

# On Learning of Spatiotemporal Patterns by Networks with Ordered Sensory and Motor Components

## 1. Excitatory Components of the Cerebellum

*By Stephen Grossberg*

### 1. Introduction

Many of our sensory and motor organs have linearly ordered components, for example the fingers on a hand, the tonotopic organization of the auditory system, the successive joints on arms and legs, the spine, etc. This paper begins a discussion of some nonlinear networks which can learn complicated spatiotemporal patterns among sensory and motor organs with linearly ordered components. These networks will ultimately resemble cerebrocerebellar systems of higher vertebrates, and can be picturesquely interpreted as an interaction via idealized "subcortical nuclei" of portions of "cerebral cortex" with "neocerebellum". To the extent that this analogy is valid, various geometrical (anatomical) and dynamical (physiological) details of the networks can be interpreted as provisions by cerebrocerebellar systems for effective learning of spatiotemporal patterns. For example, we can interpret geometrical statements concerning excitatory "cerebellar" network components to include:

- (a) mossy fiber inputs widely dispersed in a folium and ending in rosettes,
- (b) climbing fiber inputs localized among one or a small cluster of Purkinje cells, and climbing up the Purkinje cell dendritic bush,
- (c) perpendicular mossy fiber and climbing fiber somatotopic representations,
- (d) long bidirectional parallel fibers receiving inputs from mossy fiber rosettes, and delivering outputs via dendritic spines to all Purkinje dendrites by which they pass,
- (e) a segregation of parallel fibers into layers according to length with the longest fibers in the deepest layers,
- (f) fastest signal velocities in the longest parallel fibers,
- (g) control of cell size and orientation by the spatial distribution of a cell's averaged inputs via dipole forces.

A forthcoming paper on inhibitory connections will discuss network analogs of basket cells and Golgi cells.

Among "cerebral" facts, we can interpret network properties to include:

- (h) somatotopically organized primary and sensory representations separated by a central fissure,

- (i) sheets of sensory and motor cortex controlling ever higher associations as the distance from the central fissure increases,
- (j) multiple sensory and motor representations.

A later paper will discuss analogs of the laminar structure of neocortex, including analogs of pyramid cells, cells of Martinotti, stellate cells, and horizontal fibers.

We also find formal phenomena that resemble the following psychological facts:

- (k) spatiotemporal masking,
- (l) spatiotemporal consolidation,
- (m) responses controlled by temporally weighted mixtures of simple and compound associations,
- (n) a connection between input energy, signal velocity, learning speed, and behavioral reaction times,
- (o) the impossibility of spatially localizing the cells controlling complicated spatiotemporal patterns.

One of the properties needed in our construction is that of a *spatiotemporally self-similar* cell type. As interpreted, this property helps to construct behaviorally useful ensembles of cells from individual cells whose only concern is to maintain adaptive metabolic responses to the environmental signals which they receive. Compatible with the properties of spatiotemporally self-similar cells are

- (p) a cell nucleus which can set correct production levels for the entire cell based on cell body membrane excitation, and
- (q) thus cell size can increase in response to intense average excitatory inputs,
- (r) a signal velocity in axons proportional to axon length, and
- (s) thus ensembles of cells with given axonal connections can be stretched or shrunk in a growing brain or in an evolving brain wherein new cell types are emerging without destroying extant intercellular time lags,
- (t) unbiased learning conditions often arise when the spatial distribution of cells is in equilibrium relative to the spatial distribution of inputs and outputs.

A crucial assumption of the construction is that nerve cells capable of learning are "chemical dipoles". This assumption is briefly reviewed in Section 2. The dipole property underlies such phenomena as (c), (g), (n), and (p) above, and also is compatible with the variability in size of presynaptic endbulbs due to postsynaptic trophic effects.

The above results follow plausibly from behavioral principles which have been used to derive a new theory of learning called the theory of *embedding fields*. The theory is briefly reviewed in Section 2. Given its principles and equations, our present task can be stated as follows. Suppose that the input sources and output sinks of a system are arranged in a row. What is the simplest arrangement of embedding field components between the input sources and output sinks that can learn spatiotemporal patterns up to a given complexity among their inputs and outputs? Speaking mathematically, this is a boundary value problem in function space.

Even supposing that our equations qualitatively summarize important features of neural learning, our networks need not in all details agree with neural networks which carry out similar tasks. This is because new neural structures have gradually been superimposed on older structures which, in turn, have developed in response to complicated physical influences. The boundary value problems that we pose are

therefore bound to be vast simplifications of the analogous neural situations. The fact that recognizable anatomical patterns emerge from these simplified formulations suggests that the forces guiding neural evolutionary development to a remarkable degree embody a teleology of learning that can overcome irrelevant environmental trends. Such physical constraints as difficulty of transporting chemicals, difficulty of connecting widely separated areas, difficulty of constructing unbiased systems over more primitive nearest-neighbor or uniformly connected systems will not in all cases be negligible. Then a realistic anatomical pattern of cells will only arise if the boundary value problem includes these constraints.

In fact, a network has been constructed that can learn any number of essentially arbitrarily complicated space-time patterns [1], namely series of  $\Gamma$ -outstar *avalanches*. Moreover, this network uses very few components. Thus the problem solved by realistic anatomical patterns of cells is not merely one of learning complicated tasks. Rather the problem is one of "subtlety"—being able to perform these tasks adaptively in response to subtly fluctuating environmental demands, including sensory feedback due to prior motor outputs.

## 2. A brief review

The networks to be constructed obey the laws set down in [2], [3], and [4], which are summarized in [5]. Our point of departure is the following special case of equations from [5], whose intuitive meaning only will be needed to draw our qualitative conclusions. For  $i, j, k = 1, 2, \dots, n$ ,

$$\dot{x}_i(t) = \alpha_i[P_i - x_i(t)] + J_i^+(t) - J_i^-(t) + I_i(t), \quad (1)$$

and

$$\dot{z}_{jk}(t) = u_{jk}[Q_{jk} - z_{jk}(t)] + v_{jk}^+[x_j(t - \tau_{jk}^+) - \Gamma_{jk}^+]^+[x_k(t) - \Lambda_k^+]^+, \quad (2)$$

where

$$J_i^+(t) = \sum_{m=1}^n \beta_m^+[x_m(t - \tau_{mi}^+) - \Gamma_{mi}^+]^+ p_{mi}^+[z_{mi}(t) - \Omega_{mi}^+]^+ + \sum_{m=1}^n \beta_m^+[x_m(t - \tau_{mi}^+) - \Gamma_{mi}^+]^+ q_{mi}^+, \quad (3)$$

$$J_i^-(t) = \sum_{m=1}^n \beta_m^-[x_m(t - \tau_{mi}^-) - \Gamma_{mi}^-]^+ q_{mi}^-, \quad (4)$$

and the notation  $[w]^+$  denotes

$$[w]^+ = \max(w, 0), \quad (5)$$

and describes various thresholds.

$x_i(t)$  is a process fluctuating in time within a vertex (or "cell body cluster")  $v_i$  of a network  $\mathcal{M}$ , and each  $z_{jk}(t)$  is a process fluctuating in time within the arrowhead (or "endbulb cluster")  $N_{jk}$  of the directed edge (or "axon cluster")  $e_{jk}$  from  $v_j$  to  $v_k$ . See Figure 1.



Figure 1.

$I_j(t)$  is the experimental input reaching  $v_j$  at time  $t$ . This input perturbs the "average membrane potential"  $x_j(t)$  of  $v_j$ , which is transformed into a spiking frequency that travels to every vertex  $v_k$  for which at least one of the nonnegative path weights  $p_{jk}^+$ , or  $q_{jk}^+$ , or  $q_{jk}^-$  is positive. If  $q_{jk}^- > 0$ , an inhibitory signal of size

$$\beta_j^- [x_j(t) - \Gamma_{jk}^-]^+ q_{jk}^- \quad (6)$$

leaves  $v_j$  at time  $t$  and arrives at  $v_k$  at time  $t + \tau_{jk}^-$ , thereby perturbing  $x_k$ . If  $q_{jk}^+ > 0$ , an excitatory signal of size

$$\beta_j^+ [x_j(t) - \Gamma_{jk}^+]^+ q_{jk}^+ \quad (7)$$

leaves  $v_j$  at time  $t$  and arrives at  $v_k$  at time  $t + \tau_{jk}^+$ , thereby perturbing  $x_k$ . If  $p_{jk}^+ > 0$ , an excitatory signal of size

$$\beta_j^+ [x_j(t) - \Gamma_{jk}^+]^+ p_{jk}^+ \quad (8)$$

leaves  $v_j$  at time  $t$  and arrives at  $N_{jk}$  at time  $t + \tau_{jk}^+$ , thereupon activating the transmitter control process  $z_{jk}(t)$  in  $N_{jk}$  and releasing transmitter at a rate

$$\beta_j^+ [x_j(t) - \Gamma_{jk}^+]^+ p_{jk}^+ [z_{jk}(t + \tau_{jk}^+) - \Omega_{jk}^+]^+,$$

thereby perturbing  $x_k$ . Only one of these three signals can ever be positive, since we choose

$$p_{jk}^+ q_{jk}^+ + p_{jk}^+ q_{jk}^- + q_{jk}^+ q_{jk}^- = 0.$$

The signals from different  $v_j$  combine independently at  $v_k$ , as (1), (3), and (4) show.

All learning within  $\mathcal{M}$  occurs at the arrowheads  $N_{jk}$  for which  $p_{jk}^+ > 0$ , since here  $z_{jk}(t)$  cross-correlates the excitatory signal

$$\beta_j^+ [x_j(t - \tau_{jk}^+) - \Gamma_{jk}^+]^+ p_{jk}^+ \quad (9)$$

received by  $N_{jk}$  from  $v_j$  at time  $t$  with the value

$$[x_k(t) - \Lambda_k^+]^+ \quad (10)$$

of the contiguous vertex  $v_k$  at this time, as (2) shows. In particular,  $v_{jk}^+ = 0$  only if  $p_{jk}^+ = 0$ .

