RELIABILITY OF COMPUTATION IN THE CEREBELLUM

N. H. SABAH

From the Department of Biophysical Sciences, State University of New York at Buffalo, Amherst, New York 14226

ABSTRACT The mossy fiber–granule cell–parallel fiber–Purkinje cell system of the cerebellar cortex is investigated from the viewpoint of reliability of computation. It is shown that the effects of variability in the inputs to a Purkinje cell can be reduced by having a large number of parallel fibers whose activities are statistically independent. The mossy fiber–granule cell relay is shown to be capable of performing the required function of transforming the activity in a small number of mossy fibers into activity in a much larger number of parallel fibers, while ensuring that there is little correlation between the activities of individual parallel fibers. The effects of variability in the outputs of Purkinje cells may be reduced by redundancy and convergence schemes, as evidenced by the geometrical pattern of parallel fibers and Purkinje cells and the convergence of these cells onto their target neurons.

INTRODUCTION

The cerebellum is known to be intimately involved in the coordination and control of movement in vertebrates (Bell and Dow, 1967; Dow and Moruzzi, 1958; Eccles, 1969 *a*). In addition to inputs from the cerebral cortex, the cerebellum receives inputs from most, if not all, sensory modalities; its output modulates the motor outflow through the brain stem and influences both motor and sensory areas of the cerebral cortex (Bell and Dow, 1967; Eccles, Ito, and Szentágothai, 1967; Evarts and Thach, 1969; Jansen and Brodal, 1954; Provini et al., 1968; Rubia and Phelps, 1970; Snider and Mitra, 1970). Studies on the cerebella of marine and freshwater species have suggested that the function of the cerebellum might not be restricted to the control of movement (Bennett, 1969; Bullock, 1969). The cerebellum may in fact be envisaged as some type of special purpose computer that processes pertinent sensory information into a form that is utilized by the animal in formulating appropriate responses to its environment (Eccles, 1969 *b*; Eccles, Ito, and Szentágothai, 1967).

The central theme of the present paper is reliability of computation; the underlying viewpoint is that if the cerebellum is regarded as some type of neuronal computer, then whatever the nature of the computation may be, this computation must be per-

formed reliably in the presence of much variability in the input signals and in the characteristics of the constituent neuronal elements. Some general aspects of neuronal variability are considered in the Appendix. It is convenient for present purposes to distinguish between two of these aspects: (*a*) the variability in synaptic inputs as reflected at the postsynaptic membrane, and (*b*) the over-all variability in the output of the neuron. These will be referred to as "input uncertainty" and "output uncertainty", respectively. Some salient features of the architectural design of the mossy fiber–granule cell–parallel fiber–Purkinje cell system are examined from the viewpoint of reduction of input and output uncertainties. Inhibitory interneurons and climbing fibers will not be considered in this paper.

BASIC FEATURES OF CEREBELLAR CORTEX

The histology and physiology of the cerebellar cortex have been extensively studied, particularly in the cat (Bell and Dow, 1967; Eccles, Ito, and Szentágothai, 1967; Evarts and Thach, 1969; Fox and Snider, 1967; Llinás, 1969; Szentágothai, 1968). Reference may be made to these publications for detailed information and literature surveys. The salient features of the cerebellar cortex of the cat will be briefly outlined in this section, based on Eccles, Ito, and Szentágothai (1967) and Szentágothai (1968). Attention will be focused on the mossy fiber–granule cell–parallel fiber–Purkinje cell system, whose basic features are remarkably similar in all vertebrates (Eccles, 1969 *b*; Llinás and Hillman, 1969).

The cerebellar cortex consists of three layers (Fig. 1 A): the outermost molecular layer ($\sim$300 μ thick), the Purkinje cell layer (70–100 μ thick), and the granular layer (400–500 μ thick at the top of the folia and $\sim$100 μ thick in the depth of the furrows). The molecular layer has a remarkable rectangular lattice type of structure. The espalier-like dendritic arborizations of Purkinje cells occupy individual compartments, each about 300 $\mu \times 225\ \mu \times 6\ \mu$, aligned in transverse rows across the folium and separated by about 2 μ in the longitudinal direction. Basket and stellate cells are located in the molecular layer. Their dendritic arborizations are also flattened and lie in transverse planes but are sparser and much less extensive than those of Purkinje cells. The axons of basket and stellate cells make inhibitory synapses with Purkinje cells, as shown in Fig. 1 A.

The granular layer contains granule cells and Golgi cells. The dendrites of typical Golgi cells are located mainly in the molecular layer, where they expand more or less equally in all directions. The axons of granule cells ascend to the molecular layer, where they branch in a characteristic *T*-fashion and give rise to the parallel fibers (Fig. 1 A). These fibers run in a longitudinal direction and make crossing-over excitatory synapses with the dendrites of Purkinje cells, stellate cells, basket cells, and with the upper dendrites of Golgi cells. Roughly 200,000 parallel fibers cross the dendritic compartment of a Purkinje cell (Szentágothai, 1968) with practically every parallel fiber making synaptic contact with a single spine on every Purkinje cell

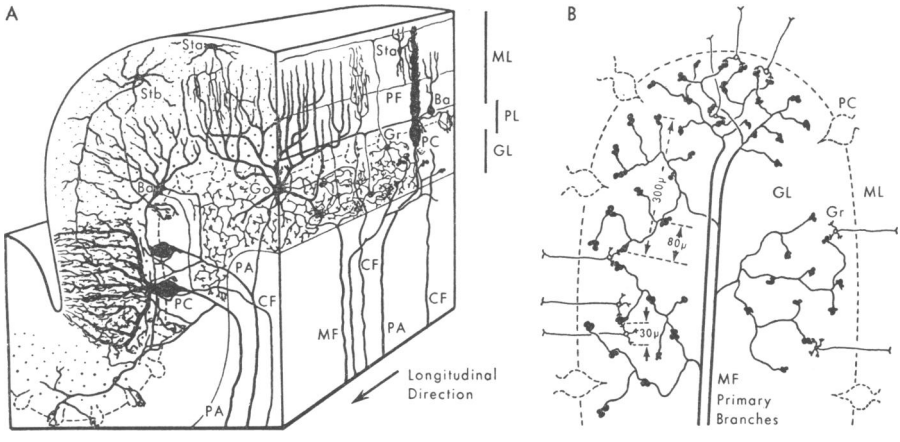

FIGURE 1 (A) A schematic perspective illustration of the neuronal elements in a folium of cerebellar cortex, the planes of the two sections being parallel to the folium axis in one case and perpendicular to it in the other. (B) Transverse section of cerebellar folium showing diagrammatically the pattern of branching of the mossy fibers in the central medullary lamina and in the granular layer. ML: molecular layer; PL: Purkinje cell layer; GL: granular layer; MF: mossy fiber; CF: climbing fiber; PA: Purkinje axon; PC: Purkinje cell; Go: Golgi cell; Gr: granule cell; PF: parallel fiber; Ba: basket cell; Sta: short axon stellate cell; Stb: long axon stellate cell. Dimensions are not to scale.

dendritic compartment it crosses, or at most with a few additional neighboring secondary spines. The length of a parallel fiber is 2–3 mm; superficial parallel fibers tend to be thinner and shorter than deeper ones.

