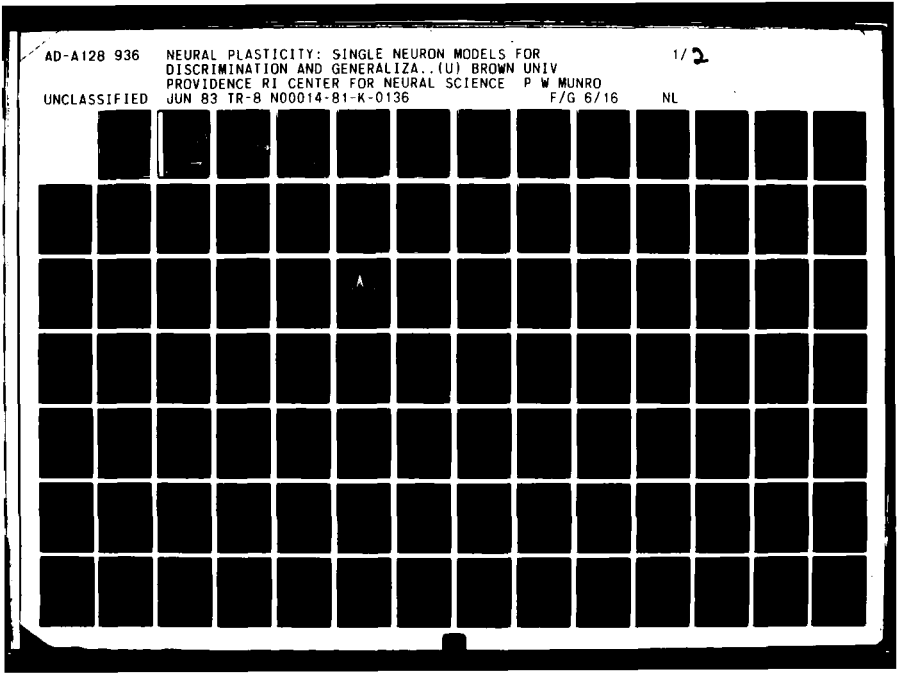
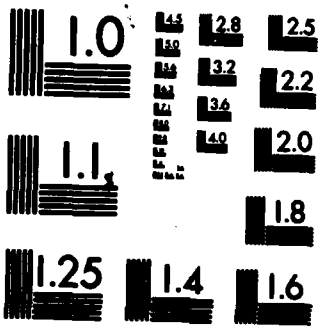
AD-A128 936 NEURAL PLASTICITY: SINGLE NEURON MODELS FOR DISCRIMINATION AND GENERALIZA..(U) BROWN UNIV PROVIDENCE RI CENTER FOR NEURAL SCIENCE P W MUNRO

1/2

UNCLASSIFIED JUN 83 TR-8 N00014-81-K-0136 F/G 6/16 NL





(12)

AD A128936

DTIC FILE COPY

83 05 26 055

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: # #8	2. GOVT ACCESSION NO.: AD-A129836
3. RECIPIENT'S CATALOG NUMBER: DT	
4. TITLE (and Subtitle): Neural Plasticity: Single Neuron Models for Discrimination and Generalization and an Experimental Ensemble Approach	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Paul Wesley Munro	8. CONTRACT OR GRANT NUMBER(s): N00014-81-K-0136
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Neural Science, Brown University, Providence, Rhode Island 02912	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-201-484
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Program, Office of Naval Research (Code 442PT), Arlington, Virginia 22217	12. REPORT DATE: April 29, 1983
	13. NUMBER OF PAGES: 100
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. Publication in whole or in part is permitted for any purpose of the United States Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Ph. D. Thesis.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Neural modeling
selectivity development
binocularity
synaptic plasticity
generalization

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A special form for modification of neuronal response properties is described in which the change in the synaptic state vector is parallel to the vector of afferent activity. This process is termed "parallel modification" and its theoretical and experimental implications are examined.

A theoretical framework has been devised to describe the complementary functions of generalization and discrimination by single neurons. This constitutes a basis for three models each describing processes for the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

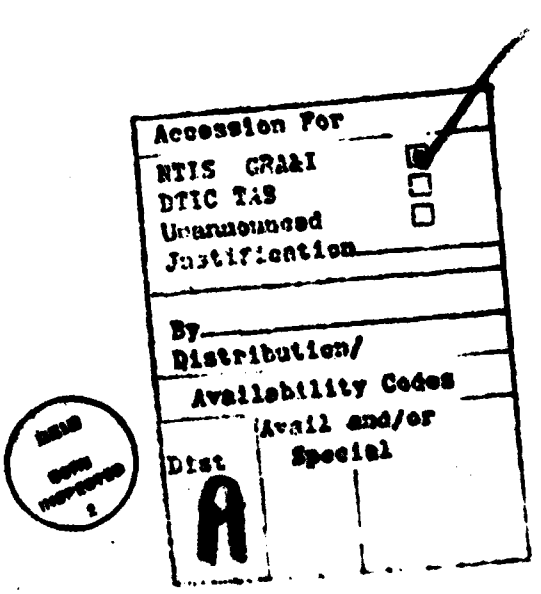
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

development of maximum selectivity (discrimination) and minimum selectivity (generalization) by neurons. Strengthening and weakening of synapses is expressed as a product of the presynaptic activity and a nonlinear modulatory function of two postsynaptic variables--namely a measure of the spatially integrated activity of the cell and a temportal integration (time-average) of that activity. Some theorems are given for low-dimensional systems and a computer simulation results from more complex systems are discussed.

Model neurons that achieve high selectivity mimic the development of cat visual cortex neurons in a wide variety of rearing condiditons. A role for low-selectivity neurons is proposed in which they provide inhibitory input to neurons of the opposite type, thereby suppressing the common component of a pattern class and enhancing their selective properties. Such contrast-enhancing circuits are analyzed and supported by computer simulation. To enable maximum selectivity, the net inhibition to a cell must become strong enough to offset whatever excitation is produced by the non-preferred patterns.

Ramifications of parallel models for certain experimental paradigms are analyzed. A methodology is outlined for testing synaptic modification hypotheses in the laboratory. A plastic projection from one neuronal population to another will attain stable equilibrium under periodic electrical stimulation of constant intensity. The perturbative effect of shifting this intensity level can yield important information regarding the form of the modification equation. The pathway from entorhinal cortex to dentate gyrus in the rat is an ideal candidate for such experiments.



Neural Plasticity:
Single Neuron Models for Discrimination and Generalization
and an Experimental Ensemble Approach

by

Paul Wesley Munro

B.S., Rensselaer Polytechnic Institute, 1977

Sc.M., Brown University, 1979

Thesis

Submitted in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in the Department of
Physics at Brown University

June, 1983

Abstract of "Neural Plasticity: Single Neuron Models for Discrimination and Generalization and an Experimental Ensemble Approach" by Paul Wesley Munro, Ph.D., Brown University, June, 1983.

A special form for modification of neuronal response properties is described in which the change in the synaptic state vector is parallel to the vector of afferent activity. This process is termed "parallel modification" and its theoretical and experimental implications are examined.

A theoretical framework has been devised to describe the complementary functions of generalization and discrimination by single neurons. This constitutes a basis for three models each describing processes for the development of maximum selectivity (discrimination) and minimum selectivity (generalization) by neurons. Strengthening and weakening of synapses is expressed as a product of the presynaptic activity and a nonlinear modulatory function of two postsynaptic variables -- namely a measure of the spatially integrated activity of the cell and a temporal integration (time-average) of that activity. Some theorems are given for low-dimensional systems and computer simulation results from more complex systems are discussed.

Model neurons that achieve high selectivity mimic the development of cat visual cortex neurons in a wide variety of rearing conditions. A role for low-selectivity neurons is proposed in which they provide inhibitory input to neurons of the opposite type, thereby suppressing the common component of a pattern class and enhancing their selective properties. Such contrast-enhancing circuits are analyzed and supported by computer simulation. To enable maximum selectivity, the net inhibition to a cell must become strong enough to offset whatever excitation is produced by the non-preferred patterns.

Ramifications of parallel models for certain experimental paradigms are analyzed. A methodology is outlined for testing synaptic modification hypotheses in the laboratory. A plastic projection from one neuronal population to another will attain stable equilibrium under periodic electrical stimulation of constant intensity. The perturbative effect of shifting this intensity level can yield important information regarding the form of the modification equation. The pathway from entorhinal cortex to dentate gyrus in the rat is an ideal candidate for such experiments.

Table of Contents

List of Illustrations and Tables

Tables

Introduction

That synaptic modification is a principal component of neural plasticity has come to be a widely held opinion among neuroscientists. Theoreticians have launched several attempts, with varying degrees of success, to deduce explicit (synaptic) principles underlying learning by examining the selforganizational consequences of various microscopic hypotheses for comparison with anatomical, physiological, and behavioral data. Such efforts have been directed toward several brain structures and pathways including sensory cortex, motor cortex, the hippocampal and reticular formations, the cerebellum, and various thalamic nuclei. It is not known whether the diverse set of subsystems that make up the CNS employ general learning mechanisms that are few in number or several specialized ones. Two plasticity phenomena motivate the ideas in the present thesis:

- the development of stimulus selectivity in visual cortex neurons
- potentiation and depotentiation of the bilateral pathway from entorhinal cortex to dentate gyrus

It is generally agreed that the instantaneous spiking frequency of an axon reflects a spatiotemporal integration of synaptic excitation and inhibition over the corresponding neuron, but the informational content of a neuron's response with respect to the stimulating activity pattern is an open issue. The very existence of

behavioral phenomena implies that large neural networks evolve to give appropriate ensemble responses to their stimuli. Each constituent neuron must therefore develop response attributes which are, in some sense, meaningful. After defining the model neuron and its stimulus environment as mathematical structures, Part I presents a mathematical framework describing two complementary forms of abstraction by individual neurons: generalization, a process of finding similarities between stimuli, and differentiation, a separation process. An analogous dichotomy in cognitive theory has been extensively studied by Tversky and Gati (1978, 1982). They develop a "contrast model" of similarity and difference perception and apply it to psychological data.

Part II extends a theory for selectivity development in visual cortex proposed by Bienenstock et al (1982) which describes a process by which cells become maximally selective over their input environments. Development of orientation specificity and ocular dominance has been subject to extensive study using many paradigms that manipulate the visual environment in various ways. Some of these investigations are described and it is shown that the theory accounts for the corresponding results. It is then shown that a simple change in the evolution equation describes cells which seek minimum, rather than maximum, selectivity. These cells tune to common features rather than distinctive features. Thus a general system is developed to account for both types of selectivity development.

In Part III an alternative model is introduced to describe the same set of visual cortex data. It also includes an implementation of the maximum-minimum selectivity framework. Here, two antagonistic mechanisms are postulated: one weakening a cell's input synapses and the other strengthening them. These mechanisms are coupled by a variable which evolves according to the same equation as the synapses. This theory rests on a more basic set of hypotheses which lead to a mathematical formulation quite similar to the previous model's starting hypothesis, but with some differences. First, the new hypotheses make certain assumptions about the biochemical mechanisms responsible for synaptic modification and second, there is a testable difference between the two models which is discussed in Part IV.

Part IV describes a population stimulus-record method for electrophysiological testing of synaptic modification models. The pre- and postsynaptic activities of certain synapses can be controlled to a degree and their efficacies roughly measured in some systems. Such measurements can give important information regarding the dependence of the change in synaptic strength on the pre- and postsynaptic variables. Since this method is applicable to neuron populations it may prove a useful tool for experimental design and data interpretation. The predicted results of the model in Part I are discussed as well as those of some other models. This approach was inspired by the hippocampal potentiation studies of Levy and

Steward (1979), so it is no accident that the anatomy and physiology in that system is ideal for this experimental framework. An outline of the biology of the entorhinal cortex to dentate gyrus projection is therefore included at the end of Part IV.

Background

The ability of organisms to process information constitutes a common basis for all behavioral phenomena. Over the past century, many empirically based theories of behavior have evolved, reshaping philosophical and clinical approaches to mental function. A nearly simultaneous accumulation of discoveries in the fields of neuroanatomy and neurophysiology has established a detailed picture of the nervous system as an enormously complex network of highly specialized cells. A full appreciation of neural function must begin with an understanding of the sub-microscopic chemical and physical membrane mechanisms of the neuron and eventually account for psychological data. Theoreticians are attempting to close the gap between neurobiology and psychology by developing mathematical theories of learning by neurons and neural networks.

The technique of stimulus-response (SR) is perhaps the most valuable protocol in all of experimental science. Response characteristics may provide the common language needed to relate electrophysiological and psychological data. We will assume that the action taken by a neural network is dependent solely upon the stimulus applied to the network and quantities internal to the network. These internal quantities constitute the state of the network. If the state can be modified by the application of stimuli, it is called plastic. Nonmodifiable states are termed

nonplastic. Learning processes are intrinsically plastic since they involve adjusting response properties in a network. This approach to the study of learning is taken by Thorndike (1932) who begins the introduction to his book The Fundamentals of Learning with this sentence: 'We are concerned in this volume with the fundamental facts of learning whereby a situation which first evokes response A later evokes response B, different from A'.

Note that nothing has been said regarding the complexity of the network. A neural network may be anything from an entire nervous system down to a single neuron or even part of a single neuron (Poggio and Koch, 1981). If the network's internal quantities and the transfer (input-output) function are known, then the state of the network has been specified such that the response to any stimulus can be calculated. Alternatively, if a sufficient (complete) set of responses to hypothetical stimuli and the transfer function are known, then the internal quantities can be deduced, again specifying the network state. Each complete set of neural stimuli is a coordinate system for the network state space. A natural or preferred coordinate system is given by the internal quantities. A nonplastic network is fully described by its state and transfer function. However, these values only partially specify plastic systems.

We are interested in the dynamics of the neural state. While the state and transfer function represent information stored

order of 10^4) sites along its membrane. These sites are called synapses, a term coined by Sherrington (1906) to indicate a junction between *neurons*. The neuron membrane maintains a "resting" electrical potential of about -60 mV (interior with respect to exterior). Each synapse transduces the incident signal to a slight disturbance in the resting potential across the membrane. If the soma membrane is sufficiently depolarized (driven above about -45 mV), a pulse is transmitted by the cell along its axon. The pulse frequency increases with membrane polarization to a maximum, while the amplitude remains constant. The frequency-coded signal is sent to many other neurons along the generally intricate arborization of the axon. This signal then either inhibits or excites other cells at synaptic sites by hyperpolarizing or depolarizing other membranes.

Thus <u>each neuron</u> operates as a stimulus-response mechanism generating a single-valued response to a multidimensional stimulus. Aggregations of nerve cells are known to cooperate functionally, receiving inputs from and projecting to common populations. So the nervous system can be sequentially broken down into subsystems of an arbitrary level down to the single-neuron system. The state of <u>any</u> of these subsystems can be characterized either by its SR properties or its electrophysiological properties (internal quantities and transfer function).

Deterministic treatments of neural function in the experimental framework of stimulus-response methodology are well suited

("knowledge"), the storage process ("learning") is expressed by a modification rule or plasticity function for the change in the neural state. This rule, together with the stimulus set delivered to the network thus determine the trajectory of the system through the state space.

Let us now consider the components of neural networks. Neurons are the basic units of neural function. As components of complex networks, these units are responsible for an extraordinary range of information processing tasks including perception, cognition, and action. Not all neurons are identical, there exist several anatomically defined categories of these cells as well as corresponding physiological and biochemical properties, but they have many features in common. All neurons contain the structures inherent to biological cells, however they are not self-replicating. Furthermore, each neuron transmits a time-varying electrical signal and receives several simultaneous electrochemical stimuli which also vary in time.

Nerve cells are irregularly shaped. Extending from the cell body (soma) are several processes - long, narrow extentions of the cell. These are of two types: dendrites, which generally conduct signals inwardly toward the cell body, and axons, which transmit frequency-coded signals to other cells. A typical neuron includes a soma, several dendrites, and a single axon which may have several branches. It receives signals from other cells at many (up to the

to synaptically based theoretical approaches. The SR technique has been used by both physiologists and psychologists for many decades. The pioneering works of Sherrington (1906) and Skinner (1938) together with their bibliographical references establish, respectively, the foundations of stimulus-response methodology in electrical and behavioral contexts. Both experimental approaches survive today, exemplified by electrophysiological data from orientation-specific cells in visual cortex (e.g., Hubel & Wiesel, 1962, 1977; Rauschecker & Singer, 1981) and the response time data of subjects learning prototypes (e.g., Posner & Keele, 1968, 1970).

Several mathematical models have been developed to account for these and other data. Examples of such theories for physiological data are Marr (1969), von der Malsburg (1973) and Nass & Cooper (1975); cognitively based models include Anderson (1972, 1973) and McClelland & Rumelhart (1981). Such models are usually based on the notion of "synaptic weights". As is described above, a neuron's instantaneous activity level is determined by a spatiotemporal summation of the influences of the activities of many other neurons. The influence of one neuron upon another is quantified as a synaptic "weight" or "efficacy". These synaptic weights along with other variables generally determine the transfer function of a neuron or a neuron network.

For example, a linear function might be used to express the output y of a neuron to its inputs x_i (1). The coefficients a_i

are the synaptic weights, representing the net influence of one cell upon another. This can easily be extended to a network consisting of a population $\{x_j\}$ projecting to a population $\{y_i\}$ (2). A simple nonlinear example is the threshold function given by (3), each neuron y_i having an activity threshold θ_i.

$$y = \sum_i \alpha_i x_i \tag{1}$$

$$y_i = \sum_j \alpha_{ij} x_j \tag{2}$$

$$y_i = K(\sum_j \alpha_{ij} x_j - \theta_i) \tag{3}$$

$$\underline{\text{where}} \quad K(x) \equiv \begin{cases} x & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

Such matrix or associative approaches began to receive attention in the 1940's (McCulloch and Pitts, 1943; Culbertson, 1948). This sort of mathematical formulation makes clear the close relationship between the transfer function and the quantification of the neural state. Modification rules (differential equations for the change in α or θ) are formulated or hypothesized in models for learning. An early hypothesis which has been a guiding principle for many theories is Hebb's (1949) "neurophysiological postulate":

> "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." [p. 62]

This states for neural function what Thorndike's (1913) Law of Use states for learning behavior:

> "The Law of Use is: When a modifiable connection is made between a situation and a response, that connection's strength is, other things being equal, increased." [p. 2]

> "Ultimately degrees of strength of a connection in behavior will be defined as degrees of some anatomical or physiological fact whereby synapses between neurones differ in intimacy." [p. 3]

Most synaptically based learning schemes extend this assumption to account for decrease in syntaptic efficacy as well. Inhibitory synapses must be separately considered as in Stent (1973) and Levy and Desmond (1982). An extensive examination of learning rules and their application is provided by Kohonen (1977).

The degree to which memory is distributed is an issue addressed by many models. Lashley (1929) sought to locate the site of learning in rat cortex by performing an extensive series of cortical lesions and measuring the ability to learn maze patterns as a function of the extent and location of each lesion. His results indicate that this ability is not localized to any particular locus in cortex:

> "The same retardation in learning is produced by equal amounts of destruction in any of the cyto-architectural fields. Hence the capacity to learn the maze is dependent upon the amount of functional cortical tissue and not upon its anatomical specialization." [p. 175]

Lashley's data are summarized in Figure 1.

The notion of distributed memory is attractive to many theoreticians. A hologram model was put forward by Longuet-Higgins (1968) since holography is a non-local storage technique -- partial destruction of a hologram reduces overall resolution of the stored image without affecting any particular region.

The opposing viewpoint is held by many researchers who point to observations (e.g. Hubel and Wiesel, 1963) of feature sensitive neurons in cortex. Several classes of selective neurons have been defined according to their modality and degree of specificity (Figure 2). The level of neuronal specificity at the most "gnostic" stage is thought to indicate the extent to which memory and cognition are localized. Barlow (1972) postulates the "cardinal cell" as a compromise between Sherrington's (1940) "pontifical cell" concept and complete distribution of memory traces.

Recent experimental data demonstrate even higher specificity than the well-documented hypercomplex neurons in visual cortex. Bruce et al (1981) find polymodal neurons in the extrastriate area they call the superior temporal polysensory area (STP). Gross et al (1972) report neurons in macaque inferotemporal cortex selective to visual inputs resembling a monkey's hand.

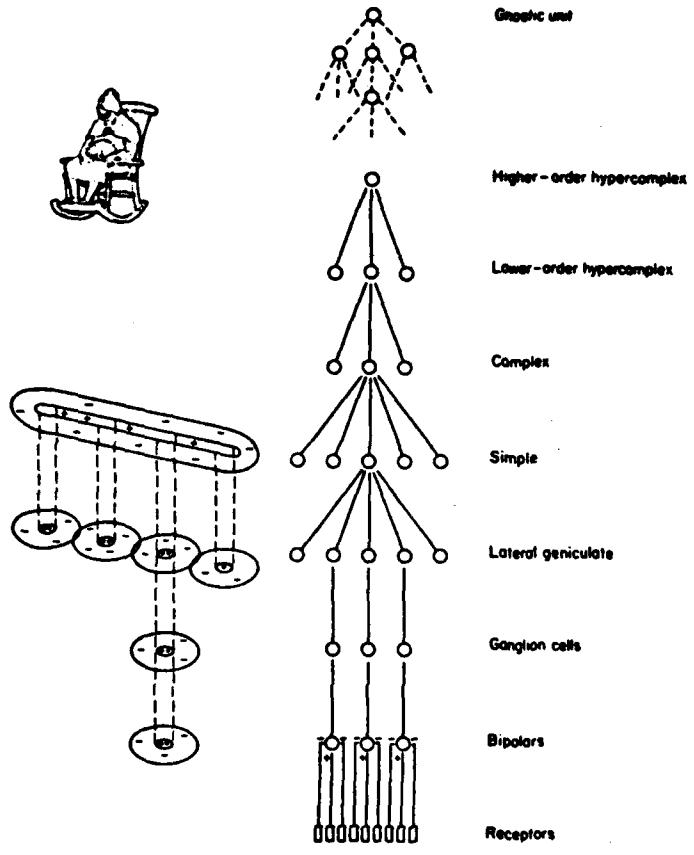


Figure 2. The hierarchical model of the visual system. Neuron classes are labelled on the right with corresponding receptive fields on the left. (From von Fieandt & Moustgaard (1977))

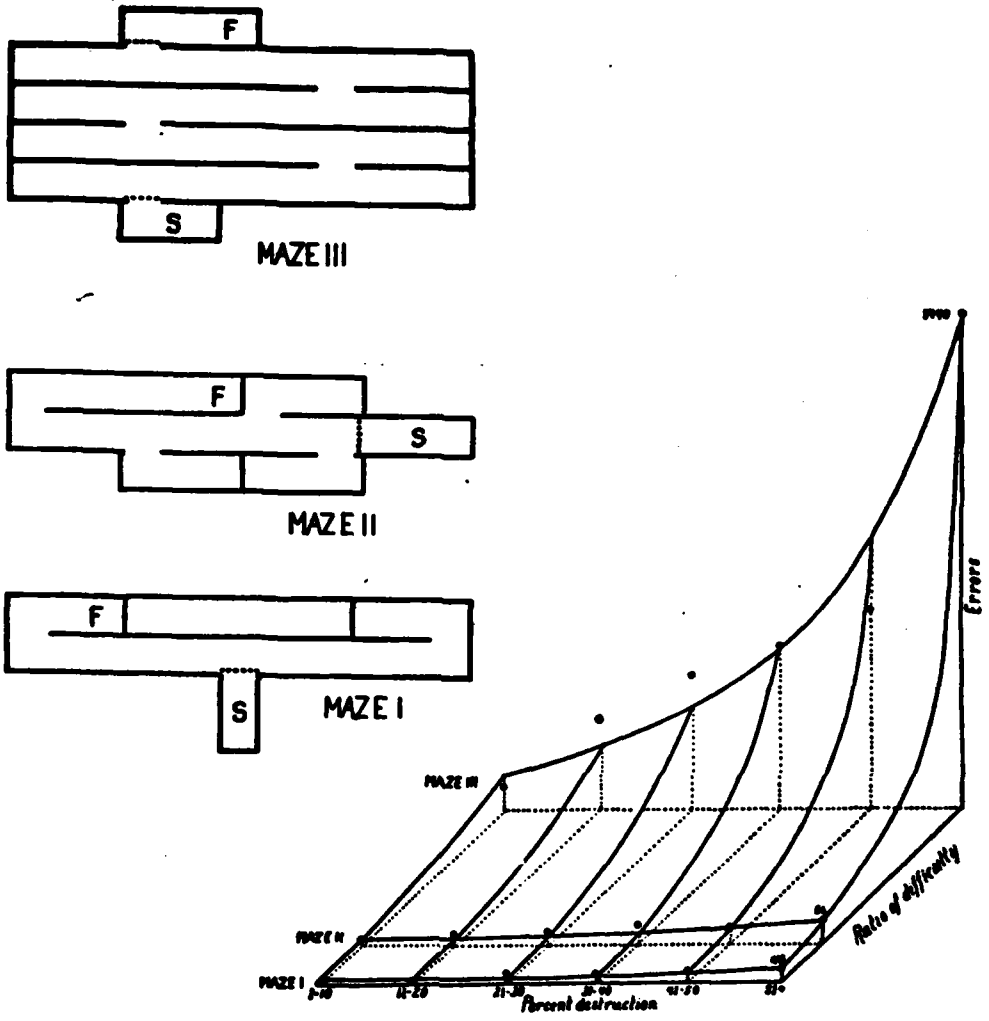


Figure 1. Lashley's experiment. Lashley's data shows a roughly continuous dependence of problem solving capability on the difficulty of the problem and on the degree of injury to the brain. (From Lashley (1929))

Finally, experiments by Yeo et al (1982) and Clark et al (1982) identify specific cerebellar nuclei that abolish classically conditioned responses in rabbits. These may be the loci sought by Lashley. It seems clear that the brain is specialized into functional regions, however the internal operation of these areas may be distributed to some extent. Also, the neural loci of individual memory traces may be distributed over several functionally distinct areas.

