

Pattern Formation, Contrast Control, and Oscillations in the Short Term Memory of Shunting On-Center Off-Surround Networks

Samuel A. Elias* ** and Stephen Grossberg***

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Received: January 30, 1975

Abstract

The transformation of spatial patterns and their storage in short term memory by shunting neural networks are studied herein. Various mechanisms are described for real-time regulation of the amount of contrast with which a pattern will be stored. Parametric studies are described for the amount of contrast in the network responses to patterns presented at variable background or overall activity levels. Mechanisms for removing spurious peak splits and other disinhibitory responses are described. Furman's (1965) results on processing of patterns by shunting networks are generalized and reanalysed. Periodic responses (stable and unstable) corresponding to the time scale of slow cortical waves can be generated if a tonic input is set between two threshold activity levels. Their frequency as a function of tonic input size is unimodal. Order-preserving limit cycles are never found in STM; hence sustained slow oscillations as a mechanism for storing a pattern in STM are ruled out in favor of steady states (i.e., fast oscillations) with spatially graded activity levels. Such slow oscillations can, nonetheless, continuously retune the network's responsiveness to the patterns that perturb it.

1. Introduction

This paper studies the global dynamics of neurons, or neuron populations, in a recurrent on-center off-surround anatomy undergoing nonlinear shunting interactions. In such an anatomy, a given population excites itself (and possibly nearby populations) and inhibits populations that are further away (and possibly itself and nearby populations also). Such an anatomical design is found in many neural structures. For example, in hippocampus (Anderson *et al.*, 1969; Dichter and Spencer, 1969; Kandel, Spencer, and Brinley, 1961; Lebovitz, Dichter, and Spencer, 1971)

* Supported in part by the Advanced Research Projects Agency under the Office of Naval Research (N00014-70-A-0362-0006).

** After January 1, 1975, Dr. Elias' address will be Department of Applied Mathematics, California Institute of Technology, Pasadena, California.

*** Supported in part by the Office of Naval Research (N00014-67-A-0204-0051) and the Advanced Research Projects Agency (DAHC 15-73-C-0320) administered by Computer Corporation of America.

the main cell type, the pyramidal cell, emits axon collaterals to interneurons. Some of these collaterals or interneurons feed back excitatory signals to nearby pyramidal cells. Other interneurons generate inhibitory feedback signals over a broad area. Variations on this anatomical theme are found in other brain areas, such as neocortex (Eccles, 1965; Phillips, 1959; Stefanis, 1969). Grossberg (1973) introduced a class of recurrent on-center off-surround networks which were shown capable of contrast enhancing significant input information; sustaining this information in short term memory; producing multistable equilibrium points that normalize, or adapt, the field's total activity; suppressing noise; and preventing saturation of population response even to input patterns whose intensities are high. These formal results were then applied to a variety of situations, including aspects of motor control, pattern discrimination, and drive-reinforcer interactions (Grossberg, 1973), as well as problems in attention and discrimination learning, such as overshadowing and reinforcement contrast effects (Grossberg, 1975).

That paper described mathematical results for the subclass of on-center off-surround networks where there are no structural biases in favor of one cell population over others. Grossberg and Levine (1975) extended those results to networks where a particular kind of bias exists; namely, where the number of excitable sites in different populations could differ. It was shown that this bias introduced a new form of contrast enhancement into the network, that could be interpreted as resulting from shifts in attention, or from developmental differences in the size of populations that code for a given sensory feature in a field of feature detectors, or even from statistical imperfections in the network's construction. Levine and Grossberg (1975) also studied cases in which the strength of excitatory and inhibitory interactions both decrease monotonically with distance, for example at Gaussian rate. Using such a network, they

simulated various sensory phenomena, such as the hysteresis effect that is seen when two binocularly superimposed lines are slowly separated until they are distinct and then slowly brought back together (Julesz, 1971). They also simulated sensory phenomena in networks where the interaction strengths decrease with distance, and the saturation levels, or equivalently the tuning curves, of different populations can differ; for example, the Gibson-Radner tilt aftereffect (Gibson and Radner, 1937) and the Blakemore angle-expansion effect (Blakemore, Carpenter, and George-son, 1970; Blakemore and Cooper, 1970).

These studies illustrate the remarkable richness of properties to be found in reverberating on-center off-surround networks undergoing shunting interactions. They do not, however, exhaust these properties, and they all make an assumption which is not always true *in vivo*, and which certainly must be investigated to achieve conceptual clarity. The assumption is that the networks are *lumpable*; in other words, that the inhibitory potentials of the network follow their excitatory inputs so quickly that they can be expressed in terms of these inputs at all times. In nonrecurrent, or feedforward, networks, this is not always true; for example, it has been shown in the mudpuppy retina (Werblin, 1971) that the horizontal cells mediate light adaptational effects, whose time course is often slow compared to the excitatory response rate to fluctuating visual signals. In recurrent networks, such as hippocampus (Eccles, 1965), inhibitory influences on the pyramidal cells can have a short latency and fast rise time, but can also have a duration (200–600 msec) that is far longer than that of simultaneously occurring excitatory influences (10–20 msec). Thus two main tasks of this work are to further delineate the capabilities of lumped shunting networks and to explore how unlumped networks differ from lumped networks. In particular, under what circumstances does an unlumped network become asymptotically lumped (i.e., act as if it is lumped as $t \rightarrow \infty$)? What capabilities for data processing does an unlumped network have that a lumped network does not have?

2. Review of Lumped Network Properties

In order to motivate our studies, we review the derivation of the networks below, using Grossberg and Levine (1975) as a basic reference. The networks describe the properties of interacting populations of cell sites. The populations can be interpreted as populations of small membrane patches on individual cells, or else as populations of whole cells. Cell sites

in a population are distributed randomly with respect to interactions within each population and between population pairs. The network can therefore be described by mass action laws governing the average potentials of its populations. The interactions between populations will be chosen multiplicative, or of shunting type, as occurs (say) in passive membranes (Hodgkin, 1964; Sperling, 1970; Sperling and Sondhi, 1968), in models of dendritic branches (MacGregor, 1968; Rall and Rinzel, 1973) and in experiments on cat motoneurons (Kuno and Miyahara, 1969).

The network discussed in Grossberg (1973) have the following form. Denote the average activity (e.g., potential) at time t of the i^{th} population v_i by $x_i(t)$, $i = 1, 2, \dots, n$. Let each population excite itself and inhibit other populations via the system of equations

$$\dot{x}_i = -Ax_i + (B - x_i)f(x_i) - x_i \sum_{k \neq i} f(x_k) + I_i, \quad (1)$$

where $i = 1, 2, \dots, n$, B is the "weight", i.e., maximum possible activity, of the i^{th} population v_i , and $x_i (0 \leq x_i \leq B)$ is the mean activity of v_i , interpretable either as a voltage or a number of active sites. Four effects determine the behavior of the i^{th} population v_i : (1) exponential decay, via the term $-Ax_i$; (2) self-excitation, via the term $(B - x_i)f(x_i)$; (3) shunting inhibition from other populations, via the term $-x_i \sum_{k \neq i} f(x_k)$; and (4) externally applied inputs, via the term I_i . The function $f(w)$ describes the mean output signal of a population as a function of its mean activity w .

What do the inhibitory surround and the reverberation accomplish? To motivate these interactions, consider what goes wrong in a system without interactions whose responses $x_i(t)$ to nonnegative inputs $I_i(t)$ have the following natural properties: (1) linearity; (2) boundedness, say by B ; (3) decay to equilibrium point, say 0, after inputs cease. Then

$$\dot{x}_i = -Ax_i + (B - x_i)I_i(t), \quad (2)$$

with $0 \leq x_i(0) \leq B$. Suppose that the relative sizes of the responses $x_i(t)$ code the relative importance of the information contained in the inputs $I_i(t)$. For example, if different populations v_i are excited by different features in a visual scene (e.g., colors, lines, edges, disparities), then the relative intensities $\Theta_i = I_i(\sum_k I_k)^{-1}$ of the inputs measure how much of that feature is in the scene. However, the equilibrium values of the responses $x_i(t)$, defined by $\dot{x}_i = 0$, satisfy

$$x_i = \frac{B\Theta_i I_i}{A + \Theta_i I_i}, \quad (3)$$

where $I = \sum_k I_k$. Thus $x_i \rightarrow B$ as I becomes large, and the measure of relative importance is lost due to saturation. If also this system contains noise, then the responses will not accurately measure relative importance when I is small. Hence the system is inadequate both at low and high total input intensities. Grossberg (1973) noted that an off-surround prevents this problem, since if

$$x_i = -A x_i + (B - x_i) I_i - x_i \sum_{k \neq i} I_k, \quad (4)$$

then the equilibrium value of x_i is

$$x_i = \Theta_i \frac{BI}{A + I}, \quad (5)$$

which is proportional to Θ_i , no matter how large I is chosen. The off-surround hereby introduces a type of adaptation to overall levels of input activity.

In system (4), the constant B is the number of excitable sites in each population v_i . Since x_i is the mean number of active sites, $B - x_i$ is the mean number of inactive sites. Thus term $(B - x_i) I_i$ says that inactive sites are activated at a rate proportional to the number of inactive sites times the excitatory input intensity. Term $-x_i \sum_{k \neq i} I_k$ says that active sites are inactivated at a rate proportional to the number of active sites times the total inhibitory input intensity. This is the meaning of "mass action" in this context.

The above example uses a nonrecurrent, or feed-forward, anatomy. A recurrent, or reverberating, anatomy is introduced when a network capable of short term memory (STM) is needed. Such a network can reverberate a pattern of activity distributed over cell populations for an indefinite interval of time. This reverberation can also be switched off rapidly by inhibitory inputs if a new pattern is delivered by external sources; the decay rates of individual cells can be large after the excitatory reverberating loop is broken by inhibition, even if the reverberation through an active excitatory loop is long lived. The use of reverberation as a mechanism of STM is, for example, suggested by operant conditioning experiments; here one is led to seek reverberatory processes that can maintain in short term storage internal representations of sequences of external events until later rewards or punishments occur and transfer the memory of these sequences to long term storage (Grossberg, 1971).

Two main themes emerge in the discussion of reverberating networks. The first is: how does the reverberation change the distribution of activity across populations through time? In particular, how

does the network suppress noise, or behaviorally irrelevant inputs, yet store behaviorally significant data in STM? This theme focusses on ways in which the *relative* sizes of population activities are transformed. The second theme is concerned with fluctuations in *total* network activity through time. In particular, when does the total activity converge rapidly to zero, so that no stable reverberation gets established? A deeper issue is illustrated by Eqs. (1) and (2). In Eq. (1), the *maximum* of the total activity $x(t) = \sum_{k=1}^n x_k(t)$ when m populations are active is mB , since each active population has a maximum activity of B . In Eq. (2), the maximum of the total activity is B , since $x(t) = \frac{BI}{A + I}$, which converges to B as $I \rightarrow \infty$.

Thus the maximum total activity is independent of the number of active populations. This result from nonrecurrent networks has an analog in recurrent networks. Grossberg (1973) has shown that if a persistent reverberation is established, then $x(t)$ will converge to a unique positive limit point as $t \rightarrow \infty$, for suitable choices of the signal function $f(w)$. It is also possible to find signal functions for which $x(t)$ converges to any one of a discrete set of limit points ("multi-stable equilibrium"), or even to a continuum of limit points. The situation in which one, or at most a few, limit points of total activity exist is typical. In all these cases, there is an upper bound on possible values of $x(\infty)$ that is independent of n . This property is called *normalization*.

An important theme about total activity normalization is the following: if the network can reverberate patterns in STM that are imposed by behaviorally relevant inputs, then what prevents the network from reverberating behaviorally irrelevant activity levels, such as noise? Grossberg (1973) shows that a proper choice of the signal function $f(w)$ overcomes this dilemma. This is seen by classifying the properties of various choices of $f(w)$. This classification is summarized below.

Case 1. Linear Signal Function: $f(w) = Cw$. Every initial pattern is preserved perfectly by this reverberation; that is, the relative activity functions $X_i(t) = x_i(t) \left[\sum_{k=1}^n x_k(t) \right]^{-1}$ are constant. Moreover the total activity $x(t) = \sum_{k=1}^n x_k(t)$ is normalized: the limit $x(\infty) = \lim_{t \rightarrow \infty} x(t)$ exists, and equals zero or a unique positive constant E . Unfortunately, $x(\infty)$ has the same value *independent of the initial data*. For example, if the system does not reverberate noise [i.e., $x(0) \cong 0$ and

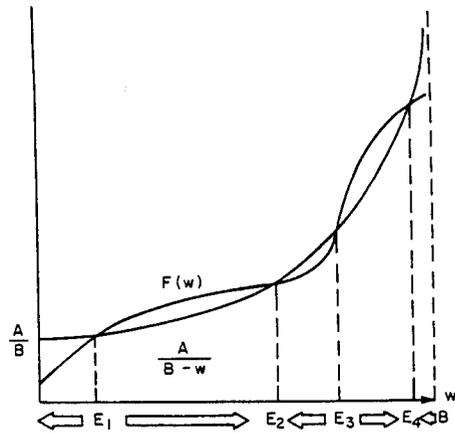


Fig. 1. Total activity normalization with multiple equilibrium points

$x(\infty)=0$], then the system does not reverberate important inputs [i.e., $x(0) \gg 0$ and $x(\infty)=0$]; or if the system does reverberate important inputs [i.e., $x(0) \gg 0$ and $x(\infty)=E$], then it also amplifies and reverberates noise [i.e., $x(0) \cong 0$ and $x(\infty)=E$]. Such a system cannot distinguish between unimportant and important inputs to be stored in STM.

This choice therefore has a good property and a bad property. The good property is that it can store any pattern without distortion. The bad property is that it amplifies noise. Case III below begins to show how to eliminate the bad property using a nonlinear signal function.

Case II. Slower-than-Linear Signal Function; i.e., $f(w) = wF(w)$, with $F(w)$ decreasing. The reverberation either dies out or is normalized. If it is normalized, the limiting distribution of pattern weights is uniform; i.e., $X_i(\infty) = \frac{1}{n}$, $i = 1, 2, \dots, n$. Thus a population that

is subjected to noise will ultimately have the same weight as a population that receives a large signal. This situation is even more pathological than the linear case, since there, in the presence of signals, populations which receive only noise will maintain a small relative weight.

Case III. Faster-than-Linear Signal Function; i.e., $f(w) = wF(w)$, with $F(w)$ increasing. The main problem posed by the previous two cases is to suppress noise, or at least to prevent noise amplification. This occurs in the present case.

Consider Fig. 1. If $x(0) < E_1$, then $x(\infty) = 0$; the value E_1 defines the level below which (total) initial activity is treated as noise and therefore suppressed. All initial values $x(0) > E_1$ lead to storage in STM. The values E_2, E_4, \dots, E_{2m} , etc. are stable equilibrium points of total activity. There can be any number of

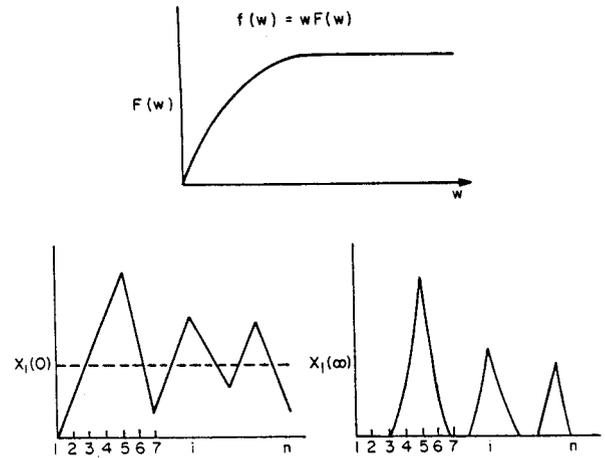


Fig. 2. Contrast enhancement and storage in STM of suprathreshold activities

these points if $F(w)$ is suitably defined. In particular, if $F(w)$ is convex, say because $f(w)$ is a sigmoid function of w , then there is a unique positive equilibrium point.

The property of noise suppression implies that the pattern weights are not preserved. Indeed, only those populations whose initial activities are maximal are reverberated in STM; the activities of all other populations are suppressed. In particular, if one population initially is more active than any other population, then the reverberation "chooses" this population for storage in STM. The faster-than-linear case suppresses too much of a pattern in order to suppress noise. We seek a way to preserve the property of noise suppression without suppressing also all but the maxima of an input pattern. Cases I and III suggest a way.

Case IV. Faster-than-Linear Levelling off to Linear. All populations whose initial activity falls below a prescribed threshold level will be quenched by the reverberation. The pattern of activity of all initially suprathreshold populations will be contrast enhanced and stored in STM (Fig. 2).

Speaking heuristically, the reason for this phenomenon is as follows. Suppose that an initial pattern of activity falls within the faster-than-linear region of $f(w)$. By Case III, contrast enhancement will begin. Were there no other region of $f(w)$, only the pattern maxima would survive. However, normalization of the total activity $x(t)$ also takes place, and can carry the pattern activities into the linear range. By Case I, any pattern in the linear range will be preserved by the reverberation; in particular, the partially contrast-enhanced pattern will be preserved. Thus the linear range terminates the contrast enhancement pro-

cedure which, if uninterrupted, would annihilate all but the pattern maxima.

Case V. Sigmoid Signal Function. *In vivo*, signal functions always have finite maxima at large activities. By levelling off the signal function of Case IV, one finds a sigmoid $f(w)$ (Fig. 3).

By Case II, we know that at activities corresponding to the slower-than-linear range of $f(w)$, the pattern will be uniformized. This unfortunate property can be eliminated by a suitable choices of parameters even if the maximal activity level B falls within the slower-than-linear range of $f(w)$.

In summary, the linear [i.e., $F(w) = \text{constant}$] range of a signal function $f(w)$ determines the network's ability to store *partially* contrast-enhanced patterns in STM. Insofar as $F(w)$ deviates only slightly from constancy in a given range, the corresponding $f(w)$ will tend to *slowly* contrast enhance or uniformize the pattern weights that it influences in this range.

Grossberg and Levine (1975) studied the system

$$\dot{x}_i = -Ax_i + (B_i - x_i)f(x_i) - x_i \sum_{k \neq i} f(x_k) + I_i \quad (6)$$

or equivalently

$$\dot{x}_i = -Ax_i + (B - x_i)f(C_i x_i) - x_i \sum_{k \neq i} f(C_k x_k) + I_i \quad (7)$$

In (6), different features coded by the populations v_i can have different saturation weights B_i . In (7), both the excitatory and inhibitory interactions of each population can be shunted by a different scaling factor C_i . The asymmetries in B_i and C_i can be attributed to developmental preferences or to shifts in attentional bias due to a shunt either of a population's maximal excitability or *all* of its interactions, both excitatory and inhibitory, by a suitable arousal signal. These authors show that such asymmetries introduce a new form of contrast enhancement into the system. For example, suppose that $B_1 \leq B_2 \leq \dots \leq B_n$ in (6). If $f(w) = Cw$, then all $x_i(\infty) = 0$ if $B_i < B_n$, and $x_i(\infty) = Kx_i(0)$ if $B_i = B_n$. In other words, all states with nonmaximal B_i are quenched, and all other traces are stored faithfully in STM. No states are quenched if all $B_i = B$. If $f(w)$ is chosen as in Fig. 2, then there is a competition between relative sizes of the B_i and the $x_i(0)$. Only the traces corresponding to one B_i can be stored in STM, but i need not equal n if the initial data of populations v_j such that $B_j = B_i$ exceeds the initial data of populations v_j such that $B_j = B_n$ (Fig. 4).

