

A Mathematical Theory of the Functional Dynamics of Cortical and Thalamic Nervous Tissue

H. R. Wilson and J. D. Cowan

Department of Theoretical Biology, The University of Chicago, Chicago, Ill., USA

Received: January 14, 1973

Abstract

It is proposed that distinct anatomical regions of cerebral cortex and of thalamic nuclei are functionally two-dimensional. On this view, the third (radial) dimension of cortical and thalamic structures is associated with a redundancy of circuits and functions so that reliable signal processing obtains in the presence of noisy or ambiguous stimuli.

A mathematical model of simple cortical and thalamic nervous tissue is consequently developed, comprising two types of neurons (excitatory and inhibitory), homogeneously distributed in planar sheets, and interacting by way of recurrent lateral connexions. Following a discussion of certain anatomical and physiological restrictions on such interactions, numerical solutions of the relevant non-linear integro-differential equations are obtained. The results fall conveniently into three categories, each of which is postulated to correspond to a distinct type of tissue: sensory neo-cortex, archi- or prefrontal cortex, and thalamus.

The different categories of solution are referred to as dynamical modes. The mode appropriate to thalamus involves a variety of non-linear oscillatory phenomena. That appropriate to archi- or prefrontal cortex is defined by the existence of spatially inhomogeneous stable steady states which retain contour information about prior stimuli. Finally, the mode appropriate to sensory neo-cortex involves active transient responses. It is shown that this particular mode reproduces some of the phenomenology of visual psychophysics, including spatial modulation transfer function determinations, certain metacontrast effects, and the spatial hysteresis phenomenon found in stereopsis.

0. Introduction

0.1. Complexity and Redundancy in Cerebral Cortex and Thalamus

One of the most striking features of the evolution of the brain is that the enormous increase in volume of cerebral cortex relative to the rest of the brain is predominantly the result of an increase in surface area rather than an increase in thickness. Table 1 shows the comparative dimensions and main cytoarchitectonic features of the cortexes of the Edible Frog (*Rana esculenta*), Rabbit (*Lepus cuniculus*), and Man. It will be seen that while there is an increase by a factor of ten of cortical thickness, there is a corresponding increase of telencephalic surface area by a factor of some 3500.

Similarly, in comparing Rabbit with Man there is an increase of cortical thickness by a factor of only 1.5–1.8 as compared to an increase of telencephalic surface area by a factor of about 100.

Table 1. Dimensions and cytoarchitectonic features of cerebral cortex

Animal species	Cortex			Neurons	
	Volume (mm ³)	Surface area (mm ²)	Thickness (mm)	Number ($\times 10^6$)	Packing density (per 0.001 mm ³)
Edible frog (<i>Rana esculenta</i>)	6	24	0.25	6	1000
Rabbit (<i>Lepus cuniculus</i>)	1470	843	1.74	645	438
Man	230000	83591	2.4–3.00	10000	30–100

The main factor responsible for the increased surface area in higher Vertebrates is, of course, the enormous increase in the number of neurons. Furthermore the neurons of higher vertebrates are larger than those in lower Vertebrates, both in nuclear volume and in the extent of dendritic ramification (Blinkov and Glezer, 1968). This increase in neuronal size is accompanied by a decrease in packing densities, as is shown in Table 1. However, this decreased packing density is compensated for by the increased cortical thickness to such an extent that the actual number of neurons located in a cylinder of given cross-sectional area, and extending radially throughout the thickness of the grey matter, shows a rather small variation from species to species. For example, consider the number of neurons contained in a cortical cylinder with a cross-sectional area of 1 mm². In Frog such a cylinder is 0.25 mm long and contains about 250000 neurons.

In Rabbit the corresponding cylinder is 1.75 mm long and contains about 750000 neurons. Finally, in Man the cylinder is about 2.7 mm long and contains an average of 2000000 neurons.

The fact that the number of neurons in such radially oriented cortical cylinders varies by less than a factor of four compared to the 10^4 variation in the number of neurons in the cortexes of these three species suggests that the complexity of radially directed interactions within the cortex is roughly constant from species to species. On the other hand, the extent and complexity of *lateral* interactions in the higher Vertebrates may be expected to increase enormously, due to the increased number and size of neurons and consequently the greater extent of their dendritic ramifications (Blinkov and Glezer, 1968).

This increase in the lateral extent and complexity of interactions within cerebral cortex is paralleled by the appearance of much more elaborate and subtle forms of behavior in higher Vertebrates, especially Primates and Man. Furthermore, all sensory stimuli are received in such species as patterns falling on receptive surfaces that are at most two-dimensional, and these receptive surfaces project topographically to cerebral-cortex (Polyak, 1957). It therefore seems very likely that individual anatomical regions of cerebral cortex are functionally organised as two-dimensional surfaces or sheets, and that the cortex may be viewed as a complex, interconnected heterarchy of such sheets. Further support for this postulate is provided by the observation that specific afferents to primary sensory cortex synapse almost entirely within the plane comprising layer IV, and that superficial and deep layers receive afferents from layer IV mainly through radially oriented inter-connexions (Lorente de Nó, 1949; Sholl, 1956; Colonnier, 1965; Szentagothai, 1967; Scheibel and Scheibel, 1969). It follows that the specific afferent stimulus to primary sensory cortex is itself two-dimensional.