I. The Theoretical Framework: Definitions and Concepts

A. The Model Neuron and Parallel Modification

It is assumed that each neuron generates a dynamic frequency-coded scalar response to a time-varying vector stimulus and that the response properties evolve with experience (learning). Each of N afferent signals d_i are transduced to membrane depolarizations and ultimately induce a potential (x) at the soma. Spatiotemporal integration is idealized here to be an instantaneous linear sum x of the N inputs d_i weighted by the corresponding N synaptic efficacies m_i (4).

$$x = \sum_{i}^{N} m_i d_i \qquad (4)$$

For now, the integrated activity level x may be considered to correspond to an arbitrary physiological measure of the cell's excitation level such as the net postsynaptic potential or the resulting firing frequency. Although it is somewhat important for the signals to be summed linearly, the neuron's output frequency may be described by certain (nondecreasing, positive, bounded, and continuous) nonlinear functions of x. Thus the response of the model

neuron is purely a function of its synaptic weights.

The state of the neuron is specified by the synaptic vector m_i together with an internal quantity q which, although it plays no role in the transfer function (4), modulates the modification or plasticity process (5). Such systems are (N+1)-dimensional: N synapses plus the internal quantity.

$$\dot{m}_i(t) = \phi(x(t),q(t))d_i(t) \tag{5a}$$

$$\dot{q}(t) = \psi(x(t),q(t)) \tag{5b}$$

Since (5b) is an autonomous equation, q reflects a temporal integration of x over some recent interval, the scope of which is determined by the relative magnitudes of ϕ and ψ. Thus the synaptic vector m is modified parallel or antiparallel to the presented stimulus vector d, the strength of the modification being a function of x and q. Of all possible input stimuli of a given magnitude, the response to the one in the direction of d is most greatly affected by the change m. Parallel modification thus has the following behavioral implication: learning induced by a particular stimulus more greatly influences future responses to that stimulus than to others of the same magnitude. The intuitive plausibility of

this statement, though it may not rigorously hold, supports the formulation (5). One must be aware that "magnitude" is a nebulous term in this context.

Certain functions ϕ and ψ that yield otherwise desirable attributes for describing certain types of neural plasticity, may cause some or all of the state components (m_i, q) to grow without bound. There are several standard approaches to problems of divergence. For example the sum of the synaptic weights $(\sum m_i)$ may be normalized to or restricted to be less than a particular value. Another approach is to introduce an exponential decay term (6). While it is not included in the bulk of this thesis, Appendix A illustrates the utility of this term to a specific application.

$$\dot{m}(t) = \phi d - \varepsilon(m) \tag{6}$$

B. The Stimulus Environment

Like the term "neuron", the "stimulus environment" is a tangible to the experimentalist and an abstract mathematical entity to the theoretician. All possible neurally encoded stimuli d can be represented by points in a hypervolume U (universe) in N-space bounded along each coordinate by a minimum firing frequency (zero) and a maximum, determined by the absolute refractory period of the

afferent. A world which presented input stimuli from a uniformly random distribution over the entire hypervolume would be completely chaotic - a maximum entropy environment containing no information (snow on television and radio static are electronic examples of high entropy environments). The complexity of an environment's information content is reflected by the topological structure of the input environment. An environment may be any probability distribution over the entire stimulus space, continuous or discrete.

Sensory experience of all modalities presents the healthy organism with a highly structured set of stimuli. In the course of processing information, the network codes, recodes, integrates, and finally generates a response to these stimuli and it must learn to respond appropriately. The precise nature of the coding will not be addressed here. The stimulus sets used in computer simulation were, in general, "fuzzy" (having a small random component) subspaces of U that share certain topological properties with the physical stimuli used in the laboratory.

The order of presentation of patterns from the environment is given by a stochastic process, which may be either discrete or continuous. In the present thesis, a stationary jump process of independent presentations is assumed for most of the analysis and computer simulation. The trajectory profile imposed by the process is illustrated by averaging the system (5) over the stimulus environment E to generate a deterministic process (7) yielding

solutions for $(m_i(t),q(t))$ from various initial states $(m_i(t_o),q(t_o))$.

$$\langle \dot{m}_i(t) \rangle_E = \langle \phi(x(t),q(t)) d_i(t) \rangle_E$$
$$= \int_E \phi(x(t),q(t)) d_i(t)\ P(dE) \qquad (7)$$
$$\langle \dot{q}(t) \rangle_E = \langle \psi(x(t),q(t)) \rangle_E = \int_E \psi(x(t),q(t))\ P(dE)$$

Corrupt Environments

Generally a stimulus space is neither entirely noisy nor is it devoid of noise -- it contains information corrupted by random or uncorrelated interference. The separation of a signal from static can be a tedious task, but in general a repeated signal reinforces itself while noise interferes with itself destructively. Thus, systems such as those discussed in this thesis are resistant to noise given sufficient exposure to the stimulus environment. Interestingly, noise can enhance the development of neural networks (Anderson et al, 1977). This is discussed in Part IIa.

In the following sections, stimuli will be generated as the sum of patterned and chaotic components. The pattern space E can therefore be considered a k-dimensional manifold (k<N) in the N-dimensional space U. The environment or stimulus space D is the sum space of the pattern space and a more homogeneous process. One

can imagine the stimulus space as an N-dimensional "fuzz" encasing the pattern subspace (Figure 3).

Selectivity

A neural response function, c(m,E), can be calculated over the pattern set E for a given synaptic state m. Bienenstock (1980) gives a precise definition of selectivity for a general stimulus density in terms of such a density function. For the purpose of this presentation, this is paraphrased (8).

$$\mathrm{Sel}(m,E) \equiv \begin{cases} 1 - \dfrac{\int_E c(m,E)\ P(dE)}{c_{max}} & c_{max} > 0 \\ 0 & c_{max} = 0 \end{cases} \qquad (8)$$

$$\underline{\text{where}} \quad c_{max} \equiv \max_E(c(m,E))$$

Note that Sel $(m,E) \in [0,1]$ with 0 indicating a uniform response across E and 1 indicating that the entire set of evocative stimuli is of measure zero with respect to E.

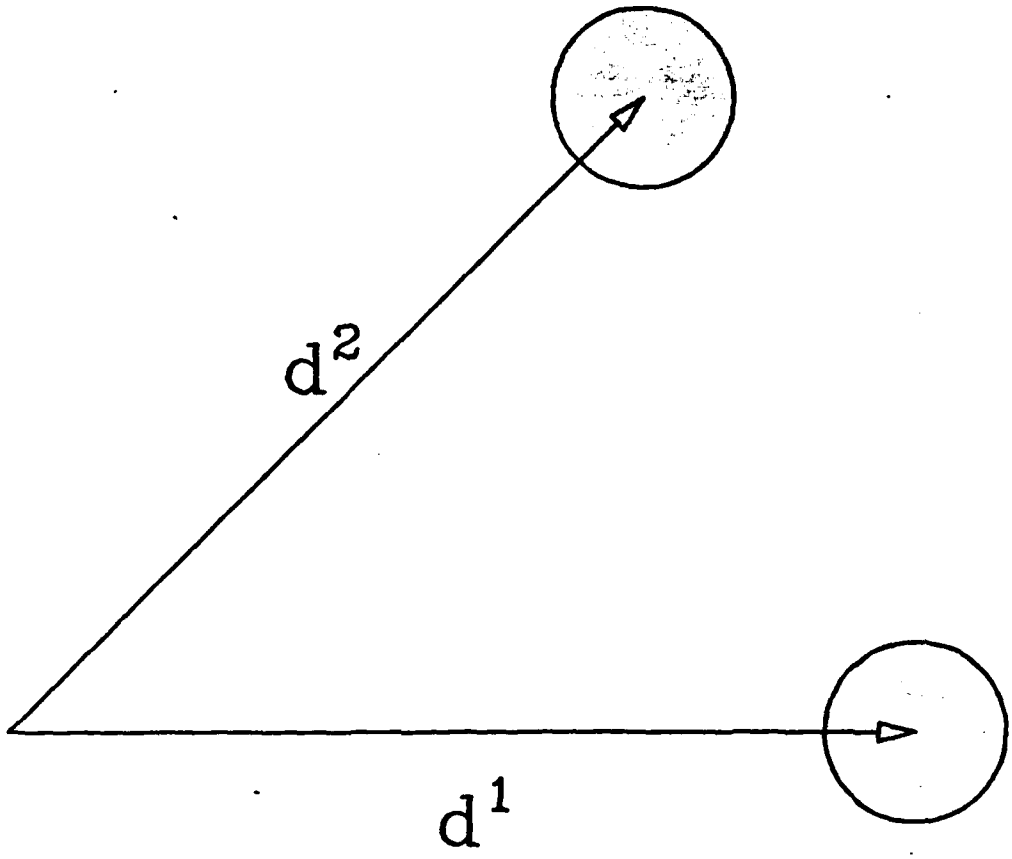


Figure 3. A two dimensional "fuzzy" stimulus space. The union of the shaded regions is a stimulus space consisting of a pattern space $\{d^1, d^2\}$ and a random vector uniformly distributed on a small circular area.

Circular and other periodic environments

A pattern set E is called circular if it describes a circle of uniform density embedded in a sphere about the origin. That is the elements d(w) of E must satisfy $|d(w)|$ = constant, $p(w) = \frac{1}{T}$ and $d(w_1) \cdot d(w_2) = f(w_2 - w_1)$ where w is a parameter of the period T and p(w) is a density function. A circular lattice pattern set is an analogous structure with finitely many elements equally spaced on a circle. The response characteristics of a synaptic state with respect to such one-parameter entities can be represented by a simple function of one variable or index from which the selectivity is easily evaluated (Figure 4).

Circular environments are idealizations of test-pattern sets used by physiologists to investigate orientation specificity development in visual cortex. The neural representations of oriented contrast bars of constant width certainly constitute a closed one-parameter (i.e., orientation) curve, if not a circle, in the axon frequency space of the optic nerve. Indeed, data from such experiments is plotted on a so-called orientation tuning curve much like the response curves in Figure 4. Thus, response data and selectivity measurements over them represent a point at which theory and experiment meet.

General and arbitrary circular environments of high dimension for computer simulation are not straightforwardly generated. This

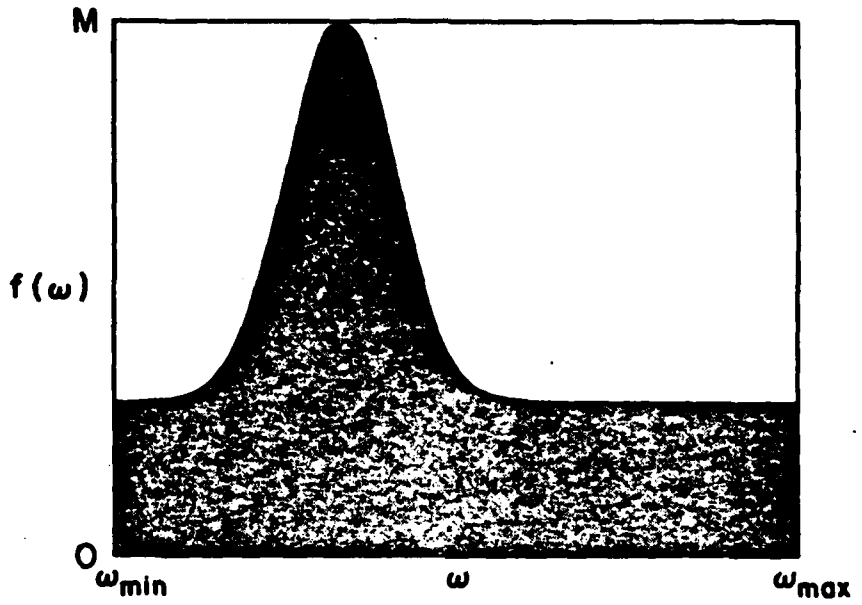


Figure 4. Selectivity. The fraction of the rectangle that is not shaded gives the selectivity of the response curve.

motivates retinal array pattern mapping (illustrated in Figure 5), a straightforward technique for simulating retinal coding of any spatiotemporal visual stimuli. An array is assembled of N units of uniform size and shape constituting the model retina, on which geometric stimuli are projected. The i-th component of the pattern vector is a measure of the "excitation density" of the i-th retinal unit. Retinal mapping distorts geometrical relationships between the visual stimuli, but fortunately the topology is conserved. Thus, applying uniform brightness patterns of constant width and intensity to such an array results in a pseudocircular pattern set for finite N.

Circular and pseudocircular environments are adequate for simulating one- parameter patterns coded by one retina. They suffice as well for cells receiving perfectly correlated patterns from several retinae. However more general treatments of multiocular coding of circular visual environments require that the destination (neural) map be a K-parameter product space. Binocular experimental paradigms invoking different stimulus sequences to the two eyes, such as strabismus or binocular disparity, can be simulated using toroidal or pseudotoroidal environments, sets of patterns characterized by two independent periodic parameters.

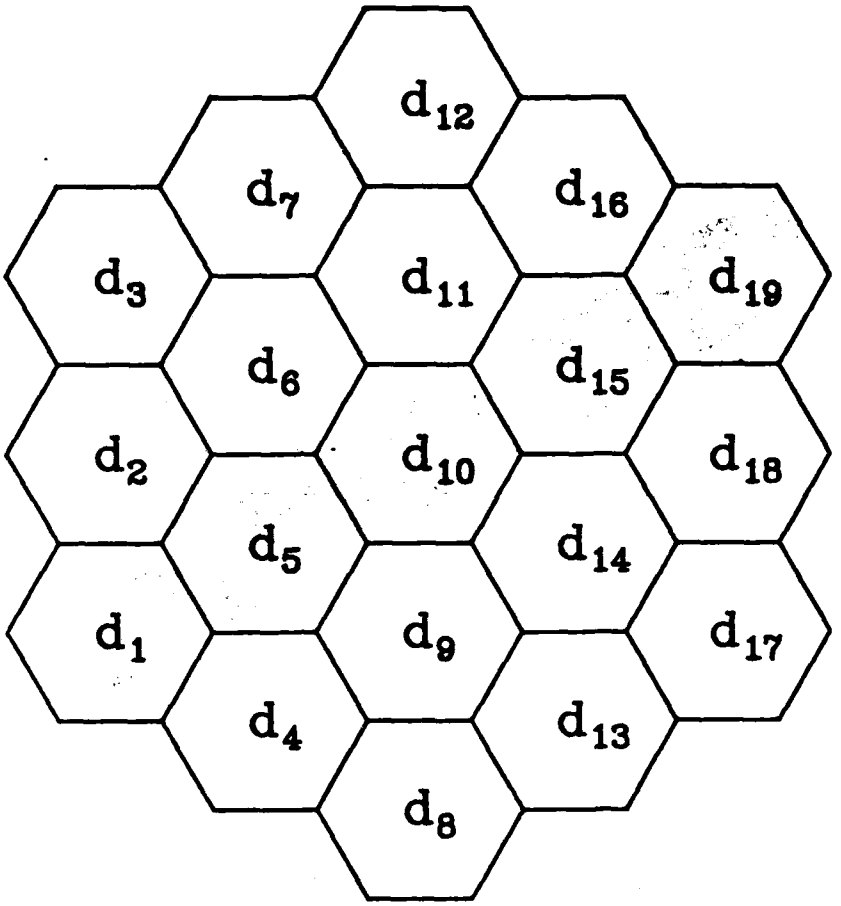


Figure 5. Retinal array pattern generation. The intersections of oriented bars with the cells in a simple geometric array was used to generate pattern vectors for computer simulation. The technique is inspired by the principle of the retina, but it is not intended to be a precise retinal model since the inputs are used for a model neuron in cortex.

C. Neuronal Abstraction

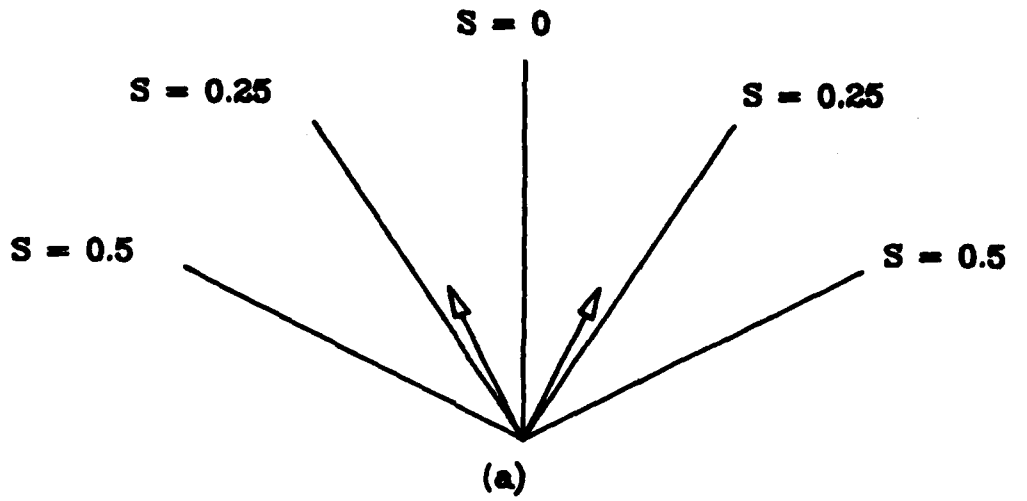
Each of the models presented in this thesis describes neurons of two kinds, which differ in the nature of the information they abstract from their stimulus environments. Neurons which seek to minimize their selectivity functions will be called similarity cells or S-cells, since they seek the pattern component most common to their environmental stimuli. The other kind of neuron, seeking maximum selectivity, will be labelled a difference cell or D-cell. These two functions should be considered elementary abstraction processes into which more complex functions can be analyzed.

Selectivity Maximization

Consider an environment of two stimuli represented by vectors d^1 and d^2. The selectivity $S(m,d^1,d^2)$ for the linear transfer function (4) is given by (9).

$$S(m,d^1,d^2) = 1 - \frac{m\cdot d^1 + m\cdot d^2}{2\max_i(m\cdot d^i)} \qquad (9)$$

Lines of constant selectivity can be plotted (Figure 6) in the plane spanned by two patterns. Maximally selective (S=1) states are orthogonal to one vector and not to the other. Minimally selective



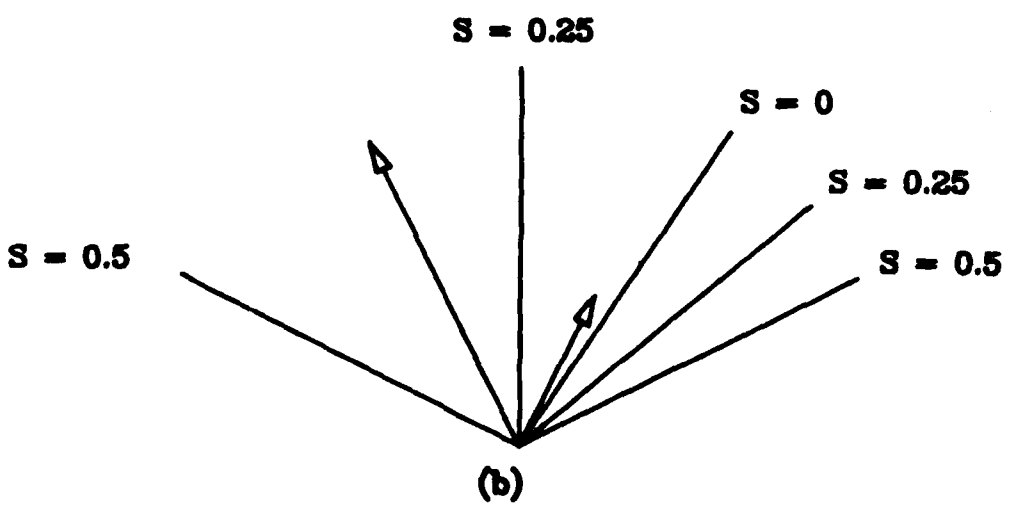


Figure 6. Curves of constant selectivity. (a) The two patterns are of equal length. (b) d^1 has a greater magnitude than d^2.

(S=0) states are orthogonal to the difference vector $d^2 - d^1$. For arbitrary S, m must satisfy (10).

$$[(2S - 1)d^i + d^j] \cdot m = 0$$
$$\underline{\text{where}} \quad m \cdot d^i \geq m \cdot d^j \qquad (10)$$

The goal of each model is to drive the states of S-cells to an S=0 point and D-cells to an S=1 point. Constraints on the synapses (such as the sign of m) may prevent a cell from realizing optimal selectivity. In such cases, the neuronal state will be driven to the constraint boundary.

Evidence for the development of neuronal abstraction properties in visual cortex

Single unit recording techniques were employed by Hubel and Wiesel (1959) to reveal the highly specific response properties of neurons in cat striate cortex. They found that these cells are sensitive to the orientation of contrast edges and bars in their visual fields. These neurons were shown (Wiesel and Hubel, 1963) to develop their selectivity attributes over the first ten weeks after birth -- the so-called critical period. The development can be influenced by manipulating the kitten's visual environment during the critical period. Results of various laboratory paradigms for such manipulations are now described.

Normal Rearing (NR): Kittens reared in normal visual environments develop (for the most part) neurons which are highly sensitive to orientation (Hubel and Wiesel, 1959, 1962). Layer IV cells[1] develop to be responsive to inputs from one eye. Other cells exhibit binocular responses which may favor one eye or the other. The preferences for ocularity and orientation vary topographically across cortex.

Monocular Deprivation (MD): Neurons in the visual cortex of a kitten, deprived of visual stimulus to one eye during the entire critical period by means of eyelid suturing, become generally nonresponsive to inputs presented to the deprived eye (Blakemore, 1976).

Reverse Suture (RS): The effects of the MD paradigm can be reversed during the critical period by opening the closed eye and closing the open one (Movshon, 1976).

Dark Rearing (DR): One would probably intuit, based on the NR and MD results, that binocular deprivation would cause complete non-responsiveness in visual cortex neurons. Moreover, while the ocular dominance statistics indicate a shift toward monocularity (Blakemore and Van Sluyters, 1975) many cells are binocularly responsive

[1] Nearly all layer IV neurons receive direct thalamic input (Hornung and Garey, 1981).

(Leventhal and Hirsch, 1980) and very few cells are non-responsive (Buisseret and Imbert, 1976).