The pattern that is stored is also contrast-enhanced. Thus attentional shunts or developmental asymmetries can bias the network in favor of one class of features

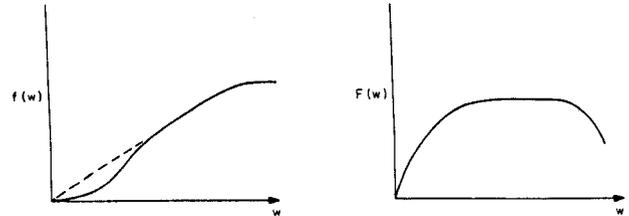


Fig. 3. Sigmoid signal function

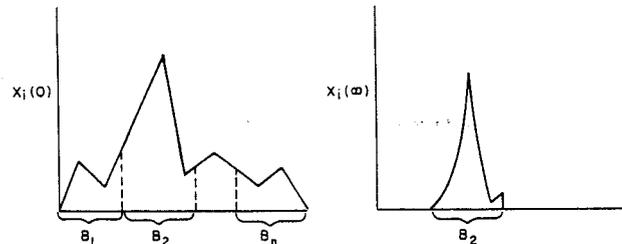


Fig. 4. Competition between number of sites B_i and initial data

(namely those with $B_i = B_n$), but if a different class of features is present with sufficient strength at a given time, it can overshadow the preferred class in STM. The reverse statement is also true. If infrequently experienced features have small B_i , then the reverberation can totally quench activity in these features. In other words, the infrequent-feature detectors are there because they might be needed, but they need not be a source of noise or other interference when different features are processed.

3. Relative Excitatory-to-Inhibitory Gain Controls Contrast

System (6) enjoys two different contrast-enhancing mechanisms:

- The nonlinear signal function $f(w)$.
- Nonuniform choice of B_i or C_i .

A third mechanism of contrast exists in systems of the form

$$\dot{x}_i = -Ax_i + (B_i - x_i)f(x_i) - x_i \sum_{k \neq i} g(x_k) + I_i \quad (8)$$

where the excitatory signal function $f(w)$ and the inhibitory signal function $g(w)$ differ. For example, set $f(w) = Cw$ and $g(w) = Dw$, where $C \neq D$. If $C = D$ and all $B_i = B$, then this system preserves patterns. If however all $B_i = B$ but $C \neq D$, then this property is dramatically changed. Thus we consider

$$\dot{x}_i = -Ax_i + (B - x_i)Cx_i - Dx_i \sum_{k \neq i} x_k \quad (9)$$

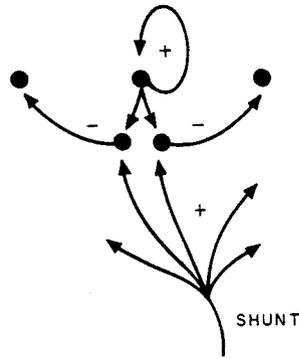


Fig. 5. Recurrent anatomy with nonspecific shunting input to inhibitory interneurons

$i = 1, 2, \dots, n$ when $C \neq D$. How do the probabilities $X_i = x_i \left(\sum_{k=1}^n x_k \right)^{-1}$ and the total activity $x = \sum_{k=1}^n x_k$ behave as $t \rightarrow \infty$? The following definitions will be used to discuss this situation.

Definition 1. The reverberation is *persistent* if $x(t)$ does not approach zero as $t \rightarrow \infty$.

Definition 2. The limiting distribution $X_i(\infty) = \lim_{t \rightarrow \infty} X_i(t)$ is *uniform* if all $X_i(\infty) = \frac{1}{n}$.

Definition 3. The limiting distribution is *0-1* if $X_i(\infty) = 1$ for some i .

Definition 4. The limiting distribution is *locally uniform* if $X_i(\infty) = 0$ or $\frac{1}{m}$ for all i , and $1 < m < n$.

Definition 5. The reverberation is *normalized* if the possible limit points of $x(\infty)$ have an upper bound that is independent of n .

In all results below, the initial data satisfies $0 \leq x_i(0) \leq B$. The following theorem describes the effect of varying the sign of $C - D$. All proofs are given in Appendix 1.

Theorem 1. Let the initial data of (9) satisfy $0 \leq x_i(0) \leq B$. Then all limits $x_i(\infty)$ exist. Suppose moreover that the reverberation is persistent, which occurs if and only if $BC > A$. If $D < C$, then the limiting distribution is uniform, and $x(\infty) = n(BC - A) \cdot [(n-1)D + C]^{-1} \leq (BC - A)D^{-1}$. If $D > C$, then the limiting distribution is 0-1 or locally uniform depending on whether or not the number m of populations that satisfy $X_i(0) = \max_k X_k(0)$ is $m = 1$ or $m > 1$. In this case, $x(\infty) = m(BC - A) [(m-1)D + C]^{-1} \leq B - AC^{-1}$, so that the reverberation is normalized.

In other words, if inhibition is weaker than excitation, then the distribution becomes uniform and all information about initial relative importance is lost. If inhibition is stronger than excitation, then

the system chooses the maximal initial data for storage in STM and suppresses all other data. If $D \neq C$ but D is close to C , then these tendencies to uniformize or contrast enhance initial data set in slowly. For example if $D > C$ and $|D - C| \cong 0$, then in a finite time interval, only partial contrast enhancement can occur.

Are these good properties or bad ones? They say that by varying the relative gain of excitation to inhibition, one can control the amount of contrast enhancement in the system. This is interesting. However, unless there is a mechanism for rebalancing the excitatory and inhibitory strengths, then the enhancement process cannot be terminated. A provocative mechanism is the following. Suppose that when the system is not active, then $D < C$; i.e., inhibition is weak. All asymmetries in system response, such as noise effects, will tend to be uniformized so that the system can respond without bias when inputs do arrive. Suppose that when an input arrives, it also activates a nonspecific arousal mechanism that strengthens the relative inhibitory-to-excitatory interaction strength in the field. This can be done, for example, by excitatory shunting (i.e., multiplying) of the responses in all the inhibitory interneurons, thereby magnifying their strength, or by inhibitory shunting of the responses in all the excitatory interneurons, thereby weakening their strength—but *not both* with equal strength, in counterdistinction to the attentional shunt described by Grossberg and Levine (1975) (Fig. 5).

If the shunt is strong enough to make $D > C$, then the field will tend to enhance asymmetries created by the input, and to suppress noise, even if the signal function is linear. After the shunting arousal is shut off, the field will actively suppress these asymmetries to prepare itself to receive new inputs without bias. It is also readily proved that if $f(w)$ grows faster-than-linearly and $D > C$, then the system will try to make a choice. This also occurs for such an $f(w)$ if $D = C$. The condition $D > C$ strengthens the tendency towards contrast enhancement.

4. Possible Anatomical Correlates of Contrast Control

The above example shows that microscopic studies of nonspecific afferents to neocortical and other sensory processing areas must carefully analyse whether these afferents reach excitatory, inhibitory, or both kinds of interneurons. In the former two cases, a variable contrast mechanism for all populations is suggested. In the last case, an attentional mechanism is anticipated.

Definition 6. The system (8) with $f \equiv g$ is said to have *matched interactions*.

The case of matched interactions has beautiful mathematical properties, but can it be realized *in vivo*? Is the system (9) with $C = D$ too singular to be realistic? Two remarks are relevant:

A. A signal function is a *statistical* property of a population. To say that $f \equiv g$ means that the populations of excitatory and inhibitory interneurons in the network have the same statistical parameters, such as spiking threshold distributions. This will occur automatically if both types of interneurons are sampled from a common cell type. When does this occur *in vivo*?

B. Let $n = 2$. Then the off-surround consists only of one population. There exist $n = 2$ situations *in vivo* where a continuum of patterns can be stored in STM; for example, the continuum of possible *postural* positions stored in the neural controls of an antagonistic pair of muscles; cf., Grossberg (1973). Such a system cannot be allowed to become 0–1 or uniform whenever STM storage occurs. How strong an argument is this for the existence of matched interactions when $n = 2$ and the anatomy is recurrent on-center off-surround? By way of comparison, see Theorem 5 below. Certainly a statistical study of the signal functions controlling postural regulation will shed valuable light on this case. Another matched $n = 2$ case is described in Grossberg (1972a, b; 1975) for the regulation of net incentive motivation in pairs of positive and negative drive centers.

5. Inhibitory Lag Can Enhance or Uniformize Even in Matched Systems

Does control of contrast by varying the relative excitatory-to-inhibitory gain also hold in unlumped systems? The simplest unlumped generalization of (9) is

$$\dot{x}_i = -A x_i + (B - x_i) C x_i - D x_i \sum_{k \neq i} y_k \quad (10)$$

and

$$\dot{y}_i = E(x_i - y_i), \quad (11)$$

where x_i is the average potential of the i^{th} excitatory population v_i^+ , and $y_i \geq 0$ is the average potential of the i^{th} inhibitory population v_i^- . Systems (10) and (11) allows the inhibitory interneuronal response to lag behind the more rapid excitatory interneuronal response (Fig. 6).

Equation (11) could also be written as

$$\dot{y}_i = -E y_i + F x_i$$

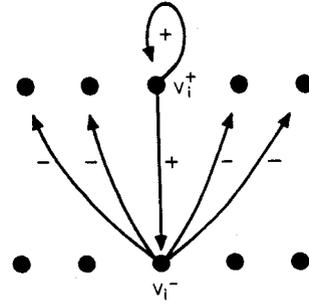


Fig. 6. Unlumped on-center off-surround anatomy

without added generality, since the substitutions $F x_i \rightarrow x_i$, $E y_i \rightarrow y_i$, and an appropriate redefinition of constants in (10) reduce the system to (10) and (11) once again. In the unlumped systems (10) and (11), we will study the probabilities $X_i = x_i \left(\sum_{k=1}^n x_k \right)^{-1}$ and $Y_i = y_i \left(\sum_{k=1}^n y_k \right)^{-1}$ as $t \rightarrow \infty$. In this system, a new possibility arises because the inhibitory population responses to inputs can lag behind the excitatory responses, especially if E is small. To illustrate this possibility, we first discuss the matched case $D = C$. In the *lumped* system with $D = C$, the probabilities $X_i(t)$ are always constant. Also, $x(\infty) = B - AC^{-1}$ and $Y_i \equiv X_i$ trivially. These tendencies are also present in the unlumped system.

Proposition 1. If $D = C$, then the critical points of (10) and (11) are characterized by the equations

$$x = B - AC^{-1} \quad (12)$$

and

$$y_i = x_i. \quad (13)$$

These critical points are asymptotically stable if $n = 2$.

Remarks: If (12) and (13) hold at $t = 0$, then the probabilities $X_i(t)$ are constant, $x(\infty) = B - AC^{-1}$, and $Y_i \equiv X_i$. Also, (12) can clearly hold for an arbitrary distribution X_i . Thus there exist initial data that preserve pattern weights, as in the lumped system. Unlike the lumped system, however, there also exist initial data that contrast enhance or uniformize pattern weights. These and some related concepts are defined below in terms of the functions $M_x(t) = \max_k X_k(t)$, $M_y(t) = \max_k Y_k(t)$, $m_x(t) = \min_k X_k(t)$, $m_y(t) = \min_k Y_k(t)$, $x_{ij}(t) = x_i(t) - x_j(t)$, and $y_{ij}(t) = y_i(t) - y_j(t)$.

Definition 3. The excitatory (inhibitory) pattern is *contrast enhanced* if $M_x(\infty) > M_x(0)$ and $m_x(\infty) < m_x(0)$ [$M_y(\infty) > M_y(0)$ and $m_y(\infty) < m_y(0)$].

Definition 4. The excitatory (inhibitory) pattern is *uniformized* if $M_x(\infty) < M_x(0)$ and $m_x(\infty) > m_x(0)$ [$M_y(\infty) < M_y(0)$ and $m_y(\infty) > m_y(0)$].

Definition 5. *Iso-order* in the pair (i, j) exists at time t if $x_{ij}(t) y_{ij}(t) \geq 0$.

Definition 6. *Anti-order* in the pair (i, j) exists at time t if $x_{ij}(t) y_{ij}(t) < 0$.

Definition 7. *Iso-pattern* exists at time t if $X_i(t) = Y_i(t)$ for all i .

The following theorem illustrates the uniformizing and contrast-enhancing tendencies that are due to the slower inhibitory interneurons. It assumes that limits exist, which has always occurred in our computer runs. It uses the fact that the functions

$$S_{ij} = \log x_i x_j^{-1} + DE^{-1}(y_i - y_j) \quad (14)$$

are constants of the motion if $D = C$.

Theorem 2. Let $D = C$ and $BC > A$. The following cases occur if all limits exist:

A. If the initial $Y_i(0)$ pattern is uniform, then the pattern X_i is uniformized. The amount of uniformization decreases monotonically as a function of E .

B. Let anti-order exist in all pairs at $t=0$, and choose $x_{ij}(0) > 0$ for definiteness. If $S_{ij}(0) > 0$, then y_{ij} changes sign once, and the X_i pattern is uniformized. If $S_{ij}(0) = 0$, then both patterns $X_i(\infty)$ and $Y_i(\infty)$ are uniform. If $S_{ij}(0) < 0$, then x_{ij} changes sign once.

C. Let $n = 2$, suppose that initial iso-order exists, and define

$$S = y(0) - (B - AC^{-1}). \quad (15)$$

If $S > 0$, contrast enhancement occurs. If $S < 0$, uniformization occurs.

Remarks: Case (A) shows that the size of $y(0)$ does not influence the amount of asymptotic uniformization if the $Y_i(0)$'s are uniformly distributed. By contrast, in Case (B), making $y(0)$ larger given a fixed inhibitory pattern $Y_i(0)$ can determine whether or not x_{ij} or y_{ij} reverses its order. Case (C) dramatizes the effect of $y(0)$ in the case of initial iso-pattern. Then if $y(0) > B - AC^{-1}$, contrast enhancement occurs, whereas if $y(0) < B - AC^{-1}$ uniformization occurs. Given general initial $Y_i(0)$, case (C) says that there is a competition between total initial inhibitory activity $y(0)$ and a critical, actually asymptotic, value $B - AC^{-1}$ of total inhibitory activity in determining uniformization or contrast enhancement. Similar effects occur in our computer runs given initial iso-pattern even if $n > 2$. Thus different choices of inhibitory pattern can dramatically change how total inhibitory activity transforms a fixed excitatory pattern.

6. Possible Physiological Correlates of Slow Inhibition

A. Amount of Contrast Depends on E

How does the system respond to an input pattern $I_i(t)$ in a time interval $0 \leq t \leq T$ if $I_1 \leq I_2 \leq \dots \leq I_n$? Start with zero initial data for definiteness in the system

$$\dot{x}_i = -Ax_i + (B - x_i)(Cx_i + I_i) - Cx_i \sum_{k \neq i} y_k \quad (16)$$

and (11). Clearly iso-order exists in all pairs for $t \geq 0$. Compare what happens for $t \geq T$ with E small or large. If E is sufficiently small, then by (11) $0 \cong y(T) \ll B - AC^{-1}$, so that by Remarks (A) and (C) in Section 5, uniformization of the $X_i(T)$ pattern will occur for $t \geq T$. If E is sufficiently large, then $x(t)$ can be driven above $B - AC^{-1}$ by the inputs, and $y(t)$ can follow rapidly behind. If $y(T) > B - AC^{-1}$, then contrast enhancement of the $X_i(T)$ pattern can occur for $t \geq T$, by Remark (C). Thus the rate with which the off-surround responds can influence whether the reverberation enhances the pattern that is created by the inputs. A similar effect occurs while the input is on. If the off-surround can respond quickly, it can, at least partially, overcome the saturation that would occur in response to intense inputs in its absence.

B. Whether a Pattern is Stored Depends on Its Duration

The anti-order case of Theorem 2 illustrates an interesting effect when E is small and the ordering of inputs reverses in time. Let a given sign of y_{ij} be established by earlier inputs, and suppose that a reversed input produces an opposite sign in x_{ij} , but is shut off before y_{ij} changes its sign to that of x_{ij} . This is the anti-order case. If E is sufficiently small, than S_{ij} can have the sign of y_{ij} . Hence x_{ij} will reverse its sign to that of y_{ij} . In other words, if the reversed input is too brief, the order stored in STM will be that of the previous input. Only inputs of sufficient duration will have a reliable effect on STM, and this duration must increase with E , other things equal.

7. Contrast Control in Unlumped Systems

If $D \neq C$, contrast control occurs in the unlumped systems (10) and (11). Theorem 1 shows that, if $D > C$ in the lumped system, then the system strives to reach a 0-1 or locally uniform limiting distribution. If $D > C$ in the unlumped system, the same tendency occurs, but if initial anti-order exists in a pair (i, j) , then the inhibitory order in y_{ij} can compete with

the excitatory order in x_{ij} . This competition creates a tendency towards uniformization in the $D > C$ unlumped system that is not found in the corresponding lumped system. Again the iso-order condition should be thought of as a consequence of an input pattern with fixed order $I_1 \leq I_2 \leq \dots \leq I_n$ that is turned on with sufficient intensity and duration before the reverberation is studied.

Theorem 3. The reverberation is persistent if and only if $BC > A$. If $BC > A$ and $D \neq C$, the limits $X_i(\infty)$ and $Y_i(\infty)$ exist, are equal, and are 0-1, locally uniform, or uniform. If $D < C$, the limiting distribution is uniform. Suppose $D > C$. If initial iso-order exists in all pairs, $x_1(0) \leq x_2 \leq \dots \leq x_n(0)$, and $x_{n1}(0) + y_{n1}(0) > 0$, then the limiting distribution is 0-1 unless $x_{ni}(0) + y_{ni}(0) = 0$ for some $i \neq n$. If some pairs exhibit initial anti-order, then the limiting distribution can be locally uniform even if $x_{ni}(0) + y_{ni}(0) > 0$ for all $i \neq n$.

8. Contrast Control with Nonlinear Signal Functions

The tendency for relative excitatory-to-inhibitory gain to control contrast can be proved for substantially more general systems than (10) and (11). Consider the systems

$$\dot{x}_i = -Ax_i + (B - x_i)f(x_i) - x_i \sum_{k \neq i} g(y_k) \quad (17)$$

$$\dot{y}_i = -Ey_i + h(x_i) \quad (18)$$

The following theorem holds.

Theorem 4. Let the nonnegative signal functions $f(z)$, $g(z)$, and $h(z)$ have positive first derivatives with $h(0) = 0$. Consider the functions

$$U(z) = (Bz^{-1} - 1)f(z) + Kh(z) \quad (19)$$

and

$$V(z) = g(z) - KEz \quad (20)$$

Suppose that there exists a K such that, for $0 \leq z \leq B$, $U(z)$ and $V(z)$ have bounded derivatives of the same sign with at least one derivative never vanishing. Then all limits $X_i(\infty)$ and $Y_i(\infty)$ exist, are equal, and are 0-1, locally uniform, or uniform. If U and V are decreasing, then the limiting distribution is uniform. If U and V are increasing, initial iso-order exists in all pairs, $x_1(0) \leq x_2(0) \leq \dots \leq x_n(0)$, and $x_{n1}(0) + y_{n1}(0) > 0$, then the limiting distribution is 0-1 unless $x_{ni}(0) + y_{ni}(0) = 0$ for some $i \neq n$. If some pairs exhibit initial anti-order, then the limiting distribution can be locally uniform even if $x_{ni}(0) + y_{ni}(0) > 0$ for all $i \neq n$.

Corollary 1. In the case of linear signal functions, U and V are increasing if $D \geq KE \geq C$ with $D > C$,

whereas U and V are decreasing if $D \leq KE \leq C$ with $D < C$. Simply choose $K = DE^{-1}$, as in (14), to get Theorem 3.

Corollary 2. Let $f(z) = \sum_{k=1}^{\infty} f_k z^k$, $g(z) = \sum_{k=1}^{\infty} g_k z^k$,

and $h(z) = \sum_{k=1}^{\infty} h_k z^k$ with all coefficients nonnegative, and with radius of convergence $\geq B$. If for all $k \geq 1$, $g_1 E^{-1} \geq f_k h_k^{-1}$ with at least one term $g_1 E^{-1} - f_k h_k^{-1}$ positive, then the 0-1 tendency exists.

