CHAPTER ONE

Theoretical Models of Neural Circuit Development

Hugh D. Simpson,* Duncan Mortimer,* *and* Geoffrey J. Goodhill*,†

Contents

Abstract

Proper wiring up of the nervous system is critical to the development of organisms capable of complex and adaptable behaviors. Besides the many experimental advances in determining the cellular and molecular machinery that carries out this remarkable task precisely and robustly, theoretical approaches have also proven

* Queensland Brain Institute, The University of Queensland, St Lucia, Australia
† School of Mathematics and Physics, The University of Queensland, St Lucia, Australia

Current Topics in Developmental Biology, Volume 87
ISSN 0070-2153, DOI: 10.1016/S0070-2153(09)01201-0

to be useful tools in analyzing this machinery. A quantitative understanding of these processes can allow us to make predictions, test hypotheses, and appraise established concepts in a new light. Three areas that have been fruitful in this regard are axon guidance, retinotectal mapping, and activity-dependent development. This chapter reviews some of the contributions made by mathematical modeling in these areas, illustrated by important examples of models in each section. For axon guidance, we discuss models of how growth cones respond to their environment, and how this environment can place constraints on growth cone behavior. Retinotectal mapping looks at computational models for how topography can be generated in populations of neurons based on molecular gradients and other mechanisms such as competition. In activity-dependent development, we discuss theoretical approaches largely based on Hebbian synaptic plasticity rules, and how they can generate maps in the visual cortex very similar to those seen *in vivo*. We show how theoretical approaches have substantially contributed to the advancement of developmental neuroscience, and discuss future directions for mathematical modeling in the field.

1. Introduction

Wiring up the developing nervous system involves a vast array of challenges. Billions of connections must be made appropriately, but rather than uniquely specifying each of these connections, nervous systems appear to adopt strategies for development based on general rules which apply across large numbers and patterns of connections. Such general rules include, for instance, how to build a topographic map based on molecular gradients and neural firing patterns, and how to change the strengths of synapses based on correlations in presynaptic and postsynaptic firing. As discussed in the other chapters of this volume, well-designed experimental approaches, combined with rapidly developing technologies for imaging the developing nervous system, have dramatically expanded our understanding of some of these general strategies, and their specific molecular implementations. However, in common with most of biology, this understanding is mostly qualitative. While the "thinking models" developed are extremely useful, they are usually not up to the task of making fully *quantitative* predictions about outcomes in different situations. Qualitative models can rapidly become overwhelmed by the huge complexity of the data, whereas quantitative models offer the potential to tease apart which variables and interactions are key to understanding the phenomena under consideration (for discussions, see Abbott, 2008; Goodhill, 2007; Mogilner *et al.*, 2006).

To illustrate these general issues, it is interesting to draw an analogy with the historical progression of our understanding of certain physical systems. Consider, for example, the effect of pressure and volume on the temperature

of gases. It is relatively straightforward to obtain qualitative data demonstrating the general relationships between these variables. More quantitative measurements are more challenging, but begin to suggest lawful relationships between the variables. But it is only when these relationships are expressed in the form of a mathematical equation (the Ideal Gas Law) that one has the power to predict behavior under a wide variety of circumstances. Furthermore, the equation provides a target for understanding *why* these relationships exist. Indeed, subsequent theoretical work showed that the Ideal Gas Law could be derived from a statistical–mechanical approach, basing it firmly in the behavior of individual molecules. However, while the Ideal Gas Law is fundamental to our understanding of gases, it ignores key effects such as intermolecular interactions, and therefore fails under certain circumstances. Thus, mathematical descriptions can still be extremely powerful tools, even when they clearly fail to include details of the real world which are known to be important. In fact, essentially all mathematical models of physical phenomena ignore details which may be important under some circumstances. It is thus not surprising that analogous limitations in scope are true of mathematical models in biology. As demonstrated by the Ideal Gas Law, this limited scope does not mean these models are not insightful and useful.

Of course, even the most optimistic computational neuroscientist would not claim that theoretical modeling of neural wiring development has so far resulted in anything like the physical laws mentioned above. Nonetheless, there are several domains where mathematical models have clarified our understanding of the capabilities and limitations of potential developmental mechanisms (van Ooyen, 2003). In this chapter, we will focus on three representative examples: axon guidance, retinotectal map formation, and activity-dependent development. In the first section, we look at the mathematical approaches that have been applied to understanding a variety of phenomena in growth cone behavior. For the most part, specific unifying principles have not emerged from these models; hence we organize models by general approach taken. In retinotectal map formation theoretical models have a long history of close engagement with experimental data, ranging from older surgically induced map manipulations, to more recent data regarding genetic manipulations of Eph and ephrin gradients. In this section, we organize models historically. In activity-dependent development, the dominant paradigm has been to apply mathematical models of Hebbian learning to a range of developmental phenomena, for instance in the visual system. In this section, we therefore start by focusing on one of the simplest mathematical models of Hebbian learning, and then discuss more sophisticated approaches based on this and other principles. Given that the rest of this volume reviews experimental data in some detail, here we assume a basic knowledge of this data and focus instead on the mathematical models.

2. Theoretical Modeling in Axon Guidance

To form the precise arrangement of connections between neurons seen in mature nervous systems, axons must be accurately directed to find their targets. This process depends heavily on the interaction between the growth cone and its environment. Current modeling of growth cone guidance and dynamics is a fragmented field. This is largely due to a dearth of quantitative data on growth cone behavior, which makes it difficult to unify experimental findings under one framework. Thus, to survey the relevant literature, we divide models into three rough categories based on the modeling approach taken: phenomenological, mechanistic, and abstract (see Maskery and Shinbrot, 2005 for another review).

2.1. Phenomenological models

Phenomenological models describe experimentally observed behavior by fitting it to a mathematical framework without making explicit reference to underlying mechanisms; much like Boyle's empirical observations of the inverse relationship between pressure and volume of a gas (though alas, seldom as mathematically concise for growth cones). This style of model has been applied to microtubule dynamics (Odde and Buettner, 1998; Odde *et al.*, 1996), filopodial dynamics (Buettner, 1995; Odde and Buettner, 1998), the "random walk" behavior of the growth cone (Katz *et al.*, 1984), and interactions between filopodia and target cells (Buettner, 1996).

Buettner's work (Buettner, 1995; Buettner *et al.*, 1994) provides a good example of this kind of model. By statistically analyzing time-lapse images of growth cones undergoing dynamic changes in morphology (Buettner *et al.*, 1994), Buettner was able to develop some probabilistic rules regarding filopodial dynamics, which were formalized as a model in Buettner (1995). Growth cone morphology is described by the instantaneous length and angle of each filopodium, and the dynamics of the filopodia are characterized by five parameters: the rate at which filopodia extend, the rate of retraction, the average rate at which new filopodia are initiated (modeled as a Poisson process), and two parameters for the scale and shape of a gamma-distribution for the time over which a filopodium extends before it begins to retract. Along with a simple probabilistic rule for where a filopodium initiates, and under the assumption that filopodia extend radially from the center of the growth cone, this model gives qualitatively realistic morphologies, which also satisfied quantitative constraints such as the correct average number of filopodia. Buettner then analyzed model growth cones with different parameter values, simulating their interaction with a contact-mediated guidance cue at different distances

from the growth cone body. From this, she was able to extract some useful quantities; for example, the ratio of extension and retraction rates, which determines the mean time required before the first filopodium on the growth cone contacts the cue. By mapping the effects of external cues such as guidance molecules onto the parameters of the model, one can hope to gain some intuition as to how those cues might operate.

Growth cone behavior and axon extension are thought to be mediated by partially independent, but related, processes. This has led to modeling studies focused on characterizing the interaction between the two. Work in the mid-1980s (Katz *et al.*, 1984) argued that under some circumstances axon elongation can be regarded as a one-dimensional uncorrelated random walk, with parameters varying with neuronal type, substrate, and chemical environment. In the mid-1990s, further work in this vein demonstrated that, rather than being uncorrelated, axon extension exhibits significant anticorrelation on short timescales—4 min, to be precise (Odde *et al.*, 1996). In other words, there is a tendency for any increases in axonal length to be followed by a retraction with a lag of about 4 min. In the same article, the authors also showed significant correlation between the dynamics of microtubule polymerization and growth cone advance (Fig. 1.1).

Such studies, while not in themselves providing explanations for why the observed behaviors occur, are nonetheless very useful for the precise constraints they suggest for growth cone behavior, and the hypotheses they

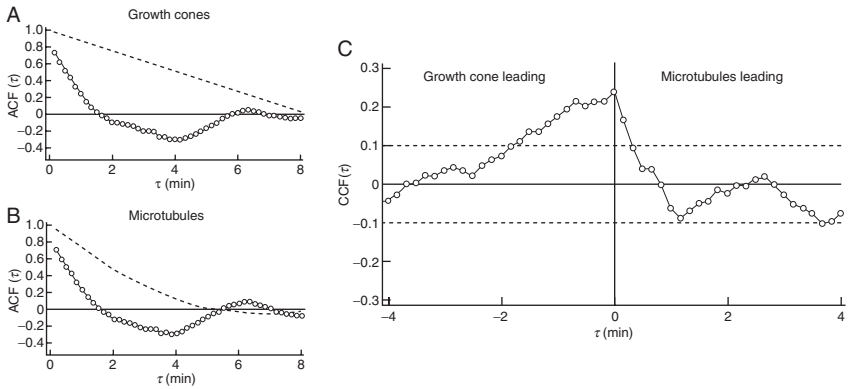

Figure 1.1 Temporal correlation in microtubule polymerization and growth cone extension. The detrended position autocorrelation functions (ACF) for growth cone and microtubule advance ((A) and (B) respectively) are plotted against time (unfilled circles). For comparison, the dashed line shows the autocorrelation function for a random walk generated by randomly shuffling measured displacements from the original data series. The observed form of the true autocorrelation function illustrates that net growth cone or microtubule advance tends to be followed by retraction about 4 min later. (C) Plotting the cross-correlation function (CCF) between growth cone and microtubule dynamics illustrates strong interactions between the two processes (from Odde *et al.*, 1996; reproduced with permission from John Wiley & Sons, Inc.).

spark concerning how underlying molecular events might contribute to the observed dynamics. For example, obtaining a better quantitative understanding of the connection between microtubule polymerization and growth cone advance through crosscorrelation analysis may aid in understanding the role these cytoskeletal elements play in growth cone motility.

2.2. Mechanistic models

Mechanistic models focus on actual biophysical mechanisms and their role in growth cone behavior; for example, the dynamics of certain cytoskeletal components and how they affect the rate of neurite growth (Hely and Willshaw, 1998; Kiddie *et al.*, 2005; Mogilner and Rubinstein, 2005), the spatiotemporal calcium concentration distributions underlying growth cone turning (Aeschlimann and Tettoni, 2001), or the patterns of activity of the Rho GTPases and their influence on growth cone guidance (Sakumura *et al.*, 2005). Such models are difficult to construct at present, partly due to the complexity of growth cone biochemistry, and partly due to the lack of experimental data on important quantities such as reaction rate constants, as well as concentrations of and interactions between molecular species.

Experimental evidence suggests that tubulin molecules are synthesized almost exclusively in the soma, and then assembled into microtubules predominantly in the growth cone (Kobayashi and Mundel, 1998). This implies that axon outgrowth is limited by the rate at which microtubules can be transported to regions of active extension. A number of theoretical models have explored this idea, and included various effects such as diffusive and active transport of tubulin monomers, competition between neurites for tubulin, viscoelastic stretching of axon segments, calcium-induced microtubule depolymerization, and varying intrinsic rates of tubulin polymerization and depolymerization within different growth cones (Graham and van Ooyen, 2001; Kiddie *et al.*, 2005; McLean *et al.*, 2004; van Veen and van Pelt, 1994; reviewed in van Ooyen, 2001). These models have been successively refined, ultimately incorporating compartment-based modeling with dynamic compartment allocation (Graham and van Ooyen, 2001; Kiddie *et al.*, 2005). Most strikingly, this modeling program has demonstrated that small variations in polymerization and depolymerization rates in the growth cones of different neurites can lead to sharp changes in elongation rate, resulting in growth cone pausing and neurite retraction.

Mogilner and Rubinstein's (2005) exploration of the physics underlying filopodial extension is an excellent, and unusually successful example of this form of modeling, applicable to many systems besides growth cones. Length, growth rate, and spacing between adjacent filopodia are analyzed by modeling the effects of membrane elasticity, g-actin diffusion, lamellipodial extension rate, and the buckling of f-actin bundles under strain. This analysis indicates that the length of a filopodium is determined by the number of

bundled actin filaments in its core. For less than about 10 bundled actin filaments, the strain exerted on the bundle by the membrane is sufficient to cause buckling for very short filopodia. As the number of included filaments increases, it becomes less likely for the filopodium to buckle, but at the same time more g-actin is required for the structure to continue extending, so that when the number of filaments is too large, the filopodium is also unable to extend. This led the authors to calculate an optimal value of approximately 30 actin filaments, giving the best trade-off between mechanical stability and g-actin diffusion (Fig. 1.2). In conclusion, the authors obtained average lengths for extending filopodia of between 1 and 10 μm, values which agree well with experiment.

This model is, however, far from being able to predict the responses of growth cones to a guidance cue gradient, and it is this task that the model of Sakumura *et al.* (2005) tackles. This model is based on the well known and widely studied Rho GTPase system. The Rho GTPases are a family of proteins known to play important roles in actin-based motility in a number of cellular systems (Giniger, 2002). The authors recognize the lack of information about the kinetic parameters characterizing the interactions

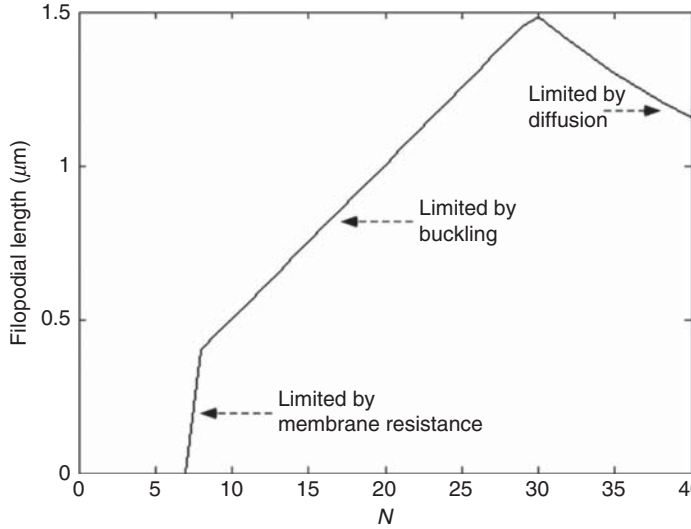

Figure 1.2 Filopodial length is limited by diffusion and buckling. Depending on the number of bundled actin fibers, *N*, various effects act to limit the length of filopodia. For too few bundled fibers, the actin polymerization force is not sufficient to overcome membrane resistance, and filopodial growth is strongly attenuated. For more than about eight bundled fibers, filopodia can overcome membrane resistance, but buckle when they reach a length dependent on the number of fibers. Finally, for greater than about 30 fibers, diffusion cannot deliver unpolymerized g-actin to the tips of the filopodia fast enough to maintain growth, and extension stalls (from Mogilner and Rubinstein, 2005; reproduced with permission from Elsevier).

between the Rho GTPases, and instead take a qualitative approach, randomly assigning values to various unknown parameters, then classifying the behavior of the dynamical system which results. The main conclusion drawn from this analysis is that for a large range of kinetic parameters, the Rho GTPase network exhibits hypersensitivity to changes in concentration of a particular class of molecules, the Rho guanosine exchange factors (Rho GEFs). As the concentration of Cdc42-GEFs increases above a threshold, the system dynamics moves from stable attractor into a rapid oscillatory regime (Fig. 1.3). The authors posit that this oscillation is characteristic of the streamlined mode of growth cone movement, while the stable attractor corresponds to the "slow-expansion" mode in which the growth cone takes on a more complex morphology. However, a more recent analysis of the Rho GTPase network, assuming a slightly different network connectivity, finds multistable rather than oscillatory solutions (Jilkine *et al.*, 2007).

At this stage, fully mechanistic models are underconstrained due to a lack of quantitative data for crucial parameters. Ideally, such models require knowledge of reaction rate constants, concentrations of key molecules inside the growth cone and an accurate map of interactions between the species of molecules present. Finally, it is worth noting that axon guidance shares many similarities with eukaryotic chemotaxis (Mortimer *et al.*, 2008; von Philipsborn and Bastmeyer, 2007), a field for which mathematical models have proved very useful.