Artificial Strabismus (AS): By cutting one of the extraocular muscles in newborn kittens, Hubel and Wiesel (1965) eliminated the normally high correlation between inputs from the two eyes. They found most cells developed monocularly.

Each of these experiments support the conjecture that some neurons seek maximum selectivity over their stimulus environments. The search for cells of low selectivity is less intense and hence the data less abundant. Nonetheless, cells in visual cortex have been classified as "circularly symmetric" (Hubel and Wiesel, 1977). Other authors (e.g., Kelly and van Essen, 1974) report difficulty in finding a cell's optimal stimulus. Barlow (1972) points out that many neurons display invariance in their response as a stimulus changes.

II. Generalization and Differentiation by Single Neurons: Two Approaches

A. The Fixed Point Model

In this section a version of the model put forward by Bienenstock et al (1982) for development of orientation specificity in visual cortex is described and then extended to include selectivity-minimizing neurons as well. Bienenstock's (1980) doctoral thesis includes formal theorems and detailed proofs in support of that theory. The extended model presented in this section will be called the fixed point (FP) model to distinguish it from the alternative formulations discussed in subsequent sections. The FP model (and those that follow as well) uses parallel modification and hence can be expressed in terms of (4) and (5).

Selectivity maximization

The function ϕ for the maximum-selectivity (D-cell) process must be continuous, bounded, and satisfy (11)

$$\text{sgn}(\phi_D(x,q)) = \text{sgn}(x)\ \text{sgn}(x - q^p) \quad \text{where } p > 1 \tag{11}$$

The rate of change of q(t) is given by (12). This is chosen so that q is a running time average over the activity x(t) (13).

$$\psi_D(x,q) = \beta(x - q) \qquad \beta > 0 \qquad (12)$$

$$q(t) = \exp[-\beta(t-t_0)]q(t_0) + \beta\int_{t_0}^{t} \exp[-\beta(t' - t_0)]x(t')\, dt' \qquad (13)$$

The (1+1) - Dimensional System

Consider a simple system of a single synapse m receiving a steady, noiseless signal d=1. The concept of selectivity is meaningless with respect to a one-element pattern environment, but aspects of the system dynamics are easily analyzed in this two-dimensional state space. The system is a completely deterministic pair of coupled autonomous equations for $\dot{m}$ and $\dot{q}$ (14).

$$\begin{aligned} \dot{m}(t) &= \phi_D(m(t),q(t)) \\ \dot{q}(t) &= \beta(m(t) - q(t)) \\ m(0) &> 0 \quad ; \quad q(0) > 0 \end{aligned} \qquad (14)$$

The critical points of the system lie at the intersection of the nullclines $\dot{m}=0$ and $\dot{q}=0$ (Figure 7), which partition the state space into regions characterized by the signs of the state velocity (sgn $\dot{m}$, sgn $\dot{q}$), the quadrant of $(\dot{m},\dot{q})$.

The asymptotic behavior of (14) is never divergent (Theorem 1). Depending on the value of β, the trajectory $(m(t),q(t))$ either converges to (1,1) or to a finite limit cycle (Figures 8a and 8b).

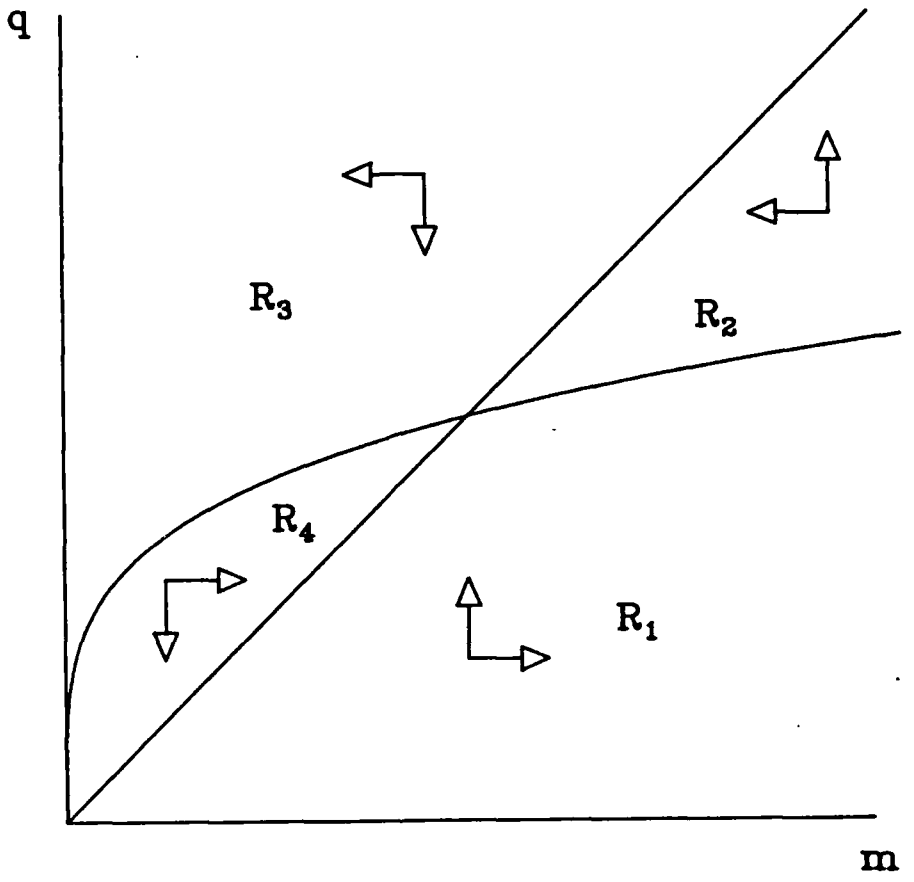
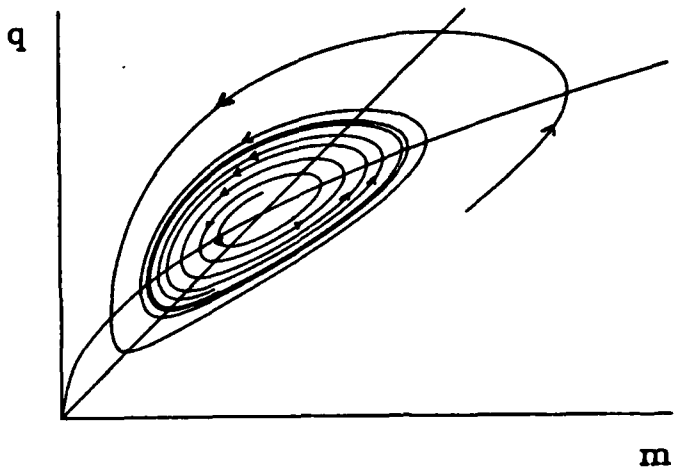
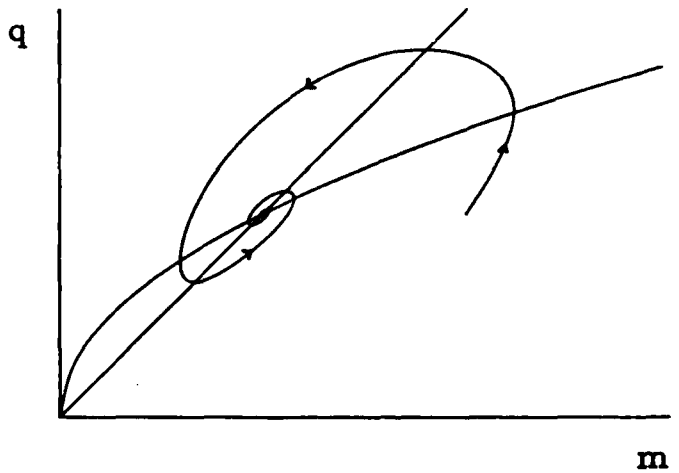


Figure 7. Critical points for the FP D-cell. The quadrant (sgn $\dot{m}$, sgn $\dot{q}$) is shown for each region R_1 - R_4.



(a)



(b)

Figure 8. Computer generated trajectories. The nullclines from Figure 7 are again shown along with sample trajectories for two values of β. (a) $\beta = 0.7$ gives limit cycle behavior -- two trajectories are shown: one spirals in from the outside and the other from the inside. (b) $\beta = 1.3$ gives convergence to the point (1,1).

Theorem 1 As $t \to \infty$, the system (14) drives the state to either the origin, the point (1,1), or to a closed path about that point.

Proof Since (14) is an autonomous system, $(\dot{m},\dot{q})$ is purely a function of (m,q) and therefore no trajectories cross at any point with the possible exception of a critical point ($(\dot{m},\dot{q}) = (0,0)$). Trajectories may only intersect at a critical point in the limit as $t \to \infty$.

Along the coordinate axes, (14) simplifies to (15) and (16)

$$\begin{aligned} \dot{m}(m=0) &= 0 \\ \dot{q}(m=0) &= -\beta q \end{aligned} \tag{15}$$

$$\begin{aligned} \dot{m}(q=0) &= \phi(m(t),0) \geq 0 \\ \dot{q}(q=0) &= \beta m \end{aligned} \tag{16}$$

Therefore if $m(0)>0$ and $q(0)>0$ then $m(t)>0$ and $q(t)>0$ for all $t>0$ (see figure 8). Now consider the four regions defined by the nullclines:

R_1 : (sgn $\dot{m}$, sgn $\dot{q}$) = (+1,+1)

R_2 : (sgn $\dot{m}$, sgn $\dot{q}$) = (-1,+1)

R_3 : (sgn $\dot{m}$, sgn $\dot{q}$) = (-1,-1)

R_4 : (sgn $\dot{m}$, sgn $\dot{q}$) = (+1,-1)

In R_1 $\dot{q}>0$, but a trajectory may not cross from R_1 to R_4. Since ϕ is bounded, a state in R_1 either asymptotically approaches (1,1) or is driven to R_2. A trajectory in R_2 must cross into R_3. In R_3, each state is driven toward the q-axis and toward R_4. However since the q-axis is a trajectory to (0,0), a state in R_3 must either enter R_4 or go asymptotically to (0,0).

A linear approximation (17) shows that if $\phi_m(0,0) > 0$, the origin is a saddle point, and hence, unstable. In this case, only the trajectory exactly along the q axis is not pushed into R_1.

$$\dot{m}(m\to 0,\ q\to 0) = \phi_m m$$
$$\dot{q} = \beta(m-q) \qquad (17)$$

The stability of the critical point at the origin depends on the derivative of ϕ and the magnitude of β. We will use a sufficiently well behaved and a sufficiently high value of β to assure convergence to (1,1).

The (2+1) dimensional D-cell

Now consider a cell with 2 synapses in an environment consisting of exactly two independent patterns $\{d^1,d^2\}$. This is the

simplest environment in which selectivity can be measured. The corresponding stochastic system of coupled ODE's (18) possesses 4 equilibrium points at which $(\underline{m}_i,q) = 0$ for both stimulus vectors.

$$\begin{aligned} \dot{m}_1(t) &= \phi_D(m\cdot\underline{d},q)\underline{d}_1 \\ \dot{m}_2(t) &= \phi_D(m\cdot\underline{d},q)\underline{d}_2 \\ \dot{q}(t) &= \beta((m\cdot\underline{d})-q) \end{aligned} \qquad (18)$$

for the random stimulus $\underline{d}\in\{d^1,d^2\}$

Since the input stimuli are independent, this condition is prerequisite to the more general deterministic criterion $\langle m_i,q\rangle_D=0$. Two of the equilibrium points have zero (minimum) selectivity. The other two each have selectivities of 0.5, the maximum for this environment. One of these is preferential to d^1, the other to d^2. A projection of the system to the m_1-m_2 plane (Figure 9) illustrates the situation. The point m = (1,1) is a saddle point and m=(0,0) is an unstable node. Only the points of maximum selectivity are stable. A detailed analysis of a very similar system is done by Bienenstock (1980).

Higher Dimensionalities

The (N+1) dimensional generalization (19) admits 2^N equilibrium points in an environment of N independent stimuli $\{d^i\}$.

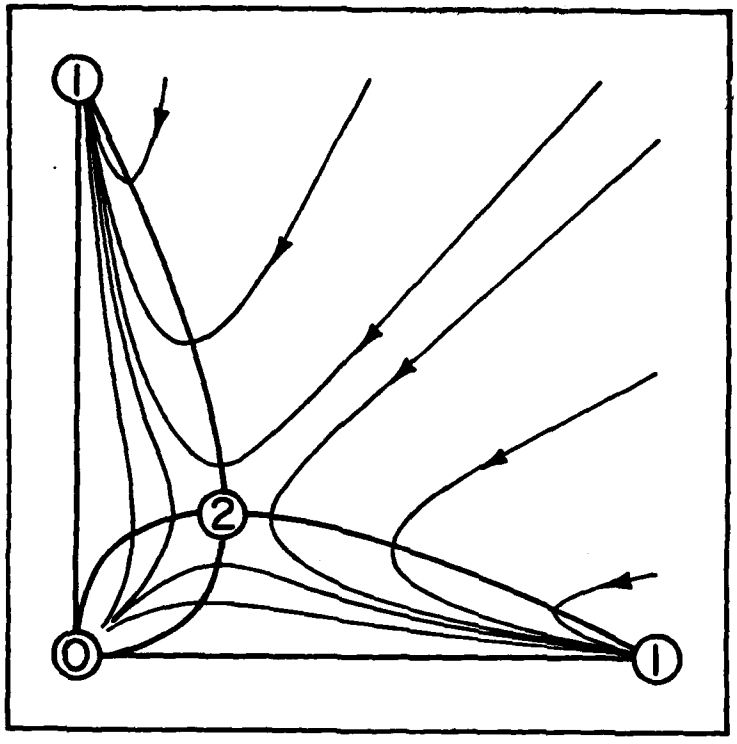


Figure 9. Projection of the (2+1) D-cell into the m-plane. The figure shows an idealized two-dimensional system in which q is the expected value of m•d evaluated over the entire environment. The parabolas are nullclines corresponding to each of the patterns: $m \cdot d^1 = q^2$ and $m \cdot d^2 = q^2$.

$$\dot{m}_i(t) = \phi_D(m \cdot \underline{d}, q)\underline{d}_i$$
$$\dot{q}(t) = \beta((m \cdot \underline{d}) - q) \qquad (19)$$

for the random stimulus $\underline{d} \in \{d^i\}$

Of these, N are maximally selective, with selectivity (N-1)/N, two (the origin and $(m_1,\ldots,m_n,q) = (1,\ldots,1,1)$) exhibit zero selectivity, and the remaining 2^N-N-2 display intermediate levels of selectivity. Each of the maximally selective points is "tuned" to a different one of the N patterns. It is conjectured on the strength of numerical simulations (Figure 10a) that only these points display asymptotic stability. This has been proven for a similar system in an orthogonal environment (Cooper et al, 1982).

Dependent Stimuli

Environments containing more than N stimuli also drive the model neuron (19) to high selectivity (Figure 10b). The maximum selectivity attainable is a function of the environment. In particular this is sensitive to the extent of separation between vectors in the pattern set (Bienenstock, 1980).

One can generally consider ϕ as a function that "reinforces" patterns giving a high response (x>q) and "suppresses" patterns that evoke low (x<q) activity (Figure 11).

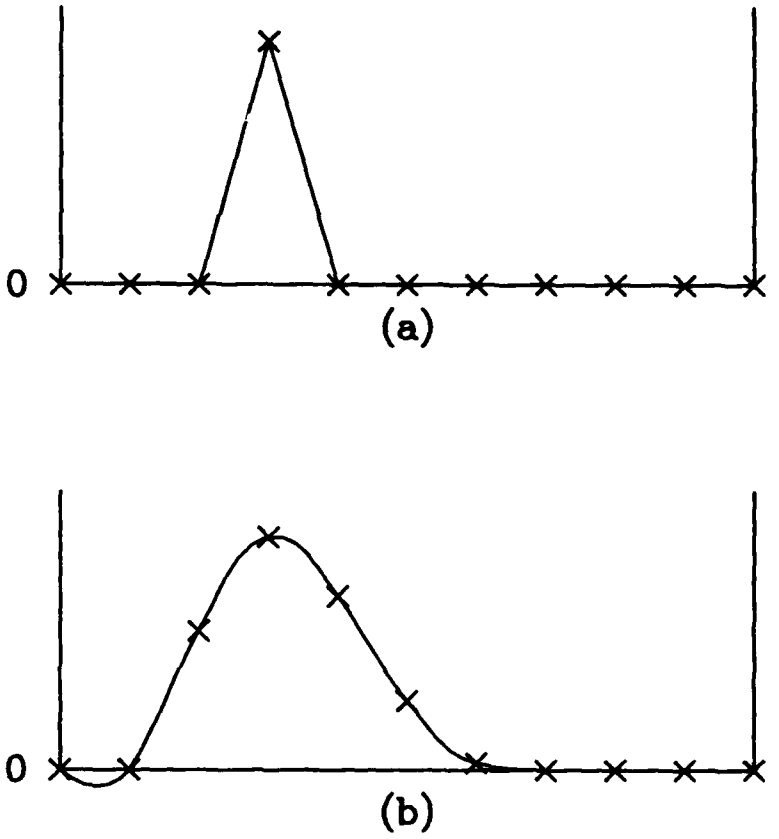


Figure 10. Tuning curves. Computer simulation shows that the system of differential equations given by (19) drive the model neuron to maximum selectivity. (a) Independent stimuli: all but one of the patterns give a null response in the final state. (b) Dependent stimuli: some patterns give intermediate responses.

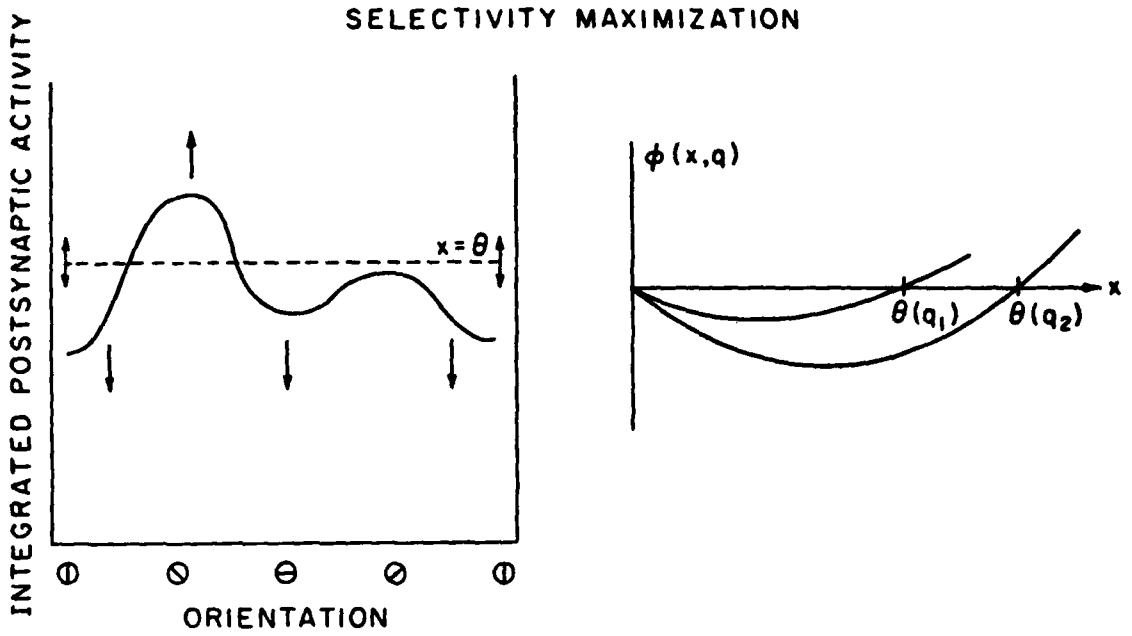


Figure 11. The D-cell in N dimensions. The function $\phi_D(x,q)$ is shown (right) plotted against x for two values of q. The effect of ϕ_D-modulated parallel modification is illustrated on the left.

Comparison of the model with various binocular interaction paradigms

The visual system is perhaps the best understood of mammalian sensory systems. This model is in good agreement with the bulk of data for binocular selectivity development in visual cortex. Below, the experimental paradigms described in Part Ic are discussed in terms of the model. Both analysis and computer simulation support the agreement between theory and experiment.

Normal Rearing (NR): The NR stimuli are represented as 2N-dimensional vectors in which the N left-eye components are highly correlated with the N components from the right eye. The environment is essentially circular since the patterns are periodic over a single parameter. Results from computer simulation are shown in Figure 12a.

Binocular Deprivation (BD): Patternless inputs ("pure" noise) drive the state of a model neuron through a random walk. Thus, the state wanders aimlessly from its initial condition. Figure 12b illustrates a typical result from computer simulation.

Monocular Deprivation (MD) and Reverse Suture (RS): Deprivation of patterned input to a single eye by way of eyelid suture results in a large shift in ocular dominance favoring the functional eye. The theoretical explanation for the MD and RS

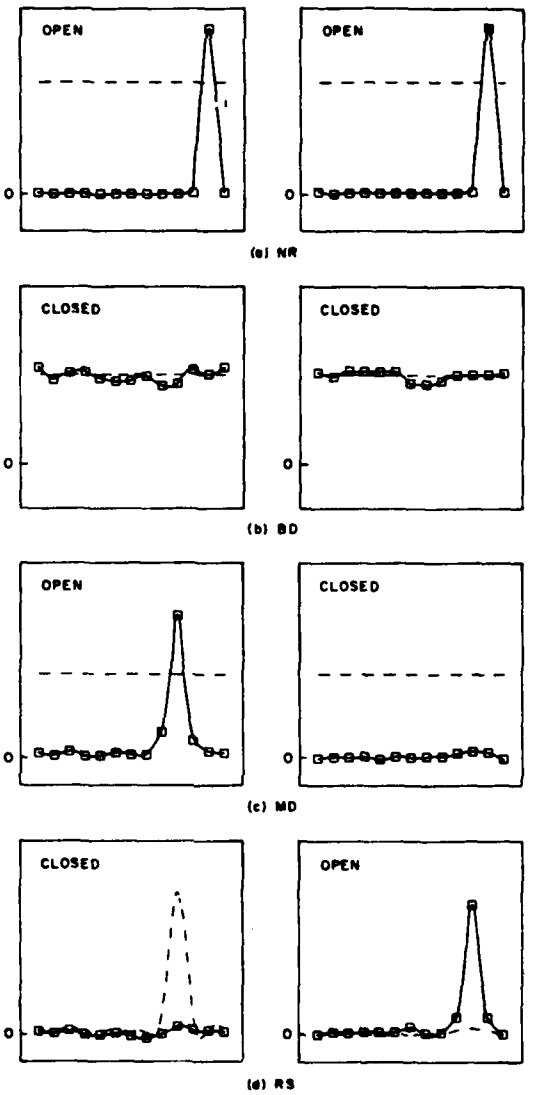


Figure 12. Computer simulation of four laboratory paradigms. See text for description.

phenomena is subtler than for NR and BD and is discussed below. The MD environment is represented by a 2N-dimensional vector generated by concatenating N noise components to an N-dimensional pattern (or pattern-plus-noise) vector. Note that computer simulation (Figures 12c and 12d) supports the success of the model in accounting for the experimental results.

Artificial Strabismus (AS): The AS environment is toroidal (or pseudo-toroidal), being periodic over two independent parameters, namely the left-eye orientation and the right-eye orientation. Thus the AS inputs are 2N-dimensional concatenations of independently selected patterns. Computer simulation suggests that only monocular states are stable. Figure 12c (MD) reflects a typical AS simulation. Occasional trials show binocular selective states (the preferred orientations of the two eyes appear to be uncorrelated) that survive for some time but eventually become monocular. Figure 13 shows the final state plotted over a toroidal environment.