A bounded version of these equations is also discussed in [5].

Given these equations, our task can be stated as follows. Let an experimentalist  $\mathcal{E}$  deliver inputs to a subset  $\mathcal{S}$  of  $\mathcal{M}$ 's vertices and receive outputs from a subset  $\mathcal{O}$  of  $\mathcal{M}$ 's vertices. Let  $\mathcal{S}$  and  $\mathcal{O}$  be ordered sets. What is the simplest geometrical arrangement of vertices and edges joining  $\mathcal{S}$  and  $\mathcal{O}$  that is capable of learning prescribed patterns of outputs from  $\mathcal{O}$  to prescribed patterns of inputs to  $\mathcal{S}$ ? That is, how should the path weights  $p_{jk}^+$ ,  $q_{jk}^+$ , and  $q_{jk}^-$ ; the time lags  $\tau_{jk}^+$  and  $\tau_{jk}^-$ ; and the other parameters of (1)–(4) be chosen? Before determining exact numerical values of these parameters, one must study how the ordering of vertices in  $\mathcal{S}$  and  $\mathcal{O}$  propagates to orderings among interpolated vertices and edges. This information will already suggest recognizable cell patterns. It will also point out that some

vertices can profitably be stretched out into complicated cell body shapes with dendrites, but these vertex deformations will be mentioned only as they are needed.

### 3. The input and output vertices are embedded in lines

To motivate the construction of our network, we will often pose our problem as follows. How do our fingers learn to move in patterns? Once our construction is completed, "fingers" can be replaced by any ordered system of components.

Conscious control can normally be exerted over the motions of each of our fingers (or finger joints) as a whole, rather than over individual finger muscles. Hence we suppose initially—and can later readily modify this assumption—that a single vertex  $v_i$  controls downward motions of the  $i$ th finger  $f_i$ ,  $i = 1, 2, 3, 4, 5$ . That is,  $v_i$  sends an excitatory edge  $e_i$  to  $f_i$  over which signals from  $v_i$  to  $f_i$  travel. To fix ideas, let  $f_i$  move downward whenever the signal received by  $f_i$  from  $v_i$  is positive (and  $f_i$  is not already completely depressed) at a velocity that increases with the size of the signal at a given finger position. Since the fingers  $f_i$  are linearly ordered, the vertices  $v_i$  that control them are also linearly ordered. For simplicity, the output (or motor) vertices  $\mathcal{O} = \{v_1, v_2, \dots, v_5\}$  will be arranged along a line  $L(\mathcal{O})$  in  $\mathcal{M}$ . Similarly, a set  $\mathcal{I} = \{V_1, V_2, \dots, V_5\}$  of input (or sensory) vertices exists, also for simplicity arranged along a line  $L(\mathcal{I})$  in  $\mathcal{M}$ , such that  $V_i$  receives excitatory inputs from  $f_i$ . See Figure 2.

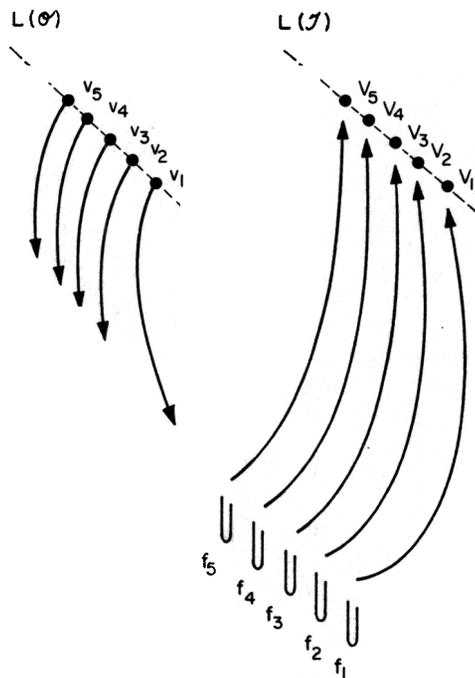


Figure 2.

#### 4. The control vertices are embedded in planes

$\mathcal{M}$  is being constructed to learn patterns of finger motion in which more than one finger participates. Thus there exist vertices and/or edges in  $\mathcal{M}$  that control the *order* in which outputs from  $\mathcal{O}$  become positive in response to particular orderings of inputs to  $\mathcal{I}$ .

The simplest way to connect inputs and outputs is to draw excitatory edges directly from  $\mathcal{I}$  to  $\mathcal{O}$ . Each connection  $V_i \rightarrow v_j$  then represents a learnable transition. Very few different patterns can simultaneously be remembered by such a network ([6], [7]), and the network's ability to make subtle temporal discriminations is unimpressive. These deficiencies can be remedied by interpolating new vertices between  $\mathcal{I}$  and  $\mathcal{O}$ , and thereby making the input-output interaction less direct. Since these vertices will control the ordering of outputs in response to a particular ordering of inputs, they will be spatially distributed in a way that clearly distinguished the orderings of  $\mathcal{I}$  and  $\mathcal{O}$ . This is simply done by placing these control vertices in rows adjacent to  $\mathcal{I}$  and  $\mathcal{O}$ , and letting the number of rows increase as the complexity of the pattern to be controlled increases. In other words, the control vertices can conveniently be embedded in *planar surfaces*  $S(\mathcal{I})$  and  $S(\mathcal{O})$  with boundaries at  $L(\mathcal{I})$  and  $L(\mathcal{O})$ , respectively.  $S(\mathcal{I})$  and  $S(\mathcal{O})$  collectively will be denoted by  $S$ . The input vertex  $V_i$  will interact—perhaps only indirectly—with the entire strip of control vertices in  $S(\mathcal{I})$  perpendicular to  $L(\mathcal{I})$  at  $V_i$ . Similarly,  $v_j$  will interact with the strip of  $S(\mathcal{O})$  vertices perpendicular to  $L(\mathcal{O})$  at  $v_j$ . Using  $S$ , the space within  $\mathcal{M}$  representing each  $f_i$  can be greatly extended, thereby opening the possibility of making dexterous motor correlations in response to subtle sensory discriminations.

There exist two main arrangements of  $S(\mathcal{I})$  and  $S(\mathcal{O})$ . These are depicted in Figure 3.

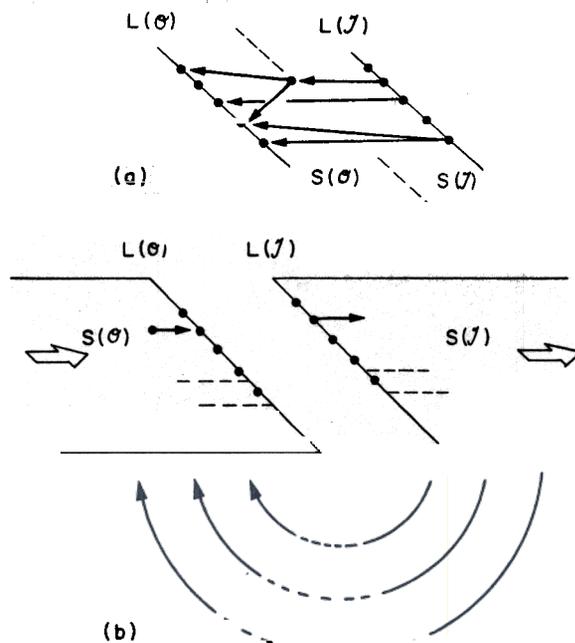


Figure 3.

In Figure 3a,  $S(\mathcal{S})$  merges with  $S(\mathcal{O})$ , and all new controls are superimposed on older and more direct  $V_i \rightarrow v_i$  connections. This arrangement would create severe space limitations and massive confusions between old and new control sequences. In Figure 3b, by contrast, the areas of  $S(\mathcal{S})$  and  $S(\mathcal{O})$  can be readily expanded if new controls are needed, and new input-output connections can be established between control regions lying ever further from  $\mathcal{S}$  and  $\mathcal{O}$  without destroying older controls. The arrangement in Figure 3b will therefore be adopted.

### 5. Central fissure and sensory-motor cortex

Figure 3b has a recognizable, though at this stage tenuously established, neural analog.  $L(\mathcal{S})$  and  $L(\mathcal{O})$  will be thought of as analogs of the primary sensory and motor somatotopic representations of the cerebral cortex. The space between them is then the "central fissure".  $S(\mathcal{S})$  is the "postcentral sensory cortex", which gradually becomes "parietal cortex" as the distance from  $L(\mathcal{S})$  increases.  $S(\mathcal{O})$  is "precentral motor cortex", which gradually becomes "frontal cortex" as the distance from  $L(\mathcal{O})$  increases ([8], p. 345).

Figure 3b suggests that the regions within  $S$  closest to  $L$  will control reflex activities, whereas those at greater distances from  $L$  will control more elaborate discriminations and learned associations. A qualitatively similar situation holds *in vivo* ([8], pp. 442, 488; [9], Chapter 16).

### 6. Impossibility of localizing learned circuits

The simplest way to control the order of  $L(\mathcal{O})$ 's outputs will now be given. This will fail to satisfy simple behavioral criteria. This failure will point the way to a better solution.

The simplest possibility is that a directed chain, or chains, of edges in  $S(\mathcal{O})$  are activated step-by-step to order the outputs from  $L(\mathcal{O})$ . Consider Figure 4. In

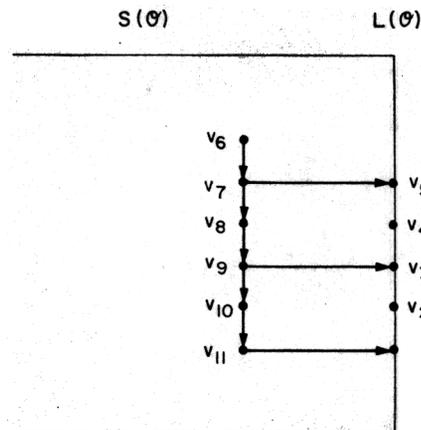


Figure 4.