In Eq. (17), iso-order arises naturally when the initial data is uniform and inputs $I_i(t)$ with a fixed order $I_1 \leq I_2 \leq \dots \leq I_n$ perturb the system for $0 \leq t \leq T$. Iso-order holds throughout this interval, and throughout the reverberation in $t \geq T$. This latter fact is a very robust property that is true in any system of the form

$$\dot{x}_i = A(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) B(x_i) + C(x_i) + D(x_i) E(y_i) \quad (21)$$

$$\dot{y}_i = F(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) G(y_i) + H(y_i) + I(x_i) J(y_i) \quad (22)$$

Proposition 2 (Preservation of Order). Let the functions A, B, C, \dots be continuous with D and J nonnegative, and E and I monotone increasing. Then the functions x_{ij} and y_{ij} change sign at most once and never for $t \geq T$ if $x_{ij}(T) y_{ij}(T) \geq 0$.

Equation (17) is put in this form by writing

$\sum_{k \neq i} g(y_k) = L - g(y_i)$, where $L = \sum_{k=1}^n g(y_k)$. The property or order preservation can profitably be thought

of as a guiding principle for designing pattern processors for the purpose of STM storage. We do not require that the network duplicate the pattern of inputs, but only that, once the input is transformed by the network, the order of the transformed pattern should be preserved in STM to provide a record of relative (albeit not necessarily absolute) importance of the features or commands that are coded in the different populations.

The above results analyse situations wherein partial contrast is eliminated if the signal functions are unmatched. In the next few sections, we consider several physically plausible mechanisms whereby partial contrast can be restored in the presence of unmatched interactions.

9. Partial Contrast Due to Saturating Inhibition

In systems (10) and (11), the off-surround produces the tendency for contrast enhancement to occur. Indeed, if no inhibitory terms existed in (10), then the

limiting distribution would be uniform. If $D > C$, the contrast enhancing tendency is so great that partial contrast is eliminated and the system tries to make a choice. Does this property depend on the exact form of (11)?

In (11), the excitatory signal $E x_i$ acts additively rather than via a shunt, as it does in (10). Were a shunt to act instead, (11) would be replaced by

$$\dot{y}_i = -E y_i + (F - y_i) C x_i. \quad (23)$$

Note that for y_i small,

$$\dot{y}_i \cong -E y_i + F C x_i,$$

which is again an approximately additive interaction. A small excitatory signal $C x_i$ to (23) can produce this additive approximation. A large excitatory signal in (23) will tend to saturate the y_i response at the level F . Intuition suggests that the tendency to produce a 0–1 distribution when $D > C$ can be offset by the shunt in (23). This is because the enhancement process will be at least partially balanced by the uniformization that is produced as the saturating level F is approached. The next theorem shows that this can indeed happen if $n = 2$, but cannot happen if $n > 2$.

Theorem 5. Let $n = 2$. Define $H \equiv DF - E - BC + A$ and $z_{12} \equiv D y_{12} - C x_{12}$. Suppose $H = 0$. If $z_{12}(0) = 0$, then all $X_i(t)$ are constant. If $x_{12}(0) z_{12}(0) \geq 0$ with either $x_{12}(0)$ or $z_{12}(0)$ nonzero, then the excitatory pattern is partially contrast enhanced if $x_{12}(0) \geq 0$ and uniformized if $x_{12}(0) < 0$. If $x_{12}(0) z_{12}(0) < 0$, the excitatory pattern is uniformized.

Let $n > 2$. Then if limits $x_i(\infty)$ and $y_i(\infty)$ exist, the $x_i(\infty)$ can assume at most one nonzero value.

Remarks: The above theorem can be extended when $n = 2$ to analyse cases where $H \neq 0$. The $H = 0$ case illustrates the main phenomena, however. For example, if $FC = E$, then (23) reduces to (11) at small y_i values. In this situation $H = 0$ is the same as $EC^{-1}(D - C) = BC - A$. In systems (10) and (11), condition $BC - A > 0$ maintains the reverberation and $D - C > 0$ makes a choice. Here D can exceed C without preventing partial contrast from occurring.

Why does partial contrast fail if $n > 2$? Speaking heuristically, it is because the inhibitory Eq. (23) does not have its own off-surround. In other words, the uniformizing tendency in (23), is either too strong or too weak if it is not balanced by an off-surround. For example, if (23) is replaced by $\dot{y}_i = -A x_i + (B - y_i) \cdot C x_i - D y_i \sum_{k \neq i} y_k$, and (10) also holds, then it is readily proved that the differences $x_i - y_i$ converge exponentially to zero, so that the system lumps itself, no matter how the positive coefficients A , B , C , and D are chosen.

10. Partial Contrast in Unmatched Fields with Distant-Dependent Interactions

The above results show that in networks whose excitatory and inhibitory signal functions are unmatched, there exists a strong tendency to approach 0–1 or uniform limits in STM. These networks are special because the strength of inhibitory connections between all populations v_i and v_k , $i \neq k$, is the same. If $n = 2$, this causes no loss of generality; hence the importance of postural control of an agonist-antagonist muscle pair in this case, or of net incentive motivational regulation by a pair of drive centers, or indeed of any dipole of antagonistic populations (Wise, Berger, and Stein, 1973). For $n > 2$, however, there are many examples of inhibitory interactions whose coefficients depend on i and k , for example as a function of $|i - k|$ (Bishop *et al.*, 1971; Colonnier, 1965; Scheibel and Scheibel, 1970; Szentagothai, 1967).

The distribution of the on-center and off-surround interactions determines the generalization gradients that join together various sensory cues, motor commands, and so on. Below we will show that partial contrast can occur given unmatched signal functions if the on-center and off-surround interaction coefficients are suitable functions of distance. Before analysing these phenomena in the recurrent case, we study properties of shunting fields that are non-recurrent, or feedforward. Different properties of analogous recurrent fields can then be unambiguously ascribed to the reverberation.

11. Input Broadening and Curvature Detection

Consider the lumped nonrecurrent system

$$\dot{x}_i = -A x_i + (B - x_i) \sum_{k=1}^n I_k C_{ki} - x_i \sum_{k=1}^n I_k D_{ki} \quad (24)$$

with inputs I_k , excitatory coefficients C_{ki} , and inhibitory coefficients D_{ki} . The equilibrium values ($\dot{x}_i = 0$) of (24) are

$$x_i = \frac{B \sum_{k=1}^n I_k C_{ki}}{A + \sum_{k=1}^n I_k E_{ki}}, \quad (25)$$

where $E_{ki} = C_{ki} + D_{ki}$. Furman (1965) considers a special case of (25) such that

$$x_i = \frac{I_i}{1 + \sum_{k=1}^n I_k E_{ki}}. \quad (26)$$

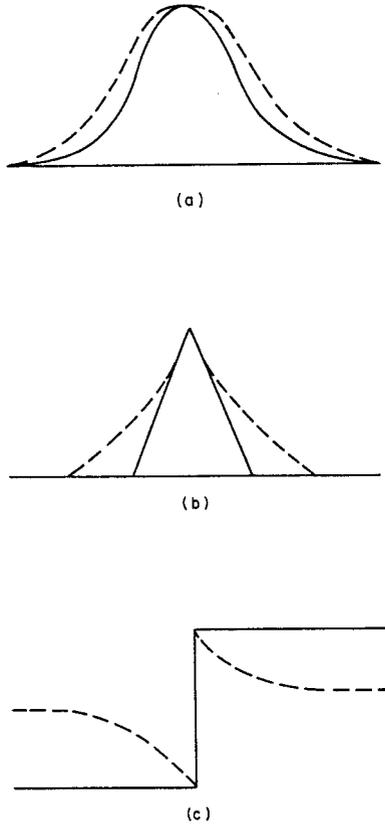


Fig. 7a-c. Normalized input and responses from Furman (1965)

He analyses the responses to three types of inputs; namely, a Gaussian input, a triangle, and a unit step function. His results are schematized in Fig. 7.

The solid line plots $M_i = I_i / \max_k I_k$, and the dotted line plots $N_i = x_i / \max_k x_k$. The network broadens the Gaussian input (Fig. 7a), broadens the triangle except at its vertex (Fig. 7b), and contrast enhances the step (Fig. 7c). Furman concludes (p. 268) that "shunting inhibition acts as a specialized discontinuity detector whereas subtractive inhibition acts more generally as a change-of-slope (i.e., curvature) detector". Below we will analyse why the shunt sometimes broadens patterns, and will also show contrary to Furman's results, that it can act as a curvature detector. In addition, various parametric studies given prescribed inputs or anatomies will be summarized. To simplify computations, we often replace the sums in (25) by integrals, namely

$$x(z) = \frac{B \int_{-\infty}^{\infty} I(w) C(w, z) dw}{A + \int_{-\infty}^{\infty} I(w) E(w, z) dw} \quad (27)$$

To discuss input broadening, we use the following definition, that is adapted from Furman.

Definition 8. Let $M(w) = I(w)/\bar{I}$ and $N(w) = x(w)/\bar{x}$, where \bar{I} is the input size at the center of a symmetric input and \bar{x} is the network response at the corresponding population. The input $I(w)$ is broadened by the network if $N \neq M$ and $N(w) \geq M(w)$ for all $w \geq 0$. In Furman's model, input I_i excites only population v_i . This motivates the following result, where $\delta(\xi)$ is the familiar delta function.

Proposition 3 (Input Broadening). Let $C(w, z) = C\delta(w-z)$ and $D(w, z) = D_1(w-z) = D_1(z-w)$, where $D_1(\xi)$ is monotone decreasing. The network broadens any differentiable input $I(w) = I(-w)$ that is monotone decreasing for $w \geq 0$.

In other words, the network broadens any differentiable unimodal input pattern if each population excites only itself. This result confirms and generalizes Furman's observations. It is not, however, true for all differentiable unimodal inputs if the condition $C(w, z) = C\delta(w-z)$ is relaxed, although it remains true for certain important inputs, such as Gaussianly distributed inputs.

Proposition 4. Let $C(w, z) = C_1(w-z) = C_1(z-w)$ and $D(w, z) = D_1(w-z) = D_1(z-w)$, where $E_1(\xi) \equiv C_1(\xi) + D_1(\xi)$ is monotone decreasing. Any Gaussian input $I(w) = I \exp(-Jw^2)$ is broadened.

On the other hand, inputs that decay quadratically with distance are not broadened, even though $e^{-w^2} \cong 1 - w^2$ in first approximation for small values of w .

Proposition 5. Let $C(w, z) = C_1(w-z) = C_1(z-w)$ where $C_1(\xi) = 0$ for $|\xi| > W$. Let $I(w)$ be any differentiable, decreasing, even function such that $I(w) = U - Vw^2$ for $|w| < W$, with U and V positive. Then if A is chosen sufficiently large, there exists a w_0 in $|w| < W$ such that $N(w_0) < M(w_0)$.

Thus unless $C(w, z) = C\delta(w-z)$, as in Furman's paper, it is possible for a shunting network not to broaden unimodal inputs. The conclusion that a hunt is insensitive to curvature information is also not generally true, even if $C(w, z) = C\delta(w-z)$.

Proposition 6 (Negative Contrast). Let $C(w, z) = C\delta(w-z)$ and choose $D(w, z) = D_1(w-z) = D_1(z-w)$ continuous, nonnegative, and such that $D_1(0) > 0$. In response to any differentiable input such that $I'(0) = 0$ and $I'(w) > 0$ if $w \neq 0$, $x'(0) < 0$ (Fig. 8).

Propositions 4-6 illustrate that care must be taken in choosing the anatomy to achieve desired input transformation properties. The next result shows that even a unimodal response to a unimodal input cannot be taken for granted.

Proposition 7 (Peak Splitting). Let $C(w, z) = C \exp[-\mu^{-1}(w-z)^2]$

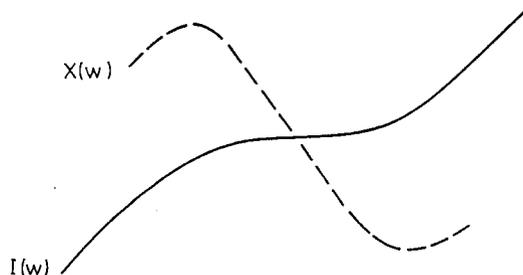


Fig. 8. Negative contrast: An example of curvature detection

and $D(w, z) = D \exp[-\nu^{-1}(w-z)^2]$. If $I(w) = I\delta(w)$, the function $x(w)$ is unimodal if

$$D(1 - \mu\nu^{-1}) + AI^{-1} \geq 0, \quad (28)$$

and is bimodal otherwise.

Condition (28) holds for all $I > 0$ only if $\nu > \mu$; that is, only if the excitatory coefficients fall off with distance faster than the inhibitory coefficients. This condition will therefore always be imposed below unless otherwise stated.

12. Variable Contrast at Variable Background Activity Levels

Consider the step input depicted in Fig. 9a. Superimposed on a baseline activity level H_1 is a step of height $H_2 - H_1$. The response shapes in Fig. 9b and c are typical if $C(w, z)$ and $D(w, z)$ are Gaussianly distributed. Two types of parametric studies are of particular interest:

A. *Difference Scale*. Let $H_2 - H_1 = H = \text{constant}$ and increase H_1 from 0.

B. *Ratio Scale*. Let $H_2/H_1 = R = \text{constant}$ and increase H_1 from 0.

Two measures of the response are convenient.

Definition 9. The *upper contrast* C_U is the ratio $[\max_v x(v)/x(\infty) - 1]$.

Definition 10. The *lower contrast* C_L is the ratio $[1 - \min_v x(v)/x(-\infty)]$.

Parametric studies of the response to a step input were carried out under the following assumptions: $C(w, z)$ and $D(w, z)$ are Gaussianly distributed with

$$\int_{-\infty}^{\infty} C(w, z) dz = \int_{-\infty}^{\infty} D(w, z) dz = 1. \quad (29)$$

We write

$$C(w, z) = [\pi\sigma_c^2]^{-\frac{1}{2}} \exp[-\sigma_c^2|w-z|^2] \quad (30)$$

and

$$D(w, z) = [\pi\sigma_d^2]^{-\frac{1}{2}} \exp[-\sigma_d^2|w-z|^2]. \quad (31)$$

Whenever $\sigma_c < \sigma_d$, (29) implies that $C(w, z)$ and $D(w, z)$ have the graphs shown in Fig. 10.

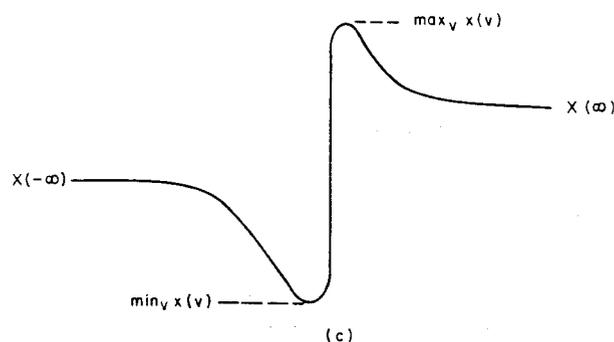
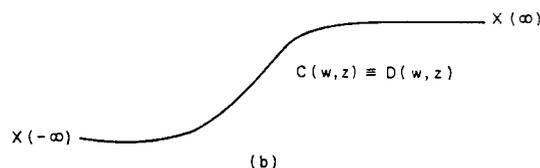
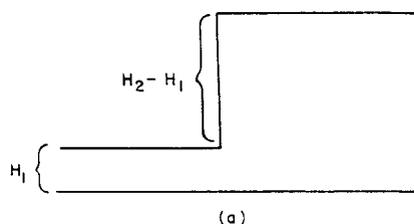


Fig. 9. Step input and steady state responses in a non-recurrent net

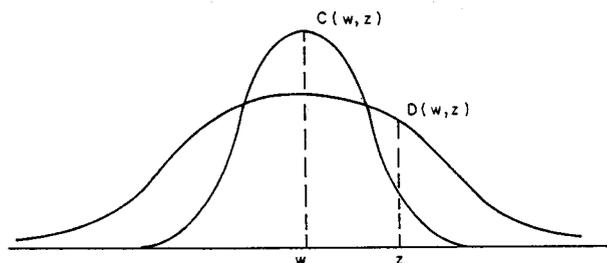


Fig. 10. Excitatory and inhibitory connection strength dependence on distance

That is, $C(w, w) > D(w, w)$ and $D(w, z)$ decreases more slowly than $C(w, z)$ as $|w-z|$ increases. In all of our computer runs, the values $\sigma_c = 0.25$ and $\sigma_d = 4$ were imposed.

The studies with $H_2/H_1 = \text{constant}$ have the following interpretation, among others. Suppose that the network represents an idealized retina being exposed to a picture of two uniformly shaded contiguous regions. How does the perceived contrast

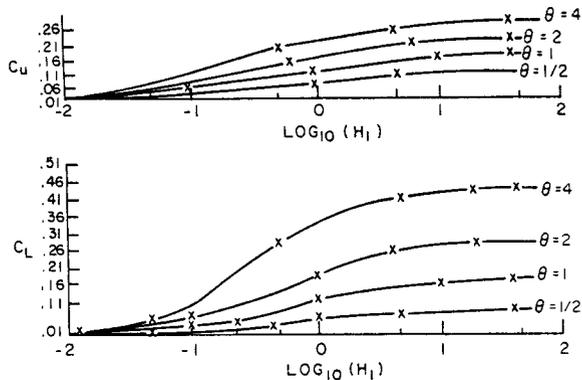


Fig. 11. Upper and lower contrast as total intensity of input is increased

between the two shaded regions change as the intensity of white light illuminating the *entire* picture is varied? This interpretation holds because a fixed fraction of light is reflected by points in the two regions, independent of the total intensity (Cornsweet, 1970). The studies with $H_2 - H_1 = \text{constant}$ have a different interpretation; namely, how easily perceived is a step of fixed size on a background of variable intensity?

The two types of study yield opposite results. The function controlling the network response is readily derived from (27) and (29). It is

$$x(z) = \frac{B(H_2 - H_1)U(z) + H_1}{A + (H_2 - H_1)V(z) + H_1}, \quad (32)$$

where

$$U(z) = \int_{-\infty}^0 C(z, w) dw$$

and

$$V(z) = \int_{-\infty}^0 E(z, w) dw.$$

Given $H_2 H_1^{-1} = 1 + \Theta = \text{constant}$, we chose four input contrast levels $\Theta = \frac{1}{2}, 1, 2, \text{ and } 4$, and varied H_1 from 0.05 to 2000. For fixed Θ , the upper contrast C_U and lower contrast C_L are monotone increasing functions of H_1 which are 0 when $H_1 = 0$ and which saturate as $H_1 \rightarrow \infty$. For fixed H_1 , C_U and C_L are increasing functions of Θ (Fig. 11).

In other words, increasing the background illumination can only increase the relative contrast in response to a step input. Saturation effects at large background values H_1 are overcome by the off-surround. Indeed, letting $H_1 \rightarrow \infty$ in (28), we find $x(z) = [\Theta B U(z) + 1][\Theta V(z) + 1]^{-1}$, which is not a uniform distribution. Saturation is not entirely operative, however, because for any fixed Θ , $C_U(H_1) < C_L(H_1)$ at all positive H_1 values.