Orientation Selectivity and Ocular Dominance

Detailed examination of theoretical behavior in an MD environment highlights a subtle relationship between orientation selectivity and ocular dominance. Here it is explained why on the one hand deprivation drives synapses to zero efficacy in an MD environment, and on the other hand deprived synapses fluctuate randomly in a BD situation. It is shown that the efficacy of

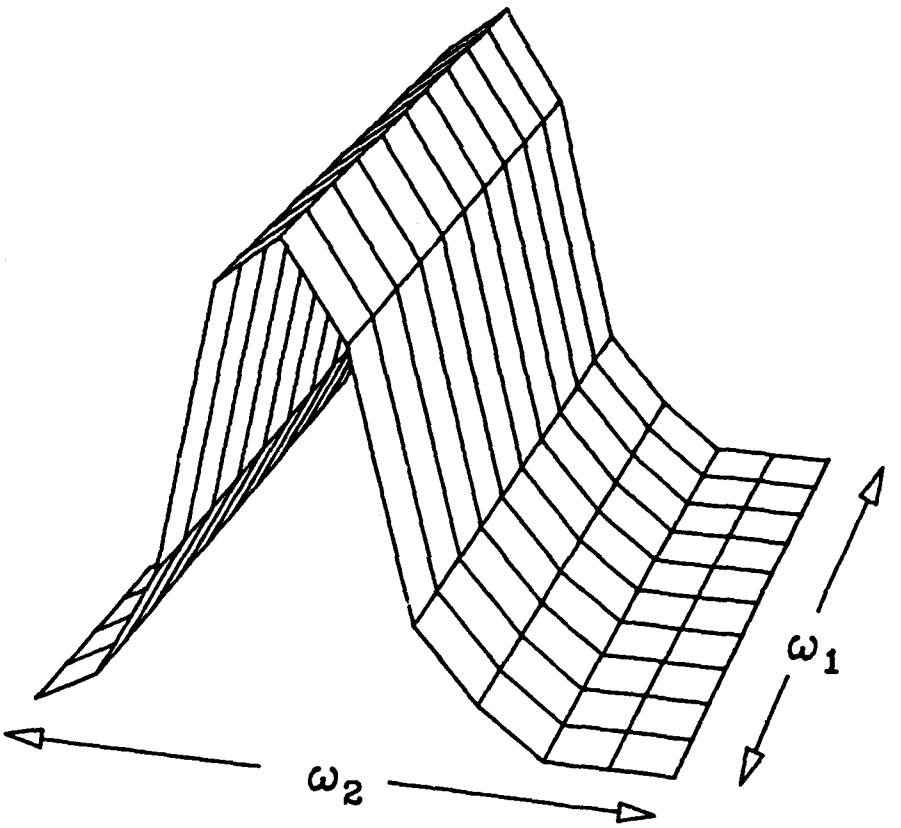


Figure 13. Tuning of a D-cell in a toroidal environment. The neuron tunes along only one parameter. This environment is analogous to an artificial strabismus paradigm.

deprived synapses decreases exponentially if the neuron is highly selective over the environment of the opposite eye. Hence it is predicted that in an MD environment, selectivity development in the open eye must precede loss of responsiveness in the closed eye (Figure 14).

Consider a neural state (m_ℓ^*, m_r^*, q), where m_ℓ^* and m_r^* are synaptic vectors corresponding to left-eye and right-eye inputs respectively, which is driven by patterned input to the left eye ($d_\ell(t) = d(\omega(t))$ periodic over ω) and pure noise to the right eye ($d_r(t) = n(t)$ uniformly distributed about 0). In general, the state m_r^* fluctuates randomly and m_ℓ^* is driven to a high-selectivity equilibrium state (m_ℓ^*, m_r^*, q) where (20) is satisfied.

$$\langle \phi(m_\ell^* \cdot d_\ell, q^*)\ d(\omega) \rangle_\omega = 0 \tag{20}$$

The expectation values of m_ℓ and m_r over the entire environment $D(\omega, n)$ can be calculated:

$$\begin{aligned} \langle \dot{m}_\ell \rangle_{\omega,n} &= \langle \phi(m_\ell^* \cdot d(\omega) + m_r^* \cdot n, q^*) d(\omega) \rangle_{\omega,n} \\ &= 0 + \langle \phi_1(m_\ell^* \cdot d(\omega), q^*)[m_r^* \cdot n] d(\omega) \rangle_{\omega,n} \\ &= 0 \end{aligned} \tag{21}$$

$$\begin{aligned} \langle \dot{m}_r \rangle_{\omega,n} &= \langle \phi(m_\ell^* \cdot d(\omega) + m_r^* \cdot n, q^*) d(\omega) \rangle_{\omega,n} \\ &= \langle \phi_1(m_\ell^* \cdot d(\omega), q^*)[m_r^* \cdot n] n \rangle_{\omega,n} \\ &= \langle \phi_1(m_\ell^* \cdot d(\omega), q^*) \rangle_\omega \ \langle (m_r^* \cdot n) n \rangle_n \end{aligned} \tag{22}$$

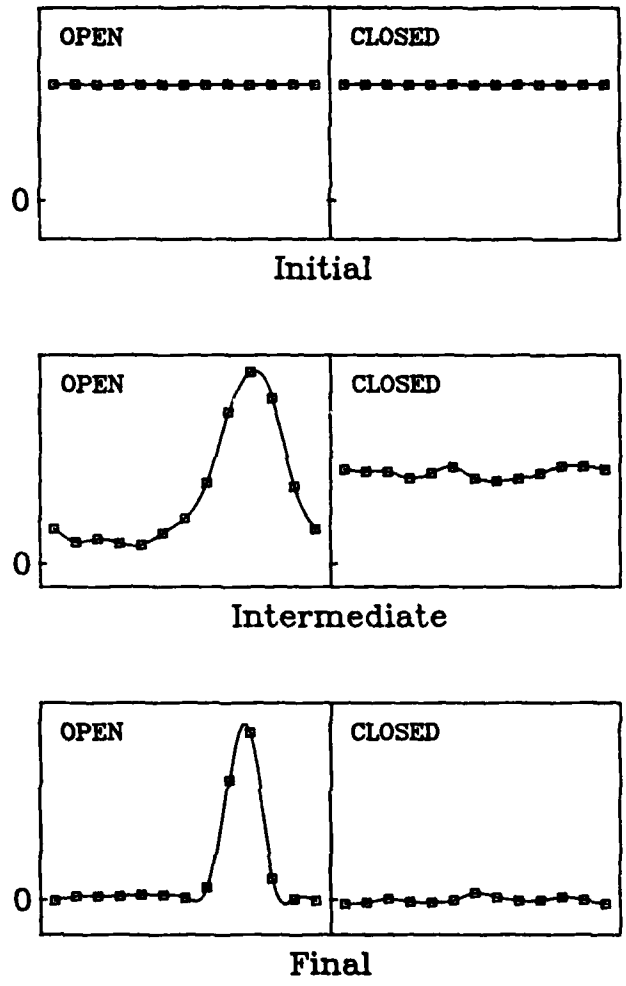


Figure 14. Monocular deprivation sequence. Computer simulation illustrates the model's prediction that the neuron must first become highly selective to patterns from the open eye before that eye dominates the neuron.

The expectation value $\langle(m_r \cdot n)n\rangle_n$ is parallel to m_r. For an N-cube with each component n_i varying between -a and +a:

$$\langle(m_r \cdot n)n_i\rangle_n = \sum_j \frac{1}{(2a)^N} \int_{-a}^{+a} \cdots \int_{-a}^{+a} m_j^r n_j n_i \, dn_1 \cdots dn_N$$

Consider a constituent integral over n_k:

$$\int_{-a}^{+a} m_j^r n_j n_i \, dn_k = \begin{cases} 2am_j^r n_j n_i & \text{if } i \neq k \text{ and } j \neq k \\ 0 & \text{if } i=k \text{ or } j=k \text{ but not both} \\ \frac{2}{3}a^3 m_i^r & \text{if } i = j = k \end{cases}$$

Hence:

$$\langle(m_r \cdot n)n_i\rangle_n = \frac{1}{(2a)^N} (2a)^{N-1} \frac{2}{3}a^3 m_i^r = \frac{a^2}{3} m_i^r \qquad (23)$$

A similar calculation is done for n distributed on an N-sphere:

$$\langle(m_r \cdot n)n\rangle_n = n_{m_r} |m_r| n_{m_r} \frac{m_r}{|m_r|} \, n = \langle n_{m_r}^2 \rangle_n \, m_r \qquad (24)$$

By spherical symmetry:

$$\langle n_{m_r}^2 \rangle_n = \frac{1}{N}\langle n \cdot n\rangle_n = \frac{1}{N} \frac{1}{V(N,a)} \int_V n \cdot n \, d^N n$$

$$= \frac{1}{N} \frac{1}{V(N,a)} \int_V x^2 \, dV$$

Since $V(N,x) = V_N x^N$, $dV = NV_N x^{N-1} dx$

$$\langle n_{m_r}^2 \rangle_n = \frac{1}{V(N,a)} \int_0^a V_N x^{N+1} \, dx = \frac{1}{V_N a^N} \frac{V_N a^{N+2}}{N+2} = \frac{a^2}{N+2}$$

$$(m_r \cdot n)n_n = \frac{a^2}{N+2} \, m_r \qquad (25)$$

Thus by (22) one sees that $\langle \dot{m}_r \rangle_{\omega,n}$ is parallel to m_r and that if $\langle \phi_1 \rangle_\omega$ is negative, then m_r decays exponentially. Since ϕ_1 is negative when $m_\ell \cdot d_\ell(\omega)$ is near zero, which is true over most of the environment when m_ℓ is a highly selective state, such a state should satisfy $\langle \phi_1 \rangle_\omega < 0$. Low selectivity states are thus unstable in two senses: (1) m_ℓ is unstable with respect to d_ℓ and (2) $\langle \phi_1' \rangle$ may be positive thereby driving m_r to higher magnitude.

Selectivity Minimization

The functions ϕ and ψ satisfying (26) drive the model neuron (5) to minimum selectivity.

$$\begin{aligned} \operatorname{sgn}(\phi_S) &= -\operatorname{sgn}(x)\cdot\operatorname{sgn}(x - q^p) \\ \psi_S &\equiv \beta(x-q) \quad ; \quad \beta > 0 \\ &\underline{\text{where}} \quad 0 < p < 1 \end{aligned} \tag{26}$$

Note $\psi_S = \psi_D$ and that (11), (12) and (26) can be reformulated into a single model (27).

$$\begin{aligned} \operatorname{sgn}(\phi^{FP}(x,q)) &= \alpha \operatorname{sgn}(x)\ \operatorname{sgn}(x - q^p) \\ \psi^{FP}(x,q) &\equiv \beta(x-q) \quad ; \quad \beta > 0 \end{aligned} \tag{27}$$

For $\alpha>0$ and $p>1$, the cell seeks maximum selectivity; for $\alpha<0$ and

$0<p<1$, the cell seeks minimum selectivity; otherwise, the system diverges.

The (1+1) Dimensional System

Again consider a one-synapse cell driven by a steady, clean signal. As in the case of the D-cell, the S-cell equation (26) reduces to a deterministic, autonomous system (28). All trajectories converge to (1,1) unless (0,0) is the initial point.

$$\begin{aligned} \dot{m}(t) &= \phi_S(m(t),q(t)) \\ \dot{q}(t) &= \beta\,(m(t),q(t)) \end{aligned} \tag{28}$$

The origin is unstable for all ϕ_S and there are no limit cycles (Theorem 2).

Theorem 2. With the exception of $(m(0),q(0)) = (0,0)$, all trajectories driven by (28) go to $(m(t),q(t)) \to (1,1)$ as $t \to \infty$.

Proof. As in theorem 1, consider four regions (Figure 15).

$$(\operatorname{sgn} \dot{m},\ \operatorname{sgn} \dot{q}) = \begin{cases} (+1,\ +1) : R_1 \\ (-1,\ +1) : R_2 \\ (-1,\ -1) : R_3 \\ (+1,\ -1) : R_4 \end{cases}$$

The point is monotonically driven to satisfy the simultaneous

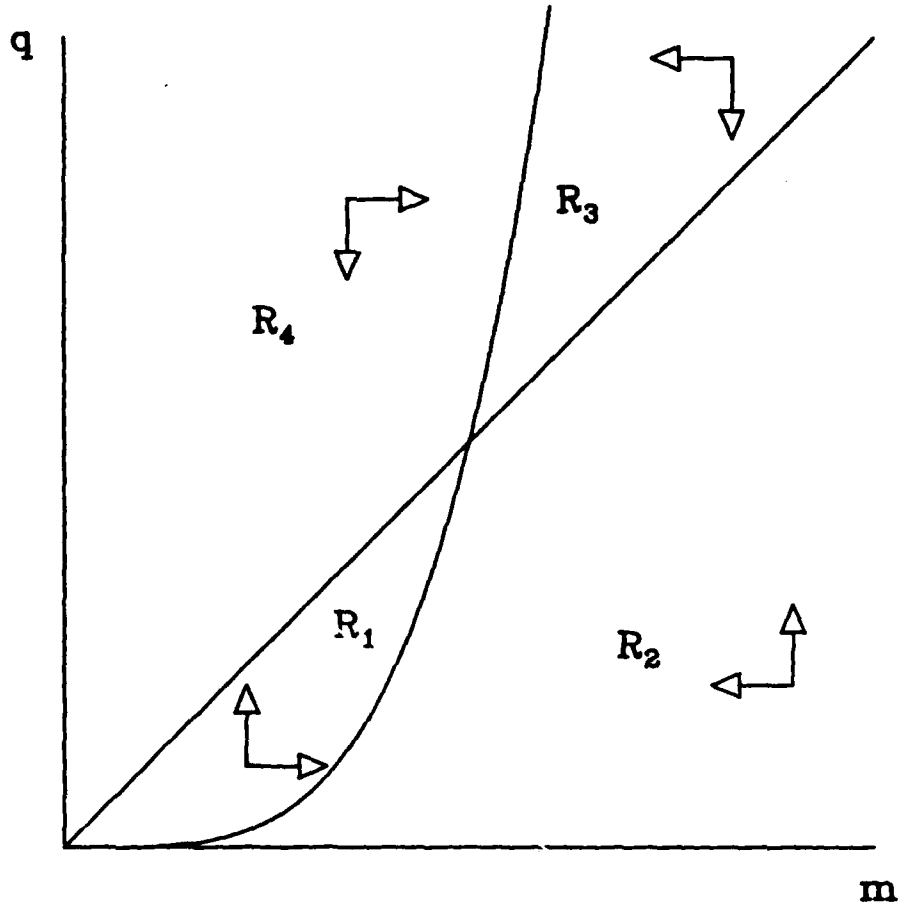


Figure 15. (1+1) - dimensional representation of an S-cell. The regions R_1 - R_4 are shown for an S-cell. In each region the quadrant (sgn $\dot{m}$, sgn $\dot{q}$) is illustrated.

conditions $\dot{m}=0$, $\dot{q}=0$ and it is always driven away from the axes $m=0$ and $q=0$. (QED)

Higher Dimensionalities

A two synapse cell in a two-pattern (independent) environment is driven to $(m_1,m_2,q) = (1,1,1)$, a zero selectivity state, with the exception (in a noiseless environment) of the initial point $(m(0),m_2(0),q(0)) = (0,0,q_0)$. The other state of zero-selectivity, namely the origin, is always unstable. The highly stable asymptotic behavior of the (1+1)-dimensional system is maintained in (2+1) dimensions (Figure 16).

Due to the parallel aspect of the modification, patterns driving the cell below q^p are "encouraged" and patterns yielding $x > q^p$ are "discouraged" (Figure 17). This, combined with the influence of q, causes the responses to the various stimuli to converge toward a common value. This can be achieved for sets of independent patterns (Figure 18a). For dependent pattern sets it may be possible, but not in general (Figure 18b). Thus, qualitatively, in the sense of Gati and Tversky (1982), the synapse vector seeks the pattern component most common in its environment.

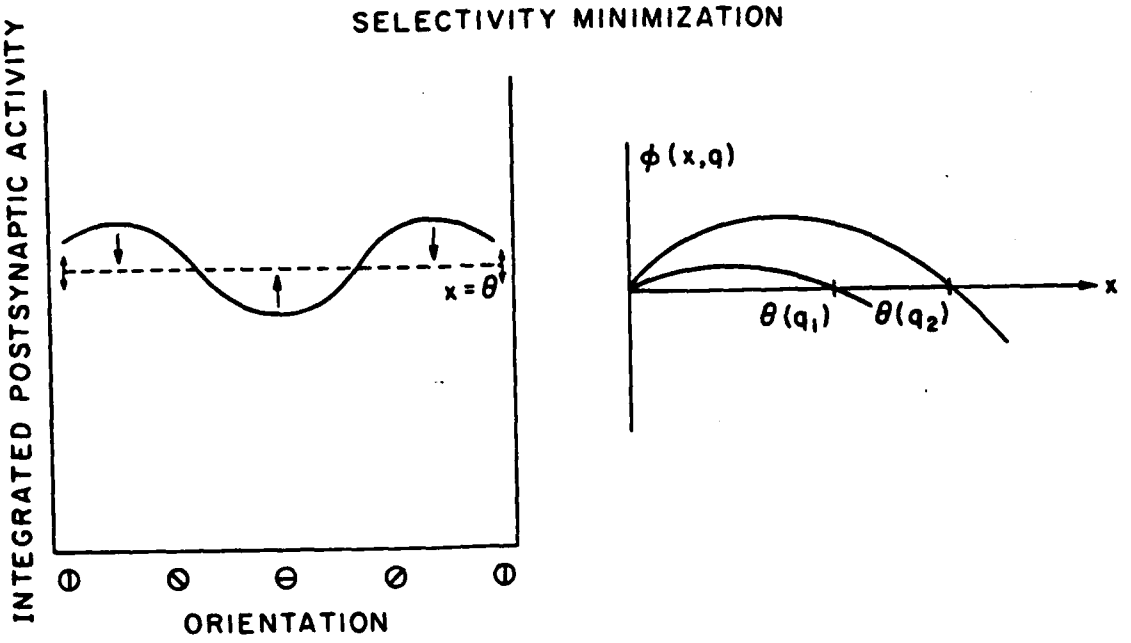


Figure 17. The S-cell in N dimensions. The function $\phi_S(x,q)$ is shown (right) plotted against x for two values of q. The effect of ϕ_S-modulated parallel modification is illustrated on the left.

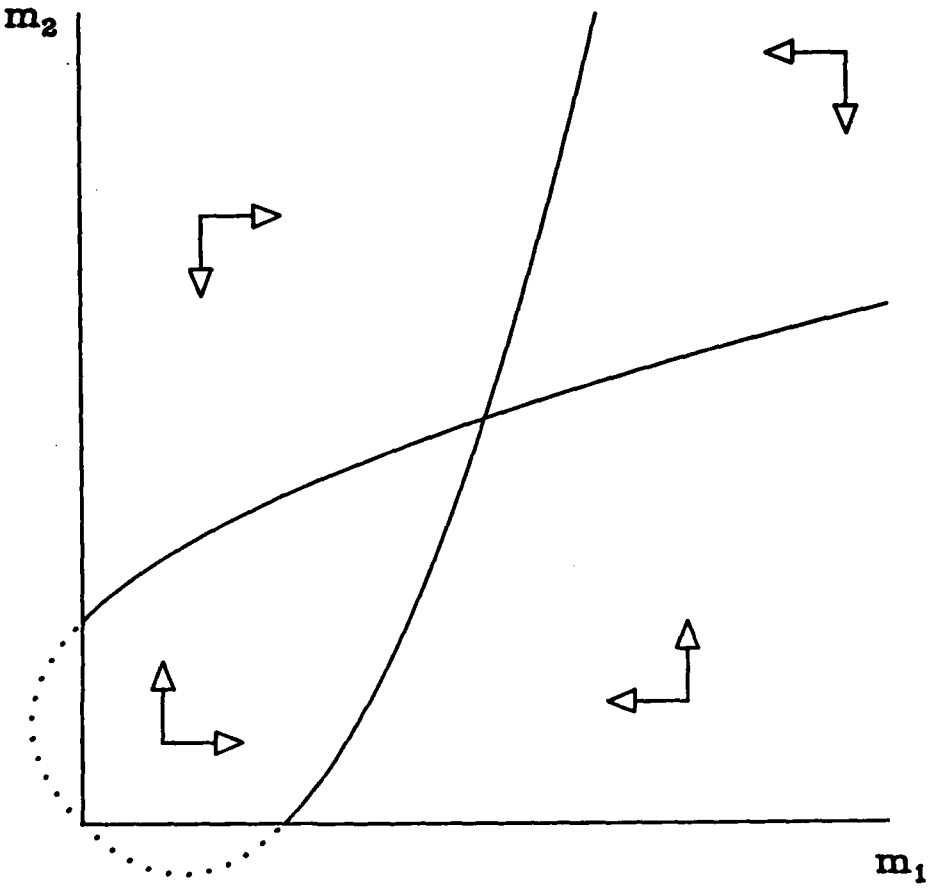


Figure 16. Projection of the (2+1) S-cell into the m-plane. As in Figure 9, q is idealized to be the expected value of m•d over the orthogonal environment ((1,0) , (0,1)). In this figure p = 0.5 and both nullclines (parabolas) are attractors.

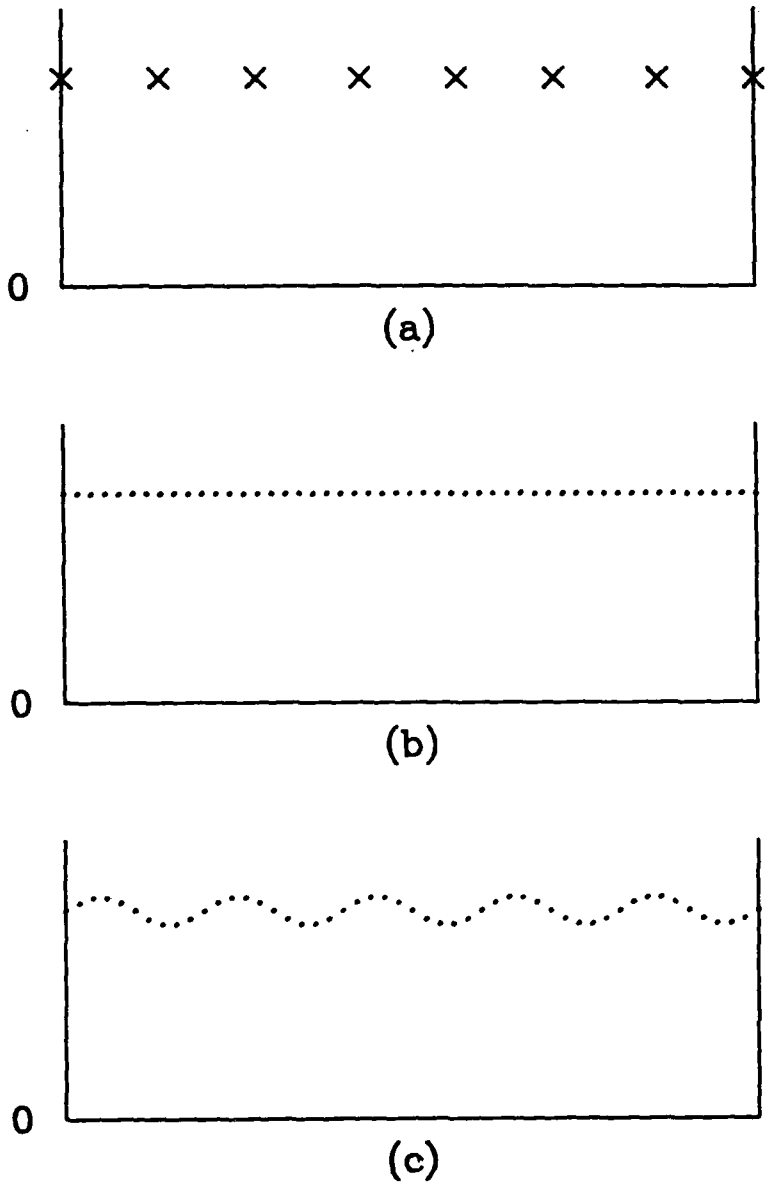


Figure 18. Selectivity minimization. (a) An environment of 7 independent patterns. (b) 50 patterns in 7 dimensions. (c) Noncircular 50 pattern environment in 7 dimensions.

