

## A Theory of Cerebellar Function

JAMES S. ALBUS

*Cybernetics and Subsystem Development Section*

*Data Techniques Branch*

*Goddard Space Flight Center*

*Greenbelt, Maryland*

Communicated by Donald H. Perkel

---

### ABSTRACT

A comprehensive theory of cerebellar function is presented, which ties together the known anatomy and physiology of the cerebellum into a pattern-recognition data processing system. The cerebellum is postulated to be functionally and structurally equivalent to a modification of the classical Perceptron pattern-classification device. It is suggested that the mossy fiber  $\rightarrow$  granule cell  $\rightarrow$  Golgi cell input network performs an expansion recoding that enhances the pattern-discrimination capacity and learning speed of the cerebellar Purkinje response cells.

Parallel fiber synapses of the dendritic spines of Purkinje cells, basket cells, and stellate cells are all postulated to be specifically variable in response to climbing fiber activity. It is argued that this variability is the mechanism of pattern storage. It is demonstrated that, in order for the learning process to be stable, pattern storage must be accomplished principally by weakening synaptic weights rather than by strengthening them.

---

### 1. INTRODUCTION

A great body of facts has been known for many years concerning the general organization and structure of the cerebellum. The regularity and relative simplicity of the cerebellar cortex have fascinated anatomists since the earliest days of systematic neuronanatomical observations. In just the past 7 or 8 years, however, the electron microscope and refined micro-neurophysiological techniques have revealed critical structural details that make possible comprehensive theories of cerebellar function. A great deal of the recent physiological data about the cerebellum come from an elegant series of experiments by Eccles and his coworkers. These data have been compiled, along with the pertinent anatomical data, in book form by Eccles *et al.* [5]. This book also sets forth one of the first reasonably detailed theories on the function of the cerebellum. Another theory,

*Mathematical Biosciences* **10** (1971), 25–61

Copyright © 1971 by American Elsevier Publishing Company, Inc.

published in 1969 by Marr [11], in many ways extends and modifies the theory of Eccles *et al.*

The theory presented here was developed independently of the Marr theory but agrees with it at many points, at least in the early sections. This article, developed from a study of Perceptrons [15] and memory model cells [1], applies these results to the structure of the cerebellum as summarized by Eccles *et al.* [5]. The theory presented here extends the Marr theory and proposes several modifications based on principles of information theory. These extensions and modifications relate mainly to the role of inhibitory interneurons in the learning process, and to the detailed mechanism by which patterns are stored in the cerebellum.

## II. DATA

To credit each piece of information presented in this section to its original source would be very tedious. Everything in this section is taken directly either from Eccles *et al.* [5] or Fox *et al.* [7]. Therefore a single reference is now made to these sources and to the extensive bibliographies that appear in them.

### A. Mossy Fibers

Mossy fibers constitute one of the two input fiber systems to the cerebellum. Input information conveyed to the cerebellum via mossy fibers is from many different areas. Some mossy fibers carry information from the vestibular system or the reticular formation, or from both. Others carry information that comes from the cerebral cortex via the pons. The mossy fiber system that has been most closely studied relays information from the various receptor organs in muscles, joints, and skin. Mossy fibers that arrive via the dorsal spinal cerebellar tract are specific as regards modality of the muscle receptor organ, from either muscle spindles or tendon organs, and have a restricted receptor field, usually from one muscle or a group of synergic muscles.

Mossy fibers from the ventral spinal cerebellar tract are almost exclusively restricted to Golgi tendon organ information but are more generalized as regards specific muscles than those from the dorsal spinal cerebellar tract. The ventral tract fibers seem to signal stages of muscle contraction and interaction between contraction and resistance to movement of a whole limb. Other mossy fibers carry information from skin pressure receptors and joint receptors. There are continuous spontaneous discharges on most mossy fibers, at rates between 10 and 30 per second, even when the muscles are completely relaxed.

Mossy fibers enter the cerebellum and arborize diffusely throughout

*Mathematical Biosciences* 10 (1971), 25-61

the granular layer of the cortex. A single mossy fiber may send branches into two or more folia. These branches travel toward the top of the folia, giving off further branches into the granular layer of the sides of the folia, finally terminating in an arborization at the top of the folia. Each branch of a mossy fiber terminates in a candelabrum-shaped arborization containing synaptic sites called mossy rosettes. There is minimum distance of 80–100  $\mu$  between rosettes from a single mossy fiber. It is estimated that each branch of a mossy fiber entering the granular layer of the cerebellum produces from 20 to 50 or more rosettes. Thus a single mossy fiber may produce several hundred rosettes considering all its branches. The mossy rosettes are the site of excitatory synaptic contact with dendrites of the granule cells. The mossy fibers also send collaterals into the intracerebellar nuclei, where they make excitatory synaptic contact with nuclear cells.

### B. Granule Cells

The granule cells are the most numerous cells in the brain. It is estimated that in humans there are  $3 \times 10^{10}$  granule cells in the cerebellum alone. Granule cells possess from one to seven dendrites, the average being four. These dendrites are from 10 to 30  $\mu$  long and terminate with a characteristic claw-shaped ramification in the mossy rosettes. In view of the spacing between rosettes on a mossy fiber, it is highly unlikely that a granule cell will contact two rosettes from the same mossy fiber. Thus an average granule cell is excited by about four different mossy fibers. Since approximately 20 granule cell dendrites contact each rosette, this means that there are about five times as many granule cells as mossy rosettes, and at least 100–250 times as many granule cells as mossy fibers. Since a mossy fiber enters several folia, there may even be four or five times this many granule cells per mossy fiber.

Each granule cell gives off an axon, which rises toward the surface of the cortex. When this axon reaches the molecular layer, it makes a T-shaped branch and runs longitudinally along the length of the folia for about 1.5 mm in each direction. These fibers are densely packed and are only about 0.2–0.3  $\mu$  in diameter. The parallel fibers make excitatory synaptic contact with Purkinje cells, basket cells, stellate cells, and Golgi cells.

### C. Golgi Cells

Golgi cells have a wide dendritic spread, which is approximately cylindrical in shape and about 600  $\mu$  in diameter (see Fig. 1). This dendritic tree reaches up into the molecular layer, where it is excited by the parallel fibers, and down into the granular layer, where it is excited by the mossy

fibers. The Golgi axon branches extensively and inhibits about 100,000 granule cells located immediately beneath its dendritic tree. Every granule cell is inhibited by at least one Golgi cell. The Golgi axons terminate on the mossy rosettes, inhibiting granule cells at this point. Fox *et al.* [7] state that the axon arborizations of neighboring Golgi cells overlap extensively, so that two or more Golgi cells frequently inhibit a single granule cell. Note the overlapping fields shown in Fig. 3. This overlapping is a point of disagreement between Eccles *et al.* [5] and Fox *et al.* [7]. It appears, however, that Golgi cells must overlap, considering their size and that there are approximately 10% as many Golgi cells as Purkinje cells.

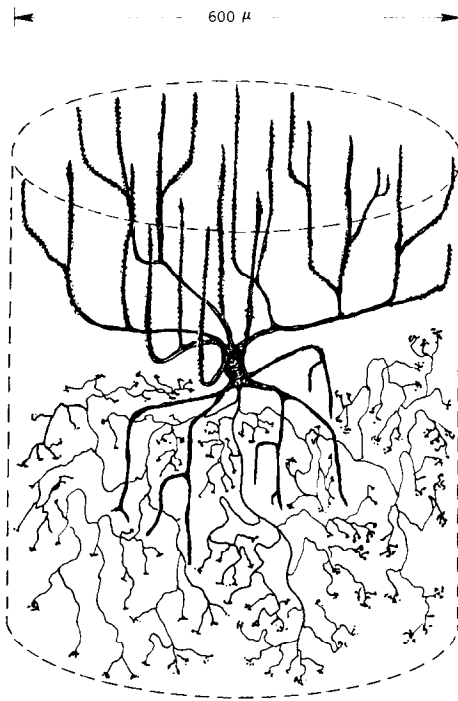


FIG. 1. A typical Golgi cell. Its arborizations extend throughout an approximately cylindrical volume  $600\ \mu$  in diameter.

The size of the dendritic spread of the Golgi cell as shown in Figs. 1 and 3 is a point of some uncertainty. Eccles *et al.* [5, page 205 and Fig. 116] state that the spread of the Golgi dendritic tree is about three times that of a Purkinje cell (i.e.,  $600\text{--}750\ \mu$ ). However, drawings by Cajal [2] and Jakob [10], and statements and drawings elsewhere in Eccles *et al.* [5, page 60 and Fig. 1] seem to indicate the dendritic spread for Golgi cells

to be only slightly larger than that of Purkinje cells (i.e.,  $250\text{--}300\ \mu$ ). However, even with a dendritic spread of only  $300\ \mu$ , the Golgi dendritic fields would still have significant overlap, as can be shown by drawing  $300\ \mu$  diameter circles around the Golgi cell bodies in Fig. 3.

#### D. *Purkinje Cells*

The Purkinje cell has a large and very dense dendritic tree. The dendritic tree of the Purkinje cell is shaped like a flat fan and measures on the average about  $250\ \mu$  across, about  $250\ \mu$  high, and only about  $6\ \mu$  thick, as shown in Fig. 2. The flat face of this fan is positioned perpendicular to the parallel fibres that course through the branches of the tree. It is estimated that around 200,000 parallel fibres pierce the dendritic tree of each Purkinje

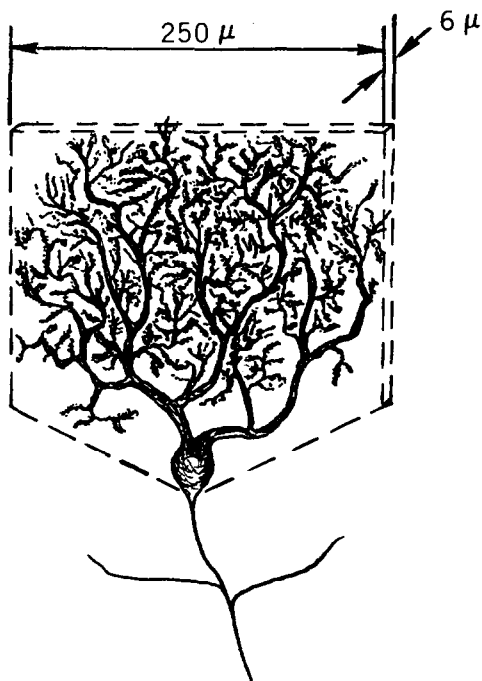


FIG. 2. A typical Purkinje cell. Its dendritic tree is restricted to a volume approximately  $250 \times 250 \times 6\ \mu$ .

cell, and that in passing virtually every parallel fiber makes a single synaptic contact with the dendrites of the Purkinje cell. At the site of a parallel fiber Purkinje dendritic synapse, the parallel fiber enlarges to about  $1\ \mu$  in diameter and is filled with synaptic vesicles. A spine grows out of the

Purkinje dendrite and is enclosed by an invagination of the enlarged part of the parallel fiber.