The cerebellar cortex has only a single type of efferent fibers—the axons of Purkinje cells—and two types of afferents: climbing fibers and mossy fibers. Climbing fibers make extensive excitatory synapses with Purkinje cells and give off collaterals that innervate inhibitory interneurons.

Primary mossy afferents may branch to two or more folia before traversing the central medullary lamina of a given folium. A primary mossy branch gives rise to two or three preterminal branches that enter the granular layer on either side (Fig. 1 B). Each preterminal branch divides in a characteristic candelabrum-shaped arborization that is roughly cylindrical, about 300 μ in diameter, the axis of the cylinder being parallel to the longitudinal direction of the folium. Each arborization bears about 10 mossy rosettes, and the two or three arborizations of the preterminal branches of a given primary branch are nonoverlapping but otherwise appear to be distributed at random.

Granule cells have four or five dendrites on the average, with each mossy rosette making excitatory synapses with the dendrites of about 20 granule cells. Since the mossy rosettes of the arborization of a given preterminal branch are 80–100 μ apart, whereas the dendritic span of granule cells is 20–30 μ (Fig. 1 B), it is considered unlikely that two or more dendrites of a granule cell synapse with the rosettes of the

same mossy afferent. Mossy rosettes also make excitatory synapses with the lower dendrites of Golgi cells. The axons of Golgi cells arborize extensively and make inhibitory synapeses with granule cell dendrites, somewhat "downstream" from the granule cell dendrite–mossy rosette synapses.

THE PARALLEL FIBER–PURKINJE CELL SYNPASES

It is argued in the Appendix that the uncertainty in the input to a neuron may be reduced by having a large number of input lines whose activities, in the ideal case, are statistically independent. These considerations are relevant to appreciating the significance of the enormous number of synapses made by the parallel fibers with a Purkinje cell. In order to have some idea of the numerical values involved, consider the case of Poisson inputs to a Purkinje cell. If it is assumed that the mean rate of firing of granule cells is 30 imp/sec (Eccles et al., 1966 *a*; Thach, 1967) and that the summation period, δ, equals 40 msec, it follows from equation A 1 that $\sigma^2 = 1.2$. According to Fig. 4, 3000–6000 input lines will be sufficient in this case to reduce input uncertainty to a low level. The actual number of input lines may differ from this because of such factors as the different effectiveness of input lines depending on their spatial distribution, correlations between the activities of input lines, and the simplifying assumptions made in the derivation. Nevertheless, the number of parallel fibers synapsing with a Purkinje cell (~200,000) is at least an order of magnitude larger than the above estimate. Presumably, this reflects the likelihood that spikes of Purkinje cells are generated at multiple zones in the dendritic tree (Eccles et al., 1966 *b*; Sabah and Murphy, 1971).

THE MOSSY FIBER–GRANULE CELL RELAY

Based on the histological description given above of the connection pattern between mossy fiber terminals and granule cells, and assuming that few, if any, granule cells share exactly the same inputs to all of their dendrites, it can be postulated that there exist a number of granule cells G and a number M of preterminal mossy fiber branches such that G equals the number of combinations of M items taken d at a time, where d is the average number of dendrites of a granule cell:

$$G = \binom{M}{d}, \tag{1}$$

since the M branches contact about $200M$ granule cell dendrites

$$G \simeq 200\frac{M}{d}. \tag{2}$$

If d is equal to 4 or 5, the following values of G and M are obtained from equations 1 and 2:

$$d = 4\text{: } M \simeq 12,\ G \simeq 600;$$

$$d = 5\text{: } M \simeq 11,\ G \simeq 440.$$

The following comments can be made concerning these values:

(a) The ratio G/M equals 40 to 50. If there are two or three preterminal branches per primary mossy branch, then the ratio of granule cells to primary branches is 80 to 150. This agrees with the figure of about 100 cited by Szentágothai (1968).

(*b*) Because of its length and the spatial arrangement of dendritic compartments of Purkinje cells, a parallel fiber would potentially contact about 400 Purkinje cells on the average (Eccles, Ito, and Szentágothai, 1967). If 200,000 parallel fibers cross the dendritic compartment of a Purkinje cell, then N, the number of parallel fibers which on the average terminate at every Purkinje cell and are replaced by a new set of N fibers, will be given by: $N \simeq 200{,}000/400 = 500$. The value of N is seen to be approximately equal to the ratio G/M. It is as if sets of about 500 parallel fibers form "unitary bundles", the activity of each bundle reflecting the activity of about a dozen preterminal mossy branches.