Figure 4, the control vertex  $v_6$  is activated. A signal flows to  $v_7$  and from  $v_7$  to the output vertex  $v_5$  and to the control vertex  $v_8$ .  $f_5$  thereupon moves.  $v_8$  sends a signal to  $v_9$ , which in turn excites  $v_3$  and  $v_{10}$ .  $f_3$  then moves. Finally  $v_{11}$  and then  $v_1$  are excited, and  $f_1$  moves.

(A) *Existence of independent patterns.* Surely each finger can be used in at least two different patterns of finger motion, but no chain of associations as in Figure 4 can independently control two patterns that use the same fingers. To see this consider Figure 5.

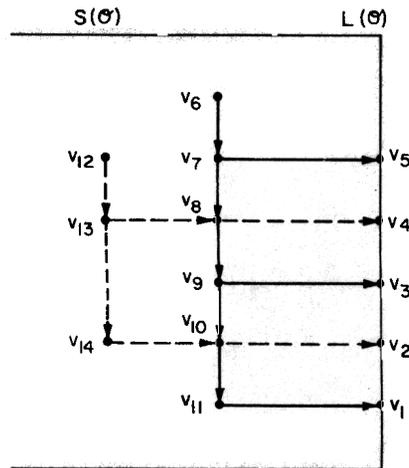


Figure 5.

Figure 5 depicts controls for two different patterns. The pattern controlled by  $v_6$  aims at moving  $f_5$ , then  $f_3$ , and then  $f_1$ , and the pattern controlled by  $v_{12}$  aims at moving  $v_4$  and then  $v_2$ . These two patterns cannot, however, be activated *independently*, since exciting  $v_6$  causes all fingers to move whereas exciting  $v_{12}$  causes all fingers but  $f_5$  to move.

Controls constructed from chains of edges (associations), as in Figure 4 and 5 are called *local controls*, since they use local flows of excitation along channels within  $S(\theta)$ . To move at least two patterns independently, either the *chainlike* character of the controls or the *locality* of the flow pattern, or both, will be violated. The chainlike character of the control will exist in some configuration since it gives rise to the ordering of the outputs. The locality of the flow pattern will therefore be violated; that is, our networks will contain edges involved in learning patterns which directly connect widely separated points in  $S$ . See Figure 6.

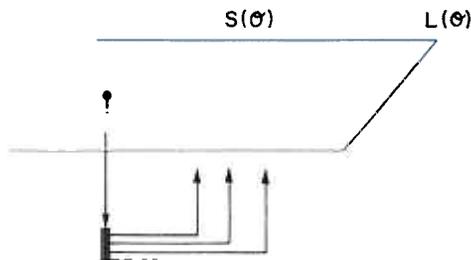


Figure 6.

(B) *Existence of spatially reordered patterns.* The need for these long edges is strengthened by noticing that signals carried locally through  $S$  impose severe constraints on the positions in  $S$  from which control vertices can generate an output pattern of prescribed timing. Consider Figure 7.

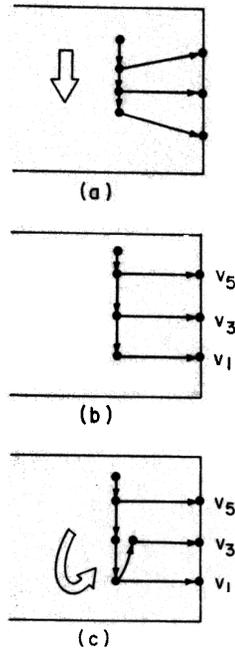


Figure 7.

Figure 7 describes three different patterns. In Figure 7a,  $f_5$ , then  $f_3$ , and then  $f_1$  move every  $w$  time units. In Figure 7b,  $f_5$ , then  $f_3$ , and then  $f_1$  move every  $2w$  time units. In Figure 7c,  $f_5$ , then  $f_1$ , and then  $f_3$  move, with  $f_3$  following  $f_1$  in half the time it takes  $f_1$  to follow  $f_3$ . Thus changing the temporal spacing in a fixed ordering of outputs changes the spatial distribution of the pattern's local control, and changing the ordering of outputs changes the distribution of time intervals between successive outputs. Other serious deficiencies of local transmission are readily imaginable.

The following sections suggest step-by-step how to violate the local surface topology of  $S$  in a way that is compatible with learning needs.  $S$  can be changed by altering either its vertices or its edges. This paper concentrates on the edges. A forthcoming paper will suggest how to alter each vertex to achieve a cell distribution reminiscent of a column of cerebral cortex. This "surface of columns" will be able to discriminate between more complicated inputs and to give rise to more adaptive outputs than can a surface of vertices.

### 7. Perpendicular somatotopic representations

By (2), learning of a transition from a vertex  $v_j$  to a vertex  $v_k$  at time  $t$  requires that the cross-correlation

$$[x_j(t - \tau_{jk}^+) - \Gamma_{jk}^+][x_k(t) - \Lambda_k^+] \quad (11)$$

be large, or that

$$x_f(t - \tau_{jk}^+) - \Gamma_{jk}^+$$

and

$$x_k(t) - \Lambda_k^+$$

separately be large. This is true if  $v_j$  creates a high average spiking frequency in  $e_{jk}$  at time  $t - \tau_{jk}^+$ , and  $v_k$  has a large average membrane potential at time  $t$ . Suppose, however, that  $v_j$  and  $v_k$  control different fingers. By Figure 3,  $v_j$  and  $v_k$  will lie in different parallel strips of either  $S$  or regions interacting with  $S$ , as we schematize in Figure 8, where  $F_i$  designates a strip representing  $f_i$ ,  $i = 1, 2, 3, 4, 5$ . Since  $v_j$  and  $v_k$  never interact, the cross-correlation (11) will never be large, and no pattern learning between fingers will occur.

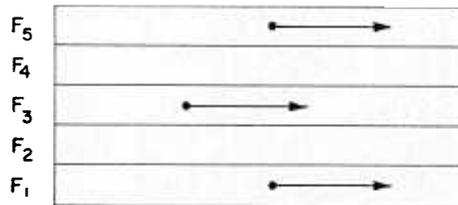


Figure 8.

In order that cross-correlations between individual finger representations be able to become large, without destroying the ordering of the representation, the existence of another finger representation that interacts with the given one at a non-parallel angle is suggested. *Multiple* somatotopic representations will hereby be implicated in learning within ordered media. Multiple sensory and motor representations are familiar *in vivo*, for example in auditory cortex, sensory-motor cortex, and the cerebellum ([8], p. 467, p. 213).

Our goal is to construct a system in which *unbiased* learning is possible. In such a system, the transitions  $f_j \rightarrow f_k, j \neq k$ , are equally easy to learn, other things equal. This property idealizes reality, since not all fingers in real life are equally easy to use, as every student of piano knows. This idealization is useful because it puts in sharp focus one extreme in the evolutionary development of learning skills from correlations between nearest-neighbor or uniformly distributed vertices to long-range correlations between highly structured arrays of vertices. Once idealizations of these extremes are understood, intermediate cases can be studied as mixtures of them.

The property of unbiased learning suggests that the two finger representations interact at right angles, as in Figure 9. This arrangement will be seen to minimize asymmetries between a given row and all columns.

The representation perpendicular to

$$F = \{F_1, F_2, F_3, F_4, F_5\}$$

will be denoted by

$$F^\perp = \{F^{(1)}, F^{(2)}, F^{(3)}, F^{(4)}, F^{(5)}\},$$

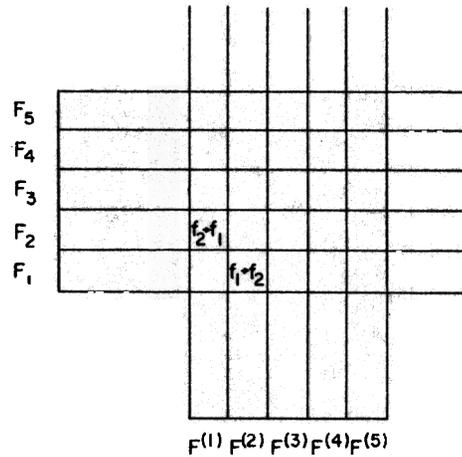


Figure 9.

or “ $F$ -perpendicular”.  $F$  and  $F^\perp$  together will be denoted by  $\mathcal{F}$ . Interactions between  $F$  and  $F^\perp$  are conveniently discussed using the *correlational matrix*

$$C = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} \end{pmatrix}.$$

The entry  $C_{jk}$  of  $C$  denotes the spatial locus of correlations corresponding to the transition  $f_j \rightarrow f_k$  between fingers.

The representations  $F$  and  $F^\perp$  need not be a part of  $S$ . It suffices that inputs to  $F$  and  $F^\perp$  be created at some stage of processing by an ordered system, and that outputs from  $F$  and  $F^\perp$  be fed back into another stage of this system in an order-preserving way. That is, the interactions within  $\mathcal{F}$  will help to learn patterns no matter how we interpret the input sources and output sinks. These interactions therefore have a “universal” character, and various copies of  $\mathcal{F}$  can be connected to different ordered bundles of input and output edges as they are needed.