A unique characteristic of the Purkinje cell is that there is virtually no intermingling of its dendritic tree with that of other cells. The Purkinje cell bodies are beet shaped and about  $35\ \mu$  in diameter. They are scattered in a single layer over the cortex at intervals of about  $50\ \mu$  along the direction of the parallel fibers, and about  $50\text{--}100\ \mu$  in the transverse direction. Thus the fan-shaped dendritic trees overlap in the transverse direction but are offset in the longitudinal direction sufficiently so as to not intermingle. Figure 3 shows a top view looking down on the packed Purkinje dendritic trees. The trees are about  $6\ \mu$  thick and are separated by about  $2\text{--}4\ \mu$ . Thus a parallel fiber encounters a different Purkinje

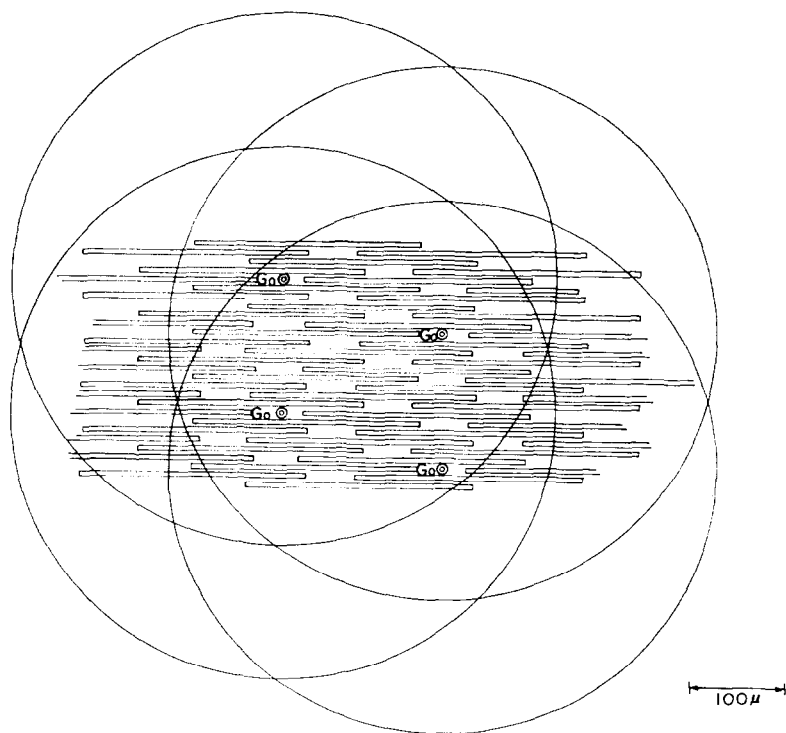


FIG. 3. View of cerebellar cortex looking down on top of Purkinje dendritic trees. Purkinje cells are shown here spaced approximately every  $50\ \mu$  in the longitudinal direction and every  $60\ \mu$  in the transverse direction. They are staggered so that the dendritic trees do not intermingle. Four Golgi cells are shown with the outline of their area of arborization traced. There is one Golgi cell to every nine Purkinje cells. Note the extensive overlapping of Golgi arborization. Each point on the cortex is subject to influence by about nine different Golgi cells.

dendritic tree every 8–10  $\mu$ . Since a parallel fiber synapses with virtually every Purkinje dendritic tree it passes, a 3-mm parallel fiber contacts about 300 Purkinje cells.

Purkinje cell axons constitute the only output from the cerebellar cortex. These axons make inhibitory synapses with the cells of the cerebellar nuclei and of the Deiters nucleus. In addition, Purkinje axons send recurrent collaterals to other Purkinje cells, basket cells, stellate cells, and Golgi cells.

#### E. *Basket Cells*

The basket cells also have flat fan-shaped dendritic trees, which extend upward in the 2–4  $\mu$  spaces between Purkinje dendritic layers. Basket dendritic trees are much less dense than those of Purkinje cells, but cover roughly the same area. Basket dendrites also receive excitatory synaptic contacts from parallel fibers via dendritic spines. Basket cell dendritic spines are much sparser, more irregularly spaced, longer, and thinner than Purkinje spines. They are very often hook shaped. Basket cell bodies, about 20  $\mu$  in diameter, are located in the lower third of the molecular layer. Basket cells are 15%–20% more numerous than Purkinje cells.

Basket cells send out axons in the transverse direction, perpendicular to the parallel fiber pathways. These axons branch and send descending collaterals, which makes strong inhibitory synapses around the preaxon portion of the Purkinje cells. They also send ascending collaterals into the Purkinje cell dendritic trees, where they form further inhibitory synapses. Each basket cell inhibits about 50 Purkinje cells over an elliptical area about 1000  $\mu \times 300 \mu$ . The basket cells do not inhibit the Purkinje cell immediately adjacent, but begin their inhibitory activity one or two cells away, and inhibit Purkinje cells out to about 1 mm away in the transverse direction. Thus any parallel fiber that excites a Purkinje cell is not likely to also inhibit the same Purkinje cell via a basket cell.

#### F. *Stellate Cells*

Stellate cells have dendritic arborization very similar to that of basket cells, although somewhat smaller. On the basis of axon distribution, there are two types of stellate cells. Stellate “a” cells send axons into Purkinje dendritic trees immediately adjacent, whereas stellate “b” cells send their axons transversely, making inhibitory contact with Purkinje dendrites in an area similar in size, shape, and relative position to that of basket cells. Functionally, the main distinction between basket cells and stellate “b” cells seem to be that stellate “b” cells are located higher in the molecular layer and send few, if any, axon collaterals to the Purkinje

preaxon, or “basket” region. However, there are many intermediate forms and the cell types seem to change progressively from basket cells in the upper granular layer to stellate “b” cells in the mid and upper molecular layer. Thus in this article the basket cells and stellate “b” cells will be assumed to perform roughly the same functions, which include receiving excitatory inputs from parallel fibers and transmitting inhibitory signals to Purkinje cells.

### G. Climbing Fibers

A second type of input fibers, the climbing fibers, also enters the cerebellum. These fibers are distinguished by the fact that each Purkinje cell receives a single climbing fiber in a 1 : 1 fashion. They are called climbing fibers because they contact the Purkinje cell at the base of its dendritic tree and climb up the trunk of the tree, making repeated strong excitatory synaptic contacts. A single spike on a climbing fiber can evoke a complex burst of Purkinje activity. The exact nature of this activity is not entirely clear. Observations by Thach [17] seem to indicate that this complex burst of activity consists of a single Purkinje axon spike followed by several milliseconds of spike-like activity propagating throughout the Purkinje dendritic tree. This dendritic activity is accompanied by intense cell depolarization and a pause in spontaneous Purkinje axon spike activity for 15–30 msec. This depolarization and pause was termed the *inactivation response* by Granit and Phillips [8].

The climbing fibers are usually thought to originate primarily in the inferior olivary nucleus and make a precise point-to-point mapping from the olivary nucleus to the cerebellar cortex. There is, however, some indication from cell counts done in the olivary nucleus [6], that either each climbing fiber branches about 15 times before reaching the cerebellum, or the majority of climbing fibers come from other sources outside the olivary nucleus.

Information carried by climbing fibers comes from a great variety of areas. The inferior olive receives afferents from proprioceptive end organs as well as all lobes of the cerebral cortex. The inferior olive also receives a strong projection from the red nucleus and the periaqueductal gray via the central tegmental tract.

The response of climbing fibers to peripheral stimulation is quite distinct from that of mossy fibers. A climbing fiber will typically respond to pinching the skin and deeper tissue anywhere within a receptive field, which may encompass an entire limb [17]. In monkeys performing a motor task it has been observed that climbing fiber spikes are correlated with quick movements made in response to external stimuli, but not with self-paced movements, such as rapidly alternating wrist motions [18, 19]. This



evidence would seem to indicate that information carried on climbing fibers is the product of a great deal of integration through higher centers.

In addition to the precise one-for-one climbing fiber contact with Purkinje cells, climbing fibers also put out three sets of collaterals; that is, (1) a climbing fiber sends collaterals to synapse on basket cells and stellate cells in the immediate vicinity of the Purkinje cell that it contacts; (2) a climbing fiber sends collaterals to one or more Golgi cells located within an elliptical region about  $1000 \times 300 \mu$  centered on the Purkinje cell that it contacts; (3) a climbing fiber sends collaterals to nuclear cells in the cerebellar nuclei and in the Deiters nucleus.

#### H. *Nuclear Cells*

The nerve cells of the cerebellar nuclei and Deiters nucleus are of at least two types. One type is large multipolar neurons, with relatively simple and irregular dendritic arborization. The axons from cells of the cerebellar nuclei go to the nucleus ventralis lateralis of the thalamus, to the red nucleus, to the pontomedullary reticular formation, and to the vestibular nuclei. Cells from the Deiters nucleus join the vestibulospinal tract. Thus some of these efferents send information toward the sensorimotor cortex, others toward the spinal motor neurons. The second type of nuclear neuron is smaller, with short axons, possibly a Golgi type II cell.

The cerebellar nuclei and Deiters nucleus cells receive excitatory inputs from climbing fiber collaterals and mossy fiber collaterals. They receive inhibitory inputs from Purkinje axons.

### III. PATTERN RECOGNITION AND THE PERCEPTRON

#### A. *The Classical Perceptron*

Since the neurophysiologist is usually not well versed in the field of pattern-recognition theory, a few short tutorial paragraphs concerning the pattern-recognition device known as the Perceptron are included to form a basis for arguments relating the cerebellum to the Perceptron. Again, rather than crediting all the many contributors to the theory of pattern-recognition and linear threshold devices, we refer the reader to the review books by Nilsson [14] and Minsky and Papert [13] for extensive references to the literature. These books contain mathematical proofs for most of the informal assertions made in following paragraphs.

The Perceptron developed by Rosenblatt [15] was inspired in large measure by known or presumed properties of nerve cells. In particular, a Perceptron possesses cells with adjustable-strength synaptic inputs of competing excitatory and inhibitory influences that are summed and compared against a threshold. If the threshold is exceeded, the cell fires.

If not, the cell does not fire. The original Perceptron was conceived as a model for the eye (see Fig. 4).