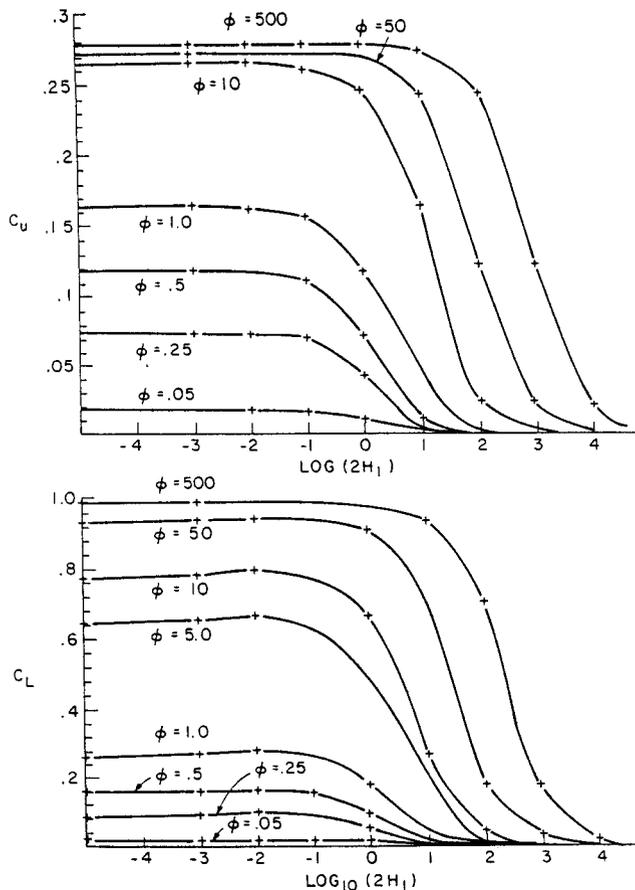


Fig. 12. Upper and lower contrast as background intensity is increased

Given $H_2 - H_1 = \phi = \text{constant}$, we chose ϕ between 0.005 and 5000 and varied H_1 from 0.05 to 2000. In this situation, for fixed ϕ , C_U and C_L are both decreasing functions of H_1 , except for a small increase in C_L before it decreases rapidly. Hence saturation is an important effect when a fixed step is placed on backgrounds of variable intensity. Given fixed ϕ , $C_U(H_1) < C_L(H_1)$, for all $H_1 > 0$, which is a secondary effect of saturation. Also the critical value of H_1 above which C_U and C_L begin to quickly decrease is a monotone increasing function of ϕ , which shows that a step of greater height is almost equally discriminable over a broader range than a step of smaller height (Fig. 12).

13. Variable Contrast and Nonrecurrent Peak Splitting

Proposition 7 shows that, in a nonrecurrent network, peak splitting can occur in the response to a spike input if the input intensity is chosen sufficiently

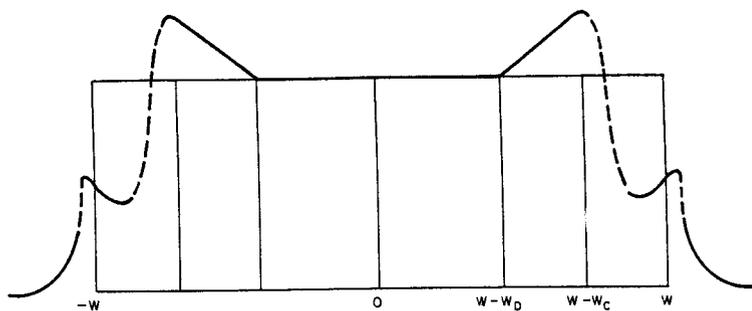


Fig. 13. Possible response to rectangular input in a nonrecurrent network

large. Here we show that a similar phenomenon can occur in the response to the boundary of a rectangle. This effect is important as a prelude to studies of recurrent networks, because peak splitting in a nonrecurrent network is clearly not due to recurrent disinhibition, but it shares some formal properties of a disinhibitory mechanism.

Proposition 8. Let $C(w, z) = C_1(w - z) = C_1(z - w)$ and $D(w, z) = D_1(w - z) = D_1(z - w)$, where $C_1(\xi) > 0$ only if $|\xi| < W_c$, $D_1(\xi) > 0$ only if $|\xi| < W_d$, and $W_c < W_d$. Let $I(w)$ be a rectangular input of intensity I , width $2W$ centered at $w=0$, and choose $W > W_d$. Then $x(w) = x(0)$ for $W_d - W \leq w \leq W - W_d$, and $x(w) > x(0)$ for both $W_c - W < w < W_d - W$ and $W - W_d < w < W - W_c$. Moreover, $x(W) < x(0)$ and

$$\text{sign} \left[\frac{dx}{dw}(W) \right] = \text{sign} \left[\frac{\int_0^{2W} C(z) dz}{C(0)} - \frac{\frac{A}{I} + \int_0^{2W} D(z) dz}{D(0)} \right]. \quad (33)$$

(Fig. 13). If $\frac{dx}{dw}(W) > 0$, then there are at least two bumps in the network response at each boundary of the rectangle. Thus making AI^{-1} small makes double boundary bumps more likely. Condition (33) strikingly resembles condition (28) for peak splitting in response to an input spike. Indeed, suppose that $C_1(w)$ and $D_1(w)$ vary in an approximately Gaussian fashion with distance, say $C_1(w) \cong C \exp(-\mu^{-1}w^2)$ and $D_1(w) \cong D \exp(-\nu^{-1}w^2)$. Then if $\nu > \mu$, (33) readily shows that $\frac{dx}{dw}(W) < 0$; the extra bump at the rectangle boundary is hereby eliminated, just as in condition (28).

14. Recurrent Peak Splitting, Disinhibition, and Graded STM

The above results show that matched signal functions are an important means for achieving

graded patterns in STM whenever $n > 2$. The $n > 2$ case is, of course, typical when the network populations model feature detectors in a sensory cortical field; here billions or greater numbers of units are present. Below we show that distance-dependent interactions are capable of graded STM even in the unmatched case when $n > 2$. The general setting for our studies is characterized by unumped recurrent networks of the form

$$\dot{x}_i = -Ax_i + (B - x_i) \left[\sum_{k=1}^n f(x_k) C_{ki} + I_i \right] - x_i \sum_{k=1}^n g(y_k) D_{ki} \quad (34)$$

and

$$\dot{y}_i = -Ey_i + \sum_{k=1}^n x_k F_{ki}. \quad (35)$$

In (35), we have generalized the additive interaction used in (11) to include excitatory signals from populations other than v_i . We have not included an excitatory shunting term, as in (23), for two main reasons: firstly, Theorem 5 illustrates its main effect, which is to partially uniformize a contrast-enhanced pattern; secondly, we wish to prove that such a term is not *necessary* to achieve graded STM given distance-dependent interactions. Since a shunt becomes approximately additive at low input intensities, we infer that graded STM can be achieved both at low and at high input intensities when the full shunting interaction replaces (35). On the other hand, a sum of signals such as $\sum_{k=1}^n x_k F_{ki}$ is sometimes needed to produce a physically plausible transformation of an input pattern into a pattern of activity in STM. We have not included an inhibitory term such as $-y_i \sum_{k=1}^n y_k G_{ki}$ in (35) for two main reasons: firstly, such a term often tends merely to lump the network, as we noted at the end of Section 9; secondly, this term describes a disinhibition effect (i.e., inhibition of the inhibitory interneurons), and we wish to prove

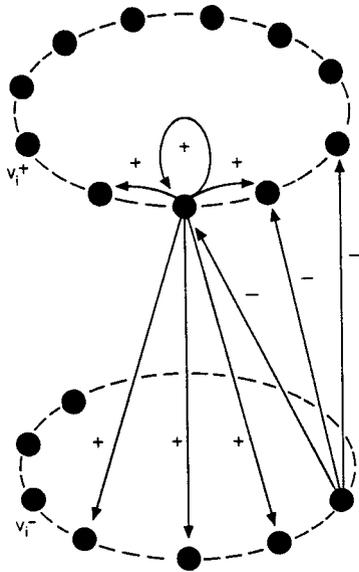


Fig. 14. Recurrent ring anatomy: eleven excitatory cells and eleven inhibitory cells

that disinhibitory effects can be produced, and can be eliminated, in the absence of such a term. Moreover, the fact that our model produces graded STM without such explicitly disinhibitory terms contrasts with the reported phenomena in the model of Wilson and Cowan (1972, 1973).

The signal functions $f(w)$ and $g(w)$ in (34) are chosen below to have the form

$$f(w) = g(w) = \max(w - H, 0), \quad (36)$$

where $H > 0$. This signal function approximates a faster-than-linear-approaching-linear signal function using a threshold cut-off at $w = H$. The coefficients C_{ki} , D_{ki} , and F_{ki} are chosen to decrease at a Gaussian rate with distance; i.e.,

$$C_{ki} = \hat{C} \exp(-\sigma_c^{-2} |k - i|^2), \quad (37)$$

$$D_{ki} = \hat{D} \exp(-\sigma_d^{-2} |k - i|^2), \quad (38)$$

and

$$F_{ki} = \hat{F} \exp(-\sigma_f^{-2} |k - i|^2). \quad (39)$$

We also permit $\sigma_c = 0$ or $\sigma_f = 0$. When $\sigma_c = 0$, $C_{ij} = 0$ unless $i = j$. This is the limiting case of a very narrow on-center. When $\sigma_f = 0$, $F_{ij} = 0$ unless $i = j$. In this case, an inhibitory population v_i^- can be excited only by a recurrent excitatory signal from populations within the on-center of v_i^+ , or by an excitatory input to v_i^+ itself. The anatomy is chosen to form a ring (Fig. 14). This anatomy eliminates boundary effects that are due merely to the finite width of a uniformly distributed background input.

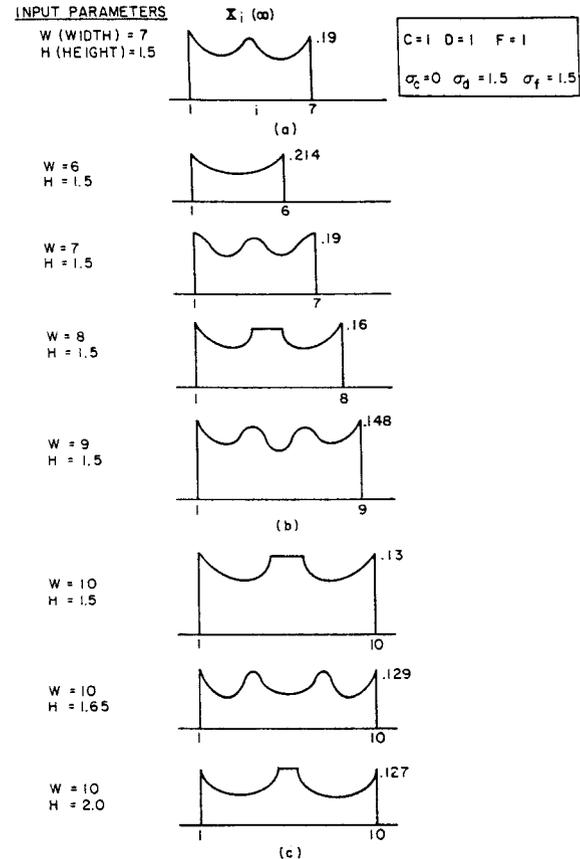


Fig. 15a-c. Effect of stimulus width and amplitude on steady state activity pattern

The results describe network responses to one or more input spikes ($I_i > 0$ at only one v_i) of variable separation and height, to rectangles of variable width and height, and to triangles. These suggest important constraints on network design that are needed to process more complex patterns. We study the asymptotic pattern of activity when the input pattern is left on indefinitely, and when the uniform background activity level is set at various values. We then shut off the input pattern to study what the asymptotic pattern of activity is in STM. All the computer studies reported in Sections 14 and 15 set $A = B = E = 1$ and hold the parameters $C = \sum_{k=1}^n C_{ki}$, $D = \sum_{k=1}^n D_{ki}$, and $F = \sum_{k=1}^n F_{ki}$ constant whenever σ_c , σ_d , and σ_f are varied. First we describe important aspects of the results when the input patterns are left on.

Suppose $\sigma_c = 0$ and $\sigma_d = \sigma_f > 0$, but σ_f is relatively small; e.g., $F_{11} F_{14}^{-1} \cong 0.02$. In response to a sustained rectangular input, one often finds a steady state pattern

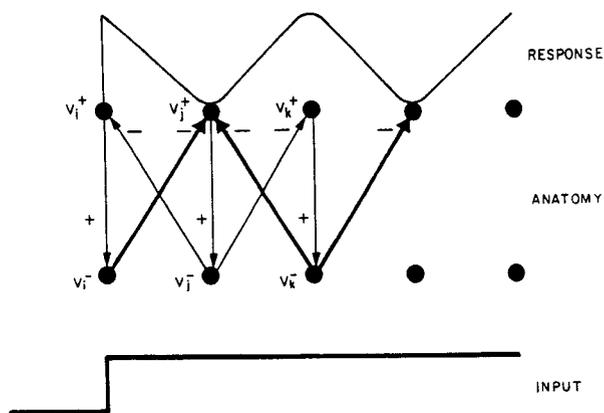


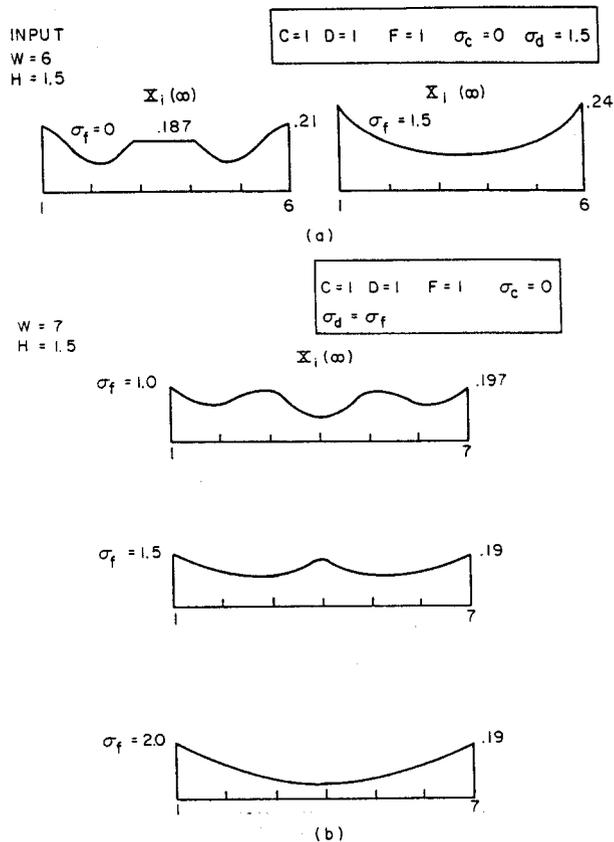
Fig. 16. Recurrent disinhibition in network response to step input

that exhibits spurious interior lumps (Fig. 15a). The number of bumps depends on the width W (Fig. 15b) and height H (Fig. 15c) of the rectangle. These bumps are due to disinhibition. They can be explained as follows. Populations within the width of the rectangular input receive equal inputs. Populations near the rectangle boundary v_i^+ are less inhibited initially by signals from other populations than populations v_j^+ near the center. Thus the populations v_i^+ inhibit their neighbors v_j^+ more than conversely. When v_j^+ is inhibited, it cannot excite its inhibitory population v_j^- . Thus adjacent populations v_k^+ towards the center are inhibited less. They strongly excite their respective inhibitory populations v_k^- , and so on (Fig. 16).

The extra bumps occur because v_j^- is inactivated whenever v_j^+ is inhibited. One way to correct this is to let nearby populations v_i^+ activate a given v_j^- more strongly by increasing σ_f . Then, in response to a rectangle, the populations v_i^+ near the boundary can excite populations v_j^- closer to the center. Even if v_j^+ is strongly inhibited by v_i^- , populations v_k^+ that are still closer to the center can receive strong inhibition from v_j^- . The effect of increasing σ_f , and σ_d correspondingly, is illustrated in Fig. 17. The spurious bumps are indeed removed.

Two questions arise from this example: (1) Does the effect depend crucially on setting $\sigma_c = 0$? (2) Does the effect occur whenever the ratio $\sigma_c \sigma_f^{-1}$ is reduced, other things equal? The next results answer question (1) negatively, and show that "yes" is often the answer to question (2).

In Fig. 18a, the input is a spike, $\sigma_c > 0$, and the response to a sustained input has a center peak and two very large and wide spurious peaks. As σ_f is increased, the spurious peaks vanish and the width of the network pattern contracts. In Fig. 18b, the

Fig. 17. Increasing σ_f can eliminate spurious bumps

input is again a spike and σ_c is increased from 0 while σ_f is kept constant. When $\sigma_c = 0$, the network response is the obvious spike, which broadens as σ_c increases and then splits into a trimodal response. Finally, no asymptotic limits exist. Instead, travelling waves propagate outward from the spike in opposite directions. Thus the disinhibition which produces a peak split in response to the spike can also destabilize the network. Compare the remarks on peak splitting and the μv^{-1} ratio in Proposition 7.

How are rectangles stored in STM? Fig. 19a shows that, even if the network response to a sustained rectangle contains spurious bumps, nonetheless, when the rectangle is shut off, only the peak activities near its boundary might be stored in STM. Interior peaks can, however, be stored in STM if they are sufficiently intense. In Fig. 19b, a centered spike input is superimposed on a rectangle. The response in STM records where the rectangle boundary was, as well as where the spike was. STM does not necessarily delete the interior of the rectangle when it is not supplemented by a spike. That happens in Fig. 19a because $\sigma_c = 0$; there is no tendency to smooth the excitatory response

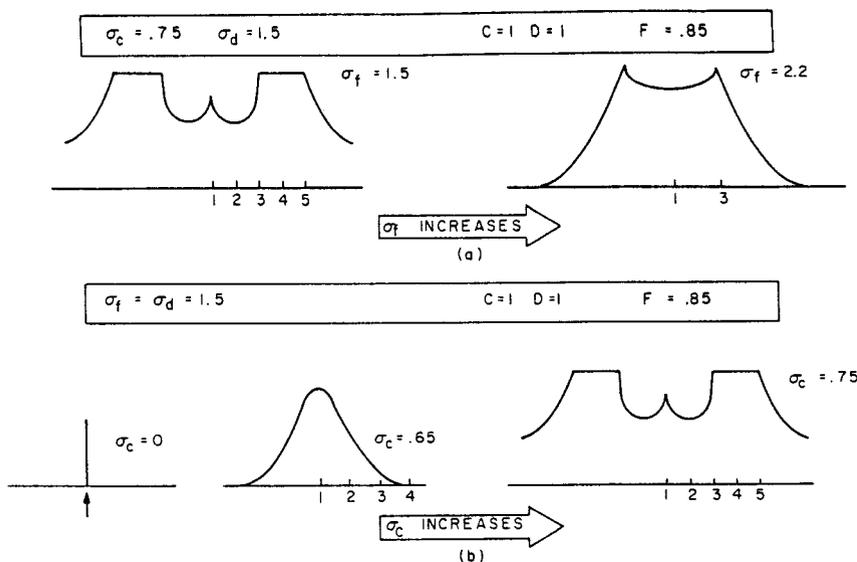


Fig. 18a and b. Increasing σ_c can produce spurious bumps

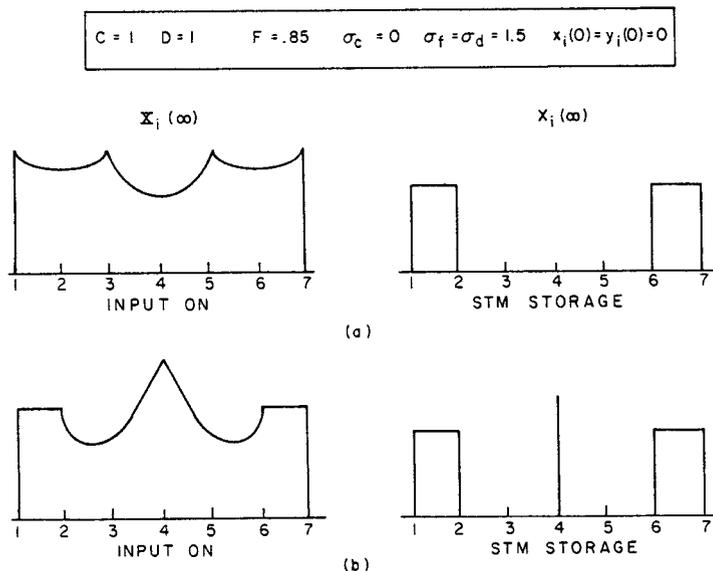


Fig. 19a and b. Storage of interior bumps depends on their height

across the rectangle using recurrent excitatory interactions. Figure 20a depicts a case in which $\sigma_c > 0$, such that the edges of the rectangle are still enhanced in STM, but the interior of the rectangle is also stored. Figure 20b shows, however, that the interior of the rectangle can be deleted from STM if the intensity of the rectangle is increased. In other words, increasing the relative contrast of input to background has decreased the relative net inhibition received by the boundary populations as compared with the center

populations. Contrast the nonrecurrent response in Fig. 13. A similar enhancement can occur when a spike is increased in intensity, as in Fig. 21. As the spike height increases, the unimodal STM response bifurcates to form a bimodal STM response.

