AMRL-TR-68-189

COMPOUND NET MODEL OF THE CEREBRAL CORTEX

E. M. HARTH

B. BEEK

T. J. CSERMELY

R. D. LINDSAY

FOREWORD

This research was conducted by the Physics Department of Syracuse University, Syracuse, New York, under contract number F33615-67-C-1413, and in support of Project 7232, "Research on the Logical Structures and Function of the Nervous System," and Task 723205 "Mathematical Theory of Neural and Mental Processes," and Unit No. 011. This research was sponsored by the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio. Professor Erich M. Harth was director of the study for Syracuse University. Dr. Hans Oestreicher, Chief, Mathematics and Analysis Branch, Biodynamics and Bionics Division, was contract monitor for the Aerospace Medical Research Laboratory. This work began on 1 February 1967 and was completed in November 1968. The authors are pleased to acknowledge partial support received from the Office of Naval Research under Contract No. N00014-67-A-0378-0001.

This technical report has been reviewed and is approved.

C. H. KRATOCHVIL, Colonel, USAF, MC
Commander
Aerospace Medical Research Laboratory

ABSTRACT

The central nervous system is believed to consist of a network of neurons, combining in its structure genetically determined design with probabilistic features. A model is described in which probabilistic nets are used as building blocks from which systems can be assembled, approximating for example the interaction between sensory projection areas in the cortex with cortical association areas. The operation of such *compound nets* is based on the dynamics of single probabilistic nets and on a set of assumptions concerning the coding of sensory information into the language of central neuronal activity. Data describing the dynamics of simple nets are presented. It is found that various types of information processing can be explained by applying the hypothesis of synaptic facilitation to compound nets. Sustained neural activity is investigated, particularly the problem of ergodicity. Highly non-ergodic types of reverberations were found both theoretically and by computer simulation for certain types of nets.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (cont'd)

LIST OF SYMBOLS

A	total number of neurons in net.
$a(t)$	number of active neurons at time t.
$\underline{a}(t)$	neuronal state, i.e., the set of active neurons at time t.
$\alpha=a/A$	fraction of active neurons.
$a_i(t)$	the state of the i-th neuron at time t (1 if active, 0 if inactive).
μ_+	the average number of neurons receiving EPSP's from an excitatory neuron.
μ_-	the average number of neurons receiving IPSP's from an inhibitory neuron.
k_+	the average EPSP.
k_-	the average IPSP (a negative number).
k_{ij}	the PSP generated at the i-th neuron at time $(n+1)\tau$ as a result of activity of the j-th neuron at time $n\tau$. This is also called the *coupling coefficient* from the j-th to i-th neuron.
$\underline{k}_{ij}$	the matrix of coupling coefficients.
θ	the firing threshold, or minimum total PSP required to trigger a neuron.
η	the minimum number of average EPSP's required to trigger a neuron if no IPSP's are present.
τ	synaptic delay.
n	an integer used for counting intervals of τ.
α_n	the fraction of active neurons at $t=n\tau$.
$\langle\alpha_{n+1}\rangle$	the expectation value for the fraction of active neurons at $t=(n+1)\tau$.
h	the fraction of inhibitory neurons.
α^*	the fraction of active neurons if activity is in stable equilibrium.

SECTION I

INTRODUCTION

In a recent paper (Harth and Edgar, 1967), which will be referred to as (I), the behavior of simple, randomly connected neural nets was investigated analytically and by computer simulation. Under appropriate conditions such nets can be made to perform a variety of cognitive functions.

The assumptions made in (I) were highly restrictive both as regards network structure and the functioning of individual neurons. Time dependencies were ignored and the connectivity was completely random. The system could be considered to represent a primitive model of one component of the association cortex of the brain, and that it would would lend itself to being extended in the future to conform with a somewhat larger body of biological reality. Such an extension will be attempted in the present paper.

Attempts to understand the functioning of the brain often take the form of contriving networks of interconnected neurons and studying the dynamics of such systems. The problems one encounters may be classified roughly under these three headings:

• Anatomy and physiology. Questions concerning the functioning of the individual neuron, the role of the synapse and all of the micro- and macrostructure of the brain belong in this category. Information concerning many of these subjects is still fragmentary. Some of the most fundamental questions, such

as the nature of the physical changes that must underlie learning have not been answered. In other respects the biologist often succeeds in uncovering a plethora of detail which the theorist at this stage is incapable of absorbing into a meaningful pattern.

• Behavior of neural nets. Calculation of the dynamics of neural activity, even under the most stringent simplifying assumptions, presents formidable mathematical difficulties. Only the most primitive types of nets, such as homogeneous, randomly connected nets, have been studied in any detail. Computer simulation is often resorted to as a means of circumventing computational difficulties. In doing so, net sizes are in practice restricted to at most a few thousand neurons.

• The meaning of neural activity. This complex of problems concerns the presumed correspondence between neuronal activity and such psychologically defined functions as *to know, to remember, to feel, to perceive, to associate,* etc. Put differently, we would like to know what are the parameters in neural activity that determine what we may loosely call our *state of mind*, a quantity observable only by a single person. Is this state completely determined by the time dependence of all of the action potentials emanating from about 10^{10} neurons? Or must a complete description also contain the specification of such variables as the subthreshold postsynaptic potentials or perhaps also glial processes?

The theorist, of course, always hopes for simplicity, so the non linear character of the response of the individual

neurons to membrane changes received within its dendritic field--the so-called all-or-none-effect--has been taken as a hint that the results of these integrations and the ensuing action potentials, are the physically significant variables. A description of neural dynamics would thus be reduced to the specification of firing times for all neurons in the net, still an enormous task. Conceivably some degeneracy still remains, that is to say, it may be possible to eliminate further details of the description of the neuronal state without attendant loss of the information carried by the net. It has been postulated, for example, that information resides in the temporal distribution of activity, and that therefore a significant amount of information could be elicited from the net by monitoring a single neuron (Bullock, 1967).

This has been an attractive postulate for a number of reasons. Not only are we used to employing a frequency code in communication, but it is also known experimentally (Granit and Renkin, 1961) that, at least in the peripheral nervous system, firing frequency is often directly related to sensory parameters such as intensity of a stimulus. Also, we know a lot more about the temporal response of single neurons than about the spatial distribution of neuronal activity. This is merely due to the fact that neurophysiologists can readily monitor a single neuron while the simultaneous detailed observation of more than a very small number of neurons presents great technical difficulties.

We shall nevertheless maintain the point of view taken in

(I) that the information we seek is expressed by the spatial distribution of activity, specifically by the *set* of neurons active at a given instant. We call this the *neuronal state* at time t: it may be expressed by the column matrix:

$$\underline{a}(t) = \begin{bmatrix} a_i(t) \\ \vdots \\ a_A(t) \end{bmatrix}$$

where A is the total number of neurons in the net and $a_i(t)$ equals 1 or 0 according to whether or not the i-th neuron fires at time t. The matrix $\underline{a}(t)$ thus describes the instantaneous state of the net. Implicit in this formulation is the assumption that axonal transmission over intra-cerebral distances is fast compared with synaptic delays, refractory periods,and other time constants, so that simultaneity may be defined not just locally, but for the entire network under study. This assumption appears to be reasonably justified (Ruch and Patton, 1966; Ochs, 1965).

SECTION II

DESIGN AND RANDOMNESS

The controversy between these two viewpoints concerning cerebral organization has a long history, and may be compared with the classic controversy over the wave or particle nature of light in physics. Both points of view have amassed an impressive repertoire of supporting evidence. The wave-particle controversy finally ended not with the victory of one and the defeat of the other, but with a reconciliation of the seemingly irreconcilable points of view--the principle of complementarity.

Unquestionably the brain has structure, that is present at birth, serves important biological functions and is presumably genetically determined. In recent years evidence has been accumulating for the existence of highly specific, inborn neural circuits in parts of the cerebral cortex, in seeming contradiction to the *mass action principle* and the concept of cortical *equipotentiality* (Lashley, 1931; 1933; 1950). The similarity of the design principles extant in nearly all sensory systems (v. Békésy, 1967) makes it tempting to postulate that this kind of genetically determined neural organization is characteristic of the entire neocortex. An extreme view of this work would be that statement that every synaptic connection throughout the nervous system is genetically determined. It has been argued that the limited information content of the DNA molecule precludes such a view and that therefore elements of randomness must be present. But this argument if specious; it is clearly possible

to specify the connectivity in every detail and yet use only a very limited amount of information, in the same way in which we can describe the space coordinates of large numbers of atoms in a crystal merely by giving the parameters of the lattice. Thus, any symmetry principle that can be found, reduces the amount of information required for a detailed specification of the system. Symmetry principles are information-saving devices. In the case of neural connectivity in the higher cortical areas we know as yet too little about symmetry laws to be able to tell whether all of the remaining information content of a particular connectivity can be supplied by our genes.

Opposed to the view of complete specification in the design of the neural net is the assumption of a *probabilistic* net in which stochastic rules form a significant part of the blueprint. Whether or not a particular network is probabilistic is of course a question that cannot be readily answered. It assumes meaning if we inquire, for example, into the processes of formation. There is however another way of defining randomness, and this is of particular interest to the theorist attempting to find a model of brain function. Suppose that we were to generate a class of neural networks by allowing certain variables to vary in a random fashion within predetermined probability distributions, and suppose that each of these nets[1] is found to

[1] Some pathological configurations whose occurrence would be rare, may be possible exceptions.

exhibit the same behavior. We would then say that the behavior is characteristic of this *class* of probabilistic nets. It is now conceivable that a brain which, by virtue of its formation, is completely deterministic, i.e., its connectivity is completely specified by a combination of genetic information and symmetry principles, is found nevertheless to be, functionally, a probabilistic net. This would merely imply that the biological solution is not the only one, but that some parameters could have been left to chance without ill effects.

As a practical matter, a probabilistic model is often used because knowledge concerning the precise connectivity of real neural nets is lacking, and because of the assumption that the behavior of the net is not affected by allowing the connectivity to vary within specified limits. There is however yet another aspect to randomness: it is the assertion that a lack of organization is a functional property of the net, or that a certain amount of disorder is required to allow that establishment of the kind of ordering we associate with learning, expecially if the tasks to be learned are unpredictable.

We can now write down sets of properties which, if not mutually exclusive, are at least antagonistic, one set loosely associated with deterministic and one with probabilistic nets (table I.).

Table I

Deterministic Nets	Probabilistic Nets
Design	Disorder, Randomness
Stability, Consistency	Adaptability, Learning
Specificity, Localization	Equipotentiality
Resettability	Hysteresis, Memory
Predictable Tasks	Unpredictable Tasks

Design, if it is of functional significance, apparently implies a predictability of tasks. In this sense the brain is genetically designed from birth to prevent the infant from falling off a ledge. On the other hand, the necessity of stopping an automobile upon a certain visual cue (red light) is a genetically unpredictable cerebral task, hence must be learned.

We argue that there exists in the brain an immense pool of specialized circuitry capable of taking care of every possible exigency, and that circuits are activated as experience requires. Bremermann (1967) showed that the "genetic cost" of such a system would be staggering. On the other hand, it was shown in (I) how a completely random net can become structured as a result of sensory inputs and perform a variety of learned tasks. We shall continue to assume that the portions of the cortex whose function it is to carry out the fundamentally unpredictable tasks, the so-called association cortex, is a relatively unstructured probabilistic net.

In summary, we believe that, as in the particle-wave duality, we have in the brain a duality of design and randomness, with design predominating in the more peripheral sensory and motor portions, and randomness in the highest, or association areas of the cortex. Possibly no area is completely designed, nor completely probabilistic.

SECTION III

DYNAMICS OF PROBABILISTIC NETS

In many quantitative studies, investigators sought to establish brain-like functions by computer simulation or mathematical analysis. This work was to a large extent inspired by Hebb's hypothesis of synaptic facilitation (Hebb, 1949) according to which the effectiveness of a synapse in triggering a postsynaptic neuron is enhanced whenever a presynaptic action potential succeeds in firing the cell. This facilitation of selected pathways was believed to constitute the physical basis of memory.

If memory is to be established in the network as a result of a single experience one may argue that mechanisms must exist that produce sustained reexcitation of the same neural pathways following the sensory input. Hebb postulated that such reexcitation can occur in a richly connected network through reverberations and concomitant synaptic facilitation of neurons involved in the reverberating circuits. As a result of this process and the interaction between the various reverberatory pathways a subset of the neural net would differentiate itself into what Hebb called a *cell assembly*. This would represent something like the emergence of meaning in the firing patterns of cortical neurons. Accordingly much subsequent work was aimed at establishing the existence of quasi-stable modes of sustained neural activity.

Most of the studies concerning the dynamics of neural nets performed to date had the following features in common: The

network was assumed to consist of randomly interconnected neuron-like elements possessing these properties:

all-or-none law of discharge;

spatial and temporal summation of stimuli;

a synaptic delay; and

a relative refractory period.

Refractoriness was assumed to be absolute for a fixed period following an action potential. In some cases, it was assumed that this was followed by a *relative refractory period*. Times were usually quantized in units of the synaptic delay. The network itself was described in terms of the following parameters:

the number of outgoing and incoming connections per neuron;

the threshold of neurons; and

the effectiveness of synaptic coupling ("coupling coefficient").

The latter parameter expressed the magnitude of the postsynaptic potential and the threshold gave the amount of total excitation needed to fire a particular neuron.

In one of the earlier studies of network dynamics, Beurle (1956) used a quasi-field theoretical approach: the neurons were assumed to be distributed continuously with a volumetric density ζ and the axonal and dendritic fields were specified by the connectional density μ. For mathematical convenience, variations were assumed to occur in the x-direction only and inhibitory connections were excluded from consideration.

Based on Scholl's findings (Scholl, 1953) an exponential form was assumed for $\mu(x)$, i.e., interactions between neurons were restricted mainly to near neighbors. The analysis showed that the fractional activity α_{n+1} at $t=(n+1)\tau$ is related to the activity α_n at $t=n\tau$ in a way depicted in figure 1. From the graph the net tends toward one of two stable states, namely complete cessation of activity or saturated activity. Thus the information content of the net is just one bit. Subsequently Beurle pointed out that the inclusion of inhibition would allow for sustained intermediate activity (Beurle, 1962a; 1962b).

In a later study, Ashby, *et al.* (1962) assumed that the activity of the net could be described by a single variable, the probability of firing of a neuron in a given interval of time. Like Beurle (1956), he did not include inhibition in his model, and the results were quite similar.

Griffith (1963; 1965) included inhibitory neurons in the net and was able to demonstrate that sustained oscillation of activity is indeed possible in the net under these conditions. Subsequently he developed a field theoretical approach and succeeded in obtaining intermediate activity in the net.

Working with so-called *discrete nets*, Rapoport (1952) showed that excitation will spread through the net until a steady state is reached provided the initial activity is larger than a certain critical value. Otherwise the activity tends to zero. Subsequently Trucco (1952) extended this work by deriving additional conditions for what he called *ignition phenomena*.