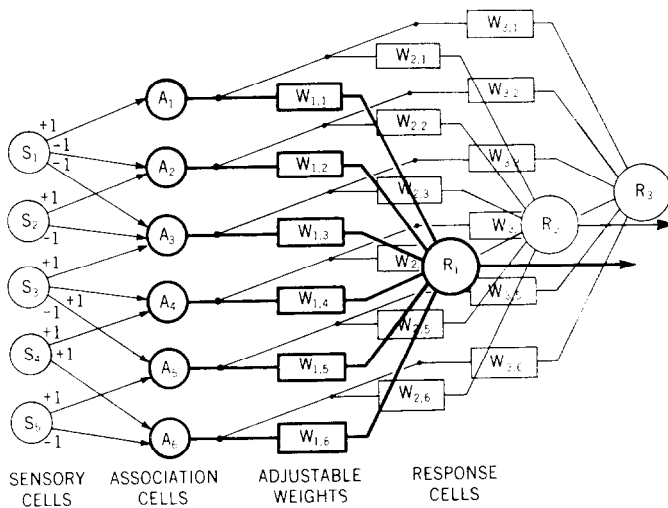


FIG. 4. Classical Perceptron. Each sensory cell receives stimulus either  $+1$  or  $0$ . This excitation is passed on to the association cells with either a  $+1$  or  $-1$  multiplying factor. If the input to an association cell exceeds  $0$ , the cell fires and outputs a  $1$ ; if not, it outputs  $0$ . This association cell layer output is passed on to response cells through weights  $W_{i,j}$ , which can take any value, positive or negative. Each response cell sums its total input and if it exceeds a threshold, the response cell  $R_j$  fires, outputting a  $1$ ; if not, it outputs  $0$ . Sensory input patterns are in class  $1$  for response cell  $R_j$  if they cause the response cell to fire, in class  $0$  if they do not. By suitable adjustment of the weights  $W_{i,j}$ , various classifications can be made on a set of input patterns.

Patterns to be recognized, or classified, are presented to a retina, or layer of sensory cells. Connections from the sensory cells to a layer of associative cells perform certain (perhaps random, perhaps feature-detecting) transformations on the sensory pattern. The associative cells then act on a response cell through synapses, or weights, of various strengths. The firing, or failure to fire, of the response cell performs a classification or recognition on the set of input patterns presented to the retina.

### B. Perceptron Learning

The Perceptron shows a rudimentary ability to learn. If a Perceptron is given a set of input patterns and is told which patterns belong in class  $1$  and which in class  $0$ , the Perceptron, by adjusting its weights, will gradually make fewer and fewer wrong classifications and (under certain rather restric-

tive conditions) eventually will classify or recognize every pattern in the set correctly. The weights usually are adjusted according to an algorithm similar to the following.

1. If a pattern is incorrectly classified in class 0 when it should be in class 1, increase all the weights coming from association cells that are active.
2. If a pattern is incorrectly classified in class 1 when it should be in class 0, decrease all the weights coming from association cells that are active.
3. If a pattern is correctly classified, do not change any weights.

Four features of this algorithm are common to all Perceptron training algorithms, and are essential to successful pattern recognition by any Perceptron-type device: (1) Certain selected weights are to be increased, others decreased. (2) The average total amount of increase equals the total amount of decrease. (3) The desired classification, together with the pattern being classified, governs the selection of which weights are varied and in which direction. (4) The adjustment process terminates when learning is complete.

The Perceptron works quite well on many simple pattern sets, and if the sensory-association connections are judiciously chosen, it even works on some rather complex pattern sets. For patterns of the complexity likely to occur in the nervous system, however, the simple Perceptron appears to be hopelessly inadequate. As the complexity of the input pattern increases, the probability that a given Perceptron can recognize it goes rapidly to zero. Alternatively stated, the complexity of a Perceptron required to produce any arbitrary classification, or dichotomy, on a set of patterns increases exponentially as the number of patterns in the set. Thus the simple Perceptron, in spite of its tantalizing properties, is not practical as a realistic brain model without significant modification.

### *C. The Binary Decoder Perceptron*

This lack of power of the conventional Perceptron can be overcome by replacing the sensory-association layer connections with a binary decoder, as shown in Fig. 5. It is then possible to trivially construct a Perceptron that will produce any arbitrary pattern classification. A binary decoder can be considered to be a recoding scheme that recodes a binary word of  $N$  bits into a binary word of  $2^N$  bits. This recoding introduces great redundancy into the resulting code. Each association cell pattern is restricted to a unique association cell in the 1 condition, all other association cells in the 0 condition.

However, a binary decoder Perceptron is seldom seriously considered

as a brain model for several reasons. First, the binary decoder requires such specific wiring connections that it is entirely too artificial to be imbedded in the rather random-looking structure of the brain. Second, the number of association cells increases exponentially as the number of inputs. Thus  $N$  input fibers require  $2^N$  association cells. Simple arithmetic thus eliminates the binary decoder Perceptron as a brain model.

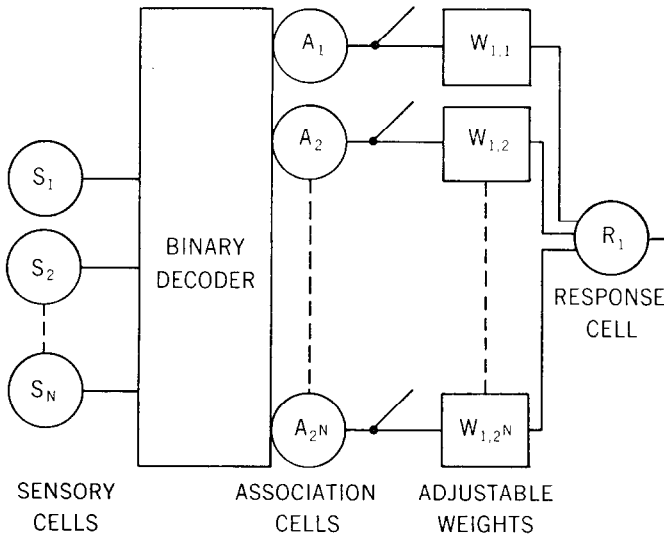


FIG. 5. Binary decoder Perceptron. Each association cell firing uniquely corresponds to one of the possible  $2^N$  input patterns. This type of Perceptron can perform any desired classification of input patterns. It has, however, no capacity for generalizing.

#### D. The Expansion Recoder Perceptron

However, there does exist a middle ground between a simple Perceptron and a binary decoder Perceptron. Assume a decoder, or rather a recoder, that codes  $N$  input fibers onto  $100N$  association cells, as shown in Fig. 6. Such a recoding scheme provides such redundancy that severe restrictions can be applied to the  $100N$  association cells without loss of information capacity. For example, it is possible to require that of the  $100N$  association cells, only 1% (or less) of them are allowed to be active for any input pattern. That such a recoding is possible without loss of information capacity is easily proven, for  $2^N \ll \binom{100N}{N}$ . That such a recoding increases the pattern-recognition capabilities of a Perceptron is certain, since the dimensions of the decision hyperspace have been expanded 100 times. The amount of this increase under conditions likely to exist in the nervous

system is not easy to determine, but it may be enormous. It can be shown that  $\binom{100N}{N} > 100^N$ . Thus  $2^N$  possible input patterns can be mapped onto  $100^N$  possible association cell patterns. If this is done randomly, the association cell patterns are likely to be highly dissimilar and thus easily recognizable. The ratio of  $100^N/2^N = 50^N$  rapidly increases as  $N$  becomes large.

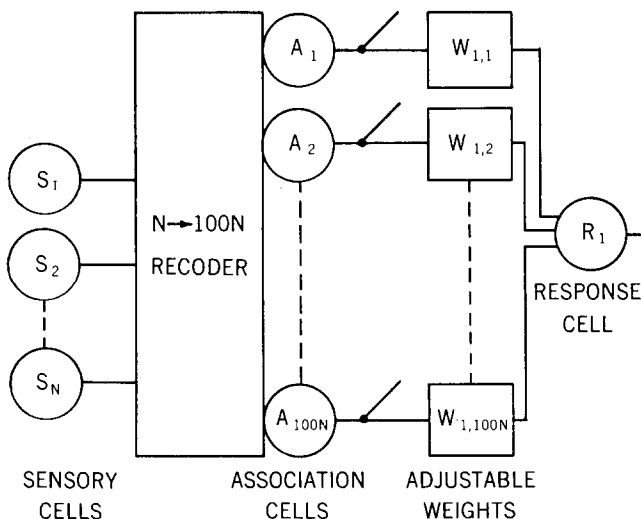


FIG. 6.  $N \rightarrow 100N$  Expansion recoder Perceptron. The association cell firing is restricted such that only 1% of the association cells are allowed to fire for any input pattern. This Perceptron has a large capacity and fast learning rate, yet it maintains the number of association cells within limits reasonable for the nervous system.

The restriction that only 1% of the association cells are allowed to be active for any input pattern means that any association cell participates in only 1% of all classifications. Thus its weight needs adjusting very seldom and there is a fairly good probability that its first adjustment is at least in the proper direction. This leads to rapid learning.

#### IV. THE CEREBELLUM AS A PERCEPTRON

##### A. Pattern Recoding in the Cerebellum

The granular layer of the cerebellum takes in information on mossy fibers and puts out information on parallel fibers. There are from 100 to 600 times as many parallel fibers as mossy fibers. Thus the granule cells can be said to be association cells that recode information from  $N$  inputs to at least  $100N$  outputs.

What can be said about the nature of this recoding? It was already noted that no granule cell receives more than one excitatory input from any one mossy fiber. It was also noted that the mossy rosettes from a single mossy fiber were widely distributed over several folia with a rather uniform random distribution. Thus, by the central limit theorem of probability, the distribution of granule cells with any given number of excitatory inputs will approach a Gaussian distribution with  $\sigma$  equal to the extent of the mossy rosette distribution. Since the mossy rosette distribution of each mossy fiber extends over several folia, the Gaussian curve will be flat, for all practical purposes, over regions large compared with a single folia, even more so compared with any individual cell.

Since virtually no granule cells are excited at two sites by the same mossy fiber, the relative abundance of granule cells simultaneously excited by  $n$  active mossy fibers will be proportional to  $1/n$ .

Thus at any instant the surface of the cerebellum should be dotted nearly uniformly randomly with granule cells whose input consists of one mossy fiber excitation. The surface of the cerebellum should also be dotted randomly, but less densely, with granule cells excited by two mossy fibers; and so on, progressively less densely with granule cells excited by three, and four, and five, up to seven mossy fibers. The total density of this dotting depends on the percentage of mossy fibers active.

The particular granule cells that actually fire as a result of various levels of mossy fiber excitation depend on the threshold levels of the granule cells. Only granule cells with enough excitatory inputs to exceed threshold will fire. This threshold for granule cells is regulated by Golgi cell activity.

