to a band of spatial frequencies in the input patterns.

Pattern matching is illustrated as follows. Suppose in A14 that each input I_i is a sum of two inputs J_i and K_i whose patterns $J = (J_1, J_2, \ldots, J_n)$ and $K = (K_1, K_2, \ldots, K_n)$ are to be matched. If J and K mismatch each other's peaks and troughs to form an almost uniform total pattern $I = (I_1, I_2, \ldots, I_n)$, then by A15 all x_i will be inhibited if $CB^{-1} \geq (n-1)^{-1}$. By contrast, if the two patterns reinforce each other, say $J_i = \alpha K_i$, then by (A15),

$$x_i = \frac{(B + C)(1+\alpha)\bar{K}}{A + (1+\alpha)\bar{K}}\left[\theta_i - \frac{C}{B + C}\right], \quad \text{(A18)}$$

where

$$\bar{K} = \sum_{i=1}^{n} K_i$$

and $\theta_i = K_i(\bar{K})^{-1}$. In other words, matching J and K amplifies each x_i without changing the pattern θ_i.

Appendix D

This section summarizes some properties of recurrent on-center off-surround networks, including normalization, contrast enhancement, quenching threshold, and STM properties.

To see how recurrent networks normalize their STM activity, we first note by Appendix C that these networks need competitive interactions to solve the noise-saturation dilemma. The simplest recurrent on-center off-surround network is defined by

$$\frac{d}{dt}x_i = -Ax_i + (B - x_i)[f(x_i) + I_i] - x_i[\sum_{k \neq i} f(x_k) + J_i], \quad \text{(A18)}$$

$i = 1, 2, \ldots, n$. As usual, x_i is the STM activity of v_i, term $(B - x_i)f(x_i)$ describes the self-excitation of v_i via a positive feedback signal $f(x_i)$—the recurrent on-center—and term

$$-x_i \sum_{k \neq i} f(x_k)$$

describes the inhibition of v_i via negative feedback signals $f(x_k)$, $k \neq i$—the recurrent off-surround. Term I_i is the ith excitatory input, and term J_i is the ith inhibitory input, for example,

$$J_i = \sum_{k \neq i} I_k$$

in A12.

Contrast Enhancement, Normalization, and Quenching Threshold

An important problem in system A18 is to choose the feedback signal function $f(w)$ as a function of activity level w in such a way as to suppress noise but contrast enhance and store in STM behaviorally important patterns. This problem was solved in Grossberg (1973). The solution is reviewed in Grossberg (1978e, Sections 14 and 15).

To understand the simplest STM properties, A18 is transformed into pattern variables $X_i = x_i x^{-1}$ and total activity variables

$$x = \sum_{k=1}^{n} x_k$$

using the notation $g(w) = w^{-1}f(w)$ and supposing that all $I_i = J_i = 0$. Then

$$\frac{d}{dt}X_i = BX_i \sum_{k=1}^{n} X_k[g(X_i x) - g(X_k x)] \quad \text{(A19)}$$

and

$$\frac{d}{dt}x = -Ax + (B - x)\sum_{k=1}^{n} f(X_k x). \quad \text{(A20)}$$

For example, if $f(w)$ is linear, namely, $f(w) = Cw$, then $g(w) = C$ and all $(d/dt)X_i = 0$ in A19. In other words, A19 can perfectly remember *any* initial pattern of reflectances. However, by A20 if $A \geq B$, then $x(t)$ approaches zero as $t \rightarrow \infty$, whereas if $B > A$, then $x(t)$ approaches $B - A$ as $t \rightarrow \infty$, whether or not a prior input pattern occurs. Thus if STM storage is ever possible, then $B > A$, and consequently noise will be amplified as vigorously as inputs. A linear signal amplifies noise, and is therefore inadequate despite its perfect memory of reflectances.

A slower-than-linear signal $f(w)$, for example, $f(w) = Cw(D + w)^{-1}$ or more generally, any $f(w)$ such that $g(w)$ is monotone decreasing, is even worse. By A19, if $X_i > X_k$, $k \neq i$, then $(d/dt)X_i < 0$ and if $X_i < X_k$, $k \neq i$, then $(d/dt)X_i > 0$. All differences in reflectances are hereby erased by the reverberation, and noise amplification also occurs. The whole network experiences a type of seizure.