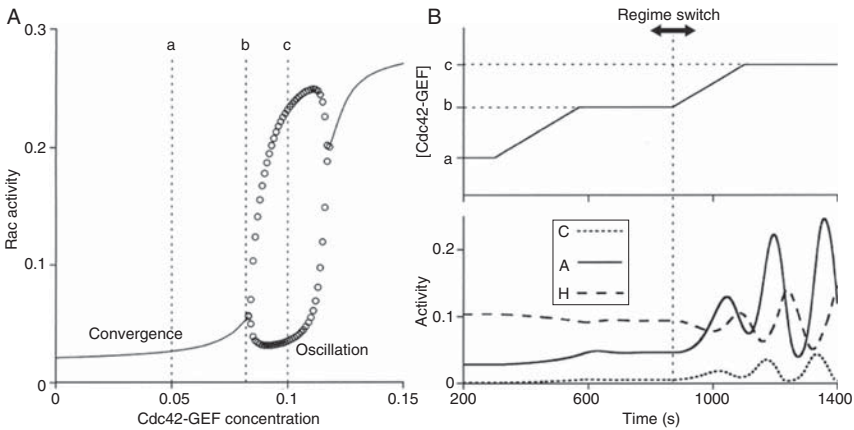

Figure 1.3 Rho GTPase network dynamics can depend sensitively on Cdc42-GEF concentration. (A) For sufficiently low concentrations of Cdc42-GEF, the network settles into a stable state. As Cdc42-GEF concentration is increased, a threshold is passed at which the network enters a large-amplitude oscillatory mode. (B) External control of Cdc42-GEF concentration (perhaps by receptor activation) can induce a regime change in the activation dynamics of Cdc42 (dashed line), Rac (solid line), and RhoA (dotted line) (from Sakumura *et al.*, 2005; reproduced with permission from Elsevier).

2.3. Abstract models

Abstract models discuss the mechanism(s) underlying a particular behavior in more general terms, without explicit reference to the molecular details of its implementation. The influential model of Meinhardt (1999) provides an archetypal example of this kind of approach; it describes a general strategy for the growth cone to amplify weak signals from its environment. In Meinhardt's model, this amplification is achieved by coupling the external signal to a pattern formation system involving local activation and long range inhibition. The system begins in a spatially symmetric, but unstable steady state. Symmetry is broken by the external signal, which pushes the system into a stable, asymmetric state reflecting the direction in which the symmetry was broken. One difficulty with this approach, recognized by Meinhardt, is that the system then becomes stuck: it is unable to respond to new inputs, such as a change in the external signal. Meinhardt works around this by postulating a second mechanism which serves to reset the system to its original, unstable state. He also makes several other generalizations of his basic model which qualitatively capture other aspects of chemotactic devices, such as the formation of filopodia.

The work of Goodhill, Urbach, and Baier in the late 1990s (Goodhill, 1997; Goodhill and Baier, 1998; Goodhill and Urbach, 1999; Urbach and Goodhill, 1999) following Berg and Purcell (1977) can also be placed in this category. Growth cones are believed to sense and respond to gradients by comparing receptor binding across their spatial extent: the side of a growth cone exposed to the highest concentration of ligand will, on average, also display the largest amount of receptor binding. If growth cones do use such a spatial-sensing strategy (cf. temporal sensing for bacteria), then in order for a growth cone to detect and reliably respond to a chemical gradient, the noise due to fluctuations in receptor binding cannot be much larger than the difference in receptor binding across its spatial extent. By modeling the physics of receptor–ligand interaction, Goodhill and Urbach estimated the limitations growth cones face when responding to chemical gradients. If the root mean-squared error in a concentration measurement is given by σ_C, then the error associated with taking the difference between two such measurements is $\sigma_{\Delta C} = \sqrt{2}\sigma_C$. This gives an order-of-magnitude lower bound on the difference in concentration, $\Delta C_{\min}$, that the growth cone can detect: $\Delta C_{\min} \approx \sqrt{2}\sigma_C$. A strong point of this approach is that the results should apply regardless of the intracellular mechanism used to extract the gradient direction from the spatiotemporal distribution of bound receptors—the details of the intracellular machinery can only reduce performance. Hence, their results give an upper bound on gradient-sensing performance of growth cones.

Abstract models have also been developed which simulate growth cone behavior for direct comparison to experiment, rather than discussing general restrictions on gradient sensing. Goodhill *et al.* (2004) examined the

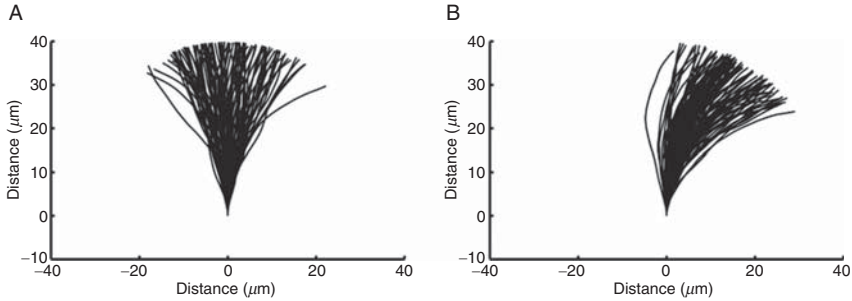

Figure 1.4 Growth cone trajectories generated by a filopodium-based model. (A) Without exposure to a guidance cue, axons project straight on average. (B) In contrast, when exposed to a 5% gradient pointing to the right, the axons turn (from Goodhill *et al.*, 2004).

implications of different models for filopodial formation and the subsequent influence of filopodia on the growth cone (Fig. 1.4). One finding from this work was that, given some simple assumptions about how an external signal influences the distribution of filopodia on the growth cone, testable experimental predictions could be made—namely, that the growth cone may display qualitatively different sensitivity curves to attractive and repulsive gradients. A further example is provided by Xu *et al.* (2005) who forgo any speculation about the role of specific structures in growth cone chemotaxis, focusing instead on the issues of adaptation, temporal averaging, and spatial averaging. These ideas are then applied to a simplified model growth cone which nonetheless displays realistic trajectories when appropriate parameters are chosen. This work demonstrated that experimentally observed behaviors displayed when growth cones are exposed to a gradient for a long period of time (Rosoff *et al.*, 2004), can be reproduced without explicitly implementing an adaptation mechanism. This suggests that experimental work probing growth cone behavior on short timescales (Ming *et al.*, 2002; Piper *et al.*, 2005) may require reinterpretation (e.g., see Mortimer *et al.*, 2008).

Such direct simulation models have also been applied to understanding the kinds of effects that come into play when more than one axon is considered. One of the earliest examples of computer modeling in axon guidance (Katz and Lasek, 1985) studied the formation of axon sheets and axon bundles in the vertebrate spinal cord. In this approach, axon extension was modeled as a simple random walk on a hexagonal lattice. Three additional constraints were investigated to determine which were necessary to produce ordered sheets of axon extension: directional persistence (achieved by limiting the degree to which an individual axon could bend), differential adhesion (in which axons preferred binding to the substrate than to other axons), and initial polarization (in which axons were assumed to initially project in a uniform direction). By systematically investigating the

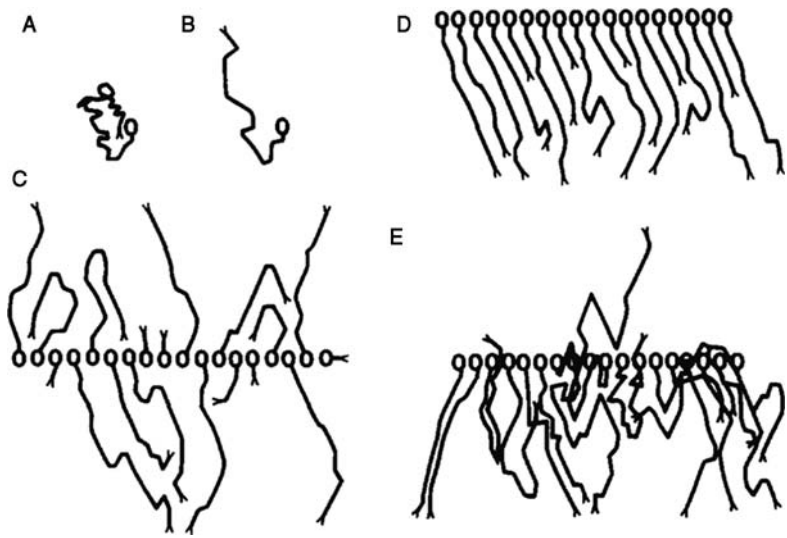

Figure 1.5 Models can be used to test the importance of various assumptions. In the model of Katz and Lasek (1985), the degree to which axons wander can vary (compare (A) with (B)); neurons can project in an unoriented or oriented manner (compare (C) with (D)), and axon crossing can be differentially inhibited (compare (D) with (E)). Only in case (D) do the axons project in ordered sheets (from Katz and Lasek, 1985; reproduced with permission from Elsevier).

effect of strengthening or weakening these constraints, the authors found necessary conditions to ensure the growth of ordered sheets of axons; in particular, that for a dense enough field of axons, the inhibition of axon crossing along with initial polarization of outgrowth direction was sufficient for ordering (Fig. 1.5). Hentschel and van Ooyen (1999) and Krottje and van Ooyen (2007) are more recent simulation models that allow for interactions between multiple axons. The latter provides a general framework for studying systems level effects in axon guidance. Here, the authors develop a set of mathematical tools allowing the simulation of the development of multiple axons through complex domains including any number of sources of guidance factors (which can even be released by the growth cones themselves).

3. Theoretical Modeling in the Retinotectal System

The ordered projection from the retina to the midbrain tectum is often referred to as the retinotectal map. (The corresponding midbrain target in mammals is the superior colliculus; hence also retinocollicular map.)

The map formed by these connections is described as topographic or retinotopic because neighborhood spatial relationships are preserved; that is, two cells that are near each other in the retina project to neighboring positions in the tectum (Fig. 1.6).

The development of midbrain retinotopy requires the axons of spatially ordered cells in the retina to navigate through a complex three-dimensional environment to their target, and re-establish their original order. We focus

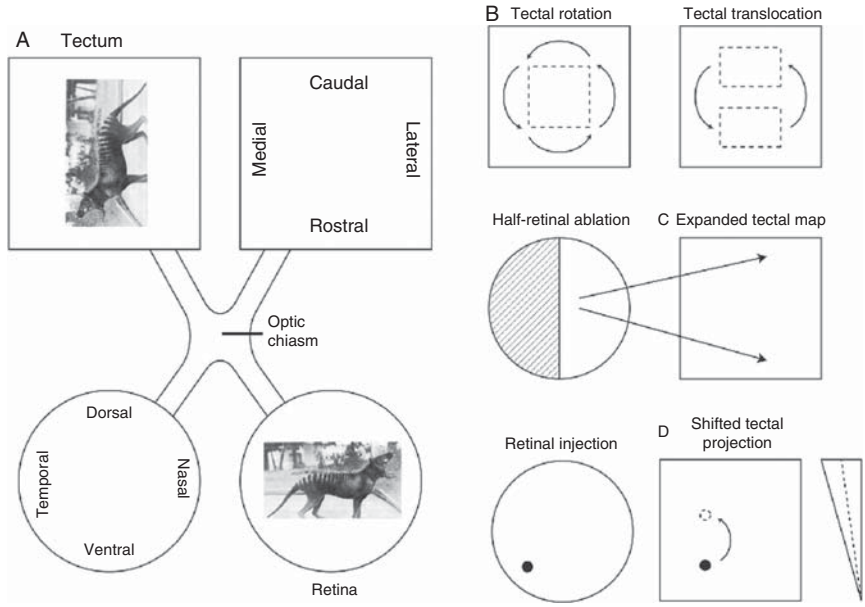

Figure 1.6 Normal and experimental variants in retinotectal maps. (A) Schematic of the normal retinotectal projection; nasal (temporal) retina maps to caudal (rostral) tectum and dorsal (ventral) retina maps to lateral (medial) tectum. In this way, an image (here the presumed extinct Tasmanian Tiger) presented to the retina is reproduced in the pattern of activity in the midbrain. (B) Schematics of tectal graft manipulations. On the left, a central piece of tectum is rotated (along with any markers on the tectum); maps formed after these procedures were also rotated. On the right, a graft translocation experiment is pictured, where two pieces of tectum have their positions exchanged. In this case, the corresponding map is often found to have similar translocations. (C) Retinal ablation experiment. Half of the retina is ablated, and instead of remaining confined to its usual half-tectum, the projection of the nonablated retina axons expands to fill the entire tectum. (D) Schematic (simplified) of a more recent misexpression study. Here, a gradient of repulsive guidance molecules (solid triangle, right) is distributed along the rostrocaudal axis of the tectum. Normally, this gradient is involved in determining where retinal ganglion cell axons project to. Here, for example, a group of ventrotemporal cells in the retina (filled circle, left) project to rostromedial tectum (filled circle, center). If this gradient is artificially weakened (dotted line, right), the reduced level of repulsion allows axons from the same retinal origin to terminate more caudally (dashed circle, center).

on the final step in this sequence: how the axonal terminations order themselves correctly on the tectum. Mathematical models of this process generally fall into the abstract category described in Section 2, and so are better categorized as belonging to one of the two main generations of modeling (and experiments) so far (1) earlier systems-based data and models and (2) more recent molecular level data and models. The large number of detailed models built around the systems results necessitates more brief descriptions, while the more recent models are considered in greater detail. This should not be taken as a suggestion that earlier models have been uniformly improved on by newer models that utilize the latest data on the retinotectal system. In fact earlier models tended to be more flexible and mathematically sophisticated than more recent ones, so that the core problems in retinotectal modeling will likely require insights from both generations. We preface the discussion of these generations of models with a brief description of hypotheses relevant to retinotectal mapping.

3.1. Hypotheses underlying map formation

A number of mechanisms have been identified that make significant contributions to retinotectal map formation. Models generally employ at least one of the following:

1. The *chemoaffinity* hypothesis of retinotectal mapping posits that gradients of molecular markers across a population of cells allow positions of cells to be uniquely identified by the level of marker on the cell, and that if there is a second population of cells with matched gradients, then this can allow for a topographic map to form between them (Sperry, 1963). Such gradients can also provide vector signals to guide ingrowing axons.
2. *Competition*, usually between retinal ganglion cell (RGC) axons for some limiting resource such as target space, has been consistently used in models from both generations (Fraser and Perkel, 1990; Gaze and Keating, 1972; Goodhill and Xu, 2005; Prestige and Willshaw, 1975; Schmidt and Easter, 1978).
3. Axon *branching* (in the form of interstitial branching, backbranching, growth cone bifurcation, or terminal arborization) contributes to retinotectal map formation, although to different degrees in different model organisms (Fujisawa *et al.*, 1982; Kaethner and Stuermer, 1992; McLaughlin and O'Leary, 2005; O'Rourke and Fraser, 1990; Roskies and O'Leary, 1994; Simon and O'Leary, 1992).

Spontaneous or correlated waves of neural *activity* in the retina and tectum contribute to map formation as well, though this is more of a secondary role. Generally speaking, activity is considered to carry out a refinement of the projection once the preceding mechanisms (above) have generated an initial (coarse) map (Cang *et al.*, 2008b; Debski and Cline, 2002; McLaughlin

et al., 2003; O'Rourke *et al.*, 1994; Ruthazer and Cline, 2004; Ruthazer *et al.*, 2003; Schmidt, 1990).

The four mechanisms already listed are well established in retinotectal map formation and are supported by a wealth of evidence. An additional hypothesized mechanism is *marker induction*, which refers to the respecification of tectal (and possibly retinal) cues during map development and regeneration. It is most commonly used to mean respecification of tectal markers due to the presence of RGC axons, and is hence often referred to as retinal induction. It was first formalized as a hypothesis for modeling in topographic map formation by von der Malsburg and Willshaw (1977). Early data suggested such a mechanism might be at work (Gaze *et al.*, 1974; Schmidt, 1978; Sharma, 1972; Yoon, 1976, 1980) but molecular demonstrations of this respecification have been limited (King *et al.*, 2003; Rodger *et al.*, 2000). There are other mechanisms that could potentially impart topographic information, such as the order of fibers within the optic nerve/tract and the temporal order of arrival of fibers on the tectum, but they do not in fact appear to make an important contribution (Holt, 1984; Stuermer, 1986).

3.2. Systems-based modeling work

Following the proposition of Sperry (1963), experiments designed to test the contribution of chemoaffinity to retinotopy were performed and were accompanied by the first formal quantitative models of map formation in this context. This generation of "systems level" experiments involved various surgical manipulations of both retina and tectum to test the chemoaffinity hypothesis and how robust the retinotectal system as a whole was to these perturbations. An illustrative example of the many experiments done in this vein are the retinal ablation (or map expansion) experiments. In one such example of these experiments (Schmidt *et al.*, 1978), half-retinal ablations were performed on goldfish eyes and the resulting map observed. In some cases, the remaining retinal fibers remained confined to their appropriate hemitectum, but it was found that they eventually shifted their connections so that the remaining fibers expanded their projection to fill the entire tectum (Fig. 1.6). These and other experiments (Goodhill and Xu, 2005; Udin and Fawcett, 1988) clearly demonstrated that while chemoaffinity is important in retinotectal mapping, there was significant plasticity in the map that could not be accounted for by chemoaffinity alone. Hence successful modeling relied on incorporating multiple mechanisms or constraints to recreate the experimental results.