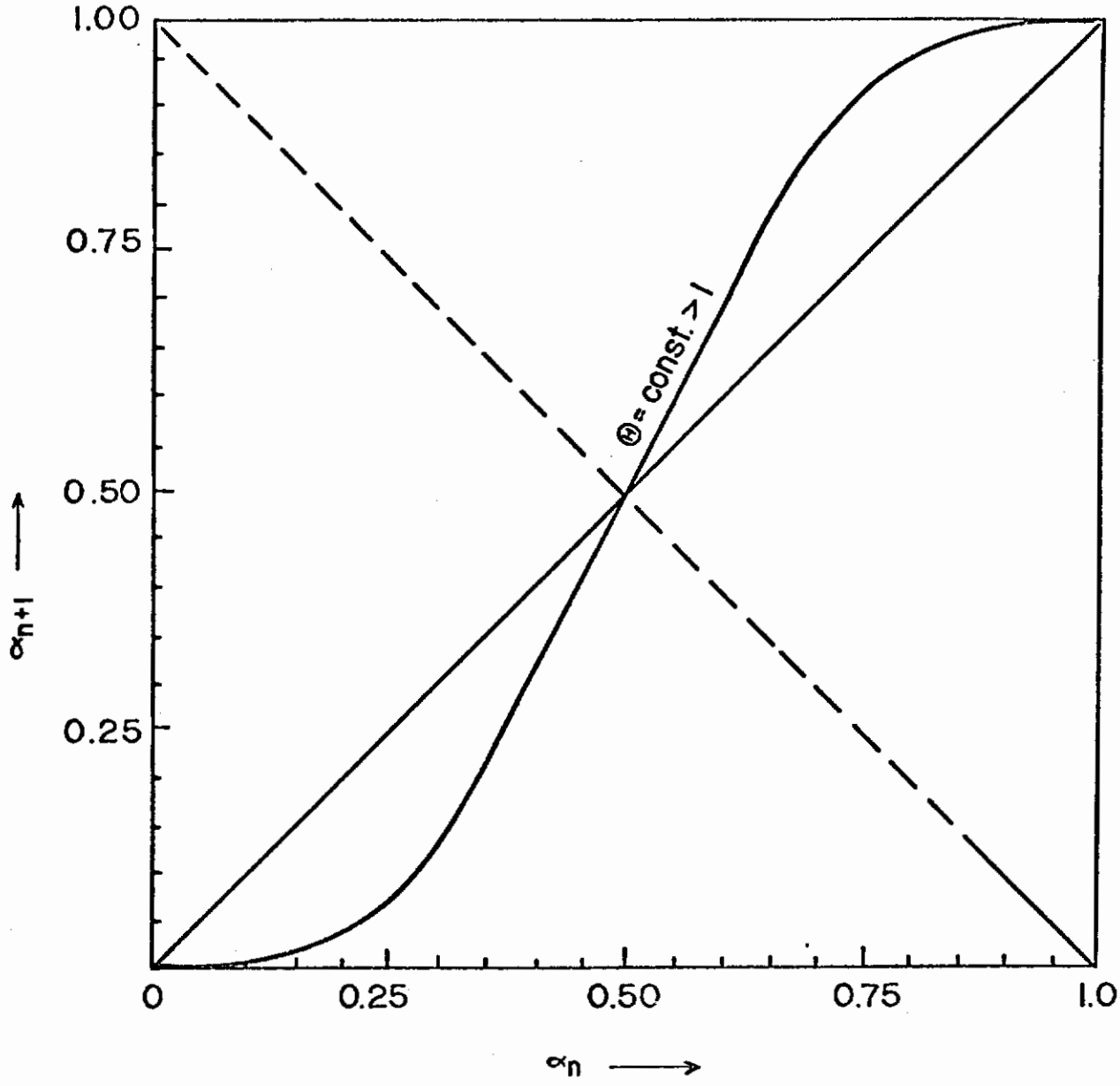

Figure 1. Relationship between fractional activities arising one synaptic delay apart in a nonrefractory network of excitatory neurons (Beurle, 1956).

In an extensive mathematical analysis Allanson (1956) investigated the dependence of net dynamics on the parameters of a randomly connected neural net. Depending on the parameters, such nets could oscillate continuously, remain quiescent, or exhibit damped oscillations. Inhibition was included in Allanson's model.

The first analysis of net dynamics using simulation by digital computer was carried out by Rochester, *et. al.* (1956). In their first study, a probabilistic network of 64 neurons was simulated on an IBM 704 digital computer. All connections were excitatory and synaptic facilitation, as outlined by Hebb, was included. It was found that such a network exhibits *diffuse reverberation*, an aperiodic activity involving virtually the entire net.

In the second phase of their work, the number of neurons was increased to 512 and inhibitory connections were permitted to occur. The originally random connectivity pattern in the net was modified by introducing a "distance bias" so that near neighbors were more likely to be coupled than distant ones. Synaptic facilitation was again assumed to occur but the additional constraint was imposed that the sum of synaptic coupling strengths for the entire net had to remain constant. This caused a degeneration of the less frequently used connections. The results showed that cell assemblies are formed under these conditions and that the activity in the assemblies is reverberatory.

Farley and Clark (1961) simulated a planar net of 1296 neurons in which the interconnections were specified by two-dimensional probability distributions again introducing a "distance bias" in the network. Inhibitories were not included. In general, their results confirmed Beurle's analysis that the activity in the net leads to either saturation or quiescence. It yielded, however, the additional information that a randomly connected network can, under certain circumstances, such as low threshold and repetitive stimulation exhibit sustained oscillations, even though the net does not contain inhibitory connections.

The above results obtained by Farley and Clark were in tightly connected nets in which the connectivity pattern favored interactions between nearest neighbors only. In loosely connected nets in which connections to remote elements were equally likely, scattered activity over the entire net was seen. Under some conditions the whole net oscillated disfusely, the oscillations either continuing indefinitely or else stopping spontaneously after a few cycles.

Smith and Davidson (1962a; 1962b) in an analytical study examined networks similar to those of Farley and Clark, but included inhibitory neurons. Without inhibition, their results were similar to those of Beurle as shown in figure 2. With inhibitory connections added, the activity curves exhibited several stable and unstable equilibrium points as shown in figure 3. Periodic firing of identical subsets of elements is possible both in stable steady-state

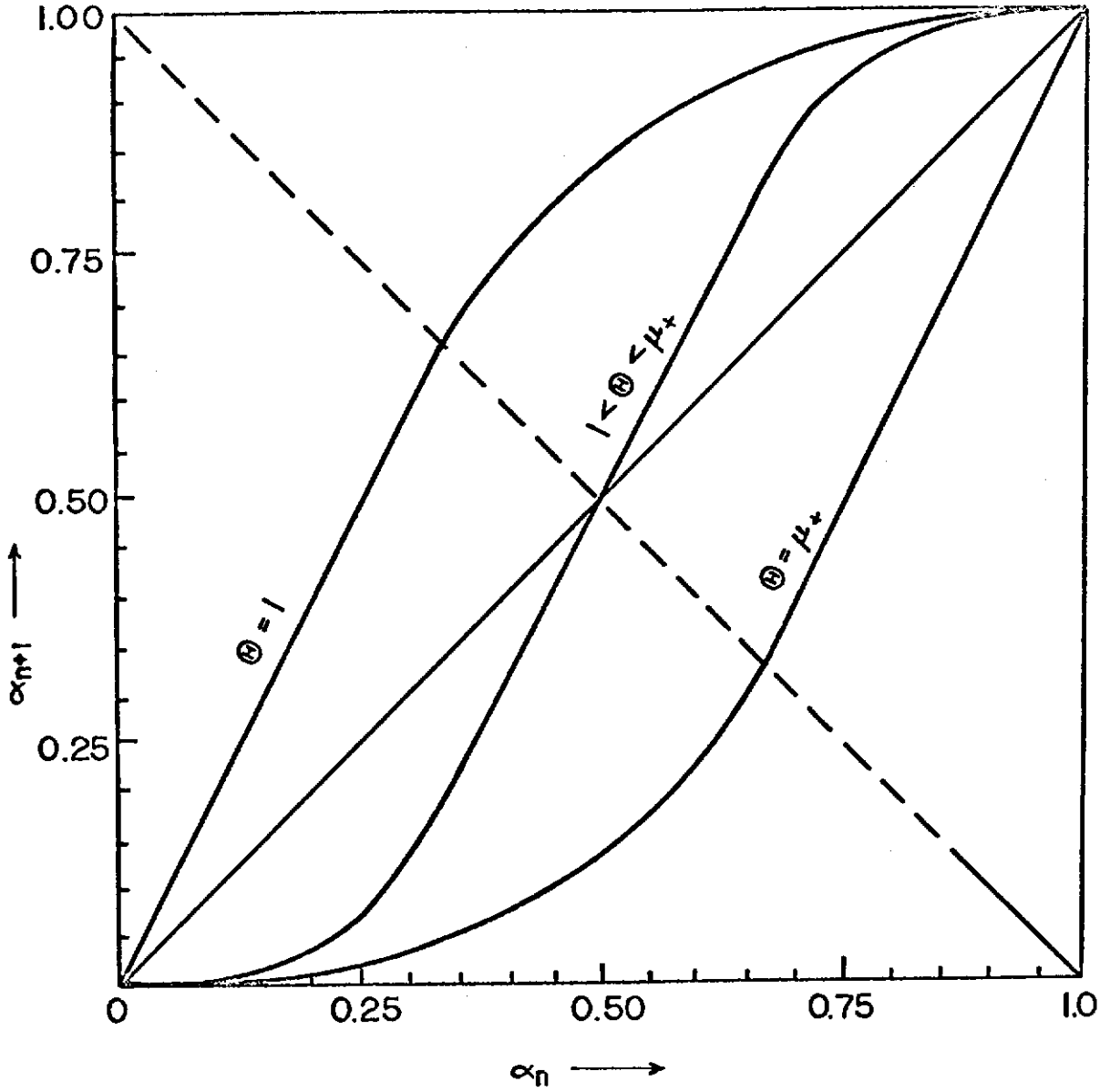

Figure 2. Relationship between fractional activities arising one synaptic delay apart in a nonrefractory network of excitatory neurons (Smith and Davidson, 1962a, 1962b).

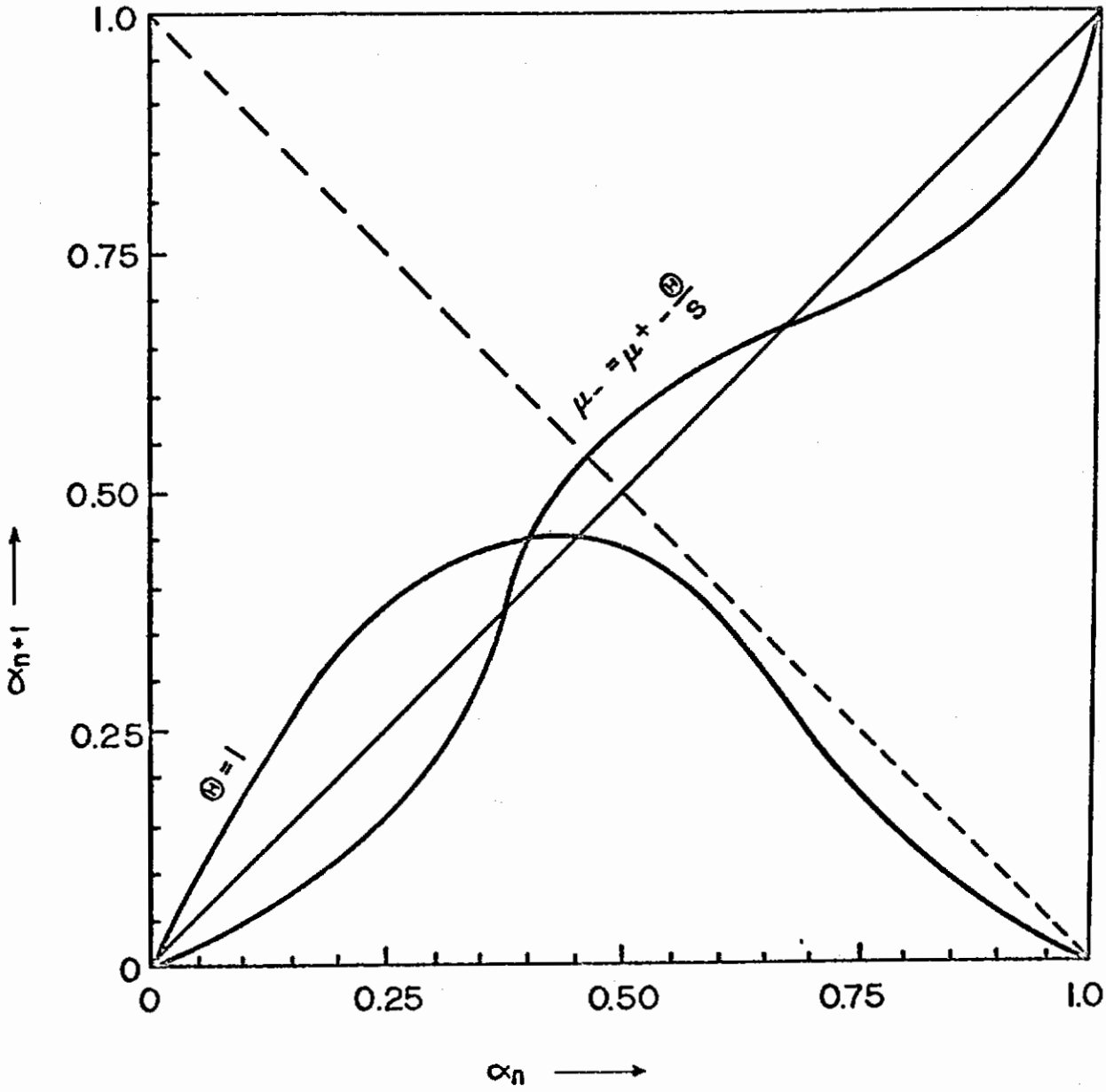

Figure 3. Relationship between fractional activities arising one synaptic delay apart in a nonrefractory network of excitatory and inhibitory neurons (Smith and Davidson, 1962a, 1962b).

activity and in oscillatory behavior.

Finally, Caianiello, *et. al.* (1967) have investigated theoretically the conditions of net structure which would lead to collective modes of excitation. Along similar lines, Ricciardi and Umezawa (1967) suggest the application of the formalism of many-body problems in physics to a description of neural nets.

We have investigated the dynamical properties of some very simple probabilistic nets. We give first a brief description of the systems to be studies and the assumptions used. A list of symbols is given in the Index.