It becomes of interest, therefore, to determine the number g out of G granule cells that become excited when m out of M preterminal branches are active, and the excitation of 1, 2, 3, 4, or 5 dendrites is required to fire a granule cell. A detailed investigation is quite complex and would require the specification of the firing patterns of mossy fibers. Limiting probability distributions are of little use because the number of inputs to a granule cell is small. Nevertheless, useful approximate results can be obtained using Vandermonde's convolution formula of combinatorial algebra (Riordan, 1968):

$$\binom{M}{d} = \binom{m}{m}\binom{M-m}{d-m} + \binom{m}{m-1}\binom{M-m}{d-m+1} + \cdots + \binom{m}{0}\binom{M-m}{d}. \quad (3)$$

This formula follows readily from the identity $(1 + x)^M = (1 + x)^{M-m}(1 + x)^m$ by equating the coefficients of x^d on both sides. The interpretation of equation 3 is as follows: Suppose a set of M items is given and consider m distinguishable items from this set. If the number of combinations of M items, d at a time, is formed, then the first term on the right-hand side of equation 3 represents all the combinations that contain all of the m items, the second term represents all the combinations that contain $(m - 1)$ of the m items, etc., and the last term represents all the combinations that do not contain any of the m items. For any value of m, $(m = 1, \cdots, M)$, and any threshold θ, $(\theta = 1, \cdots, d)$, the number of granule cells activated will be given by the sum of the first r terms on the right-hand side of equation 3 for which $m - r + 1 \geqq \theta$. In making numerical estimates it must be remembered, of course, that the factorial of a negative integer is infinite, so that terms in equation 3 involving negative integers will be zero. Calculated values of g for different values of m and θ are given in Table I and plotted in Fig. 2 as g/G vs. m/M. These calculations are based on the values: $d = 5$, $M = 11$, $G = 462$, according to equation 1.

TABLE I

NUMBER OF GRANULE CELLS ACTIVATED BY MOSSY FIBERS

m	$\theta = 1$	$\theta = 2$	$\theta = 3$	$\theta = 4$	$\theta = 5$
1	210	0	0	0	0
2	336	84	0	0	0
3	406	196	28	0	0
4	441	301	91	7	0
5	456	381	181	31	1
6	461	431	281	81	6
7	462	455	371	161	21
8	462	462	424	266	56
9	462	462	462	378	126
10	462	462	462	462	252
11	462	462	462	462	462

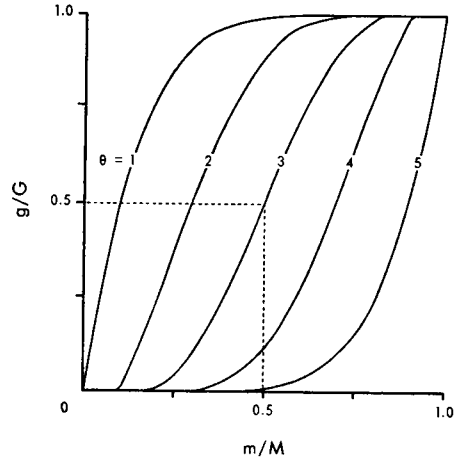

FIGURE 2 Number of granule cells (g out of G) excited when m out of M mossy fibers are active. $M = 11$, $G = 462$ as explained in text; θ is the threshold of a granule cell expressed in the number of dendrites that should be simultaneously activated to fire a granule cell.

It follows from Fig. 2 that provided θ equals 2, 3, or 4, the value of g is roughly proportional to that of m, with some "restoration" at extreme values of m. By restoration (von Neumann, 1956) is meant that a low level of activity of the mossy fibers, which may be just "noise", is converted to an even lower level of activity of the granule cells and, conversely, a high level of activity of the mossy fibers is converted to an even higher level of activity of the granule cells. Significantly, it is found experimentally that an impulse in a single mossy fiber is not sufficient to fire a granule cell (Eccles, Sasaki, and Strata, 1967; Eccles, Táboříková, and Tsukahara, 1970).

The mossy fiber–granule cell relay therefore gives rise to an increased number of parallel fibers compared to mossy fibers, while roughly preserving the level of activity of the mossy fibers. In other words, this relay functions as a divergence mecha-

nism, as postulated by Eccles et al. (1966 *a*). However, this does not by itself justify having a synaptic relay, for in principle the greater divergence of afferents could be realized simply by branching of mossy fibers. But this would result in highly correlated activity over a large number of input lines to Purkinje cells. As mentioned above, such a high correlation would be detrimental to the reduction of uncertainty in the inputs to Purkinje cells.

It can be easily shown in a qualitative manner that the mossy fiber–granule cell relay results in less correlated activity over output lines, compared to the case where the input lines simply branched to become equal in number to the output lines. Consider, for example, the case where six mossy fiber preterminal branches are active, i.e., $m = 6$, and $\theta = 3$. The nonzero terms in equation 3 will then be:

$$\binom{11}{5} = \binom{6}{5}\binom{5}{0} + \binom{6}{4}\binom{5}{1} + \binom{6}{3}\binom{5}{2} + \binom{6}{2}\binom{5}{3} + \binom{6}{1}\binom{5}{4} + \binom{6}{0}\binom{5}{5},$$

or:

$$462 = 6 \times 1 + 15 \times 5 + 20 \times 10 + 15 \times 10 + 6 \times 5 + 1 \times 1.$$

The interpretation of these numbers is as follows:

(*a*) Each of six granule cells receives a different combination of five active lines.

(*b*) Each of 75 granule cells receives four active lines with groups of five granule cells receiving exactly the same input activity.

(*c*) Each of 200 granule cells receives three active lines, with groups of 10 granule cells receiving exactly the same input activity. The remaining granule cells receive two active lines or less, and since $\theta = 3$, they will not be activated.

Without a granule cell relay, each preterminal branch would have had to give about 44 branches, for the same number of input lines to Purkinje cells. Groups of 44 lines would therefore have correlated activity. The granule cell relay in effect breaks up or "randomizes" the mossy fiber activity. In practice, the randomization can be more effective than indicated by the above example, because of the following factors:

(*a*) fluctuations in the thresholds of granule cells, enhanced by their small size ($\sim 5\ \mu$ diameter), their dense packing ($\sim 5 \times 10^6/\text{mm}^3$) and by local variability in Golgi cell inhibition;

(*b*) modifications of the above scheme. Suppose, for example, that the basic scheme is that of having combinations of M items 4 at a time but that the granule cell has 5 dendrites, the additional dendrite in different granule cells being connected at random to neighboring mossy fibers. The activities in the additional dendrites would in effect eliminate groups having the same input activity, as in (*b*) and (*c*) of the

above example. The number g would still be roughly proportional to m, particularly for higher values of θ, and the solution to equations 1 and 2 becomes: $M \simeq 12$, $G \simeq 500$.

It may be concluded, therefore, that the mossy fiber–granule cell relay increases the number of parallel fibers as compared to mossy fibers, while ensuring that there is little correlation between the activities of individual parallel fibers. This arrangement can be effective in reducing the uncertainty in the inputs to Purkinje cells. The mossy fiber–granule cell relay may be described as a randomizing divergence mechanism.