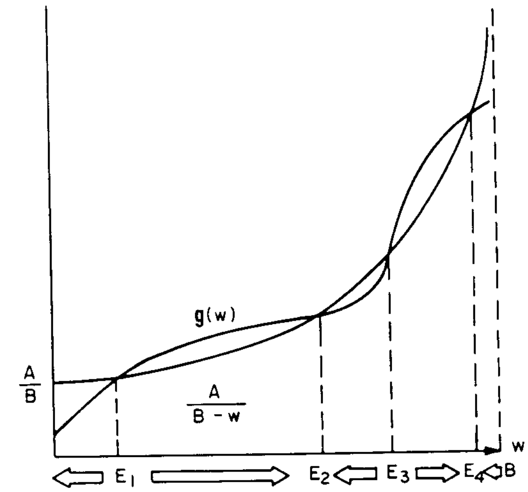

Figure A2. The even solutions E_0, E_2, ... of $g(w) = A(B - w)^{-1}$ are stable equilibrium points of $x(\infty) = \lim_{t \to \infty} x(t)$. (Since $g(w) = w^{-1}f(w)$, these points are solutions of $f(w) = Aw(B - w)^{-1}$. If $x(0) < E_1$, then $x(\infty) = 0$; thus E_1 defines the level below which $x(t)$ is treated as noise and quenched. All equilibrium points satisfy $E_i \leq B$; hence, short-term memory is normalized.)

If $f(w)$ is faster than linear, then the situation is better: for example, $f(w) = Cw^n$, $n > 1$, or more generally any $f(w)$ such that $g(w)$ is monotone increasing. In this case, if $X_i > X_k$, $k \neq i$, then $(d/dt)X_i > 0$, and if $X_i < X_k$, $k \neq i$, then $(d/dt)X_i < 0$. Consequently, this network chooses the population with the initial maximum in activity and totally suppresses activity in all other populations. This network behaves like a finite state, or binary choice machine. The same is true for total activity, since as $t \to \infty$, A20 becomes approximately

$$(d/dt)x \cong x[-A + (B - x)g(x)]. \quad \text{(A21)}$$

Thus the equilibrium points of $x(t)$ as $t \to \infty$ are $E_0 = 0$ and all the solutions of the equation

$$g(x) = A(B - x)^{-1}. \quad \text{(A22)}$$

If $g(0) < A/B$, then the smallest solution E_1 of A22 is unstable (Figure A2) so that small activities $x(t)$ are suppressed as $t \to \infty$. This is noise suppression due to recurrent competition. Every other solution E_2, E_4, ... of A22 is a stable equilibrium point of $x(t)$ as $t \to \infty$ (total activity quantization) and all equilibria are smaller than B (normalization).

The faster-than-linear signal contrast enhances the pattern so violently that the good property of noise suppression is joined to the extreme property of binary choice. This latter property is weakened by contructing a hybrid signal function that is chosen faster than linear at small activities to achieve noise suppression, but which levels off at high activities if only because all signal functions must be bounded. In the simplest case, $f(w)$ is a sigmoid, or **S**-shaped signal function. Then there exists a quenching threshold (QT). If v_is initial activity $x_i(0)$ falls below the QT, then its STM activity is quenched, or laterally masked: $x_i(\infty) = 0$. All the $x_i(0)$s that exceed the QT are contrast enhanced and stored in STM. Simultaneously, the total STM activity is normalized. Speaking intuitively, the QT exists because the faster-than-linear range starts to contrast enhance the pattern. Simultaneously, normalization shifts the activities into the intermediate linear range that stores any pattern, in particular the partially contrast-enchanced pattern. Because a QT exists, the network is a tunable filter. For example, a nonspecific arousal signal that multiplicatively inhibits all the recurrent inhibitory interneurons will lower the QT and facilitate storage of inputs in STM. Grossberg and Levine (1975) mathematically studied how such attentional shunts alter the resultant STM pattern by differentially sensitizing prescribed subfields of feature detectors that are joined together by competitive feedback interactions. The privileged subfields mask the activities in less sensitive subfields.

Such examples, either taken separately or linked together by feedback, provide insight into how interactions between continuously fluctuating quantities can sometimes generate discrete collective properties of the system as a whole. More generally, Grossberg (1978c) proves that every competitive system induces a decision scheme that can be used to globally characterize its pattern transformations as time goes on.

Appendix E

This section summarizes how the simplest transduction law realizable by a depletable chemical generates properties of antagonistic rebound due to specific cue offset and to nonspecific arousal onset when two parallel transduction pathways compete.