The output of the granule cells is sampled by the Golgi cells via synapses with parallel fibers. This sampling is over an area approximately  $250\text{--}650\ \mu$  in diameter. Each Golgi cell feeds back inhibitory influences to about 100,000 granule cells. Neighboring Golgi cells overlap extensively in their dendritic fields and in their axon arborization. This very broad general feedback system suggests the function of an automatic gain control. Thus it is argued that the Golgi cells serve to maintain granule cell, and hence parallel fiber, activity fixed at a relatively constant rate. If few parallel fibers are active, Golgi inhibitory feedback decreases, allowing granule cells with lower numbers of excitatory inputs to fire. If many parallel fibers become active, Golgi feedback increases, allowing only those few granule cells with many active mossy inputs to fire.

The Golgi cells also have input from mossy fibers directly, a so-called feed-forward inhibition. This input tends to raise granule cell threshold levels when mossy fiber activity is large, and decrease granule thresholds when mossy fiber activity is small. This effect is also such as to stabilize the amount of parallel fiber activity.

To obtain a quantitative feel for what is occurring via these two types of Golgi cell inputs, consider Fig. 7. From the figure we can write

$$P = (M - I + Sp)Gr \quad (1)$$

$$I = (KM + P)Go \quad (2)$$

where  $P$  is the expected value of the spike rate for a parallel fiber,  $M$ , the expected value of the spike rate for a mossy fiber,  $I$ , the expected value of the spike rate for a Golgi cell,  $Gr$ , the average transfer gain of granule cells,  $Go$ , the average transfer gain of Golgi cells,  $K$ , the relative strength of mossy fiber input on Golgi cells to that of parallel fiber input, and  $Sp$ , the expected value of the spontaneous rate for a granule cell.

Combining (1) and (2) and differentiating with respect to  $M$  gives

$$\frac{dP}{dM} = \frac{Gr(1 - KGo)}{1 + GrGo}. \quad (3)$$

From Eq. (3) it can be seen that by proper adjustment of parameters (i.e.,  $KGo \approx 1$ ) it is possible to make  $P$ , the expected value of the spike rate for a parallel fiber, very nearly constant despite variations in mossy fiber input rate  $M$ .

It might not be unreasonable to assume values for  $Go$  and  $Gr$  as follows.

$$Gr = \frac{1 \text{ granule spike}}{1 \text{ mossy spike}} \times \text{divergence of } 100 = 100.$$

$$Go = \frac{1 \text{ Golgi spike}}{1000 \text{ parallel spikes}} \times \text{divergence of } 100,000 = 100.$$

These values substituted in (3) give

$$\frac{dP}{dM} \approx \frac{1 - 100K}{100}. \quad (4)$$

Thus if  $K \approx 0.01$  (i.e. 1 Golgi spike/10<sup>5</sup> mossy fiber spikes), the expected value of parallel fiber activity rate  $P$  is nearly constant. This, of course, does not mean that parallel fiber patterns would be independent of mossy fiber patterns, but merely that the overall level of activity (i.e., spikes per second) of parallel fibers could be constant in spite of what percentage, or at what rate, the mossy fibers are firing.

The mossy fiber inputs to Golgi cells probably also serve to stabilize parallel fiber rates under transient conditions. The feedback path via parallel fibers involves delays. The feed-forward path is undoubtedly faster acting. The net result of Golgi cell activity seems therefore to be to stabilize the level of parallel fiber activity to a nearly constant value under all conditions.

It will thus be hypothesized that the surface of the cerebellum is dotted randomly with active parallel fibers and that the density of this activity is

very nearly uniform, both spatially and temporally. It was noted earlier that if this density of parallel fiber activity is 1 % or less, patterns are easily recognized and quickly learned. Furthermore, a 1 % activity level is more than adequate from an information theory standpoint. Therefore, it will be further hypothesized that the density of parallel fiber activity is on the order of 1 %.

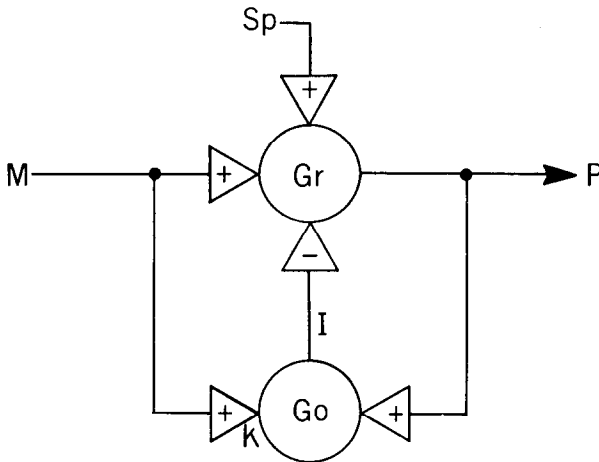


FIG. 7. Parallel fiber rate control circuit.  $M$ , expected value of mossy fiber input in spikes per second;  $P$ , expected value of parallel fiber output;  $I$ , expected value of Golgi cell rate;  $Sp$ , expected value of spontaneous granule cell rate;  $Gr$ , transfer gain of granule cell network;  $Go$ , transfer gain of Golgi cell network;  $K$ , relative strength of mossy fiber input on Golgi cells to that of parallel fiber input.

As was shown previously, recoding from  $N$  fibers to  $100N$  fibers, under the restriction that only 1 % of the output fibers are active for any input pattern, expands the number of possible patterns from  $2^N$  to about  $100^N$ , or an expansion of around  $50^N$ . In the cerebellum the number of input mossy fibers is approximately  $5 \times 10^4$  per square millimeter. Thus the pattern-expansion capacity of 1 mm<sup>2</sup> of cerebellar cortex is on the order of  $50^{50000}$ . Just what this means in increased pattern-recognition capability is unclear, but we get the feeling it is quite significant. This argument is even more compelling when it is realized that the mossy fiber system undoubtedly carries only a very restricted subset of the  $2^N$  (really  $R^N$  where  $R$  is the number of distinguishable levels of fiber firing rate) possible input patterns. Thus the recoding from  $N$  fibers to  $100N$  fibers may well produce an enormous increase in classification capability of cells in the cerebellum functioning as pattern-recognition response cells.

If this hypothesis of mossy fiber recoding by granule cells is correct, it implies that, to a neurophysiologist probing with an electrode, any



parallel fiber should appear to fire uncorrelated with neighboring parallel fibers, at least in an unanesthetized awake preparation.

An intuitive feel for why this recoding process is advantageous can be obtained from a simple example. Consider a Perceptron with only two association cells. There are then at most four different patterns of association cell firings. Suppose now it is desired for the response cell to fire whenever a sensory pattern occurs that produces an association cell pattern of 01 or 10, and it is desired for the response cell not to fire for any association cell pattern of 00, and 11. Try as we might it is impossible to find any combination of weights that can cause the response cell to have this behavior. It is rather simple to make the response cell fire on 01, and 10, and to not fire on 00. However, the 11 pattern creates a problem.

If, however, an expansion recoder is put between the sensory cells and the association cells, so that there are, for example, five association cells, the problem is much easier. The sensory pattern that previously produced the association cell pattern:

01 now might produce 00100;

10 now might produce 01001;

00 now might produce 10000;

11 now might produce 00010.

It is trivial to adjust weights so that association cell patterns 00100 and 01001 cause the response cell to fire, and the patterns 10000 and 00010 cause the response cell not to fire. The training procedure would consist of at the most one adjustment for each pattern.

A computer simulation of this type of recoding process has been run for a more complicated case. Twenty (20) mossy fibers were modeled. An expansion recoder of mossy rosettes, granule cells, and Golgi cells was modeled that transformed 20 mossy fiber firing rates into 2000 granule cell firing rates. Golgi cell feedback restricted the granule cells so that only about 1 % of them could fire. The result was that for two very similar mossy fiber patterns the granule cell firing patterns were similar in some respects but quite distinguishable in others. Some granule cells responded exactly the same for both mossy patterns, but other granule cells responded entirely differently. This implies that mossy fiber input patterns that would be very difficult to distinguish if put directly into a Perceptron response cell are easily distinguishable after passing through the pattern recoder.

#### B. *The Purkinje Response Cell*

It has been argued that the parallel fibers contain information coded in an ideal manner to serve as the input to a Perceptron response cell.

It will now be argued that the Purkinje cells serve a function similar to Perceptron response cells.

From a purely structural standpoint, the Purkinje cell certainly is related to granule cells very similarly to the way a Perceptron response cell is related to association cells. Each Purkinje cell has an enormous fan-in; each granule cell has a large fan-out. It is hard to conceive a more efficient parts layout for this type of circuit than the parallel fiber-Purkinje dendrite arrangement. A flat tree with input fibers piercing it at right angles creates the maximum possible fan-in for each Purkinje cell. The flat, closely stacked Purkinje dendritic trees allow the maximum possible fan-out for each parallel fiber. Any other arrangement would almost certainly decrease the ratio of computational elements to the brain tissue mass.

We may reasonably ask why this same structure does not exist in the cerebral cortex. The answer may well lie in the differences between the functions required of the cerebrum and of the cerebellum. The portion of the cerebral cortex that is best understood from a functional standpoint is the visual cortex. Here it is well known that a great amount of feature detection [9] takes place, such as line detection, edge detection, motion detection, and binocular correlation. Many of these transformations are translationally invariant over certain fields of view; that is, cells in the visual cortex respond to certain global features of the visual input irrespective of small changes in retinal coordinate position. It would appear, then, that in the cerebrum considerable feature-detection processing precedes, and perhaps is intermingled with, the expansion recoding circuitry. The geometrical requirements of translationally invariant global feature detection require elaborate plexuses of fibers crisscrossing in the cerebral cortex, and cells with their dendritic fields geometrically positioned to extract feature-dependent inputs from these fiber plexuses. Any pattern-recoding and pattern-recognition circuitry interspersed in this tangle would certainly be less compact and regular than that found in the cerebellar cortex.

On the other hand, in the cerebellum, granule cell receptive fields [17] show no evidence of feature detection analogous to that found in cerebral cortical cells. This is not too surprising since there should be no need for translationally invariant feature detection in a system that senses body conditions and controls motor commands. The problem of the cerebellum is merely to recognize patterns of information from proprioceptive receptors and to generate the appropriate motor command signals. The circuitry to do this is arranged as compactly as possible. The result is the beautiful regularity of the cerebellum.

Large portions of the cerebellum receive inputs from and project back

toward the cerebral cortex. Since the anatomy of this portion of the cerebellum is not appreciably different from the portion that interacts with the periphery, it is reasonable to assume that the transfer function is similar (i.e., a mossy fiber pattern input producing a Purkinje cell pattern output).