3.2.1. Competition and types of affinity

One of the first computational studies on the formation of retinotectal map (Prestige and Willshaw, 1975) demonstrated that assumptions underlying chemoaffinity needed to be clarified, and introduced a classification of type I

and type II chemoaffinity mechanisms. The authors considered a mapping between a pair of one-dimensional arrays of cells; each projecting cell was allowed to make multiple contacts (a representation of branches) in the target cell array, and target cells were similarly each allowed to accept multiple contacts. Strengths of connections were expressed as lifetimes of the contact (cf. synaptic strength). Type I chemoaffinity (or ''rigid matching'') assumes that each projecting cell has maximum affinity for only one (or other small subset) of the target cell array, and each cell in the target similarly has maximum affinity for only a subset of projecting cells. Type II matching describes *graded* affinity, whereby all projecting cells have maximum affinity for one end of the target (e.g., rostral tectum), and all cells in the target have maximum affinity for one end of the projecting cell array. Type I mechanisms can certainly form an ordered map, but not as immediately clear was whether type II mechanisms could do the same, and under what conditions. In simulations the authors were able to show that type II mechanisms were indeed able to form a retinotopic map, but only if limits were placed on the number of contacts a projecting cell could form with target cells, and vice versa. It was observed that this constraint could be realized in, for example, competition for a target resource such as space (''competition by exclusion''). The authors acknowledged that while this model could form normal maps and certain systems manipulations, others (such as map expansion and compression) required further assumptions that were not well justified. Although the simulations were only carried out in one dimension, the ideas easily extend to two; as such, competition has been a common feature of retinotectal models from this time on.

3.2.2. Marker induction models

Type I matching has the advantage of a more straightforward implementation and does not require competition to form an ordered map, but is problematic in that it cannot alone account for the plasticity of maps demonstrated in systems manipulations. To account for these results in a type I scheme, another mechanism such as marker induction needs to be employed. It was this idea was explored computationally in a series of models (von der Malsburg and Willshaw, 1977; Willshaw, 2006; Willshaw and von der Malsburg, 1979) best described as marker induction models (aka Tea Trade model, retinal induction model). These models connected retinal axons to tectal cells by contacts or synapses, whose weights were modulated depending on the similarity of markers between the projecting cell and the receiving cell. New contacts could be formed near other contacts in areas more likely to yield successful connections, and competition was introduced by normalizing the synaptic output of RGCs. Induction occurs as markers either diffuse out of (von der Malsburg and Willshaw, 1977; Willshaw and von der Malsburg, 1979) or are upregulated by (Willshaw, 2006) RGC axon terminals; neighboring tectal cells can also

influence each other through this effect. Coupled differential equations were defined for changes in marker level and synaptic strength. Through this mechanism, markers on small regions of tectum become similar to markers on small regions of retina, which can lead to topography. Willshaw (2006) updated the model to include newer molecular data (see Section 3.3), and also measured map precision in terms of mean receptive field size and separation.

The marker induction model can be described as a type II chemoaffinity model, which also employs competition and branched axons. This was the first use of branching axons (albeit abstract ones) as a kind of pathfinding device in retinotectal modeling. The synaptic modulation approach used is similar to Hebbian models of activity-dependent synaptic plasticity, but uses similarity of markers instead of correlations in activity. (Indeed an activity-based version of the model subsequently formed the basis for the elastic net model (Durbin and Willshaw, 1987), which had comparatively greater success in influencing modeling cortical feature maps; this is discussed further in Section 4.) An important property of this model that was demonstrated in simulations (and has been found in subsequent models employing similar activity-based Hebbian rules) is that it cannot generate topography without some sort of overall polarity guide (such as a weak initial gradient of markers). Although some of the assumptions of the model, particularly details of the induction mechanism, require more experimental justification, these same assumptions give it a great deal of flexibility. It is able to explain a wide range of systems-based results (see, e.g., Fig. 1.7), and is also one of the few models that has also been applied to more recent molecular data (Willshaw, 2006).

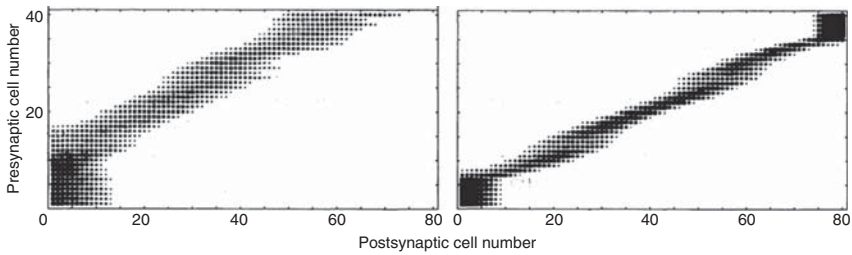

Figure 1.7 Simulation of retinal ablation experiments using the marker induction model. RGCs retinal position plotted against tectal position. Each dot represents a synapse between the corresponding presynaptic and postsynaptic cell, with area proportional to synaptic strength. The left panel shows the simulation early after retinal ablation, where the RGCs retain some of their original ordering prior to ablation. The panel on the right shows the final positions, replicating the map expansion results seen in experiments (from Willshaw and von der Malsburg, 1979; reproduced with permission from the Royal Society).

3.2.3. Sorting-based models

The Arrow model (Hope *et al.*, 1976) took a more abstract approach to modeling axon–axon interaction, treating it as a kind of sorting mechanism. An initially unordered array of RGC axon terminations on a tectal grid was considered. Pairs of neighboring array members were compared and their positions exchanged if they were incorrect with respect to both their retinal origin and orientation compared to local tectal polarity or ''arrow.'' This exchange process (a sort of type II affinity) could be iteratively alternated with a random walk step to vacant array sites, although this was not done in the simulations presented. The model was able to generate ordered maps and recreate some systems data, such as tectal graft rotation experiments, but not tectal graft translocation experiments.

To account for the remainder of the systems data Overton and Arbib (1982b) updated the Arrow model to XBAM (eXtended Branch Arrow Model), which in itself was an updated version of a previous model (BAM) by the same authors (Overton and Arbib, 1982a). There were a number of significant changes; firstly, the tectum was treated as continuous, and the Arrow model exchange rules converted to movement vectors induced in RGC axons by other nearby axons. The boundaries of the tectum as well as graft edges also contributed to the movement vector. A softened type I affinity rule was added to reflect results of translocation experiments, so that the model now employed a mixed type I/II affinity. Each axon consisted of a number of branches, and it was assumed that movements of branches from the same axon were allowed to influence each other through an averaging process (independent of their proximity to each other). XBAM successfully explained a wide range of plasticity results through its multiple constraint approach and mixed type I/II affinity. Its consideration of branched axons and motion over a continuous tectum were also quite advanced for its time in terms of realism.

3.2.4. Dual gradient chemotaxis models

A more explicit use of molecular marker ideas, and a prototypical type I affinity model, was proposed by Alfred Gierer in a series of models in the 1980s (Gierer, 1981, 1983, 1987). These models considered molecular gradients (which were largely speculative at the time) in the retina and tectum, which exerted their effects by producing a substance p within individual growth cones, which is in general a function of retinal origin and current tectal position. It was proposed that growing axons seek a maximum (or minimum) of p, by moving in the direction where the rate of change of p was maximal (or most negative). To achieve the required maxima and minima, dual gradients (generally in opposite directions) were required. These are sometimes known as ''countergradients,'' though this term has also been used to describe pairs of gradients oriented in opposite

directions, where one gradient is in the retina and the other is in the tectum. Acknowledging the inability of this type I mechanism to reproduce plasticity phenomena in the map, Gierer (1983) included a form of retinal induction ("regulation") and used this to explain some systems experiments (compression and expansion) but not others (e.g., translocations, rotations, etc.). Later versions of the model (Gierer, 1987) presented limited simulations of how branching and axonal trajectories might be expected to follow from *p*-guidance model (Fig. 1.8), but did not include regulation.

These models were unusually realistic in their consideration of responses to gradients (they are similar to more modern axon guidance models), branching, and trajectories. However, these concepts were not combined into a cohesive model in simulations of map development, so that only limited experimental data was replicated. Notably though, realistic branching patterns and trajectories were reproduced, which has been unusual in the modeling literature. The use of dual gradients represents a rigid matching approach, which is unable to account for many systems experiments,

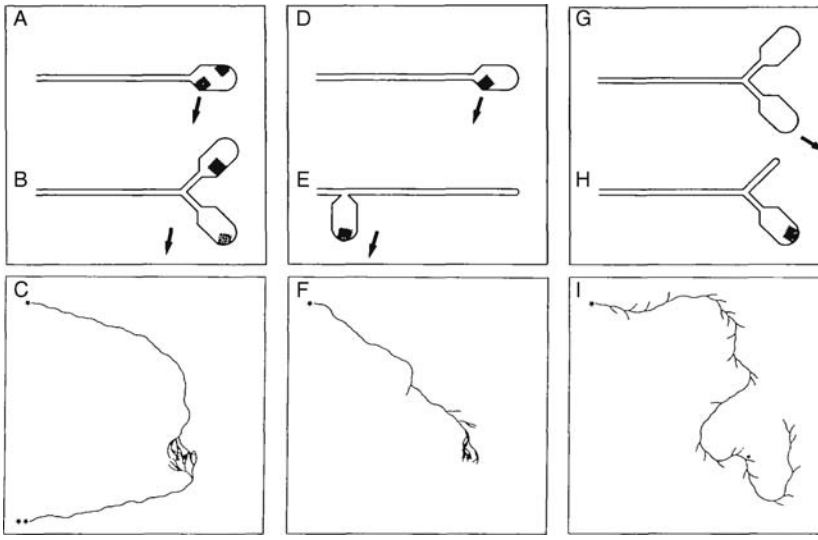

Figure 1.8 Gierer's branching and trajectory simulations. Top row: Examples of how branching might occur using the p-guidance model. Bottom row: simulations using these models. (A–C) If the growth cone encounters a negative slope (below a certain threshold), it becomes "proximally activated" and initiates "forward branching" (or growth cone bifurcation). In (C), two axons are simulated to show the different trajectories that can be observed from different starting points (filled stars) but the same target (filled circle). (D–F) Similar to the above, but proximal activation of the growth cone now causes axon growth to stop, and branching to occur behind the growth cone (backbranching). (G–I) Now branching occurs randomly, and lateral inhibition between branches causes the branch with the higher level of p to win, while the other branch stops growing (from Gierer, 1987; reproduced with permission from the Company of Biologists Ltd.).

despite the inclusion of a form of marker induction to remedy this. There are also constraints on the gradient shapes that can form maps under the proposed scheme; that is, ones for which $dp/dx = 0$ implies retinotopy (Goodhill, 1998). Despite these drawbacks, Gierer's models provide a relatively detailed and realistic framework, and their general assumptions form the primary example of what are known as *dual gradient* or *countergradient* models.

3.2.5. Adhesive energy minimization models

The idea of minimizing a certain quantity to form a map formed the basis of a series of models based around the adhesive properties of retinal and tectal cells (Fraser, 1980; Fraser and Perkel, 1990). Disks representing terminal arbors were arranged in an initially random configuration on the tectum. An "adhesive free energy" was calculated for each disk in its current location based on contributions from molecular markers, correlated activity, and competition (hence it is also known as Fraser's multiple constraint model). The disk was randomly moved to a new position, and if this resulted in a reduction in free energy, the move was accepted; otherwise further random movements were attempted. Through this approach, the map was evolved to stability.

Similarly to Gierer's model, gradients of molecular markers in a type I affinity scheme were used. In addition, two separate axon–axon interactions were modeled; one activity-based, and the other a chemospecific interaction whereby axons with similar levels of markers had a lower adhesive energy associated with the interaction (cf. retinal induction models). This latter property, in the context of this model, represents a type II mechanism. Hence, the model is a multiple constraint model with mixed type I/II affinity. It was able to successfully simulate a wide range of experimental results (Fig. 1.9), but the explanatory power of the model was reduced by the abstract nature of rules for movement and minimization.

3.2.6. Chemospecific synaptic modulation

A series of models developed by Cowan and colleagues were largely based on synaptic modulation rules, in which the change in strength of a synapse between a pair of retinal and tectal cells is considered to be proportional to some combination of chemospecific adhesiveness between them and the correlation in their activity (Cowan and Friedman, 1990, 1991; Weber *et al.*, 1997; Whitelaw and Cowan, 1981). Although models in this series had much in common, each focused on different elements in their implementations. Whitelaw and Cowan (1981) modeled the growth of the synapse as proportional to the product of chemospecific adhesion and correlated activity and used a type II chemoaffinity. Cowan and Friedman (1990, 1991) added axon–axon interactions and changed from type II to type I affinity. Weber *et al.* (1997) made the chemospecific and activity-based contributions independent, and added two axon–axon interactions in a multiple constraint

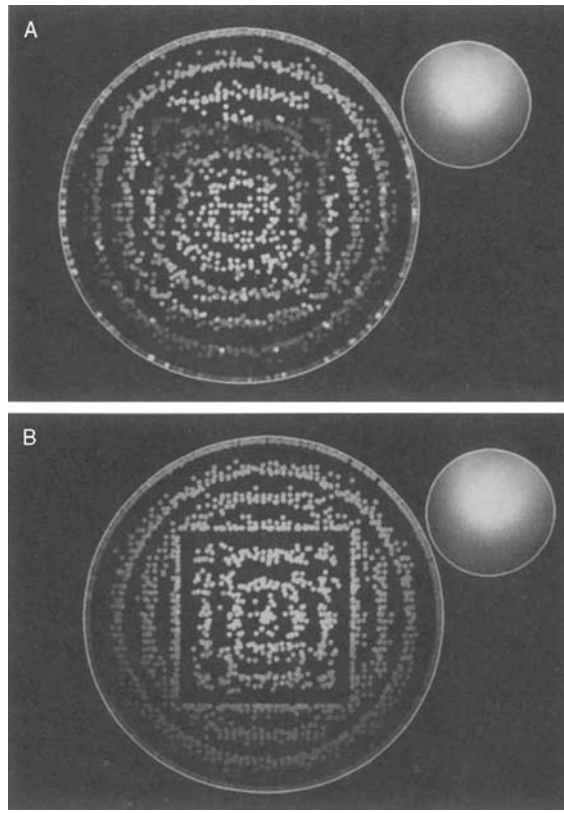

Figure 1.9 Simulation of tectal rotation experiments in the adhesive energy minimization model. The two smaller circles correspond to the retina, while the larger circles represent the tectum. Locations of RGC arbors are indicated by dots within the larger circle. In (A) a rotated map forms, following the rotated tectal cues, while in (B) a normal map results that ignores the rotated cues. The normal map tended to form when the degree of initial order was high (from Fraser and Perkel, 1990; reproduced with permission from John Wiley & Sons, Inc.).

framework, similar to the energy minimization models discussed above. These models, and in particular Weber *et al.* (1997), were quite detailed and successful in explaining experimental results, but were limited in realism by not having specific mechanisms for growth and movement of axons.

3.3. Chemoaffinity and the ephrins

A new wave of enthusiasm in experimental and theoretical exploration of the retinotectal map began when molecular effectors for Sperry's hypothesized gradients were discovered and characterized in the mid-1990s.

The primary example of these molecules are the Eph receptors and their ligands the ephrins (although other molecules may also contribute to some extent). Gradients of Eph receptors in the retina and ephrin ligands in the tectum are matched as suggested by Sperry, and were shown to be intimately involved in retinotopy. (For reviews of the contributions of the Eph–ephrin molecules to retinotopic mapping and many other areas of biology, see Flanagan and Vanderhaeghen (1998), Halloran and Wolman (2006), Kullander and Klein (2002), McLaughlin and O'Leary (2005), Pasquale (2005), Poliakov *et al.* (2004), and Wilkinson (2001).)

The discovery of the Eph and ephrin gradients led to a spate of sophisticated experiments to characterize their contribution to retinotectal mapping (a simple example is shown in Fig. 1.6), while new models were created and old ones modified to grapple with the data. We suggest that, ironically, this resulted in a step back in theoretical understanding, as all mechanisms except chemoaffinity were de-emphasized, and sometimes ignored completely. As a result, models of this generation generally have not replicated the systems-based results of the previous generation. Indeed certain features of the newer molecular-based experiments remain unexplained by these recent models as well, so that incorporating features from earlier models may be of benefit (e.g., see Goodhill, 2000; Wilkinson, 2000).