A neural net consists of A neurons. A neuron may be either excitatory or inhibitory. The fraction of inhibitory neurons in the net is denoted by h. Each neuron is connected synaptically to a number of other neurons to which it will transmit signals, the analogues of the postsynaptic potentials (PSP's) identified by neurophysiologists. The PSP produced at the i-th neuron as a result of the *firing* of the j-th neuron is called the *coupling coefficient*[2] k_{ij}. If the ℓ-th neuron is excitatory, then all coefficients $k_{i\ell}$ are positive; if the m-th neuron is inhibitory, all the coefficients k_{im} are negative. We speak of *excitatory* and *inhibitory postsynaptic potentials* respectively (EPSP's and IPSP's). The

---

[2] A connection k_{ij} as defined here represents the sum total of all couplings from the j-th to the i-th neuron, hence may be the combined effect of any number of synapses.

average EPSP shall be denoted by k_+, the average IPSP by k_-. We also denote by μ_+ the average number of neurons receiving EPSP's from an excitatory neuron, by μ_- the average number of neurons receiving IPSP's from an inhibitory neuron. It is assumed that for each neuron the outgoing connections are randomly distributed over the entire net.

The connectivity of the net is completely described by the matrix $\underline{k}_{ij}$ of coupling coefficients. In this matrix each non-vanishing element represents a one-way connection; the absence of synaptic links *from* the ℓ-th *to* the m-th neuron is characterized by $k_{m\ell}=0$.

If the neurons are numbered in an arbitrary fashion from 1 to A, the network described above is represented by a square matrix which is uniformly *seeded* with non-vanishing elements.

The minimum number of EPSP's required on the average such that their sum is equal to or greater than the firing threshold Θ, turns out to be a useful parameter. If we call this number η, and if the function $u(x)$ is defined to be the largest integer equal to or smaller than x, then

$$\eta = u(\Theta/k_+) \tag{1}$$

Dynamical Assumptions

The number of neurons firing at any given moment is called a. It is convenient to define also the fraction of active neurons $\alpha=a/A$. Each neuron active at time t is assumed to produce the appropriate PSP's after a fixed time interval τ, the *synaptic delay*. All PSP's arriving at a neuron are summed

instantly and, if they exceed the threshold, will cause the neuron to fire without further delay. Firing is momentary and causes the neuron to be insensitive to further stimulation for a period of time called the *refractory period*. Postsynaptic potentials, of below threshold, will persist with or without decrement for a period of time, called the *summation time*. For our immediate purposes it will be assumed that the refractory period is greater than the synaptic delay, but less than twice the synaptic delay, and that the summation time is less than the synaptic delay. These assumptions lead to the following effect: if a number of neurons are fired synchronously at time t, than all neural activity resulting from this initial activity will be restricted to times $t+\tau$, $t+2\tau$, etc. Furthermore the activity at $t+n\tau$, call it α_n, uniquely determines the activity α_{n+1} occuring exactly one synaptic delay later.

In Appendix I, the expectation value $\langle\alpha_{n+1}\rangle$ for the activity generated by an activity α_n is given by

$$\langle\alpha_{n+1}\rangle \simeq (1-\alpha_n)\exp(-\alpha_n h\mu_-) \sum_{m=0}^{M} (\alpha_n h\mu_-)^m/m! \cdot$$

$$\{1-\exp[-\alpha_n(1-h)\mu_+] \sum_{\ell=0}^{\eta'-1} [\alpha_n(1-h)\mu_+]^\ell/\ell!\} \qquad (2)$$

where $\eta' = u[(\theta-mk_-)/k_+]$ and M is an integer sufficiently large so that the addition of higher terms in the summation may be neglected.

Curves of $\langle\alpha_{n+1}\rangle$ vs. α_n are shown in figure 4 for $\mu_+=5$, and $h=0$. A number of important features are apparent. All

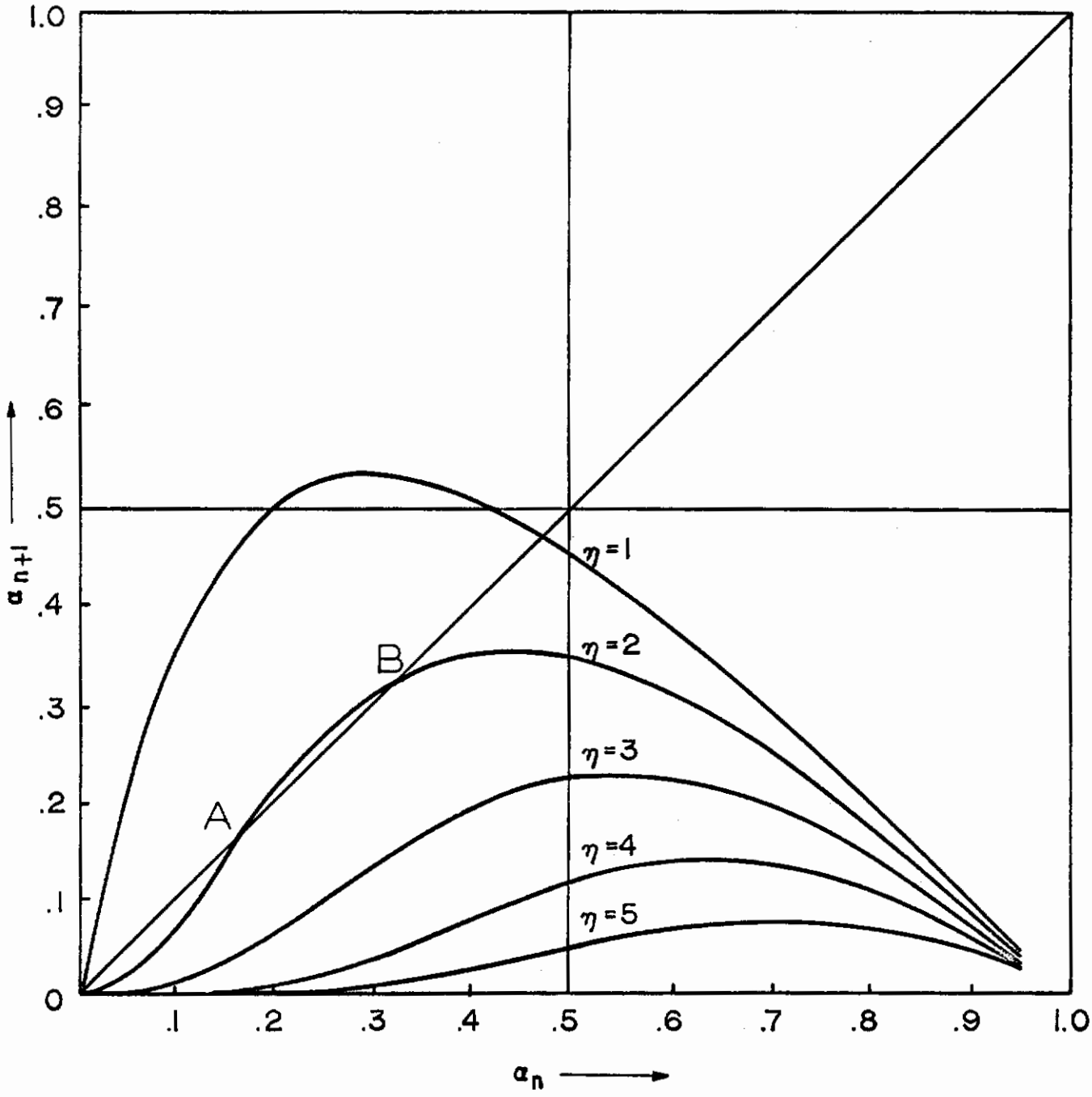

Figure 4. Curves of $\langle\alpha_{n+1}\rangle$ vs α_n for nets with $\mu_+=5$, $h=0$. *A* is a point of unstable, *B* a point of stable equilibrium.

curves go to zero for $\alpha_n=0$ and $\alpha_n=1$; the last is a saturation effect having to do with the refractoriness of the neurons.

The slopes of the curves at the origin are of interest. If we take $(\partial<\alpha_{n+1}>/\partial\alpha_n)_{\alpha_n=0}$ from Equation 2, we find that all terms of the summation over m vanish with the exception of $m=0$. Accordingly the parameter η' in the second summation reduces to $\eta'=\eta$ according to (2) and we obtain

$$(\partial<\alpha_{n+1}>/\partial\alpha_n)_{\alpha_n=0} \simeq (1-h)\mu_+ - \{\partial/\partial\alpha_n \sum_{\ell=0}^{\eta-1} [\alpha_n(1-h)\mu_+]^{\ell}/\ell!\}_{\alpha_n=0}$$

But the derivative of the summation over ℓ has only one non vanishing term, corresponding to $\ell=1$, which gives $(1-h)\mu_+$. Accordingly we obtain

$$(\partial<\alpha_{n+1}>/\partial\alpha_n)_{\alpha_n=0} = \begin{cases}(1-h)\mu_+ \text{ for } \eta=1\\ 0 \text{ for } \eta\geqslant 2\end{cases} \qquad (3)$$

The curve for $\eta=1$ in figure 4 is clearly a special one. For $(1-h)\mu_+>1$ and for the smallest activities α_n we have according to (3) the condition that $\alpha_{n+1}>\alpha_n$, i.e. the activity will build up. The point $\alpha_n=0$ is therefore a point of unstable equilibrium, since the smallest fluctuation would cause the net to go into sustained activity. The equilibrium activity is the point near $\alpha_n\leq 0.5$ where the curve crosses the 45° line (figure 4).

The fact that all curves of $<\alpha_{n+1}>$ vs. α_n for $\eta\geqslant 2$ start with zero slope, which is also apparent in figure 4, has an

interesting consequence. In all these nets, for small initial activities, $\langle\alpha_{n+1}\rangle$ will always be less than α_n, hence activity in the net will rapidly be extinguished. If α_n is sufficiently small, α_{n+1} will be zero for practical purposes. In figure 4 for $\mu_+=5$, $\eta=5$ this condition is satisfied for initial activities as high as about 0.25. We call such nets *highly damped.* For some nets the curve $\langle\alpha_{n+1}\rangle$ vs. α_n will cross into the region above the 45° line where $\langle\alpha_{n+1}\rangle > \alpha_n$. This is illustrated in figure 4 by point *A* on the $\eta=2$ curve. Eventually the curve must descend again into the lower region because of the assumed refractoriness; this will happen at $\alpha_n<0.5$ (point *B* in figure 4). The first crossing is clearly a point of unstable equilibrium, below which activity decreases with time and above which it will build up to the value of the stable equilibrium *B*. The entire net has thus the property of a highly nonlinear decision element, not unlike the single neuron, having two stable states and a threshold for being excited from the inactive to the active state. The existence of such thresholds was pointed out by Rapoport (1952) and by Trucco (1952), and the bistable character of nets similar to ours was discussed by Ashby, *et al.* (1962) and Griffith (1963).

Finally for η sufficiently high, e.g., $\eta\geq 3$, $\mu_+=5$ (figure 4), the curve never rises above the 45° line; $\langle\alpha_{n+1}\rangle$ will be less than α_n for all values of α_n, hence any initial activity will die out in time. As a convenient classification we shall

denote by *class A* all nets which will produce sustained activity for any initial activity α_n, however small, by *class B* those which have a threshold $\alpha_n > 0$ for being triggered into sustained activity, and by *class C* those which will show monotonically decreasing activity for all initial activities. The effect of inhibitory neurons is shown in figure 5 in which $\langle\alpha_{n+1}\rangle$ is plotted against α_n for a net having the same multiplicity of excitatory connections $\mu_+ = 5$ but in which 5% of all neurons are inhibitory, with an average multiplicity $\mu_- = 6$.

Time Course of Neural Activity

The formalism of the preceding section may be used to predict the development of neural activity in the net following a given initial excitation. The assumption that the state α_{n+1} is determined only by the preceding state α_n, rests of course on the assumption that the summation time for PSP's does not exceed the period τ of one synaptic delay, and that the refractory period is greater than τ but less than 2τ.

A graphic way of obtaining the time course of activity is shown in figure 6 for a typical class *B* net. A time axis, labeled *n*, extends down from the abscissa. Here *n* refers to the number of synaptic delays that have elapsed since the initial activity; the lines with arrows illustrate the method by which each successive activity may be found. The time course of activity is shown for three different initial

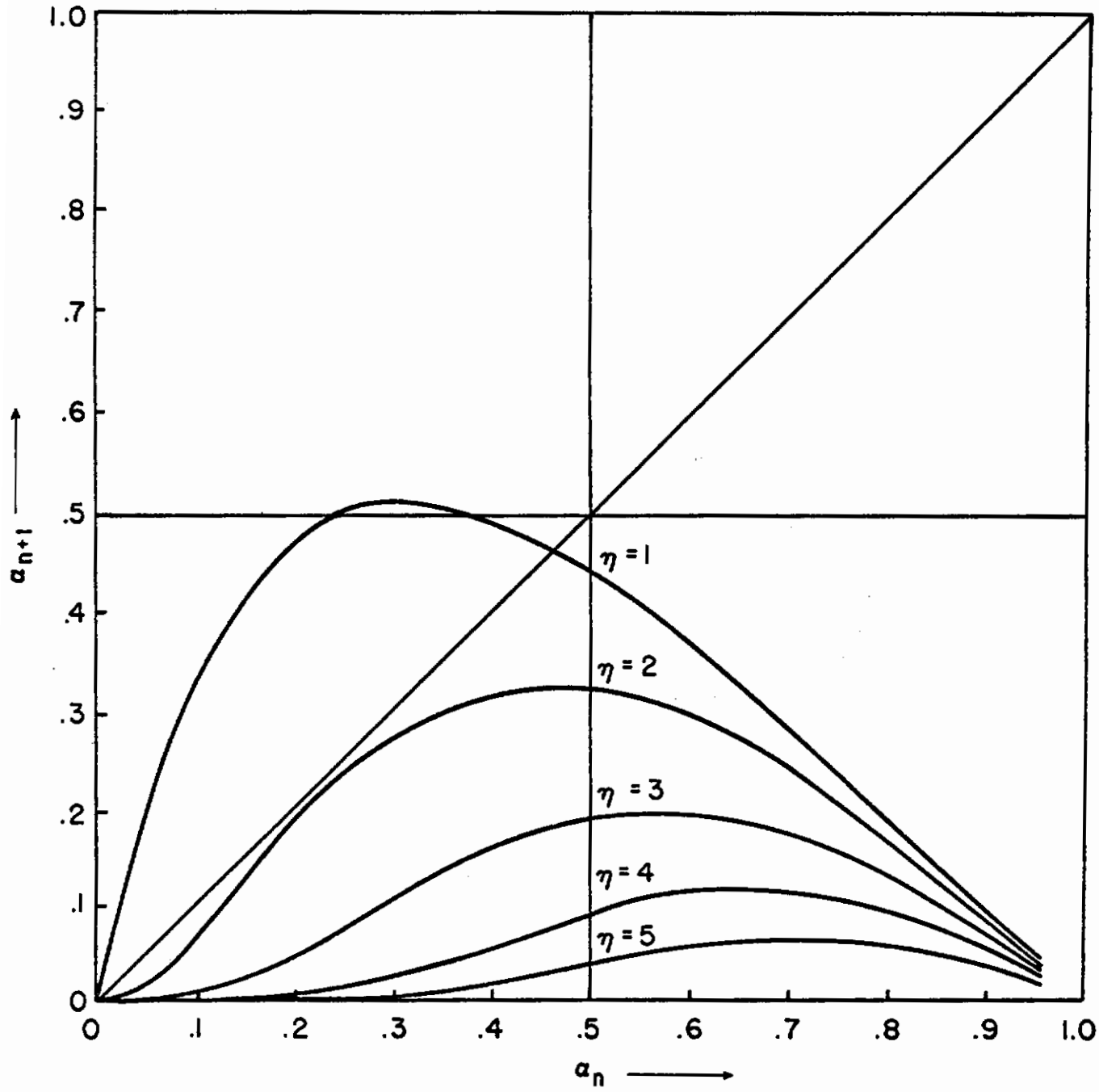

Figure 5. Curves of $\langle \alpha_{n+1} \rangle$ vs. α_n for nets with $\mu_+=5$, $\mu_-=6$, $h=0.05$ and $k_+=-k_-$