CONVERGENCE OF PURKINJE CELL OUTPUT

The axons of 20–50 Purkinje cells converge onto a single target neuron in Deiters' nucleus or in the cerebellar nuclei (Bell and Dow, 1967; Eccles, Ito, and Szentágothai, 1967). The precise distribution in the cerebellar cortex of the Purkinje cells that converge onto a single target neuron is uncertain, mainly because the problem of locating these Purkinje cells is complicated by the possibility of stimulating Purkinje axon collaterals in the same folium or in neighboring folia (Bell and Dow, 1967; Eccles, Ito, and Szentágothai, 1967). On the basis of available evidence (Ito et al., 1968; Ito et al., 1970), it seems justifiable to assume that the 20–50 Purkinje cells are located in at most a few restricted foci not necessarily confined to the same folium.

Let the number of Purkinje cells that converge onto a single target neuron be C. As a first approximation under steady state conditions, the activities of these Purkinje cells may be considered independent (Bell and Grimm, 1969; Sabah and Murphy, 1971). If the mean rate of firing of the jth Purkinje cell is q_j imp/sec, then the arrival of inhibitory postsynaptic potentials (IPSP) at a target neuron may be approximated by a Poisson process of rate $\lambda = \sum_{j=1}^{c} q_j$ imp/sec (Sabah and Murphy, 1971). If b and s^2 denote, respectively, the mean and variance of the hyperpolarizing potential of the postsynaptic membrane due to the activities of the C convergent Purkinje cells, then it follows from Campbell's theorem (Papoulis, 1965; Rice, 1954): $b = \lambda \int_{-\infty}^{\infty} h(t)\, dt$ and $s^2 = \lambda \int_{-\infty}^{\infty} h^2(t)\, dt$, where $h(t)$ denotes the time course of the IPSP. Approximating the IPSP by $ue^{-t/\tau}H(t)$, where u is the amplitude of the IPSP including its sign, τ is the time constant of the IPSP, and $H(t)$ is the unit step function, it follows that (Papoulis, 1965; Stein, 1967): $b = u\lambda\tau$ and $s^2 = u^2\lambda\tau/2$. The ratio $s/b = 1/\sqrt{2\lambda\tau}$ indicates, as to be expected, that the voltage level will be "smoother" the larger τ, C, or the mean rate of firing of Purkinje cells. If $\lambda = 1200$ imp/sec and $\tau = 30$ msec, then $s/b \simeq 0.12$. There would probably be no point in reducing the value of this ratio much further, since random fluctuations in membrane potential may amount to about 10% of threshold (Calvin and Stevens, 1968; McCulloch, 1959).

If variations in the outputs of Purkinje cells are to have a significant effect on the membrane voltage of the target neuron, the magnitude of the change in voltage

must be large compared to the standard deviation of the random fluctuations from all sources. In order to significantly modify the output of the target neuron, therefore, more than a single Purkinje cell will have to be involved in "concerted action". In general, it may be envisaged that the smallest number of Purkinje cells involved in such concerted action is a fraction "a_k" of a rank of Purkinje cells, R_k, where $0 < a_k \leqslant 1$, $\sum_{k=1}^{L} R_k = C$, and L takes integer values greater than or equal to 1. In other words, the number of Purkinje cells C that converge onto a single target neuron may be divisible into one or more subsets of R_k Purkinje cells each. The members of each rank of R_k Purkinje cells share a common input and would ideally have the same output, thereby giving rise to redundancy of computation. A fraction $a_k R_k$ from one or more ranks would participate in causing a significant change in the output of the target neuron.

The existence of a redundancy scheme involving Purkinje cells is suggested by the geometrical configuration of parallel fibers and Purkinje cells. Thus a bundle of $N = 500$ parallel fibers contacts about 400 Purkinje cells. If the dendritic compartments of Purkinje cells were aligned in the longitudinal direction, then KN bundles will be shared by $(401 - K)$ Purkinje cells. The redundancy number R_k will be smaller than this because the dendritic compartments of Purkinje cells are not aligned in the longitudinal direction (Eccles, Ito, and Szentágothai, 1967). It appears that R_k varies with K, so that there will be more redundancy for weak inputs involving a small number of bundles than for stronger inputs involving a larger number of bundles.

The above considerations apply "on-beam" to the excitation of Purkinje cells by parallel fibers and to their inhibition by short axon stellate cells. Similar considerations apply "off-beam" to inhibition of Purkinje cells by basket and type *b* stellate cells, since the excitatory input in a beam of parallel fibers is common to a number of these cells; the axon of each of these cells potentially reaches a matrix of 50–70 Purkinje cells (Eccles, Ito, and Szentágothai, 1967).

By analogy to the "majority decisions" of logic elements (Verbeek, 1960 *a*), redundancy schemes involving Purkinje cells can be effective in reducing the uncertainty in the outputs of these cells. Let p_v be the probability of a given random fluctuation in the output of a single cell in a rank of R Purkinje cells. It is assumed for simplicity that this probability is the same for all the R cells, although this need not be the case. If the fluctuations occur independently between the R cells, then the probability P_v that these fluctuations occur in at least aR of the neurons will be given by the cumulative binomial distribution:

$$P_v = \sum_{i=aR}^{R} \binom{R}{i} p_v^i (1 - p_v)^{R-i}.$$

P_v is plotted against p_v in Fig. 3 for various values of "a" and R. As to be expected, the effectiveness of such a redundancy scheme depends on the values of "a" and R.