3.3.1. Servomechanism models

The servomechanism model (Honda, 1998, 2003, 2004) is a computational model based on the observations by Nakamoto *et al.* (1996). The model assumes that a signal S is generated in growth cones, which is a function of its fixed receptor level R (determined by retinal soma position) and local ligand level L (determined by current position on the tectum). This was expressed as $S = R \cdot L$, using mass action principles at equilibrium, and it was proposed that RGC axons seek out a standard signal S_0, by comparing the current signal S with the standard value. Starting at the rostral tectal border, axons were chosen at random to be moved in a probabilistic fashion to a new position. The chance of moving to a new position is higher if that position reduces the value $|S_0 - R \cdot L|$, and greater reductions in this value give greater probabilities of moving to the new location. This, combined with a general tendency for axons to move caudally and dorsally, allowed for topographic map development. The "set point" rule employed here is a type I affinity model in the mould of Gierer's models (Section 3.2.4), and is not easily able to recreate systems experiments. As expected, another mechanism was required to demonstrate plasticity in the mapping. Honda (2003) added competition in the form of a rule for moving axons away from more dense areas to less dense areas, until the density of terminations was uniform or below a critical density n_c (although this value was set to zero in Honda (2004) so that competition takes place everywhere). The axons that are chosen to be moved away from more dense areas to less dense areas are those

that have with the smallest $|S_0 - R \cdot L|$, so that this competition step also involves some sorting. This competition was computationally separate from the servomechanism phase; that is, first an initial mapping is made with the servomechanism rules, then the density of terminals is evened out with the competition rules. The model was then able to recreate some limited systems manipulations (Honda, 1998, 2004) and some ephrin misexpression data (Honda, 2003). Variations of the model have been considered by Löschinger *et al.* (2000) and Thivierge and Balaban (2007).

The servomechanism model is the only modern model to attempt to simulate some limited systems-based data in addition to the newer molecular data, and the only computational model that has been applied to stripe assay results (see, e.g., Fig. 1.10). Effectively, the servomechanism model involves a noisy gradient ascent algorithm (including a stop signal), followed by a smoothing/sorting algorithm (i.e., type I followed by type II affinity). This scheme, where an initial chemoaffinity phase is followed by a spreading

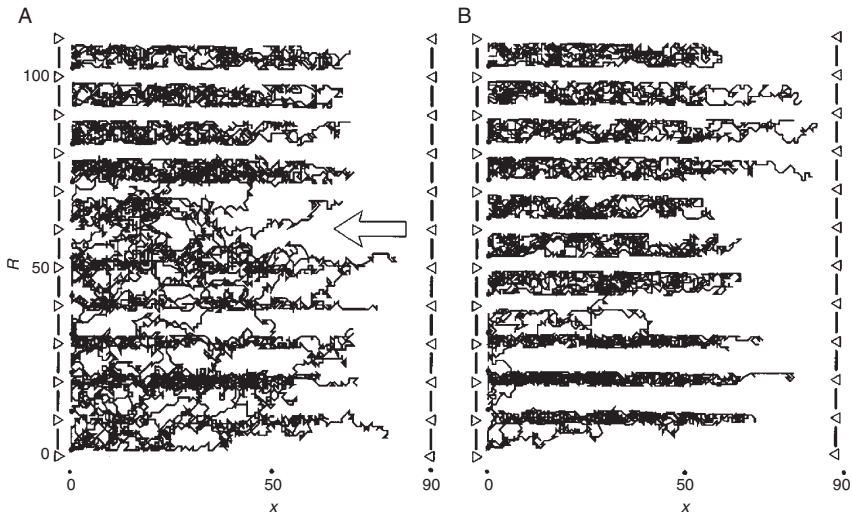

Figure 1.10 Servomechanism model simulations of a stripe assay. In this particular kind of stripe assay, alternating stripes of anterior and posterior tectal membranes are laid down, and axons from nasal to temporal retina are placed at one end (Walter *et al.*, 1987a, b). Here, anterior stripes are indicated by short vertical bars and posterior stripes are marked by open triangles. The vertical axis shows receptor level R, which corresponds to position of retinal origin, and the horizontal axis represents the distance x that axons have "grown" along the stripe. (A) Stripes composed of "crude" membrane fragments are simulated (i.e., differences in ligand level between stripes is small). (B) Stripes composed of fractionated membrane fragments (large difference in ligand between the two stripes). Experimentally a mix of striped and nonstriped growth is seen; usually from temporal and nasal axons, respectively. Sometimes sharp transitions from striped to nonstriped growth are seen, and a possible correlate of this in Honda's simulations is shown by the large open arrow (from Honda, 1998; reproduced with permission from Elsevier).

out of connections, was initially postulated by Prestige and Willshaw (1975). Although the use of type I affinity followed by competition gives the model some flexibility, this type of sequential process is somewhat artificial, and is not biologically well supported. There is some evidence for local maxima of adhesion, cell attachment, and/or outgrowth in gradients (Hansen *et al.*, 2004; Huynh-Do *et al.*, 1999), but it is not clear whether or how this translates into a "set point" rule (or stop signal) *per se*; at least not without further assumptions. Some of the movement rules, such as the tendency for dorsocaudal movement, also represent rather arbitrary assumptions. These issues aside, the model produces some detailed simulations which are comparable with a wide variety of experimental data.

3.3.2. Probabilistic sorting models

Koulakov and Tsigankov developed models involving probabilistic sorting of positions in a discrete array, similar to the sorting approach of the Arrow model, but now with explicit reference to Eph–ephrin gradients (Koulakov and Tsigankov, 2004), and later also including activity (Tsigankov and Koulakov, 2006). From initially unordered conditions, a randomly chosen pair of axons is considered and their positions exchanged if their current ordering is incorrect with respect to the direction of tectal gradient and relative retinal position. Although neighboring axons were considered in one-dimensional versions of the model, two-dimensional versions exchanged axons that were not necessarily neighboring. Differing probabilities are assigned to this exchange based on absolute differences in Eph/ephrin level, and whether the interaction is attractive or repulsive. Tsigankov and Koulakov (2006) included activity as well as this type II affinity, and a different algorithm was used for calculating exchanges. Dual gradients (countergradients) were included implicitly through functional inactivation or "masking" (Hornberger *et al.*, 1999). Repulsive and attractive gradients of Eph and ephrin molecules (in the rostrocaudal and mediolateral tectal axes, respectively), and correlated activity were assumed to contribute to a form of adhesive energy. Pairs of axons were again considered, and their positions exchanged if this reduced the adhesive energy. This equates to minimizing repulsion, while maximizing attraction and correlated activity, and hence bears resemblance to the adhesive energy minimization approach in Section 3.2.5. Calculations were presented to analytically predict features of the doubled and collapsed maps of Reber *et al.* (2004), making this one of the few models to present analytical arguments in addition to simulation data.

The use of type II affinity with implicit competition (and in the later version, activity) gives this two-dimensional model enough constraints to be applied to various experimental results, but so far only some ephrin misexpression results have been explored (Fig. 1.11). The authors used arguments based on noise in gradient-sensing and the spatial limits of activity-based mechanisms to explain different experimentally observed phenomena.

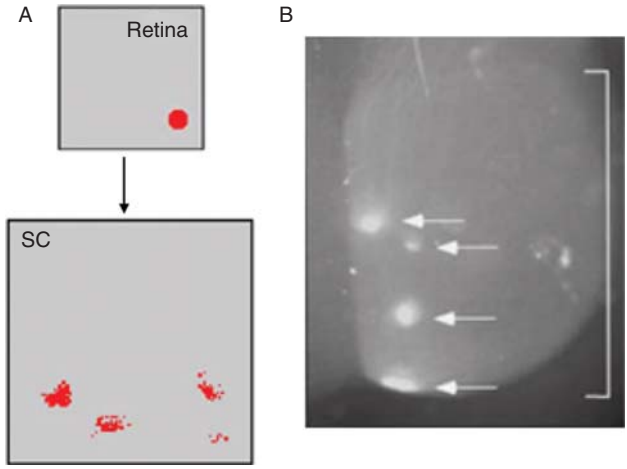

Figure 1.11 Simulations and experimental results in ephrin misexpression studies. (A) Simulated injection of tracer dye projects to the superior colliculus (SC) in the probabilistic sorting model (Tsigankov and Koulakov, 2006) applied to an ephrin misexpression study. (B) Corresponding experimental results for the ephrin misexpression study (Feldheim *et al.*, 2000) where focal retinal injections of DiI into the retina (similarly to the simulations shown in (A) but the experimental retinal injection is not shown) resulted in multiple termination zones (arrows) in knockout animals ((A) from Tsigankov and Koulakov, 2006; adapted with permission from Springer Science + Business Media. (B) from Feldheim *et al.*, 2000; reproduced with permission from Elsevier).

Although they are commonly used mathematical tools, minimization procedures in the context of retinotectal mapping have the problem that there does not appear to be a simple biological explanation for how this minimization is performed; in particular, how neural activity and molecular cues can both contribute to an energy that can be minimized by the growth cone. It is also not clear how actual movement results from this kind of scheme, nor how non-neighboring axons can interact and exchange positions in all circumstances (although the effect of this latter point was considered in Koulakov and Tsigankov (2004)). Despite these limitations in terms of biological realism, this probabilistic sorting model represents a concerted effort to engage with the ephrin misexpression literature, both in simulations and using more analytical methods.

3.3.3. Dual gradient branching model

A rather different approach was taken in the computational model of Yates *et al.* (2004) in which the focus was placed on biased branching of RGC axons in generating topography. This model followed the observations that branching along the length of an axon that initially overshoots its target

appeared to be particularly important in mammalian and avian topographic map development (Roskies and O'Leary, 1994; Yates *et al.*, 2001). The particular pattern of interstitial branching observed is regulated in part by Eph–ephrin interactions (McLaughlin and O'Leary, 2005; Rashid *et al.*, 2005; Simon and O'Leary, 1992; Yates *et al.*, 2001). In this model, dual gradients of Eph–ephrins were assumed to be present in both retina and tectum (a type I affinity model, with assumptions similar to Gierer's). Probabilistic rules were introduced which allowed branching to occur below a certain threshold of repulsion, and set branching to be maximal where repulsion was minimal. Additional rules allowed advancing, pausing, and retraction of branches. New branches contributed additional molecular markers to the tectal gradients (hence axon–axon interactions are incorporated implicitly), in some cases sharpening the gradients and therefore the resulting map. To obtain a better match of the behavior of the model to biology, an extra "branch density" term needed to be added to the model, which was time-varying and acted to increase branching probability in areas of already high branch density.

Simulations of ephrin misexpression studies were presented and are generally comparable to the results seen in experiment. However, this model does not attempt to address the inability of its type I assumptions to explain plasticity results (unlike Gierer's model which included regulation, and the servomechanism model which includes competition). One of the basic assumptions of the model, that branching is inhibited by repulsive markers on the tectum and from added branches, seems to be in contradiction with the assumption of the added "branch density" term, that branching is *promoted* by the presence of high densities of branches (i.e., high density of repulsive markers). It is also unclear how this time-varying extra term relates to the biology (e.g., activity-based mechanisms as suggested by the authors), how the specific choices of free parameters (e.g., threshold for branching) affect modeling outcomes, or whether the use of other types of gradients as suggested (e.g., dual attractive or mixed attractive and repulsive) could operate in the same way. However, this model was the first to show that branching alone, biased by gradients of molecular markers, is in principle capable of generating retinotopy.

3.3.4. Relative signaling models

Reber *et al.* (2004) presented a model based on quantitative descriptions of EphA misexpression data (Brown *et al.*, 2000; Reber *et al.*, 2004) showing that relative, rather than absolute, levels of EphA were important in mapping. This relative signaling was proposed to occur through local comparisons between RGC axons (a type II affinity), and it was assumed that a discrimination limit of relative signaling operated in this context. Calculations based on these concepts accurately predicted the position of map collapse (or lack thereof) in different phenotypes (Fig. 1.12).

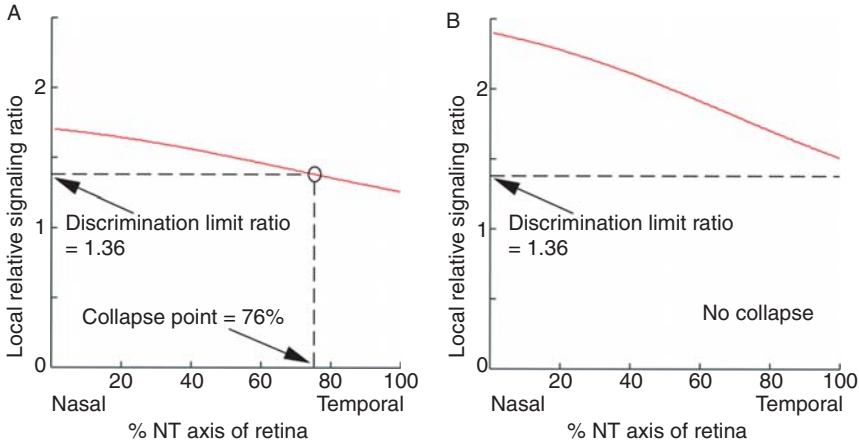

Figure 1.12 Relative signaling discrimination limit. In Brown *et al.* (2000) and Reber *et al.* (2004), the level of Eph receptor was artificially increased in half of all RGCs (distributed randomly throughout the retina). In some cases, the knockin RGCs display a second "doubled map" in addition to that formed by the wild-type RGCs. Here, the relative signaling ratio (the ratio of Eph receptor level on wild-type RGCs to receptor level of knockin RGCs) is plotted against retinal position. (A) The heterozygote case where the doubled map collapses to a single map at 76% of the map axis (giving a predicted relative signaling ration of 1.36). (B) The homozygous case shows no collapse of the doubled map using the same ratio, and this mirrors the experimental outcome in homozygotes (from Reber *et al.*, 2004; adapted with permission from Macmillan Publishers Ltd.).

To account for wild-type maps where the relative signaling ratio for neighboring RGCs is ≈1, and well below the calculated discrimination limit of 1.36, a mechanism of comparison with a universal reference level was invoked. Although no actual simulations of map development were performed, suggestions were presented for how relative signaling may be carried out in this context.

The relative signaling model replicates specific important results that other models have been largely unable to account for, particularly the presence or absence, and position of, map collapse points. Certain features of how this algorithm leads to mapping are problematic though, such as the idea of a comparison with a universal reference in wild-type maps, as this does not appear have a simple biological explanation. It is possible that including additional constraints, such as competition, could avoid the need for this assumption.

3.4. Summary of models

Table 1.1 summarizes the main types of mechanism used by the models discussed above.

Table 1.1 Model features summary table

Model	Version	Chemoaffinity	Branching	Competition	Activity	Induction
Competitive affinity	1975	Type I/II	–	Implicit	–	–
Marker induction	All	Type II	Exploratory	Explicit	–	Yes
Sorting-based: Arrow	1976	Type II	–	Implicit	–	–
Sorting-based: BAM	1982a	Type II	Averaging	Explicit	–	–
Sorting-based: XBAM	1982b	Mixed I/II	Averaging	Explicit	–	–
Dual gradient chemotaxis	1983	Type I	–	–	–	Yes
	1987	Type I	Exploratory	–	–	–
Adhesive energy minimization	All	Mixed I/II	–	Explicit	Yes	–
Chemospecific synaptic modulation	1981–1991	Type II	–	Explicit	Yes	–
	1997	Type I	–	Explicit	Yes	–
Servomechanism	1998	Type I	–	–	–	–
	2003, 2004	Mixed I/II	–	Explicit	–	–
Probabilistic sorting	2004	Type II	–	Implicit	–	–
	2006	Type II	–	Implicit	Yes	–
Dual gradient branching	2004	Type I	Exploratory	–	–	–
Relative signaling	2004	Type II	–	–	–	–

The main mechanisms used in models of retinotectal mapping, and the computational models that use them. Models are referred to by their main mechanism as described in the text. Different versions of the same model series are referred to by year; each corresponds to the citations in the main text. Implicit competition refers to models where competition is implied by the assumptions of the model (e.g., in 1–1 mappings) rather than described by explicit equations.

4. Activity-Dependent Development

Theoretical models for the development of activity-dependent wiring generally assume that an initial connectional architecture has been established by molecular cues, using the mechanisms discussed in the previous sections. The models then posit "learning rules" for how synaptic strengths change within this fixed architecture. Usually, these rules specify how particular combinations of presynaptic and postsynaptic activity lead to incremental increases or decreases in the synaptic strength, or "weight," of the connection. It is generally assumed that these changes lead to the organism becoming better adapted to its environment than if all synaptic strengths were fixed in advance. Historically this formalism has been shared with the field of artificial neural networks, which experienced a strong resurgence of interest in the 1980s (e.g., Hinton, 1989, 1992). Sometimes the goal of developing ANNs which can solve difficult practical learning problems in the fields of, for instance, perception, classification, and action selection has been taken as synonymous with understanding how nervous system wiring develops. While there has certainly been a rich flow of insight in both directions, here we will mostly focus on theoretical models which engage more directly with neuroscientific data (Dayan and Abbott, 2001).