### 8. Long axons: learning speeds vs. reaction times

An effective learning machine will learn material that is presented at many different speeds. An  $\mathcal{M}$  will be constructed that can learn lists at essentially any positive speeds less than some fixed finite, but otherwise arbitrary, constant  $W$ . This construction exploits the intimate connection between learning speeds, behavioral reaction times, and the transit times of excitatory signals within edges [3]. This connection follows readily from (2), since learning of a transition from  $v_j$  to  $v_k$  at time  $t$  requires formation of a large cross-correlation  $z_{jk}(t)$ . The cross-correlational process occurs at the arrowhead  $N_{jk}$  of  $e_{jk}$ , where it receives at time  $t$  the presynaptic signal (9) that left  $v_j$  at time  $t - \tau_{jk}^+$ , and compares it with the value (10) derived

from the contiguous  $v_k$  vertex. Suppose that the pre- and post-synaptic values are made large by input pulses corresponding to successive motions of  $f_j$  and  $f_k$  at a time lag  $\xi$ . These motions can be effectively learned only if  $\xi$  is approximately equal to  $\tau_{jk}^+$ .  $\xi$  need not exactly equal  $\tau_{jk}^+$  for two reasons.

- (1) The input pulses are spread out in time. Hence the vertex functions are also spread out in time.
- (2) The pre- and post-synaptic values are only correlated at suprathreshold values ( $\Gamma_{jk}^+$  and  $\Lambda_k^+$ , respectively), and these values are achieved shortly after the input pulses occur, if at all.

Thus a machine capable of learning  $f_j \rightarrow f_k$  at many different speeds will send edges between representations of  $f_j$  and  $f_k$  whose time lags  $\tau_{jk}^+$  take on many different values.

### 9. Resondant conditioning via intermittent axonal connections to output cells

Two main ways of connecting edges to output vertices are diagrammed in Figure 10. The first way provides the richer collection of learning speeds given a fixed number of edges. In Figure 10a, the excitatory correlational edge, or axon, sends signals at regularly spaced intervals along its entire length to output vertices via short axon collaterals. In Figure 10b, the excitatory edge sends signals only to an output vertex at its terminal point. In both cases, an independent source of inputs to each output vertex exists. If a large input reaches a given output cell as a large signal from the excitatory axon passes by, a large cross-correlation will begin to form at the arrow-head of the axon collateral.

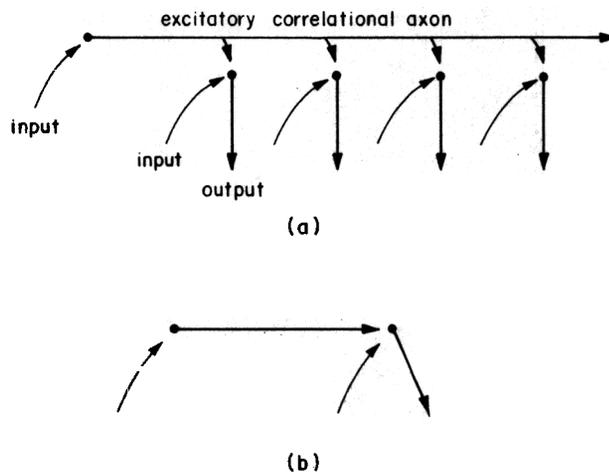


Figure 10.

Suppose that a signal travels within a given edge with uniform velocity, and that the same spatial distribution of edges occurs in Figures 10a and 10b. Then clearly Figure 10a provides at least as many learning speeds as Figure 10b. Figure 10a has several other important advantages, which will be discussed in forthcoming sections; also see [1].

The networks of Figure 10 learn according to a respondent conditioning paradigm in which the conditioned stimulus sends inputs to excitatory correlational axons and the unconditioned stimulus sends inputs directly to output cells [1].

### 10. Multiple associations in collaterals of long axons:

$$(ABC) \rightarrow D \neq (A)(B)(C) \rightarrow D$$

Axon collaterals allow a single long axon to learn sequences of associations (or a "compound" association). See Figure 11. In Figure 11, inputs perturb  $v_1$  at time  $t$ ,  $v_2$  at time  $t + \tau_{12}$ , and  $v_3$  at time  $t + \tau_{12} + \tau_{23}$ , where  $\tau_{12}$  is the (approximate) time lag for a signal to pass from  $v_1$  to  $v_2$ , and  $\tau_{12} + \tau_{23}$  is the approximate time lag for a signal to pass from  $v_1$  to  $v_3$ . As a result, large cross-correlations form at the

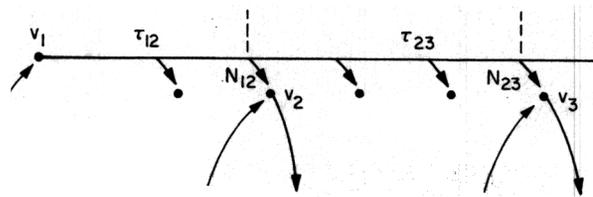


Figure 11

arrowheads  $N_{12}$  and  $N_{13}$ . Thus a subsequent input to  $v_1$  alone will create outputs from  $v_2$  and  $v_3$  approximately  $\tau_{12}$  and  $\tau_{12} + \tau_{23}$  time units later. A single edge can therefore learn a compound association and readily reproduce it. See [1] for some mathematically rigorous examples.

All edges from  $v_1$  through  $v_2$ , or from  $v_1$  through  $v_3$ , or from  $v_2$  through  $v_3$  with the proper time lags will also learn the *simple* associations

$$v_1 \rightarrow v_2, \text{ OR } v_1 \rightarrow v_3, \text{ OR } v_2 \rightarrow v_3,$$

respectively. Thus the output from  $\mathcal{F}$  will result from a temporally weighted mixture of simple and compound associations activated by previous inputs. Only the longest edges can, however, learn multiple associations in a single sweep between widely separated vertices. See [1] and [3], Sections 10 and 11, for a related discussion of multiple associations.

### 11. Spatiotemporally self-similar vs. uniform cell types

A virtually unlimited number of learning speeds for  $\mathcal{F}$  can be guaranteed by using

- (i) axon collaterals, and
- (ii) axons whose signal velocity is proportional to axon length.

The proper setting for understanding this fact is imposed by the following question. Among cell types capable of transmitting *undistorted* signals between input and output vertices, which *single* cell type assures the largest number of learning speeds for a fixed number of edges? [4] and [10] discuss two main idealized cell types that can transmit undistorted signals between vertices. Geometrically speaking, these two types have the following properties.

(1) *Spatiotemporally self-similar (STSS) cell type*

- (a) The shapes of any two cells in a given STSS type differ only by a multiplicative factor.
- (b) The velocity of signal transmission in a given edge is proportional to the length of the edge (and also, by (a), to the diameter of the edge). Thus the time lag for signal transport through the entire edge is independent of edge length.

STSS cells can, in principle, shrink or grow without disturbing extant cell connections and time lags. They can therefore be used to coordinate the activities of widely separated cell groups while the system grows in size or makes room for new correlational regions. Moreover, production levels set in the cell body by membrane excitation are the correct production levels for the entire cell.

(2) *Uniform (U) cell type*

These cells differ from STSS cells primarily by violating (b) as follows:

- (c) The velocity of signal transmission in a given cell type is independent of edge length, and thus the transit time of a signal through an edge is proportional to edge length.

Cell type STSS provides more learning speeds than *U* for the following reasons. Let the distance between vertices  $v_j$  and  $v_k$  be denoted by  $d(v_j, v_k)$ . By (c), type *U* can learn a transition  $v_j \rightarrow v_k$  only with time lag proportional to  $d(v_j, v_k)$  whether or not axon collaterals exist. Moreover, if

$$d(v_j, v_k) \neq d(v_j, v_m), \quad (12)$$

then the transitions  $v_j \rightarrow v_k$  and  $v_j \rightarrow v_m$  must be learned at different speeds.

By (b), type STSS can learn transitions  $v_j \rightarrow v_k$  and  $v_j \rightarrow v_m$  satisfying (12) at the same speed in either of two ways:

- (1) Send axons from  $v_j$  to  $v_k$  and  $v_m$  that terminate at  $v_k$  and  $v_m$ , respectively.
- (2) Sends axons from  $v_j$  through  $v_k$  and  $v_m$  with lengths proportional to  $d(v_j, v_k)$  and  $d(v_j, v_m)$ , respectively, and connect these axons to  $v_k$  and  $v_m$  using axon collaterals.

Although axon collaterals do not add to the learning speeds of type *U*, they do the following for type STSS. Let  $M$  be the maximum edge length. Then a transition  $v_j \rightarrow v_k$  with

$$d(v_j, v_k) < M$$

can be learned with any time lags between

$$\frac{d(v_j, v_k)}{\theta M}, \quad \text{and} \quad \frac{1}{\theta}, \quad (13)$$

approximately, where  $\theta$  is the signal velocity in an edge of unit length. These advantages suggest the use of STSS cells wherever possible.

**12. Unbiased learning speeds require spatial dispersion of input**

Equation (13) shows that unbiased learning cannot hold if the input to  $F_i$  arrives at a localized cluster of vertices. Consider Figure 12. By (13),  $C_{ik}$  correlations can

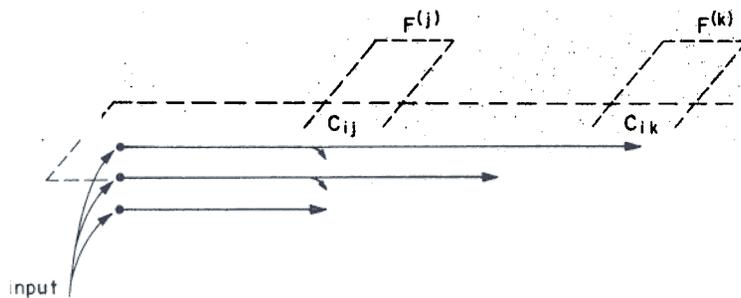


Figure 12.

be learned at fewer speeds than can  $C_{ij}$  correlations. This bias in learning speeds can be overcome by distributing the input uniformly in space and time to all correlational axons of a given length *throughout* the strip  $F_i$ , as in Figure 13.