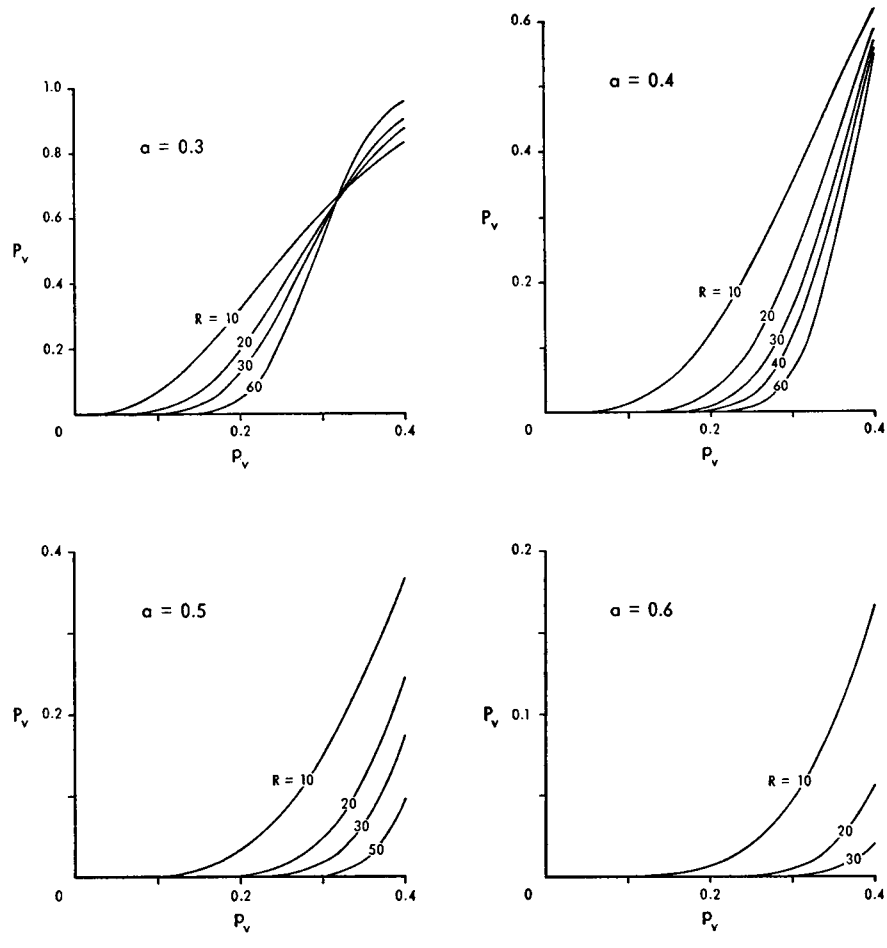

FIGURE 3 Effects of convergence of a rank of R neurons. p_v is the probability of a certain fluctuation in the output of a single neuron in the rank of R neurons, and P_v is the probability of this fluctuation in the outputs of aR neurons of the same rank.

Provided "a" exceeds 0.5, a value of R between 20 and 50 appears adequate to reduce P_v to a low level. If "a" is 0.6 or more, then R need only equal about 10.

The salient features of the mossy fiber–granule cell–parallel fiber system may be summarized as follows:

(a) The mossy fiber–granule cell synaptic relay gives rise to a much larger number of parallel fibers, compared to mossy fibers, and ensures that little correlation occurs between the activities of individual parallel fibers. This arrangement can be effective in reducing the uncertainty in inputs to Purkinje cells.

(b) The design of this synaptic relay appears to be such that the activity in "unitary bundles", of about 500 parallel fibers each, reflects the activity of about a dozen mossy afferents.

(c) The length of the parallel fibers ensures redundancy of input to an adequate

number of Purkinje cells. Such redundancy may be utilized at the level of the target neurons to reduce the uncertainty in outputs of Purkinje cells.

(*d*) The geometrical pattern of the branching of mossy fibers, the length of the parallel fibers, and their staggering because a number of them terminate on every Purkinje cell, to be replaced by other parallel fibers, and the nonalignment of dendritic compartments of Purkinje cells in the longitudinal direction ensure that the mossy fiber input is relayed to Purkinje cells in a wide variety of combinations.

(*e*) The bifurcation of the parallel fibers at their midpoints ensures a minimum time delay between given sets of granule cells and Purkinje cells. Thus if a set of 400 Purkinje cells is to be contacted by a set of 500 granule cells, then for a given location of the Purkinje cells, the longest delay involved will be minimized if the granule cells were located near the middle of the set of Purkinje cells and their axons branched in both directions.

DISCUSSION

In considering the significance of a large number of inputs to a neuron (see Appendix) it was assumed initially that firing of the neuron is primarily determined by the threshold of the neuron and by the number of excitatory inputs arriving within a summation period δ. It followed that if the inputs are independent and large in number, then the significant parameter of the input is the mean excitation level ϵ over the input lines; in other words, the precise spatiotemporal pattern of the input is of little importance (Segundo et al., 1966). In the context of reliability of computation Von Neumann (1956) suggested the possibility that messages might be conveyed in the nervous system by excitation levels over bundles of fibers. At least in the cerebellum, activity in single fibers is probably of little significance when considered in isolation, as evidenced by the fact that the firing of a granule cell generally requires the convergence of two or more mossy fiber inputs (Eccles, Sasaki, and Strata, 1967; Eccles, Táboříková, and Tsukahara, 1970).

It is interesting in this regard that in the efferent pathway of the cerebellar cortex, interspike intervals have been found to be independent in the case of Purkinje cells (Sabah and Murphy, 1971) and spinal motoneurons (Calvin and Stevens, 1968). This implies that the precise sequence of impulses in a spike train does not carry information, at least under stationary conditions. In the final pathway of motor units, the mean firing rate of the motoneuron is believed to be the significant parameter controlling muscular contraction (Magdaleno et al., 1968). On general grounds it does not seem likely that the operation of an organ like the cerebellum depends on strict temporal relationships and precise sequences of impulses. Man-made digital devices are generally designed to operate with such signals. But such devices operate under closely controlled conditions, and any appreciable departures from these conditions usually lead to malfunctions. On the other hand, the nervous systems of higher organisms function reliably over long periods of time and in the presence of much variability in the neuronal signals and in the operating physiological condi-

tions. Therefore, when considering the operation of neuronal networks, the most attractive and probably the most plausible schemes are those that do not depend on rigid timing or sequencing of impulses and which allow the network to cope with variability, i.e., compute reliably. While such schemes are usually not the most efficient in terms of information capacity and economy of components, these considerations must be weighed against the requirements of reliability of computation. The enormous number of 10–100 billion granule cells in the human cerebellar cortex (Braitenberg, 1967) is hardly in favor of any argument that economy of components plays a major role in the design of the cerebellar cortex.