4.1. Types of learning

From a computational point of view, it is common to distinguish three types of learning: supervised, reinforcement, and unsupervised learning. In *supervised* learning, the goal is to learn a mapping between given input and output vectors, for instance classifying the identity of faces in a set of images. This requires a way of propagating information about errors in the output for a given input through the network so that the strength of all relevant synapses can be changed appropriately. While over the past several decades powerful methods have been developed for such learning (e.g., Ackley *et al.*, 1985; Hinton, 2007; Minsky and Papert, 1969; Rumelhart and Zipser, 1986), the relevance of these to experimental data regarding neuronal wiring development is only at an abstract level. It is also hard to imagine how the detailed "teaching" signal that such algorithms require to provide error information could be present biologically. However, it remains possible that some of the computational principles underlying effective learning identified by these approaches might in the future find direct application for understanding biological wiring development.

In *reinforcement* learning, the goal is to learn a mapping between a set of inputs or actions in a particular environment and a (possibly temporally delayed) scalar-valued output representing some measure of "reward." Computationally this is much harder than supervised learning, since it

requires a way of assigning credit or blame to individual actions which may have occurred some time before the reward signal arrives. Good progress has been made on this problem computationally through the development of methods such as temporal difference learning (Sutton, 1988), and there is direct evidence that such algorithms are indeed employed in reward processing in the adult nervous system (e.g., Schultz *et al.*, 1997). The extent to which these algorithms can be used to understand nervous system development remains an intriguing unanswered question.

In *unsupervised* learning, the network is provided with no feedback at all. Rather, synaptic strength changes occur according to a learning rule based only on pre- and postsynaptic activity, with no explicit specification of the desired or "best" output for each input. The outcome (i.e., the pattern of synaptic strengths developed) depends on the nature of the learning rule and the statistical structure of the inputs presented. For example, the learning rule might imply finding clusters in the inputs, without specifying what meaning the clusters found should have for the animal. This approach can be very effective for understanding the development of neuronal wiring patterns at early stages of sensory processing, where it is reasonable to hypothesize that representations are driven more by just the statistical structure of the input data than the specific goals of the animal. Indeed, unsupervised learning methods have proven highly successful at reproducing the development of early stages of visual processing, as we discuss in detail below.

4.2. Linear Hebbian learning

A particularly fruitful hypothesis for how synaptic strengths change during both development and learning is the rule of Hebb (1949):

> When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.

A direct biological implementation of such "coincidence detection" emerged with the discovery of long-term potentiation (LTP), first in the hippocampus and subsequently in other brain regions including the cortex (Malenka and Bear, 2004). Qualitatively, Hebbian-type rules appear to explain many phenomena in neural wiring development; for instance the development and plasticity of ocular dominance columns in primary visual cortex (Katz and Shatz, 1996). Quantitatively, there are many different mathematical equations consistent with Hebb's qualitative statement. A large amount of research in models of neural wiring development has been devoted to understanding the outcomes of different types of Hebbian equations, and how to relate these outcomes to biological phenomena.

Although historically the first models were nonlinear, it is easier to explain some basic concepts in the context of linear models. We will first present in some detail a simple but highly illustrative example: a two-layer network with a single output neuron and a linear activation rule (e.g., see Dayan and Abbott, 2001; Hertz *et al.*, 1991).

In particular consider the simple network shown in Fig. 1.13. We assume the output activity y is given by

$$y = \sum_i w_i x_i, \tag{1.1}$$

where the w_is are the weights and the x_is are the inputs. This can be expressed more compactly in vector form as $y = \boldsymbol{w} \cdot \boldsymbol{x}$. In this model, the detailed temporal dynamics of neural activity are not considered; instead, the focus is on the timescale of weight changes, which is here assumed to be much slower. One of the simplest mathematical expressions of Hebb's rule is

$$\Delta \boldsymbol{w} \propto y\, \boldsymbol{x}, \tag{1.2}$$

which says that a small change in the weights occurs which is proportional to the product of the output activity and the input activity. A slightly more subtle and biologically realistic version (Sejnowski, 1977) is

$$\Delta \boldsymbol{w} \propto (y - \langle y \rangle)(\boldsymbol{x} - \langle \boldsymbol{x} \rangle),$$

where the angle brackets denote time averaging. In this equation, we have subtracted the average value from both the input and output activities. We can now combine this with Eq. (1.1), take a time average over all possible input patterns, and assume that the weights change slowly relative

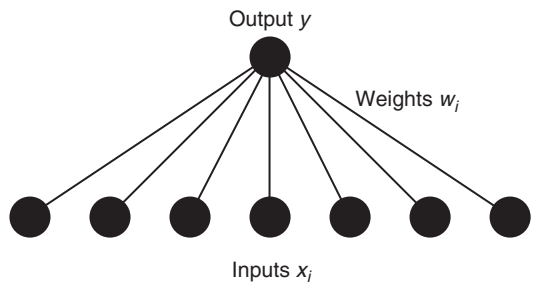

Figure 1.13 A simple neural network. Activity x_i in input neurons i is conveyed by fibers with synaptic weights w_i to an output neuron with activity y. A linear Hebbian learning rule for the weights w_i is sufficient to produce interesting behavior of the network.

to the rate at which input patterns are presented. That is, what matters for the weight dynamics is not the properties of any one individual input pattern, but rather the average statistical properties of the entire set of input patterns. This leads straightforwardly to the following equation:

$$\langle \dot{\boldsymbol{w}} \rangle = \alpha \mathbf{Q} \boldsymbol{w}, \tag{1.3}$$

where $\mathbf{Q}$ is the covariance matrix of the inputs.

This differential equation specifies temporal dynamics for the weights: how they change in response to the average correlations in the inputs. It is mathematically straightforward to immediately predict the outcome of this dynamics by imagining the weight development in the basis of the eigenvectors of $\mathbf{Q}$. Since the eigenvectors are the set of axes along which each component grows independently, with rate determined by the eigenvalue for each axis, eventually the eigenvector with the largest eigenvalue will dominate. That is, $\boldsymbol{w}$ tends to the principal eigenvector of $\mathbf{Q}$. Remarkably, this simple rule achieves a sophisticated computational goal. The principal eigenvector (also known as the principal component) of $\mathbf{Q}$ represents the direction along which the input data is most spread out in the sense of having maximum variance (e.g., Krzanowski, 1988). That is, the Hebbian rule above naturally tends to find patterns of weights which produce the biggest range in activity for the output neuron, given the statistical structure of the inputs presented. Even more remarkably, this learning rule also turns out to maximize the amount of information about the input that is conserved in the output (Linsker, 1990).

4.3. Constraining the weights

A problem with the simple Hebbian rule described above is that it is unstable: eventually all the weights go to infinity. One way to avoid this is to add extra terms to the right-hand side of Eq. (1.3). An example is Oja's rule (Oja, 1982, 1989), which again finds the principal component of the data but maintains the weights within bounds. A second method is to simply saturate the weights at upper and lower limits, which can be thought of as constraining the weights inside a box in the weight space. However, this makes it much harder to analyze the learning rule, since once the weight vector hits the walls of the box its behavior will be highly nonlinear (Feng *et al.*, 1997). A third method is weight normalization, which is motivated by the idea that the total strength of synaptic inputs that a postsynaptic neuron can support depends on some limited resource (reviewed in Miller, 1996). The most common way to implement this is to assume that either the sum of weights or the sum of squares of weights is maintained at a limiting value. The latter is equivalent to saying that the weight vector is constrained to be

of constant length. Whichever normalization constraint is used, there are two common ways to enforce it (Goodhill and Barrow, 1994; Miller and MacKay, 1994). In divisive (aka multiplicative) enforcement each weight is scaled proportional to its size, while in subtractive (aka additive) enforcement the same amount is subtracted from each weight to enforce the constraint.

4.4. Application of linear Hebbian learning to visual system development

Hebbian learning rules very closely related to those described above have been used to model both ocular dominance and orientation column development in the primary visual cortex. Linsker (1986) used a generalized form for Eq. (1.3) in a model consisting of a series of layers, with learning in each layer driven by correlations in the layer below. Neurons were assumed to have spatially localized receptive fields, and random activity (white noise) was applied to the input layer. This simple model was capable of producing several different types of receptive fields including all-excitatory, all-inhibitory, center-surround, and oriented. Furthermore, adding fixed Mexican-hat type lateral interactions between neurons in the final layer (discussed further below) induced the organization of the orientation-selective neurons into an orientation map (Fig. 1.14). Using an eigenvector analysis, MacKay and Miller (1990) subsequently demonstrated analytically how the type of receptive field formed in Linsker's model depends on the parameters of the model. Miller *et al.* (1989) applied a similar approach to the formation of ocular dominance columns (Fig. 1.15). Now each output neuron initially receives input from small regions of both the left and right eyes. It was found that the subtractive enforcement of a weight normalization constraint, and assuming zero or negative correlations in activity between the two eyes, was sufficient to drive receptive fields to become strongly dominant for only one eye or the other. Again adding lateral interactions between neurons caused the left- and right-eye dominant receptive fields to become clustered into columnar structures. This model was subsequently extended to the formation of orientation columns (Miller, 1994) and the joint formation of both ocular dominance and orientation columns (Erwin and Miller, 1998).

4.5. Nonlinear Hebbian learning

Learning in the models described above is essentially linear: the weights for all neurons are updated proportionally to their amount of activation. However, the first Hebbian models of the formation of visual maps were actually nonlinear. In particular von der Malsburg (1973) proposed a two-layer model, where the stimuli in the input layer consisted of oriented edges

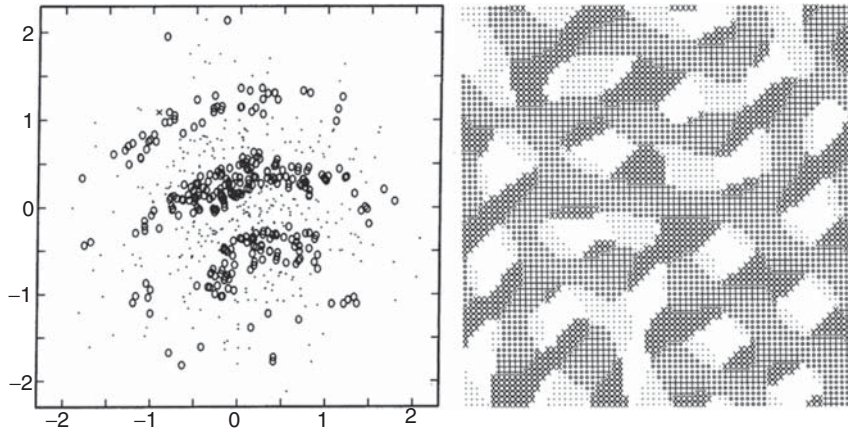

Figure 1.14 Results from the Linsker's (1986) model. Left panel: The receptive field of a V1 neuron generated from Linsker's model. Circles represent negative weights and dots represent positive weights. It can be seen that this cell is selective for both orientation and spatial frequency. Right panel: An orientation preference map generated by Linsker's model. Each different symbol corresponds to a different preferred orientation. Several pinwheels can be seen (from Linsker, 1986).

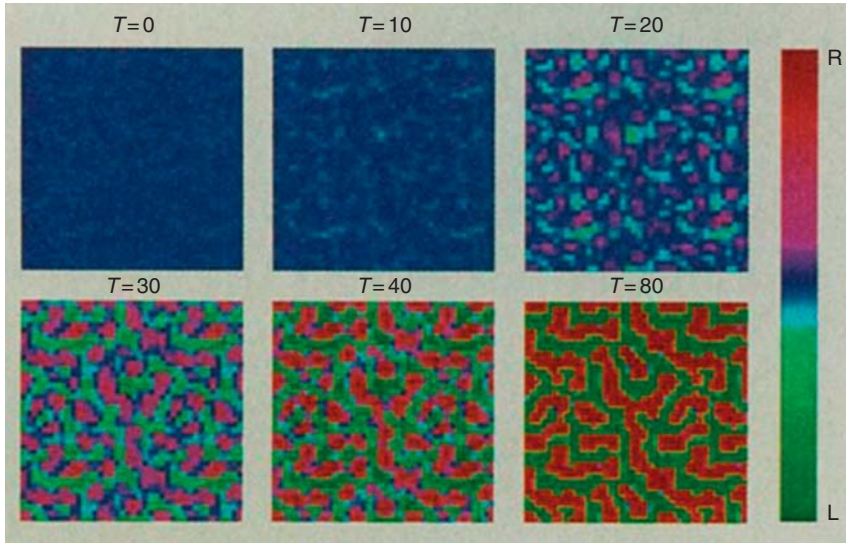

Figure 1.15 Time development of ocular dominance columns in the model of Miller *et al.* (1989). Segregation from an initially unsegregated state can be seen as development proceeds. Developmental time is measured in arbitrary units (from Miller *et al.*, 1989; reproduced with permission from AAAS).

or bars. The output (cortical) layer consisted of both excitatory and inhibitory neurons, with specific patterns of lateral connectivity between them. When an input pattern was presented to the network the input to each cortical neuron was calculated according to a rule similar to Eq. (1.1), but the neuron was only active if its total input exceeded a threshold. The response of each cortical neuron to the input was then iteratively recalculated, taking into account the additional inputs coming from the lateral connections. Once activity in the cortical layer had settled to a stable state the afferent weights were then updated by a Hebbian rule similar to Eq. (1.2). However, the crucial difference with the linear models described above is the nonlinearity of the activation rule: mathematical analysis is now much harder, and in particular the outcome cannot be predicted from the principal components of the correlation matrix of the inputs. However (and remarkably given the computing resources available at the time), von der Malsburg (1973) was able to show by simulation that this model produces oriented receptive fields in the cortical layer, organized in a map-like structure (Fig. 1.16). A very similar approach was then applied to the formation of topography (Willshaw and von der Malsburg, 1976) and ocular dominance columns (von der Malsburg and Willshaw, 1976).

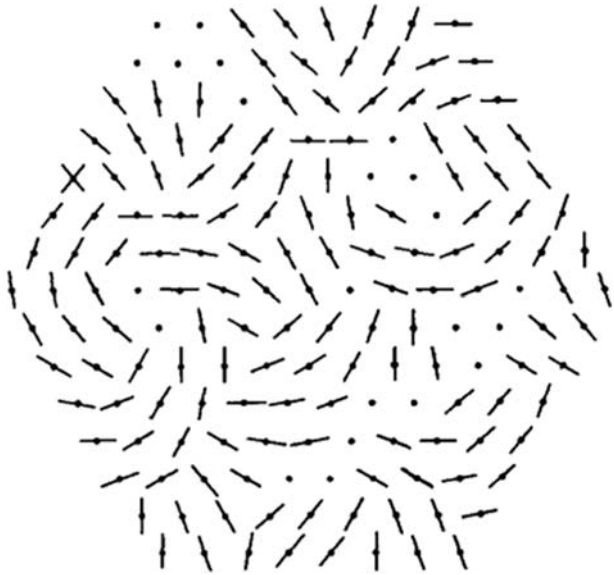

Figure 1.16 Orientation preference map generated from the model of von der Malsburg (1973). This was the first such map produced by a computational model, predating by a decade the detection of such overall map structure experimentally (from von der Malsburg, 1973; reproduced with permission from Springer Science + Business Media).

Another important nonlinear learning rule is that of Bienenstock *et al.* (1982). This has a form similar to that of Eq. (1.2), except that the size of the weight change is also scaled by a factor comparing the current activity of the neuron with a threshold activity θ:

$$\Delta \boldsymbol{w} \propto \gamma \boldsymbol{x}(\gamma - \theta).$$

This allows positive and negative changes in the weights, and adapting θ appropriately allows the learning to be stable. This rule can produce a variety of receptive fields matching those seen in reality, and direct experimental evidence for such a sliding threshold has subsequently been obtained (reviewed in Bear, 2003).

4.6. Competitive learning

In models such as that of von der Malsburg (1973), the presentation of an input pattern leads to the gradual emergence of one or a few hotspots of activity in the output layer, as the activity is iterated through the lateral connections. This iteration process is time consuming, and convergence requires the careful control of several parameters. Kohonen (1982) proposed a drastic but effective simplification of this approach, whereby it is assumed *a priori* that the pattern of activity in the output layer that eventually results from each input pattern will be a "hump" centered on the "winning" neuron; that is, the neuron which received the largest feedforward activation initially. Standard Hebbian learning is then performed, which amounts to adapting the feedforward weights of each neuron as a function of their distance in the output layer from the winning neuron. This is a computationally efficient and highly robust method for forming topographic maps, and has found wide application both in modeling map formation in the brain (discussed below) and in technological applications (Kohonen, 1995).

An alternative way of viewing Kohonen's algorithm is as an extension of another nonlinear approach to Hebbian learning, competitive learning (Rumelhart and Zipser, 1986). In competitive learning, inputs are presented to a layer of output neurons which are assumed to laterally inhibit each other, so that the neuron with the largest initial activity suppresses the activity of all other output neurons. The weights of this "winning" neuron are then updated as usual. This type of algorithm is highly effective for finding clusters in the input data: each input pattern is effectively assigned to one neuron, whose weights then provide the exemplar for that type of input. Barrow (1987) used this approach to model the development of oriented receptive fields in visual cortex. The activity γ of each neuron was calculated as in Eq. (1.1), but then only the neuron with the largest γ value had its weights updated according to Eq. (1.2). When given fragments of natural images as input patterns, this model produced oriented receptive

fields resembling those in V1. Orientation map development using a competitive (Kohonen) approach was modeled by Obermayer *et al.* (1990), and a related competitive model was proposed for the development of ocular dominance columns by Goodhill (1993) (Fig. 1.17).