The nervous system has one constraint that does not exist in the Perceptron. In the nervous system a particular type of cell is either excitatory or inhibitory. Any single granule cell thus cannot be excitatory on one Purkinje cell and inhibitory on another. The basket and stellate b cells appear to provide a means of overcoming this deficiency. Basket and stellate b cells receive excitation from parallel fibers and inhibit Purkinje cells located transversely. This arrangement allows any parallel fiber to excite a number of Purkinje cells along its length, and to inhibit another group of Purkinje cells located on its flanks. As noted before, a parallel fiber is not likely both to excite a Purkinje cell directly and also to inhibit the same Purkinje via basket or stellate b cells. Thus, as shown in Fig. 8,

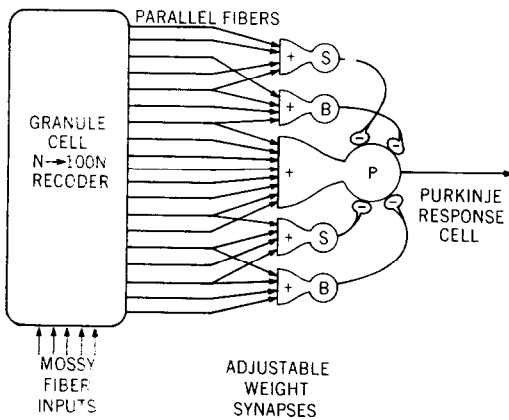


FIG. 8. Cerebellar Perceptron: *P*, Purkinje cell; *B*, basket cells; *S*, stellate b cells. Each Purkinje cell has inputs of the type shown.

the Purkinje cell looks very much like a Perceptron response cell. The only logical difference is that the inhibitory input to the Purkinje cell is collected and summed by flanking basket and stellate b cells before being relayed to the Purkinje cell. The inhibitory input of each basket and stellate b cell is also sent to many other Purkinje cells, but this fact is immaterial to any individual Purkinje. It is influenced only by the inputs it receives, not by the other places those inputs may go. In order to complete the analogy between Purkinje cells and Perceptron response cells, it is necessary to introduce adjustable synaptic strengths.

### C. The Hypothesis of Variable Synapses

The fundamental hypothesis of this article is that parallel fiber synapses are adjustable on both Purkinje cell dendrites and stellate and basket cell dendrites. The mechanism of change in both cases is hypothesized to be closely related to climbing fiber input activity. It will be argued that both excitatory and inhibitory influences on Purkinje cells are specifically modified under the control of climbing fiber activity patterns.

Each Purkinje cell is contacted by a single climbing fiber. In a conscious animal the climbing fibers fire in short bursts of one or more spikes at a rate of about 2 bursts/sec [5, 18]. Each climbing fiber burst causes a single spike on the Purkinje axon followed by a complex burst of spike-like activity in the Purkinje dendritic tree and intense depolarization of the Purkinje cell. The single axon spike is followed by a pause in the spontaneous Purkinje axon spike activity for 15–30 msec. This pause, accompanied by intense depolarization, was first observed by Granit and Phillips [8] and was termed the *inactivation response* to distinguish it from a normal pause in activity resulting from hyperpolarization. After the 15- to 30-msec inactivation response, the cell gradually recovers its spontaneous firing rate over a period of 100–300 msec [3]. As it approaches normal, the cell becomes once again responsive to parallel fiber input activity.

It is now hypothesized that the inactivation response pause in Purkinje spike rate is an unconditioned response (UR) in a classical learning sense caused by the unconditioned stimulus (US) of a climbing fiber burst. It is further hypothesized that the mossy fiber activity pattern ongoing at the time of the climbing fiber burst is the conditioned stimulus (CS). If this is true, the effect of learning should be that eventually the particular mossy fiber pattern (CS) should elicit a pause (CR) in Purkinje activity similar to the inactivation response (UR) that previously had been elicited only by the climbing fiber burst (US). In order to accomplish this result it is necessary to postulate that the climbing fiber input to the Purkinje cell not only causes the Purkinje cell to pause momentarily but also *weakens* any parallel fiber synapses that are tending to cause the Purkinje to fire during the inactivation response.

A possible mechanism for such weakening might be that there exists a critical interval near the end of the inactivation response after the effect of the climbing fiber burst has worn off sufficiently so that the cell can be fired by parallel fiber input but before the dendritic membrane has returned completely to normal. If the Purkinje cell fires in this interval, this firing is an error signal that signals every active parallel fiber synapse to be weakened.

The amount of weakening of each synapse is proportional to how

*Mathematical Biosciences* 10 (1971), 25–61

strongly that synapse is exciting the Purkinje cell at the time of error signal. The effect of this mechanism would be to train the Purkinje cell to pause at the proper times, that is, at climbing fiber burst times. After learning is complete, the Purkinje knows when to pause because it recognizes the mossy-parallel fiber pattern that occurred previously at the same time as the climbing fiber burst. Later, since each parallel fiber active synapse was weakened by the error signal, if the same mossy-parallel fiber pattern occurs again, the Purkinje will pause even without the climbing fiber burst. Thus, the Purkinje is forced to perform in a certain way by the climbing fiber teacher. After learning is complete, however, it behaves in that same way, under the same mossy fiber conditions, even in the teacher's absence.

Note that this mechanism corresponds closely with the Perceptron training algorithm in that (1) if the response cell fires (or tends to fire) when it should not fire, then all synapses coming from active parallel fibers will be decreased or weakened; (2) if the response cell does not fire improperly, no adjustments are made.

It is now possible to consider many climbing fibers, each firing at different rates in some spatial pattern  $C_1$  at time  $t_1$ . This climbing fiber firing pattern will elicit a Purkinje firing pattern  $C'_1$ . Assume at time  $t_1$  the mossy fibers have some firing pattern  $M_1$ . Each climbing fiber will train its respective Purkinje cell (or cells) to recognize the mossy fiber input pattern  $M_1$  that was present when  $C_1$  occurred. If during training  $M_1$  on the mossy fibers occurs in coincidence with  $C_1$  on the climbing fibers, after training the occurrence of  $M_1$  on the mossy fibers will elicit  $C'_1$ , from the Purkinje cells whether or not  $C_1$  appears on the climbing fibers. It can then be said that climbing fiber pattern  $C_1$  has been imprinted, or stored, on mossy fiber pattern  $M_1$ . In the same way a second climbing fiber firing pattern  $C_2$  can be stored on another mossy fiber pattern  $M_2$ , and so on.

An important feature of this hypothesis is that the  $C'_i$  patterns coming out of the Purkinje cells are not necessarily binary patterns;  $C'_i$  represents the relative rates of firing of all the Purkinje cells. Thus relative patterns are stored and relative patterns are recalled.

#### D. *Variable Inhibitory Synapses*

Since variation of parallel fiber Purkinje cell synapses is sufficient to cause patterns to be stored in the cerebellum, we might well suggest [11] that no further mechanism of variable inhibitory synapses is necessary. However, there are good reasons to further hypothesize variable inhibitory synapses.

First, if only the excitatory inputs to a cell are caused to decrease, while

the inhibitory inputs are held fixed, eventually the cell fails to fire in response to any input pattern. Second, a pattern-recognition device based on only excitatory weight adjustment has inherently low capacity. Marr [11] estimates that a Purkinje cell capable of only excitatory synaptic adjustment has the capacity to make about 200 mossy fiber pattern dichotomies. However, a Perceptron with both positive and negative weight adjustments has the capacity to make about twice as many dichotomies as there are adjustable weights [4]. Thus, if both excitatory and inhibitory synapse adjustment is possible in the cerebellum, each Purkinje cell would have the capacity to make on the order of 200,000 pattern dichotomies. The adjustment of inhibitory weights thus results in a thousandfold increase in recognition capacity. Third, any pattern-recognition system capable of varying weights in only one direction is necessarily very slow to learn. An example of the learning difficulties encountered by such a system can be seen by referring to Fig. 4. Assume a pattern  $M$  causes only association cell  $A_2$  to fire. This will affect the response cell  $R_1$  through weight  $W_{1,2}$ . Four possible situations can exist when pattern  $M$  is first presented:

- case 1  $M$  desired in class 1,  $R_1 = 1$ ;
- case 2  $M$  desired in class 1,  $R_1 = 0$ ;
- case 3  $M$  desired in class 0,  $R_1 = 1$ ;
- case 4  $M$  desired in class 0,  $R_1 = 0$ .

In case 1 and case 4,  $M$  is already in the proper class and no adjustment of weights is necessary. In case 3, the weight  $W_{1,2}$  needs to be decreased so as to force the  $R_1$  cell below threshold. In case 2, the weight  $W_{1,2}$  needs to be made more positive so as to raise the  $R_1$  cell above threshold. If such a positive adjustment is not allowed, another means is available. All the weights to  $R_1$  except  $W_{1,2}$  can be decreased, and the threshold of the  $R_1$  cell somehow decreased accordingly. This would have the same result as an increase in  $W_{1,2}$ . As a mechanism likely to occur in the cerebellum, however, this scheme has several serious difficulties:

1. Decreasing all weights except one is cumbersome. It is inconceivable to decrease 199,999 weights in order to increase 1.
2. It is very difficult to suggest a mechanism with such abilities. The mechanism must, in case 3, decrease the synaptic strength of all active parallel fibers, but in case 2, decrease the synaptic strength of all except the active parallel fibers.
3. If the threshold of the  $R_1$  cell is to be lowered along with all the weights except  $W_{1,2}$ , this in itself implies that variable inhibitory synapses are necessary in the cerebellum.

4. If basket and stellate cells have no variable synapses, it is hard to imagine why they are so numerous, or what is the purpose of their peculiar

axon distributions. If these inhibitory interneurons merely serve the purpose of general threshold regulators, it would seem that a few cells should do as well. For example, only a few Golgi cells are necessary to set general threshold levels for an enormous number of granule cells. Yet there are about twice as many basket and stellate b cells as Purkinje cells. Surely these cells have a more sophisticated function than general threshold regulation. Variable inhibitory synapses could explain why basket and stellate cells are so numerous.

#### *E. Site of Inhibitory Synaptic Change*

Inhibitory synaptic strength variation could occur at two sites. One site is where basket and stellate b cells synapse on the Purkinje cells. This is perhaps an obvious first candidate. However, the amount of convergence is small. Certainly less than 1000 different basket and stellate b cells synapse on each Purkinje. The actual figure is probably less than 100. This is a far cry from the parallel fiber convergence of about 200,000 variable excitatory synapses. The addition of 100 variable inhibitory synapses would seem to add little to the recognition capacity of the Purkinje cell.

The second site where inhibitory inputs to Purkinje cells might be varied is at the parallel fiber synapses on basket and stellate b dendrites. A decrease in strength of the excitatory parallel fiber synapses on basket and stellate b cells results in a decrease in inhibitory input to the related Purkinje cells. The basket and stellate b dendritic trees are sparser than those of Purkinje cells, but they do contact perhaps 5% of the parallel fibers coursing through them. When account is taken of the fact that about 100 of these cells then synapse on a single Purkinje, the result is a convergence of variable inhibitory inputs to the Purkinje cell of the same order of magnitude as that of variable excitatory inputs. Thus the Purkinje recognition capacity is on the order of 200,000 patterns rather than 200 patterns as suggested by Marr [11].

