Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 8779–8783, November 1986
Neurobiology

From basic network principles to neural architecture: Emergence of orientation columns*

(modular self-adaptive networks/visual system/feature-analyzing cells/orientation-selective cells)

RALPH LINSKER

IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598

Communicated by Richard L. Garwin, July 7, 1986

ABSTRACT **Orientation-selective cells—cells that are selectively responsive to bars and edges at particular orientations—are a salient feature of the architecture of mammalian visual cortex. In the previous paper of this series, I showed that such cells emerge spontaneously during the development of a simple multilayered network having local but initially random feedforward connections that mature, one layer at a time, according to a simple development rule (of Hebb type). In this paper, I show that, in the presence of lateral connections between developing orientation cells, these cells self-organize into banded patterns of cells of similar orientation. These patterns are similar to the "orientation columns" found in mammalian visual cortex. No orientation preference is specified to the system at any stage, none of the basic developmental rules is specific to visual processing, and the results emerge even in the absence of visual input to the system (as has been observed in macaque monkey).**

This series of papers explores, with reference to the mammalian visual system, the structures that emerge in a network consisting of several layers of cells with connections of initially random strength, which develop according to a Hebb-type rule that "rewards" correlated activity of connected cells. In papers 1 and 2 (1, 2), I showed the emergence of spatial-opponent and orientation-selective cells in a layered system with parallel feedforward connections only and with random spontaneous uncorrelated activity (no environmental input) in the first layer.

In primate visual cortex, orientation-selective cells are organized, prior to any visual experience, into banded regions ("columns"), such that the preferred cell orientation tends to vary monotonically, but with frequent breaks and reversals, as one traverses these regions (3, 4). In this paper, we will explore the self-organization of orientation-selective cells that occurs when lateral connections between cells of the orientation-selective cell-forming layer are added to the purely feedforward network of papers 1 and 2. I will demonstrate a resulting columnar organization that agrees with the qualitative observations, and I will show why this organization is irregular (exhibits breaks and reversals in orientation sequence). The approach, and some of the early results, were described in IBM Research Report RC11642, January 1986 (R.L., unpublished). The present series of three papers is, however, self-contained.

The System Through Layer F. To summarize the state of the network through layer F, as derived in papers 1 and 2: There is random spontaneous activity in layer A. The A-to-B connections are all excitatory (1). The cells of layers C, D, E, and F are approximately circularly symmetric spatial-opponent cells (1, 2). The character of layers A–F affects layer-G development only through a function $Q^F(s)$ which describes the correlation of signaling activities of a pair of F cells as a function of the distance s between them. For the present case, I have found that $Q^F(s)$ is of "Mexican-hat" form: positive for small s, negative (implying anticorrelation of activities) for intermediate s, and near zero for large s (2).

Layer G in Absence of Lateral Connections. For this case, it was shown (2) that there is a parameter regime for which the cells of layer G mature to become "bilobed" cells, each such cell having a bar-shaped excitatory central region that extends to the periphery and is flanked by two inhibitory lobes. These cells have approximate bilateral symmetry. Each cell develops an arbitrary orientation that is independent of its neighbors' orientations.

Let us choose the same illustrative parameter values used in paper 2; namely, $n_{EG} = 0.5$, $r_G/r_F = 1.8$, $k_1 = 0.6$, $k_2 = -3$ (see paper 2 for definitions). In the limit of a large number N_G of feedforward synaptic inputs to each G cell, random variations in synaptic density (due to random synaptic placement) become arbitrarily small. The mature cell morphology can then be obtained by solving for the development of the connection-strength values, on a polar grid having a Gaussian density of sites (see Fig. 1*a* and paper 2). The number of synapses lying within the grid box represented by each site is the same for all sites in the large-N_G limit.

In paper 2, I derived an essentially unique "energy" or "objective function" corresponding to the ensemble-averaged development equation and showed that the mature states obtained by explicitly solving the development equation are the states having globally near-minimal values of this energy function. I called such states "nearly Hebb-optimal" (2) and calculated them using the method of simulated annealing (5). Fig. 1*a* shows a symmetric bilobed cell that is nearly Hebb-optimal for the parameter values given above. I shall refer to this cell, when rotated counterclockwise through angle θ, as the "standard cell" of orientation θ (for these parameter values).