Although the problem of reliability of computation has been considered quite extensively by automata theorists (cf. references in Appendix), analogous problems in the nervous system have received little attention, with some exceptions, (Eccles, Faber, et al., 1970). In fact, it has sometimes been assumed that the variability in the outputs of single neurons limits the "accuracy" of the nervous system, even though simple convergence schemes can be effective in reducing the effects of such variability. The problems of coping with variability are probably relevant to most parts of the nervous system, although not to the same extent. The cerebellum seems to be highly developed for this purpose, irrespective of what its actual functions may be. Any attempt at estimating quantitatively how much variability can be tolerated in the output of the cerebellum must take into account the reduction in the effects of variability that arises from the integration of movement in whole muscles or sets of muscles and from the presence of dynamic feedback loops that continually signal the progression of an evolving movement (Eccles, 1969 *a*).

The purpose of investigating the question of reduction of uncertainties in the cerebellum is twofold. Firstly, these considerations provide some insight into the possible meaning of some of the architectural features of the cerebellar cortex; secondly, and perhaps more significantly, the understanding that is thereby gained serves to reduce the order of complexity of the system so that some other selected aspect of cerebellar operation can then be considered with advantage. Rather than attempting to understand the operation of the whole system all at once, a phased approach may be used, attention being focused at each stage on some well defined aspect of system operation. Such an approach is often used in the study of physical systems, particularly in engineering. It must be kept in mind, however, that biological systems often perform integrated functions, so that the same neuronal structure may conceivably serve more than one purpose.

The author is indebted to Professor J. C. Eccles for his advice and criticism, and to Dr. R. A. Spangler for many stimulating discussions.

This work was supported in part by grant no. 5 R01NS-0822-102, National Institutes of Neurological Disease and Stroke, and by NASA grant no. NGR33-015-016.

Received for publication 19 June 1970 and in revised form 9 September 1970.

APPENDIX

Variability of Neuronal Activity

A neuron is essentially an information processing and transmitting device, the soma-dendritic region being specialized for the former role and the axon for the latter. The output of a neuron may be formally described as some function of its inputs; the nature of this function depends in general on the assumed character of inputs and output. In broad terms therefore, a neuron can be considered as an element that computes some function of its inputs.

Computation in neuronal networks is subject to numerous sources of variability. The thresholds of nerve membranes are affected by changes in the ionic environment, temperature, pH, partial pressures of oxygen and carbon dioxide, as well as by pharmacological agents, metabolic factors, and local fields (Blair and Erlanger, 1933; Burns, 1955; Chalazonitis, 1968; Cole, 1968; Kerkut and Thomas, 1965; McCulloch, 1959; McCulloch et al., 1962; Pecher, 1939). Membrane potential is subject to random fluctuations (Fatt and Katz, 1950), a significant source of which appears to be related to the net flow of potassium ions through the membrane (Poussart, 1969; Verveen and Derksen, 1968). Another possible source of membrane potential fluctuations in neurons is spontaneous miniature PSPs that might occur randomly, irrespective of activity in presynaptic terminals (Katz and Miledi, 1963; Blankenship and Kuno, 1968). In some motoneurons, at least, synaptic inputs apparently account for most of the observed variability in the output of the neuron (Calvin and Stevens, 1968). Additional variability arises in synaptic transmission (Chalazonitis, 1968; Eccles, 1964; Burke, 1967), possible blockage of impulse propagation in presynaptic terminal branches (Allanson, 1958) and possible generation of impulses due to extraneous factors, such as increased concentration of potassium ions (Eccles et al., 1969; Fertziger and Ranck, 1970).

If it is assumed that neuronal systems must cope with variability, the problem of reducing the effects of variability in neuronal networks becomes similar to that of reliability of computation considered by several authors (Allanson, 1958; Cowan, 1960 *a*, 1960 *b*; Elias, 1958; Lofgren, 1960; McCulloch, 1959; Muroga, 1962; Verbeek, 1960 *a*, 1960 *b*; Von Neumann, 1956; Winograd and Cowan, 1963). These treatments, however, have been almost exclusively concerned with the synthesis of reliable automata from logic elements or formal neurons of the McCulloch-Pitts type (Blum, 1960) that are subject to error. The methods used in these studies, although suggestive, are not directly applicable to structures of real neurons. Moreover, it can be argued that the concept of "error" cannot be properly applied to inherently probabilistic systems, such as neuronal networks, without some a priori assumptions as to what the "correct" states are.

Significance of a Large Number of Inputs to a Neuron

It will be shown in this section that input uncertainty can be reduced by having a large number of input lines whose activities are statistically independent. The problem under consideration can be formulated with some rigor in terms analogous to those of "shot-noise" theory (Rice, 1954; Stein, 1967) and taking into account weighted summation of excitatory postsynaptic potentials (EPSP) having specified time courses and amplitude distributions. For present purposes, however, a less rigorous but more illustrative approach will be used which, although greatly simplified, is more applicable to real neurons than the formalism of logic elements employed by Allanson (1958) and Muroga (1962).

Consider a neuron having n excitatory input lines, and choose a time interval δ such that synaptic excitations by impulses occurring in the input lines during this interval can be con-

sidered to summate spatially and/or temporally. The number of impulses occurring in the ith line during the interval δ will be considered as a random variable s_i. Let S be the sum of these random variables, i.e., $S = \sum_{i=1}^{n} s_i$. Neuronal firing would only occur when $S > T$ where T is the threshold expressed in total number of impulses on the excitatory input lines during the interval δ.

The probability distribution functions of the random variables s_i need not be known; but if these random variables are independent, then according to the central limit theorem, S expressed in standard form will be normally distributed for large values of n (Gnedenko, 1967; Gnedenko and Kolmogorov, 1968).

Let the expectation of s_i be ϵ_i and its variance be σ_i^2, calculated over the duration of any given "message". Define:

$$\epsilon = \frac{1}{n} \sum_{i=1}^{n} \epsilon_i$$

and

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2.$$

Thus ϵ and σ^2 are the mean expectation and variance, respectively, of the number of impulses on the n input lines during the time interval δ. In the special case of Poisson inputs with a mean rate λ_i on the ith input line:

$$\epsilon = \sigma^2 = \frac{\delta}{n} \sum_{i=1}^{n} \lambda_i. \tag{A 1}$$

According to the central limit theorem, the variable $Y = (S - n\epsilon)/(\sigma \sqrt{n})$ will be normally distributed for large values of n. The probability of firing, P_f, will be given by:

$$P_f = \frac{1}{\sqrt{2\pi}} \int_{Y_T}^{\infty} e^{-Y^2/2}\, dY \tag{A 2}$$

where

$$Y_T = \frac{T - n\epsilon}{\sigma\sqrt{n}} = \frac{(T/n) - \epsilon}{\sigma/\sqrt{n}}.$$

It is of interest to investigate how P_f varies with ϵ for different values of n, T, and σ. The salient features of these relations can be easily worked out. From equation A 2

$$\frac{dP_f}{d\epsilon} = \frac{1}{\sqrt{2\pi}} \cdot \frac{\sqrt{n}}{\sigma} \cdot e^{-Y_T^2/2}.$$

For constant values of T/n a family of curves is obtained for different values of $\sigma/\sqrt{n}$ (Fig. 4); all the curves of the family have their maximum slope at their common intersection point whose coordinates are: $\epsilon = T/n$ and $P_f = 0.5$. If n and σ are constant, then as T is varied, the curve corresponding to the value of $\sigma/\sqrt{n}$ will shift in the ϵ direction so that the point $(T/n, 0.5)$ lies on the curve.