It is interesting that lower forms, such as frogs, have no basket cells. A cerebellar Perceptron with no variable inhibitory weights is certainly possible. Its only shortcoming would be a very limited capacity for discrimination.

Several other facts support the hypothesis that the parallel fiber synapses on basket and stellate b cells are the sites of variable inhibitory weights. First, the basket and stellate cells contact the parallel fibers with dendritic spines similar to those of the Purkinje cells. Second, each climbing fiber, in addition to synapsing strongly on a single Purkinje cell, also sends collaterals, which synapse on the soma of adjacent basket and stellate cells. Since the climbing fiber input is assumed to be intimately

related with varying parallel fiber synapses on Purkinje cells, it is perhaps reasonable to suggest that the same climbing fiber may also vary parallel fiber synapses on basket and stellate cells. The mechanism of variation could be identical or at least very similar. In other words it is argued that on every cell contacted by an active climbing fiber, each active parallel fiber synapse is weakened by the same mechanism regardless of whether the cell is Purkinje, basket, or stellate b. This hypothesis has the elegant feature that a single event causes a change in both excitatory and inhibitory influences. The fact that climbing fibers do not contact dendrites of basket and stellate cells may be accounted for by the fact that their dendritic arborization is less extensive than that of Purkinje cells.

In order to satisfy the Perceptron training conditions that excitatory and inhibitory changes be equal on the average, it is merely necessary to assume that the size of the decrement in each synapse is such that the expected value of the excitatory change be equal to the expected value of the inhibitory change.

#### *F. Pattern Storage on Excitatory and Inhibitory Synapses*

The effect in terms of pattern storage of this scheme can be seen by referring to Fig. 9. Assume the climbing fiber firing pattern  $cf_1 = 1, cf_2 = 0$  occurs. In this case  $P_1$  pauses and  $P_2$  is released from inhibitions by  $B_1$  pausing. Further, assume a mossy fiber pattern occurs such that  $Pf_1 = 1, Pf_2 = 1$ . The coincidence of these two patterns will tend to decrease weights  $W_{P_1}$  and  $W_{B_1}$  but leave unchanged  $W_{P_2}$  and  $W_{B_2}$ . At a later time when the climbing fibers are silent,  $cf_1 = cf_2 = 0$ ; if the same mossy fiber pattern recurs such that  $Pf_1 = Pf_2 = 1$ ,  $P_1$  will pause because of decreased  $W_{P_1}$  and  $P_2$  will be disinhibited because of decreased  $W_{B_1}$ . Thus, the original climbing fiber response,  $P_1$  pause,  $P_2$  disinhibited, can be recalled by the mossy fiber pattern, which causes  $Pf_1 = Pf_2 = 1$ . It can thus be said that the climbing fiber pattern is imprinted on the mossy fiber pattern.

Note that all the adjustment of the variable synapses takes place in the immediate vicinity of the Purkinje cell excited by an active climbing fiber, even though the disinhibitory effects are felt by Purkinje cells far removed in the transverse direction.

In order to satisfy the requirement that the expected value of the change in excitation equals the expected value of the change in inhibition it is necessary to assume some things concerning the relative amount by which  $W_{P_1}$  and  $W_{B_1}$  are changed. The synapse of  $Pf_1$  on  $P_1$  occurs with a probability of nearly 1. The synapse of  $Pf_1$  on  $B_1$  occurs with a probability of around 0.05 or less. However, the effects of  $W_{B_1}$  are distributed to 30–50 Purkinje cells, whereas the effects of  $W_{P_1}$  are confined to one



Purkinje cell. In addition, the strength of  $W_{B1}$  is multiplied by a gain factor governed by the strength of the basket cell synapses on Purkinje cells. Since this is a rather strong synapse, the gain factor is probably greater than 1. Thus in order for the total average decrease in excitation to equal the total average decrease in inhibition, the following equation must be satisfied.

$$\Delta W_{B1} \times P_{B1}(Pf_1) \times D_{B1} \times G_{B1} = \Delta W_{P1} \times P_{P1}(Pf_1) \quad (5)$$

where  $\Delta W_{B1}$  is the change in  $W_{B1}$ ,  $\Delta W_{P1}$  the change in  $W_{P1}$ ,  $P_{B1}(Pf_1)$  the probability  $B_1$  contacts  $Pf_1$ ,  $P_{P1}(Pf_1)$  the probability  $P_1$  contacts  $Pf_1$ ,  $D_{B1}$  the number of Purkinje cells  $B_1$  contacts,  $G_{B1}$  the strength of  $B_1$  synapse in Purkinje cells.

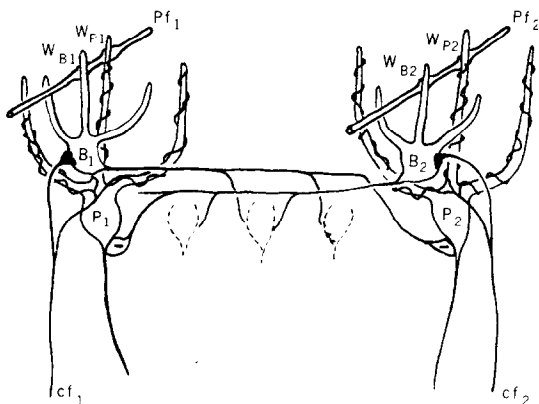


FIG. 9. Climbing fiber input. Each climbing fiber contacts a single Purkinje cell and several nearby basket cells or stellate cells, or both. If  $Pf_1$  is active when  $P_1$  or  $B_1$ , or both, fire in the critical interval during a  $cf_1$  inactivation response, then  $W_{P1}$  or  $W_{B1}$ , or both, are altered. This change in synaptic strength can later be read out in the form of Purkinje postsynaptic potentials by firing  $Pf_1$  again.

Everything considered, it is likely that  $\Delta W_{B1}$  is less than  $\Delta W_{P1}$ . This judgment seems to be supported by the experimental fact that the effect of a climbing fiber on a basket cell is less strong than on a Purkinje cell [5]. Presumably a smaller climbing fiber effect produces less synaptic weakening.

This cerebellar system now has most of the characteristics of a Perceptron; that is, it corrects errors by adjusting weights positively and negatively; the average total increase equals the average total decrease; the pattern being stored, in coincidence with the pattern on which it is stored, governs which weights are increased and which are decreased; and the adjustment procedure terminates as learning asymptotically approaches completion. In addition, the hypothesized cerebellar system exhibits the

capacity to store information concerning the relative firing rates of climbing fiber patterns.

#### *F. Defense of the Synaptic Weakening Argument*

The argument synaptic weights are weakened by learning rather than strengthened is counterintuitive and contrary to most, if not all, theories of synaptic learning that have appeared in the literature. Thus it perhaps should be examined in more detail. There are three main reasons why synaptic weakening rather than strengthening is hypothesized to take place in the cerebellum. First, the experimental data that are available seem to suggest it. Climbing fiber inputs cause Purkinje cells to pause. If the Purkinje is to learn to pause, parallel fiber excitation must be decreased. Second, Perceptron theory proves that the most effective training algorithms are error correcting in nature. Thus, firing at erroneous times should reduce the tendency to fire again. Firing at the proper times requires no adjustment. This algorithm implies weakening of synapses that contribute to erroneous firings. It is possible to conceive an error-correcting scheme that would operate by strengthening synapses but the mechanism seems quite unlikely. There are only two possible error conditions: (1) Cell fires when it should not. This condition can be corrected by weakening erroneous excitatory synapses (as suggested) or by strengthening erroneous inhibitory synapses. On the Purkinje cell the excitatory spine synapses seem much more likely candidates for variability than the inhibitory synapses. There are relatively few inhibitory synapses. Learning capacity would be quite low if on the Purkinje the inhibitory synapses rather than the excitatory were the site for variability. (2) Cell does not fire when it should. This condition can be corrected by strengthening erroneous excitatory synapses or by weakening erroneous inhibitory synapses. In this case it is difficult to suggest how the individual synapses know when an error has occurred. The absence of postsynaptic cell firing may be the correct response as far as each synapse knows. An additional piece of information is needed—the information that an error has occurred. It is difficult to imagine how this information is conveyed to synaptic sites in the absence of postsynaptic activity. Thus, if the Purkinje cell learns by error correction, the most probable mechanism is synaptic weakening in the presence of erroneous firing.

The third reason synaptic weakening is hypothesized to occur in the cerebellum is that there are serious stability problems of learned responses under conditions of overlearning if synaptic activity causes synaptic facilitation. Consider Fig. 10:  $C_1$  and  $C_2$  are climbing fibers synapsing with synapses of fixed strength on Purkinje cells  $P_1$  and  $P_2$ . A parallel fiber pf synapses on  $P_1$  and  $P_2$  with variable-strength synapses of weights

$W_1$  and  $W_2$ . If it is now assumed that the synaptic weights are strengthened by coincidence of pre- and postsynaptic activity, it is possible to write

$$\Delta_i W_1 = f_{P1} \cdot f_{pf} \quad \text{at } t = i \quad (6)$$

where  $\Delta_i W_1$  is the increase in  $W_1$  at time  $t = i$ ,  $f_{P1}$  the frequency of spikes on  $P_1$ , and  $f_{pf}$  the frequency of spikes on pf. Let  $W_1$  originally equal  ${}_0W_1$ . As learning takes place, the following situation obtains. At

$$\begin{aligned} t = 0, \quad f_{P1} &= kf_{C1} + {}_0W_1 f_{pf}; \\ t = 1, \quad f_{P1} &= kf_{C1} + ({}_0W_1 + \Delta_0 W_1) f_{pf}; \\ t = 2, \quad f_{P1} &= kf_{C1} + ({}_0W_1 + \Delta_0 W_1 + \Delta_1 W_1) f_{pf}; \\ t = 3, \quad f_{P1} &= kf_{C1} + ({}_0W_1 + \Delta_0 W_1 + \Delta_1 W_1 + \Delta_2 W_1) f_{pf}; \\ &\vdots \\ &\vdots \\ &\vdots \end{aligned}$$

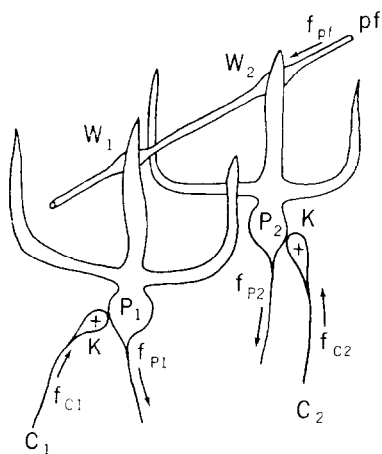


FIG. 10. Two Purkinje cells contacting the same parallel fiber.