Introduction of Lateral Connections. I now treat the development of the same system, except that each G cell now receives lateral inputs from a surrounding neighborhood of other G cells, as well as the feedforward inputs from cells of the predecessor layer F. Each of the lateral and feedforward connections may in general be excitatory or inhibitory and have initially random strength. The distribution of these connections exhibits no directional preference. The cell-response and development rules are the same as in papers 1 and 2 (1, 2). Because there is a new class of connections, however, the mathematical form of these rules is slightly different. (See *Appendix* for equations.)

Assume that each set of input activities from layer F [called a "presentation" (1)] persists long enough so that a G cell is still receiving a presentation from layer F at the same time that it is receiving from other G cells *their* responses to the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*This is paper no. 3 in a series. Paper no. 2 is ref. 2.

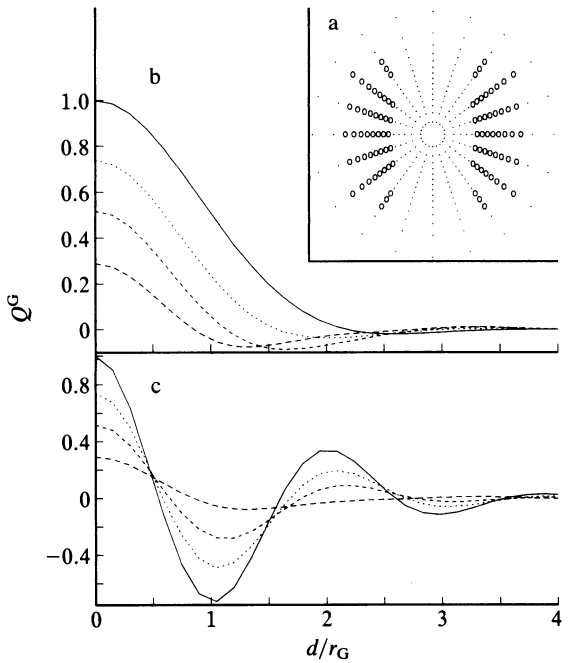

FIG. 1. (*a*) Arrangement of excitatory and inhibitory regions in the arborization of inputs to a "bilobed" G cell that is nearly Hebb-optimal, for parameters $n_{EG} = 0.5$, $r_G/r_F = 1.8$, $k_1 = 0.6$, $k_2 = -3$. Site coordinates on the polar grid shown are (r, ϕ), with $\phi = 0.1\pi n$ ($n = 1, \ldots, 20$) and with r given by $z \equiv \exp(-r^2/r_G^2)$, for 15 uniformly spaced values of z between 0 and 1 [$z = (n - 0.5)/15$, $n = 1, \ldots, 15$]. [The outermost ($n = 1$) ring has $r = 1.844r_G$.] Sites having c values at their excitatory limit are denoted by dots; at their inhibitory limit, by ovals. We refer to this as the "standard cell" for the above parameter values. (*b*) The function $Q^G(\theta, \theta', \boldsymbol{d})$ of Eq. **9** vs. distance $d = |\boldsymbol{d}|$ (in units of r_G) for $\theta = 0°$ and for vertical (north–south) displacements $\boldsymbol{d}$. Values of θ' are 0° (solid curve), 36° (dotted curve), 54° (short-dashed curve), and 90° (long-dashed curve). (*c*) Same as *b* but for horizontal (east–west) displacements $\boldsymbol{d}$.

same presentation. That is, ignore (for simplicity) possible "contamination" by input from a previous presentation.

Development of Layer G with Lateral Connections. I shall describe the system's development here with minimal reference to the mathematical details, which are provided in the *Appendix*. The development Eqs. **5–7** can be used to treat the general case in which F-to-G and G-to-G connections are developing simultaneously from initially random excitatory or inhibitory values. In this paper, however, I discuss two specific cases—the first briefly, the second in some detail.