B. An (N+1) - parallel Model.

In this section, an (N+1)-parallel model is presented (29). Again $\phi(x,q)$ is bounded and continuous. The function $\theta(q)$ gives the value of x for which the signs of ϕ and ψ change.

$$\begin{aligned} \operatorname{sgn}(\phi) &= \alpha \operatorname{sgn}(x) \operatorname{sgn}(x - \theta(q)) \\ \psi &= \alpha \beta \phi(x,q) x \\ \alpha &= \pm 1 \quad ; \quad \beta > 0 \end{aligned} \tag{29}$$

It need not be specified as a power function as was done in the FP model, however it is subject to certain restrictions. The α parameter determines whether the cell generalizes (α=-1) or seeks maximum selectivity (α=+1). Not only is this model parallel in the sense of (5) ($\dot{m} \| d$), but the change in the neuronal state $(\dot{m},\dot{q})$ is parallel to the (N+1) - activity vector $(\alpha d, \beta x)$. Thus the system can be expressed more concisely in terms of an "(N+1) formalism" (30). The model is given in a general form in that $\theta(q)$ is not specified, so that overall properties of the system (30) can be examined. A more specific example is presented in Part IV.

$$\begin{pmatrix} \dot{m}_i \\ \dot{q} \end{pmatrix} = \phi(x,q) \begin{pmatrix} \alpha d_i \\ \beta x \end{pmatrix} \tag{30}$$

where $\operatorname{sgn}(\phi(x,q)) = \operatorname{sgn}(x) \operatorname{sgn}(x - \theta(q))$

$\alpha = \pm 1 \quad ; \quad \beta > 0$

$\theta(q)$ nondecreasing & continuous

(1+1) - Dimensions

The one synapse - one pattern situation is examined first. Synapse m receives a noncorrupted steady signal $d \equiv 1$, and so $x \equiv m$ giving (31).

$$\begin{pmatrix} \dot{m} \\ \dot{q} \end{pmatrix} = \Phi(m,q) \begin{pmatrix} \alpha \\ \beta m \end{pmatrix} \tag{31}$$

$$\alpha = \begin{cases} +1 : \text{D-cell} \\ -1 : \text{S-cell} \end{cases}$$

The trajectories in the m-q plane are seen to be parabolas (32) since (31) gives the path derivative $\frac{dq}{dm} = \alpha\beta m$ (Figure 19).

$$q = q_0 + \frac{\alpha\beta}{2}(m^2 - m_0^2) \tag{32}$$

For the S-cell, the conditions on $\theta(q)$ are that it be continuous, nondecreasing and positive on some domain $[q_a, q_b]$. If $q_a > 0$, then $\theta(q_a) = 0$ and $\Phi(x, q_a) > 0$ for all x. Remember that $x \in [0, \infty)$. If q_b is finite then $\theta(q_b)$ is not and $\Phi(x, q > q_b) < 0$ for all x.

The D-cell needs additional restrictions on $\theta(q)$ for global stability of (31). If $q(t=0)$ is very large, the synapses will "die" - i.e., (m,q) converges to a point on the q-axis. This is not a serious problem. If necessary, restrictions can be put on the initial state. Possible divergence of the system is more problematic. If $\theta(q)$ does not increase sufficiently fast as q gets

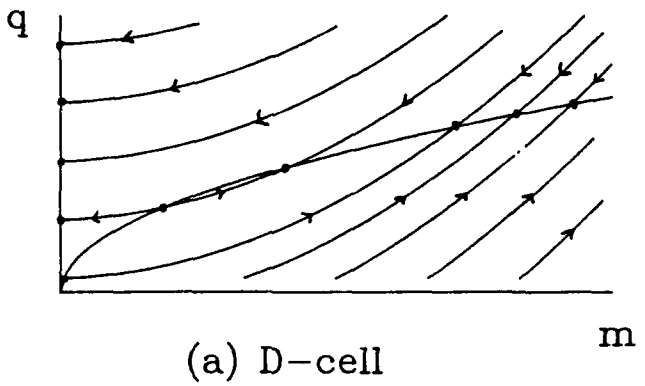
large, some trajectories (m(t),q(t)) will grow without bound. Therefore, we impose (33) on $\theta(q)$. Note that there are many functions which can be used for both cell types.

$$\lim_{q\to\infty} \frac{q^{1/2}}{\theta(q)} = 0 \qquad (33)$$

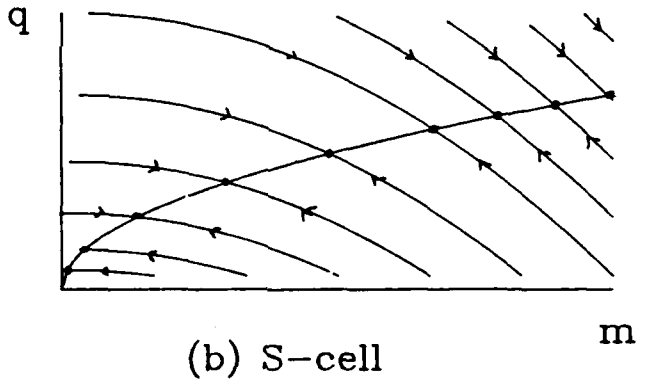
Unlike the FP model, this system does not have a finite number of isolated equilibrium points. Instead, the equilibrium points define a locus consisting of the q-axis (m=0) and the curve $m=\theta(q)$. All points satisfying $m=\theta(q)$ are the attractors with the possible exception of $(0,q_a)$. Points on the q-axis can be either stable (attractors) or unstable equilibria. For $\alpha q>\alpha q_a$ they are stable, otherwise they "repel" states.

Stability in noisy environments

Isolated attractors are generally stable in environments corrupted by a not-too-large noise level. However, continuous equilibrium loci are subject to noi e-driven "creep" effects. Chaotic perturbations may preferentially drive the neuronal state along the equilibrium surface. Consider a one-synapse cell receiving a signal $d(t) = 1+ \varepsilon(t)$ where $\varepsilon(t)$ is a stochastic variable uniform on [-a,a] where $a<<1$. The stability of the system (34) can be analyzed by expanding about an equilibrium point (m_0,q_0) where $m_0 = \theta(q)$.



(a) D−cell



(b) S−cell

Figure 19. Parabolic trajectories. The (1+1) neuron follows a parabolic path in the m-q plane for any initial point. The function $\theta(q)$ determines at what point along the trajectory the path direction changes. Here $\theta(q) = q^2$. (a) D-cell. (b) S-cell.

$$\begin{pmatrix} \dot{m} \\ \dot{q} \end{pmatrix} = \Phi(m+m\varepsilon(t),q) \begin{pmatrix} \alpha + \alpha\varepsilon \\ \beta m + \beta m\varepsilon \end{pmatrix} \tag{34}$$

The effect of noise on the stability of q-axis equilibrium points is not relevant to this thesis. The expected value of $(\dot{m}, \dot{q})$ is calculated over ε in (35). For either value of α, the expected perturbation direction is along a trajectory and hence the neuronal state tends to return to (m_0, q_0).

$$\begin{aligned} \left\langle\!\left\langle \begin{pmatrix} \dot{m} \\ \dot{q} \end{pmatrix} \right\rangle\!\right\rangle_{\varepsilon} &= \frac{1}{2a} \int_{-a}^{+a} \Phi_m(m_0,q_0)\ m_0\varepsilon \begin{pmatrix} \alpha(1+\varepsilon) \\ \beta m_0(1+\varepsilon) \end{pmatrix} d\varepsilon \\ &= \Phi_m\ m_0 \begin{pmatrix} \alpha \\ \beta m_0 \end{pmatrix} \frac{1}{2a} \int_{-a}^{+a} \varepsilon\ (1+\varepsilon)\ d\varepsilon \\ &= \frac{a^2}{3} \Phi_m\ m_0 \begin{pmatrix} \alpha \\ \beta m_0 \end{pmatrix} \end{aligned} \tag{35}$$

(2+1) - Dimensions

The selectivity properties of this (N+1)-parallel model for a two-synapse cell in a two-pattern environment $\{d^1, d^2\}$ resemble those of the FP model. However, we again have a locus of connected, rather than isolated, equilibrium points. These points lie along four curves intersecting at the point $(m_1, m_2, q) = (0, 0, q_a)$. The four curves correspond to the following intersections of surfaces (Figure 20):

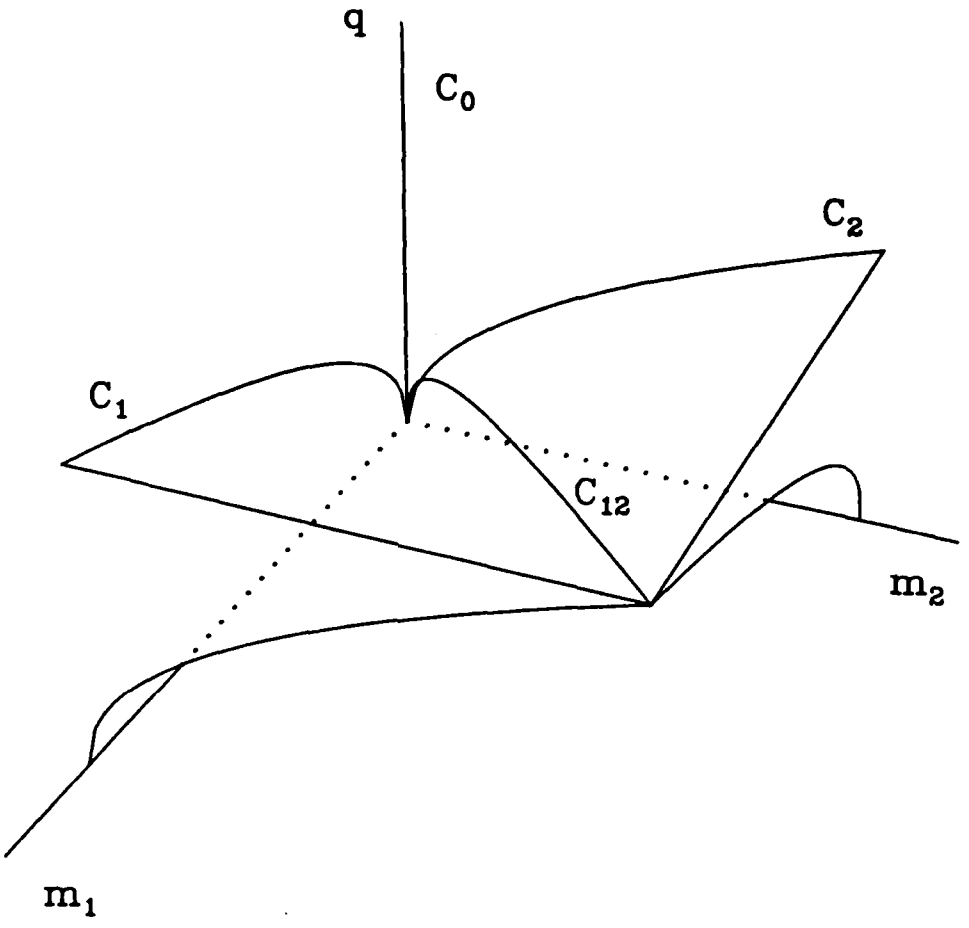


Figure 20. (2+1) - space. The four equilibrium curves and the four regions defined by the surfaces $m \cdot d = \theta(q)$ and $m \cdot d = 0$ for $d \in \{d^1, d^2\}$ are shown for $\theta(q) = q^2$ in an orthogonal environment.

C_0 : $m \cdot d^1 = m \cdot d^2 = 0$ (the q-axis)

C_1 : $m \cdot d^1 = \theta(q)$; $m \cdot d^2 = 0$ } (maximum selectivity)

C_2 : $m \cdot d^1 = 0$; $m \cdot d^2 = \theta(q)$ }

C_{12}: $m \cdot d^1 = m \cdot d^2 = \theta(q)$ (minimum selectivity)

The surfaces partition the space into four regions, according to the sign of Φ resulting from each of the two patterns. For $m \cdot d^1 > \theta(q)$ and $m \cdot d^2 > \theta(q)$, q increases until the one or both patterns cannot stimulate the neuron above $\theta(q)$. Neuronal states in the opposite region ($m \cdot d < \theta(q)$ for both patterns) are likewise forced out of that region. Therefore the state generally finds itself in a region "between" the surfaces $m \cdot d^1 = \theta(q)$ and $m \cdot d^2 = \theta(q)$ or, in the case of the D-cell, possibly on the q-axis (if q(t=0) is "too large").

Once the trajectory is between the two surfaces, that is for $m \cdot d^i < \theta(q) < m \cdot d^j$ where $i \neq j$, its behavior depends on α. For α=+1, the trajectory converges to some point on C_j and hence the neuron becomes maximally selective whereas for α=-1 the trajectory goes to a minimally selective point on C_{12}.

Higher Dimensionalities

A linearly independent environment of N patterns admits 2^N equilibrium curves in the (N+1) - state space of the neuron. These curves lie at the intersections of the surfaces $m \cdot d^i = \theta(q)$ and

$m \cdot d^i = 0$ for $i = 1, \ldots, N$. Of these, two (one trivial and one non-trivial) represent states of minimum selectivity and N represent states of maximum selectivity -- one per pattern.

The behavior of the model in more complex environments is difficult to analyze. Numerical techniques suggest that for circular environments, *D*-cells seek maximum selectivity and S-cells seek minimum selectivity. Toroidal environments have also been simulated for which the pattern depends on two stochastic parameters ω_1 and ω_2. Computer results indicate that a D-cell tunes with respect to one parameter only (see Figure 13) and that S-cells seek a uniform response across the entire environment.

III. Construction of an (N+1)-parallel Model using Antagonistic Mechanisms

A function satisfying the conditions on Φ in (30) can be constructed (36) as a q-coupled difference of two non-negative monotone increasing saturating functions satisfying certain criteria (Figure 21).

$$\Phi(x,q) \equiv \sigma_2(x) - q\sigma_1(x) \tag{36}$$

The subtractive nature of Φ suggests an antagonism between two synaptic modification mechanisms - one potentiating (strengthening), and the other depotentiating. It is seen that as q varies, σ_2 and $q\sigma_1$, intersect at different values $x=\theta(q)$.

There are two conditions that both σ_1 and σ_2 must satisfy to let $\Phi(x,q)$ fulfill the conditions of the (N+1) parallel model: First, $\sigma_1(0) = \sigma_2(0) = 0$ and second, the sign of the Wronskian $\text{sgn}(W(\sigma_1(x),\sigma_2(x))) = \alpha$ for $x>0$ (Theorem 3).

Theorem 3. If $\text{sgn}(W(\sigma_1(x), \sigma_2(x))) = \alpha$ for $x>0$ and $\sigma_1(0) = \sigma_2(0) = 0$, then $\text{sgn}(\Phi(x,q)) = \alpha\ \text{sgn}(x)\ \text{sgn}(x-\theta(q))$.

Proof. The proof is as follows: First, since $\Phi(0,q) = 0$ for all q, $\sigma_1(0) = \sigma_2(0) = 0$ must be satisfied. The other zero of $\Phi(x,q)$ is at $x= \theta(q)$, or $q = \theta^{-1}(x) = \sigma_2(x)/\sigma_1(x)$. The slope of

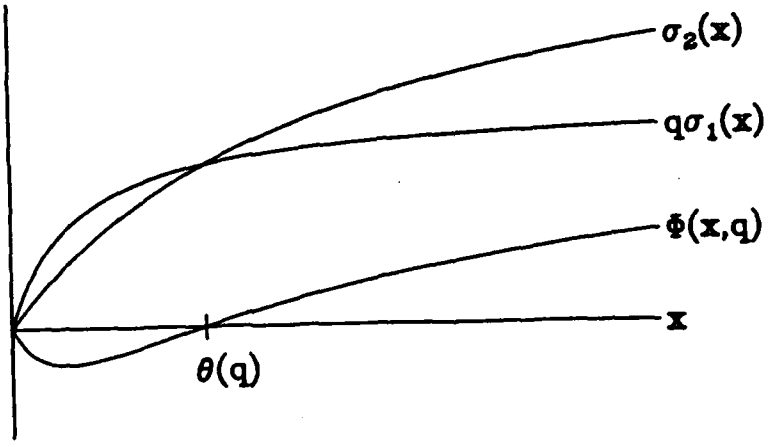


Figure 21. $\Phi(x,q)$ as the difference of two saturating functions. The modification threshold θ depends on the coupling q.

ϕ with respect to x is evaluated (37) at this point and shown to be of the same sign as the Wronskian $W(\sigma_1,\sigma_2)$.

$$\begin{aligned}\phi_x(x=\theta(q)) &= \sigma_2'(x) - \frac{\sigma_2(x)}{\sigma_1(x)}\ \sigma_1'(x)\\ &= \frac{1}{\sigma_1(x)}\ [\sigma_1\sigma_2' - \sigma_1'\sigma_2]\\ &= \frac{1}{\sigma_1(x)}\ \begin{vmatrix}\sigma_1(x) & \sigma_2(x)\\ \sigma_1'(x) & \sigma_2'(x)\end{vmatrix}\\ &= \frac{1}{\sigma_1(x)}\ W(\sigma_1(x),\sigma_2(x))\end{aligned} \qquad (37)$$

Therefore if $\mathrm{sgn}(W(\sigma_1,\sigma_2))$ is positive or negative definite for $x>0$, so is $\phi_x(x=\theta(q))$ and $\phi = \sigma_2 - q\sigma_1$ has at most two zeros: $x=0$ and $x=\theta(q)$. D-cells and S-cells can be characterized by $\mathrm{sgn}(\phi_x(x=\theta(q)))$ as well as by α since $\alpha = \mathrm{sgn}\ (\phi_x(x=\theta(q)))$, hence $\alpha=\mathrm{sgn}(W(\sigma_1,\sigma_2))$.

q as a time average

As in (13), the quantity q(t) is a kind of time average of the the post-synaptic activity. Equation (38) can be solved for q(t) using an integrating factor to give the solution (39).

$$q(t) = \beta(\sigma_2(x) - q\sigma_1(x))\, x \tag{38}$$

$$\begin{aligned} q(t) &= q_0 \exp[-\beta\textstyle\int_0^t x(t'')\sigma_1(x(t''))\, dt''\,] \\ &+ \beta\textstyle\int_0^t \exp[-\beta\int_0^{t'} x(t'')\sigma_1(x(t''))]\, \sigma_2(x(t'))\, x(t')\, dt' \end{aligned} \tag{39}$$

Note that the first term vanishes as $t \to \infty$. Again, $q(t)$ is seen to be an exponentially damped average of some measure of the cell's activity (in this case $x\sigma_2(x)$).

Some allowed functions $\sigma(x)$

For simplicity's sake, only functions σ_1 and σ_2, which are identical in form while having different horizontal and vertical scales, are considered (40).

$$\text{let } \sigma_i(x) \equiv \mu_i \sigma\left(\frac{x}{\eta_i}\right) \qquad \mu_i, \eta_i > 0 \tag{40}$$

where $\sigma(0) = 0$

$\sigma(1) = \frac{1}{2}$ ($x=\eta_i$ gives $\frac{1}{2}$-max of σ_i)

$\lim_{x\to\infty} \sigma(x) = 1$ (μ is max(σ_i))

$\sigma'(x) > 0$ for all x

Since the product $\mu_1\mu_2$ can be factored from the Wronskian, W is independent of μ_1 and μ_2. Thus, consider the determinant $W(\sigma(\frac{x}{\eta_1}),\sigma(\frac{x}{\eta_2}))$. The following Theorem (4) establishes a sufficient condition on σ so that $\Phi(x,q)$ will have at most one zero for $x>0$.

Theorem 4. If $\sigma(x)$ satisfies $x(\sigma')^2 - x\sigma\sigma'' - \sigma\sigma' > 0$, then for a given value of q, there will exist either one or zero values $x>0$ for which $\Phi(x,q) = 0$, depending on the value of q.

Proof. $W(\sigma(\frac{x}{\eta_1}),\sigma(\frac{x}{\eta_2}))$ is evaluated and expressed so that its sign depends upon a difference of two terms, separated according to η (41).

$$W(\sigma(\tfrac{x}{\eta_1}),\sigma(\tfrac{x}{\eta_2})) = \tfrac{1}{\eta_2}\sigma'(\tfrac{x}{\eta_2})\sigma(\tfrac{x}{\eta_1}) - \tfrac{1}{\eta_1}\sigma'(\tfrac{x}{\eta_1})\sigma(\tfrac{x}{\eta_2})$$
$$= \sigma(\tfrac{x}{\eta_1})\sigma(\tfrac{x}{\eta_2})\left[\frac{\sigma'(\frac{x}{\eta_2})}{\eta_2\sigma(\frac{x}{\eta_2})} - \frac{\sigma'(\frac{x}{\eta_1})}{\eta_1\sigma(\frac{x}{\eta_1})}\right] \qquad (41)$$

Hence, the condition (42) is sufficient to our purpose and we see that sgn $(\Phi_x(\theta(q),q))$ depends on the sgn $(\eta_2-\eta_1)$.

$$\frac{d}{d\eta}\left[\frac{\sigma'(\frac{x}{\eta})}{\eta\sigma(\frac{x}{\eta})}\right] > 0 \qquad (42)$$

for all $x>0$, $\eta>0$

The derivative with respect to η is evaluated and the substitution $s \equiv \frac{x}{\eta}$ completes the proof (43).

$$\frac{d}{d\eta}\left[\frac{\sigma'(\frac{x}{\eta})}{\eta\sigma(\frac{x}{\eta})}\right] = \frac{\eta\sigma(\frac{x}{\eta})(-\frac{x}{\eta^2})\sigma''(\frac{x}{\eta}) - \sigma'(\frac{x}{\eta})[\eta(-\frac{x}{\eta^2})\sigma'(\frac{x}{\eta}) + \sigma(\frac{x}{\eta})]}{(\eta\sigma(\frac{x}{\eta}))^2} \qquad (43)$$

$$= \frac{-s\sigma(s)\sigma''(s) + s(\sigma'(s))^2 - \sigma'(s)\sigma(s)}{(\eta\sigma(s))^2}$$

$$\Rightarrow \quad x(\sigma')^2 - x\sigma\sigma'' - \sigma'\sigma > 0 \quad \text{is sufficient!} \quad \text{QED}$$

Under this formulation the feature abstracting quality (S or D) of a neuron depends on which of two saturating functions increases more rapidly rather than on the sign of a variable in the model's differential equation. A mechanistic interpretation poses two biochemical processes as antagonists modulating synaptic plasticity, each increasing with the activity of the postsynaptic cell. The S/D-nature of the cell then depends on certain details of the functional relationships relating each process to the cell's activity.

Among functions satisfying the condition (43) are those sigmoidal functions belonging to the class (44).

$$\sigma(x) = \frac{x^p}{x^p + 1} \qquad p \geq 1 \qquad (44)$$

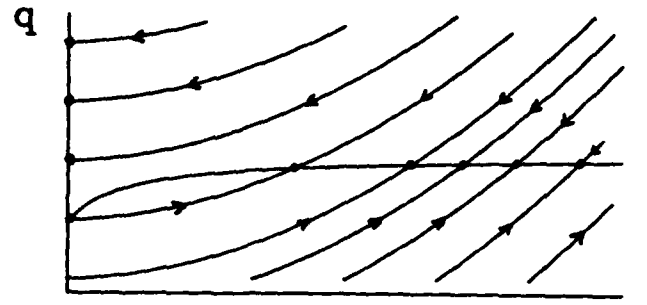
The parameter p determines the "sigmoidocity" of the curve. Near x=0, the function behaves like the power function x^p. This suggests that one or both of the functions σ_i may be linked to the firing rate of the cell, which is seen (Chapman, 1966) to be an increasing bounded function of the summed potential with a threshold.

This system is no longer precisely of the form (30). While the properties of the D-cell model are maintained, the S-cell version of this variation is somewhat different. Here, $(\frac{\dot{m}_i}{d_i})$ and $(\frac{\dot{q}_i}{x})$ are always the same sign: that is, there is no α. Thus, the system is written in a more appropriate form (45), where σ satisfies (43). For a D-cell, $\eta_2 > \eta_1$ whereas the opposite holds for an S-cell.