The competitive approaches described so far are often described as "hard," in the sense of being "winner-take-all." In "soft" competitive learning, *all* neurons in the output layer are updated, by an amount that takes into account both their feedforward activation and the activity of other output neurons. An example is the elastic net, originally developed as a method for solving the *Travelling Salesman Problem* (Durbin and Willshaw, 1987). Here, each input pattern is assumed to have the same amount of efficacy, but this is shared out in a nonlinear way between the output neurons. The elastic net was subsequently shown to be a highly effective tool for modeling the development of ocular dominance (Goodhill and Willshaw, 1990) and orientation (Durbin and Mitchison, 1990) maps in V1. Together with related approaches based on the Kohonen algorithm (Obermayer *et al.*, 1992), these competitive methods have not been surpassed in their ability to predict the fine details of the overall geometric structure of feature maps in V1, for both normal and abnormal development (Carreira-Perpiñán and Goodhill, 2004; Carreira-Perpiñán *et al.*, 2005; Erwin *et al.*, 1995; Farley *et al.*, 2007; Giacomantonio and Goodhill,

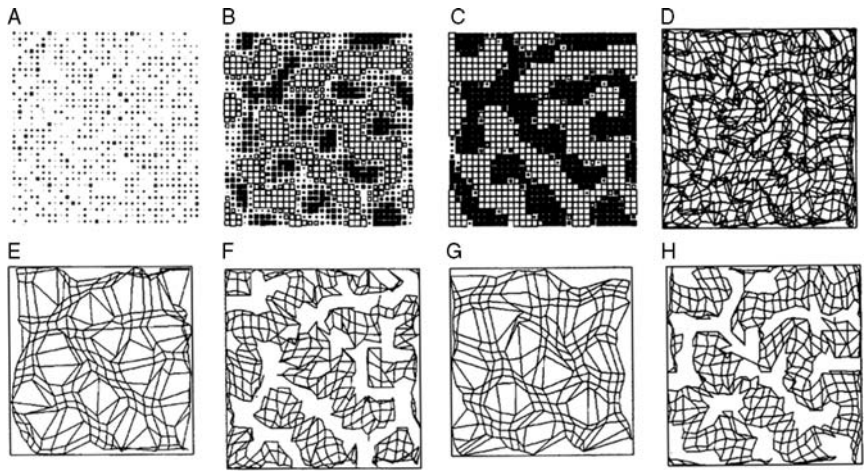

Figure 1.17 Simultaneous formation of ocular dominance columns and a topographic map in the model of Goodhill (1993). (A)–(C) A progression of time points, where the size of each square represents its degree of selectivity. (D)–(H) The center of mass of the weights of each cortical or retinal neuron is represented as a point in the retinal or cortical space, respectively, and neighboring lines are connected to form a grid. (D) Representation of cortical topography for both eyes. (E) and (G) Representation of retinal topography for the right and left eyes, respectively. (F) and (H) Representation of cortical topography for right and left eye, respectively (from Goodhill, 1993; reproduced with permission from Springer Science + Business Media).

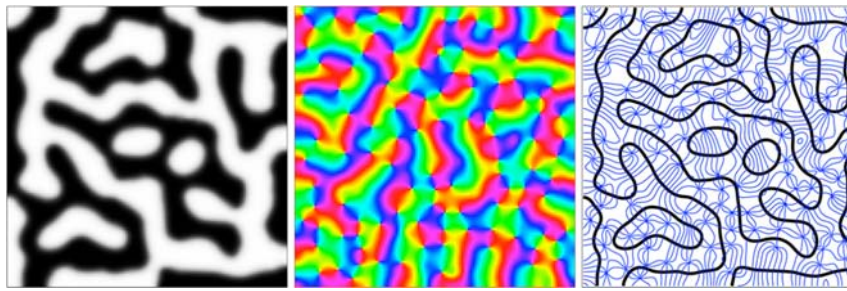

Figure 1.18 Elastic net simulations of visual map development. Left panel: Ocular dominance map. Middle panel: Orientation preference map. Right panel: Superimposed contours of ocular dominance and orientation. Note the tendency to intersect at steep angles, and for pinwheels to lie at the center of ocular dominance columns (see Carreira-Perpiñán *et al.*, 2005 for more details).

2007; Goodhill, 2007; Goodhill and Cimponeriu, 2000; Goodhill *et al.*, 1997; Swindale, 1996; Yu *et al.*, 2005). Figure 1.18 shows ocular dominance and orientation preference maps generated by the elastic net simulations in Carreira-Perpiñán *et al.* (2005).

4.7. Spike-timing-dependent plasticity

All the mathematical instantiations of Hebb's rule we have considered so far were developed prior to the general realization that the relative timing of the presynaptic and postsynaptic spike can crucially affect whether synapses increase or decrease—so-called spike-timing-dependent plasticity (STDP) (Bi and Poo, 1998; Markram *et al.*, 1997; reviewed in Dan and Poo, 2006). While mathematical models that do not address fine-scale timing issues in synaptic plasticity can still be extremely useful for addressing many problems (see for instance the examples above), the inclusion of spike timing opens up a large array of new computational issues to consider (Brette *et al.*, 2007; Burkitt *et al.*, 2004; Kepecs *et al.*, 2002; Morrison *et al.*, 2008).

A relatively simple mathematical model of such a learning rule was proposed by Song *et al.* (2000), which can be expressed in terms of a function $F(t)$ which specifies the amount of increase or decrease in synaptic strength when pre- and postsynaptic spikes occur at times t_{pre} and t_{post}, respectively:

$$F(\Delta t) = \begin{array}{ll} +A_+ e^{\Delta t/\tau_+}, & \text{if} \quad \Delta t < 0, \\ -A_- e^{-\Delta t/\tau_-}, & \text{if} \quad \Delta t > 0, \end{array}$$

where $\Delta t = t_{\text{pre}} - t_{\text{post}}$, and A_+ and A_- are positive scaling factors. In the completely symmetric case $A_+ = A_-$ and $\tau_+ = \tau_-$. A variety of asymmetries

are now known experimentally (Dan and Poo, 2006), many of which can be captured at least roughly by varying the ratios A_+/A_- and τ_+/τ_-. This rule is innately competitive, meaning that the rather unphysiological devices such as weight normalization required in many of the models discussed previously are no longer necessary to achieve competition. It also has the properties of causing postsynaptic neurons to reduce their latency, and tending to make weights go to their maximum or minimum values—a bimodal distribution. (In another version of this rule, where the size of the weight change depends on the current synaptic strength, a unimodal distribution of weights results; discussed further in Morrison *et al.* (2008).) There are also complex issues involved in deciding exactly which postsynaptic spikes the rule applies to; two extremes being all postsynaptic spikes, and just the one closest to the presynaptic spike.

Song and Abbott (2001) used the rule of Song *et al.* (2000) to model cortical map formation, including the development of ocular dominance columns. In this model lateral as well as feedforward weights were adapted by the same STDP rule. A key advantage of this approach over the non-STDP models discussed above is that maps can reach a stable state, but are then still free to change their structure if the input statistics change. In particular, Song and Abbott (2001) showed that "lesioning" some of the inputs in the adult state causes appropriate map rearrangement. Subsequent work by Young *et al.* (2007) showed directly that STDP but not more standard Hebbian learning rules were capable of reproducing a particular type of postlesion plasticity observed in cat visual cortex. Both these authors and Song and Abbott (2001) highlighted the "didactic" properties of STDP learning rules: if neuron A spikes before neuron B and they are reciprocally connected, then STDP tends to cause neuron B to develop similar response properties to neuron A. STDP learning rules have also been applied to, for instance, the development of directionally selective cells in primary visual cortex (Wenisch *et al.*, 2005) (Fig. 1.19).

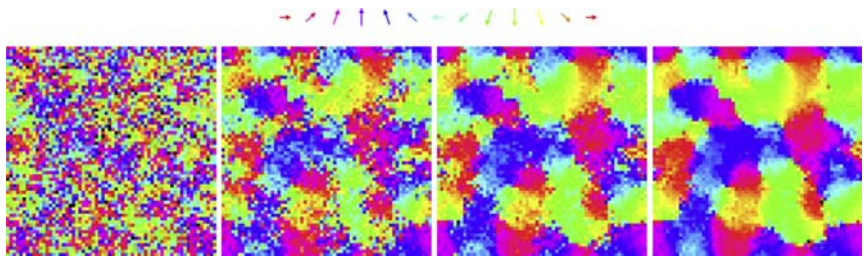

Figure 1.19 Development of a direction preference map in the STDP-based model of Wenisch *et al.* (2005). Sequence of map development at four different time points. Colors corresponding to different direction preferences are shown above the simulations (from Wenisch *et al.*, 2005; reproduced with permission from Springer Science + Business Media).

4.8. Functional modeling

A quite different theoretical approach to neural wiring development is functional modeling. The methods we have discussed up to now generally start with equations intended to capture at least something of the biology of Hebbian synaptic plasticity. By contrast in functional models, the focus is on the underlying computational goal of the developmental process. In particular, functional models generally propose that the purpose of development (at least in certain contexts) is to optimize some information-theoretic measure of the ''quality'' of the representation (for discussions, see Linsker, 1990; Simoncelli, 2003). Although in some cases the optimization process has a Hebbian interpretation, this is less important for these approaches than a good match between real and simulated receptive fields.

One popular approach in this class is independent component analysis (ICA) (Bell and Sejnowski, 1997; Caywood *et al.*, 2004; Doi *et al.*, 2003; Hsu and Dayan, 2007; Hyvärinen and Hoyer, 2001; Hyvärinen and Köster, 2007; van Hateren and Ruderman, 1998; van Hateren and van der Schaaf, 1998). We saw earlier how a linear Hebbian rule can lead to weight distributions which find the principal components of the input patterns. Rather than maximizing the variance as in PCA, in ICA the goal is to find weight distributions which maximize the statistical independence of the different components. Intuitively, the idea is that sensory inputs (e.g., images) are produced by summing several independent causes (e.g., some basis functions, such as Gabor patches), so that by maximizing independence one can hopefully recover the underlying causes of the input.

A closely related approach, motivated by energetic considerations, is to maximize the sparseness of the representation, that is, minimize the number of neurons that fire in response to any particular input (e.g., Olshausen and Field, 2004). A seminal example of this ''sparse-coding'' approach is the work of Olshausen and Field (1996). They trained a set of simulated V1 neurons, using a learning rule which maximized response sparseness, on natural images (somewhat analogously to the work of Barrow, 1987 in competitive learning mentioned earlier). They recovered a set of receptive fields that were both orientation selective and localized in space, in particular Gabor-like (see Fig. 1.20). For a much more extensive discussion of these types of approaches to learning see Chapter 10 of Dayan and Abbott (2001).

5. Discussion

Our main conclusions are as follows:

(1) **Axon guidance**. A complete model of axon guidance will ultimately require us to understand the detailed biophysical mechanisms

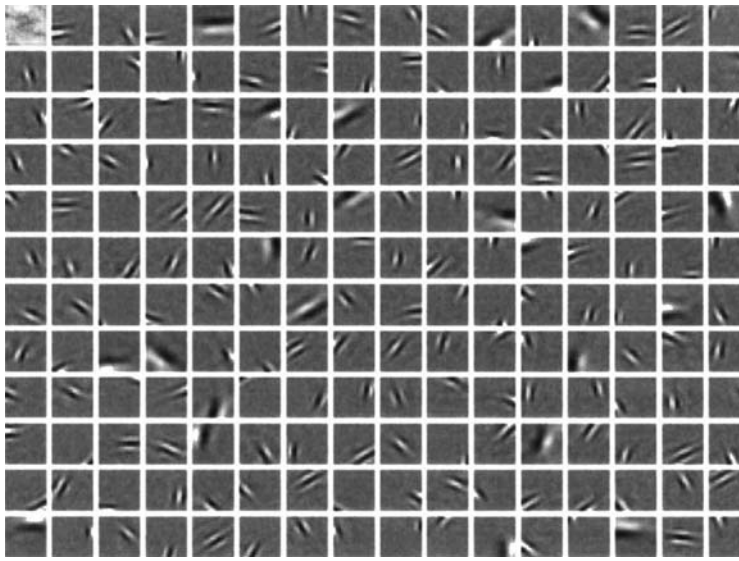

Figure 1.20 Receptive fields of simulated V1 neurons. One hundred and ninety-two basis functions trained on fragments of natural scenes using a sparse-coding algorithm (Olshausen and Field, 1996). These basis functions resemble the receptive fields of neurons in primary visual cortex in that they are localized, oriented, Gabor-like, and display a range of spatial frequency. In this case no topographic structure was imposed on the map (from Olshausen and Field, 1996; reproduced with permission from Macmillan Publishers Ltd.).

underlying growth cone behavior, which will in turn require access to quantitative data to constrain models. Modeling axon guidance is still in its early stages, perhaps because these mechanisms have not yet been well described experimentally (compared with, for example, bacterial chemotaxis), and because quantitative kinetic data is still lacking. Despite these challenges, progress has been made at the broad level of understanding physical constraints faced by growth cones (e.g., in sensing shallow gradients, and filopodial protrusion).

(2) **Retinotectal mapping**. Models in retinotectal mapping have established a number of mechanisms that are capable of producing retinotopic maps, and have described how they can do so. However, in general models have tended to be abstract rather than mechanistic, and perhaps because of this fail to explain important *in vitro* and *in vivo* results. Addressing this should be one focus of future modeling work. Another is to meet challenges presented by new data, which suggest that assumptions regarding known mechanisms (such as competition, chemoaffinity, and activity) and the interactions between them may need to be revisited (Cang *et al.*, 2008a; Gosse *et al.*, 2008; Hua *et al.*, 2005; Nicol *et al.*, 2007).

(3) **Activity-dependent development**. The unfolding of the genetic program alone does not, of course, fully determine nervous system structure. We have seen that by tying theory closely with results from neuroimaging and neurophysiology, a relatively simple set of assumptions about activity in the developing visual system can be used reproduce important features of visual maps seen *in vivo*. Hence by modeling the influence of activity in the development of neural circuitry, we can obtain a quantitative understanding of the influence of the environment on brain wiring. One challenge for future modeling in this field is to use these ideas to tease apart the relative importance of environmentally driven and intrinsic mechanisms for particular problems (e.g., whether certain properties of visual maps are under genetic control or are due to higher order statistics of natural scenes).

We have canvassed some important features of these three fields individually, but how do they relate to each other? We can think of retinotectal mapping as an applied or population version of axon guidance models, in that it should resemble models of individual axon guidance in the limit of small numbers of axons. In turn, activity-dependent processes depend on initial connections being set up by processes such as those in retinotectal mapping. So these areas are certainly related, yet they are generally treated as being quite separate. One reason may be that traditional reductionist modeling usually involves choosing just one particular time and length scale that is representative for the particular problem, and working within that. However, the development of brain wiring, like many biological processes of interest, occurs in regimes where multiple time and length scales interact. As a result, traditional approaches lead to fragmenting of models of these processes, making them appear more distinct than they actually are. For this reason, and almost certainly for many other reasons, biology is difficult to describe in as simple terms as the Ideal Gas Law discussed in the introduction. It may be that unifying principles for biology take on a different form than the laws of physics, so that different approaches may be necessary to solve these problems. For instance, ideal observer analysis (Körding, 2007) and wire minimization approaches (Chklovskii *et al.*, 2002) have proven useful and may drive further progress in understanding neural development. These are examples of normative methods, which seek to find the optimal solution to a given computational task or goal, given certain constraints (i.e., what computational strategy the organism should take to achieve a desired outcome).

There are other challenges for modelers besides the multiscale nature of problems in neural development. Generalizing models of more specific systems (e.g., modeling in the visual system) is not straightforward, but may inform us about brain wiring in general. We have already suggested that a lack of quantitative data is a barrier both to furthering modeling in

individual areas, and for marrying together apparently disparate fields. Better communication between experimental and theoretical disciplines will not only facilitate the gathering of this kind of data, but it will also maximize their synergy. There are many reasons why such synergy would be of mutual benefit. Theoretical approaches can allow us to identify hidden assumptions, rapidly generate and test hypotheses, and make accurate predictions; however, modeling is ultimately a solipsistic activity without experimental input and verification. Similarly, the kinds of questions one can ask experimentally are driven strongly by the theoretical framework within which one poses those questions. We suggest that in the future, understanding in neuroscience will not be driven by either field independently, but by thoughtful and respectful collaboration. In this respect, we are motivated by the great success achieved by similar modeling work in eukaryotic and bacterial cell chemotaxis (Paliwal *et al.*, 2007; Tindall *et al.*, 2008a,b), morphogen gradients (Bollenbach *et al.*, 2008; Gregor *et al.*, 2007a,b), gene regulatory networks (Jaeger *et al.*, 2004), and a range of other areas (Mogilner *et al.*, 2006). We look forward to similar success in the field of neural wiring development, as detailed and realistic mathematical models are vital to exploring exciting potential applications in the emerging fields of bioengineering and regenerative medicine; for example, the creation of organic self-wiring computers and bioelectrical interfaces, or guidance and rewiring in cell therapies.