If differing anatomical regions of cerebral cortex are indeed functionally two-dimensional, what is the role of neurons distributed in depth, i.e., in the third (radial) dimension? One role for these neurons may be to provide reliability through *local redundancy*. It is well known to neurophysiologists that the response of any single neuron is variable, even under carefully controlled experimental conditions: only through the averaging of perhaps twenty to thirty trials are reliable results to be obtained. This variability has two sources: uncontrolled or extraneous signals from neurons in other parts of the brain, i.e., ambiguity and noise in the signal itself, and intrinsic fluctuations of membrane potential within the neuron. Mathematical

investigations of reliable information processing by systems composed of unreliable components have shown that any prescribed level of fidelity may be obtained so long as there is sufficient redundancy within these systems (v. Neumann, 1956), and that the most efficient systems are those which comprise complex components rather than simple ones (Winoograd and Cowan, 1963). It is therefore suggested that one reason for the distribution of neurons in depth might be to provide the local redundancy necessary for reliable operation, and that the increased complexity of radially oriented neuronal interactions in mammals as compared to lower vertebrates might serve to increase the efficiency with which reliable operation is obtained (see § 5.2.1).

Physiological evidence for such redundancy is provided by extracellular microelectrode studies of single neurons in somato-sensory cortex (Mountcastle, 1957) and striate and peri-striate cortex of Cats and Monkeys (Hubel and Wiesel, 1963, 1965, 1968). As is now well known, within a single radially organised cortical column the majority of neurons respond optimally to very nearly identical stimuli. Further support for the local redundancy hypothesis is provided by the existence of strong radial connectivity within cortex (Lorente de Nó, 1949; Colonnier, 1965; Szentagothai, 1967; Scheibel and Scheibel, 1970) and several anatomists have speculated on its significance in similar terms.

This view of cortex as a laterally organised net of radially redundant columns carries with it the implication that such columns are functional units, and that it is the over-all activity of neurons within a column which is functionally significant rather than the responses of any single neuron (see § 5.2.2). In the terminology of a previous paper, a column may be thought of as a spatially localised neuronal aggregate (Wilson and Cowan, 1972). However in the present paper columns will not be treated explicitly. Rather, it will be seen that localised neuronal activity consistent with the physiological concept of a column arises as a natural consequence of the interplay between topographic afferent projections and self-organising interactions within cortex. This view of cortical organisation is very similar to that of Scheibel and Scheibel (1970) in which cortex is conceived of as a matrix of overlapping radially oriented cylindrical modules, each module being defined by the extent of specific afferent terminal arborisations and the lateral spread of the dendritic branches of large cortical pyramids.

Entirely similar considerations obtain for sub-cortical structures such as thalamus. There are concomitant increases in the surface areas of these

structures in relation to receptive surfaces from Frog to Man, although not by such large factors (Blinkov and Glezer, 1968). Moreover there is clear evidence of the existence of functional columns, e.g., in the lateral geniculate nucleus (Bishop, Kozak, Levick, and Vakkur, 1962; Sanderson, 1971) correlated with receptive field location, and visual direction. It appears that equal numbers of retinal ganglion cells project to equal volumes of the lateral geniculate nucleus (Sanderson, 1971), and *a fortiori* to equal numbers of lateral geniculate neurons distributed in columns. It will be assumed that radial redundancy exists in these sub-cortical structures, and that like cortex, they are functionally two-dimensional.

It follows that the brain as a whole may be thought of as a complex hierarchy of functionally two-dimensional sheets of nervous tissue. In this paper a mathematical model will be developed in which cortical (or thalamic) tissue is represented as a single two-dimensional sheet. It will be assumed that neurons of various types are uniformly distributed within the sheet, and that lateral connectivity is a function of distance only, i.e., it will be assumed that the sheet is homogeneous and isotropic. The theory as developed in this paper is not specially tailored to any particular area of cortex or thalamus other than through the choosing of certain parameters to specify patterns of interconnection, and it may therefore serve as a basis for the analysis of very generalised cortex-like tissue. Certain details of cortical organisation, such as the orientational specificity of columns (Hubel and Wiesel, 1963, 1965, 1968) have clearly been omitted. It should be noted, however, that several neurophysiologists have questioned the functional significance of orientational specificity and have pointed out that it may be just an incidental manifestation of the asymmetry which must obtain if cortical neurons receive projections from two to three thalamic neurons (Burns, 1968; Creuzfeldt and Ito, 1968).