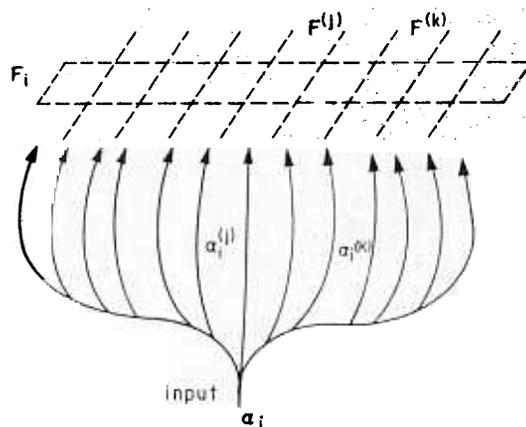


Figure 13.

### 13. Curved gray matter vs. fast signals in long axons

The various  $F^{(j)}$  representations might lie at different distances from the input source  $\alpha_i$  to  $F_i$ . Yet all these representations should receive simultaneous signals from  $\alpha_i$  to avoid temporal biases. This spatial bias can be overcome in either of two ways.

- (1) Bend the strip  $F_i$  until all branches  $\alpha_i^{(j)}$  of the  $\alpha_i$  input have equal length.

A simpler way is given by

- (2) Let the axon collaterals  $\alpha_i^{(j)}$  obey STSS. Then the longest collaterals have the fastest signal velocities, and so the time lags of all  $\alpha_i^{(j)}$  can be made the same.

### 14. Postsynaptic control of axon branches vs. rosettes

The existence of different signal velocities in axons of different length can create temporal biases. If  $\alpha_i^{(j)}$  is long (short), it will disperse a rapidly (slowly) travelling signal among the correlational axons with vertices in  $C_{ij}$ . This temporal bias can be overcome in either of two ways.

- (1) *Axon collaterals of uniform thickness*: Let all axons  $\alpha_i^{(j)}$  break up into branches of *uniform* thickness (and hence the same velocity) once they enter  $C_{ij}$ . This possibility suggests that the average activity of the *postsynaptic* cells to which each  $\alpha_i^{(j)}$  sends signals can help to determine the distribution of  $\alpha_i^{(j)}$  terminal branches. This suggestion is compatible with the idea that average postsynaptic activity helps to set various presynaptic production levels (e.g., of transmitter [4]), and thereby presumably to control the volume of the chemical machine regulating these levels and/or the storage space for the chemicals produced. This postsynaptic influence on a presynaptic cell can violate STSS in the presynaptic cell, especially in the cell's most distal axons and (reversing pre- and post-synaptic labels) dendrites.
- (2) *Rosettes*: Let each  $\alpha_i^{(j)}$  terminate in a spatially localized region  $R_i^{(j)}$ , or "rosette" ([11], pp. 43, 129), to which several correlational cells can send their dendrites. Then the problem of distributing inputs without temporal bias to each  $C_{ij}$  vanishes, since no longer is the rate of input arrival dependent on the presynaptic axonal length. Rather it will be seen to depend on the network analogs of granule cell dendrites. A spatial bias can, however, occur, since wider axons will deliver large signals, other things equal. This difficulty is readily overcome by letting volume of  $R_i^{(j)} \propto$  diameter of  $\alpha_i^{(j)}$ , or effective surface area for synaptic contacts with  $R_i^{(j)} \propto$  diameter of  $\alpha_i^{(j)}$ .

Then the excitation density in all rosettes will be the same, other things equal. This constraint on rosette size does *not* violate STSS. A hypothetical rosette is diagrammed in Figure 14.

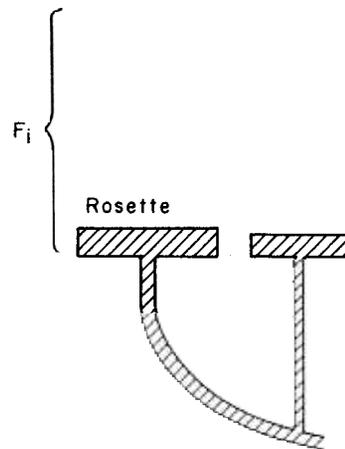


Figure 14.

### 15. Bidirectional axons: unbiased arrow in time and boundary correlations

In preceding sections, all correlational edges have been drawn facing the right. This assumption is inadmissible for two reasons.

- (1) No right-handed correlational edges send signals to the left-hand boundary of  $F_i$ . Thus only weak correlations can be made near this boundary.

(2) No right-handed correlational edge can learn a multiple association  $f_i \rightarrow f_j \rightarrow f_k$  if  $F^{(k)}$  is to the left of  $F^{(j)}$ .

Thus edges facing both right and left are needed to avoid biases. A simple way to achieve this is to suppose that bidirectional correlational edges exist, as sketched in Figure 15.

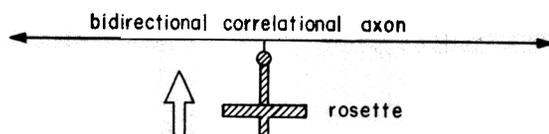


Figure 15.

### 16. Folia, parallel fibers, mossy fibers, and rosettes

A similarity exists between some components of  $\mathcal{F}$  and known cerebellar objects. The following correspondence is suggested.

bidirectional correlational axon = parallel fiber

vertex of correlational axon = granule cell

axons  $\alpha_i$  and  $\alpha_i^{(j)}$  = mossy fibers

$R_i^{(j)}$  = rosettes

collection of all cells in  $F_i$  = a folium

See ([11], p. 5). We will sometimes use this neural interpretation of  $\mathcal{F}$  to make suggestive statements about the cerebellum. For example: the needs of unbiased learning between ordered sensory and motor media suggest that mossy fibers terminate diffusely within a given folium in rosettes to which vertically displaced granule cells with long bidirectional fibers send their dendrites.

### 17. Purkinje cells and climbing fibers

At least two further analogs of cerebellar components exist in  $\mathcal{F}$ . Consider Figure 10a. We adopt the following nomenclature:

output cell = Purkinje cell,

and

input axon to

an individual

output cell = climbing fiber.

Our construction thus suggests that *two* sources of inputs are needed to guarantee learning: mossy fibers and climbing fibers. In Figure 10, each climbing fiber has a localized projection to a given Purkinje cell, or to a localized Purkinje cell cluster in a given folium. These statements are borne out experimentally [11], Chapters VII and VIII. Moreover, the climbing fiber somatotopic representation is *perpendicular* to the mossy fiber somatotopic representation to guarantee the possibility of forming unbiased cross-correlations between the two representations. The perpendicularity of mossy fiber and climbing fiber representations has been recently observed experimentally [12].

### 18. Layers of parallel fibers segregated by length

Suppose as above that parallel fibers of equal length have the same signal velocity. To avoid temporal biases among fibers of equal length, the time needed to deliver inputs to any of them should be the same, and the time needed to deliver outputs from them should be the same. Both criteria can be simultaneously achieved by arranging parallel fibers in layers according to length.

It remains only to decide whether long or short axons occur closest to the mossy fiber input source and to the Purkinje cell body output sink. Since the longest axons can form the most complicated associations, it is natural to let them occur closest to the input source, where they can receive input information "first-hand" and before it is spread more diffusely to higher layers. The longest axons will then also pass closest to the Purkinje cell body so that the most subtle learned correlations will most powerfully influence the output.

### 19. Dendritic spines and learning

In Section 8, we noted that each parallel fiber makes functional contact with the Purkinje cell over which it passes. This contact was assumed for simplicity to occur via short axon collaterals at regular distances along the axon. Given the separation of parallel fibers into layers by length, the axon collaterals passing from the uppermost layer to the Purkinje cell body could interfere substantially with the parallel course of the long fibers, if a substantial number of fibers exist, and thereby bias these most useful correlators. To avoid this unpleasant possibility, the mechanism of parallel fiber and Purkinje cell contact should have three properties.

- (1) It should be sufficiently spread out among the parallel fibers to collect signals from all fibers above a given Purkinje cell.
- (2) It should nonetheless be so thin that parallel fibers can pass by it to other Purkinje cells without deviating from their parallel course.
- (3) It should collect excitation in a way that preserves the temporal ordering sought by segregating parallel fibers in layers according to length.

Clearly the extensive bush of dendrites found above each Purkinje cell body ([11], p. 71) can fulfill all these requirements.

Short axon collaterals from a parallel fiber to an adjacent Purkinje cell dendrite will cause less interference than before, but even these collaterals will limit the packing density of parallel fibers. Since the existence of the largest possible number of learning speeds is a basic requirement, it is not surprising that *in vivo* the analogous functional contact is made via *dendritic spines* that invaginate the parallel fiber and therefore minimally limit the parallel fiber packing density ([11], p. 52). Thus the

parallel fiber → dendritic spine

contact is implicated as a region in which new cross-correlations can be formed. In particular, presynaptic transmitter concentrations might be expected at such contacts, and indeed synaptic vesicles have been observed in these regions ([11], p. 52). Moreover the number of spines would then increase either as the motor learning capacities of various species increase through the phyla, or as an individual organism's motor skills mature. [11] (p. 314) briefly makes a similar suggestion, but goes on to say that "as yet, of course, we have no knowledge of the structural and

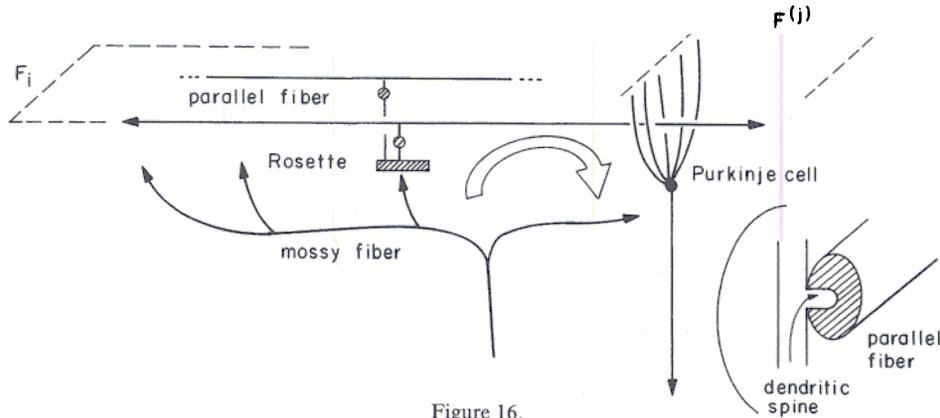


Figure 16.

functional changes that form the basis of this learned response". We can now assert that the

mossy fiber → parallel fiber → Purkinje cell

circuit diagrammed in Figure 16 is suggested by simple formal prerequisites of unbiased learning within a system of type (1)–(5) with ordered sensory and motor components. Section 22 will, however, suggest that  $\mathcal{F}$  interacts with another system  $\mathcal{G}$  that contains a richer learning structure, and that  $\mathcal{F}$ 's primary task is to reduce behaviorally destructive background noise in  $\mathcal{G}$ .