First suppose that the F-to-G connections mature before the G-to-G connections have formed. Then each G cell becomes a bilobed cell of arbitrary orientation, and the strength of each connection from one G cell to another subsequently reaches its excitatory (or inhibitory) limit, depending upon whether the correlation between the activities of the two G cells is greater (or less) than a parameter-specified value. The G-to-G connections thus play a passive developmental role in this case.

The second case we consider is more interesting. Suppose that the G-to-G connections mature to become all-excitatory before the F-to-G connections have developed. Then any given G cell will affect its neighbors' development—through the Hebb-type rule—in such a way as to tend to increase the correlations between the activity of the given G cell and the activities of the neighboring G cells to which it provides excitatory input via the lateral connections. (See *Appendix* following Eq. **9**.)

The direct solution of Eqs. **5–7**, for an assembly of several thousand G cells (to reveal columnar organization) with at least several hundred inputs to each cell, is possible in principle but is computationally very demanding. Instead, I shall show how to calculate the nearly Hebb-optimal configurations for such an assembly of cells.

METHODS

We first consider the system in the large-N_G limit. Random deviations from rotationally symmetric synaptic density (due to random synaptic placement) go to zero in this limit. Therefore, the energy function (E_n in paper 2) for a G cell indexed by n, in the absence of lateral connections (2), is rotationally symmetric. [This does not mean that the lowest-energy states are rotationally symmetric (2).] In particular, the energy of a bilobed cell is independent of cell-axis orientation.

The energy of a G cell varies with its morphology (the arrangement of excitatory and inhibitory regions in its input arborization). We focus here on the parameter regime for which approximately bilaterally symmetric bilobed cells are of globally near-minimal energy. Cells of different morphology [substantially asymmetric bilobed cells, trilobed cells (2), etc.] have energies that are greater than the global minimum by at least an amount of an order that we denote ΔE_{morph}.

In the presence of lateral connections, the energy of a G cell acquires an additional, lateral-interaction component (see *Appendix*). This component is proportional to κ, where κ describes the relative contribution (e.g., the relative total number) of lateral vs. feedforward connections. The value of this lateral-interaction energy *does* vary with cell orientation. We are free to choose κ small enough (e.g., lateral connections sparse enough) so that the lateral-interaction energy is small compared with ΔE_{morph}. Then the lateral interaction will cause a particular orientation of bilobed cell to be favored at a given site (depending upon the orientations of the cell's neighbors) but will not cause non-bilobed cells to become nearly Hebb-optimal (their energies will still lie above the global minimum by at least ΔE_{morph}). Thus a nearly Hebb-optimal state of the entire assembly of G cells—a state of near-minimal E value (see *Appendix*)—will be composed of mature bilobed cells having a structured arrangement of orientation preferences that we shall calculate explicitly.

In terms familiar from theoretical physics, the unperturbed ($\kappa = 0$) energy function for the assembly of cells [$E^{(0)}$ in the *Appendix*] is invariant under rotation of each cell (about its axis) through an arbitrary angle chosen independently for each cell. This invariance is broken by the lateral-interaction terms. The latter terms do not, however, mix bilobed states (of near-minimal energy) with states having different morphology, provided the perturbation is sufficiently weak (κ small).

It suffices, therefore, to calculate nearly Hebb-optimal states of an assembly of cells, each of which is a mature bilobed cell of arbitrary orientation. (Cells having other morphologies need not be considered.) To make this problem tractable, we adopt the following *Ansatz*. (*i*) Represent layer G by using a square grid. An arbitrary configuration is characterized by assigning a value of orientation angle to each grid box. All cells within a given box share the same orientation. (We shall see that each iso-orientation region indeed comprises many lattice sites, justifying the assignment of a single orientation to the cells at each site.) For practical computational reasons, allow the orientation at each grid box to take on any of ten values, spaced every 18°. (*ii*) The lateral interaction between any pair of cells is computed as if both cells were "standard cells" of the appropriate orientations, located at their appropriate grid sites. (In other words, we approximate the actual Q^G value, between two cells that are each approximately bilaterally symmetric bilobed cells similar to that represented in Fig. 1*a*, by the value for two cells that are each exactly bilaterally symmetric "standard cells."

Analysis of the limits of validity of this approximation is beyond the scope of this paper.)