In other words, by increasing the relative contrast of input to background, the trend towards disinhibition is strengthened, sometimes showing itself as a suppression of nonmaximal responses, and sometimes as a peak split. This interpretation is supported by

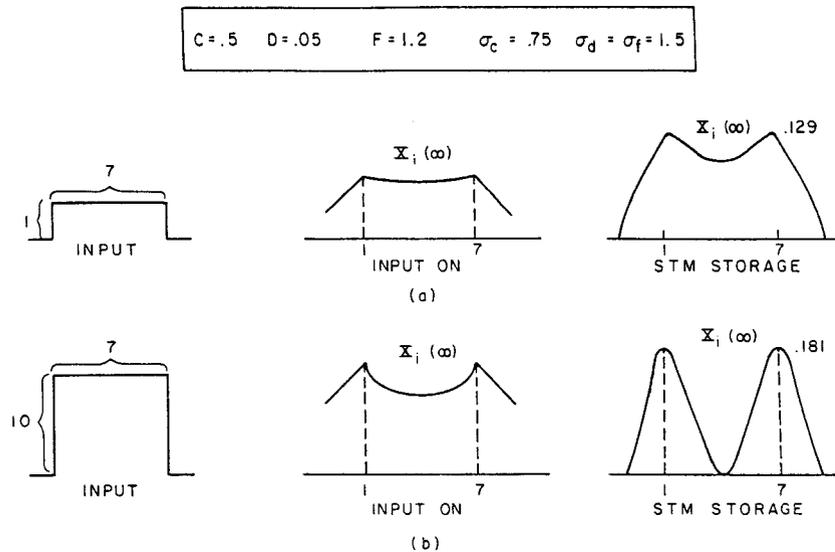


Fig. 20a and b. Effect of input intensity on STM of rectangles

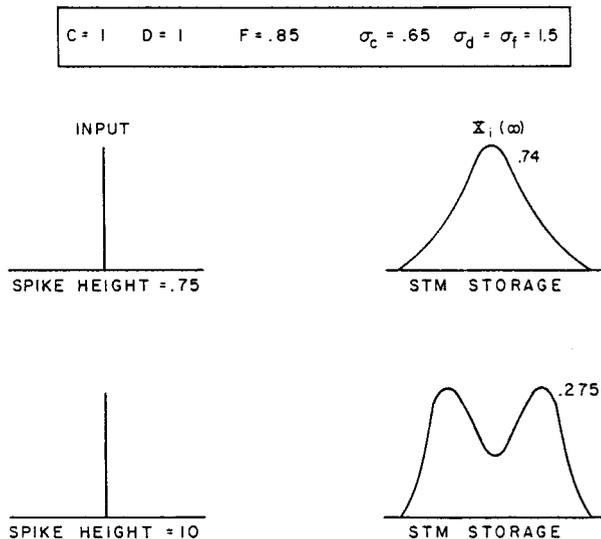


Fig. 21. Increasing spike intensity can induce peak splitting

Fig. 22, in which a peak splitting tendency in response to a rectangle of increasing height is illustrated. Propositions 7 and 8 showed that peak splitting can occur even in a nonrecurrent network. Here we see a peak split due to a recurrent interaction. As in the nonrecurrent case, the tendency towards a peak split can be eliminated by increasing the relative rate with which excitatory interaction strength diminishes with distance; cf. decreasing μv^{-1} in (28) and decreasing $\sigma_c \sigma_f^{-1}$ herein. The recurrent case can also produce suppression of certain activities in STM, although this does not occur in the nonrecurrent case.

Are the extra disinhibitory bumps that are produced by increasing rectangle intensity always stored in STM? Figure 22 shows that they are not. This is analogous to the effect in Fig. 18a.

Increasing the uniform background activity level can uniformize the response to a rectangle that is superimposed on it; compare the results in Section 12 on the monotone decrease of C_U and C_L as H_1 increases and $H_2 - H_1$ is kept fixed (Fig. 23a). On the other hand, increasing the tonic input can also produce spurious bumps when the rectangle is on (Fig. 23b) by activating populations outside the rectangle and thereby creating new sources of disinhibition. These spurious bumps can also be erased when the pattern is stored in STM (Fig. 23c).

Given suitable parameters, the disinhibitory bumps can be so large that they are stored in STM, whereas the original pattern is not. This is illustrated in the response to a spike (Fig. 24a), to a spike superimposed on a uniform background (Fig. 24b), and to a triangle (Fig. 24c). These disinhibitory bumps do not help to store the input pattern in STM, but they are interesting in themselves as a kind of complementary *afterimage*, in which the highest peaks of the input pattern are suppressed, and the lowest troughs are enhanced.

15. Conjoint Discrimination by Synergism in STM

Suppose that an animal is confronted by a problem in which two cues *A* and *B* together predict a food reward for pressing a lever, but the cues separately

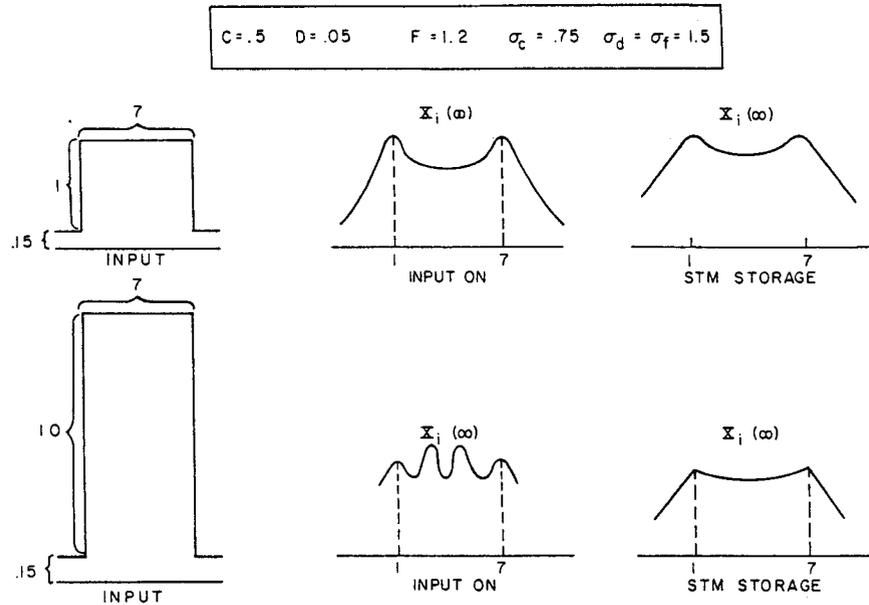


Fig. 22. Influences of increasing rectangular input intensity

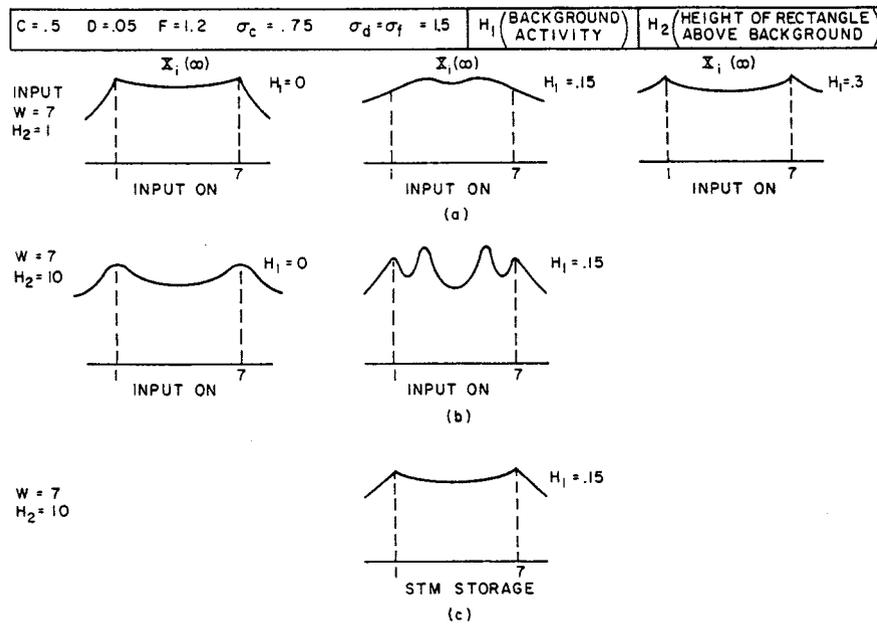


Fig. 23a-c. Effects of tonic background level

predict a shock for pressing the lever. How does the animal make the discrimination? Herein we do not presume to attack this problem frontally, but merely note an interesting formal analog. Suppose that a spike input to population v_A (v_B) corresponds to the occurrence of cue A (B). Suppose that storage of

activity in STM is necessary for the network to learn a response to the cue. See Grossberg (1972a, b, 1974, 1975) for a relevant theory of reinforcement and discrimination learning. Let the populations excited by A or B alone be conditioned to negatively reinforcing sites. Can the joint occurrence of A and B

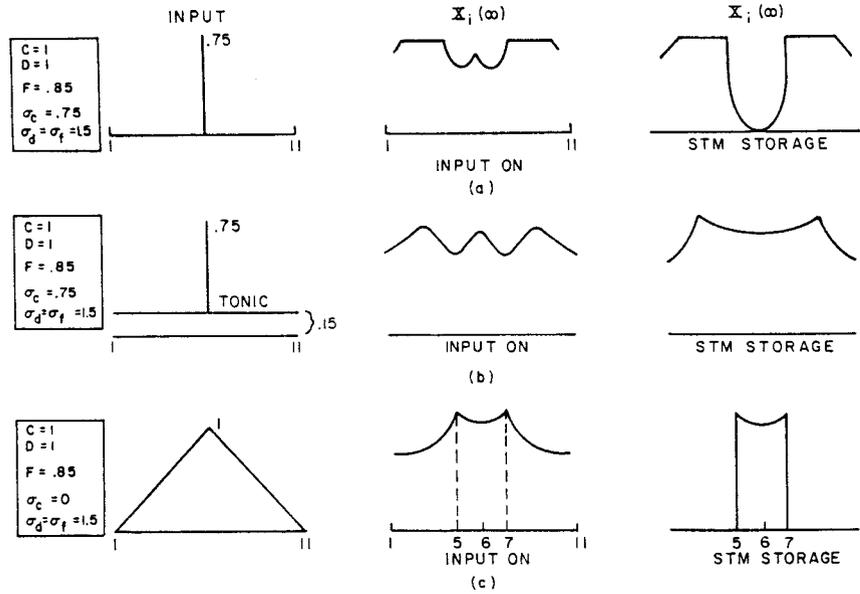


Fig. 24a-c. Disinhibitory bumps can be stored in STM

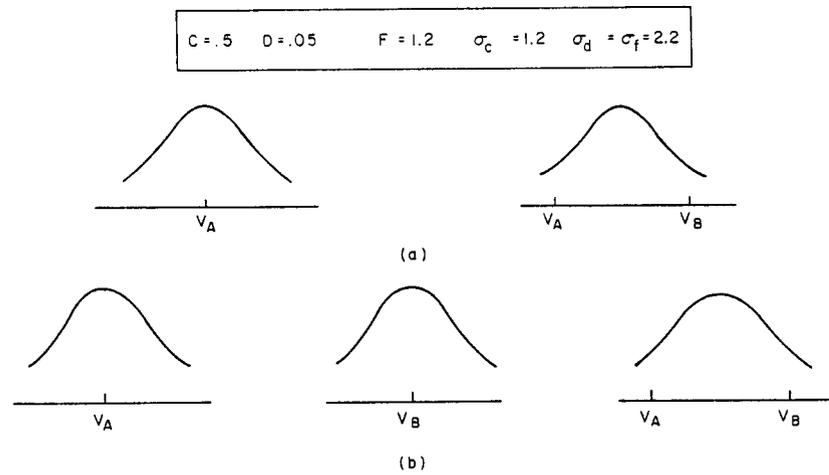


Fig. 25a and b. Synergisms in STM due to subthreshold and suprathreshold spikes

together bring new populations into STM which can be conditioned to positively reinforcing sites? The answer is "yes" (Fig. 25a). In Fig. 25a, an input spike to v_A generates in STM the depicted generalization gradient. The weaker input spike to v_B generates no response in STM. The two spikes simultaneously produce a gradient whose peak is between v_A and v_B . In Fig. 25b, an input spike to v_A or v_B separately generates a STM generalization gradient centered at the respective population. The gradient produced by both spikes simultaneously has a maximum between v_A and v_B , and the location

of the peak response of the synergistic gradient is not at either v_A or v_B . Grossberg (1975) develops a theory of attention in which the incentive motivational feedback generated by a coupling between the recurrent network described herein and formal reinforcement mechanisms will tend to enhance the peak response of the synergistic gradient and to suppress, or overshadow, submaximal responses. Thus populations which are activated by v_A and v_B together, but not separately, can control network performance to make the necessary discrimination between approach and avoidance.

16. The Existence of Stable and Unstable Limit Cycles in Unlumped Networks

In all of the above results, the unlumped network lumps itself as $t \rightarrow \infty$; i.e., limits of network variables exist globally as $t \rightarrow \infty$. This is not necessarily true; periodic solutions can exist, as well as solutions that approximate periodic solutions as $t \rightarrow +\infty$ or $t \rightarrow -\infty$ (limit cycles). We study these solutions to see whether they possess useful STM capabilities that differ from the asymptotically lumped ones. First we examine system

$$\dot{x}_i = -Ax_i + (B - x_i) \left[\sum_{k=1}^n f(x_k) C_{ki} + I_i \right] \quad (40)$$

$$-x_i \sum_{k=1}^n f(y_k) D_{ki}$$

and

$$\dot{y}_i = -Ey_i + \sum_{k=1}^n x_k F_{ki} \quad (41)$$

given uniform initial data $x_i(0) = x > 0$, $y_i(0) = y > 0$, and uniform inputs $I_i = I$, subject to the natural constraints $C = \sum_{k=1}^n C_{ki}$, $D = \sum_{k=1}^n D_{ki}$, and $E = \sum_{k=1}^n F_{ki}$ for all $i = 1, 2, \dots, n$. Then the solution is uniform for all $t \geq 0$, and (40) and (41) reduces to the two-dimensional system

$$\dot{x} = -Ax + (B - x)[Cf(x) + I] - Dx f(y) \quad (42)$$

and

$$\dot{y} = E(x - y). \quad (43)$$

Equation (43) can be generalized to

$$\dot{y} = -Ey + Fx;$$

i.e., to $F = \sum_{k=1}^n F_{ki}$ for all i , without altering the discussion below in any essential way. Thus we study first the effect of a uniform background intensity I on the production of periodic phenomena in a uniform field. The mathematical results below prove the existence of stable and unstable limit cycles using two different methods. Then a series of computer studies in uniform and nonuniform fields will be summarized and interpreted.

A. Unique Unstable Critical Point

We will first prove that, given suitable parameters, there exists a unique critical point of systems (42) and (43), and that this critical point is unstable. Because the system is 2-dimensional and bounded, this implies that a limit cycle exists which is stable from within (Hahn, 1967, p. 66). Such a limit cycle is illustrated

in Fig. 27c. Our first result is stated in terms of the functions

$$P(I, w) = \frac{BI - (A + I)w}{(C + D)w - BC}, \quad (44)$$

$$Q(I, w) = \frac{1}{C(B - w)} \left[E + \frac{B(DI - AC)}{(C + D)w - BC} \right], \quad (45)$$

and

$$R(I, w) = \frac{A + I + (C + D)f(w)}{BC - (C + D)w}. \quad (46)$$

Theorem 6. Let $f(w)$ be nondecreasing and differentiable for $0 \leq w \leq B$. Suppose that there exists an $I > 0$ and an x , $0 \leq x \leq B$, such that

A. (Existence).

$$f(x) = P(I, x), \quad (47)$$

B. (Instability).

$$f'(x) > Q(I, x), \quad (48)$$

C. (Uniqueness). Either

$$I > ACD^{-1} \quad (49)$$

or

$$f'(x) < R(I, x) \text{ for } 0 < x < BC(C + D)^{-1}. \quad (50)$$

Then there is a unique critical point P^* of system (42) and (43) for $0 \leq x \leq B$ and $0 \leq y \leq B$, which is either an unstable node or spiral point. Hence a limit cycle exists around P^* that is stable from within.

Theorem 6 can be used to construct signal functions that produce limit cycles given prescribed background activity levels. This procedure is illustrated with the help of Fig. 26.

In Fig. 26a, values of I such that $0 < I < ACD^{-1}$ are considered. For fixed I in this range, $P(I, x)$ is graphed. It is positive only for $BI(A + I)^{-1} < x < BC(C + D)^{-1}$. Fix an x in this range, and begin the construction of f by requiring that (47) hold. Given this choice of I and x , we wish to define $f'(x)$ such that (48) and (50) hold simultaneously. This can be done if and only if

$$R(I, x) > Q(I, x), \quad (51)$$

where $f(x)$ in $R(I, x)$ satisfies (47). A routine computation shows that (48) and (50) can hold if and only if

$$BDx(AC - DI) > E[BC - (C + D)x]^2. \quad (52)$$

For any x that satisfies (52), an $f(x)$ defined by (47) and an $f'(x)$ defined by (48) and (50) will produce a limit cycle. For fixed I , there exist x values that satisfy (52) because when $x = BC(C + D)^{-1}$ the inequality holds. If inequality (49) holds, then Fig. 26c and d show the range in which $f(w)$ and $f'(x)$ can be chosen to satisfy (47) and (48).

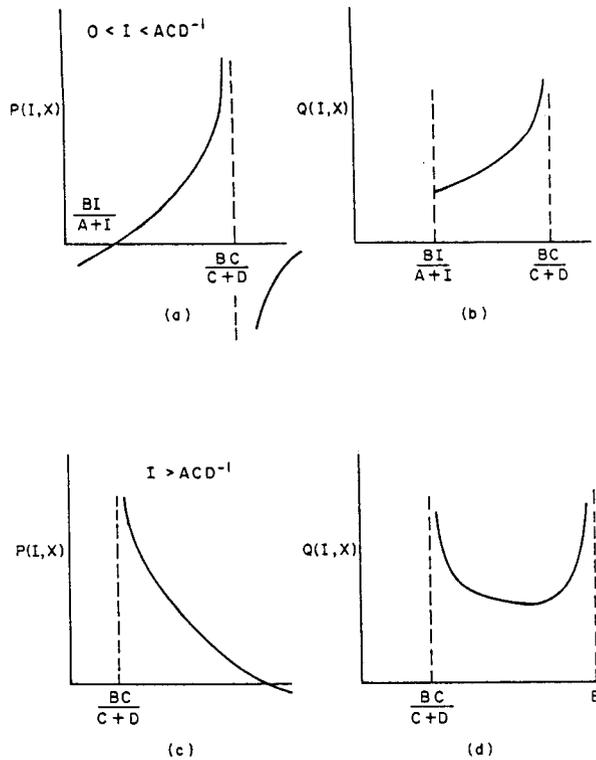


Fig. 26a-d. Some functions that determine stable limit cycle in uniformly distributed unlumped network

B. Hopf Bifurcation Theorem

In 1942, Eberhard Hopf wrote a remarkable paper concerning the existence and stability of periodic solutions that depend on a system parameter. In our case, this parameter will be the size of the uniform background intensity level. Below we briefly describe the general setting for the Hopf theorem in two dimensions, and then apply it to our special case, where it proves the existence of an unstable limit cycle, as in Fig. 28e.