### 20. Climbing fibers around Purkinje dendrites

By Figure 10a, the climbing fiber creates an input that can be correlated with signals passing through all parallel fibers over its Purkinje cell. Since parallel fibers make their contacts via Purkinje cell dendrites, the climbing fiber input should reach all dendrites. Since the lowest layers receive the most direct mossy

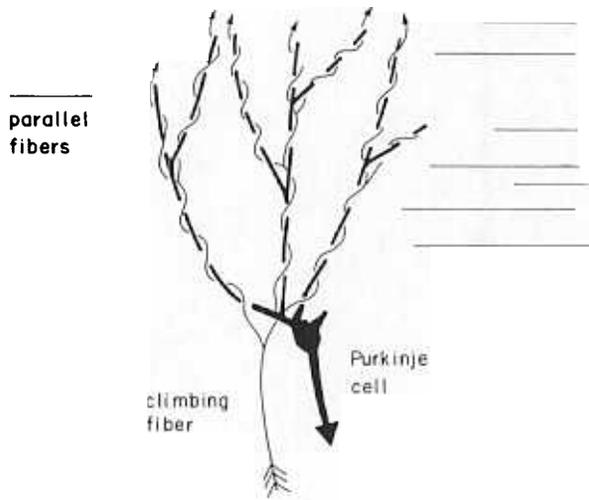


Figure 17

fiber inputs, the best correlations will be formed if the lowest Purkinje dendrites receive the most direct climbing fiber inputs. Figure 17 describes the resultant situation, which is similar to the arrangement between climbing fiber and Purkinje cell that exists *in vivo*.

### 21. Equilibrium cell configurations: control of cell size and orientation by dipole forces

The very fact that only finitely many cells exist often tends to create biased learning conditions. Two cases of this are illustrated in Figure 18. In Figure 18a, parallel fibers are uniformly distributed in a strip, and in Figure 18b, edges are distributed

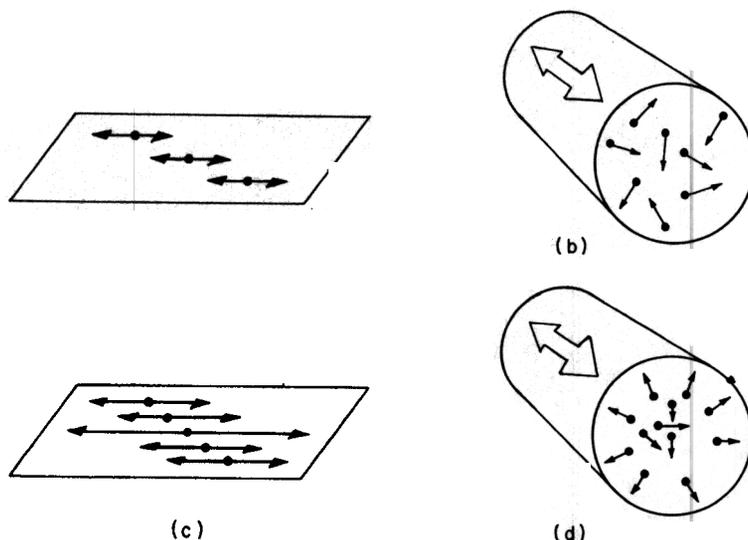


Figure 18.

with random orientations perpendicular to a bundle of parallel fibers. In both cases, the boundaries  $B$  receive less excitation (or inhibition) than the interiors  $I$ , on the average. These boundary biases can be overcome as follows. The boundary bias in Figure 18b is corrected in 18d by orienting the cells near  $B$  with their edges facing  $B$  and their vertices facing  $I$ . Speaking teleologically, this orientation maximizes the excitatory input which the vertex can receive at a given cell position, and then disperses the input in signals to  $B$ , where they are most needed. The attraction of vertices to excitation can be qualitatively understood as follows.

Reference [4] suggests that various nerve cells are "chemical dipoles". The cell body (and dendrites) and the synaptic knobs contain the two ends of the dipole. The cell body end of the dipole thrives on various electrical and/or ionic fluctuations created by presynaptic excitation, since this excitation helps to set various production levels within the cell. Presumably the spatial orientation of such a cell can be determined by the spatial gradient of these electrical and/or ionic densities. In particular, the sources of excitation are often the synaptic knobs (or axons) of other cells, and in this sense opposite ends of the dipole attract and like ends repel.

In 18a, the orientation of the cells is fixed, but their size is variable. [4] suggests that the size of STSS cells can, in principle, covary with the average intensity of inputs to the cell body. Then if cell bodies near *I* receive more excitation than cells near *B*, they will send out long axons towards *B* and thereby help to bring more excitation to *B*.

Dipole cells satisfying STSS can in these ways adapt to changing patterns of inputs by altering their size and orientation. These geometrical changes presumably seek a "maximally stable" configuration of cells relative to a given input mechanism. Fortunately, this equilibrium configuration also helps to lessen biases in learning.

This discussion provides an example of how the laws for individual cells might help to determine useful arrangements of many cells. Dipole forces, for example, let each cell body maintain production levels needed to keep up with prior input demands and to transmit signals between cells that reproduce these demands. STSS guarantees that the production levels of the cell body are the right ones for the entire cell. Such a cell is ignorant of all production needs but its own. Nonetheless, by satisfying its own production needs, it automatically helps to create behaviorally useful anatomical patterns with other equally egocentric cells.

Note that the mechanism for reducing boundary biases in 18c introduces a temporal bias by creating longer parallel fibers from *I* than from *B*. To avoid these boundary biases, the length of a given set *F* of folia should be substantially (at least three times) longer than the width of its perpendicular representation  $F^\perp$ .

Keeping the widest and longest correlational axons closest to input sources and output sinks helps maintain an equilibrium configuration of cells. The largest mossy fiber inputs are then carried through the largest parallel fibers to the largest Purkinje cell dendritic branches. Less direct inputs are carried through finer parallel fibers to finer Purkinje cell dendrites. See Figure 19.

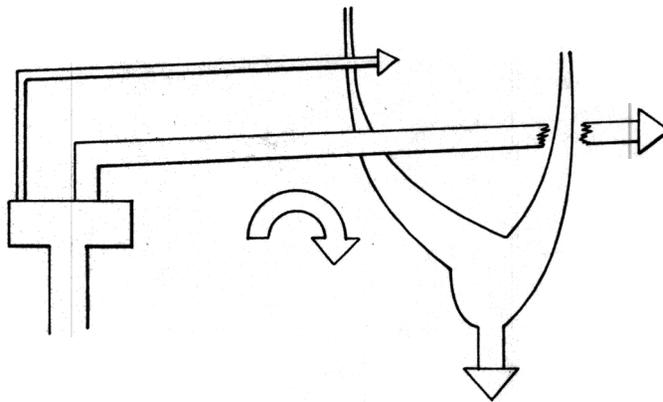


Figure 19.

## 22. Possible learning speeds

For which presentation speeds of  $f_j \rightarrow f_k$ ,  $j \neq k$ , will learning occur if

- (1) STSS holds in parallel fibers, and
- (2) no boundary biases occur in the spatial distribution of parallel fibers.

Let  $f_j \rightarrow f_k$  occur with a time lag of  $\xi$ , and suppose for simplicity that inputs representing these events arrive at *F* and  $F^\perp$  with the same time lag  $\xi$ . The parallel fibers

that can learn this transition will have large signals in  $C_{jk}$  when the climbing fibers of  $C_{jk}$  are active. These parallel fibers will have received inputs from mossy fiber rosettes approximately  $\xi$  time units earlier.