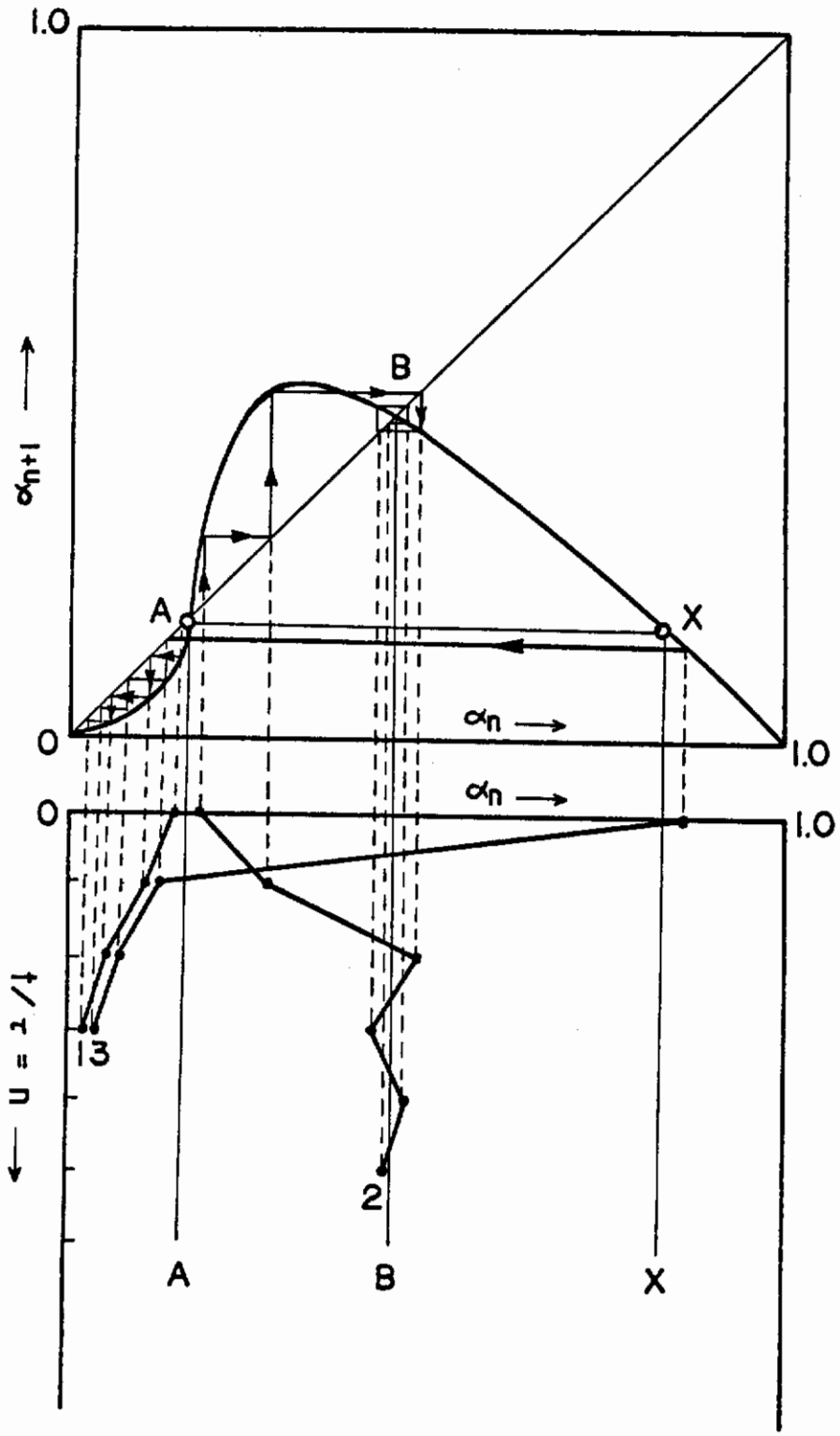

Figure 6. Graphic determination of time dependence of activity in a class *B* net. Upper half: $\langle\alpha_{n+1}\rangle$ vs. α_n; Lower half: α_n vs. n for three initial activities, (1) $\alpha_0 < A$; (2) $A < \alpha_0 < X$; (3) $\alpha_0 > X$

activities, one in the damped portion below point *A*, one in the *ignition* region, and one above a point marked *X*, which would have *A* as its successor. Such a high initial activity would produce a subsequent state below the threshold *A*, hence would not produce sustained activity.

In the sustained activity shown as the second case, we distinguish two phases: an initial monotonic rise, followed by a damped oscillation about the equilibrium point *B*.

Computer generated curves of α_n vs. n are shown in figures 7, 8, and 9 for $h=0$, $\mu_+=10$. Three values of η were chosen which make the nets class *A*, *B*, and *C* respectively. Each figure shows the time course of activity for different initial activities. These curves represent not simulation experiments, but computations based on equation 2. Some difference between the two is expected for reasons described below. Simulation experiments on the IBM System/ 360-50 digital computer are now being carried out in this laboratory. They will enable us also to extend the study to nets in which summation times and refractory times can be varied at will, and neuronal threshold following an absolute refractory period can be made to follow any arbitrary time dependence. In addition the net can be given structure, as explained in Section V of this report. The calculations carried out above, despite their very restrictive assumptions, are providing interesting limiting cases and guidelines for the more general and, we believe, biologically more realistic, systems which are now being studied by computer simulation

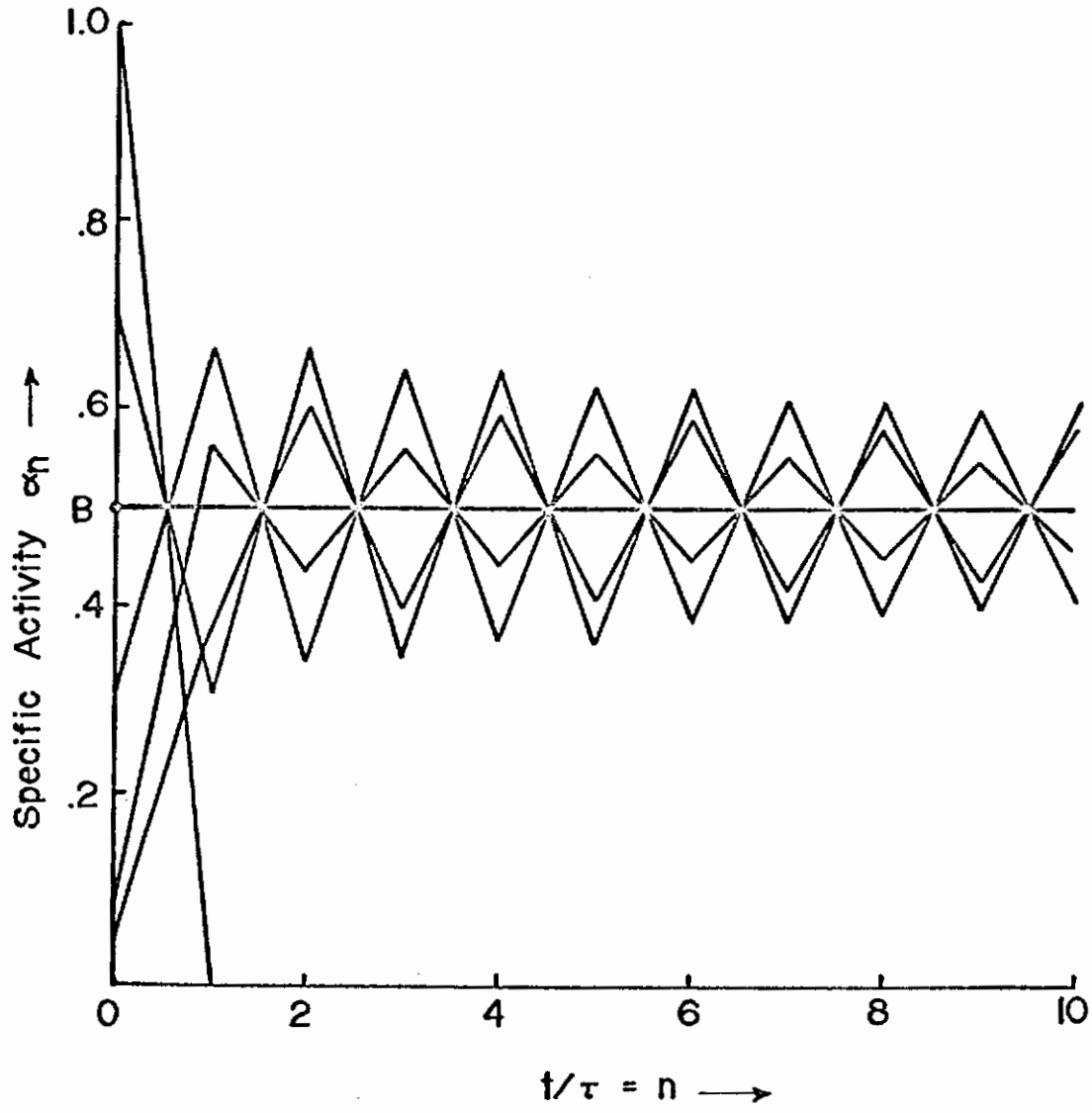

Figure 7. Time dependencies of a_n for various initial values, computed from neuron gas model for h=0, μ_+=10, and η=1 (class A net).

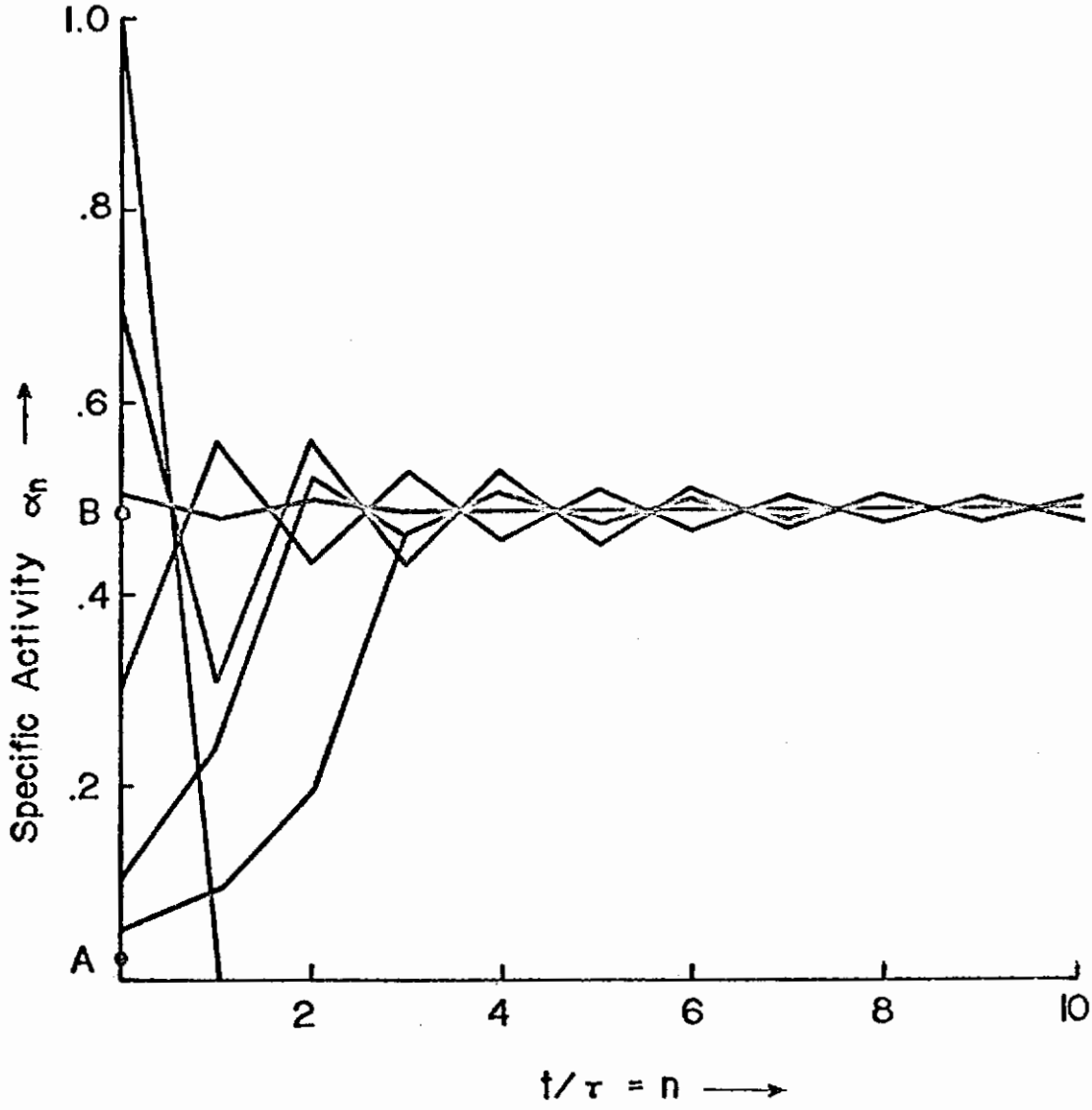

Figure 8. Time dependencies of a_n for various initial values, computed from neuron gas model for h=0, μ_+=10, and η=2 (class *B* net).