The energy to be near-minimized (over all sets of grid-site orientation assignments) is then E' of Eq. **8**.

Before proceeding with this analysis, I comment on the large-N_G limit invoked above. As N_G is decreased, random variations in synaptic density give rise to a random rotationally asymmetric component of the energy function for each cell. When N_G is too small, the orientation preference for each cell is determined by this random component of the energy, rather than by the lateral interaction terms (which are of order κ). Let us consider here the regime in which N_G is sufficiently large that this does not occur.

RESULTS

The Q^G Function for a Pair of Mature Bilobed Cells. To avoid possible confusion, I emphasize that we are considering the case in which the excitatory lateral connections are formed before the feedforward connection strengths develop. The actual development process depends upon the interactions (i.e., the values of the Q^G function) between pairs of G cells as they mature. But by transforming the development problem into a global near-optimization problem, we avoid having to compute the details of the maturation process. The only functions that need to be computed in order to find the nearly Hebb-optimal states of the system are the Q^G values between mature pairs of bilobed cells. This enormously reduces the computational burden, at the cost of introducing the above *Ansatz*.

The correlation between the activities of two "standard cells"—mature cells of the type shown in Fig. 1*a*—depends upon their orientations θ and θ' and the displacement $\boldsymbol{d}$ of the second cell position relative to the first. (We measure θ counterclockwise from the vertical or y axis.) This correlation function $Q^G(\theta, \theta', \boldsymbol{d})$ was calculated by use of Eq. **9** and plotted in Fig. 1 *b* and *c* for $\theta = 0°$ and for several values of θ', as a function of distance $d = |\boldsymbol{d}|$ along the y axis (Fig. 1*b*) or the x axis (Fig. 1*c*). Rotating θ, θ', and the $\boldsymbol{d}$ vector through the same angle leaves Q^G unchanged.

Fig. 1 *b* and *c* shows that to maximize the activity correlation Q^G, a given "standard cell" X having a vertical (north–south) axis ($\theta = 0°$) will favor its northern and southern neighbors (out to $d = 2.3r_G$) and its eastern and western near-neighbors ($d \leq 0.5r_G$) to have $\theta' = 0°$ and its eastern and western midrange neighbors ($0.5r_G \leq d \leq 1.5r_G$) to have $\theta' = 90°$.

Why does this happen? Consider the type of input presentation that maximally stimulates cell X: it has (in the vicinity of X) a vertical band of high activity, centered on X and flanked by low-activity regions that overlie the inhibitory lobes of cell X. The cells north and south of X are centered on the high-activity band (for this presentation), and their activity will be correlated maximally with that of X if they have $\theta' = 0°$. The cells east and west of X (at intermediate distance) are centered on a low-activity flank. If they had $\theta' = 0°$, their activities would be anticorrelated with X. They are therefore favored to be as unresponsive as possible to this type of presentation and, hence, to have $\theta' = 90°$. This shows heuristically why the favored axis orientation choices implied by Fig. 1 *b* and *c* arise. We emphasize that Fig. 1 *b* and *c* gives the correct Q^G—which is based on an average over an entire ensemble of presentations of random layer-A activity to the system—and not just the result for a particular presentation.

But notice that the set of neighbor orientations favored by our given cell X conflicts with the set favored by its midrange neighbor to the east (or west): the latter is oriented at 90° and favors *its* eastern and western near and midrange neighbors (including cell X) to be oriented also at 90°, not at 0°. That is, Fig. 1 *b* and *c* shows that any given cell "wants" to be at the center of a band of like-oriented cells, flanked by perpendicularly oriented cells, with the long axis of the band being aligned with the cell's own orientation.

Computation of Hebb-Optimal States. To see how this contention is resolved, we explicitly compute arrangements of G-cell orientations that are nearly Hebb-optimal; i.e., that globally near-minimize the objective function E' (Eq. **8**), which is the arrangement-dependent portion of the total E. We use the technique of simulated annealing (5). We start with a random assignment of site orientations and carry out a series of passes, each pass at a "temperature" T (5). During each pass (*i*) each site $\boldsymbol{x}$ is considered in random sequence, (*ii*) the energy $E'(\theta)$ of the configuration is computed for each of the ten possible orientation assignments θ at site $\boldsymbol{x}$, and (*iii*) the new orientation at site $\boldsymbol{x}$ is assigned to be θ_{new} with probability proportional to $\exp[-E'(\theta_{new})/T]$. The T value is chosen to be large (compared with the differences between E' values for different values of θ) initially and then is gradually reduced until it reaches zero in the final passes.