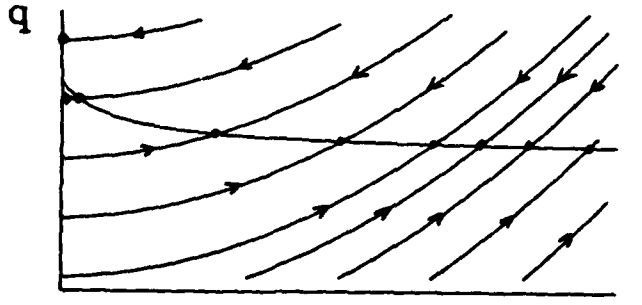
$$\begin{pmatrix} \dot{m}_i \\ \dot{q} \end{pmatrix} = \sigma(\frac{x}{\eta_2}) - q\sigma(\frac{x}{\eta_1}) \begin{pmatrix} d_i \\ x \end{pmatrix} \qquad (45)$$

Examination of the N=1 case

For N=1, the system is always bounded. As in (30), the synaptic weights may vanish altogether if the initial value of q is too large. Sample trajectories are shown (Figure 22) to converge to points on either the nullcline or on the q-axis above the nullcline



(a) $\eta_2 > \eta_1$



(b) $\eta_2 < \eta_1$

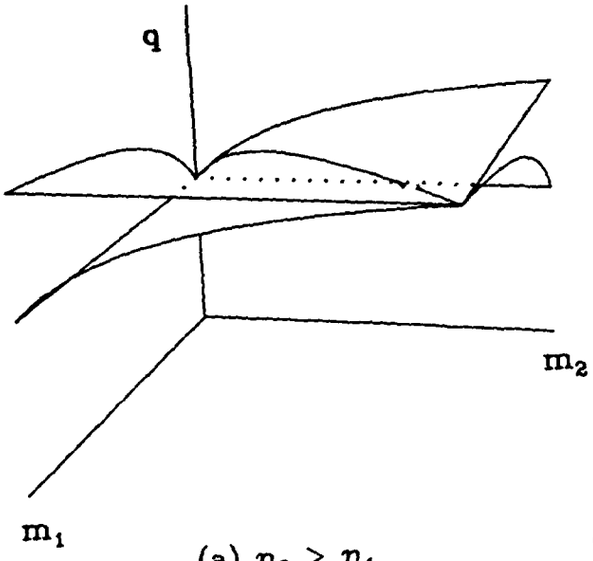
Figure 22. Trajectories of the system (45). The function $\sigma(x) = x/(x+1)$ was used here. Note that for the two cases, the trajectories lie along the same parabolas. (a) $\eta_1 > \eta_2$: D-cell. (b) $\eta_2 < \eta_1$: S-cell.

locus which depends on η_1 and η_2. Note that, in contrast to Figure 19, the trajectories for the two cell types are the same in both cases and the nullcline is not.

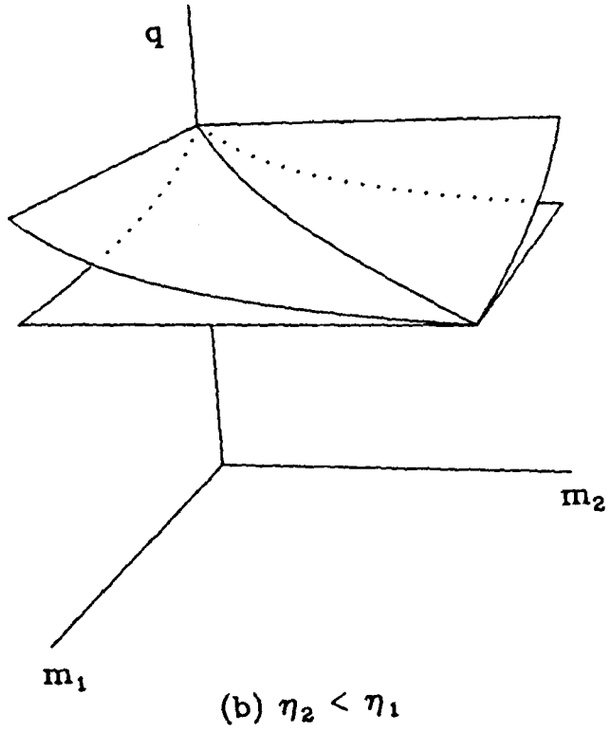
N=2

The surfaces $\sigma_2(m \cdot d^i) - q\sigma_1(m \cdot d^i) = 0$ for $i = 1,2$ are sheets defined by a q-translation of the nullcline above. The pairwise intersections of these nullclines and the planes $m \cdot d^i = 0$ with each other consitute the locus of points in equilibrium with the environment $\{d^1, d^2\}$. Whether the function $x = \theta^{-1}(q) = \sigma_2(x)/\sigma_1(x)$ is monotone increasing or monotone decreasing is seen to determine the feature abstracting properties of the neuron.

Figure 23 illustrates the (2+1) dimensional case. For $\theta(q)$ between $m \cdot d^1$ and $m \cdot d^2$ the neuronal state is driven toward maximum selectivity for $\eta_2 > \eta_1$ and toward minimum selectivity for $\eta_2 < \eta_1$. For η_2 fixed, as η_1 increases from less than η_2 to greater than η_2, the surfaces "pass through" one another, thereby reversing the directions of the trajectories in the m-projection. At $\eta_1 = \eta_2$ the surfaces coincide and the neuron seeks neither maximum nor minimum selectivity.



(a) $\eta_2 > \eta_1$



(b) $\eta_2 < \eta_1$

Figure 23. (2+1)-Space. (a) $\eta_2 > \eta_1$. (b) $\eta_2 < \eta_1$.

N>2

In higher dimensional systems, computer simulation indicates that for $\eta_2 > \eta_1$, the model neuron always seeks maximum selectivity and for $\eta_2 < \eta_1$, it seeks minimum selectivity. Simulation of binocular stimulus environments to D-cells give results consistent with the experimental data reviewed in Part Ic.

IV. An Ensemble Method for Electrophysiological Testing of Synaptic Modification Hypotheses

A. Asymptotic Stability and Investigation of Equilibrium States

Both experimental and theoretical considerations allow elimination of several hypothetical excitatory synaptic modification rule formulations. In particular a good formulation should (1) allow for both strengthening and weakening of a synapse, (2) not allow synaptic strength to grow without bound, and (3) not allow the sign of a synaptic strength to change (a synapse must be always excitatory or always inhibitory).

Properties (2) and (3) are particularly relevant to the present note. Experiments show that, with repetition, conditioning stimulation that initially produces increased synaptic strength is eventually ineffective for producing change. That is, the continued stimulation drives the synapse to an equilibrium state (the approach to equilibrium may be asymptotic). Likewise experimental conditions which initially depress synaptic function are eventually ineffective and, particularly from a theoretical viewpoint, provide no reason to posit synaptic influence reversing modes from excitatory to inhibitory or vice-versa. Of course formulations that predict asymptotic neural states that change due to shifts in environmental statistics are desirable.

Based on these considerations, we now restrict our concern to formulations that yield this asymptotic stability. In fact this criterion is the basis for the proposed experiments and yields a method that is insensitive to many of the problems of other classes of laboratory paradigms. By performing all experiments after an initial *asymptote has* been reached, actual values of postsynaptic excitation and presynaptic frequency can be ignored.

Experimental data (see section IVe) have revealed that the strength m of a synapse can be increased or decreased depending, at least in part, on the presynaptic activity x and the postsynaptic activity y. As a plastic synapse adjusts to a constant activity pattern (x_0, y_0), in its pre- and postsynaptic elements, its efficacy levels off ($\dot{m} \to 0$, where the dot indicates derivative with respect to time). This equilibrium state can be perturbed, shifting $\dot{m}$ to either positive or negative values, by altering the constant activity pattern by an amount (dx,dy) to (x_1, y_1). The dependence of the sign of m on the signs of dx and dy reveals primary information regarding the form of synaptic modification rule (46) and secondary clues to plasticity mechanisms.

$$\dot{m} = \dot{m}(x, y, \ldots) \tag{46}$$

The condition $\dot{m} = 0$ thus implies that some quantities varying with x and/or y have achieved a static equilibrium which can be disrupted by perturbing the activity pattern (x,y). Thus by appropriately adjusting (x,y), one ought to be able to induce potentiation ($\dot{m} > 0$) or depression ($\dot{m} < 0$).

By investigating the dependence of the sign of $\dot{m}$ on the various manipulations (dx,dy), one can determine the approximate form of the modification function (46) in the neighborhood of (x_0,y_0). Let a given experiment be characterized by (sgn dx, sgn dy) where each component is a member of the set (+, 0 ,+) and let each possible result be written as (sgn dx, sgn dy; sgn dm). For example, the paradigm $x_1 > x_0$ and $y_1 = y_0$ is expressed as (+,0) and if the result is $\dot{m} < 0$, then we would write (+,0 ; -).

There are nine possible experiments of this type. The (0,0) experiment is assumed to have a null result, so considering 3 possible results for each of the remaining eight paradigms gives 3^8 or 6561 combinations of possible outcomes. Certain (seemingly inconsistent) combinations imply very complex modification functions. We will confine our analysis to predictions of some simpler theories.

Consider a set of (bi)linear approximations (47) to $\dot{m}$ in the neighborhood of an equilibrium state in which a constant activity pattern (x_0,y_0) has driven the synaptic efficacy to a value m^*.

$$\dot{m} = K^{(a)}(x_0, y_0, m^*)\ (x - b_x) \qquad (47a)$$

$$\dot{m} = K^{(b)}(x_0, y_0, m^*)\ (y - b_y) \qquad (47b)$$

$$\dot{m} = K^{(c)}(x_0, y_0, m^*)\ (x - b_x)\ (y - b_y) \qquad (47c)$$

The equilibria in these hypothetical systems result from a balance in (a) presynaptic, (b) postsynaptic, or (c) both pre- and postsynaptic (independently) terms. In each formula, K is positive (or negative) semi-definite and considered constant in these approximations. The zeros, or balance points, b_x and b_y vary according to their own rules following a time course assumed to be fast enough that we can reach equilibrium in the laboratory and slow enough, that we can measure the immediate change in m resulting from the perturbation.

These formulae represent classes of theories for various neural systems including probability learning (47a) (e.g. Levy & Desmond, 1982), development of specificity in visual cortex (47b) (Bienenstock et al, 1982), and associative memory (47c) (Kohonen, 1977; Sejnowski, 1977a,b). The predictions of the paradigm for each of these models are given in Table 1.

x \ y	↓	0	↑
↓	−	−	−
0	0	0	0
↑	+	+	+

(a)

x \ y	↓	0	↑
↓	−	0	+
0	−	0	+
↑	−	0	+

(b)

x \ y	↓	0	↑
↓	+	0	−
0	0	0	0
↑	−	0	+

(c)

Table 1

B. Extension to Population Studies

The above analysis for individual synapses can be extended to population stimulus-response paradigms. Instead of measuring the effect of a single afferent upon a neuron, consider the macroscopic effects of a stimulus pattern over an axon bundle on the target population. Just as the response of a single cell to a test pulse in an afferent axon is a measure of synaptic efficacy, the net response potential of a population to a test pulse pattern is similarly a measure of the net efficacy of the synapse ensemble corresponding to the afferent bundle.

As all neurons in the postsynaptic population approach an equilibrium steady state in response to a constant stimulus pattern, so must the synapse ensemble efficacy (SEE). The relationship between the SEE and the individual synaptic strengths is closely related to the concepts of microstate and macrostate introduced by Amari (1974) in his method of statistical neurodynamics. It will be noted that Amari is concerned with activity levels rather than synaptic states, hence one should be aware that the test stimulus evokes an activity level that reflects the synaptic state (either of a single synapse or a synapse ensemble). The SEE concept can be generalized to a vector or matrix quantity[2] describing the pattern

[2] The terms "vector" and "matrix" are meant in the multi-dimensional sense. Tensorial mappings between coordinate systems do not hold in general.

of ensemble efficacies of several bundles to a population or to several populations.

Consider the special case in which several bundles ($S_0, S_1, \ldots S_N$) project to a common population R, one projection, S_0, being much denser than the others (Figure 24). That is the activity in S_0 can completely dominate the response in the target population thereby providing a convenient mechanism for experimental control of postsynaptic activity. Thus the activities both pre- and postsynaptic to the synapses <u>from the sparse projections</u> can be imposed by the experimenter. However, the dense or modulatory projection modifies <u>passively</u>, that is the activity postsynaptic to this projection is principally dependent on the projection's (presynaptic) signal. The synapses from the sparse projection undergo <u>active</u> modification in the sense that they are driven by activities with no strong causal relation.

Caution must be exercised with respect to the passively modifying modulatory projection. Assuming this projection is plastic, one must recognize that a constant stimulus to this bundle *does not necessarily* elicit a constant postsynaptic response. Therefore before one proceeds with an equilibrium perturbation paradigm in a dense-sparse system, both the sparse SEE under examination <u>and</u> the dense SEE must achieve sufficiently stable activity levels. Presynaptic perturbations, that is perturbations in the stimuli to the sparse bundles, can be used to provide

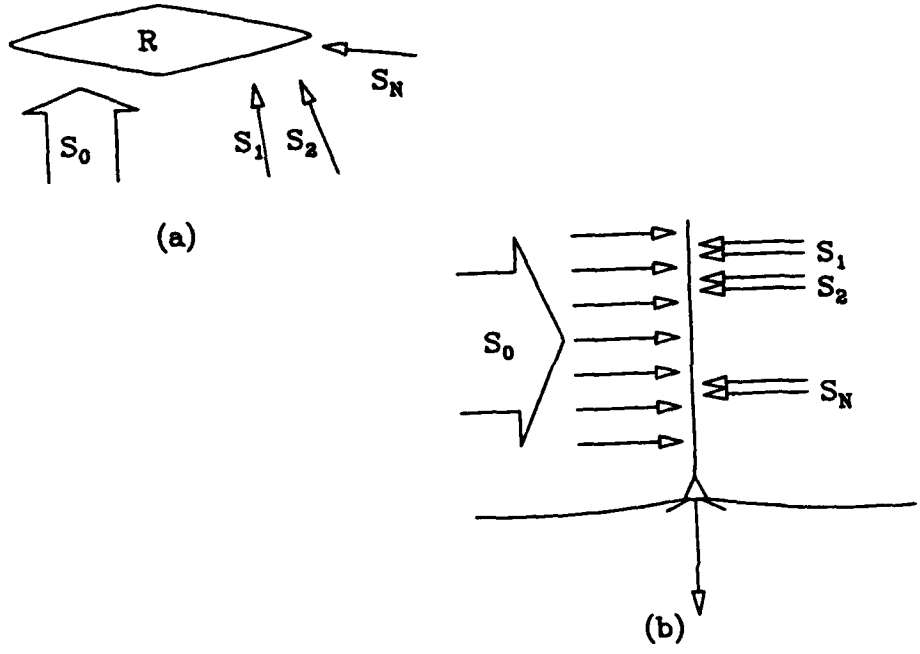


Figure 24. The afferent vector S projects to a population R. (a) The population R is the common target of several axon bundles S_i. The projection S_0 is much denser than the others. (b) A single cell in R.

information about the x dependence of $\dot{m}$ (the middle columns in Table 1).

The investigation of postsynaptic effects is not so straightforward. To alter the steady-state activity in R, we must change the stimulus delivered to S_0, thereby forcing a change in the corresponding SEE. This unfortunate aspect of the experimental design is not fatal. The procedure's validity can be ensured by monitoring the SEE from S_0. As long as the change in the efficacy of the S_0 ensemble follows the change in the S_0 stimulus, it is valid to infer that the activity in the target population is similarly affected.

However, if the S_0-SEE changes in the direction (+ or -) opposite to the perturbation in the S_0 stimulus, it is difficult or impossible to ascertain the change in the activity in R. If the connection is excitatory, the <u>initial</u> change in the S_0-SEE will most likely follow the change in the stimulus. On the other hand, as is seen in the next section, certain models predict that the change in the net asymptotic efficacy value <u>opposes</u> the change in the stimulus after initially following it.

C. Passive Potentiation

Investigation of passive potentiation is not only a necessary

control stage for the experimental framework we have outlined. It may also be used as a self-contained technique for distinguishing between certain associative models. As a control for active synapse modification experiments, passive potentiation in the dense projection S_0 must be understood for the reasons stated in the above section.

In this section we will examine the utility of a passive-potentiation study as a stand-alone experiment to reveal traits of the modification process that are important from a mathematical standpoint. Consider two models both having postsynaptic equilibrium points $y = b_y$ as in (47b), but differing in their evolution equations for the balance point itself (48). The time scale for the modification of b_y is assumed to be slower than that for m. Notice that equilibrium can be attained only when both m and b_y vanish.[3]

$$\dot{b}_y = \beta(y - f(y,b_y)) \tag{48}$$

If $f(y,b_y) = b_y$, then there is a single criterion for equilibrium, namely when $y = b_y$. Otherwise (as in the FP model), the system has a much more restricted set of equilibrium points. In this case y must simultaneously satisfy two conditions and so in general only certain isolated values of y are allowed.

[3] The expectation values of m and b_y over their pattern environments should be zero at equilibrium. In the experimental paradigms discussed here, only one pattern is presented for the duration over which the expectation values are integrated.

Consider an experiment similar to that described in the last section, in this case for the measurement of the SEE of a dense (controlling) projection. As before, the synapse ensemble is driven to asymptotic efficacy. Hypothetically, (47b) has been accepted as a description of $\dot{m}$ local to equilibrium. The next step is to determine whether (48) can be expressed with $f(y,b_y) = b_y$. When the system settles down, the constant stimulus to the axon bundle is shifted and again held constant. If the condition $y = b_y$ alone is sufficient for equilibrium, the prediction of the coupled equations (47b,48) is that the SEE will <u>monotonically</u> shift to a new asymptote.

On the other hand, if y is forced to the same equilibrium value by both constant stimulus levels then the SEE must initially follow the stimulus change, peak at an extreme (relative minimum or maximum), and asymptotically seek a net shift opposite to the stimulus shift. To appreciate this, consider a more detailed example (Figure 25). The synapse ensemble is initially potentiated to a steady state value M_1 by a stimulus of intensity X_1 over a time period T_1, after which a stronger stimulus X_2 is delivered for a period T_2 driving the SEE to a new asymptotic value M_2. Following (47b) in the short time scale, an initial synaptic increase is expected. The weak criterion, under which $y = b_y$ is sufficient for equilibrium, allows the system to achieve a second steady state without reversing the trend. Otherwise both asymptotic population responses must be equal, and since $X_2 > X_1$ the model predicts $M_2 < M_1$.

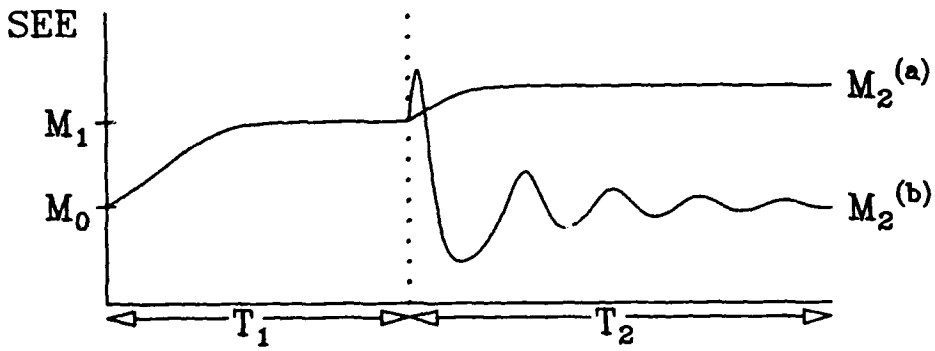


Figure 25. Theoretical predictions for a population experiment. Whereas nearly all functions $f(y,b_y)$ may give the same initial behavior (T_1), the prediction for the period following a post-equilibrium stimulus shift (T_2) is different for (a) $f = b_y$ than for (b) other functions f.

These predicted behaviors of the SEE result from the interaction between the coupled differential equations for $\dot{m}$ and $\dot{b}_y$. This is illustrated by the global state diagrams in Parts II & III (Figures 8 and 19). Even though these figures represent single synapse cells, the SEE of a multisynaptic neuron should behave accordingly. In Figure 26, we see that a shift in the constant stimulus value causes a corresponding shift in the nullclines. Thus new "rules" are established in the state space and the synaptic strengths must obey them. Note the direct approach allowed in (a) as opposed to the spiral trajectory in (b). The projections of these paths on the m-coordinate give the T_2 time courses illustrated in Figure 25.

D. Testing Nonhebbian Models

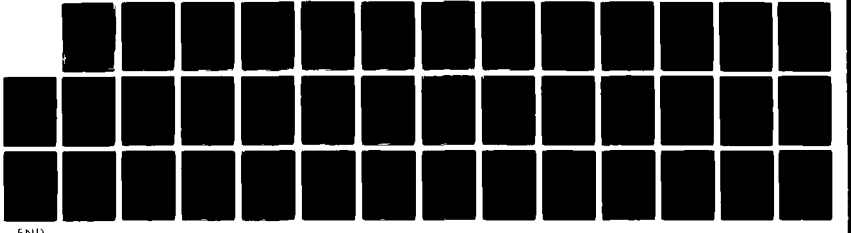
Nonhebbian (in the sense that they are not multiplicative) models are not described by any of the formulations in (47), hence they predict combinations of experimental results not included in Table 1. These theories are by no means excluded from our experimental approach. The only requirement is that one be able to predict qualitative behavior of the SEE near equilibrium for various paradigms. In this section we again contrast the predictions of two variants of a class of models.

Recall the presynaptically balanced modification rule (47a). As

AD-A128 936 NEURAL PLASTICITY: SINGLE NEURON MODELS FOR DISCRIMINATION AND GENERALIZA..(U) BROWN UNIV PROVIDENCE RI CENTER FOR NEURAL SCIENCE P W MUNRO

2/2

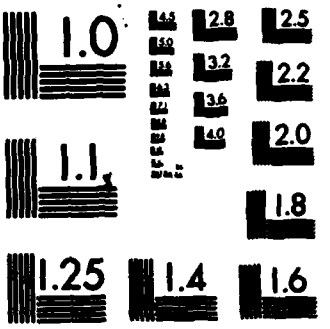
UNCLASSIFIED JUN 83 TR-8 N00014-81-K-0136 F/G 6/16 NL

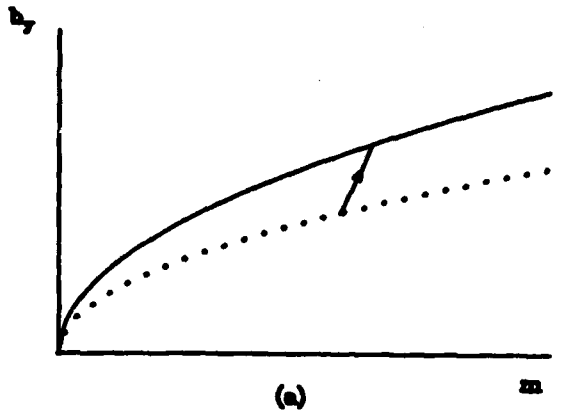


END

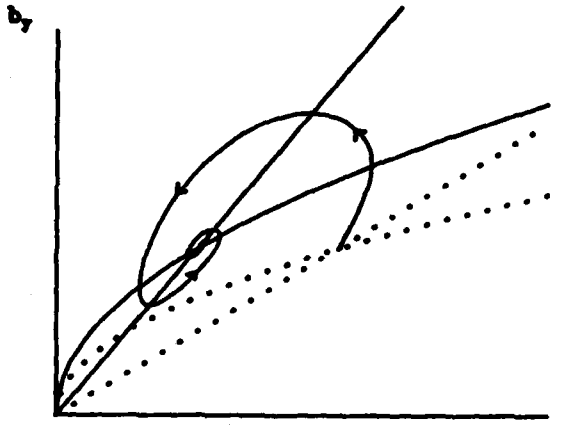
DATE FILMED

DTIC





(a)



(b)

Figure 26. State space diagrams. Two cases are illustrated: (a) $f(y,b_y) = b_y$, and (b) other functions f. Note that the m-projections of these curves give the predictions for M in the previous figure.

a hebbian rule, the expression for $\dot{m}$ should be factorable into pre- and postsynaptic components, so the term $(x - b_x)$ should be independent of postsynaptic variables. Consider a nonhebbian generalization of (47a) in which the equilibrium condition is expressed as a balance between pre- and postsynaptic activities (49).

$$\dot{m} = K(x_0, y_0, m^*)\ (x - f(y)b_x) \qquad (49)$$

Prediction combinations for various functions f(y) are given in Table 2: (a) f(y) increases monotonically ; (b) f(y) is constant (hebbian case) ; (c) f(y) decreases monotonically.