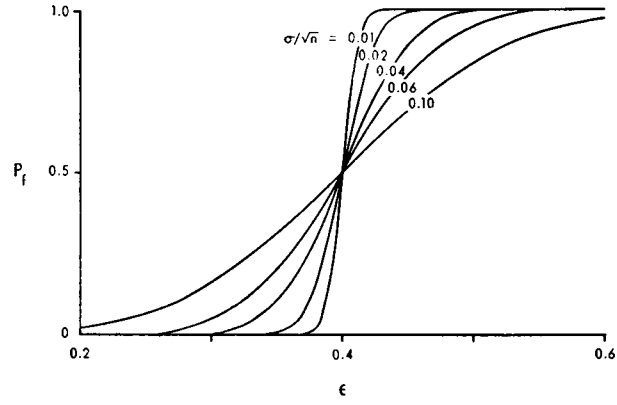

FIGURE 4 Probability of firing (P_f) of a neuron having many inputs. ϵ and σ^2 are, respectively, the mean expectation and mean variance of the number of impulses occurring on the n excitatory input lines during a summation period δ. The curves are drawn for T/n = 0.4, where T is the threshold of the neuron expressed in total number of impulses on the excitatory input lines during the interval δ.

In order to reduce input uncertainty, it would clearly be desirable to have the probability of firing almost zero for $S < T$, almost unity for $S > T$, with a sharp transition between the two states. If input activity is highly regular, σ will be small and there will be little input uncertainty, even with low values of n. The curves of Fig. 4 indicate that for any value of σ input uncertainty can be reduced if the value of n is sufficiently large.

P_f, as used above, denotes the probability of one or more impulses being discharged by the neuron during the interval δ. In principle, the above argument can be extended to determine P_k, the probability of k or more impulses being discharged during the interval δ, where k can take integer values between 1 and h, the maximum number of impulses that the neuron can discharge during the interval δ. P_k will be given by equation A 2 but with T_k used instead of T, where $T_k > T$ is a "threshold" such that if $S > T_k$ then k or more impulses will be discharged during the interval. In effect, the curves of Fig. 2 will be shifted to the right so that the abscissa of the common intersection point becomes $\epsilon = T_k/n$; the same conclusions apply as to the reduction of input uncertainty by large values of n.

It can be seen that correlated activity on the input lines is detrimental to the reduction of input uncertainty. In the extreme case, if all inputs are perfectly correlated, they will be equivalent to a single line, as far as input uncertainty is concerned, and no measures on the input side can be effective in reducing this uncertainty. It may be concluded, therefore, that input uncertainty can be reduced by having a large number of input lines whose activities, in the ideal case, are statistically independent. This conclusion is intuitively plausible, since the variability due to a large number of statistically independent inputs will in effect be "smoothed out". In addition, the threshold must not be too low; otherwise, the neuron might be fired by stray inputs or "noise" alone. This would partially defeat the purpose of having a large number of inputs.

REFERENCES

ALLANSON, J. T. 1958. Proceedings 1st International Congress on Cybernetics, Namur. Gauthier-Villars, Paris.

BELL, C. C., and R. S. DOW. 1967. *Neurosci. Res. Program Bull.* 5(2):517.