We can readily see that the weight  $W_1$  continuously increases at each learning interval. In fact, since  $\Delta_i W_1$  is the product of  $f_{P1} \cdot f_{pf}$ , and since  $f_{P1}$  increases during each learning interval,  $\Delta_0 W_1 < \Delta_1 W_1 < \Delta_2 W_1 < \dots$ . Therefore  $W_1$  grows at an exponential rate, and of course so does  $f_{P1}$ . Certainly  $W_1$  must eventually saturate. Now suppose that during the same learning sequence a spike train also appears on  $C_2$  at half the frequency of that on  $C_1$ :

$$f_{C2} = \frac{1}{2} f_{C1}.$$

Until  $W_1$  saturates,

$$W_1 \simeq 2W_2.$$

Eventually, however,

$$W_1 = W_2 = \text{saturation value.}$$

Thus, after a sufficiently long period, all parallel fiber synapses will eventually become saturated. The very active ones will saturate first, but over a long time virtually every synapse will saturate. Synaptic facilitation suggests learning is exponential. Synaptic weakening suggests learning is asymptotic.

This problem could possibly be averted by proposing some sort of decay rate for all synaptic strengths. Thus synaptic strengths would not remain saturated. However, such a mechanism would need to be very exotic to prevent continued learning from degrading performance and, at the same time, to preserve learned patterns over long time periods. It is common experience that memories of motor skills are preserved rather well over periods of many years. It is also common experience that repeated practice of motor skills leads to improved motor performance, even when the practice sessions are intensive and of short duration (on the order of minutes or hours). It is difficult to conceive of a decay system that could preserve memory over periods of years and at the same time prevent saturation over periods of minutes.

It is an obvious fact that continued training in motor skills improves performance. Extended practice improves dexterity and the ability to make fine discriminations and subtle movements. This fact strongly indicates that learning has no appreciable tendency to saturate with over-learning. Rather, learning appears to asymptotically approach some ideal value. This asymptotic property of learning implies that the amount of change that takes place in the nervous system is proportional to the *difference* between actual performance and desired performance. A difference function in turn implies error correction, which requires a decrease in excitation upon conditions of incorrect firings.

This argument is not meant to suggest that synaptic facilitation does not occur anywhere in the nervous system. In fact the stellate cells will shortly be conjectured to undergo synaptic facilitation. Synaptic facilitation very probably plays an important role in many places in the nervous system. However, in situations where saturation would degrade performance, and particularly in the cerebellar cortex, where other evidence points to weakening, synaptic weakening seems very likely to be the principle learning mechanism. It might be argued that the saturation argument holds equally well in the opposite sense, that is, that all synapses would eventually be reduced to zero. One answer to this is that the synaptic strengths tend toward zero asymptotically. Therefore the weaker a synapse becomes, the less is its contribution to any erroneous firings and the less it is weakened by any correction. Another answer is that new variable spiny synapses may be hypothesized to spontaneously and randomly grow and mature into active efference synapses. The result of

this would not be to destroy learning but to mask it over a period of time by background noise. To clarify this point, no synapse that has undergone any decrementing is hypothesized to grow back in strength. However, new synapses are hypothesized to grow to full size and then mature into an effective state. From this point they are then decremented, perhaps all the way to zero. There may be some evidence for such a phenomenon in the visual cortex of the mouse. Ruiz-Marcos and Valverde [16] note that the density of spines on pyramidal cells in mouse visual cortex rises to a maximum shortly after the mouse opens its eyes. From that time the density of spines decreases asymptotically to a smaller value. Light deprivation considerably reduces the spine density. This might suggest that spines develop randomly under tropic influence of presynaptic nerves and are specifically decremented in the process of learning.

#### *G. Response Speedup via Stellate a Cells*

The notion that occurrence of a particular mossy fiber pattern causes a decrease in excitation of Purkinje, basket, and stellate b cells, and that this decrease in excitation causes the proper response of the Purkinje cell, raises a question of response speed. The decrease in excitation resulting from a decay of synaptic transmitter substance is not generally considered to occur as quickly as a buildup of excitation resulting from release of transmitter substance. Thus a system that operates solely on decay of excitation may lack the speed necessary for quick movements. It will now be suggested that stellate a cells are ideally situated for providing a speedup mechanism.

The main structural difference between stellate a and stellate b cells is in their axon arborization. The stellate a cells send synaptic contacts to Purkinje cells in their immediate vicinity and to adjacent Purkinje cells in the longitudinal direction. Thus it is quite likely for a parallel fiber to excite a particular Purkinje cell and to inhibit the same Purkinje via a stellate a cell. Climbing fiber collaterals also contact stellate a cells. Thus, following the same reasoning used for Purkinje, basket, and stellate b cells, it is not unreasonable to assume that coincidence between climbing fiber and parallel fiber activity effects a change in synaptic strength of stellate a cells also. It would seem, however, that in order to perform a useful function, the synaptic change in this case should be a strengthening rather than a weakening. It will be conjectured that coincidence of a climbing fiber spike with parallel fiber activity on a stellate a cell will cause an *increase* in the synaptic strength of the parallel fiber–stellate a cell synapse. Thus the stellate a synapses are conjectured to change in the opposite direction from all the other variable synapses under the same coincidence conditions.

Consider parallel fiber pattern  $M_1$  to be imprinted positively on stellate a cells, but negatively on an immediately adjacent Purkinje cell. Occurrence of pattern  $M_1$  causes the Purkinje cell to receive less excitation. Pattern  $M_1$  causes the stellate a cell to receive more excitation, and hence actively inhibit the Purkinje. The result would be an increase in speed of the Purkinje cell response.

The stellate a cell variable synapses would of course be subject to the saturation problem discussed previously. However, if the stellate a contribution to the Purkinje input were small compared to the other inputs from basket and stellate b cells and parallel fibers, the saturation effect would be small in the steady state. The stellate a input would be significant only in the first few milliseconds following a transient. In this interval the stellate a cell would get the Purkinje response going in the proper direction. Later the other inputs to the Purkinje would predominate to set the proper final value. The same effect would obtain if the stellate a response were not necessarily small but merely of short duration.

Note that in the arguments concerning stellate a cells the word conjecture was used rather than hypothesis. Very little is known concerning the behavior of stellate a cells and any confident prediction concerning their function is certainly premature. Stellate a cells may have nothing at all to do with memory or variable synapses. In the next section it is suggested that perhaps stellate a cells may have rather to do with attention mechanisms.

#### *H. The Function of Recurrent Purkinje Collaterals*

The fact that the cerebellum is spontaneously active allows it to achieve a high degree of sensitivity and precision. A spontaneously active system is essentially linear, at least for small inputs. Thus any small input will produce an output whose size will depend on both the size of the input and the gain of the system. If the system is not spontaneously active, small signals do not have any effect on the output until they exceed a certain threshold. This is usually not a desirable trait for a feedback control system.

As was discussed earlier, the mossy fiber  $\rightarrow$  granule cell  $\rightarrow$  Golgi cell interconnection network appears to work so as to maintain granule cell activity at some relatively constant level. In addition, the Purkinje cell axons put out recurrent collaterals that are known to contact Golgi cells, basket cells, and other Purkinje cells. These Purkinje recurrent collaterals send inhibitory impulses over a wide-ranging area, even into adjacent folia. The Purkinje recurrent collateral synapses on other Purkinje cells have the effect of maintaining the average Purkinje cell activity fixed at a relatively constant level over the entire cortex. If the average Purkinje activity rises

too high, the inhibitory effect of the recurrent collaterals drives it back down. If Purkinje cell activity drops too low, the decrease in inhibition will let it rise again. Thus a relatively constant spontaneous discharge rate will be maintained despite rather large variations in cell conditions, such as nutrition or fatigue.

Another effect of the recurrent collateral inhibition on Purkinje cells is the contrast enhancement effect of lateral inhibition. Thus any local increase in activity will be accompanied by a surrounding field of depressed activity. There also appears to be some specific contralateral inhibition produced by Purkinje recurrent collaterals.

The existence of Purkinje recurrent collateral synapses on Golgi cells is very interesting. The effect is that of both positive and negative feedback since the affected parallel fibers both excite the Purkinje cells directly and inhibit them via basket and stellate cells. The total effect may be that when a general area of the cerebellar cortex is actively engaged in processing information, the Golgi cells limiting the input to that area are suppressed, thus allowing input to that area more free access. This would then constitute a crude form of attention mechanism. Any area actively engaged in processing information would be given priority over other areas that are inactive at the time. This of course is quite speculative, but a rather pregnant possibility.

The function of Purkinje recurrent collateral synapses with basket cells is not clear. The effect is certainly that of positive feedback. Positive feedback is commonly used in electronic circuitry to produce one or the other of two effects: either oscillatory behavior or bistable switching behavior. There is no evidence of any oscillatory effects in the cerebellum that are likely to be mediated by Purkinje recurrent collaterals. There is, however, a curious bistable effect in the firing rate of Purkinje cells that may be caused by the Purkinje recurrent collateral interaction with the various interneurons. Although a Purkinje cell sometimes is spontaneously active, at other times the same cell is completely quiet except for climbing fiber responses. This rather implies that Purkinje cells have at least two stable states, one spontaneously active, the other completely silent. The transition between states seems to be somewhat correlated with climbing fiber activity [3]. We might speculate that certain parts of the cerebellum are switched on by an attention mechanism when they are needed, and switched off again when they are not in use. The Purkinje collateral-basket cell or Golgi cell circuit may provide the positive feedback necessary to switch between states. Specific climbing fiber patterns could provide the trigger signal to initiate the switching. Climbing fiber inputs to Golgi cells may be the means by which climbing fibers trigger Purkinje cells into an active state. Climbing fiber inputs to stellate a (or basket and





The firing rate of the nuclear cell, which is also spontaneously active, is given by

$$f_N = f_c k_{cN} - f_P k_{PN} + f_{m1} k_{mN} + f_{0N} \quad (8)$$

where  $f_{0N}$  is the spontaneous firing rate of the nuclear cell and  $k_{cN}$  is the climbing fiber input–nuclear cell output transfer function. Substitution of (7) in (8) gives

$$f_N = f_c(k_{cN} - k_P) + f_m k_{mN} + X_i(f_{m1}, \dots, f_{mN}) + f_0 \quad (9)$$

where  $k_P$  is the combined effect of  $k_{PN}$  and  $k_{cP}$  and  $f_0$  is the combined effect of  $f_{0P}$  and  $f_{0N}$ .