Fig. 2 illustrates a nearly Hebb-optimal assembly of "standard cells." We find that cells of similar orientation are organized into band-like regions. An "electrode" passing tangentially across this layer will often, though not always, measure a progression of orientation angles that is generally monotonic but has breaks and reversals.

I have intentionally chosen d_0 (see Fig. 2 legend) to lie in a regime such that the midrange region ($0.5r_G < d < 1.5r_G$) in Fig. 1*c* contributes to E'. If d_0 is too small ($<<0.5r_G$, for example), a solution with all G-cell orientations identical would maximize Q^G for all pairs of G cells and would therefore be favored.

Vortices and Fractures. To characterize further the qualitative structure of Fig. 2, let us define a positive (resp., negative) "half-vortex" as a point such that, as one traces a small clockwise circular path around the point, the orienta-

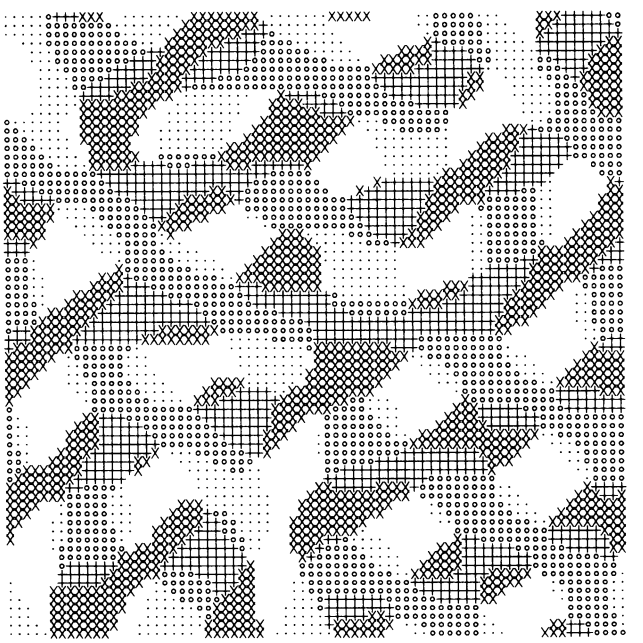

FIG. 2. A nearly Hebb-optimal solution for an array of orientation cells. Each cell is "stained" according to its orientation preference θ (measured counterclockwise from the vertical): 9–45° (dot); 45–81° (small circle); 81–117° (+); 117–153° (×); 153–189° (blank). Adjacent grid positions are separated by distance $0.1493r_G$. The function $\rho(d)$ (see Eq. **8**) is taken proportional to $\exp(-d^2/d_0^2)$ for $d_0 = 1.194r_G$ (= 8 grid positions). Array is 72 by 72 with periodic boundary conditions. [Qualitatively similar patterns result if $\rho(d)$ is constant out to, for example, 11 grid positions and zero beyond that, or if an 80-by-80 array is used.]

tion angle at the sites lying on this path rotates clockwise (resp., counterclockwise) by 180°. Fig. 2 contains many half-vortices, which are in general located wherever four or all five types of orientation marker symbols meet near a single point. In addition, as one proceeds around the boundary of an iso-orientation region, the half-vortices that one encounters are generally of alternating (positive and negative) sign. There are also occasional "fracture" lines across which the orientations at two adjacent sites differ substantially. In Fig. 2 a fracture is indicated by the abutment of two regions whose marker symbols (e.g., dot and +) correspond to orientation angles differing by at least 36°. A break in orientation sequence is recorded when an electrode track passes through such a fracture.

DISCUSSION

The Hebb-optimal arrangements of orientation cells in layer G, in the presence of excitatory G-to-G connections, consist of bands of cells of the same or similar orientation. The generally monotonic progression of cell orientation—with frequent breaks and reversals—as one traverses the layer, agrees qualitatively with the observations in primates (3, 4).