Let

$$\dot{v} = F(v, \mu) \quad (53)$$

be a two-dimensional system where F is analytic in the 2-vector v and the parameter μ , v varies in a domain \mathcal{D} and $|\mu| < C$. Thus $F(v, \mu)$ can be expanded as

$$F(v, \mu) = L_{\mu}(v) + Q_{\mu}(v, v) + C_{\mu}(v, v, v) + \dots, \quad (54)$$

where the successive terms are multilinear forms in v of successively higher degree. Suppose that for $\mu=0$, there exists a stationary solution $v^*(0)$ of system (53). If moreover, none of the eigenvalues $\lambda(\mu)$ of the linearized system

$$L_{\mu}(a) = \lambda a \quad (55)$$

equals 0 when $\mu=0$, then there exists a unique stationary solution $v^*(\mu)$ in a suitable neighborhood of $v^*(0)$ for all sufficiently small μ , and $v^*(\mu)$ is analytic in μ at $\mu=0$. Hopf also assumes that for $\mu=0$, $\lambda(0)$ is pure imaginary (i.e., the system has a center), but that the continuous extension $\lambda(\mu)$ of this eigenvalue satisfies

$$\operatorname{Re} \left[\frac{d\lambda}{d\mu}(0) \right] \neq 0. \quad (56)$$

Under this condition, he proves the existence of a family of real nonconstant periodic solutions $v = v(t, \varepsilon)$, for all sufficiently small ε by writing $v = \varepsilon w$ in (54) to find

$$\frac{dw}{dt} = L_{\mu}(w) + \varepsilon Q_{\mu}(w, w) + \varepsilon^2 C_{\mu}(w, w, w) + \dots$$

and also expanding $\mu = \mu(\varepsilon)$ and $t = t(\varepsilon)$ in power series with variable ε . Hopf also presents a method for testing the stability of these periodic solutions.

Theorem 7. Given system (42) and (43) with initial data $0 \leq x(0) \leq B$ and $0 \leq y(0)$, suppose that there exists an x and an intensity I such that

A. (Critical Point).

$$f(x) = P(I, x), \quad (47)$$

B. (Zero Trace).

$$f'(x) = Q(I, x), \quad (57)$$

C. (Positive Determinant).

$$x f(x) > ED^{-1} \quad (58)$$

and (56) hold. Then the conclusions of the Hopf theorem are true.

Corollary 3. Consider the system

$$\begin{aligned} \dot{x} = & -x + (1-x)[20 \max(x-0.4, 0) + I] \\ & - 33.3x \max(y-0.4, 0) \end{aligned} \quad (59)$$

and

$$\dot{y} = x - y. \quad (60)$$

For $I^* \cong 2.43$ and $x^* = y^* \cong 0.5032$, all the conditions of Theorem 7 hold. Therefore there exists a family of periodic solutions for I near I^* . Moreover these solutions are unstable.

17. Oscillation Thresholds, Hysteresis, and Frequency Dependence on Background Intensity

Parametric studies of network response to a uniform background input intensity have been carried out, and illustrate both Theorems 6 and 7. A typical result is described below for the system (59) and (60). A similar series of network responses occurred as I

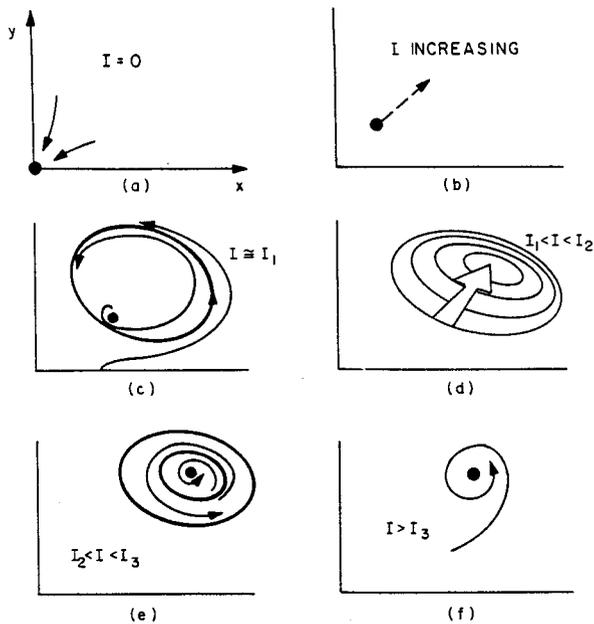


Fig. 27a-e. An observed sequence of phase portraits as tonic input I is increased

was parametrically increased given various other choices of network parameters. This series is summarized in Fig. 27.

When $I = 0$, the system is globally asymptotically stable; i.e., $x(\infty) = y(\infty) = 0$ for all initial data (Fig. 27a). As I increases, the unique stable critical point migrate from the origin into the first quadrant (Fig. 27b). At a critical value of $I = I_1$, a stable limit cycle emerges, surrounding an unstable critical point (Fig. 27c). As I increases further, the stable limit cycle migrated upward and contracts, always enclosing an unstable critical point (Fig. 27d). At a second critical value of $I = I_2$, an unstable limit cycle emerges within the stable limit cycle, and the unstable critical point becomes stable (Fig. 27e). As I increases further, finally at a third critical value of $I = I_3$, the stable and unstable limit cycles "collide" and vanish, leaving only a unique stable limit point. Thus at both low and high input intensities, no periodic phenomena exist; at intermediate intensities, either the motion always approaches a periodic solution, or it can be asymptotically stable or periodic, depending on the initial data. Thus the two values I_1 and I_3 are lower and upper thresholds for generating a periodic response. Between these critical values, the frequency of oscillation of the stable limit cycle varies unimodally with I , increasing for $I < I_2$ and decreasing for $I > I_2$.

The frequency of oscillation also varies with the choice of numerical parameters. Figure 28 describes

A	E	PERIOD
1.0	.05	18.0
1.0	.10	9.5
1.0	.14	7.0
1.4	.14	8.0
0.5	1.0	1.7

B = 1.0 C = 20 D = 33.33

Fig. 28a-e. Dependence of limit cycle period on decay parameters

a series in which varying the decay rate of $x(t)$ and/or of $y(t)$ by a factor ~ 3 also varies the period of response by a factor ~ 3 . It is possible to generate periods which are even 20 times as large as the excitatory decay rate. The oscillations can therefore be thought of as "slow" compared to the time scale of the recurrent on-center off-surround interaction.

Hysteresis is possible in this system because of the existence of the unstable limit cycle. For example, if $I_1 < I < I_2$ and the system starts on the stable limit cycle $\mathcal{L}(I)$, then as I is slowly increased, it remains close to $\mathcal{L}(I)$, even after the unstable limit cycle is produced. Finally, for $I > I_3$, the system is attracted to the unique stable critical point. If it starts at this point and I is slowly decreased, then it cannot escape from the region bounded by the unstable limit cycle until the cycle vanishes when $I < I_2$. Hysteresis has also been found in a system of this type when two nonzero stable critical points exist, even though no limit cycles were observed for any input value in this case. A phenomenon similar to this latter case was reported by Wilson and Cowan (1972) in their model.

18. Nonexistence of Order-Preserving Limit Cycles in STM

We studied systems (40) and (41) with nonuniform initial data, Gaussian distribution of coefficients, and $f(w) = \max(w - G, 0)$, $G > 0$, to test whether limit cycles can encode patterned activity in STM. We were especially interested in order-preserving periodic phenomena, namely oscillations which preserve an activity index of the relative importance of each population through time (Fig. 29).

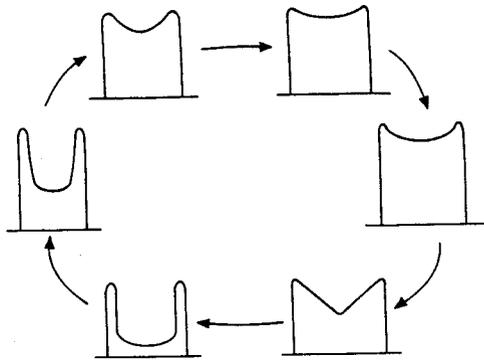


Fig. 29. Order preserving limit cycle in unlumped system

Extensive parameter searches failed to reveal a single example of this phenomenon. A summary of what was found is given below.

It was possible to find a periodic response that is order-preserving if the input pattern (e.g., spike or rectangle) is left on, but whenever this happened, the reverberation decayed to zero when the pattern was turned off. The parameters that produced this response were close to parameters that produced an order-preserving asymptotic steady state, as in Fig. 20a.

Either a decrease in σ_c , keeping $\sum_{k=1}^n C_{ki} = C$ fixed,

an increase in σ_d keeping $\sum_{k=1}^n D_{ki} = D$ fixed, or an

increase in σ_f keeping $\sum_{k=1}^n F_{ki} = F$ fixed, could change the limit cycle into an asymptotic steady state. This confirms the idea that a decrease in $\sigma_c \sigma_f^{-1}$ stabilizes the network. Periodic phenomena also exist while the input pattern is left on, such that the order reverses itself at various phases of the period. Here too the reverberation died when the input was shut off. Finally, there exist persistent reverberations in STM which are not order preserving. In these cases, a uniform background activity level was left on, but the initial data was chosen nonuniformly. Persistent travelling waves were generated, and their form depended on the initial data. Waves travel in one direction in response to a right triangle of initial data. Waves travel outward in both directions in response to an isosceles triangle of initial data. These reverberations indeed "remember" their initial data, in the sense that they generate different wave structures. But it is hard to see how they could possibly retrieve the initial pattern during performance. It is also hard to see how they could sustain any clear index of initial pattern if the pattern were chosen in a more complicated way. In no case could we generate a

persistent STM reverberation unless background activity level was positive.

19. A Spatial STM Code

The above results for lumped and unlumped networks show that a spatial code for STM reverberation has remarkable richness and flexibility, especially when nonspecific inputs (e.g., "arousal") can modify such processes as contrast control and attention using shunting interactions. The results also argue against using slow periodic phenomena, such as the limit cycle behavior in Theorems 6 and 7, or the computer runs of Section 18, for STM storage, as opposed to the fast reverberations which define our spatial code. Slow periodic phenomena are not, however, atypical in the nervous system, for example the various waves, such as the alpha wave and thalamo-cortical reverberations (Anderson and Eccles, 1962), and these waves are usually most evident when the system is not being driven by external inputs. The present results argue against these slow periodic waves as mechanisms for the actual carrying of patterns that are being stored in STM, even though slow periodic waves can carry a pattern in response to sustained inputs. The results do not, however, prevent slow periodic phenomena from acting as shunts that continually retune the structures that are capable of STM storage. For, if recurrent networks can maintain a STM response in the absence of inputs, then there must exist mechanisms capable of terminating the reverberation quickly when new input data arrives. Once again we must make the distinction between the specific cues that are being processed and the nonspecific mechanisms that sustain, transform, or suppress them.

Appendix 1

Proof of Mathematical Results

Theorem 1. First note that $BC \leq A$ implies $x(\infty) = 0$ because (9) implies $\dot{x}_i \leq -Cx_i^2$. Hence assume $BC > A$ below. Rewrite (9) as

$$\dot{x}_i = x_i [BC - A - Dx + (D - C)x_i]. \quad (A.1)$$

By (A.1), there exists a fixed ordering of x_i 's for all $t \geq 0$, which can be written as

$$0 \leq x_1 \leq x_2 \leq \dots \leq x_{n-1} \leq x_n \quad (A.2)$$

without loss of generality. Also the functions $R_{ij} = \log x_i x_j^{-1}$ satisfy

$$\dot{R}_{ij} = (D - C)x_{ij}, \quad (A.3)$$

where $x_{ij} = x_i - x_j$. Since the sign of x_{ij} is constant by (A.2), R_{ij} is monotonic, and therefore limits $R_{ij}(\infty)$, possibly infinite, exist.

The case $D = C$ was treated in Grossberg (1973). Suppose $D \neq C$. If $|R_{ij}(\infty)| < \infty$, then since \dot{R}_{ij} is bounded, $\dot{R}_{ij}(\infty) = 0$. By (A 3), $x_{ij}(\infty) = 0$.

Let $D > C$ and $x_n(0) > x_i(0)$. [If all $x_i(0) = x_n(0)$, then the limiting distribution is uniform.] Then by (A 3), R_{ni} is positive and monotone increasing. Suppose that $R_{ni}(\infty) < \infty$. A contradiction is derived as follows. Write $R_{ni}(\infty) = \log(1 + r_i)$ with $0 < r_i < \infty$. Rewrite

$$\lim_{t \rightarrow \infty} [x_n x_i^{-1} - \exp R_{ij}(\infty)] = 0$$

as

$$\lim_{t \rightarrow \infty} \frac{x_{ni}}{x_i} = r_i.$$

The limit $x_{ni}(\infty) = 0$ because $|R_{ni}(\infty)| < \infty$. Consequently $x_i(\infty) = 0$ as well. Again use the equation $x_{ni}(\infty) = 0$ to conclude that $x_n(\infty) = 0$, and thus that $x(\infty) = 0$. The equation $x_{ni}(\infty) = 0$ holds for any i such that $x_n(0) > x_i(0)$. For any i such that $x_n(0) = x_i(0)$, $x_n \equiv x_i$. Consequently (A 1) can be rewritten with $i = n$ as

$$\dot{x}_n = x_n [U - V_n x_n + 0(1)], \quad (\text{A 4})$$

where $U = BC - A$ and $V_n = C + (n-1)D$. Proceeding in the usual way, let $z = x_n^{-1}$ and obtain the linear equation

$$\dot{z} + Wz = V_n, \quad (\text{A 5})$$

where $W = U + 0(1)$. Integrate this equation and show that $z(\infty) = UV_n^{-1}$. Consequently $x_n(\infty) = V_n U^{-1} > 0$, which contradicts $x(\infty) = 0$. In other words, $BC > A$ and $D > C$ implies that $R_{ni}(\infty) = \infty$ whenever $x_n(0) > x_i(0)$. But then $x_i(\infty) = 0$ for any such index i . Supposing that there are m indices i such that $x_n(0) = x_i(0)$, $1 \leq m < n$, this shows that (A 4) holds with n replaced by m . The limiting distribution is therefore 0-1 or locally uniform depending on whether $m = 1$ or $1 < m < n$.

Let $D < C$ and $x_i(0) > x_j(0)$. By (A 3), R_{ij} decreases to $R_{ij}(\infty) \geq 0$. In particular, $|R_{ij}(\infty)| < \infty$ and thus $x_{ij}(\infty) = 0$. Indeed $x_{ij}(\infty) = 0$ for all i and j . Again (A 4) holds, and shows that $x_n(\infty) = V_n U^{-1}$. Since all $x_{ni}(\infty) = 0$, the limiting distribution is uniform.

Proposition 1. Setting $\dot{y}_i = 0$ in (11) shows that $y_i = x_i$. Setting $\dot{x}_i = 0$ in (10), substituting $y_i = x_i$, and rearranging terms yields $x_i [BC - A - x] = 0$. Unless all $x_i = 0$, $x = BC - A$. This proves (12) and (13).

We now prove local asymptotic stability of the critical points when $n = 2$. First eliminate y_1 by using the fact (see Theorem 2) that

$$\log x_1 x_2^{-1} + CE^{-1}(y_1 - y_2) = S_{12} = \text{constant}.$$

Let $(x_1^*, x_2^*, y_1^*, y_2^*)$ be a critical point with a fixed value of S_{12} . Define the variables $\tau_1 = \log x_1 (x_1^*)^{-1}$, $\tau_2 = \log x_2 (x_2^*)^{-1}$, and $\tau_3 = y_2 - y_1^*$. Steady state is characterized by $\tau_1 = \tau_2 = \tau_3 = 0$. Introduce the notation $U = BC - A$, and find that the variables τ_i satisfy the equations

$$\dot{\tau}_1 = U - Cx_1^* e^{\tau_1} - C(\tau_3 + y_2^*)$$

$$\dot{\tau}_2 = U - Cx_2^* e^{\tau_2} - C(\tau_3 + y_2^*)$$

$$-E[S_{12} + \tau_2 - \tau_1 + \log x_2^* (x_1^*)^{-1}]$$

and

$$\dot{\tau}_3 = E(x_2^* e^{\tau_2} - \tau_3 - y_2^*).$$

Use the identities

$$U - Cx_1^* - Cy_2^* = 0,$$

$$U - Cx_2^* - Cy_1^* = 0,$$

$$x_2^* - y_2^* = 0,$$

and

$$\log x_1^* (x_2^*)^{-1} + CE^{-1}(x_1^* - x_2^*) = S_{12}$$

to rewrite these equations as

$$\dot{\tau}_1 = Cx_1^*(1 - e^{\tau_1}) - C\tau_3$$

$$\dot{\tau}_2 = Cx_2^*(1 - e^{\tau_2}) - C\tau_3 - E(\tau_2 - \tau_1)$$

and

$$\dot{\tau}_3 = E[x_2^*(e^{\tau_2} - 1) - \tau_3].$$

To study the behavior of this system near $\tau_1 = \tau_2 = \tau_3 = 0$, we compute its Jacobian matrix and find

$$\begin{pmatrix} -Cx_1^* & 0 & -C \\ E & -Cx_2^* - E & -C \\ 0 & Ex_2^* & -E \end{pmatrix}.$$

The characteristic polynomial corresponding to this matrix is

$$p_0 \lambda^3 + p_1 \lambda^2 + p_2 \lambda + p_3 = 0,$$

where $p_0 = 1$,

$$p_1 = C(x_1^* + x_2^*) + 2E,$$

$$p_2 = C^2 x_1^* x_2^* + 2CE(x_1^* + x_2^*) + E^2,$$

and

$$p_3 = 2C^2 E x_1^* x_2^* + CE^2(x_1^* + x_2^*).$$

To show that all eigenvalues have negative real parts, we check whether the Routh Hurwitz criteria (Gantmacher, 1959, p. 226) are satisfied. These conditions are $p_0 > 0$, $p_1 p_2 > p_3$, and $p_3 > 0$. Only the second inequality requires checking, and it reduces to the obvious inequality

$$2CE(x_1^* + x_2^*) + C^2 x_1^* x_2^* + E^2 > 0.$$

Theorem 2. We first show that the functions

$$S_{ij} = \log x_i x_j^{-1} + DE^{-1} y_{ij} \quad (\text{A 6})$$

obey the equations

$$\dot{S}_{ij} = (D - C) x_{ij}. \quad (\text{A 7})$$

Write (10) as

$$\dot{x}_i = x_i [BC - A - Dy_i + Cy_i - Cx_i] \quad (\text{A 8})$$

using the notation $y = \sum_k y_k$. From (A 8) follows

$$\dot{R}_{ij} = -Cx_{ij} + Dy_{ij}, \quad (\text{A 9})$$

where again we define $R_{ij} = \log x_i x_j^{-1}$. By (11),

$$\dot{y}_{ij} = Ex_{ij} - Ey_{ij}. \quad (\text{A 10})$$

Equations (A 9) and (A 10) together yield (A 7). These equations also show that if inequality

$$x_{ij}(t) y_{ij}(t) \geq 0 \quad (\text{A 11})$$

holds for some $t = T$, then it holds for all $t \geq T$.