Let  $W$  be the longest time lag of any parallel fiber. Let  $w$  be the time lag between onset of the shortest parallel fiber signal and passage of this signal through the nearest Purkinje cell dendrites. Given (1) and (2), learning is (approximately) equally easy for any  $\xi$  such that

$$w \leq \xi \leq W,$$

but becomes harder as  $\xi$  decreases below  $w$  or increases above  $W$ . The importance of  $w$  and  $W$  is seen in Figure 20. Figure 20 shows that as  $\xi$  decreases below  $w$  or

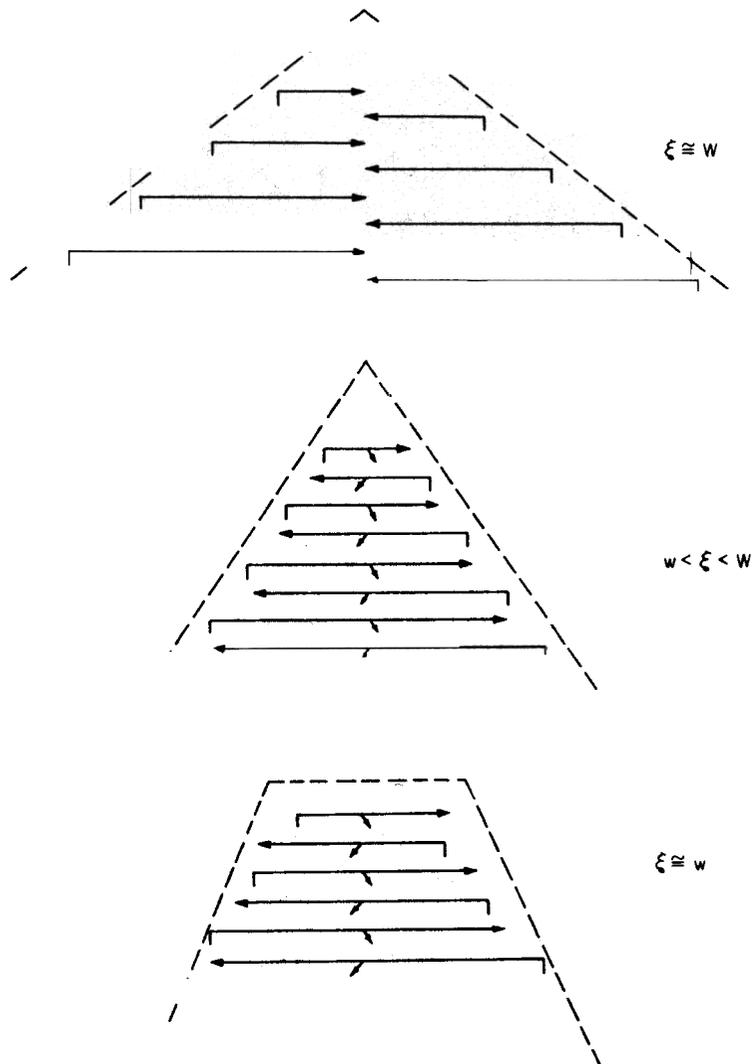


Figure 20.

increases above  $W$ , fewer parallel fibers can carry excitation to  $C_{jk}$  when the climbing fibers in  $C_{jk}$  are active.

### 23. Spatiotemporal masking and consolidation

Spatiotemporal masking and consolidation are two consequences of

- (1) STSS in parallel fibers,
- (2) arranging parallel fibers in layers according to length (i.e., of arranging parallel fibers in an equilibrium configuration).

Figure 20 shows one kind of consolidation as the speed of  $1/\xi$  of the sequence increases. The volume of the dotted triangular region containing the segments of parallel fibers used in learning decreases as  $1/\xi$  increases.

Consolidation also occurs in another way. Consider learning of the sequence

$$f_i \rightarrow f_j \rightarrow f_k$$

presented with a time lag  $\xi$  between successive items. If  $\xi \cong W$ , then no single parallel fiber is long enough to form a multiple association  $f_i \rightarrow f_j \rightarrow f_k$ . At slow speeds, only sequences of simple associations can be learned. Each of these simple associations will be formed among a triangular region of parallel fiber segments. For smaller values of  $\xi$ , by contrast, the entire chain  $f_i \rightarrow f_j \rightarrow f_k$  of associations can be encoded in individual long parallel fibers. Only parallel fibers  $w_{ijk}$  with the following properties can form the multiple associations  $f_i \rightarrow f_j \rightarrow f_k$ :

- (1)  $w_{ijk}$  lies in  $F_i$ ,
- (2)  $w_{ijk}$  faces from  $F^{(j)}$  to  $F^{(k)}$ , and
- (3)  $w_{ijk}$  is at least as long as the distance from  $F^{(j)}$  to  $F^{(k)}$ .

See Figure 21.

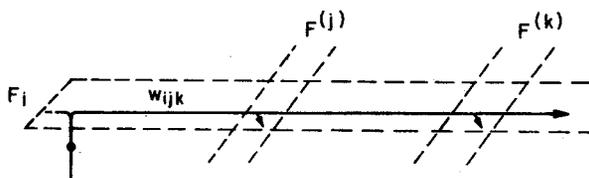


Figure 21.

Suppose the long fibers  $w_{ijk}$  have learned the multiple association. Then activation of  $F_i$  by mossy fibers creates a signal in  $w_{ijk}$  that travels from  $F^{(j)}$  to  $F^{(k)}$  Purkinje cells faster than the shorter fibers carrying single associations. Moreover the  $w_{ijk}$  fibers end closer to the Purkinje cell body than the shorter fibers. Successive outputs from  $F^{(j)}$  and  $F^{(k)}$  are hereby created by  $w_{ijk}$  before the simple associations can create outputs. That is, *spatiotemporal masking* of the simple associations by the multiple association has occurred. Similarly, the spatial locus of the  $w_{ijk}$  association is a small region of long fibers that lie maximally close to the input and output cells of  $F_i$  rather than in two separate large triangular regions of parallel fibers in  $F_i$  and  $F_j$ . *Spatiotemporal consolidation* of the memory trace has hereby occurred.

## 24. Subliminal excitation in upper layers

Since the fastest and most accurate correlations occur in the longest fibers of the lowest layers, the most diffuse subliminal inputs should send signals to the upper layers. This diffuse excitation would subliminally excite the total length of Purkinje dendrites via short parallel fibers to prepare the Purkinje cell and dendrites for later signals from the long parallel fibers. For example, reticulocerebellar mossy fibers seem to contact parallel fibers in the upper layers ([11], p. 34–5).

## 25. Intermodality interactions and inhibitory output from Purkinje cells

Thusfar our construction has been motivated by the special case in which  $F$  and  $F^\perp$  represent the same set of peripheral organs. This need not hold true in general, because a particular pattern of motor outputs is often determined by sensory feedback from several modalities. Consider, for example, the motion of a pianist's fingers over the piano keyboard. Surely the particular sequence of notes played by a given ordering of fingers will help to determine the next note to be played. We thus consider the general case in which  $F$  and  $F^\perp$  can represent different modalities, and in particular will often interpret inputs to  $F^\perp$  as intermodality sensory feedback created by motor outputs from the modality represented by  $F$ . In this more general case, our previous arguments suggest that  $\mathcal{F}$  interacts with another region  $\mathcal{G}$  of  $\mathcal{M}$  to produce learning. This is because  $C_{jk}$  no longer represents a region that once activated by  $f_j$  gives rise to outputs to  $f_k$ . Rather,  $F_j$  and  $F^{(k)}$  activation focussed at  $C_{jk}$  means, for example, that “the finger corresponding to  $F_j$  has played the note corresponding to  $F^{(k)}$ ”. Whereas activating the “finger row”  $F_j$  can activate the “frequency column”  $F^{(k)}$ , outputs from  $C_{jk}$  produce no motor output from  $\mathcal{M}$ —although they can create internal feedback to a tonotopic representation within  $\mathcal{M}$ . Playing the sequence of “notes”  $F^{(1)}, F^{(2)}, F^{(3)}, F^{(4)}, F^{(5)}$  using the “fingers”  $F_1, F_2, F_3, F_4, F_5$  now creates maximal excitation successively in the *diagonal row*  $C_{11}, C_{22}, C_{33}, C_{44}, C_{55}$  of regions. Associations such as

$$C_{11} \rightarrow C_{22} \rightarrow C_{33} \rightarrow C_{44} \rightarrow C_{55}$$

should therefore form, and in general non-horizontal associations such as  $C_{i_1j_1} \rightarrow C_{i_2j_2} \rightarrow C_{i_3j_3} \rightarrow \dots$  should be possible. But our ordering constraints have precluded the existence of non-horizontally oriented correlational edges in  $\mathcal{F}$ ! Hence we are led to consider the possible existence of another region  $\mathcal{G}$  interacting with  $\mathcal{F}$  that has these non-horizontally oriented correlational edges. Such a  $\mathcal{G}$  has correlational axons which are distributed in essentially all directions, for example radially around their source vertices, and joint inputs to  $\mathcal{G}$  and  $\mathcal{F}$  will presumably preserve within  $\mathcal{G}$  the ordering of the modalities represented by  $F$  and  $F^\perp$ . But then  $\mathcal{G}$  can readily be endowed with at least as rich a correlational structure as  $\mathcal{F}$  possesses. What purpose, then can  $\mathcal{F}$  serve that cannot be better served by  $\mathcal{G}$ ?

$\mathcal{G}$  faces a formidable problem that can be eliminated by interaction with  $\mathcal{F}$ . Let  $G_i$  and  $G^{(j)}$  be the row and column in  $\mathcal{G}$  corresponding to  $F_i$  and  $F^{(j)}$  in  $\mathcal{F}$ . Let  $G_i$  receive an input which is also passed along to  $F_i$ . In  $F_i$ , this input creates signals that remain in  $F_i$ . In  $G_i$ , the input creates signals that are carried throughout large sectors of  $\mathcal{G}$  by its non-horizontally distributed correlational axons. Although this wide dispersion of the input is needed for  $\mathcal{G}$  to be able to make intermodality

correlations of the above type,  $\mathcal{G}$  would quickly become hopelessly flooded by background noise if these signals were not rapidly suppressed. Part II of this paper will suggest that these widespread signals within  $\mathcal{G}$  can be partially controlled by *inhibitory* signals from the Purkinje cells of  $\mathcal{F}$ . With this interpretation, the cross-correlated mossy and climbing fiber inputs pick out the spatial foci in  $\mathcal{G}$  that will receive maximal inhibition. Recent experiments have, indeed, shown that the Purkinje cell output is inhibitory ([11], p. 71). We will suggest that internal inhibitory interactions within suitable  $\mathcal{G}$ 's can partially control  $\mathcal{G}$ 's large membrane potentials when excitation is spread rather uniformly over  $\mathcal{G}$ 's cells; i.e., when the *entropy* of the potential distribution in space is large. When the entropy of the potential distribution in space becomes small, however, Purkinje cell inhibition will become an especially important factor in eliminating residual background noise.