Blasdel and Salama (8) have recently presented an analysis of experimental data that leads them to a picture of orientation columns in macaque monkey (see figures 5*a* and 6*a* of ref. 8). These figures are strikingly similar to the theoretical results presented here (Fig. 2) and in IBM Research Report RC11642, January 1986 (R.L., unpublished). In particular, the features I have identified as half-vortices of alternating positive and negative sign (see previous section), as well as fracture lines (figure 6*c* of ref. 8), are prominent features of their orientation maps.

How Excitatory Lateral Connections Yield "Columnar" Organization. It may be surprising that *excitatory* lateral connections give the progression of preferred orientations that I have demonstrated. One tends to think of a "Mexican-hat" form of interaction within the developing layer—excitatory at short range and inhibitory at midrange—as inducing the assignment of different "labels" (e.g., orientations) to different groups of cells (6). What I have shown is that like-orientation G cells, displaced by an intermediate distance in the direction perpendicular to their orientation axes, have anticorrelated firing activities. (This is because Q^F is of Mexican-hat form, which in turn is because layers C–F are composed of opponent cells.) An excitatory G-to-G cell interaction tends to maximize activity correlations and, hence, to prevent two such G cells from having parallel orientations. At the "label" level of discussion, there thus appears to be a kind of "lateral inhibition" between like labels at intermediate distance, but this results from an excitatory interaction at the level of the physical G-to-G connections in our system. This by no means implies that inhibitory feedback or lateral connections have no important role in developing systems.

Comparison with an Earlier Model of Orientation "Column" Formation. Swindale (6) studied the problem in which one is given an assembly of cells that are to be assigned a label representing orientation preference and an ad hoc interaction that favors like label assignment to near-neighbor cells and unlike assignment to pairs of cells having intermediate separation, regardless of direction. No mechanism for such an interaction was implied, nor was the question addressed of how orientation cells might develop. Cells having similar labels were found to become organized into bands, with these bands themselves lying at a variety of apparently random orientations.

For comparison with the results of ref. 6, I have repeated my analysis using, instead of our correct Q^G function, a function $Q^{iso}(\Delta\theta,d)$, which is defined as the average of $Q^G(\theta, \theta', d)$ over all θ and θ', keeping $\Delta\theta = \theta - \theta'$ constant. This Q^{iso} has a Mexican-hat form and is isotropic (i.e., independent of the direction of the displacement d). The nearly Hebb-optimal states for Eq. **8** using Q^{iso} consist of regular, generally parallel iso-orientation bands, with a generally monotonic progression of orientation as one moves transverse to the bands. The overall orientation of the parallel-banded pattern is arbitrary. I find much greater parallelism of adjacent bands than is shown in ref. 6, probably because the results of ref. 6 were generated by a method that does not ensure Hebb-optimality.

The correct anisotropic Q^G function for our system leads to Hebb-optimal states having band patterns (as in Fig. 2) that are not disposed in a parallel fashion, in contrast to those generated by Q^{iso}. This is because the Q^G function favors the formation of vertically elongated regions of vertical-orientation cells, horizontally elongated regions of horizontal-orientation cells, etc. Such preferences mutually conflict. The Q^{iso} function, on the other hand, favors the formation of parallel bands (in an arbitrary direction) for regions of all labels.

The experimentally observed degree of regularity and parallelism of orientation bands varies from an irregular arrangement (with frequent breaks and reversals of orientation sequence) in macaque (3, 4, 8) to a more regular arrangement of locally parallel bands in tree shrew (7). Since I have not explored the parameter space for orientation column formation in detail, I am not in a position to place limits on the degree of band irregularity to be expected in general in a modular self-adaptive network.

APPENDIX

Let c_{nu} denote the feedforward-connection strength from cell u in layer F to cell n in layer G, and let f_{nm} denote the lateral-connection strength from cell m in G to cell n in G. There can be more than one connection between two given cells; all such connections are summed over. For ease of notation, suppress the additional index that would distinguish such connections from one another. Index pairs u,n or m,n not corresponding to any connection are understood to be omitted from the sums below. [The equations can easily be rewritten in terms of the more exact "pre(ni)" notation of paper 2, but this adds possibly confusing detail that I wish to avoid here.]