E. Application to the EC-DG Pathway

The principal input to the hippocampus arises from the entorhinal cortex, an adjacent structure. Figure 27 gives a schematic view of the bilateral pathway projecting from the entorhinal cortex (EC) to the dentate gyrus (DG) in hippocampus. Plasticity in the EC-DG pathway has been demonstrated by many investigators. The response of DG granule cell populations has been shown to increase as a result of strong electrical stimulation applied directly to the pathway (Bliss & Lomo, 1973). The so-called long-term potentiation (LTP) effect has been shown to last for a

Table 2

(a)

x / y	←	0	→
←	?	+	+
0	1	0	+
→	1	1	?

(b)

x / y	←	0	→
←	1	0	+
0	1	0	+
→	1	0	+

(c)

x / y	←	0	→
←	1	1	?
0	1	0	+
→	?	+	+

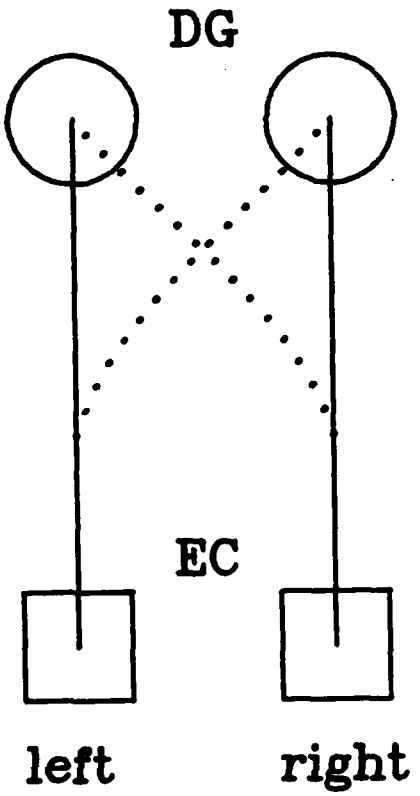


Figure 27. The bilateral EC-DG pathway. This diagram is based on the data of Steward & Vinsant (1978) in which they induced sprouting in the crossed pathway and found all new crossed projections to be collaterals of ipsilateral projections. The dotted lines indicate the sparse, crossed pathways. Solid lines are the denser ipsilateral projections.

period of months (Douglas & Goddard, 1975). Levy & Steward (1979) have simultaneously stimulated contra- and ipsilateral pathways for Hebb-like modification. The above development of the SEE approach in this section was motivated by this system.

In the proposed experiment, stimulating electrodes are placed (see Levy & Steward (1979) for details of the setup) in the left and right angular bundles and at least one recording electrode is placed in the dentate gyrus, either unilaterally or bilaterally. When both stimulators are active, the DG population response is under the nearly complete control of the ipsilateral electrode. Thus, the experiment is to observe the change in the contralateral SEE (cSEE). A measure of the cSEE is the population voltage response in DG to a single contralateral pulse.

One must be careful to follow the cSEE to equilibrium for each driving stimulus. It is important to be aware of the possibility of a change in the iSEE. Hence it may be a worthwhile precaution to measure the iSEE as well as the cSEE as a test of equilibrium.

V. Discussion

A. The S-cell as a Contrast-Enhancing Unit

Tonic inhibition can easily be shown to increase specificity in a selective neuron. Experimental evidence for this phenomenon is provided by Sillito (1975). The role of inhibitory activity in a model neuron is discussed by Biemenstock (1980). In that paper it is shown that cells lose specificity if their "ideal" synaptic inputs are confined to positive values. Anatomical data does not support the notion of the ideal synapse. Feed-forward inhibition by an S-cell to several D-cells all having roughly the same receptive field supplies them with a generalized, tonic, negative input.

Consider a circuit (Figure 28) in which a cortical D-cell receives strictly excitatory patterned thalamic inputs and is inhibited by a neighboring cell receiving the same input. Assuming linear response properties, the D-cell responds according to (50), where m and n are respectively the states of the D and I (inhibitory) cells and k is the magnitude of the inhibitory I-D coupling.

$$x = m \cdot d - k(n \cdot d) = (m - kn) \cdot d \qquad (50)$$

By restricting m, n, and k to be non-negative, the model is more consistent with contemporary anatomical theory. For n and k

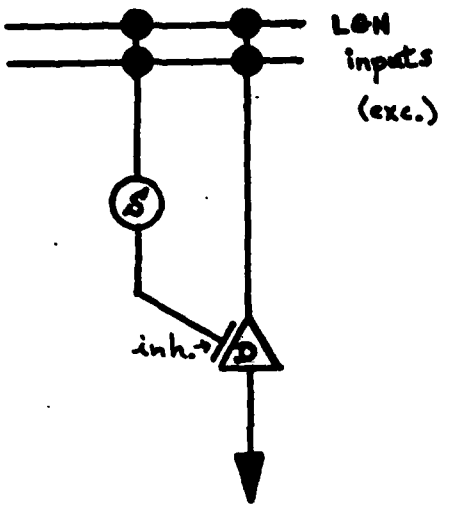


Figure 2a. The S-cell as a contrast enhancer to a D-cell. In this simple circuit, a generalizing unit (S-cell) inhibits a D-cell uniformly over their common pattern environment thereby separating the patterns for the D-cell.

nonmodifiable, the product kn provides an arbitrary offset since $d/dt(m-kn) = dm/dt$. Provided kn is sufficiently large for all patterns, there exist positive synaptic values m_i for which the D-cell has maximum selectivity. If the inhibitory contribution kn is too large the D-cell will never fire. Hence, the "allowed" values of kn over the pattern space are restricted by these two considerations.

Taking a more robust approach, let the inhibitory neuron seek minimum selectivity and let the k-synapse be modifiable. The S-cell receives only excitatory input and therefore develops according to the analysis in previous sections. Note the response x of the D-cell (50). As the S-cell approaches equilibrium, n goes to zero and $d/dt(m_i-kn_i) = dm_i/dt - n_i dk/dt$. To maintain the properties of D-cell parallel modification, k should modify <u>opposite</u> to m (51), since they share the same post-synaptic cell.

$$\begin{aligned} \dot{k} &= -\phi(x,q) \cdot [\text{S-cell output}] \\ &= -\phi(x,q) \cdot (n \cdot d) \end{aligned} \tag{51}$$

Hence m and k change such that $d/dt[(m-kn)\cdot d]$ increases if $x>\theta(q)$ for a given d, and decreases if $0<x<\theta(q)$. This circuit will tolerate a much broader range of initial conditions than if the inhibitory cell is nonplastic. The time course of both S and D response characteristics for a (monocular) circular environment is given in Appendix C using the FP model.

Note that a single S-cell may inhibit several D-cells receiving inputs from the same receptive field (Figure 29), each tuning independently. Recent experimental data from Hubel and Livingstone (1982) and Tootell et al (1982) support this kind of circuit. They find periodic clusters of nonselective cells centered in ocular dominance columns. Also, Hendrikson et al (1981) report having found the GABA-synthesizing enzyme gluatamic axid decarboxylase (GAD) in localized regions arranged in a strikingly similar pattern. Their studies in monkey visual cortex reveal these GAD-rich areas to run parallel to the centers of the ocular dominance columns.

This approach to resolving the ideal synapse has been developed in cooperation with Chris Scofield who is applying it to a theory for the development of the topography of orientation preference across cortex.

B. Comments Regarding the Antagonistic Mechanisms Approach

q as a synaptic value

As is mentioned in Part III, either of the saturating functions $\sigma(x)$ may be related to the firing response of the neuron. Consider the term $q\sigma_1(x)$. Since q modifies in a hebb-like fashion, it may reflect the efficacy of a synapse. Figure 30 depicts q as the efficacy of an autapse (a synaptic junction having pre- and

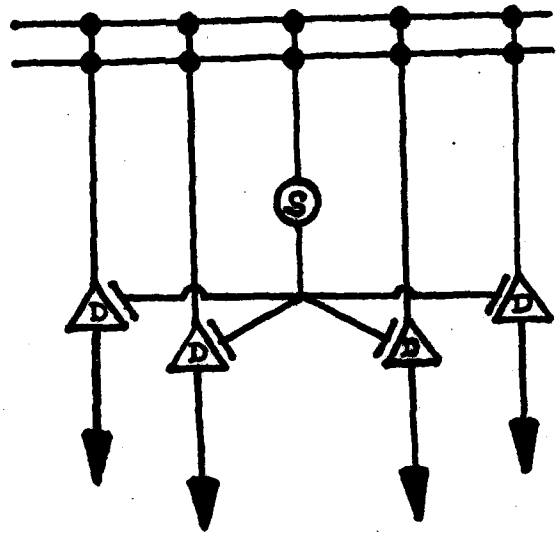


Figure 29. An S-cell enhances several D-cells. A single S-cell can provide contrast enhancing feed-forward inhibition to several D-cells.

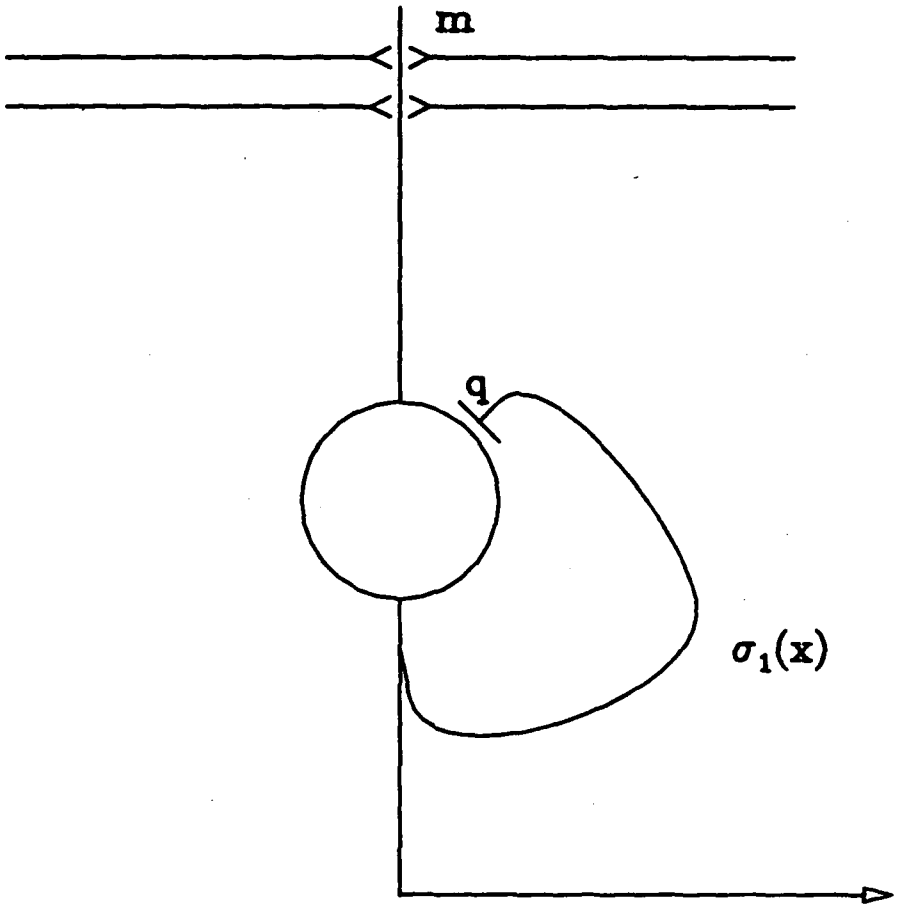


Figure 30. Autaptic modulation of Φ. In this circuit the neuron's output is fed back to the cell as the function σ_1. This may be accomplished via a special autaptic connection of strength q.

postsynaptic elements that are processes of a common neuron (van der Loos and Glaser, 1972)).

Whether $q\sigma_1(x)$ inhibits the firing of the cell and whether it is mediated by an actual autaptic structure are details not considered here except for the following observation. If an explicit structure mediates the subtractive term $q\sigma_1(x)$, such structures are expected to be axosomatic so they can interact with the quantity $\sigma_2(x)$. (Note that x is a somatic quantity). Calvet & Calvet (1979) report the existence of axosomatic autapses in Purkinje cells from young kittens. Such a structure must either function differently than other synapses or give the neuron negative feedback. In the latter case, the autapse would be inhibitory and therefore this mechanism would only be expected on inhibitory neurons.

The function $f(x) = 27\ \ln(x+1)$ where x is the total membrane current in μamp/cm has been computed by Agin (1964) from the equations of Hodgkin and Huxley (1952) to give the response frequency (pulses/sec) of an axon. Strictly speaking, this function is not bounded, however it increases very slowly for large x and behaves like a saturating function in other respects. Keep in mind that x is bounded in practice. Note that by (52) this function satisfies the condition (43) and hence by Theorem 4 it gives a ϕ with only one nontrivial zero.

$$\sigma(x) = \ln(1+x) \qquad \sigma'(x) = \frac{1}{1+x} \qquad \sigma''(x) = -\frac{1}{(1+x)^2}$$

$$x(\sigma')^2 - x\sigma\sigma'' - \sigma'\sigma = \frac{x}{(1+x)^2} + \frac{x\ln(1+x)}{(1+x)^2} - \frac{\ln(1+x)}{1+x} \qquad (52)$$

$$= \frac{1}{(1+x)^2}(x - \ln(1+x)) > 0$$

for $x > 0$

The equation for the system in Figure 30 can be rewritten in the N+1 dimensional form of (53). A formalism inspired by the four-vector approach to relativity theory in physics can now be presented for the D-cell.

$$\begin{pmatrix} \dot{m} \\ \dot{q} \end{pmatrix} = [\sigma_2(x) - q\sigma_1(x)] \begin{pmatrix} d \\ \sigma_1(x) \end{pmatrix} \qquad (53)$$

Consider a linear range for $\sigma_2(x)$ such that $\sigma_2(x) = x$. Then the (N+1)-vectors (m,q) and $(d,\sigma_1(x))$ can be considered to consist of N components describing the <u>spatial</u> pattern of afferent activity and connectivity to the neuron and a single <u>temporally</u> varying spatial integration and self-connectivity. The metric g (54) allows Φ $(= x - q\sigma_1(x))$ to be written as the inner product of (m,q) and $(d,\sigma_1(x))$ giving (55).

$$g = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & -1 \end{array}\right) \qquad (54)$$

$$\dot{M} = (M \cdot D)D \qquad (55)$$

$$\underline{\text{where}} \quad M \equiv (m,q) \quad D \equiv (d,\sigma_1)$$

Note further that all components of M are 'synaptic' and those of D are measures of axonal activities. The postsynaptic activity x is then just the N-product of m and d while the modulatory signal ϕ is the (N+1) 'scalar' M·D.

Antagonistic Mechanisms

The simplicity of the approach in Part III, namely that two mechanisms work against one another to give the modulatory function ϕ, is important. Together with the parallel aspect of modification ($m_{\parallel} d$), this permits the system equation to be separated into a few straightforward principles of synaptic modification:

1. The efficacy of an afferent's signal to a given target neuron varies over time in proportion to that (instantaneous) signal. (Parallel modification)

2. Two mechanisms, which are dependent on the postsynaptic activity, determine the ratio of efficacy change to input signal. One strengthens the synapses; the other weakens them.

3. The mechanisms are coupled by a variable quantity

which varies linearly with the postsynaptic activity according to the same ratio.

As long as the functional dependence of each mechanism on the cell's activity satisfies (37), the neuron will seek either minimum or maximum selectivity, depending on the sign of the Wronkskian $W(\sigma_1,\sigma_2)$ alone. In terms of the earlier presentations of the model in Part II, the sign of $W(\sigma_1,\sigma_2)$ determines whether the monotone function $x = \theta(q)$ is increasing or decreasing. Thus it is observed that the feature-abstracting character of a neuron can be determined either by the relative signs of m and q (the parameter α) or the monotonic trend of $\theta(q)$.

C. Cognitive Applications

Four neuronal types can be straightforwardly defined based on the models described here. The afferents of each neuron can be tuned for either minimum (S-cell) or maximum (D-cell) selectivity over a single stimulus parameter and the efferent action is either excitatory or inhibitory. It is hypothesized that a network consisting exclusively of these four cell types can be designed to perform arbitrarily complex similarity-difference perceptual tasks. That is, such a network should be able to construct a taxonomy appropriate to any stimulus environment. Certainly more easily said than done, this hypothesis is nevertheless difficult to disprove. It

is offered as a challenge to provoke the realization of the full potential of such circuits. In this subsection some initial steps to test this hypothesis are outlined.

It must be emphasized from the outset that this discussion is limited to the cognitive phenomenon of hierarchical abstraction, i.e. classification of stimuli by their similarities and differences. The activity of each neuron simultaneously modulates the modification of all its input synapses in proportion to the activities of their individual afferents relative to an internal time-varying scalar parameter. This mechanism allows individual units to evolve such that they abstract simple (scalar) information from their stimulus environments. In this sense, the above hypothesis and the mathematical model given in this thesis are consistent with the single-unit approach of Barlow (1972). His five dogmas rely on the notion of a neuron's 'trigger feature' -- that aspect of a stimulus which excites the neuron. In the picture of the nervous system presented here, stimuli drive neuronal states to become maximally sensitive to certain trigger features. These may be specific 'redundant patterns of stimulation' (D- cells) as in Barlow's third dogma or generalized prototypical aspects common to various stimulus patterns that are rarely or never observed as pure inputs (S-cells).

As is commonly stated (implicitly or explicitly), the nervous system is thought to be organized in stages, progressing in perceptual complexity from sensation to recognition, association, and

cognition. So, it is proposed, should the hypothetical feature abstracting network be organized. A genetically coded macrostructure lends efficacy and relevance to environmentally influenced learning at the cellular level in that whereas each neuron is modifiable and adapts to informational qualities in its environment, the overall design of the system is genetically programmed so that increasingly complex information is abstracted by successive stages of neural processing.

Shepard and Podgorny (1975) point out that symbolic stimuli are discretely coded while physical (nonsymbolic) stimuli vary continuously along one or several dimensions: "For, whereas we can continuously shift a color (for example, blue) in brightness, hue, and saturation until it becomes as similar as we wish to any other color (for example, green), we cannot continuously deform a word 'blue' into another word 'green' without passing through intermediate configurations that are not words at all". Thus it is expected that inputs to the higher level (more "gnostic") stages are more distinctly separated than early (sensory) input patterns. Two candidates for implementing such separation between inputs are strategic coding and tonic inhibition. Vectors of very high dimensionality tend to be nearly orthogonal and are hence intrinsically well-separated. Tonic inhibition can enhance separation by quashing responses to mediocre excitation.

The perception of differences and similarites by human subjects

has been studied by several researchers (e.g. Rosch, 1975; Tversky and Gati, 1978, 1982). The considerable amount of data these studies have generated ought to provide tests for theoretical hierarchy-abstracting networks. For example, the response to stimulus pattern A of a neuron broadly tuned to stimulus pattern B may measure the perceived similarity of A to B. Such a definition is consistent with Tversky's (1977) notion of asymmetry judgements expressed by a contrast model (56). Such mathematical formulations provide a means of comparing the performances of model networks with behavioral data.

$$D(i,j) = \alpha f(i-j) + \beta f(j-i) - \theta f(i \cap j) \qquad \alpha,\beta,\theta > 0 \qquad (56)$$

As a final point, the value of this approach to understanding the relationships between neural and cognitive phenomena is considered. The complex information-processing capabilities of biological networks can only be understood via the increased application of sophisticated mathematical tools to neuroscience. Such efforts may provide brain research with the long-awaited unravelling of the neural mechanisms underlying learning and cognition. The mapping of a mathematical structure for neural network processing onto a mathematical model of cognitive performance is likely to be the unifying event in this undertaking.

D. The Synapse Ensemble Method

Data from a variety of investigative modes (linguistic, psychological, electrophysiological, and neuropharmachological) have been subjected to analytical treatment of an increasingly sophisticated mathematical nature over the last few decades. As synaptic modification models evolve, a growing interaction between model development and laboratory research is anticipated. The method for electrophysiological testing put forward in Part IV is suited to the current state of laboratory technique. It promises to narrow the allowed classes of models for certain synaptic networks by revealing aspects of their plasticity properties near equilibrium states.

The notion of measuring a *synapse ensemble efficacy* (SEE) is an important one. It is based on the following assumption: As the net response of a neural population to a constant stimulus pattern from a projecting population or populations approaches an asymptotic value, the activity of each individual neuron converges to a limit and furthermore, so does the efficacy of each synapse. This should be considered as nothing more than a working hypothesis until it is either proven false or no longer necessary as single-cell methods improve. The notions of convergence and asymptotic approach must be carefully considered by the experimenter, for they have a strong impact on data interpretation.

The terms "active" and "passive" modification respectively correspond to the notions of supervised and unsupervised learning used in theories of knowledge acquisition (see Cooper, 1973). Both types require that the organism (student) abstract information from a stimulus environment (the world). Passive learning is a fully <u>self-organizing</u> process while active learning requires a teacher or supervisor to rectify incorrect conclusions drawn by the inductive process, thereby directing and accelerating the organization of the system. Analogously, individual neurons or neuron populations are students undergoing a process of synaptic modification so that they "learn" to respond appropriately to their stimulus environment. If a projection has sufficient strength to control a population it can perform the teacher role in associatively modulating the plasticity of weaker projections. However the strong projection must modify passively unless it receives very weak stimulation while another (convergent) bundle "teaches".

Appendix A. The Model Neuron as a Signal Detector

In multipattern stimulus environments, D-cells seek maximum selectivity and S-cells seek minimum selectivity. In the context of a one-pattern environment, selectivity is no longer well-defined. The parallel modification models (both S and D) described in this thesis drive the synaptic state parallel to the pattern until (A1) is satisfied (Figure 31a).

$$\mathbf{m}\cdot\mathbf{d} = \theta(q) \tag{A1}$$

The addition of a decay term (antiparallel to m) to the function driving synaptic modification causes the components of m not parallel to d to vanish (Figure 31b), thereby forcing m to a final state parallel to d. The decay term must be carefully chosen to ensure that not m does not vanish completely. Consider a linear decay (A2).

$$\dot{\mathbf{m}} = \phi(x,q)\mathbf{d} - \varepsilon\mathbf{m} \tag{A2}$$

This works for m_0 sufficiently large, however the state m = 0 is unstable. By using a decay term that goes as a higher power of m, this problem is solved (A3).

$$\dot{\mathbf{m}} = \phi(x,q)\mathbf{d} - \varepsilon|\mathbf{m}|\mathbf{m} \tag{A3}$$

Random or noisy components to the signal cause perturb the state, but if each modification to m is not too large (i.e. for small time steps in the difference equation), the state remains nearly parallel to d forever (Figure 31c).