Several interesting observations can be made from Eq. (9). First, the output of the nuclear cell is directly affected by mossy fiber input. Thus the nuclear cell may be part of a reflex arc. Second, the strength of this reflex arc is modulated by patterns arriving on the mossy fibers corresponding to patterns previously stored by climbing fibers. Third, the effect of climbing fiber activity  $f_c$  on the nuclear cell depends on the factor  $(k_{cN} - k_P)$ ;  $k_P$  is a negative quantity since  $k_{PN}$ , the effect of the Purkinje on the nuclear cell, is inhibitory, and  $k_{cP}$ , the effect of the climbing fiber on the Purkinje, is the inactivation response. Thus the factor  $(k_{cN} - k_P)$  is always positive.

Since the climbing fiber pattern is stored in the  $X_i$  pattern, the effect of the mossy fiber  $X_i$  pattern associated with the climbing fiber pattern reinforces the climbing fiber's effect on the nuclear cell. Thus, as learning takes place, less and less input from the climbing fiber is necessary to produce the same amount of nuclear cell response.

Fourth, the effect of an input on mossy fibers through the function  $X_i(f_{m1}, \dots, f_{mN})$  is a positive response. The  $X_i$  function in (7) decreases the output of the Purkinje cell and hence in (9) increases the output of the nuclear cell.

## V. IMPLICATIONS

It is reasonably certain that patterns of activity on mossy fibers represent to the cerebellum the position, velocity, tension, and so on of the muscles, tendons, and joints. This is feedback information that is required to control precise or sequential movements, or both. This information must modulate signals to the muscles to achieve precise movement under varying load conditions. This feedback information must also be able to generate the next command in a sequence of muscle commands in order to produce sequential motor activity at a subconscious level. The functioning of the cerebellum, as hypothesized in this article, seems rather well suited for either or both of these behaviors.

Assume, for example, that the red nucleus sends a command  $C_1$  through the inferior olive and thence via climbing fibers through Purkinje cells and nuclear cells to the muscles. At this time the muscles and joints in their resting state are sending pattern  $M_1$  to the cerebellum via mossy fibers. Thus  $C_1$  is imprinted on  $M_1$ . Now when  $C_1$  reaches the muscles, they respond by moving to a new position. This generates a new mossy fiber pattern  $M_2$ . By this time a second command  $C_2$  is sent from the red nucleus. Command  $C_2$  will be imprinted on  $M_2$ . In a similar manner  $C_3$  is imprinted on  $M_3$ ,  $C_4$  on  $M_4$ , and so on. This process may be continued for a lengthy sequence of motor commands  $C_1 C_2 C_3 \cdots$  and resulting body positions  $M_1 M_2 M_3 \cdots$ . Upon repetition of the sequence of motor commands  $C_1 C_2 C_3 \cdots$ , the signals from the red nucleus will be reinforced at the nuclear cells by output from Purkinje cells responding to feedback mossy fiber patterns  $M_1 M_2 M_3 \cdots$ . Upon each repetition more and more of the muscle control can be assumed by the output of the Purkinje cells, and less attention is required by higher motor centers.

Once learning is complete, the sequence of motor commands  $C_1 C_2 C_3 C_4$  can be elicited entirely from the Purkinje cells via the mossy fiber input patterns  $M_1 M_2 M_3 M_4 \cdots$ . Little input is required from higher centers except perhaps to initiate or terminate the sequence.

The theory so far has no means of initiating or terminating such a sequence. It is possible that this operation takes place in the intracerebellar nuclei or outside the cerebellum altogether. Lack of detailed anatomical and physiological data makes it difficult to conjecture how this function is accomplished. However, it is perhaps not unreasonable to speculate that the Schiebel collaterals of climbing fibers to Golgi cells or stellate cells, or to both, may be related to initiation or termination of sequence generation in the cerebellar cortex. The Golgi cells control the mossy fiber input pathway, which is a vital link in sequence generation. Excitation of Golgi cells via Schiebel collaterals could cut off mossy fiber input to the cerebellum and terminate a sequence. Inhibition of Golgi cells by Purkinje recurrent collaterals, on the other hand, would lower Golgi inhibition, possibly in response to specific patterns. This might initiate sequences upon certain key commands. Golgi cells may also have variable synapses, since they possess both spine synaptic contacts with parallel fibers and input from climbing fibers. However, more data are necessary before confident predictions are possible on these points.

The circuit described can also function as a modulator of conscious motor activity on climbing fibers. Assume that a sequence of motor commands from higher centers  $C_1 C_2 C_3 \cdots$  had been imprinted on a series of mossy fiber patterns  $M_1 M_2 M_3 \cdots$  as before. If the muscles upon receipt of conscious command  $C_1$  were to encounter greater than usual resistance,

this would delay or prevent the appearance of  $M_2$  at the cerebellum, and instead a pattern  $M'_2$  would appear, signaling the existence of extraordinary resistance to motion. The pattern  $M'_2$  would modify pattern  $C_2$  in a manner different from  $M_2$ , perhaps calling for additional force or some other modification. What  $M'_2$  produces is governed by what previously had been imprinted on  $M'_2$ . If previously  $C'_2$ , an additional force command, had been imprinted on  $M'_2$ , the  $C'_2$  would be substituted for  $C_2$  automatically when the  $M'_2$  feedback signal was received instead of the usual  $M_2$ . By this means a sequence of conscious commands can be modified at the reflex level by cerebellar activity. This perhaps is the means by which motor activity such as running or skating can be under conscious control in a general sense but under reflex feedback control at the individual muscle level.

The implication, then, is that climbing fibers carry from higher centers control patterns that are to be stored. In this form the cerebellar memory becomes a form of conditioned reflex. If the climbing fibers are cut, we would expect deficiencies primarily in conscious motor control and further conditioning. This may in some measure account for data of Mettler [12], which noted a lack of obvious severe effects when climbing fibers were cut.

Marr [11] suggests an interesting analogy of the cerebellum as a language translator between data in the cerebrum and command sequences needed by the muscles. The cerebellum thus becomes analogous to a computer compiler that translates source language instructions into machine language instructions for execution by the machine hardware. Following the same analogy, the cerebellum becomes a subroutine library in which subroutines can be stored from above and cycled from below.

## VI. PREDICTIONS

The theory of cerebellar function set forth in this article makes possible a number of predictions that are subject to experimental verification:

1. Parallel fibers do not fire in coordinated beams in a conscious active animal, but rather in a widely scattered, apparently random fashion.
2. One percent or less parallel fibers are active simultaneously, and this activity level is quite constant.
3. Parallel fiber synapses with dendritic spines on Purkinje cells, basket cells, and stellate cells are modifiable synapses.
4. The Purkinje cell response can be conditioned by climbing fiber inputs. Climbing fiber spikes are the unconditioned stimulus (US). Mossy fiber activity patterns are the conditioned stimulus (CS). The climbing fiber inactivation response is the unconditioned response (UR).

5. The conditioning mechanism is a three-way coincidence between the inactivation response, a cell spike due to parallel fiber excitation, and parallel fiber synaptic activity.

6. Parallel fiber synapses on Purkinje cells, basket cells, and stellate cells are weakened by incorrectly firing during climbing fiber activity.

7. Climbing fibers are essential for acquisition of certain types of motor skills, and for cerebellar feedback control of conscious motor activity. They are less necessary for conditioned reflex behavior.

8. Some of the mechanisms hypothesized in the cerebellum will almost certainly also occur in other parts of the brain. The expansion recoding system; the imprinting of patterns from specific fiber inputs onto synapses of nonspecific fibers; the use of laterally coursing inhibitory interneurons to achieve both positive and negative synaptic weight adjustment; the weakening of synaptic weights during training to achieve convergence; these are all basic principles of data processing likely to occur elsewhere in the nervous system.

#### ACKNOWLEDGMENT

The author thanks Mr. Anthony J. Barberra for his valuable criticism and suggestions.

#### REFERENCES

- 1 J. S. Albus, A model of memory in the brain, *Cybernetica* (1970) (in press).
- 2 R. S. Cajal, *Histologie du système nerveux de l'homme et des Vertébrés*, Tome II. Maloine, Paris, 1911.
- 3 C. D. Bell and R. J. Grimm, Discharge properties of Purkinje cells recorded on single and double microelectrodes, *J. Neurophysiol.* **32**(1969), 1044–1055.
- 4 T. M. Cover, Classification and generalization capabilities of linear threshold units, Rome Air Development Center Tech. Documentary Rept. RADC-TDR-64-32 (1964).
- 5 J. C. Eccles, M. Ito, and J. Szentagothai, *The cerebellum as a neuronal machine*. Springer, Berlin, 1967.
- 6 A. Escobar, E. D. Sampedro, and R. S. Dow, Quantitative data on the inferior olivary nucleus in man, cat and vampire bat, *J. Comp. Neurol.* **132**(1968), 397–403.
- 7 F. A. Fox, D. E. Hillman, K. A. Sugessmund, and C. R. Dutta, The primate cerebellar cortex: A Golgi and electron microscope study, *Progr. Brain Res.* **25**(1967), 174–225.
- 8 R. Granit and C. G. Phillips, Excitatory and inhibitory processes acting upon individual Purkinje cells of the cerebellum in cats, *J. Physiol. (London)* **133**(1956), 520–547.
- 9 D. H. Hubel and T. N. Wiesel, Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex, *J. Physiol. (London)* **160**(1962), 106–154.
- 10 A. Jakob, Das Kleinhirn, in *Handbuch der mikroskopischen Anatomie des Menschen* IV/I (W.v. Mollendorf, ed.). Springer, Berlin, 1928.
- 11 D. Marr, A theory of cerebellar cortex, *J. Physiol. (London)*, **202**(1969), 437–470.
- 12 F. A. Mettler, (1967), In a discussion following a paper by J. C. Eccles in *Neurophysiological basis of normal and abnormal motor activities* (M. D. Yahr and D. P. Purpura, eds.), pp. 411–414, Raven Press, N.Y., 1967.

- 13 M. Minsky and S. Papert, *Perceptrons: An introduction to computational geometry*. MIT Press, Cambridge, Massachusetts, 1969.
- 14 N. J. Nilsson, *Learning machines: Foundations of trainable pattern-classifying systems*. McGraw-Hill, New York, 1965.
- 15 F. Rosenblatt, *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. Spartan Books, Washington, D.C., 1961.
- 16 A. Ruiz-Marcos and F. Valverde, Temporal evolution of the distribution of dendritic spines in the visual cortex of normal and dark-raised mice, *Exptl. Brain Res.* **8**(1969), 284–294.
- 17 W. T. Thach, Jr. Somatosensory receptive fields of single units in cat cerebellar cortex, *J. Neurophysiol.* **30**(1967), 675–696.
- 18 W. T. Thach, Discharge of Purkinje and cerebellar nuclear neurons during rapidly alternating arm movements in the monkey, *J. Neurophysiol.* **31**(1968), 785–797.
- 19 W. T. Thach, Discharge of cerebellar neurons related to two maintained postures and two prompt movements, II: Purkinje cell output and input, *J. Neurophysiol.* **33**(1970), 537–547.