Eqs. **1** and **2** give the linear-summation rule for the postsynaptic activity $F_n^{G\pi}$ of the G cell labeled n, to first order [hence the superscript "(1)"] in the lateral (G-to-G) interaction, in terms of the input activities from F and G cells. Eqs. **3** and **4** give the Hebb-type modification rule (1) for the c and f connections, to lowest required order in the lateral interaction for each case.

$$F_n^{G\pi(1)} = R_a + R_b \times (\Sigma_u c_{nu} F_u^{F\pi} + \Sigma_m f_{nm} F_m^{G\pi(0)}) \quad [1]$$

$$F_m^{G\pi(0)} = R_a + R_b \times \Sigma_u c_{mu} F_u^{F\pi} \quad [2]$$

$$(\Delta c_{nu})^\pi = k_a + k_b \times (F_n^{G\pi(1)} - F_0^G) \times (F_u^{F\pi} - F_0^F) \quad [3]$$

$$(\Delta f_{nm})^\pi = k_c + k_d \times (F_n^{G\pi(0)} - F_1^G) \times (F_m^{G\pi(0)} - F_2^G). \quad [4]$$

Here $R_{a,b}$, $k_{a,b,c,d}$, F_0^F, and $F_{0,1,2}^G$ are constants (R_b, k_b, $k_d > 0$). The "presentation," or set of signaling activities at a given time, is indexed by π.

By ensemble-averaging Eqs. **3** and **4**, and using Eqs. **1** and **2**, we obtain the development Eqs. **5** and **6** for the time

derivatives of c_{nu} and f_{nm}, in terms of the Q^F function defined by Eq. 7.

$$\dot{c}_{nu} = k_1 + \frac{1}{N_G}\Sigma_v(Q^F_{uv} + k_2)c_{nv} + R_b\Sigma_m f_{nm}\left[k_{1a} + \frac{1}{N_G}\Sigma_v(Q^F_{uv} + k_2)c_{mv}\right] \quad [5]$$

$$\dot{f}_{nm} = k_3[k_4 + k_5\Sigma_u c_{nu} + k_6\Sigma_u c_{mu} + \Sigma_u\Sigma_v(Q^F_{uv} + k_7)c_{nu}c_{mv}] \quad [6]$$

$$Q^F_{uv} \propto \langle(F^{F\pi}_u - \overline{F}^F) \times (F^{F\pi}_v - \overline{F}^F)\rangle_\pi. \quad [7]$$

Angle brackets denote the ensemble average, k_1–k_7 are simple functions of the constants introduced in Eqs. **1–4** (see paper 1 for the explicit definitions of k_1 and k_2 in the absence of lateral interactions), N_G is the number of c inputs to each G cell, and $\overline{F}^F$ is the ensemble-averaged activity at any point in layer F (1). Here Q^F_{uv} is a function, $Q^F(s)$, only of the distance s between u and v (2) and is normalized to $Q^F(0) = 1$. We use Eqs. **5** and **6** except when a c or f value reaches one of its saturation limits (1), in which case c or f is "pinned" at that limit for that time step.

From Eq. **1**, we see that the relative contribution (to G-cell response) of the lateral (f) vs. the feedforward (c) connections is proportional to the relative numbers of f and c connections. We characterize the perturbation expansion used in Eqs. **1–6** as an expansion in a small parameter denoted by κ, whose value can be controlled (for example) by adjusting the relative numbers of f and c connections.

Special Case of All-Excitatory f Values. Consider an assembly of G cells distributed with uniform density in layer G, and with the number $\rho(d)$ of excitatory ($f = 1$) connections between any two G cells depending only upon the distance d between the cells. [This latter condition is unnecessarily restrictive. For the *Ansatz* used here, only the number of connections between pairs of grid boxes is relevant to the function (Eq. **8**) we shall need to optimize. The random variations in the number of box-to-box connections are much smaller than the variations in the number of connections between pairs of individual cells, if there are many cells per box.]

