Early functional neural networks in the developing retina. *Nature*, 374(6524):716–718.

Wong, R. O. L., Meister, M., and Shatz, C. J. (1993). Transient period of correlated bursting activity during development of the mammalian retina. *Neuron*, 11:923–938.

Zahs, K. R. and Stryker, M. P. (1985). The projection of the visual field onto the lateral geniculate nucleus of the ferret. *Journal of Comparative Neurology*, 241:210–224.