REFERENCES

Abbott, L. F. (2008). Theoretical neuroscience rising. *Neuron* **60,** 489–495.

Ackley, D., Hinton, G., and Sejnowski, T. (1985). A learning algorithm for Boltzmann machines. *Cogn. Sci.* **9,** 147–169.

Aeschlimann, M., and Tettoni, L. (2001). Biophysical model of axonal pathfinding. *Neurocomputing* **38,** 87–92.

Barrow, H. (1987). Learning receptive fields. *Proc. IEEE 1st Annu. Conf. Neural Networks* **IV,** 115–121.

Bear, M. F. (2003). Bidirectional synaptic plasticity: From theory to reality. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358,** 649–655.

Bell, A. J., and Sejnowski, T. J. (1997). The ''independent components'' of natural scenes are edge filters. *Vision Res.* **37,** 3327–3338.

Berg, H. C., and Purcell, E. M. (1977). Physics of chemoreception. *Biophys. J.* **20,** 193–219.

Bi, G. Q., and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18,** 10464–10472.

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

Bollenbach, T., Pantazis, P., Kicheva, A., Bökel, C., González-Gaitán, M., and Jülicher, F. (2008). Precision of the Dpp gradient. *Development* **135,** 1137–1146.

Brette, R., Rudolph, M., Carnevale, T., Hines, M., Beeman, D., Bower, J. M., Diesmann, M., Morrison, A., Goodman, P. H., Harris, F. C., Zirpe, M., Natschlgër, T., *et al.* (2007). Simulation of networks of spiking neurons: A review of tools and strategies. *J. Comput. Neurosci.* **23,** 349–398.

Brown, A., Yates, P. A., Burrola, P., Ortuño, D., Vaidya, A., Jessell, T. M., Pfaff, S. L., O'Leary, D. D., and Lemke, G. (2000). Topographic mapping from the retina to the midbrain is controlled by relative but not absolute levels of EphA receptor signaling. *Cell* **102,** 77–88.

Buettner, H. M. (1995). Computer simulation of nerve growth cone filopodial dynamics for visualization and analysis. *Cell Motil. Cytoskel.* **32,** 187–204.

Buettner, H. (1996). Analysis of cell-target encounter by random filopodial projections. *AIChE J.* **42,** 1127–1138.

Buettner, H. M., Pittman, R. N., and Ivins, J. K. (1994). A model of neurite extension across regions of nonpermissive substrate: Simulations based on experimental measurement of growth cone motility and filopodial dynamics. *Dev. Biol.* **163,** 407–422.

Burkitt, A. N., Meffin, H., and Grayden, D. B. (2004). Spike-timing-dependent plasticity: The relationship to rate-based learning for models with weight dynamics determined by a stable fixed point. *Neural Comput.* **16,** 885–940.

Cang, J., Niell, C. M., Liu, X., Pfeiffenberger, C., Feldheim, D. A., and Stryker, M. P. (2008a). Selective disruption of one Cartesian axis of cortical maps and receptive fields by deficiency in ephrin-As and structured activity. *Neuron* **57,** 511–523.

Cang, J., Wang, L., Stryker, M. P., and Feldheim, D. A. (2008b). Roles of ephrin-As and structured activity in the development of functional maps in the superior colliculus. *J. Neurosci.* **28,** 11015–11023.

Carreira-Perpiñán, M., and Goodhill, G. (2004). Influence of lateral connections on the structure of cortical maps. *J. Neurophysiol.* **92,** 2947–2959.

Carreira-Perpiñán, M., Lister, R., and Goodhill, G. (2005). A computational model for the development of multiple maps in primary visual cortex. *Cereb. Cortex* **15,** 1222–1233.

Caywood, M. S., Willmore, B., and Tolhurst, D. J. (2004). Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning. *J. Neurophysiol.* **91,** 2859–2873.

Chklovskii, D. B., Schikorski, T., and Stevens, C. F. (2002). Wiring optimization in cortical circuits. *Neuron* **34,** 341–347.

Cowan, J., and Friedman, A. (1990). Development and regeneration of eye–brain maps: A computational model. *In* "Advances in Neural Information Processing Systems II" (D.S. Touretzky, Ed.), pp. 92–99. Morgan Kaufman, San Mateo, CA.

Cowan, J., and Friedman, A. (1991). Studies of a model for the development and regeneration of eye–brain maps. *In* "Advances in Neural Information Processing Systems III" (D.S. Touretzky, Ed.), pp. 3–10. Morgan Kaufman, San Mateo, CA.

Dan, Y., and Poo, M. M. (2006). Spike timing-dependent plasticity: From synapse to perception. *Physiol. Rev.* **86,** 1033–1048.

Dayan, P., and Abbott, L. (2001). "Theoretical Neuroscience" MIT Press, Cambridge, MA.

Debski, E. A., and Cline, H. T. (2002). Activity-dependent mapping in the retinotectal projection. *Curr. Opin. Neurobiol.* **12,** 93–99.

Doi, E., Inui, T., Lee, T. W., Wachtler, T., and Sejnowski, T. J. (2003). Spatiochromatic receptive field properties derived from information-theoretic analyses of cone mosaic responses to natural scenes. *Neural Comput.* **15,** 397–417.

Durbin, R., and Mitchison, G. (1990). A dimension reduction framework for understanding cortical maps. *Nature* **343,** 644–647.

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

Erwin, E., and Miller, K. D. (1998). Correlation-based development of ocularly matched orientation and ocular dominance maps: Determination of required input activities. *J. Neurosci.* **18,** 9870–9895.

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

Farley, B. J., Yu, H., Jin, D. Z., and Sur, M. (2007). Alteration of visual input results in a coordinated reorganization of multiple visual cortex maps. *J. Neurosci.* **27,** 10299–10310.

Feldheim, D. A., Kim, Y. I., Bergemann, A. D., Frisén, J., Barbacid, M., and Flanagan, J. G. (2000). Genetic analysis of ephrin-A2 and ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping. *Neuron* **25,** 563–574.

Feng, J., Pan, H., and Roychowdhury, V. P. (1997). A rigorous analysis of Linsker's unsupervised Hebbian learning. *Neural Networks* **10,** 705–720.

Flanagan, J. G., and Vanderhaeghen, P. (1998). The ephrins and Eph receptors in neural development. *Annu. Rev. Neurosci.* **21,** 309–345.

Fraser, S. E. (1980). Differential adhesion approach to the patterning of nerve connections. *Dev. Biol.* **79,** 453–464.

Fraser, S. E., and Perkel, D. H. (1990). Competitive and positional cues in the patterning of nerve connections. *J. Neurobiol.* **21,** 51–72.

Fujisawa, H., Tani, N., Watanabe, K., and Ibata, Y. (1982). Branching of regenerating retinal axons and preferential selection of appropriate branches for specific neuronal connection in the newt. *Dev. Biol.* **90,** 43–57.

Gaze, R. M., and Keating, M. J. (1972). The visual system and "neuronal specificity" *Nature* **237,** 375–378.

Gaze, R. M., Keating, M. J., and Chung, S. H. (1974). The evolution of the retinotectal map during development in Xenopus. *Proc. R. Soc. Lond. B Biol. Sci.* **185,** 301–330.

Giacomantonio, C. E., and Goodhill, G. J. (2007). The effect of angioscotomas on map structure in primary visual cortex. *J. Neurosci.* **27,** 4935–4946.

Gierer, A. (1981). Development of projections between areas of the nervous system. *Biol. Cybern.* **42,** 69–78.

Gierer, A. (1983). Model for the retino-tectal projection. *Proc. R. Soc. Lond. B Biol. Sci.* **218,** 77–93.

Gierer, A. (1987). Directional cues for growing axons forming the retinotectal projection. *Development* **101**(3), 479–489.

Giniger, E. (2002). How do Rho family GTPases direct axon growth and guidance? A proposal relating signaling pathways to growth cone mechanics. *Differentiation* **70,** 385–396.

Goodhill, G. J. (1993). Topography and ocular dominance: A model exploring positive correlations. *Biol. Cybern.* **69,** 109–118.

Goodhill, G. J. (1997). Diffusion in axon guidance. *Eur. J. Neurosci.* **9,** 1414–1421.

Goodhill, G. J. (1998). Gradients for retinotectal mapping. *In* "Advances in Neural Information Processing Systems" (M. I. Jordan, M. J. Kearns, and S. A. Solla, Eds.), vol. 10, pp. 152–158. MIT Press, Cambridge, MA.

Goodhill, G. J. (2000). Dating behavior of the retinal ganglion cell. *Neuron* **25,** 501–503.

Goodhill, G. J. (2007). Contributions of theoretical modeling to the understanding of neural map development. *Neuron* **56,** 301–311.

Goodhill, G. J., and Baier, H. (1998). Axon guidance: Stretching gradients to the limit. *Neural Comput.* **10,** 521–527.

Goodhill, G., and Barrow, H. (1994). The role of weight normalization in competitive learning. *Neural Comput.* **6,** 255–269.

Goodhill, G. J., and Cimponeriu, A. (2000). Analysis of the elastic net model applied to the formation of ocular dominance and orientation columns. *Network* **11,** 153–168.

Goodhill, G. J., and Urbach, J. S. (1999). Theoretical analysis of gradient detection by growth cones. *J. Neurobiol.* **41,** 230–241.
Goodhill, G., and Willshaw, D. (1990). Application of the elastic net algorithm to the formation of ocular dominance stripes. *Network* **1,** 41–59.
Goodhill, G. J., and Xu, J. (2005). The development of retinotectal maps: A review of models based on molecular gradients. *Network* **16,** 5–34.
Goodhill, G. J., Bates, K. R., and Montague, P. R. (1997). Influences on the global structure of cortical maps. *Proc. R. Soc. Lond. B Biol. Sci.* **264,** 649–655.
Goodhill, G. J., Gu, M., and Urbach, J. S. (2004). Predicting axonal response to molecular gradients with a computational model of filopodial dynamics. *Neural Comput.* **16,** 2221–2243.
Gosse, N. J., Nevin, L. M., and Baier, H. (2008). Retinotopic order in the absence of axon competition. *Nature* **452,** 892–895.
Graham, B. P., and van Ooyen, A. (2001). Compartmental models of growing neurites. *Neurocomputing* **38–40,** 31–36.
Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W., and Tank, D. W. (2007a). Stability and nuclear dynamics of the Bicoid morphogen gradient. *Cell* **130,** 141–152.
Gregor, T., Tank, D. W., Wieschaus, E. F., and Bialek, W. (2007b). Probing the limits to positional information. *Cell* **130,** 153–164.
Halloran, M. C., and Wolman, M. A. (2006). Repulsion or adhesion: Receptors make the call. *Curr. Opin. Cell Biol.* **18,** 533–540.
Hansen, M. J., Dallal, G. E., and Flanagan, J. G. (2004). Retinal axon response to ephrin-As shows a graded, concentration-dependent transition from growth promotion to inhibition. *Neuron* **42,** 717–730.
Hebb, D. O. (1949). "The Organization of Behaviour." Wiley, New York.
Hely, T. A., and Willshaw, D. J. (1998). Short-term interactions between microtubules and actin filaments underlie long-term behaviour in neuronal growth cones. *Proc. Biol. Sci.* **265,** 1801–1807.
Hentschel, H. G., and van Ooyen, A. (1999). Models of axon guidance and bundling during development. *Proc. Biol. Sci.* **266,** 2231–2238.
Hertz, J., Krogh, A., and Palmer, R. (1991). Introduction to the theory of neural computation. *In* "Lecture Notes in the Santa Fe Institute Studies in the Sciences of Complexity" Addison-Wesley, Redwood City, CA.
Hinton, G. (1989). Connectionist learning procedures. *Artif. Intell.* **40,** 185–234.
Hinton, G. E. (1992). How neural networks learn from experience. *Sci. Am.* **267,** 144–151.
Hinton, G. E. (2007). Learning multiple layers of representation. *Trends Cogn. Sci.* **11,** 428–434.
Holt, C. E. (1984). Does timing of axon outgrowth influence initial retinotectal topography in Xenopus? *J. Neurosci.* **4,** 1130–1152.
Honda, H. (1998). Topographic mapping in the retinotectal projection by means of complementary ligand and receptor gradients: A computer simulation study. *J. Theor. Biol.* **192,** 235–246.
Honda, H. (2003). Competition between retinal ganglion axons for targets under the servomechanism model explains abnormal retinocollicular projection of Eph receptor-overexpressing or ephrin-lacking mice. *J. Neurosci.* **23,** 10368–10377.
Honda, H. (2004). Competitive interactions between retinal ganglion axons for tectal targets explain plasticity of retinotectal projection in the servomechanism model of retinotectal mapping. *Dev. Growth Differ.* **46,** 425–437.
Hope, R. A., Hammond, B. J., and Gaze, R. M. (1976). The arrow model: Retinotectal specificity and map formation in the goldfish visual system. *Proc. R. Soc. Lond. B Biol. Sci.* **194,** 447–466.

Hornberger, M. R., Dtting, D., Ciossek, T., Yamada, T., Handwerker, C., Lang, S., Weth, F., Huf, J., Wessel, R., Logan, C., Tanaka, H., and Drescher, U. (1999). Modulation of EphA receptor function by coexpressed ephrinA ligands on retinal ganglion cell axons. *Neuron* **22,** 731–742.

Hsu, A. S., and Dayan, P. (2007). An unsupervised learning model of neural plasticity: Orientation selectivity in goggle-reared kittens. *Vision Res.* **47,** 2868–2877.

Hua, J. Y., Smear, M. C., Baier, H., and Smith, S. J. (2005). Regulation of axon growth *in vivo* by activity-based competition. *Nature* **434,** 1022–1026.

Huynh-Do, U., Stein, E., Lane, A. A., Liu, H., Cerretti, D. P., and Daniel, T. O. (1999). Surface densities of ephrin-B1 determine EphB1-coupled activation of cell attachment through $\alpha_v\beta_3$ and $\alpha_5\beta_1$ integrins. *EMBO J.* **18,** 2165–2173.

Hyvärinen, A., and Hoyer, P. O. (2001). A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Res.* **41,** 2413–2423.

Hyvärinen, A., and Köster, U. (2007). Complex cell pooling and the statistics of natural images. *Network* **18,** 81–100.

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Manu, E., Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., and Reinitz, J. (2004). Dynamic control of positional information in the early drosophila embryo. *Nature* **430,** 368–371.

Jilkine, A., Marée, A. F. M., and Edelstein-Keshet, L. (2007). Mathematical model for spatial segregation of the Rho-family GTPases based on inhibitory crosstalk. *Bull. Math. Biol.* **69,** 1943–1978.

Kaethner, R. J., and Stuermer, C. A. (1992). Dynamics of terminal arbor formation and target approach of retinotectal axons in living zebrafish embryos: A time-lapse study of single axons. *J. Neurosci.* **12,** 3257–3271.

Katz, M. J., and Lasek, R. J. (1985). Early axon patterns of the spinal cord: Experiments with a computer. *Dev. Biol.* **109,** 140–149.

Katz, L. C., and Shatz, C. J. (1996). Synaptic activity and the construction of cortical circuits. *Science* **274,** 1133–1138.

Katz, M. J., George, E. B., and Gilbert, L. J. (1984). Axonal elongation as a stochastic walk. *Cell Motil.* **4,** 351–370.

Kepecs, A., van Rossum, M. C. W., Song, S., and Tegner, J. (2002). Spike-timing-dependent plasticity: Common themes and divergent vistas. *Biol. Cybern.* **87,** 446–458.

Kiddie, G., McLean, D., Ooyen, A. V., and Graham, B. (2005). Biologically plausible models of neurite outgrowth. *Prog. Brain Res.* **147,** 67–80.

King, C. E., Wallace, A., Rodger, J., Bartlett, C., Beazley, L. D., and Dunlop, S. A. (2003). Transient up-regulation of retinal EphA3 and EphA5, but not ephrin-A2, coincides with re-establishment of a topographic map during optic nerve regeneration in goldfish. *Exp. Neurol.* **183,** 593–599.

Kobayashi, N., and Mundel, P. (1998). A role of microtubules during the formation of cell processes in neuronal and non-neuronal cells. *Cell Tissue Res.* **291,** 163–174.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43,** 59–69.

Kohonen, T. (1995). "Self-Organizing Maps." Springer, Berlin.

Körding, K. (2007). Decision theory: What "should" the nervous system do? *Science* **318,** 606–610.