If $D = C$, (A 7) says that S_{ij} is constant. If limits exist, then the boundedness of \dot{y}_i shows $\dot{y}_i(\infty) = 0$, and thus by (11) that $Y_i(\infty) = X_i(\infty)$. Similarly, the boundedness of \dot{x}_i shows $\dot{x}_i(\infty) = 0$, and thus by (A 8) that $y(\infty) = x(\infty) = B - AC^{-1}$.

Rewrite (A 6) in the form

$$S_{ij} = \log X_i X_j^{-1} + DE^{-1} y(Y_i - Y_j). \quad (\text{A 12})$$

Since $S_{ij}(0) = S_{ij}(\infty)$,

$$\begin{aligned} & \log X_i X_j^{-1}(0) + DE^{-1} y(0) [Y_i(0) - Y_j(0)] \\ &= \log X_i X_j^{-1}(\infty) + (B - AC^{-1}) DE^{-1} [X_i(\infty) - X_j(\infty)]. \end{aligned} \quad (\text{A 13})$$

Always suppose that $X_i(0) > X_j(0)$ for definiteness.

If the initial y_i pattern is uniform, then

$$\begin{aligned} & \log X_i X_j^{-1}(0) \\ &= \log X_i X_j^{-1}(\infty) + (B - AC^{-1}) DE^{-1} [X_i(\infty) - X_j(\infty)]. \end{aligned} \quad (\text{A 14})$$

Since $x_{ij}(0)y_{ij}(0) = 0$, $X_i(\infty) \geq X_j(\infty)$. Thus

$$\log X_i X_j^{-1}(0) > \log X_i X_j^{-1}(\infty),$$

which shows that uniformization has occurred. The amount of uniformization varies inversely with E .

Let initial anti-order occur with $S_{ij} > 0$ and $x_{ij}(0) > 0$. Since $Y_i(\infty) = X_i(\infty)$,

$$\begin{aligned} \log X_i X_j^{-1}(0) &> \log X_i X_j^{-1}(\infty) \\ &+ (B - AC^{-1})DE^{-1}[X_i(\infty) - X_j(\infty)] > 0 \end{aligned}$$

and thus $\log X_i X_j^{-1}(0) > \log X_i X_j^{-1}(\infty) > 0$. Hence the y_{ij} order reverses itself and the excitatory pattern is uniformized. If initial anti-order occurs with $S_{ij} < 0$ and $x_{ij}(0) > 0$, the same argument shows that the x_{ij} order reverses itself. If $S_{ij}(0) = 0$, the equations $Y_i(\infty) = X_i(\infty)$ along with (A12) shows that the limiting pattern is uniform.

Let $n = 2$ and suppose that initial iso-pattern exists. By (A12) and the constancy of S_{12} ,

$$\begin{aligned} \log X_1(0)[1 - X_1(0)]^{-1} + 2DE^{-1}y(0)[X_1(0) - \frac{1}{2}] & \quad (A15) \\ = \log X_1(\infty)[1 - X_1(\infty)]^{-1} + 2(B - AC^{-1})DE^{-1}[X_1(\infty) - \frac{1}{2}]. \end{aligned}$$

Equation (A15) can be rewritten as

$$\begin{aligned} 2DE^{-1}[S(X_1(0) - \frac{1}{2}) + (B - AC^{-1})(X_1(0) - X_1(\infty))] & \quad (A16) \\ = \log \left\{ \frac{X_1(\infty)[1 - X_1(0)]}{X_1(0)[1 - X_1(\infty)]} \right\}. \end{aligned}$$

Thus contrast enhancement occurs if $S > 0$ and uniformization occurs if $S < 0$, as arguing by contradiction shows.

Theorem 3. Again it is easily seen that $BC \leq A$ implies $x(\infty) = 0$ because (10) implies $\dot{x}_i \leq -Cx_i^2$. If $BC > A$, then the origin in $2n$ -dimensional $(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n)$ space is a saddle point, since the linearization of (10) and (11) has the form $\dot{w} = Lw$, where

$$w = (\xi_1, \xi_2, \dots, \xi_n, \eta_1, \eta_2, \dots, \eta_n)^T, \text{ and } L = \begin{pmatrix} (BC - A)I & 0 \\ EI & -EI \end{pmatrix} \text{ with}$$

I the $n \times n$ identity matrix. Thus the eigenvalues are $BC - A$ (n times) and $-E$ (n times). The system cannot approach zero just so long as $x(0) > 0$. Assume $BC > A$ below.

Equations (A7), (A9), and (A10) still hold. They imply again that if inequality (A11) is true at $t = T$, then it is also true for all $t \geq T$. In particular x_{ij} changes sign at most once, and an ordering $x_1 \leq x_2 \leq \dots \leq x_n$ is established for all large t . Inequality (A11) also shows that either of two cases hold:

- A. $x_{ij}y_{ij} < 0$ for all large t , or
- B. $x_{ij}y_{ij} \geq 0$ for all large t .

Since we are interested only in limiting distributions, we will below consider our system only after either case (A) or (B) is entered. Since the system is autonomous, we can assume without loss of generality that $x_{ij} \geq 0$ for all $t \geq 0$.

Suppose $D \neq C$ and let case (A) hold. For definiteness, let $x_{ij} \geq 0 \geq y_{ij}$ for all $t \geq 0$. By (A9) and (A10), $\dot{R}_{ij} \leq 0 \leq \dot{y}_{ij}$ for all $t \geq 0$. Consequently limits $R_{ij}(\infty)$ and $y_{ij}(\infty)$ exist such that

$$y_{ij}(0) \leq y_{ij}(\infty) \leq 0 \leq R_{ij}(\infty) \leq R_{ij}(0). \quad (A17)$$

Since \dot{y}_{ij} is bounded, $\dot{y}_{ij}(\infty) = 0$. Equation (A10) is a sum of two nonnegative terms. Thus $x_{ij}(\infty) = y_{ij}(\infty) = 0$.

Suppose $D \neq C$ and note by (A7) that S_{ij} is monotonic for all large t , so $S_{ij}(\infty)$ exists but is possibly infinite. If $|S_{ij}(\infty)| < \infty$, then since \dot{S}_{ij} is bounded, it follows that $\dot{S}_{ij}(\infty) = 0$. By (A7), also $x_{ij}(\infty) = 0$, and then (A10) shows that $y_{ij}(\infty) = 0$.

In particular, suppose that $D < C$ and let case (B) hold. If $x_{ij} \geq 0$ for all large t , then (A7) shows that S_{ij} decreases to a limit

$S_{ij}(\infty) \geq 0$. In particular, $|S_{ij}(\infty)| < \infty$, so that $x_{ij}(\infty) = y_{ij}(\infty) = 0$. Pasting together the results for cases (A) and (B), it follows that whenever $D < C$, all $x_{ij}(\infty) = y_{ij}(\infty) = 0$ no matter how the initial data is chosen. Can we conclude that $X_i(\infty) = Y_i(\infty) = \frac{1}{n}$ from

this information? Suppose $S_{ij}(0) > 0$ and $x_{ij} \geq 0$. Then (A7) shows that S_{ij} decreases to $S_{ij}(\infty)$. From $x_{ij} \geq 0$ and $y_{ij}(\infty) = 0$ follows that $\log x_i x_j^{-1}$ decreases to $S_{ij}(\infty) \geq 0$. If $S_{ij}(\infty) > 0$, write $S_{ij}(\infty) = \log(1 + r_{ij})$ with $0 < r_{ij} < \infty$, and conclude that $\lim_{t \rightarrow \infty} x_{ij} x_j^{-1} = 1 + r_{ij} > 0$. Since $x_{ij}(\infty) = 0$, $x_j(\infty) = 0$.

But all $x_{kj}(\infty) = 0$, hence all $x_k(\infty) = 0$, which contradicts the persistence of the reverberation. Consequently $\log X_i X_j^{-1}$ decreases to $S_{ij}(\infty) = 0$, and thus $X_i X_j^{-1}(\infty) = 1$. If $S_{ij}(0) < 0$ and $x_{ij} \geq 0$, then again $S_{ij}(\infty) \geq 0$ because $y_{ij}(\infty) = 0$. Consequently, if $D < C$, $X_i X_j^{-1}(\infty) = 1$ for all i and j . The limiting distribution is consequently uniform.

Suppose $D > C$. If initial iso-order exists in the pair (i, j) with $S_{ij}(0) > 0$, then S_{ij} increases to $S_{ij}(\infty) \leq \infty$. If $S_{ij}(\infty) < \infty$, then as above we can show that $x_{ij}(\infty) = y_{ij}(\infty) = 0$. But then

$$\log x_i x_j^{-1}(\infty) = S_{ij}(\infty) > 0, \quad (A18)$$

which along with $x_{ij}(\infty) = 0$ shows that $x_i(\infty) = x_j(\infty) = 0$. Thus if the reverberation is persistent and initial iso-order exists in all pairs, then $S_{ni}(\infty) = \infty$ whenever $S_{ni}(0) > 0$. The equation $S_{ni}(\infty) = \infty$ implies that $x_n x_i^{-1}(\infty) = \infty$. In summary, either $x_n \equiv x_i$ or $x_n x_i^{-1}(\infty) = \infty$. The ratio $X_n = \left(1 + \sum_{i \neq n} x_i x_n^{-1}\right)^{-1}$ therefore converges to m^{-1} ,

where m is the number of i such that $x_i \equiv x_n$.

Suppose again that $D > C$, but that anti-order exists in the pair (i, j) at large times with $x_{ij} \geq 0$. If $S_{ij} > 0$ at such times, then the above argument can be repeated to show that either $x_i x_j^{-1}(\infty) = \infty$ or that $x_i(\infty) = x_j(\infty) = 0$. If $S_{ij} < 0$, then S_{ij} increases to $S_{ij}(\infty) \leq \infty$. If $S_{ij}(\infty) < \infty$, then as above, we can show that $x_i x_j^{-1}(\infty) = \infty$. If $S_{ij}(\infty) = \infty$, then again $x_{ij}(\infty) = y_{ij}(\infty) = 0$, so that inequality $x_{ij} \geq 0$ implies $S_i(\infty) \geq 0$. If $\infty > S_{ij}(\infty) > 0$, we can argue as above to conclude that $x_i(\infty) = x_j(\infty) = 0$. The case $S_{ij}(\infty) = 0$ remains, and here $x_i x_j^{-1}(\infty) = 1$ and $x_{ij}(\infty) = y_{ij}(\infty) = 0$ despite the possibility that $x_{ij}(0) > 1$.

Theorem 4. The method is a direct generalization of the method used to prove Theorem 3. The equations that generalize (A7), (A9), and (A10) use the old definitions $R_{ij} = \log x_i x_j^{-1}$, $y_{ij} = y_i - y_j$, and the modified definition

$$S_{ij} = R_{ij} + K y_{ij}. \quad (A19)$$

Note that Theorem 4 reduces to Theorem 3 in the linear signal function case if $K = DE^{-1}$. These functions obey the equations

$$\dot{S}_{ij} = U(x_i) - U(x_j) + V(y_i) - V(y_j), \quad (A20)$$

$$\dot{R}_{ij} = (Bx_i^{-1} - 1)f(x_i) - (Bx_j^{-1} - 1)f(x_j) + g(y_i) - g(y_j), \quad (A21)$$

$$\text{and } \dot{y}_{ij} = -E y_{ij} + h(x_i) - h(x_j). \quad (A22)$$

From (A21) and (A22), it is clear that if (A11) holds at some time $t = T$, then it holds for all $t \geq T$. Again cases (A) and (B) in the proof of Theorem 3 are exhaustive. The remainder of the proof imitates the proof of Theorem 3, using such facts as $x_{ij}[U(x_i) - U(x_j)] \geq 0$ and $y_{ij}[V(y_i) - V(y_j)] \geq 0$. In Case (A), y_{ij} is again monotonic, hence $y_{ij}(\infty)$ exists, and since it and \dot{y}_{ij} are bounded, $\dot{y}_{ij}(\infty) = 0$, from which it follows that $y_{ij}(\infty) = 0$ and that $h(x_i) - h(x_j) \rightarrow 0$ as $t \rightarrow \infty$. Since $h(w)$ is strictly monotonic, $x_{ij}(\infty) = 0$. Now we can complete the arguments for this case as in proof of Theorem 3.

In case (B), (A22) shows that S_{ij} is monotonic at large times, and hence $S_{ij}(\infty)$ exists. If $|S_{ij}(\infty)| < \infty$, then $\dot{S}_{ij}(\infty) = 0$ as usual, and

$$\lim_{t \rightarrow \infty} [U(x_i) - U(x_j)] = \lim_{t \rightarrow \infty} [V(y_i) - V(y_j)] = 0.$$

If $U(w)$ is strictly monotonic, then $x_{ij}(\infty) = 0$, which by integration of (A22) shows that $y_{ij}(\infty) = 0$. If $V(w)$ is strictly monotonic, then $y_{ij}(\infty) = 0$. Again integration of (A22) shows that $x_{ij}(\infty) = 0$. To see this, write (A22) as

$$y_{ij}(t) = y_{ij}(0) e^{-Et} + \frac{\int_0^t h_{ij}(v) e^{Ev} dv}{e^{Et}}, \quad (\text{A23})$$

with $h_{ij} = h(x_i) - h(x_j)$. Suppose $x_{ij} \rightarrow 0$ as $t \rightarrow \infty$. Then $h_{ij} \rightarrow 0$ as $t \rightarrow \infty$. Given that case (B) holds, we can let $x_{ij} \geq 0$, and thus $h_{ij} \geq 0$, without loss of generality. By (17), \dot{x}_{ij} is bounded. Thus there exist positive δ and ε such that $h_{ij}(t) \geq \varepsilon$ if $t \in \bigcup_{k=1}^{\infty} [t_k - \delta, t_k + \delta]$

for suitable $t_1 < t_2 < \dots < t_k \rightarrow \infty$. Consequently $\int_0^{\infty} h_{ij}(v) e^{Ev} dv = \infty$.

Analysis of (A13) as $t \rightarrow \infty$ therefore requires L'Hospital's rule, from which we find that $h_{ij}(\infty) = y_{ij}(\infty) = 0$, or that $x_{ij}(\infty) = 0$. The remainder of the proof requires that we separately analyse the cases $|S_{ij}(\infty)| = \infty$ and $|S_{ij}(\infty)| < \infty$. This can be done just as in Theorem 3.

Proposition 2. Compute equations for \dot{x}_{ij} and \dot{y}_{ij} . Show that if $x_{ij}(t) = 0$ then $\dot{x}_{ij}(t) y_{ij}(t) \geq 0$, and that if $y_{ij}(t) = 0$ then $\dot{y}_{ij}(t) x_{ij}(t) \geq 0$.

Theorem 5. A routine computation shows that if $n = 2$,

$$\dot{R}_{12} = z_{12}, \quad (\text{A24})$$

and

$$\dot{z}_{12} = -(E + Cx) z_{12} + CH x_{12}, \quad (\text{A25})$$

where $x = x_1 + x_2$. If $H = 0$, (A24) can be integrated to yield

$$z_{12}(t) = z_{12}(0) \exp\left(-Et - C \int_0^t x dv\right). \quad (\text{A26})$$

Equations (A24) and (A26) show that R_{12} is monotonic. If $z_{12}(0) = 0$, they show that $\dot{R}_{12} \equiv 0$, or that X_1 and X_2 are constant. If $x_{12}(0) \cdot z_{12}(0) \geq 0$ with $z_{12}(0) > 0$, they show that $X_1 X_2^{-1}$ is strictly increasing. However by (A26),

$$\dot{R}_{12}(t) \leq z_{12}(0) e^{-Et}, \quad (\text{A27})$$

so that $R_{12}(\infty) < \infty$, and the contrast enhancement is partial. If $x_{12}(0) z_{12}(0) \geq 0$ with $z_{12}(0) < 0$, then (A24) and (A26) show that $\dot{R}_{ij} < 0$, so that the pattern is uniformized. It is also uniformized if $x_{12}(0) z_{12}(0) < 0$, since then $x_{12}(t) \dot{R}_{12}(t) < 0$ for all $t \geq 0$.

If $n > 2$, then at equilibrium, $F(x_i(\infty)) = F(x_j(\infty))$ for all i and j such that $x_i(\infty) x_j(\infty) > 0$, where $F(w) = w[1 - DF(E + Cw)^{-1}]$ is a strictly increasing function.

Proposition 3. Using $C(w, z) = C\delta(w - z)$, we find that

$$x(w) = \frac{BCI(w)}{A + CI(w) + \int_{-\infty}^{\infty} I(z) D_1(z - w) dz}, \quad (\text{A28})$$

Since the input is even, $\bar{x} = x(0)$. Defining $P(w) = N(w) M^{-1}(w)$, we note that

$$P(w) = \frac{A + CI(0) + \int_{-\infty}^{\infty} I(z) D_1(z) dz}{A + CI(w) + \int_{-\infty}^{\infty} I(w - z) D_1(z) dz}. \quad (\text{A29})$$

Since $D_1(z)$ and $I(z)$ are even functions, it suffices to show that $P(w) \geq 1$ if $I'(w) \leq 0$ for $w \geq 0$. Define

$$Q(w) = \int_{-\infty}^{\infty} D_1(z) [I(z) - I(w - z)] dz. \quad (\text{A30})$$

Since $I(0) \geq I(w)$, (A29) shows that it suffices to prove that $Q(w) \geq 0$. Obviously $Q(0) = 0$. The proof will be completed by showing that $z = 0$ is a global minimum of $Q(z)$. We therefore compute

$$\begin{aligned} Q'(w) &= - \int_{-\infty}^{\infty} D_1(z) I'(w - z) dz \\ &= \int_{-\infty}^{\infty} D_1(w - z) r(z) dz, \end{aligned}$$

where $r(w) = -I'(w)$. Since $I(w)$ is even, $r(w) = -r(-w)$, and thus

$$Q'(w) = \int_0^{\infty} S(w, z) r(z) dz, \quad (\text{A31})$$

where $S(w, z) = D_1(w - z) - D_1(w + z)$. By the evenness of $D_1(\xi)$, $Q'(0) = 0$. Furthermore, since $D_1(\xi)$ is monotone decreasing, $S(w, z) \geq 0$ if $w \geq 0$. Likewise $S(w, z) \leq 0$ if $w < 0$. (A31) therefore shows that $w = 0$ is a global minimum of Q , and hence $P \geq 1$ for all w .

Proposition 4. In this case, we define $P = NM^{-1}$, $\alpha = AI^{-1}$,

$$F(w) = C(w) \exp(-Jw^2), \quad (\text{A32})$$

$$P_1(w) = \frac{\alpha + \int_{-\infty}^{\infty} E_1(z) \exp(-Jz^2) dz}{\alpha + \int_{-\infty}^{\infty} E_1(z) \exp[-J(w - z)^2] dz}, \quad (\text{A33})$$

and

$$P_2(w) = \frac{\int_{-\infty}^{\infty} F(z) \exp(2Jwz) dz}{\int_{-\infty}^{\infty} F(z) dz} \quad (\text{A34})$$

so that we can write

$$P(w) = P_1(w) P_2(w). \quad (\text{A35})$$

We will prove that $P \geq 1$ by separately proving that $P_1 \geq 1$ and $P_2 \geq 1$.