Construct an objective function E of all the $\{c_{nu}\}$ in a manner similar to that used in paper 2 (but here comprising all c for an entire assembly of G cells), having the property that $\dot{c}_{nu} = -N_G\,\partial E/\partial c_{nu}$. This property means that the system, which develops according to Eq. **5**, is always following the path of locally steepest (gradient) descent of the E function. To obtain E (full result omitted to save space), multiply each term on the right-hand side of Eq. **5** by $(-1/N_G)c_{nu}$, then multiply terms that are quadratic in the c variables by 1/2, then sum over n and u.

The resulting E is the sum of two parts, $E^{(0)} + E^{(1)}$. The dominant part (since the contribution of lateral connections is taken to be weak), $E^{(0)}$, is just the sum over all G cells of the E_n function of paper 2 for each G cell n in the absence of lateral connections. Every arrangement of orientations (of cells of given morphology) has the same globally near-minimal value of $E^{(0)}$ (apart from random variations in synaptic density, which are arbitrarily small for N_G large). We can therefore ignore this term when we calculate Hebb-optimal arrangements of orientations. The $E^{(1)}$ part results from the lateral interaction; its orientation-dependent portion for our *Ansatz* is proportional to

$$E' = -\Sigma_x\Sigma_{x'}\rho(|x - x'|)Q^G(\theta_x, \theta_{x'}, x' - x), \quad [8]$$

where the sum is over all pairs of sites x and x' on a square lattice, and θ_x is the orientation of the cells at site x. In our calculations, we take $\rho(d)$ to be decreasing with d, either gradually (as a Gaussian) or abruptly (constant out to some distance, and zero beyond). Here

$$Q^G(\theta, \theta', d) \propto \Sigma_i\Sigma_j Q^F(|d + t_j - t_i|)c^\theta_i c^{\theta'}_j, \quad [9]$$

where i indexes the F-to-G connections of a standard cell (see Fig. 1*a*) whose axis is oriented at angle θ measured counterclockwise from the vertical; j indexes the F-to-G connections of a standard cell that is displaced d relative to the first cell and has orientation θ'; t_i and t_j are the locations of connections i and j relative to the center of their respective bilobed cells; each c value is +0.5 or −0.5, depending upon whether the corresponding connection lies in an excitatory or inhibitory domain of Fig. 1*a*; and we choose the normalization $Q^G(\theta, \theta, 0) = 1$. The Q^G function is in fact proportional to the autocorrelation function of the activity of a pair of G cells, if one computes only the leading-order terms (i.e., those that are independent of the lateral connections). To see this, replace superscripts "F" by "G" in Eq. **7** and use Eq. **2** (see also paper 2).

That same calculation shows (2) that for two arbitrary G cells at *any* stage of development (not just for two standard cells), we have $Q^G_{nm} \propto \Sigma_u\Sigma_v Q^F_{uv}c_{nu}c_{mv}$. Also, Eq. **5** shows that the $f_{nm}Q^F_{uv}c_{mv}$ term tends to drive c_{nu} in the direction of $Q^F_{uv}c_{mv}$ for fixed positive f_{nm}, during c-value development. This explains why each G cell tends to affect its neighbors' development in such a way as to increase Q^G_{nm}, as stated in the text.

I thank Dr. D. C. Van Essen for detailed comments on the three papers in this series and Drs. N. Fisch, W. D. Knowles, and D. H. Weingarten for their comments on an earlier version of these manuscripts.

1. Linsker, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 7508–7512.
2. Linsker, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8390–8394.
3. Wiesel, T. N. & Hubel, D. H. (1974) *J. Comp. Neurol.* **158,** 307–318.
4. Hubel, D. H., Wiesel, T. N. & Stryker, M. P. (1978) *J. Comp. Neurol.* **177,** 361–380.
5. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983) *Science* **220,** 671–680.
6. Swindale, N. V. (1982) *Proc. R. Soc. London Ser. B* **215,** 211–230.
7. Humphrey, A. L., Skeen, L. C. & Norton, T. T. (1980) *J. Comp. Neurol.* **192,** 549–566.
8. Blasdel, G. G. & Salama, G. (1986) *Nature (London)* **321,** 579–585.