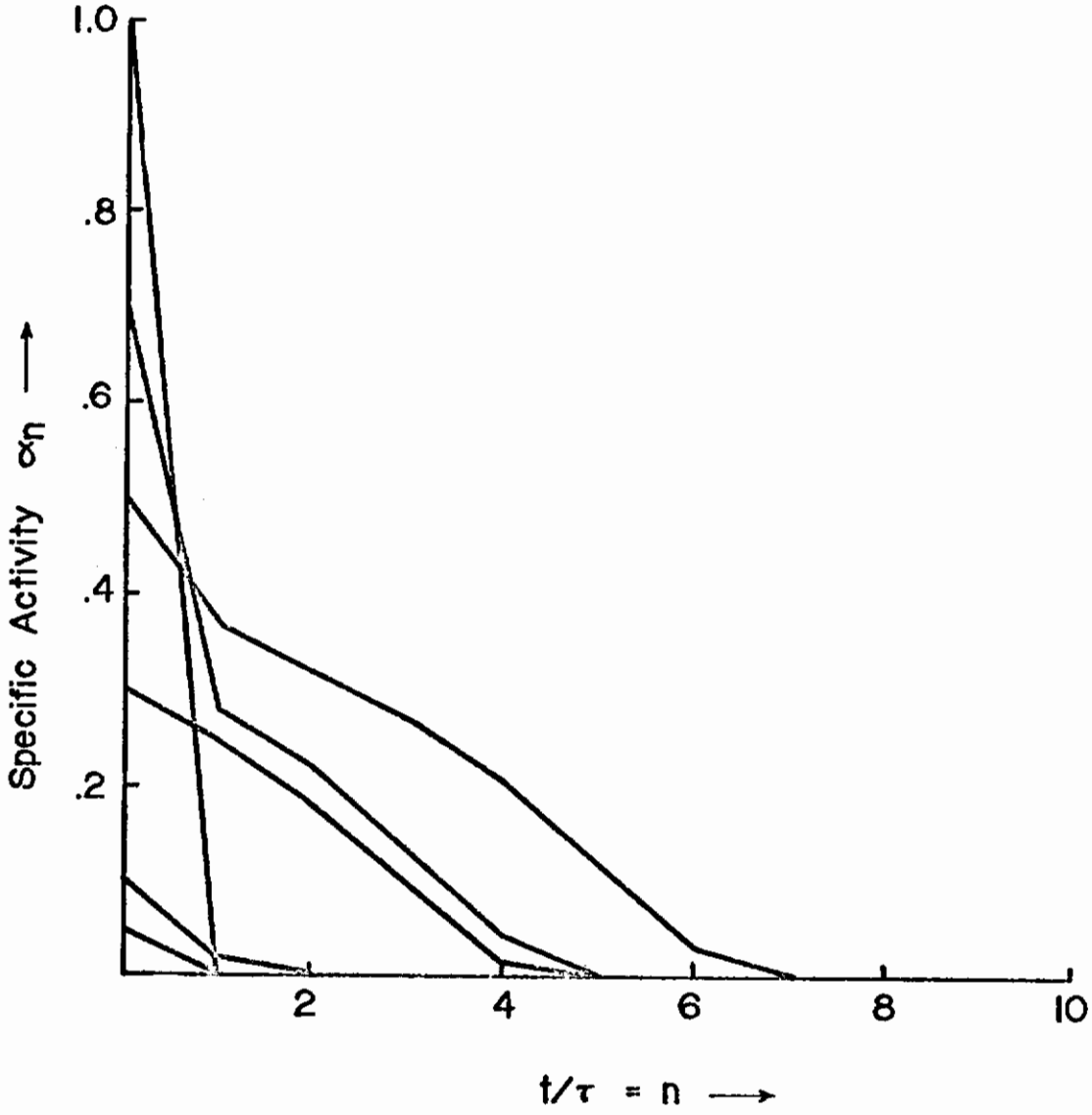

Figure 9. Time dependencies of a_n for various initial values, computed from neuron gas model for h=0, μ_+=10, and η=4 (class C net).

methods.

It should be made clear that Equation 2 gives only the expectation values of the levels of neuronal activity, but does not tell *which* neurons are being excited. Also, implicit in the calculations is the assumption of complete *incoherence* in the temporal development of the activity, by which we mean that the connections are not only random, but are re-randomized after every interval τ. They apply thus, strictly speaking, not to a fixed net, but to something that could be described as a *neuron gas*. Earlier studies by Rochester, *et. al.* (1956) seem to indicate that such coherence effects are not significant beyond a few cycles of activity.

Ergodicity

We wish to raise the following two questions concerned with what may be called the ergodic behavior of nets:

1. Given a *class B* net of the type described above, and assuming sustained equilibrium activity α^*, what fraction of all the *A* neurons in the net will fire at least once if the activity is indefinitely sustained?

2. What fraction of the total number of possible neuronal states of size α^* are reached by the system which sustains equilibrium activity of size α^*?

We shall try to answer these questions first by using the neuron gas model and then consider the coherence effects. The differences, as we shall see, are very significant.

Neuron Gas Model. We consider first the question of the number of neurons involved in sustained activity. Let the successive neuronal states be $\underline{a}_0$, $\underline{a}_1$, $\underline{a}_2$, etc. and assume that each state has precisely the size a^*, where $a^*=\alpha^* A$. The assumption of refractoriness requires that $\underline{a}_0 \cap \underline{a}_1 = \underline{0}$, and in general $\underline{a}_n \cap \underline{a}_{n+1} = \underline{0}$. If we denote by b_i the number of those neurons that were active in the i-th state but have not been active previously, then clearly $b_0=a^*$, $b_1=a^*$, and $b_2=a^*(A-2a^*)/(A-a^*)$. It follows that

$$b_\ell = a^*(A-2a^*-\sum_{i=2}^{\ell-1} b_i)/(A-a^*)$$

Introducing $\beta_i=b_i/A$, we readily obtain

$$\begin{aligned} \beta_0&=\beta_1=\alpha^* \\ \beta_2&=\alpha^*(1-2\alpha^*)/(1-\alpha^*) \end{aligned} \tag{4}$$

In general

$$\beta_\ell=\alpha^*(1-2\alpha^*-\sum_{i=2}^{\ell-1}\beta_i)/(1-\alpha^*) \tag{5}$$

and

$$\beta_{\ell+1}=\beta_\ell(1-2\alpha^*)/(1-\alpha^*) \quad \text{for } \ell \geqslant 1 \tag{6}$$

Finally from (4), (5), and (6)

$$\beta_\ell=\alpha^*\{(1-2\alpha^*)/(1-\alpha^*)\}^{\ell-1} \tag{7}$$

The total number of neurons firing at least once during

an equilibrium activity, sustained for N intervals τ, expressed as a fraction α_N of the total number of neurons in the net, is now given by the expression

$$\alpha_N = 2\alpha^* + \sum_{i=2}^{N} \beta_i \qquad (8)$$

Substitution of (7) into (8) shows that α_N will approach 1 as N goes to infinity. This result indicates that in time every neuron in the net will partake in the activity.

The most obvious coherence effect is *cycling*, the continuous repetition of a given sequence of states. With our assumptions of summation time and refractory period, the precise repetition of a single neuronal state $\underline{a}$, will cause the system to repeat indefinitely the sequence of states between the first appearance of $\underline{a}$ and its first repetition. (With longer summation times or refractory periods, a *sequence* of states will have to be precisely repeated before cycling is assured. Such nets would therefore be expected to approximate the gas model behavior for much longer periods of time.)

We shall now make an approximate calculation of the average length of time it will take before a state is repeated. There are $\binom{A}{\alpha^* A}$ different states of size α^*, hence the probability of repeating in one trial one of r previous states is $r/\binom{A}{\alpha^* A}$, and the probability P_r that in a sequence of r successive states there have been no repetitions, is given by

$$P_r = \prod_{\ell=1}^{r} \left[1 - \ell/\binom{A}{\alpha^* A}\right] \approx 1 - \frac{r(r+1)}{2} \binom{A}{\alpha^* A}$$

The approximation holds for $\binom{A}{\alpha^* A} >> 1$. Now let N be the number of successive states such that there is a 50% chance for a repetition. We have therefore $P_N = 0.5$ and

$$N(N+1) \simeq N^2 \simeq \binom{A}{\alpha^* A} \tag{9}$$

With Stirling's approximation this becomes

$$N^2 \simeq (1-\alpha^*)^{-A(1-\alpha^*)} (\alpha^*)^{-\alpha^* A} / \sqrt{2\pi A} \tag{10}$$

We are now in a position to give an approximate answer to the first of the questions raised above. The fraction of the net that will ultimately be involved, if the cycling begins after N synaptic delays, is obtained by substituting equation (7) into equation (8). We obtain

$$\alpha_N \simeq 1-(1-\alpha^*)\{(1-2\alpha^*)/(1-\alpha^*)\}^{N-1} \tag{11}$$

The parameter N in this model is given by (10); it will in general be a large number, so that, according to (11), α_N will be very close to unity.

The second question, regarding the fraction of possible states reached can also be answered now. Let this fraction be f. Evidently

$$f = N/\binom{A}{\alpha^* A} \simeq 1/N \tag{12}$$

This would predict that even though many states will in general be passed before cycling sets in, the fraction of states accessible from a given initial state is small.

Fixed Net. In a neural net in which the connectivity remains fixed, the refractoriness of neurons will cause strong correlations between successive states; thus the number of available states is reduced and the probability for repeating a previous state should be enhanced. With a refractory period equal to one synaptic delay, any two successive states $\underline{a}_n$ and $\underline{a}_{n+1}$ will be disjoint. If these states have sizes α^* close to 0.5, there would thus be an increased probability for a state $\underline{a}_{n+2}$ to be identical to $\underline{a}_n$. This probability might be expected to be about $(\frac{(1-\alpha^*)A}{\alpha^* A})^{-1}$ which goes to unity as α^* approaches 0.5. The average number of successive states preceding repetition due to this short-range order would be of the order of the reciprocal of this expression:

$$N \simeq (\frac{(1-\alpha^*)A}{\alpha^* A}) \tag{13}$$

Comparing equations (13) and (9) we see that the short range order should contribute a negligible effect to the probability for repeating a state as long as α^* is small, but will become the dominant factor as α^* approaches 0.5.

However the correlation between states $\underline{a}_n$ and $\underline{a}_{n+2}$ is even stronger than indicated by equation (13). This may be shown in the following manner. In figure 10 we show in a Venn diagram four successive neuronal states $\underline{a}_{n-1}$, $\underline{a}_n$, $\underline{a}_{n+1}$, and $\underline{a}_{n+2}$. We calculate the fractional overlap between $\underline{a}_n$ and $\underline{a}_{n+2}$, called $f_{n,n+2}$. The calculation will be based on two assumptions: a) if a state $\underline{a}_i$ generates the state $\underline{a}_{i+1}$, then a state $\underline{a}_i'$ which is a proper subset of $\underline{a}_i$ will generate a state $\underline{a}_{i+1}'$

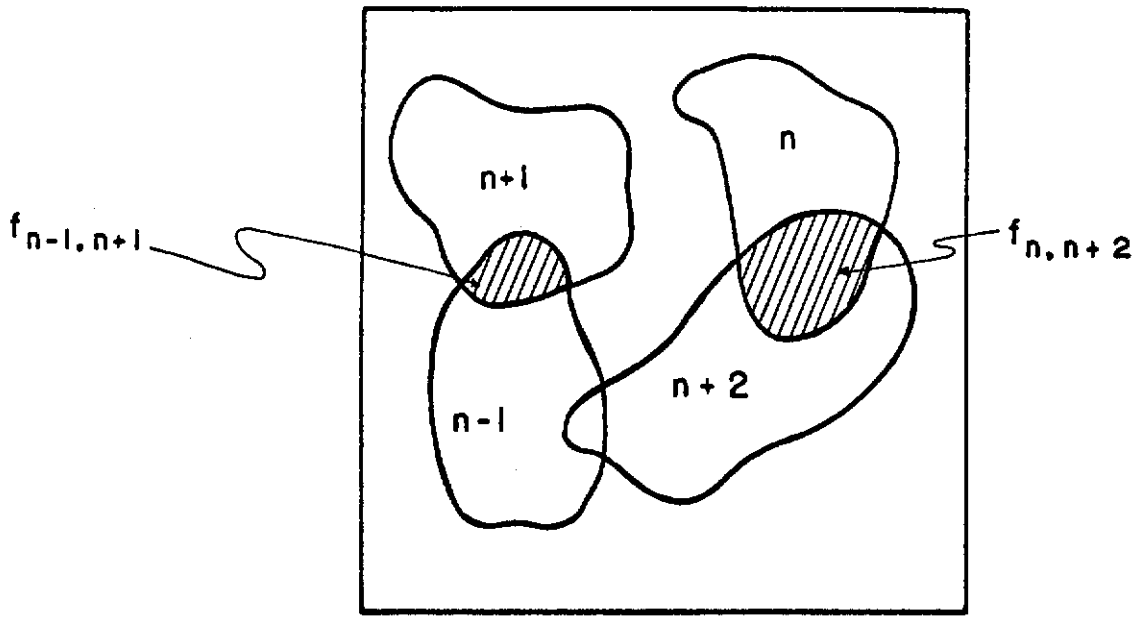

Figure 10. Venn diagram of four successive neuronal states

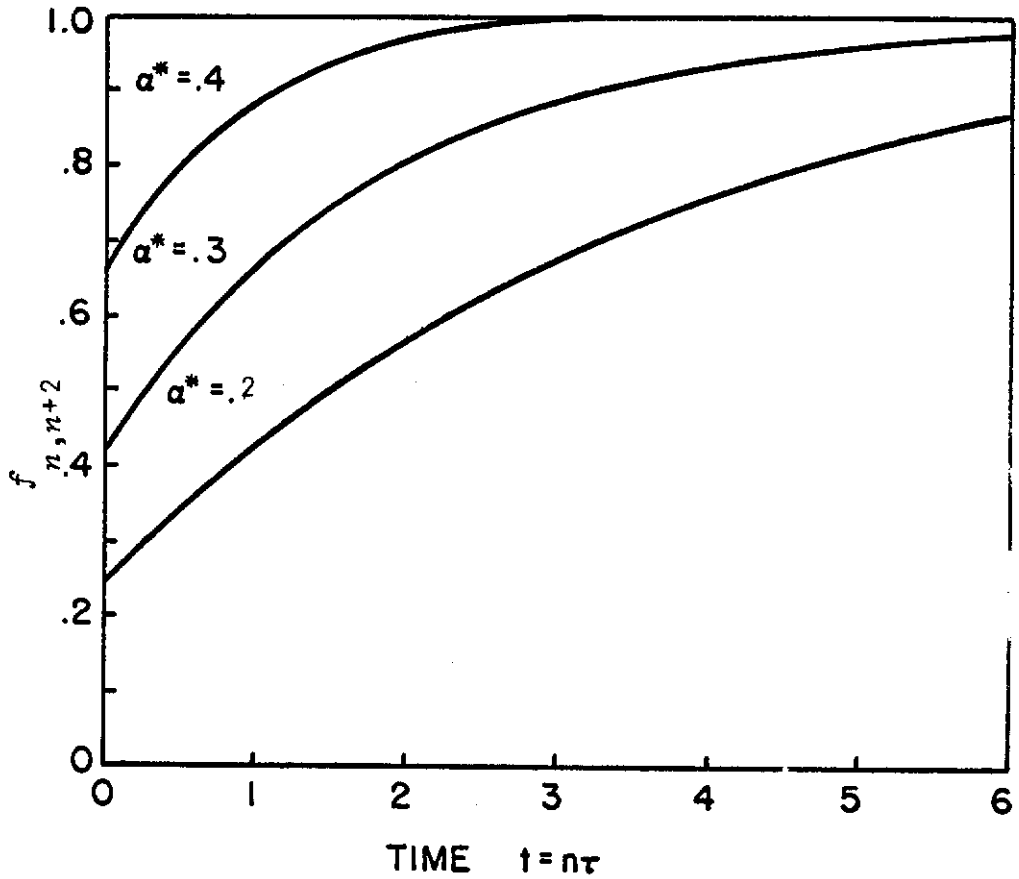

Figure 11. Fractional overlap between alternating states vs. time for sustained activity α^*

which is a proper subset of $\underline{a}_{i+1}$; and b) the size of the proper subset $\underline{a}'_{i+1}$ will be proportional to the size of the proper subset $\underline{a}'_i$.