First we study the function

$$p_1(w) = \text{sign}[P_1(w) - 1],$$

where

$$\text{sign}(\xi) = \begin{cases} 1 & \text{if } \xi > 0 \\ 0 & \text{if } \xi = 0 \\ -1 & \text{if } \xi < 0. \end{cases}$$

Clearly

$$p_1(w) = \text{sign } r_1(w),$$

where

$$r_1(w) = \int_{-\infty}^{\infty} E_1(z) [e^{-Jz^2} - e^{-J(w-z)^2}] dz.$$

Noting that $r_1(0) = 0$, it suffices to show that $r_1(w)$ has a global minimum at $w = 0$ to prove that $P_1 \geq 0$. Compute

$$r_1'(w) = 2J \int_{-\infty}^{\infty} E_1(z) (w - z) e^{-J(w-z)^2} dz.$$

Changing variables and using the evenness of the Gaussian, we define

$$T(w, z) = E_1(w - z) - E_1(w + z)$$

to conveniently write

$$r_1'(w) = 2J \int_0^{\infty} z T(w, z) e^{-Jz^2} dz. \quad (\text{A36})$$

An argument similar to the one used on (A31) proves our assertion.

To study P_2 , we define

$$p_2(w) = \text{sign}[P_2(w) - 1].$$

Clearly

$$p_2(w) = \text{sign } r_2(w),$$

where

$$r_2(w) = \int_{-x}^{\infty} F(z) (e^{2Jwz} - 1) dz.$$

Again $r_2(0) = 0$, and to show that $r_2(w)$ has an absolute minimum at $w = 0$ we compute

$$r_2'(w) = 2J \int_0^x z F(z) (e^{2Jwz} - e^{-2Jwz}) dz,$$

from which the proof is readily completed.

Proposition 5. In this case, we define

$$P_1(w) = \frac{A + \int_{-\infty}^{\infty} I(z) E_1(z) dz}{A + \int_{-\infty}^{\infty} I(w-z) E_1(z) dz} \quad (\text{A 37})$$

and

$$P_2(w) = \frac{I^{-1}(w) \int_{-\infty}^{\infty} I(w-z) C(z) dz}{I^{-1}(0) \int_{-\infty}^{\infty} I(z) C(z) dz} \quad (\text{A 38})$$

in order to write

$$P(w) = P_1(w) P_2(w).$$

It is possible to prove that $P_1 \geq 1$ in the manner of Proposition 3. We will show, however, that $P_2 < 1$ for $I(w)$ chosen as stated in the Proposition. By choosing parameter A sufficiently large, we can therefore force $P_1(w_0)$ so close to 1 that $P(w_0) < 1$ for some value w_0 of w .

First we expand $I(z)$ in Taylor's series about $z = 0$ and $z = w$ in (A 38). Since $I(z)$ and $C_1(z)$ are even, all odd derivatives of $I(z)$ vanish in this expansion. Letting $C = \int_{-w}^w C(z) dz$, what remains is

$$P_2(w) = \frac{C + I^{-1}(w) \sum_{k=1}^{\infty} \frac{I^{(2k)}(w)}{(2k)!} \int_{-w}^w z^{2k} C(z) dz}{C + I^{-1}(0) \sum_{k=1}^{\infty} \frac{I^{(2k)}(w)}{(2k)!} \int_{-w}^w z^{2k} C(z) dz} \quad (\text{A 39})$$

Assume $I(w) = U - Vw^2$ for $|w| < W$. Letting $M = \int_{-w}^w z^2 C(z) dz$, we find

$$P_2(w) = \frac{C - VM(U - Vw^2)^{-1}}{C - VMU^{-1}} < 1.$$

Fixing a $w = w_0$ with $|w_0| < W$, we can therefore increase parameter A in (A 38) until $P(w_0) < 1$.

Proposition 6. Computing $x'(w)$ from (A 28) we find that

$$\text{sign } x'(w) = \text{sign } r(w),$$

where

$$r(w) = I'(w) \left[A + \int_{-\infty}^{\infty} I(w+z) D_1(z) dz \right] - I(w) \int_{-\infty}^{\infty} I'(w+z) D_1(z) dz.$$

Let $w = 0$. Using $I'(0) = 0$, $I'(z) > 0$ for $z \neq 0$, and $D_1(0) > 0$, we conclude that

$$r(0) = -I(0) \int_{-\infty}^{\infty} I'(z) D_1(z) dz < 0$$

and finally that $x'(0) < 0$.

Proposition 7. From (27) with $I(w) = I\delta(w)$, we find

$$x(w) = \frac{BC(w)}{\alpha + C(w) + D(w)}$$

where $\alpha = AI^{-1}$. Consequently

$$\text{sign } x'(w) = \text{sign} \{ C'(w) [\alpha + D(w)] - C(w) D'(w) \}.$$

Given Gaussian coefficients, $x'(0) = 0$ and

$$\text{sign } x''(w) = \text{sign} [D(w)(\mu\nu^{-1} - 1) - AI^{-1}].$$

Consequently $x''(0) \leq 0$ if and only if (28) holds; otherwise there must be at least two peaks in the graph of $x(w)$ because $\lim_{w \rightarrow \pm\infty} x(w) = 0$.

There are no more than two peaks because the function $D(w) \cdot (\mu\nu^{-1} - 1) - AI^{-1}$ is strictly monotone decreasing for $w \geq 0$, and thus at most one critical point w such that $x'(w) = 0$ can exist for $w \geq 0$.

Proposition 8. From (27) follows that

$$x(w) = \frac{B \int_{w-W}^{w+W} C_1(z) dz}{\alpha + \int_{w-W}^{w+W} E_1(z) dz} \quad (\text{A 40})$$

where $\alpha = AI^{-1}$. Define $y(w) = x(w) x^{-1}(0)$,

$$y_1(w) = \frac{\alpha + \int_{-w}^w E_1(z) dz}{\alpha + \int_{w-W}^{w+W} E_1(z) dz} \quad (\text{A 41})$$

and

$$y_2(w) = \frac{\int_{-w}^{w+W} C_1(z) dz}{\int_{-w}^w C_1(z) dz} \quad (\text{A 42})$$

in order to be able to write

$$y(w) = y_1(w) y_2(w). \quad (\text{A 43})$$

For $W_d - W \leq w \leq W - W_d$, both $y_1(w)$ and $y_2(w)$ equal 1, since no boundary effects appear in any of the integrals. Hence $x(w) = x(0)$ in this region. For either $W_c - W < w < W_d - W$ or $W - W_d < w < W - W_c$, $y_2(w) = 1$ for the same reason, whereas $y_1(w) > 0$ because the shifted integral in the denominator of (A 41) does not include part of the inhibitory interactions. Thus $x(w) > x(0)$ in this region.

To evaluate $y(W) = y_1(W) y_2(W)$, note that

$$y_1(W) = \frac{\alpha + 2 \int_0^W E_1(z) dz}{\alpha + \int_0^W E_1(z) dz} < 2,$$

and

$$y_2(W) = \frac{\int_0^W C_1(z) dz}{\int_{-W}^0 C_1(z) dz} = \frac{1}{2}.$$

Thus $x(W) < x(0)$. To evaluate $x'(W)$, define

$$U(x) = \int_{w-W}^{w+W} C_1(z) dz,$$

and

$$V(w) = \int_{w-W}^{w+W} D_1(z) dz.$$

Then by (A 40),

$$\text{sign } x'(w) = \text{sign} \{ U'(w)[x + V(w)] - U(w)V'(w) \},$$

where whenever $w \geq 0$,

$$U'(w) = C_1(w+W) - C_1(w-W) = -C_1(w-W)$$

and

$$V'(w) = D_1(w+W) - D_1(w-W) = -D_1(w-W).$$

Thus setting $w = W$,

$$\text{sign } x'(w) = \text{sign} \{ D_1(0)V'(W) - C_1(0)[x + V(W)] \},$$

from which (33) readily follows.

Theorem 6. The existence of a critical point is an obvious consequence of (47). Uniqueness is established as follows. Either $f(w)$ has a positive threshold F or not. If it does, and a critical point $x = y \leq F$ exists, then by (42) it is clearly unique since $f(x) = 0$. If a critical point satisfies $x = y > F$, or there is no threshold, then we can define $\Theta(x) = x$ and

$$\phi(x) = f^{-1} \left\{ \frac{-Ax + (B-x)[Cf(x) + I]}{Dx} \right\}$$

and study the values of x for which $\Theta(x) = \phi(x)$. To prove uniqueness, it suffices to show that

$$\frac{d\phi}{dx}(x) < 1 \text{ whenever } \Theta(x) = \phi(x). \quad (\text{A } 44)$$

That this proves uniqueness is seen by defining $g(x) = \phi(x) - \Theta(x)$, and supposing that there exist two consecutive critical points x_1 and x_2 that satisfy (A 44). Then $g(x_1) = g(x_2) = 0$ whereas $g'(x_i) < 0$, $i = 1, 2$. This contradicts the continuity of $g(x)$ for $x_1 \leq x \leq x_2$. To prove (A 44), we note that

$$\text{sign} \left[\frac{d\phi}{dx}(x) - 1 \right] \\ = \text{sign} \{ [BC - (C + D)x] f'(x) - [A + I + (C + D)f(x)] \}$$

whenever $\Theta(x) = \phi(x)$. Condition (50) therefore implies (A 44). If (49) holds, then also all critical points satisfy $x > BC(C + D)^{-1}$, and again (A 44) is proved.

To study the stability of the unique critical point, we first rule out the possibility of a saddle point. This can be done by noting that the direction field of the system (42) and (43) on the Jordan curve defined by

$$\{x = 0, 0 \leq y \leq B\} \cup \{y = 0, 0 \leq x \leq B\} \cup \{x = B, 0 \leq y \leq B\} \\ \cup \{y = B, 0 \leq x \leq B\}$$

always points inward. The Index Theorem states that in this case, the index of the curve is -1 , but the index of a saddle point is $+1$ (Coddington and Levinson, 1955). Hence the critical point is a node or a spiral point. It therefore suffices to show that the trace $G - E$ of the linearized system

$$\frac{d}{dt} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} G & H \\ E & -E \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (\text{A } 45)$$

of (42) and (43) is positive, where

$$G = (B - x)Cf'(x) - (C + D)f(x) - A - I, \quad (\text{A } 46)$$

$$H = -Dxf'(x), \quad (\text{A } 47)$$

and $(x, x)^T$ is the unique critical point. But $G > E$ follows immediately from (48) with $f(x)$ defined by (47).

Theorem 7. The proof of Theorem 6 shows that conditions (47), (57), and (58) have the meaning indicated in (A), (B), and (C)

alongside their statement, where the trace and determinant correspond to the linearized system (A 45).

In particular, (57) and (58) guarantee the existence of a pair of conjugate imaginary eigenvalues for the linearized system (A 45). This information, along with (56) in the 2-dimensional case, implies that the Hopf theorem holds.

Corollary 3. The values of I^* and x^* satisfy (47), (57), and (58). To show that (56) holds, we compute that near I^* ,

$$\text{Re } \lambda(I) = 39.3328 - 73.332x - I$$

where

$$x = \frac{25,208 - 625I - \sqrt{390,625I^2 + 51,821,250I - 31,206,736}}{66,665}.$$

From this it can be shown that

$$\text{Re } \frac{d\lambda}{dI}(I^*) < 0.$$

Instability is proved using a procedure described by Hopf. This procedure is summarized below.

Let a be an eigenvector of $\lambda(0)$ in (55). Define a vector e that satisfies

$$a \cdot e = \frac{1}{\lambda(0) - \bar{\lambda}(0)}. \quad (\text{A } 48)$$

Solve the linearized system derived from (53) and (54), namely

$$\dot{U}_0 = L_0(U_0), \quad U_0(0) \cdot e = 0, \quad \dot{U}_0(0) \cdot e = 1. \quad (\text{A } 49)$$

Solve the linear inhomogeneous system

$$\dot{U}_1 = L_0(U_1) + Q_0(U_0, U_0), \quad U_1(0) \cdot e = \dot{U}_1(0) \cdot e = 0. \quad (\text{A } 50)$$

Solve the adjoint system

$$\dot{V}_0 = -L_0^*(V_0), \quad \int_0^{T_0} \dot{U}_0 \cdot V_0 d\tau = 0, \quad \int_0^{T_0} U_0 \cdot V_0 d\tau = 1, \quad (\text{A } 51)$$

where $T_0 = 2\pi|\lambda(0)|^{-1}$. Given this information, we compute the value

$$R = \int_0^{T_0} V_0 \cdot [2Q_0(U_0, U_1) + C_0(u_0, U_0, U_0)] d\tau. \quad (\text{A } 52)$$

If $R > 0$ the periodic solution is unstable, whereas if $R < 0$ the periodic solution is stable. Moreover, the equation

$$\mu_2 \text{Re}(\lambda'(0)) = -R \quad (\text{A } 53)$$

holds, where μ_2 is defined by the expansion $\mu(\epsilon) = \sum_{k=0}^{\infty} \mu_k \epsilon^k$. If $\mu_2 > 0$ (< 0), then the periodic solutions exist only if $\mu > 0$ (< 0).

In the present case, the following equations show that $R > 0$ and thus that $\mu_2 > 0$:

$$x^* \cong 0.5032.$$

$$\lambda \equiv \lambda(I^*) = i\sqrt{\frac{5}{3}x^*f'(x^*) - 1} \cong 3.972i,$$

$$a = \left(\frac{1}{2} - \frac{i}{2\lambda}, -\frac{i}{2\lambda} \right),$$

$$e = (0, 1)^T,$$

$$U_0 = (\cos \lambda t + \lambda^{-1} \sin \lambda t, \lambda^{-1} \sin \lambda t)^T,$$

$$U_1 = -0.741 - 2.636 \sin \lambda t - 0.9834 \cos \lambda t - 1.198 \sin 2\lambda t \\ + 1.725 \cos 2\lambda t - 0.741 - 0.390 \sin \lambda t + 0.5656 \cos \lambda t \\ + 0.195 \sin 2\lambda t + 0.1754 \cos 2\lambda t,$$

and

$$V_0 = \left(\frac{\lambda}{2\pi} \cos \lambda t, \frac{\lambda}{2\pi} [\lambda \sin \lambda t - \cos \lambda t] \right)^T.$$

Then to compute R , we need the equations

$$Q_0(m, n) = (-20[m_1 n_1 + \frac{5}{6}(m_1 n_2 + m_2 n_1)], 0)^T,$$

for any vectors $m = (m_1, m_2)$ and $n = (n_1, n_2)$, and $C_0 \equiv 0$. Therefore, letting $Q_0 = (Q_{01}, Q_{02})$, we find

$$Q_{01}(U_0, U_1) = (A_1 \cos \lambda t + A_2 \sin \lambda t + A_3 \sin \lambda t \cos \lambda t \\ + A_4 \cos^2 \lambda t + A_5 \cos \lambda t \sin 2\lambda t + A_6 \cos \lambda t \cos 2\lambda t \\ + A_7 \sin^2 \lambda t + A_8 \cos \lambda t \sin 2\lambda t + A_9 \sin \lambda t \cos 2\lambda t,$$

where the A_k , $k = 1, 2, \dots, 8$, are constants; e.g.,

$$A_1 = 296.4, \quad A_6 = -26.958, \quad \text{and} \quad A_8 = 12.04.$$

By (A 52),

$$R = \frac{1}{2\pi} \int_0^{2\pi} \left(A_1 \cos^2 t + A_6 \cos^2 t \cos 2t + \frac{A_8}{2} \sin^2 2t \right) dt$$

from which the equality $R > 0$ is immediate.

References

- Anderson, P., Eccles, J.C.: *Nature (Lond.)* **196**, 645 (1962)
- Anderson, P., Gross, G.N., Lomo, T., Sveen, O.: In: Brazier, M. (Ed.): *The interneuron*, p. 415. Los Angeles: Univ. of California Press 1969
- Bishop, P. O., Coombs, J.S., Henry, G.M.: *J. Physiol.* **219**, 625 (1971)
- Blakemore, C., Carpenter, R.H.S., Georgeson, M.A.: *Nature* **228**, 37 (1970)
- Blakemore, C., Cooper, G.F.: *Nature* **228**, 477 (1970)
- Coddington, E.A., Levinson, N.: *Theory of ordinary differential equations*, p. 398. New York: McGraw-Hill Book Co., Inc. 1955
- Colonnier, M.L.: In: Eccles, J.C. (Ed.): *Brain and conscious experience*. Berlin-Heidelberg-New York: Springer 1965
- Cornsweet, T.N.: *Visual perception*. New York: Academic Press 1970
- Dichter, M., Spencer, W.A.: *J. Neurophysiol.* **32**, 663 (1969)
- Eccles, J.C.: In: Kimble, G.A. (Ed.): *Anatomy of memory*, p. 57. Palo Alto, Calif: Science and Behavior Books, Inc. 1965
- Eccles, J.C., Ito, M., Szentagothai, J.: *The cerebellum as a neuronal machine*. Berlin-Heidelberg-New York: Springer 1967
- Furman, G.G.: *Kybernetik* **6**, 257 (1965)
- Gantmacher, F.F.: *The theory of matrices*. New York: Chelsea 1959
- Gibson, J.J., Radner, M.: *J. exp. Psychol.* **20**, 453 (1937)
- Grossberg, S.: *J. theor. Biol.* **33**, 225 (1971)
- Grossberg, S.: *Math. Biosci.* **15**, 39 (1972a)
- Grossberg, S.: *Math. Biosci.* **15**, 253 (1972b)
- Grossberg, S.: *Stud. appl. Math.* **52**, 213 (1973)
- Grossberg, S.: *Progr. theor. Biol.* **3**, 51 (1974)
- Grossberg, S.: In: Pfeiffer, C.C. (Ed.): *Int. Rev. Neurobiol.* **18**, (1975)
- Grossberg, S., Levine, D.: Some developmental and attentional biases in the contrast enhancement and short term memory of recurrent neural networks. *J. theor. Biol.* (in press 1975)
- Hahn, W.: *Stability of motion*. Transl.: A. P. Baartz. Berlin-Heidelberg-New York: Springer 1967
- Hodgkin, A.L.: *The conduction of the nervous impulse*. Springfield, Ill.: C. C. Thomas 1964
- Hopf, E.: *Berichte der Mathematisch-Physischen Klasse der Sächsischen Akademie der Wissenschaftler zu Leipzig*, XCIV. Band (1942)
- Julesz, B.: *Foundations of cyclopean perception*. Chicago: University of Chicago Press 1971
- Kandel, E.R., Spencer, W.A., Brinley, F.J.: *J. Neurophysiol.* **24**, 225 (1961)
- Kuno, M., Miyahara, J.T.: *J. Physiol.* **201**, 465 (1969)
- Lebovitz, R.M., Dichter, M., Spencer, W.A.: *Int. J. Neurosc.* **2**, 99 (1971)
- Levine, D., Grossberg, S.: Visual illusions in neural nets: Line neutralization tilt after-affect, and angle expansion. *Subm. for publ.* (1975)
- MacGregor, R.J.: *Biophys. J.* **8**, 305 (1968)
- Phillips, G.C.: *Quart. J. exp. Physiol.* **44**, 1 (1959)
- Rall, W., Rinzel, J.: *Biophys. J.* **13**, 648 (1973)
- Schiebel, M.E., Schiebel, A.B.: In: Schmitt, F.O. (Ed.): *The neurosciences—second study program*. New York: Rockefeller University Press 1970
- Sperling, G.: *Percept. Psychophys.* **8**, 143 (1970)
- Sperling, G., Sondhi, M.M.: *J. Opt. Soc. Amer.* **58**, 1133 (1968)
- Stefanis, C.: In: Brazier, M. (Ed.): *The interneuron*, p. 497. Los Angeles: University of California Press 1969
- Werblin, F.S.: *J. Neurophysiol.* **34**, 228 (1971)
- Wilson, H.R., Cowan, J.D.: *Biophys. J.* **12**, 1 (1972)
- Wilson, H.R., Cowan, J.D.: *Kybernetik* **13**, 55 (1973)
- Wise, D.C., Berger, B.D., Stein, L.: *Biol. Psychiat.* **6**, 1 (1973)

Prof. Dr. Stephen Grossberg
Department of Mathematics
Boston University
Boston, Mass. 02215, USA