The various  $\mathcal{G}$ 's interacting with  $\mathcal{F}$  will be interpreted as various subcortical nuclei and neocortical regions. Dispersion of correlational axons in  $\mathcal{G}$  will clearly be richest whenever  $\mathcal{G}$  can make very complicated associations. The veritable jungle of converging and diverging associational axons in various thalamic nuclei and between neocortical columns is a manifestation of this need. See [3], Section 20, for a discussion of how convergence and divergence of correlational pathways creates subtle associations.

The interpretation of  $\mathcal{F}$  as an inhibitory mechanism for  $\mathcal{G}$  shifts the burden of learning from  $\mathcal{F}$  to  $\mathcal{G}$ . For example, if a given column  $F_i$  is associated at randomly distributed times with all columns  $F^{(j)}$  (the same "finger" plays different "notes"), then the strength of correlations in the regions  $C_{ij}$  will be uniformly distributed throughout  $F_i$ , on the average, and only a general *facilitation* effect for usage of the "finger"  $f_i$  will remain. The correlational axons of  $\mathcal{G}$  can be studied much as  $\mathcal{F}$ 's parallel fibers have been studied above, but the richer distribution of axons in  $\mathcal{G}$  will permit a much wider variety of discriminations to be made.  $\mathcal{F}$  itself needs some internal inhibitory cells to prevent massive outputs that do not represent behaviorally significant alternatives. These will be discussed in Part II of this paper.

## 26. Convergence and divergence of representations

Juxtaposition of various copies of  $F$  and  $F^+$  can yield subtle graded effects in space and time. In particular, varying the convergence of signals due to overlap of mossy fiber projections and of the parallel fibers to which they project can yield significantly different effects, some of whose main controlling parameters are listed below.

Let

- (a) the length of a single mossy fiber projection in a given folium equal  $K$ ,
- (b) the number of mutually independent mossy fiber sources in a folium length of  $K$  equal  $L$ ,
- (c) the longest parallel fibers in the folium have length  $M$ ,
- (d) the width of each independent source of climbing fiber inputs equal  $N$ ,  
and
- (e) the number of adjacent mutually independent climbing fiber sources equal  $P$ .

See Figure 22.

Overlap of mossy fiber projections occurs unless  $L \cong 1$ . If  $L \gg 1$ , this overlap can have the following effects.

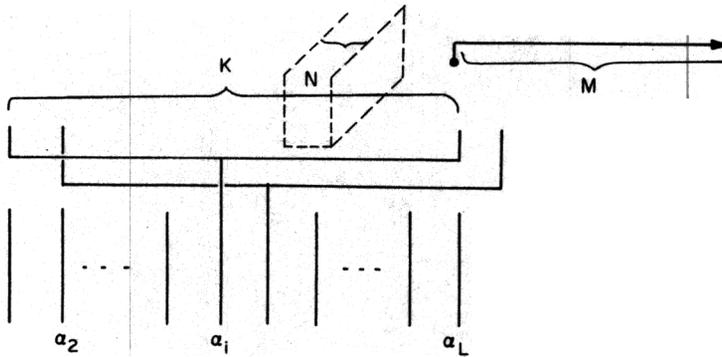


Figure 22.

(1) *Input context affects associational growth.* Let the mossy fibers  $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_m}$  be excited at almost identical times. Then the parallel fibers  $P(i_1, \dots, i_m)$  that send dendrites to the rosettes of all these mossy fibers will receive the largest total input. Their ability to correlate suitably timed climbing fiber inputs is thereby enhanced. If a later input to even *one* mossy fiber  $\alpha_{i_j}$  occurs, the strongest outputs will be activated via  $P(i_1, \dots, i_m)$ .

The number of fibers in  $P(i_1, \dots, i_m)$  will depend sensitively on  $m$  and the distances  $|\alpha_{i_j} - \alpha_{i_k}|$ . Consider Figure 23. Figure 23(a, b) shows that exciting increasing numbers of adjacent mossy fibers can limit the granule cells of  $P(i_1, \dots, i_m)$  to a narrow band of cells, whereas exciting even a large number of mossy fibers for which  $|\alpha_{i_j} - \alpha_{i_k}| > K$  has no such effects.

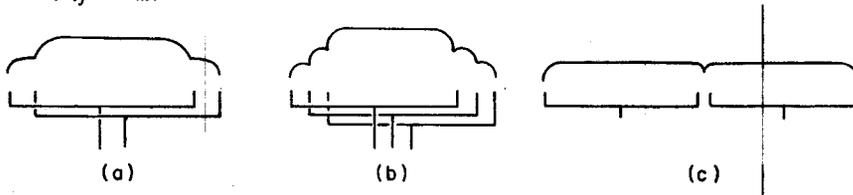


Figure 23.

(2) *Temporal biases due to spatial summation.* If a single mossy fiber source  $\alpha_i$  perturbs a given folium region, then temporal biases in that region can be eliminated by the methods of Section 11–19, say by choosing  $K \geq 3PN$  and  $L \geq PN$ , etc. Then all transitions  $f_j \rightarrow f_k, j \neq k$ , can be correlated at the same speeds.

Summation of inputs from several contiguous mossy fiber sources  $\alpha_{i_j}, j = 1, 2, \dots, m$ , with  $|\alpha_{i_j} - \alpha_{i_k}| \ll K$ , can reintroduce temporal biases. For example, the focus of excitation in Figure 23b has much the same effect as the localized excitation of Figure 12. Thus effective correlations can be made at the highest speeds with the climbing fibers terminating closest to the excitation focus, by (13). In particular, if adjacent mossy fiber representations are rapidly excited in succession, a predisposition is created for them to make the most rapid correlations with an adjacent climbing fiber representation.

### 25. Spatially disjoint multiple representations

Useful consequences derive from orienting disjoint multiple somatotopic representations at different angles relative to the folia. Consider Figure 24 ([8], p. 213).

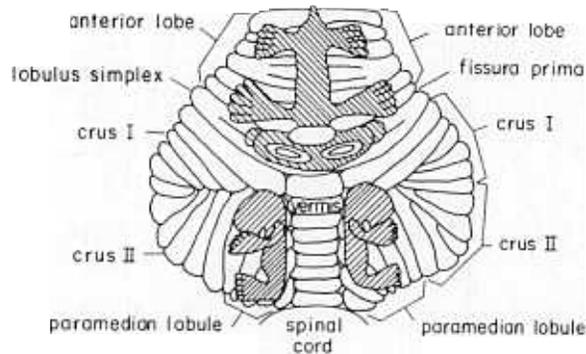


Figure 24.

(Figure 133 from Snider, 1952, *Res. Publ. Ass. nerv. ment. Dis.*, 30: 267–281). Three somatotopic representations *A*, *B*, and *C* appear in Figure 24. The representations *A* and *B* represent one half of the body each, whereas *C* represents both halves. The somatotopic representations in *A* and *B* have an orientation relative to the folia that is approximately perpendicular to the relative orientation of *C*. For example, in *A* the folia run perpendicular to the finger representations of the right hand, and in *C* the folia run approximately parallel to the representation of each finger. Whereas parallel fibers join different fingers in *A* and *B*, parallel fibers join successive joints from finger to hand, to arm, to body, and to the contralateral representation of successive joints in *C*. These multiple representations presumably facilitate both learning and performance of motor sequences by “crispening” the waves of excitation along their respective motor control paths, say from finger to finger in *A* and *B* or from joint to successor joint in *C*.

#### Acknowledgment

The preparation of this work was supported in part by the National Science Foundation (GP 9003) and the Office of Naval Research (N00014-67-A-0204-0016).

#### References

1. S. GROSSBERG, “Some Networks that Can Learn, Remember, and Reproduce Any Number of Complicated Space-Time Patterns, I, II”, *J. Math. and Mech.*, July (1969), in press. *SIAM J. of Applied. Math.*, submitted for publication.
2. S. GROSSBERG, “Embedding Fields: A Theory of Learning with Physiological Implications”, *J. Math. Psych.*, June (1969), in press.
3. S. GROSSBERG, “On Learning, Information, Lateral Inhibition, and Transmitters”, *Math. Biosci.* (1969), in press.
4. S. GROSSBERG, “On the Production and Release of Chemical Transmitters and Related Topics in Cellular Control”, *J. Theoret. Biol.*, 22, 325 (1969).
5. S. GROSSBERG, “Some Physiological and Biochemical Consequences of Psychological Postulates”, *Proc. Natl. Acad. Sci.*, 60 (3), 758–765 (1968).
6. S. GROSSBERG, “Nonlinear Difference-Differential Equations in Prediction and Learning Theory”, *Proc. Natl. Acad. Sci.*, 58 (4), 1329–1334 (1967).
7. S. GROSSBERG, “Some Nonlinear Networks Capable of Learning a Spatial Pattern of Arbitrary Complexity”, 59 (2), 368–372 (1968).

8. E. C. CROSBY, T. HUMPHREY, and E. W. LAUER, "Correlative Anatomy of the Nervous System". New York: The Macmillan Co., (1962).
9. W. J. S. KRIEG, "Functional Neuroanatomy". Philadelphia: The Blakiston Co., (1947).
10. S. GROSSBERG, "On Spatiotemporal Self-Similarity as an Organizing Principle in Neural Interactions".
11. J. C. ECCLES, M. ITO, and J. SZENTAGOTHAI, "The Cerebellum as a Neuronal Machine". New York: Springer-Verlag, (1967).
12. O. OSCARSSON and N. UDDENBERG, "Somatotopic Termination of Spino-Olivocerebellar Pathway", *Brain Res.*, 3, 204-207, (1967).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(Received November 5, 1968)