The first of these assumptions can be shown to be strictly true if there are no inhibitories, and will be approximately true as long as the fraction h of inhibitories is small. Assumption b) is an approximation which is good for small activities.

If we now let $f_{n,n+2}$ be the fractional overlap between states $\underline{a}_n$ and $\underline{a}_{n+2}$ (Figure 10), then the above assumptions easily lead to the following

$$f_{n,n+2}=f_{n-1,n+1}+(1-f_{n-1,n+1})(\alpha^*/1-\alpha^*) \tag{14}$$

from which

$$f_{n,n+2}=k\sum_{i=0}^{n}(1-k)^i \quad \text{where } k=\alpha^*/1-\alpha^* \tag{15}$$

but

$$\lim_{n\to\infty}\sum_{i=0}^{n}(1-k)^i=1/k \tag{16}$$

It follows that for large n, Equation 15 gives

$$f_{n,n+2} \simeq 1$$

This implies that alternating states would become identical; the activity of the system would switch back and forth between two fixed states. Figure 11 shows that equation 15 would predict the system to approach quite rapidly this non ergodic state. Computer simulations have indicated that this indeed

is the valid approximation. In a net of 1000 neurons, two-state cycling was observed to occur every time, shortly after the net reached sustained activity near the point of stable equilibrium.

SECTION IV

INFORMATION PROCESSING IN THE ASSOCIATION NET

It was shown in (I) that a completely unstructured, randomly connected network of neurons was capable of performing a variety of cognitive tasks under the following assumptions:

the network in its naive state is highly damped, i.e. its operation would correspond to what we termed class *C*, operated in the portion of the graph where $\langle \alpha_{n+1} \rangle \approx 0$;

synaptic facilitation occurs whenever presynaptic action potentials coincide with postsynaptic firing;

information is carried by the network in the form of *neuronal states* that specify the *set* of neurons active in the net at one instant. The significance of the neuronal states is described by the correspondences M ("means"), I ("implies"), and S ("suggests") between these states on the one hand and a *sensory* state s on the other. We say that a state a *means* s, written as aMs, if the sensory input s incident on the quiescent associator net produces in the latter the state a. A state b *implies* s, written bIs, if $b \supset a$ and aMs. A state c *suggests* s, written cSs, if $c \cap a \neq 0$, aMs , and $(c \cap a) \subset c$. These correspondences are shown schematically in figure 12.

The necessity of the strong damping became quite apparent in early computer simulation runs. Whenever significant spreading of activity took place in the naive net following stimulation, the net performed poorly upon learning. This can be understood in the light of the results of

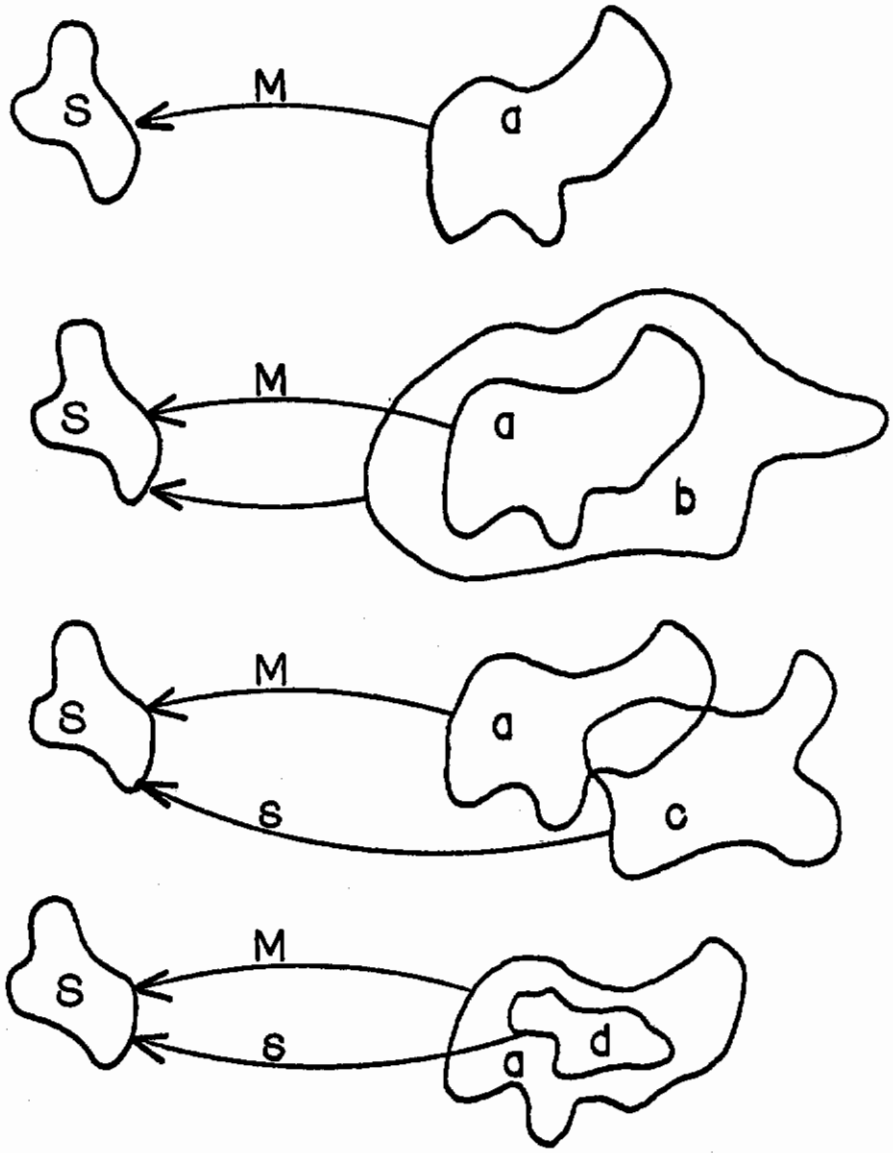

Figure 12. Venn diagram of the three correspondences M, I, and S between states in sensory net and states in association net

Rochester, *et al* (1956), Allanson (1956), Beurle (1956), Farley and Clark (1961), Smith and Davidson (1962a; 1962b), and Griffith (1963; 1965), all of which demonstrate that random nets rapidly lose information when they are reverberating. Only in so-called *local* nets, in which there existed a strong preference for connections to nearest neighbors, did sustained activity remain confined spatially, and hence preserve information concerning its origin (Rochester, *et. al.*, 1956). Such a net however would be expected to perform poorly as an associator net, in which strong interconnections must exist between different modalities. We have here an example of the complementarity discussed in Section II: the design of strong local bias in the connectivity (which is information-preserving) is antagonistic to the capability to carry out unpredictable tasks. This complementarity prevents the spontaneous formation of information-preserving *cell assemblies* predicted by Hebb (1949) in a reverberating random associator net.

The above reasoning suggests that, if synaptic facilitation is the basis of memory formation and learning, and if synaptic facilitation is brought about by repeated transmission of activity across the synapse, then the repeated stimulation must be generated outside the association area. The sources of such sustained activity in the association area may be either repetition of the sensory input, or neuronal reverberations occuring in structured, information-preserving portions of the brain. We need only postulate that each of

of the states in this association net is in turn triggered by a specific sensory input--the *meaning* of that state--and our model of learning will have a close biological counterpart in what is known as sensory-sensory cortical conditioning (Morrell, 1957; Yoshii, *et. al.*, 1957, 1960). Here paired stimuli of different modalities become functionally connected in such a way that the application of only one of the stimuli will produce neuronal activity characteristic of both.

The effect of applying the reinforcement rules to particular states in the association net was discussed in (I). The fundamental operation involved in learning was defined to be the association of neural states. In brief, this consisted of applying the reinforcement rules over a domain of neurons activated by the simultaneous presentation of two or more stimuli. It results in the enhancement of coupling coefficients k_{ij} over an area in the connectivity matrix. Schematic diagrams, the so-called association diagrams, were introduced in (I) as an aid in visualizing the effect.

In one of the examples the system was used to simulate what Asratian (1965) has called the *switching reflex*. An animal learns to associate a certain conditioned stimulus C with an unconditioned stimulus U_1 until C alone will elicit the reflex R_1. The training is done under certain environmental conditions we call A_1.

Next, the animal is placed in a different environment, where the same conditioned stimulus C is associated with a different unconditioned stimulus, U_2, which produces the

reaction R_2. After appropriate training the conditioned stimulus C will elicit the conditioned reflex R_1 under conditions A_1 and R_2 under conditions A_2. This is called ***switching.***

In our interpretation of the conditioned reflex C, U_1, and A_1 produce neuronal activity in distinct neuron pools in the association net and we let $\underline{C}$, $\underline{U}_1$, and $\underline{A}_1$ be the cortical representations of these stimuli. Also it was assumed that strong inborn neural pathways lead from $\underline{U}_1$ to a pool of motor neurons $\underline{R}_1$ such that whenever a significant fraction of neurons in $\underline{U}_1$ is active, the reflex R_1 will ensue. By the ternary association of $\underline{C}$, $\underline{U}_1$, and $\underline{A}_1$ we mean the simultaneous activation of these three neuronal states and the consequent strengthening of synaptic connections between any two neurons which belong to either of these states. The switching reflex may now be pictured in our interpretation as resulting from two ternary associations $(\underline{C},\underline{A}_1,\underline{U}_1)$ and $(\underline{C},\underline{A}_2,\underline{U}_2)$ which is shown schematically in figure 13 where the heavy lines represent inborn pathways, all other lines facilitated pathways between neuron pools. The reader is referred to (I) for detailed discussions of this and other cognitive tasks.

It is instructive to look at the network parameters used above, and compare them with the curve presented in Section III. The naive net had the parameters $\mu_+=4$, $\mu_-=6$, $k_+=0.19$, $k_-=-0.19$, $\Theta=1$, and $h=0.05$. The increment used in learning was $\delta=0.38$. Comparison with the curves for the (less strongly damped) net $\mu_+=5$, $\eta=5$, $h=0.05$ (figure 5) shows that the system had been

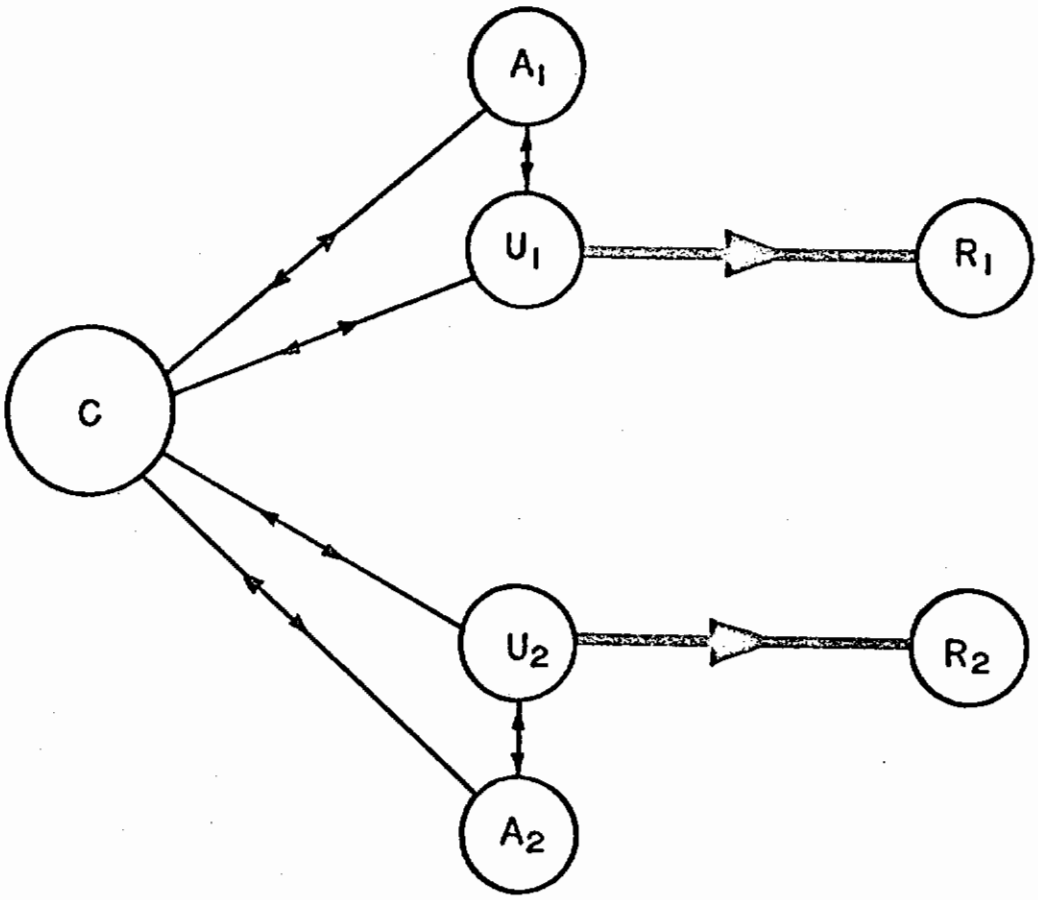

Figure 13. ASSOCIATIONS MADE IN SWITCHING REFLEX. Heavy lines are inborn connections leading to motor neurons controlling the reflexes R_1 and R_2. Thin lines represent reinforced pathways between the states $\underline{C}$, $\underline{U}_1$ and $\underline{A}_1$ and between $\underline{C}$, $\underline{U}_2$ and $\underline{A}_2$, respectively.

what we called above a class *C* net and that with activities up to about 0.2 it was in the highly damped mode of operation. On the other hand, learning will establish a net-within-a-net, this one with $\eta=2$ which should be on the verge of being a reverberating class *B* net. The data suggest that it either just missed being class *B*, or was operated just below the ignition point. It seems clear now, that slight modifications in the design parameters could have produced much more impressive effects than those reported in (I).