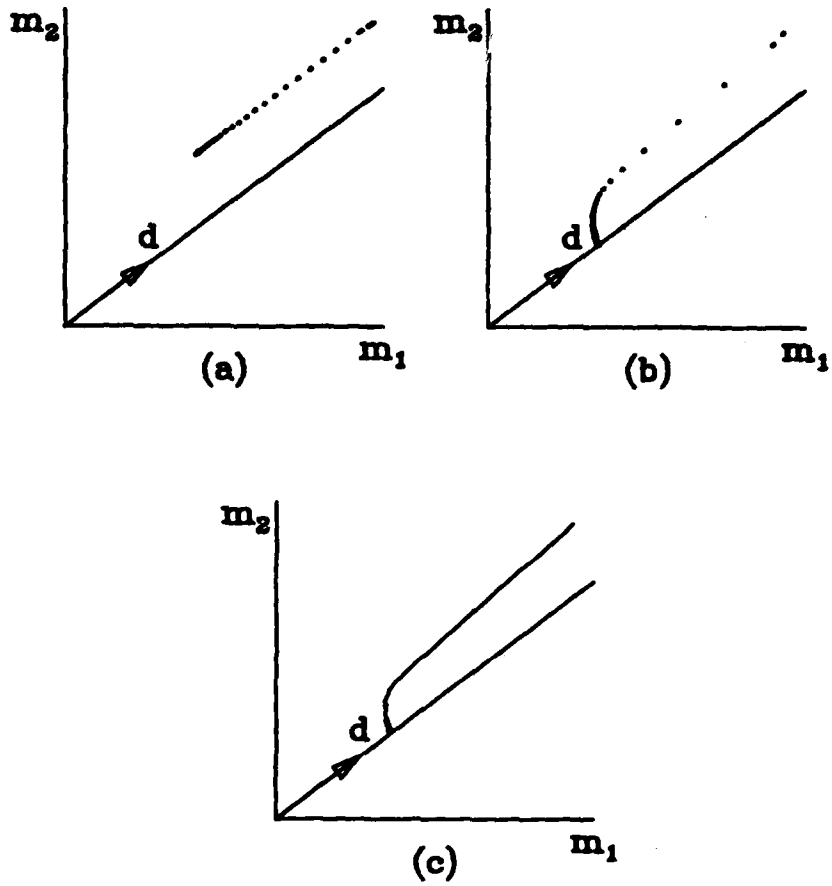


Figure 31. Extraction of a signal from noise. A single pattern is presented to a D-cell in this 2-dimensional simuliation. (a) no decay, no noise. (b) linear decay, no noise. (c) linear decay with noise.

So for $\beta < \phi_m$, the real parts of both roots are positive and therefore a limit cycle exists. For $\beta > \phi_m$, the trajectories converge to the critical point. Note that for Figure 8, $\phi_m = \frac{\pi}{4} = 0.785$ was used.

Appendix B. Stability of the (1+1)-dimensional FP-model D-cell about the Critical Point (m,q) ≡ (1,1)

The behavior of the system (14) is analyzed about the critical point (1,1) in this appendix. From Theorem 1 it is understood that if trajectories diverge locally, they must eventually converge to a limit cycle on a more global scale. Consider the perturbation $(m,q) = (1+x,1+y)$. This gives a linear first order system (B1) for x and y.

$$\begin{aligned} \dot{x} &= \phi_m(m=1,q=1)\,[x - 2y] \\ \dot{y} &= \beta\,[x - y] \end{aligned} \tag{B1}$$

The solutions of (B1) are of the form (B2):

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} A_x \\ A_y \end{pmatrix} \exp(r_1 t) + \begin{pmatrix} B_x \\ B_y \end{pmatrix} \exp(r_2 t) \tag{B2}$$

where r_1 and r_2 (both non-zero) satisfy (B3):

$$r^2 + (\beta - \phi_m(1,1))r + \phi_m\beta = 0 \tag{B3}$$

therefore

$$r = \tfrac{1}{2}(\phi_m - \beta \pm \sqrt{\phi_m^2 + \beta^2 + 6\phi_m\beta}) \tag{B4}$$

Appendix C. Computer Simulation of an S-cell Inhibiting a D-Cell.

The circuit used in this simulation is described in Part Va (Figure 27).

ENVIRONMENT (10 patterns in 7 dimensions):

pattern	components						
1	1.518	0.076	0.000	1.722	0.000	0.076	1.518
2	1.677	0.000	0.016	1.729	0.016	0.000	1.677
3	0.917	0.000	0.417	1.713	0.417	0.000	0.917
4	0.224	0.000	1.235	1.717	1.235	0.000	0.224
5	0.000	0.000	1.732	1.732	1.732	0.000	0.000
6	0.000	0.224	1.235	1.717	1.235	0.224	0.000
7	0.000	0.917	0.417	1.713	0.417	0.917	0.000
8	0.000	1.677	0.016	1.729	0.016	1.677	0.000
9	0.076	1.518	0.000	1.722	0.000	1.518	0.076
10	0.644	0.644	0.000	1.711	0.000	0.644	0.644

Initial state of S - cell : 1 0 0 1 0 0 1
Initial state of D - cell : 1 1 1 1 1 1 1
Initial value of k : 0

pattern	1	2	3	4	5	6	7	8	9	10
t = 0 k = 0.00 q(D-cell) = 0.00 q(S-cell) = 0.00										
D-cell:	4.91	5.11	4.38	4.63	5.20	4.63	4.38	5.11	4.91	4.29
S-cell:	4.76	5.08	3.55	2.16	1.73	1.72	1.71	1.73	1.87	3.00
t = 100 k = 1.48 q(D-cell) = 0.99 q(S-cell) = 0.99										
D-cell:	1.34	1.53	0.67	0.04	0.00	0.00	0.00	0.00	0.00	0.39
S-cell:	1.06	1.10	0.95	0.96	1.04	0.94	0.90	1.03	0.99	0.92
t = 200 k = 1.11 q(D-cell) = 2.47 q(S-cell) = 0.99										
D-cell:	7.77	8.57	4.69	0.98	0.00	0.00	0.00	0.00	0.16	3.26
S-cell:	1.02	1.05	0.93	0.97	1.07	0.97	0.93	1.05	1.02	0.91
t = 300 k = 1.30 q(D-cell) = 2.90 q(S-cell) = 0.99										
D-cell:	9.63	10.73	5.41	0.36	0.00	0.00	0.00	0.00	0.00	3.46
S-cell:	1.02	1.05	0.93	0.97	1.06	0.97	0.93	1.05	1.02	0.92

```
t =  400  k =  1.77  q(D-cell) =  3.15  q(S-cell) =  0.99
D-cell: 10.47 11.69  5.76  0.09  0.00  0.00  0.00  0.00  0.00  3.58
S-cell:  1.02  1.05  0.93  0.97  1.06  0.97  0.93  1.05  1.02  0.92

t =  500  k =  2.25  q(D-cell) =  3.23  q(S-cell) =  0.99
D-cell: 10.83 12.13  5.82  0.00  0.00  0.00  0.00  0.00  0.00  3.49
S-cell:  1.02  1.05  0.94  0.97  1.06  0.97  0.94  1.05  1.02  0.92

t =  600  k =  2.75  q(D-cell) =  3.27  q(S-cell) =  0.99
D-cell: 11.08 12.45  5.81  0.00  0.00  0.00  0.00  0.00  0.00  3.35
S-cell:  1.02  1.05  0.94  0.98  1.06  0.98  0.94  1.05  1.02  0.92

t =  700  k =  3.25  q(D-cell) =  3.31  q(S-cell) =  0.99
D-cell: 11.33 12.77  5.79  0.00  0.00  0.00  0.00  0.00  0.00  3.20
S-cell:  1.02  1.05  0.94  0.98  1.06  0.98  0.94  1.05  1.02  0.93

t =  800  k =  3.76  q(D-cell) =  3.35  q(S-cell) =  0.99
D-cell: 11.59 13.09  5.77  0.00  0.00  0.00  0.00  0.00  0.00  3.05
S-cell:  1.01  1.04  0.94  0.98  1.06  0.98  0.94  1.04  1.01  0.93

t =  900  k =  4.26  q(D-cell) =  3.39  q(S-cell) =  0.99
D-cell: 11.84 13.41  5.74  0.00  0.00  0.00  0.00  0.00  0.00  2.89
S-cell:  1.01  1.04  0.94  0.98  1.05  0.98  0.94  1.04  1.01  0.93

t =  1000  k =  4.77  q(D-cell) =  3.43  q(S-cell) =  0.99
D-cell: 12.10 13.74  5.72  0.00  0.00  0.00  0.00  0.00  0.00  2.73
S-cell:  1.01  1.04  0.95  0.98  1.05  0.98  0.95  1.04  1.01  0.93

t =  2000  k = 9.76  q(D-cell) =  3.81  q(S-cell) =  1.00
D-cell: 14.63 17.00  5.42  0.00  0.00  0.00  0.00  0.00  0.00  1.11
S-cell:  1.01  1.03  0.96  0.98  1.04  0.98  0.96  1.03  1.01  0.95

t =  3000  k =  13.38  q(D-cell) =  4.12  q(S-cell) =  1.00
D-cell: 16.55 19.49  5.14  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.01  1.02  0.97  0.99  1.03  0.99  0.97  1.02  1.01  0.97

t =  4000  k =  16.32  q(D-cell) =  4.20  q(S-cell) =  1.00
D-cell: 17.20 20.53  4.32  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.01  1.02  0.98  0.99  1.02  0.99  0.98  1.02  1.01  0.97

t =  5000  k =  19.28  q(D-cell) =  4.29  q(S-cell) =  1.00
D-cell: 17.85 21.58  3.48  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.01  0.99  0.99  1.02  0.99  0.99  1.01  1.00  0.98

t =  6000  k =  22.20  q(D-cell) =  4.38  q(S-cell) =  1.00
D-cell: 18.50 22.62  2.66  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.01  0.99  1.00  1.01  1.00  0.99  1.01  1.00  0.99

t =  7000  k =  24.98  q(D-cell) =  4.47  q(S-cell) =  1.00
D-cell: 19.15 23.64  1.89  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.01  0.99  1.00  1.01  1.00  0.99  1.01  1.00  0.99
```

```
t =  8000  k =  27.48  q(D-cell) =  4.57  q(S-cell) =  1.00
D-cell: 19.80 24.64  1.25  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.01  0.99  1.00  1.01  1.00  0.99  1.01  1.00  0.99

t =  9000  k =  29.58  q(D-cell) =  4.68  q(S-cell) =  1.00
D-cell: 20.44 25.56  0.78  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 10000  k =  31.25  q(D-cell) =  4.79  q(S-cell) =  1.00
D-cell: 21.03 26.38  0.49  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 20000  k =  38.65  q(D-cell) =  5.12  q(S-cell) =  1.00
D-cell: 22.44 28.71  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 30000  k =  44.92  q(D-cell) =  5.14  q(S-cell) =  1.00
D-cell: 22.24 29.16  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 40000  k =  51.30  q(D-cell) =  5.17  q(S-cell) =  1.00
D-cell: 22.05 29.63  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 50000  k =  57.78  q(D-cell) =  5.20  q(S-cell) =  1.00
D-cell: 21.87 30.12  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 100000  k = 91.23  q(D-cell) =  5.38  q(S-cell) =  1.00
D-cell: 21.04 32.78  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 150000  k = 125.77  q(D-cell) =  5.60  q(S-cell) =  1.00
D-cell: 20.32 35.68  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 200000  k = 160.84  q(D-cell) =  5.84  q(S-cell) =  1.00
D-cell: 19.66 38.70  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 250000  k = 196.27  q(D-cell) =  6.07  q(S-cell) =  1.00
D-cell: 18.98 41.75  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 300000  k = 232.13  q(D-cell) =  6.31  q(S-cell) =  1.00
D-cell: 18.29 44.81  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 350000  k = 267.99  q(D-cell) =  6.54  q(S-cell) =  1.00
D-cell: 17.56 47.84  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99
```

```
t = 400000  k = 304.61  q(D-cell) =  6.77  q(S-cell) =  1.00
D-cell: 16.78 50.89  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 450000  k = 341.23  q(D-cell) =  6.99  q(S-cell) =  1.00
D-cell: 15.97 53.89  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 500000  k = 377.72  q(D-cell) =  7.20  q(S-cell) =  1.00
D-cell: 15.11 56.85  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 550000  k = 413.29  q(D-cell) =  7.40  q(S-cell) =  1.00
D-cell: 14.25 59.70  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 600000  k = 448.89  q(D-cell) =  7.59  q(S-cell) =  1.00
D-cell: 13.37 62.53  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 650000  k = 484.52  q(D-cell) =  7.78  q(S-cell) =  1.00
D-cell: 12.46 65.34  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 700000  k = 520.39  q(D-cell) =  7.96  q(S-cell) =  1.00
D-cell: 11.51 68.12  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 750000  k = 557.01  q(D-cell) =  8.15  q(S-cell) =  1.00
D-cell: 10.53 70.95  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 800000  k = 593.40  q(D-cell) =  8.33  q(S-cell) =  1.00
D-cell:  9.54 73.74  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 900000  k = 663.15  q(D-cell) =  8.67  q(S-cell) =  1.00
D-cell:  7.60 79.07  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1000000  k = 730.29  q(D-cell) =  8.99  q(S-cell) =  1.00
D-cell:  5.71 84.17  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1100000  k = 797.08  q(D-cell) = 9.30  q(S-cell) =  1.00
D-cell:  3.83 89.22  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1200000  k = 856.50  q(D-cell) = 9.59  q(S-cell) =  1.00
D-cell:  2.18 93.75  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99
```

```
t = 1300000  k = 903.06  q(D-cell) = 9.82  q(S-cell) =  1.00
D-cell:  0.91 97.33  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1400000  k = 926.71  q(D-cell) = 9.95  q(S-cell) =  1.00
D-cell:  0.28 99.17  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1500000  k = 934.69  q(D-cell) = 9.99  q(S-cell) =  1.00
D-cell:  0.07 99.79  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1600000  k = 935.10  q(D-cell) = 9.99  q(S-cell) =  1.00
D-cell:  0.06 99.82  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1700000  k = 935.10  q(D-cell) = 9.99  q(S-cell) =  1.00
D-cell:  0.06 99.82  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1800000  k = 935.10  q(D-cell) = 9.99  q(S-cell) =  1.00
D-cell:  0.06 99.82  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 1900000  k = 935.10  q(D-cell) = 9.99  q(S-cell) =  1.00
D-cell:  0.06 99.82  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99

t = 2000000  k = 935.10  q(D-cell) = 9.99  q(S-cell) =  1.00
D-cell:  0.06 99.82  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
S-cell:  1.00  1.00  1.00  1.00  1.01  1.00  1.00  1.00  1.00  0.99
```

Bibliography

Agin, D. (1964) Hodgkin-Huxley equations: Logarithmic relation between membrane current and frequency of repetitive activity. Nature 201: 625-626.

Amari, S. (1974) A method of statistical neurodynamics. Kybernetik 14: 201-215.

Anderson, J. A. (1972) A Simple Neural Network Generating an Interactive Memory. In: Mathematical Biosciences. vol. XIV. American Elsevier, New York.

Anderson, J. A. (1973) A theory for the recognition of items from short memorized lists. Psych. Rev. 80: 417-438.

Anderson, J. A. (1977) Distinctive features, categorical perception, and probability learning: Some applications of a neural model. Psych. Rev. 84: 413-451.

Barlow, H. B. (1972) Single units and sensation: A neuron doctrine for perceptual psychology? Perception 1: 371-394.

Bienenstock, E. (1980) A theory of development of neuronal

selectivity. Ph.D. Thesis. Brown University.

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. 28: 32-48.

Blakemore, C. (1976) The conditions required for the maintenance of binocularity in the kitten's visual cortex. J. Physiol. 261: 423-444.

Blakemore, C. and Van Sluyters, R. C. (1975) Innate and environmental factors in the development of the kitten's visual cortex. J. Physiol. 248: 663-716.

Bliss, T. V. P. and Lomo, T. (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforent path. J. Physiol. (Lond.) 232: 331-356.

Bruce, C., Desimone, R. and Gross, C. G. (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J. Neurophysiol. 46: 369-384.

Buisseret, P. and Imbert, M. (1976) Visual cortical cells. Their developmental properties in normal and dark reared kittens. J. Physiol. 255: 511-525.

Calvet, M.-C. and Calvet, J. (1979) Horseradish peroxidase iontophoretic intracellular labelling of cultured Purkinje cells. Brain Res. 173: 527-531.

Chapman, R. A. (1966) The repetitive responses of isolated axons from the crab, Carcinus Maenas. J. Exp. Biol. 45: 475-488.

Clark, G. A., McCormick, D. A., Lavond, D. G., Baxter, K., Gray, W. J. and Thompson, R. F. (1982) Effects of electrolytic lesions of cerebellar nuclei on condtioned behavioral and hippocampal neuronal responses. Soc. Neurosci. Abstr. 8: 22.

Cooper, L. N (1973) A possible organization of animal memory and learning. In: Proceedings of the Nobel Symposium on Collective Properties of Physical Systems Vol. 24, pp. 252-164. B. Lindquist & S. Lindquist, eds. Academic Press, New York.

Cooper, L. N, Munro, P. W. and Scofield, C. L. (to be published) Neuron selectivity: Single neuron and neuron networks.

Culbertson, J. T. (1948) A mechanism for optic nerve conduction and form perception: I. Bull. Math. Biophys. 10: 31-40.

Douglas, R. M. and Goddard, G. V. (1975) Long-term potentiation of the perforant path - granule cell synapse in the rat hippocampus.

Brain Res. 86: 205-215.

Gati, I. and Tversky, A. (1982) Representations of qualitative and quantitative dimensions. J. Exp. Psych. 8: 325-340.

Gross, C. G., Rocha-Miranda, C. E. and Bender, D.B. (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J. Neurophysiol. 35: 96-111.

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

Hendrickson, A. E., Hunt, S. P. and Wu, J.-Y. (1981) Immunocytochemical localization of glutamic acid decarboxylase in monkey striate cortex. Nature 292: 605-607.

Hodgkin, A. L. and Huxley, A. F. (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. 117: 500-544.

Hornung, J. P. and Garey, L. J. (1981) The thalamic projection to cat visual cortex: Ultrastructure of neurons identified by Golgi impregnation or horseradish peroxidase transport. Neuroscience. 6: 1053-1068.

Hubel, D. H. and Livingstone, M. S. (1982) Cytochrome oxidase blobs

in monkey area 17: Response properties and afferent connections. Soc. Neurosci. Abstr. 8: 706.

Hubel, D. H. and Wiesel, T. N. (1959) Receptive fields of single neurons in the cat striate cortex. J. Physiol. 148: 574-591.

Hubel, D. H. and Wiesel, T. N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160: 106-154.

Hubel, D. H. and Wiesel, T. N. (1963) Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. J. Neurophysiol. 26: 994-1002.

Hubel, D. H. and Wiesel, T. N. (1965) Binocular interaction in striate cortex of kittens with artificial squint. J. Neurophysiol. 28: 1041-1059.

Hubel, D. H. and Wiesel, T. N. (1977) Functional architecture of macaque monkey visual cortex. Proc. R. Soc. Lond. B. 198: 1-59.

Kelly, J. P. and van Essen, D. C. (1974) Cell structure and function in the visual cortex of the cat. J. Physiol. 238: 515-547.

Kohonen, T. (1977) Associative Memory: A System Theoretical Approach. Springer, Berlin / Heidelberg / New York.

Lashley, K. S. (1929) Brain Mechanisms and Intelligence. University of Chicago Press.

Leventhal, A. G. and Hirsch, H. V. B. (1980) Receptive field properties of different classes of neurons in visual cortex of normal and dark-reared cats. J. Neurophysiol. 43: 1111-1132.

Levy, W. B. and Desmond, N. L. (to be published) The rules of elemental synaptic plasticity.

Levy, W. B. and Steward, O. (1979) Synapses as associative memory elememts in the hippocampal formation. Brain Res. 175: 233-245.

Longuet-Higgins, H. C. (1968) The non-local storage of temporal information. Proc. Roy. Soc. B. 171: 327-334.

McClelland, J. L. and Rumelhart, D. E. (1981) An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. Psych. Rev. 88: 375-407.

McCulloch, W. S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5: 115-133.

Marr, D. (1969) A theory of cerebellar cortex. J. Physiol. 202: 437-470.

Movshon, J. A. (1976) Reversal of the physiological effects of monocular deprivation in the kitten's visual cortex. J. Physiol. 261: 125-174.

Nass, M. M. and Cooper, L. N (1975) A theory for the development of feature detecting cells in visual cortex. Biol. Cybernetics 19: 1-18.

Poggio, T. and Koch, C. (1981) Nonlinear interactions in a dendritic tree: localization, timing and role in information processing. M.I.T. A.I. Memo No. 657.

Posner, M. I. and Keele, S. W. (1968) On the genesis of abstract ideas. J. Exp. Psych. 77: 353-363.

Posner, M. I. and Keele, S. W. (1970) Retention of abstract ideas. J. Exp. Psych. 83: 304-308.

Rauschecker, J. P. and Singer, W. (1981) The effects of early visual experience on the cat's visual cortex and their possible explanation by Hebb synapses. J. Physiol. 310: 215-240.

Rosch, E. (1975) Cognitive reference points. Cognitive Psychology 7: 532-547.

Sejnowski, T. J. (1977a) Storing covariance with nonlinearly interacting neurons. J. Math. Biol. 4: 303-321.

Sejnowski, T. J. (1977b) Statistical constraints on synaptic plasticity. J. Theor. Biol. 69: 385-389.

Shepard, R. N. and Podgorny, P. (1975) Cognitive processes that resemble perceptual processes. In: Handbook of Learning and Cognitive Processes, vol. 5. pp. 189-237. W. K. Estes, ed. Lawerence Erlbaum Associates, Hillsdale, N. J.

Sherrington, C. (1906) The Integrative Action of the Nervous System. Charles Scribner's Sons. Reprinted by Cambridge University Press, 1948.

Sherrington, C. (1940) Man on his Nature. Cambridge University Press.

Sillito, A. M. (1975) The contribution of inhibitory mechanisms to the receptive field properties of neurons in the striate cortex of the cat. J. Physiol. 250: 305-329.

Skinner, B. F. (1938) The Behavior of Organisms. Appleton-Century-Crofts, Inc. New York.

Stent, G. S. (1973) A physiological mechanism for Hebb's postulate of

learning. Proc. Nat. Acad. Sci. USA 70: 997-1001.

Steward, O. and Vinsant, S. L. (1978) Collateral projections of cells in the surviving entorhinal area which reinnervate the dentate gyrus of the rat following unilateral entorhinal lesions. Brain Res. 149: 216-222.

Thorndike, E. L. (1913) The Psychology of Learning. Teachers College, Columbia University, New York.

Thorndike, E. L. (1932) The Fundamentals of Learning. Teachers College, Columbia University, New York.

Tootell, R. B. H., Silverman, M. S., Switkes, E. and De Valois, R. L. (1982) The organization of cortical modules in primate striate cortex. Soc. Neurosci. Abstr. 8: 707.

Tversky, A. (1977) Features of similarity. Psych. Rev. 84:327-352.

Tversky, A. and Gati, I. (1978) Studies of similarity. In: Cognition and Categorization. E. Rosch and B. B. Lloyd eds. Lawerence Erlbaum Associates, Hillsdale, N. J.

Tversky, A. and Gati, I. (1982) Similarity, separability and the triangle inequality. Psych. Rev. 89: 123-154.

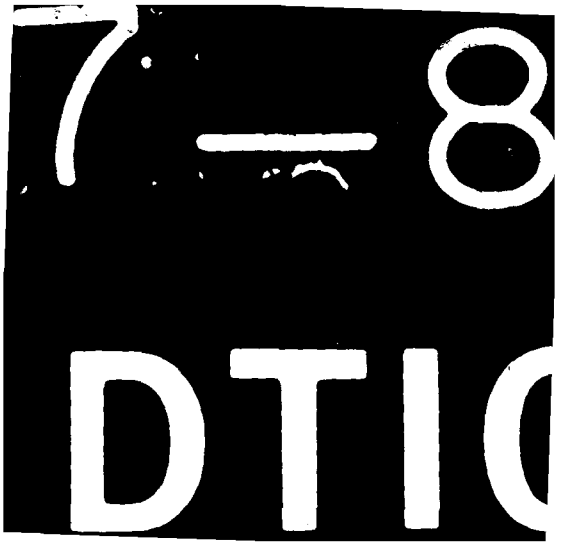
van der Loos, H. and Glaser, E. M. (1972) Autapses in neocortex cerebri: synapses between a pyramidal cell's axon and its own dendrites. Brain Res. 48: 355-360.

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

von Fieandt, K. and Moustgaard, I. K. (1977) The Perceptual World. Academic Press Inc., New York.

Wiesel, T. N. and Hubel, D. H. (1963) Single-cell responses in striate cortex of kittens deprived of vision in one eye. J. Neurophysiol. 26: 1003-1017.

Yeo, C. H., Hardiman, M. J., Glickstein, M. and Russell, I. (1982) Lesions of cerebellar nuclei abolish the classically conditioned nicitating membrane response. Soc. Neurosci. Abstr. 8: 22.


7—8
DTIC