BELL, C. C., and R. J. GRIMM. 1969. *J. Neurophysiol.* **32**:1044.
BENNETT, M. V. L. 1969. *In* Neurobiology of Cerebellar Evolution and Development. R. Llinás, editor. American Medical Association Education and Research Foundation, Chicago. 242.
BLAIR, E. A., and J. A. ERLANGER. 1933. *Amer. J. Physiol.* **106**:524.
BLANKENSHIP, J. E., and M. KUNO. 1968. *J. Neurophysiol.* **31**:195.
BLUM, M. 1960. *In* Principles of Self-Organization. H. von Foerster and G. Zopf, editors. Pergamon Press, Inc., New York. 95.
BRAITENBERG, V. 1967. *Progr. Brain Res.* **25**:334.
BULLOCK, T. H. 1969. *In* Neurobiology of Cerebellar Evolution and Development. R. Llinás, editor. American Medical Association Education and Research Foundation, Chicago. 536.
BURKE, R. E. 1967. *J. Neurophysiol.* **30**:1114.
BURNS, B. D. 1955. *J. Physiol.* (*London*). **127**:168.
CALVIN, W. H., and C. F. STEVENS. 1968. *J. Neurophysiol.* **31**:574.
CHALAZONITIS, N. 1968. *Ann. N. Y. Acad. Sci.* **147**:421.
COLE, K. C. 1968. Membranes, Ions and Impulses. University of California Press, Berkeley, Calif.
COWAN, J. D. 1960 *a*. *In* Principles of Self-Organization. H. von Foerster and G. Zopf, editors. Pergamon Press, Inc., New York. 135.
COWAN, J. D. 1960 *b*. Bionics Symposium. WADD Technical Report 60-600. 93.
DOW, R. S., and G. MORUZZI. 1958. The Physiology and Pathology of the Cerebellum. University of Minnesota Press, Minneapolis.
ECCLES, J. C. 1964. The Physiology of Synapses. Academic Press Inc., New York.
ECCLES, J. C. 1969 *a*. *In* Information Processing in the Nervous System. K. N. Leibovic, editor. Springer-Verlag New York, Inc., New York. 245.
ECCLES, J. C. 1969 *b*. *Naturwissenschaften*. **56**:525.
ECCLES, J. C., D. S. FABER, J. T. MURPHY, N. H. SABAH, and H. TÁBOŘÍKOVÁ. 1970. *In* Excitatory Synaptic Mechanisms. P. Andersen and J. K. S. Jansen, editors. Oslo University Press, Oslo.
ECCLES, J. C., M. ITO, and J. SZENTÁGOTHAI. 1967. The Cerebellum as a Neuronal Machine. Springer-Verlag New York, Inc., New York.
ECCLES, J. C., H. KORN, H. TÁBOŘÍKOVÁ, and N. TSUKAHARA. 1969. *Brain Res.* **15**:276.
ECCLES, J. C., R. LLINÁS, and K. SASAKI. 1966 *a*. *Exp. Brain Res.* **1**:82.
ECCLES, J. C., R. LLINÁS, and K. SASAKI. 1966 *b*. *J. Physiol.* (*London*). **182**:316.
ECCLES, J. C., K. SASAKI, and P. STRATA. 1967. *Exp. Brain Res.* **3**:58.
ECCLES, J. C., H. TÁBOŘÍKOVÁ, and N. TSUKAHARA. 1970. *Brain Res.* **17**:87.
ELIAS, P. 1968. *I.B.M.* (*Int. Bus. Mach. Corp.*) *J. Res. Develop.* **2**:346.
EVARTS, E. V., and W. T. THACH. 1969. *Physiol. Rev.* **49**:451.
FATT, P., and B. KATZ. 1950. *Nature* (*London*). **166**:597.
FERTZIGER, A. P., and J. B. RANCK, JR. 1970. *Exp. Neurol.* **26**:571.
FOX, C. A., and R. S. SNIDER, editors. 1967. *Progr. Brain Res.* **25**.
GNEDENKO, B. V. 1967. The Theory of Probability. 4th edition. Chelsea Publishing Co., Bronx, New York.
GNEDENKO, B. V., and A. N. KOLMOGOROV. 1968. Limit Distributions for Sums of Independent Random Variables. Addison-Wesley Publishing Co., Inc., Reading, Mass.
ITO, M., N. KAWAI, and M. UDO. 1968. *Exp. Brain Res.* **4**:310.
ITO, M., M. YOSHIDA, K. OBATA, N. KAWAI, and M. UDO. 1970. *Exp. Brain Res.* **10**:64.
JANSEN, J., and A. BRODAL. 1954. Aspects of Cerebellar Anatomy. Johan Grundt Tanum Forlag, Oslo, Norway.
KATZ, B., and R. MILEDI. 1963. *J. Physiol.* (*London*). **168**:389.
KERKUT, G. A., and R. C. THOMAS. 1965. *Comp. Biochem. Physiol.* **14**:167.
LLINÁS, R., editor. 1969. Neurobiology of Cerebellar Evolution and Development. American Medical Association Education and Research Foundation, Chicago.
LLINÁS, R., and D. E. HILLMAN. 1969. *In* Neurobiology of Cerebellar Evolution and Development. R. Llinás, editor. American Medical Association Education and Research Foundation. 43.
LOFGREN, L. 1960. *In* Principles of Self-Organization. H. von Foerster and G. Zopf, editors. Pergamon Press, New York. 181.

McCULLOCH, W. S. 1959. *In* Mechanization of Thought Processes. National Physical Laboratory Symposium. H.M.S.O., Teddington, England. 2(10): 611.
McCULLOCH, W. S., A. ARBIB, and J. D. COWAN. 1962. *In* Self-Organizing Systems. M. C. Yovits, G. T. Jacobi, and G. D. Goldstein, editors. Spartan Books, New York. 49.
MAGDALENO, R. E., D. T. McRUER, and G. P. MOORE. 1968. *NASA Contract Rep.* CR-1212.
MUROGA, S. 1962. *In* Self-Organizing Systems. M. C. Yovits, G. T. Jacobi, and G. D. Goldstein, editors. Spartan Books, New York. 243.
PAPOULIS, A. 1965. Probability, Random Variables and Stochastic Processes. McGraw-Hill Book Company, New York.
PECHER, C. 1939. *Arch. Int. Physiol. Biochem.* **49**:129.
POUSSART, D. J. M. 1969. *Proc. Nat. Acad. Sci. U.S.A.* **64**:95.
PROVINI, L., S. REDMAN, and P. STRATA. 1968. *Exp. Brain Res.* **6**:216.
RICE, S. O. 1954. *In* Selected Papers on Noise and Stochastic Processes. N. Wax, editor. Dover Publications, Inc., New York. 133.
RIORDAN, J. 1968. Combinatorial Identities. John Wiley & Sons, Inc., New York.
RUBIA, F. J., and J. B. PHELPS. 1970. *Pfluegers Arch.* **314**:68.
SABAH, N. H., and J. T. MURPHY. 1971. *Biophys. J.* **11**:414.
SEGUNDO, J. P., D. H. PERKEL, and G. P. MOORE. 1966. *Kybernetik.* **3**:67.
SNIDER, R. S. 1967. *Progr. Brain Res.* **25**:322.
SNIDER, R. S., and J. MITRA. 1970. *In* The Cerebellum in Health and Disease. W. S. Fields and W. D. Willis, Jr., editors. Warren H. Green, Inc., St. Louis, Mo. 319.
STEIN, R. B. 1967. *Biophys. J.* **7**:37.
SZENTÁGOTHAI, J. 1968. *Proc. I.E.E.E.* (*Inst. Elec. Electron. Eng.*). **56**:960.
THACH, W. T. 1967. *J. Neurophysiol.* **30**:675.
VERBEEK, L. A. M. 1960 *a*. *In* Principles of Self-Organization. H. von Foerster and G. Zopf, editors. Pergamon Press, Inc., New York. 121.
VERBEEK, L. A. M. 1960 *b*. Bionics Symposium. WADD Technical Report 60–600.
VERVEEN, A. A., and H. E. DERKSEN. 1968. *Proc. I.E.E.E.* (*Inst. Elec. Electron. Eng.*). **56**:906.
VON NEUMANN, J. 1956. *In* Automata Studies. C. E. Shannon and J. McCarthy, editors. Princeton University Press, Princeton, New Jersey. 43.
WINOGRAD, S., and J. D. COWAN. 1963. Reliable Computation in the Presence of Noise. M.I.T. Press, Cambridge, Mass.