In summary, learning in the association net was considered to involve the establishment through synaptic facilitation of reverberating or almost reverberating subnets, embedded in an otherwise unstructured and highly damped matrix of neurons.

SECTION V

COMPOUND NET MODEL

Our next objective is to incorporate into a single dynamic model both the design known to exist in the sensory system up to and including the sensory areas of the cortex, and the presumed lack of organization at the higher cortical levels. We shall not be concerned here with the role of other, subcortical, structures.

The sensory system that has been explored in greatest detail is probably the visual system. Lateral inhibition was first studied in the eye of the horseshoe crab (Hartline, 1949), and receptive fields of retinal ganglion cells were plotted for the frog (Lettwin, *et. al.*, 1959) and for the cat (Hubel and Wiesel, 1962). The results all demonstrate neural organizations designed to cope with tasks that are reasonably well understood and of genetically predictable value to the organism. In a series of studies, Hubel and Wiesel extended their work to neurons in the lateral geniculate body (1963) and in areas 17, 18, and 19 of the visual cortex (1965). These studies again revealed a remarkable degree of structure and design, characterized by spatially circumscribed pools of neurons with identical or nearly identical receptive fields. Thus there exist in area 17 columnar structures, comprising probably many thousands of neurons, "wired" to act as detectors of a particular cognitive element such as a directed line or edge of light located in a particular portion of the visual field. Besides such simple fields, Hubel and Wiesel also observed

more complex structures such as detectors for directed lines but differing from the above-mentioned fields in that the precise location of the line was immaterial. Other pools of neurons were triggered when a directed line moved across the field. Such and other complex fields were rare in area 17 but increased in frequency in areas 18 and 19. The fact that all of these functions and the requisite neural designs were inborn was definitely established (Hubel and Wiesel, 1966).

Compound Nets

In (I) the structure of a neural net was represented by the connectivity matrix $\underline{k}_{ij}$ of coupling coefficients. Figure 14 represents a schematic diagram of the connectivity matrix of a net which incorporates some of the features discussed above. Here $\underline{s}_1$, $\underline{s}_2$, $\underline{s}_3$, etc. are pools of neurons in the sensory cortex. Each pool consists of neurons of identical receptive fields; we may therefore consider that the set $\underline{s}_1$, for example, is simultaneously activated when its receptive field is stimulated.

We next define the association net designated by A. Unlike the sensory cortex, this net is to be initially unstructured, and highly damped. Non vanishing coupling coefficients k_{ij} are seeded uniformly and randomly over the entire area A. In our model, only the coupling coefficients in this area are assumed to be subject to the rules of synapse facilitation. All other areas in the matrix $\underline{k}_{ij}$ have constant elements. Thus memory is contained not in the sensory or

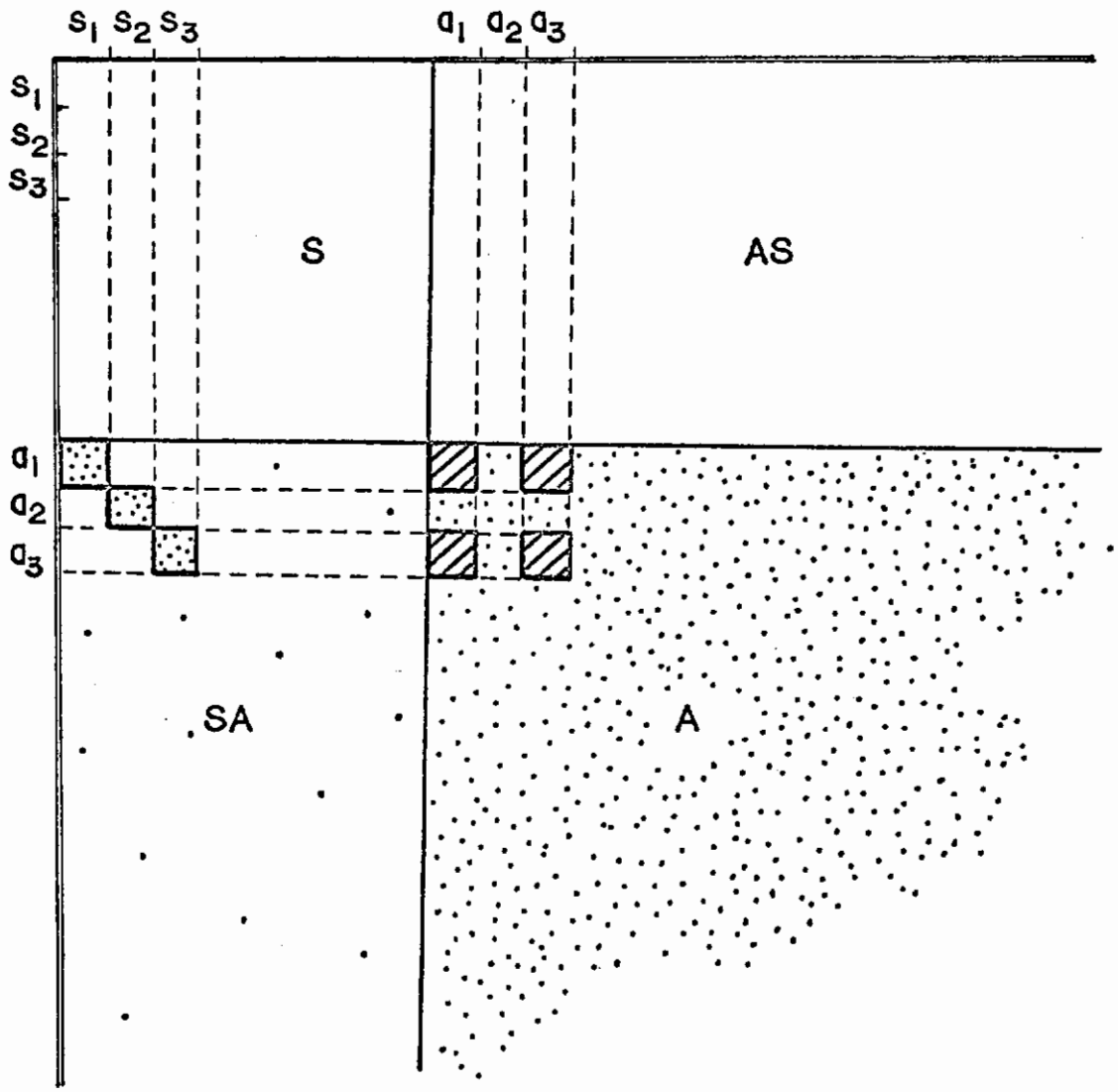

Figure 14. CONNECTIVITY MATRIX k_{ij} OF COMPOUND NET. Here $\underline{s}_1$, $\underline{s}_2$, $\underline{s}_3$ are states in the sensory net, $\underline{a}_1$, $\underline{a}_2$, $\underline{a}_3$ in the association net. Nonvanishing matrix elements are shown by dots. Reinforced areas are cross-hatched. The diagram shows the result of associating $\underline{a}_1$ and $\underline{a}_3$.

motor areas of the cortex but in a nonspecific region, the association area. Some evidence for this has been deduced by Asratian (1961) from the so-called "switching" experiments.

The area *SA* in the connectivity matrix (figure 14) represents synaptic connections originating in the sensory cortex and terminating in the association cortex. We make the assumption that the neurons in *S* have their axonal endings randomly distributed among the neurons in *A*. Thus *SA* is also intrinsically an area that is uniformly seeded with nonvanishing coupling coefficients.

We shall now rearrange the order of the neurons in *A* as follows: let $\underline{a}_1$ be the set of neurons in *A* which are triggered whenever all of the neurons in $\underline{s}_1$ are firing. Similarly $\underline{a}_2$ is the set of neurons in *A* triggered by $\underline{s}_2$, etc. The order of neurons in the association cortex may now be changed in the diagram without any loss of generality. Specifically, if $\underline{a}_1$, $\underline{a}_2$, etc are disjoint states[3],[4] the reordering may be carried out in such a way as to lead to an arrangement like the one shown in figure 14. Area *SA* now

---

[3]This condition can always be satisfied in practice by assuming the number of neurons in the association net to be very large compared with the size of the states $\underline{a}_1$, $\underline{a}_2$, etc.

[4]The states $\underline{a}_1$, $\underline{a}_2$, etc are what we called in (I) *fundamental cortical states*. We drop this nomenclature here as superfluous. Nevertheless if $\underline{s}_1$ sends inputs into an otherwise completely quiescent association net and triggers $\underline{a}_1$, the correspondence $\underline{a}_1 \underline{M} \underline{s}_1$ is invariant under our assumptions even with learning taking place.

appears as though it were structured, having blocks of dense coupling superimposed on a background of more diffuse coupling. It is important to realize that the re-labeling of neurons in the association net will not change the uniform and random distribution of coupling coefficients in area *A*.

We try first a simple scheme, which may perhaps not be very realistic. Assume that a coupling coefficient k_{ij} is increased slightly whenever the *i*-th and the *j*-th neuron fire simultaneously[5]. It will be convenient to adopt also the so-called *adiabatic learning hypothesis* (Caianiello, 1961) according to which these changes are so slow that they may be neglected if we compute the activity in the net resulting from a single sensory input.

The following mode of operation of our compound net model may now be contemplated: consider two stimuli for which the neuron pools $\underline{s}_1$ and $\underline{s}_3$ are the specific detectors. If these two stimuli occur simultaneously for some time, reinforcement will occur in the cross-hatched areas of the association diagram (figure 14). We had previously shown that the set of neurons $\underline{a}_1 \bigcup \underline{a}_3$, after sufficient reinforcement, may be considered as a net-within-a-net; it will be relatively independent of the embedding net.*A* if the latter is highly damped. The association of $\underline{s}_1$ and $\underline{s}_3$, thus accomplished will have the effect that hereafter either one

---

[5] This assumption differs from the Hebbian hypothesis in that here the presynaptic action potential, because of the synaptic delay, does *not* contribute to the firing of the postsynaptic neuron.

of the stimuli will produce a neural activity in the association net which *suggests* (according to our definition of this term) the simultaneous presence of the other stimulus.

So far we have followed the static treatment, first presented in (I). The system described above has thus all the capabilities of carrying out the various cognitive tasks described in (I).

We shall now go a step further. In figure 15 we show a neural net which differs from the previous one in that the neuron pools $\underline{s}_1$, $\underline{s}_2$, $\underline{s}_3$ each form a small interconnected net as shown by the blocks in S (figure 15). The connections from S to A are the same as before. We wish to reinvestigate now the mechanisms that may lead to the reinforcement of coupling coefficients within a given domain, say $\underline{a}_1$, in the association net. For this purpose we shall employ the Hebbian reinforcement rule in its original form, which states that a coupling coefficient k_{ij} between two neurons in the association net increases in value whenever the firing of the j-th neuron is, by itself or in conjunction with others, responsible for triggering the i-th neuron. Detection of a given cognitive element may now be pictured as activity in a specific neuron pool $\underline{s}_1$ in the sensory cortex. This activity may be damped or sustained, depending on the parameter of the net, the size of the initial excitation, and the question whether or not the sensory input is sustained. A sustained activity in $\underline{s}_1$ clearly will trigger some sustained activity in $\underline{a}_1$ by virtue of the connections in SA (figure 15). Since $\underline{a}_1$

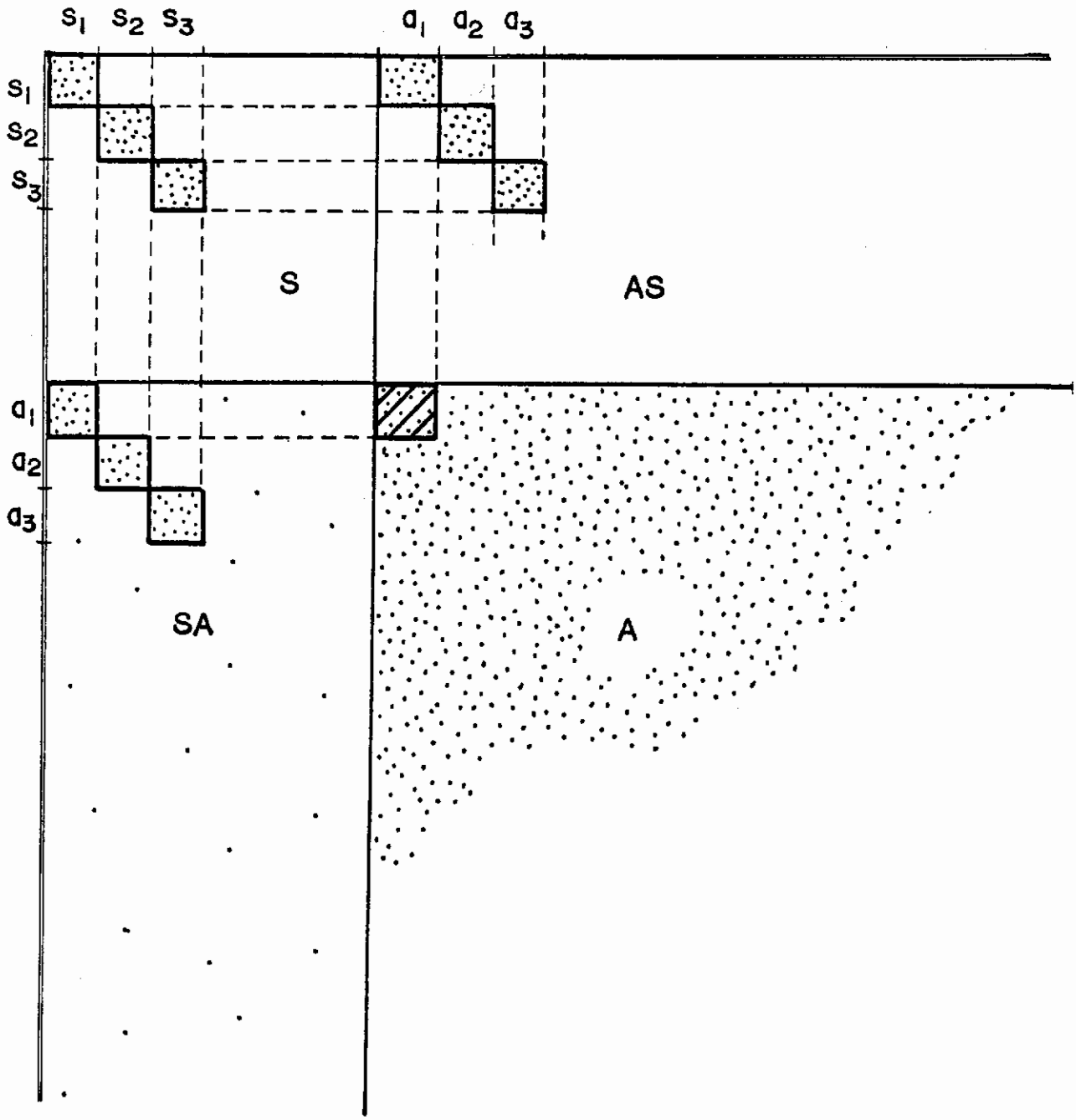

Figure 15: CONNECTIVITY MATRIX OF COMPOUND NET. Here $\underline{s}_1$, $\underline{s}_2$, $\underline{s}_3$ are small reverberating nets supplying activity to and receiving activity from sets of neurons $\underline{a}_1$, $\underline{a}_2$, $\underline{a}_3$ respectively in the association net.

is assumed to be initially highly damped (as is the net A in which it is embedded); the activity in $\underline{a}_1$ will be just that due to these inputs from $\underline{s}_1$. Let the fraction of neurons in $\underline{a}_1$, that are activated in every successive interval, be α^*. For the neuron gas model the $\alpha^* a$ neurons in $\underline{a}_1$ are chosen completely at random for every interval τ. The average firing frequency of a neuron in $\underline{a}_1$ will thus be $\langle v \rangle = \alpha^*/\tau$. The probability for a particular k_{ij} $(i,j \varepsilon A)$ to become strengthened in the $(n\tau)$-th interval will be the probability that the i-th neuron fires in the $(n\tau)$-th interval, multiplied by the probability that the j-th neuron has fired in the $(n-1)\tau$-th interval. But that is just $(\alpha^*)^2$. This means, for instance, that for $\alpha^* = 1/10$ and $\tau = 10^{-3}$ sec., each coupling coefficient in $\underline{a}_1$ will receive an average of 10 increments per second of sustained activity. Clearly, in time the entire domain $\underline{a}_1$ will have become reinforced.