Koulakov, A. A., and Tsigankov, D. N. (2004). A stochastic model for retinocollicular map development. *BMC Neurosci.* **5,** 30.

Krottje, J. K., and van Ooyen, A. (2007). A mathematical framework for modeling axon guidance. *Bull. Math. Biol.* **69,** 3–31.

Krzanowski, W. (1988). "Principles of Multivariate Analysis: A User's Perspective." Oxford Statistical Science Series V3. Oxford University Press, New York.
Kullander, K., and Klein, R. (2002). Mechanisms and functions of Eph and ephrin signalling. *Nat. Rev. Mol. Cell Biol.* **3,** 475–486.
Linsker, R. (1986). From basic network principles to neural architecture (series). *Proc. Natl. Acad. Sci. USA* **83,** 7508–7512, 8390–8394, 8779–8783.
Linsker, R. (1990). Perceptual neural organization: Some approaches based on network models and information theory. *Annu. Rev. Neurosci.* **13,** 257–281.
Löschinger, J., Weth, F., and Bonhoeffer, F. (2000). Reading of concentration gradients by axonal growth cones. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355,** 971–982.
MacKay, D., and Miller, K. (1990). Analysis of Linsker's application of Hebbian rules to linear networks. *Network* **1,** 257–297.
Malenka, R. C., and Bear, M. F. (2004). LTP and LTD: An embarrassment of riches. *Neuron* **44,** 5–21.
Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275,** 213–215.
Maskery, S., and Shinbrot, T. (2005). Deterministic and stochastic elements of axonal guidance. *Annu. Rev. Biomed. Eng.* **7,** 187–221.
McLaughlin, T., and O'Leary, D. D. M. (2005). Molecular gradients and development of retinotopic maps. *Annu. Rev. Neurosci.* **28,** 327–355.
McLaughlin, T., Torborg, C. L., Feller, M. B., and O'Leary, D. D. M. (2003). Retinotopic map refinement requires spontaneous retinal waves during a brief critical period of development. *Neuron* **40,** 1147–1160.
McLean, D., van Ooyen, A., and Graham, B. P. (2004). Continuum model for tubulin-driven neurite elongation. *Neurocomputing* **58–60,** 511–516.
Meinhardt, H. (1999). Orientation of chemotactic cells and growth cones: Models and mechanisms. *J. Cell Sci.* **112**(Pt. 17), 2867–2874.
Miller, K. D. (1994). A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. *J. Neurosci.* **14,** 409–441.
Miller, K. D. (1996). Synaptic economics: Competition and cooperation in synaptic plasticity. *Neuron* **17,** 371–374.
Miller, K., and MacKay, D. (1994). The role of constraints in Hebbian learning. *Neural Comput.* **6,** 100–126.
Miller, K. D., Keller, J. B., and Stryker, M. P. (1989). Ocular dominance column development: Analysis and simulation. *Science* **245,** 605–615.
Ming, G. I., Wong, S. T., Henley, J., Yuan, X. B., Song, H. J., Spitzer, N. C., and Poo, M. M. (2002). Adaptation in the chemotactic guidance of nerve growth cones. *Nature* **417,** 411–418.
Minsky, M., and Papert, S. (1969). "Perceptrons." MIT Press, Cambridge, MA.
Mogilner, A., and Rubinstein, B. (2005). The physics of filopodial protrusion. *Biophys. J.* **89,** 782–795.
Mogilner, A., Wollman, R., and Marshall, W. F. (2006). Quantitative modeling in cell biology: What is it good for? *Dev. Cell* **11,** 279–287.
Morrison, A., Diesmann, M., and Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. *Biol. Cybern.* **98,** 459–478.
Mortimer, D., Fothergill, T., Pujic, Z., Richards, L. J., and Goodhill, G. J. (2008). Growth cone chemotaxis. *Trends Neurosci.* **31,** 90–98.
Nakamoto, M., Cheng, H. J., Friedman, G. C., McLaughlin, T., Hansen, M. J., Yoon, C. H., O'Leary, D. D., and Flanagan, J. G. (1996). Topographically specific effects of ELF-1 on retinal axon guidance *in vitro* and retinal axon mapping *in vivo*. *Cell* **86,** 755–766.

Nicol, X., Voyatzis, S., Muzerelle, A., Narboux-Nême, N., Südhof, T. C., Miles, R., and Gaspar, P. (2007). cAMP oscillations and retinal activity are permissive for ephrin signaling during the establishment of the retinotopic map. *Nat. Neurosci.* **10,** 340–347.

Obermayer, K., Ritter, H., and Schulten, K. (1990). A principle for the formation of the spatial structure of cortical feature maps. *Proc. Natl. Acad. Sci. USA* **87,** 8345–8349.

Obermayer, K., Blasdel, C. G., and Schulten, K. (1992). Statistical–mechanical analysis of self-organization and pattern formation during the development of visual maps. *Phys. Rev. A* **45,** 7568–7589.

Odde, D. J., and Buettner, H. M. (1998). Autocorrelation function and power spectrum of two-state random processes used in neurite guidance. *Biophys. J.* **75,** 1189–1196.

Odde, D., Tanaka, E., Hawkins, S., and Buettner, H. (1996). Stochastic dynamics of the nerve growth cone and its microtubules during neurite outgrowth. *Biotechnol. Bioeng.* **50**(4), 452–461.

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

Oja, E. (1989). Neural networks, principal components, and subspaces. *Int. J. Neural Syst.* **1,** 61–68.

Olshausen, B. A., and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381,** 607–609.

Olshausen, B. A., and Field, D. J. (2004). Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* **14,** 481–487.

O'Rourke, N. A., and Fraser, S. E. (1990). Dynamic changes in optic fiber terminal arbors lead to retinotopic map formation: An *in vivo* confocal microscopic study. *Neuron* **5,** 159–171.

O'Rourke, N. A., Cline, H. T., and Fraser, S. E. (1994). Rapid remodeling of retinal arbors in the tectum with and without blockade of synaptic transmission. *Neuron* **12,** 921–934.

Overton, K. J., and Arbib, M. A. (1982a). Systems matching and topographic maps: The branch-arrow model (BAM). *In* "Competition and Cooperation in Neural Nets", Lecture Notes in Biomathematics (S. Amari and M.A. Arbib, Eds.), vol. 45, pp. 202–225. Springer, Berlin.

Overton, K. J., and Arbib, M. A. (1982b). The extended branch-arrow model of the formation of retino-tectal connections. *Biol. Cybern.* **45,** 157–175.

Paliwal, S., Iglesias, P. A., Campbell, K., Hilioti, Z., Groisman, A., and Levchenko, A. (2007). MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature* **446,** 46–51.

Pasquale, E. B. (2005). Eph receptor signalling casts a wide net on cell behaviour. *Nat. Rev. Mol. Cell Biol.* **6,** 462–475.

Piper, M., Salih, S., Weinl, C., Holt, C. E., and Harris, W. A. (2005). Endocytosis-dependent desensitization and protein synthesis-dependent resensitization in retinal growth cone adaptation. *Nat. Neurosci.* **8,** 179–186.

Poliakov, A., Cotrina, M., and Wilkinson, D. G. (2004). Diverse roles of Eph receptors and ephrins in the regulation of cell migration and tissue assembly. *Dev. Cell* **7,** 465–480.

Prestige, M. C., and Willshaw, D. J. (1975). On a role for competition in the formation of patterned neural connexions. *Proc. R. Soc. Lond. B Biol. Sci.* **190,** 77–98.

Rashid, T., Upton, A. L., Blentic, A., Ciossek, T., Knöll, B., Thompson, I. D., and Drescher, U. (2005). Opposing gradients of ephrin-As and EphA7 in the superior colliculus are essential for topographic mapping in the mammalian visual system. *Neuron* **47,** 57–69.

Reber, M., Burrola, P., and Lemke, G. (2004). A relative signalling model for the formation of a topographic neural map. *Nature* **431,** 847–853.

Rodger, J., Bartlett, C. A., Beazley, L. D., and Dunlop, S. A. (2000). Transient up-regulation of the rostrocaudal gradient of ephrin A2 in the tectum coincides

with reestablishment of orderly projections during optic nerve regeneration in goldfish. *Exp. Neurol.* **166,** 196–200.
Roskies, A. L., and O'Leary, D. D. (1994). Control of topographic retinal axon branching by inhibitory membrane-bound molecules. *Science* **265,** 799–803.
Rosoff, W. J., Urbach, J. S., Esrick, M. A., McAllister, R. G., Richards, L. J., and Goodhill, G. J. (2004). A new chemotaxis assay shows the extreme sensitivity of axons to molecular gradients. *Nat. Neurosci.* **7,** 678–682.
Rumelhart, D., and Zipser, D. (1986). Feature discovery by competitive learning. *In* "Parallel Distributed Processing: Explorations in the Microstructure of Cognition" (J. L. McClelland and D. E. Rumelhart, Eds.), pp. 151–193. MIT Press, Cambridge, MA.
Ruthazer, E. S., and Cline, H. T. (2004). Insights into activity-dependent map formation from the retinotectal system: A middle-of-the-brain perspective. *J. Neurobiol.* **59,** 134–146.
Ruthazer, E. S., Akerman, C. J., and Cline, H. T. (2003). Control of axon branch dynamics by correlated activity *in vivo*. *Science* **301,** 66–70.
Sakumura, Y., Tsukada, Y., Yamamoto, N., and Ishii, S. (2005). A molecular model for axon guidance based on cross talk between Rho GTPases. *Biophys. J.* **89,** 812–822.
Schmidt, J. T. (1978). Retinal fibers alter tectal positional markers during the expansion of the retinal projection in goldfish. *J. Comp. Neurol.* **177,** 279–295.
Schmidt, J. T. (1990). Long-term potentiation and activity-dependent retinotopic sharpening in the regenerating retinotectal projection of goldfish: Common sensitive period and sensitivity to NMDA blockers. *J. Neurosci.* **10,** 233–246.
Schmidt, J. T., and Easter, S. S. (1978). Independent biaxial reorganization of the retinotectal projection: A reassessment. *Exp. Brain Res.* **31,** 155–162.
Schmidt, J. T., Cicerone, C. M., and Easter, S. S. (1978). Expansion of the half retinal projection to the tectum in goldfish: An electrophysiological and anatomical study. *J. Comp. Neurol.* **177,** 257–277.
Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* **275,** 1593–1599.
Sejnowski, T. J. (1977). Statistical constraints on synaptic plasticity. *J. Theor. Biol.* **69,** 385–389.
Sharma, S. C. (1972). Reformation of retinotectal projections after various tectal ablations in adult goldfish. *Exp. Neurol.* **34,** 171–182.
Simon, D. K., and O'Leary, D. D. (1992). Development of topographic order in the mammalian retinocollicular projection. *J. Neurosci.* **12,** 1212–1232.
Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Curr. Opin. Neurobiol.* **13,** 144–149.
Song, S., and Abbott, L. F. (2001). Cortical development and remapping through spike timing-dependent plasticity. *Neuron* **32,** 339–350.
Song, S., Miller, K. D., and Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci.* **3,** 919–926.
Sperry, R. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proc. Natl. Acad. Sci. USA* **50,** 703–710.
Stuermer, C. A. (1986). Pathways of regenerated retinotectal axons in goldfish. I. Optic nerve, tract and tectal fascicle layer. *J. Embryol. Exp. Morphol.* **93,** 1–28.
Sutton, R. (1988). Learning to predict by the methods of temporal differences. *Mach. Learn.* **3,** 9–44.
Swindale, N. V. (1996). The development of topography in the visual cortex: A review of models. *Network* **7,** 161–247.
Thivierge, J. P., and Balaban, E. (2007). Getting into shape: Optimal ligand gradients for axonal guidance. *Biosystems* **90,** 61–77.

Tindall, M. J., Porter, S. L., Maini, P. K., Gaglia, G., and Armitage, J. P. (2008a). Overview of mathematical approaches used to model bacterial chemotaxis I: The single cell. *Bull. Math. Biol.* **70,** 1525–1569.

Tindall, M. J., Maini, P. K., Porter, S. L., and Armitage, J. P. (2008b). Overview of mathematical approaches used to model bacterial chemotaxis II: Bacterial populations. *Bull. Math. Biol.* **70,** 1570–1607.

Tsigankov, D. N., and Koulakov, A. A. (2006). A unifying model for activity-dependent and activity-independent mechanisms predicts complete structure of topographic maps in ephrin-A deficient mice. *J. Comput. Neurosci.* **21,** 101–114.

Udin, S. B., and Fawcett, J. W. (1988). Formation of topographic maps. *Annu. Rev. Neurosci.* **11,** 289–327.

Urbach, J., and Goodhill, G. (1999). Limitations on detection of gradients of diffusible chemicals by axons. *Neurocomputing* **26–27,** 39–43.

van Hateren, J. H., and Ruderman, D. L. (1998). Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* **265,** 2315–2320.

van Hateren, J. H., and van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* **265,** 359–366.

van Ooyen, A. (2001). Competition in the development of nerve connections: A review of models. *Network* **12,** R1–R47.

van Ooyen, A. (Ed.), (2003). ''Modeling Neural Development.'' MIT Press, Cambridge, MA.

van Veen, M. P., and van Pelt, J. (1994). Neuritic growth rate described by modeling microtubule dynamics. *Bull. Math. Biol.* **56,** 249–273.

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

von der Malsburg, C., and Willshaw, D. (1976). A mechanism for producing continuous neural mappings: Ocularity dominance stripes and ordered retino-tectal projections. *Exp. Brain Res.* Suppl. **1,** 463–469.

von der Malsburg, C., and Willshaw, D. J. (1977). How to label nerve cells so that they can interconnect in an ordered fashion. *Proc. Natl. Acad. Sci. USA* **74,** 5176–5178.

von Philipsborn, A., and Bastmeyer, M. (2007). Mechanisms of gradient detection: A comparison of axon pathfinding with eukaryotic cell migration. *Int. Rev. Cytol.* **263,** 1–62.

Walter, J., Kern-Veits, B., Huf, J., Stolze, B., and Bonhoeffer, F. (1987a). Recognition of position-specific properties of tectal cell membranes by retinal axons *in vitro*. *Development* **101,** 685–696.

Walter, J., Henke-Fahle, S., and Bonhoeffer, F. (1987b). Avoidance of posterior tectal membranes by temporal retinal axons. *Development* **101,** 909–913.

Weber, C., Ritter, H., Cowan, J., and Klaus Obermayer, K. (1997). Development and regeneration of the retinotectal map in goldfish: A computational study. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **352**(1361), 1603–1623.

Wenisch, O. G., Noll, J., and van Hemmen, J. L. (2005). Spontaneously emerging direction selectivity maps in visual cortex through STDP. *Biol. Cybern.* **93,** 239–247.

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

Wilkinson, D. G. (2000). Topographic mapping: Organising by repulsion and competition? *Curr. Biol.* **10,** R447–R451.

Wilkinson, D. G. (2001). Multiple roles of Eph receptors and ephrins in neural development. *Nat. Rev. Neurosci.* **2,** 155–164.

Willshaw, D. (2006). Analysis of mouse EphA knockins and knockouts suggests that retinal axons programme target cells to form ordered retinotopic maps. *Development* **133,** 2705–2717.

Willshaw, D. J., and von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proc. R. Soc. Lond. B Biol. Sci.* **194,** 431–445.

Willshaw, D. J., and von der Malsburg, C. (1979). A marker induction mechanism for the establishment of ordered neural mappings: Its application to the retinotectal problem. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **287,** 203–243.

Xu, J., Rosoff, W. J., Urbach, J. S., and Goodhill, G. J. (2005). Adaptation is not required to explain the long-term response of axons to molecular gradients. *Development* **132,** 4545–4552.

Yates, P. A., Roskies, A. L., McLaughlin, T., and O'Leary, D. D. (2001). Topographic-specific axon branching controlled by ephrin-As is the critical event in retinotectal map development. *J. Neurosci.* **21,** 8548–8563.

Yates, P. A., Holub, A. D., McLaughlin, T., Sejnowski, T. J., and O'Leary, D. D. M. (2004). Computational modeling of retinotopic map development to define contributions of EphA–ephrinA gradients, axon–axon interactions, and patterned activity. *J. Neurobiol.* **59,** 95–113.

Yoon, M. G. (1976). Progress of topographic regulation of the visual projection in the halved optic tectum of adult goldfish. *J. Physiol.* **257,** 621–643.

Yoon, M. G. (1980). Retention of topographic addresses by reciprocally translocated tectal re-implants in adult goldfish. *J. Physiol.* **308,** 197–215.

Young, J. M., Waleszczyk, W. J., Wang, C., Calford, M. B., Dreher, B., and Obermayer, K. (2007). Cortical reorganization consistent with spike timing-but not correlation-dependent plasticity. *Nat. Neurosci.* **10,** 887–895.

Yu, H., Farley, B. J., Jin, D. Z., and Sur, M. (2005). The coordinated mapping of visual space and response features in visual cortex. *Neuron* **47,** 267–280.