The above argument is equally well applied to the association of two or more stimuli. Thus again, all of the associative functions which have been shown in (I) to be the properties of highly damped nets apply to the present model. We have merely added a mechanism, lacking there, for generating the sustained stimulation necessary for producing the slow changes in the coupling coefficients. This mechanism is believed to be the existence of reverberating pools of neurons in the sensory cortex which act of detectors of specific cognitive elements. Also in the present treatment, a more rigorous application of the Hebbian reinforcement rule

(Hebb, 1949) has been possible.

An interesting possibility is the existence of fibers returning activity from the regions $\underline{a}_1$, $\underline{a}_2$, $\underline{a}_3$ to the sensory pools $\underline{s}_1$, $\underline{s}_2$, and $\underline{s}_3$ respectively, shown in area *AS* in figure 15. In this case reverberations may occur between the two nets, particularly after reinforcements have taken place in the association net. Evidence for this may be deduced from the sensory-sensory cortical conditioning experiments of Morrell (1957) mentioned above. Here, following the frequent pairing of a visual and an auditory stimulus, the auditory stimulus by itself would elicit in the visual cortex an activity characteristic of the visual stimulus. We believe that the study of the dynamics of coupled probabilistic nets detailed in this section will contribute to a physical understanding of such neurological processes in cortical conditioning. Such processes are currently being studied by us using computer simulation.

SECTION VI

SUMMARY AND CONCLUSIONS

The existence of various neuronal structures in the cerebral cortex is a well-known fact. These circuits and their afferents are evidently genetically designed and fulfill readily discernible functions. It was argued in Section II that probabilistic elements should play an increasing role as higher and higher cortical areas are being considered. We have investigated a model of cortical functioning in which the lower cortical structures interact with a large probabilistic net, the association net. The latter is completely unstructured initially. Synaptic learning of the type postulated by Hebb, is assumed to be operative within the association net only.

In a preliminary study the dynamic properties of simple probabilistic nets was investigated (Section III). Curves are presented that show the dependence of net behavior on various parameters.

In Section V, we discuss the coupling of the sensory structures to the association net. In particular, *cell assemblies* emerge in a natural way in the association net if we assume the sensory structures to consist of pools of interconnected neurons that are miniature reverberating nets.

The information processing that takes place in the association net is that discussed in some detail in a previous

paper. A brief resumé of this work is presented in Section IV.

The calculations on net dynamics carried out in Section III apply to a specific set of assumptions concerning the individual neurons. Among these are that the refractory period of a neuron is greater than the synaptic delay but less than twice the synaptic delay. At the end of the refractory period the neuron reaches full sensitivity. We are now in the process of extending these studies by means of computer simulation techniques. This will allow us to consider a great variety of assumptions on absolute and relative refractoriness and other neuronal parameters. The appropriate computer programs, written for the most part in Fortran IV, will be published in a forthcoming report. The same programs are also capable of simulating the compound neuronal structures we have discussed in Section V.

APPENDIX

Derivation of Equation 2.

The expression for the activity $\langle\alpha_{n+1}\rangle$ generated by the preceding activity α_n in our probabilistic net, Equation 2, is obtained by adding the probabilities of all those combinations of excitatory and inhibitory inputs to a neuron which give a total PSP exceeding the threshold θ.

Let p_ℓ be the probability that a neuron will receive ℓ excitatory inputs, and q_m the probability of receiving m inhibitory inputs.

Using the Poisson approximation, which will be reasonably good especially for small values of α_n, we obtain

$$p_\ell = exp\left[-\alpha_n(1-h)\mu_+\right]\cdot\left[\alpha_n(1-h)\mu_+\right]^\ell/\ell! \tag{17}$$

$$q_m = exp(-\alpha_n h\mu_-)\cdot(\alpha_n h\mu_-)^m/m! \tag{18}$$

Let $P(\theta)$ be the probability of triggering a particular neuron. We have

$$P(\theta) = \sum_{m=0}^{m_{max}} \sum_{\ell=\eta'}^{\ell_{max}} p_\ell q_m \tag{19}$$

Here η' is the minimum number of excitatory inputs necessary to trigger a neuron, which has received m inhibitory inputs. It is given by $\eta' = u\left[(\theta - mk_-)/k_+\right]$. The upper limits ℓ_{max} and m_{max} in the double sum are the total number of active excitatory and inhibitory connections respectively, i.e.

$$\ell_{max}=A\alpha_n(1-h)\mu_+$$
$$m_{max}=A\alpha_n h\mu_- \qquad (20)$$

From (20) we see that ℓ_{max} will always be very much larger than the exponent in (17), hence we can write

$$\sum_{\ell=\eta'}^{\ell_{max}} p_\ell \simeq \sum_{\ell=\eta'}^{\infty} p_\ell = 1-exp\left[-\alpha_n(1-h)\mu_+\right] \cdot$$
$$\cdot \sum_{\ell=0}^{\eta'-1} \left[\alpha_n(1-h)\mu_+\right]^\ell/\ell!$$

Substitution into (19) now yields

$$P(\Theta)\simeq \sum_{m=0}^{m_{max}} exp(-\alpha_n h\mu_-)\cdot(\alpha_n h\mu_-)^m/m! \cdot$$
$$\cdot\{1-exp\left[-\alpha_n(1-h)\mu_+\right]\cdot \sum_{\ell=0}^{\eta'-1} \left[\alpha_n(1-h)\mu_+\right]^\ell/\ell!\}$$

According to our assumptions of refractoriness there are exactly $(1-\alpha_n)A$ neurons which are not in a refractory state at time $(n+1)\tau$. Hence the expectation value of α_{n+1} is given by

$$<\alpha_{n+1}>\simeq(1-\alpha_n)P(\Theta)$$

which is Equation 2.

REFERENCES

Allanson, J.T., 1956, Some properties of a randomly connected neural network, *In* Information Theory (Third London Symposium). C.Cherry, ed., Butterworth & Co. Ltd., London, pp. 303-313

Ashby, W.R., H. von Foerster, and C.C. Walker, 1962, Instability of pulse activity in a net with threshold, *Nature*, 196, pp. 561-562

Asratian, E.A., 1961, *Ann. N.Y. Acad. Sci.*, 92, p. 1141

Asratian, E.A., 1965, Compensatory Adaptations, Reflex Activity and the Brain, English translation by S.A. Corson, Pergamon Press, Ltd., Oxford, England

von Békésy, G., 1967, Sensory Inhibition, Princeton University Press, Princeton, N.J.

Beurle, R.L., 1956, Properties of a mass of cells capable of regenerating pulses, *Trans. Roy. Soc. London*, Ser. B, 240, pp. 55-94

Beurle, R.L., 1962a, Functional organization in random networks, *In* Principles of Self-Organization, H. von Foerster and G.W. Zopf Jr., eds., Pergamon Press, Ltd., New York, pp. 291-314

Beurle, R.L., 1962b, Storage and manipulation of information in random networks, *In* Aspects of the Theory of Artificial Intelligence, C.A. Muses, ed., Plenum Press, New York, pp. 19-42

Bremermann, H., 1967, Quantitative aspects of goal-seeking self-organizing systems, *In* Progress in Theoretical Biology Vol. I., F.M. Snell, ed., Academic Press, New York, pp. 59-77

Bullock, T.H., 1967, Signals and neuronal coding, *In* The Neurosciences, G.C. Quanton, T. Melnechuk and F.O. Schmitt, eds., The Rockefeller University Press, New York

Caianiello, E.R., 1961, *J. Theoret. Biol.*, 1, p. 204

Caianiello, E.R., A. deLuca, and L.M. Ricciardi, 1967, *Kybernetik*, 4, p. 10

Farley, B.G. and W.A. Clark, 1961, Activity in networks of neuron-like elements, *In* Information Theory (Fourth London Symposium), C.Cherry, ed., Butterworth & Co., Ltd., London, pp. 242-251

Granit R. and B. Renkin, 1961, *J. Physiol.*, 158, p. 461

Griffith, J.S., 1963, On the stability of brain-like structures, *Biophys. J.*, 3, pp. 299-308

Griffith, J.S., 1965, A field theory of neural nets: II. Properties of the field equations, *Bull. Math. Biophys.*, 27, pp. 187-195

Harth, E.M. and S.L. Edgar, 1967, Association by synaptic facilitation in highly damped neural nets, *Biophys. J.*, 7, pp. 689-717

Hartline, H.K., 1949, *Fed. Proc.* , 8, p. 69

Hebb, D.O., 1949, The Organization of Behavior, John Wiley & Sons, Inc., New York, N.Y.

Hubel, D.H. and T.N. Wiesel, 1962, Receptive fields, binocular interactions and functional architecture in the cat's visual cortex, *J. Physiol.*, 160, 106-154

Hubel, D.H. and T.N. Wiesel, 1963, Shape and arrangement of columns in cat's striate cortex, *J. Physiol.*, 165, pp. 559-568

Hubel, D.H. and T.N. Wiesel, 1965, Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat, *J. Neurophysiol.*, 28(a), pp. 229-289

Hubel, D.H. and T.N. Wiesel, 1966, Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex, *In* Pattern Recognition--Theory, Experiment, Computer Simulation and Dynamic Models of Form, Perception and Discovery, John Wiley & Sons, Inc., New York, N.Y.

Lashley,K.S., 1931, Mass action in cerebral function, *Science*, 73, pp. 245-254

Lashley,K.S., 1933, Integrative functions of cerebral cortex, *Physiol. Rev.*, 13, pp. 1-42

Lashley,K.S., 1950, In search of the engram, *Symp. Soc. Exp. Biol.*, 4, pp. 454-482

Lettvin, J.Y., H.R. Maturana, W.S. McCulloch, and W.H. Pitts, 1959, What the frog's eye tells the frog's brain, *Proc. Inst. Radio Engs.*, 47, pp. 1940-1951

Morrell, F., 1957, An anatomical and physiological analysis of electrocortical conditioning, *Proc. First Internat. Cong. Neurol. Sci.*, Brussels, Belgium

Ochs, S., 1965, Elements of Neurophysiology, John Wiley & Sons, Inc., New York, N.Y.

Ricciardi, L.M. and H. Umezawa, 1967, *Kybernetik*, 4, pp. 44

Rapoport, A., 1952, Ignition phenomena in random nets, *Bull. Math. Biophys.*, 14, pp. 35-44

Rochester, N., J.H. Holland. L.M. Haibt, and W.L. Duda, 1956, Tests on a cell assembly theory of the actions of the brain using a large digital computer, *IRE Trans. Inform. Theory*, 2, pp. 80-93

Ruch, T.C. and H.D. Patton, eds., 1966, Physiology and Biophysics, Saunders, Philadelphia, Pa.

Sholl, C.A., 1953, *J. Anat. London*, 87, pt. 4, pp. 387-406

Smith, D.R. and C.H. Davidson, 1962a, Maintained activity in neural nets, *J. Assoc. Comput. Mach.*, 9, pp. 268-279

Smith, D.R. and C.H. Davidson, 1962b, Activity levels and oscillation modes in neural nets, *In* Biological Prototypes and Synthetic Systems, E.E. Bernard and M.R. Kare, eds., Vol. I, Plenum, New York, N.Y., pp. 148-159

Trucco, E., 1952, The smallest value of the axon density for which ignition can occur in a random net, *Bull. Math. Biophys.*, 14, pp. 365-374

Yoshii, N., P. Pruvot, and H. Gestaut, 1957, *Electroencephalog. Clin. Neurophysiol.*, 9, p. 595

Yoshii, N., M. Shimokichi, and Y. Yamaguchi, 1960, *Med. J. Osaka Univ.*, 10, p. 375

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Department of Physics
Syracuse University
Syracuse, New York 13210

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP
N/A

3. REPORT TITLE

COMPOUND NET MODEL OF THE CEREBRAL CORTEX

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Final Report, February 1967 - November 1968

5. AUTHOR(S) *(First name, middle initial, last name)*
E. M. Harth, B. Beek, T. J. Csermely and R. D. Lindsay

6. REPORT DATE
June 1969

7a. TOTAL NO. OF PAGES
61

7b. NO. OF REFS
39

8a. CONTRACT OR GRANT NO.
F33615-67-C-1413

b. PROJECT NO. 7232

c. Task No. 723205

d. Work Unit 723205011

9a. ORIGINATOR'S REPORT NUMBER(S)

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
AMRL-TR-68-189

10. DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Aerospace Medical Research Laboratory
Aerospace Medical Div., Air Force Systems Command, Wright-Patterson AFB, OH 45433

13. ABSTRACT
The central nervous system is believed to consist of a network of neurons, combining in its structure genetically determined design with probabilistic features. A model is described in which probabilistic nets are used as building blocks from which systems can be assembled, approximating for example the interaction between sensory projection areas in the cortex with cortical association areas. The operation of such compound nets is based on the dynamics of single probabilistic nets and on a set of assumptions concerning the coding of sensory information into the language of central neuronal activity. Data describing the dynamics of simple nets are presented. It is found that various types of information processing can be explained by applying the hypothesis of synaptic facilitation to compound nets. Sustained neural activity is investigated, particularly the problem of ergodicity. Highly non-ergodic types of reverberations were found both theoretically and by computer simulation for certain types of nets.

DD FORM 1 NOV 65 1473

Security Classification

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Bionics						
Modeling of the Brain						
Information Processing						
Probabilistic Networks						
Neural Networks						

Security Classification