0.2. List of Symbols

$\alpha(t)$	Post-synaptic membrane potential (<i>psp</i>).
α	Maximum amplitude of <i>psp</i> .
t	Time.
μ	The neuronal membrane time constant.
ϑ	Threshold value of membrane potential.
r	Absolute refractory period.
τ	Synaptic operating delay.
v	Velocity of propagation of action potential.
x	Cartesian coordinate.
$\beta_{jj'}(x)$	The probability that cells of class j' are connected with cells of class j a distance x away.
$b_{jj'}$	The mean synaptic weight of synapses of the jj' -th class at x .
$\sigma_{jj'}$	The space constant for connectivity.

q_e	Surface density of excitatory neurons in a one-dimensional homogeneous and isotropic tissue.
q_i	Surface density of inhibitory neurons in a one-dimensional homogeneous and isotropic tissue.
$E(x, t)$	Excitatory Activity, proportion of excitatory cells becoming active per unit time at the instant t , at the point x .
$I(x, t)$	Inhibitory Activity, proportion of inhibitory cells becoming active per unit time at the instant t , at the point x .
δx	A small segment of tissue.
δt	A small interval of time.
$P(x, t)$	Afferent excitation or inhibition to excitatory neurons.
$Q(x, t)$	Afferent excitation or inhibition to inhibitory neurons.
$\bar{N}_e(x, t)$	Mean integrated excitation generated within excitatory neurons at x .
$\bar{N}_i(x, t)$	Mean integrated excitation generated within inhibitory neurons at x .
$\mathcal{S}_e[\bar{N}_e]$	Expected proportion of excitatory neurons receiving at least threshold excitation per unit time, as a function of \bar{N}_e .
$\mathcal{S}_i[\bar{N}_i]$	Expected proportion of inhibitory neurons receiving at least threshold excitation per unit time, as a function of \bar{N}_i .
$G(\vartheta_e)$	Distribution function of excitatory neuronal thresholds.
$G(\vartheta_i)$	Distribution function of inhibitory neuronal thresholds.
ϑ_1	A fixed value of neuronal threshold.
$h(\bar{N}_e; \vartheta_1)$	Proportion per unit time of excitatory neurons at x reaching ϑ_1 with a mean excitation \bar{N}_e .
$1[]$	Heaviside's "step-function".
$R_e(x, t)$	Number of excitatory neurons which are sensitive at the instant t .
$R_i(x, t)$	Number of inhibitory neurons which are sensitive at the instant t .
r_e	Refractory period of excitatory neurons.
r_i	Refractory period of inhibitory neurons.
$\langle E(x, t) \rangle$	Time coarse-grained excitatory activity.
$\langle I(x, t) \rangle$	Time coarse-grained inhibitory activity.
\otimes	Spatial convolution.
θ	Threshold of a neuronal aggregate.
v	Sensitivity coefficient of response of a neuronal aggregate.
$\langle E(t) \rangle$	Time coarse-grained spatially localised excitatory activity.
$\langle I(t) \rangle$	Time coarse-grained spatially localised inhibitory activity.
L_1, L_2, L, Q	See § 2.2.1, § 2.2.7, § 3.1.
v	Velocity with which retinal images are moved apart.
σ	Stimulus width.
$\langle E_0 \rangle, \langle I_0 \rangle$	Spatially homogeneous steady states of neuronal activity.
k_e, k_{ij}, S_{ij}	See § 5.1.
c_1, c_2, c_3, c_4	

1. The Mathematical Model

1.0. Relevant Properties of Neurons and Their Interconnexions

Cortical neurons can be classified in two very general ways. They are either excitatory or else inhibitory in their effects on other neurons (Eccles,

1964) and they are either efferent cells, or "inter-neurons". Efferent neurons possess long axons that penetrate the white matter beneath the cortex, whereas interneurons possess short axons that remain within the cortical grey matter. One can of course introduce further distinctions such as the differing size and shape of perikarya or of dendritic and axonal arborisations, but it will be assumed that such features may be accounted for in terms of the mean pattern of neuronal interconnexions at various locations within the cortex. It will therefore be assumed that there are two types of neurons in the cortex, excitatory and inhibitory, interspersed in depth and uniformly distributed relative to the pial surface. That is, it is assumed that equal numbers of excitatory and inhibitory neurons co-exist within the grey matter under equal areas of cortical surface, at least within a particular anatomical region. It is also assumed that the great majority of excitatory cells are efferents and that inhibitory neurons are interneurons, i.e., it is the activity of excitatory neurons which is projected from one anatomical region to another.

Individual neurons will be assumed to summate incoming excitation both spatially and temporally, in a linear time-invariant fashion. Thus the response to a unit excitation is represented by the function $\alpha(t)$ for an excitatory post-synaptic potential and $-\alpha(t)$ for an inhibitory post-synaptic potential. The usual form for $\alpha(t)$ is $\alpha \exp(-t/\mu)$, a decaying exponential with a time constant μ of several milliseconds. Neurons will also be assumed to have excitation thresholds ϑ for the generation of action potentials, measured in units of membrane potential, and absolute refractory periods of duration r following activation during which they cannot re-fire (see § 5.2.3). Finally neurons will be assumed to possess a latent period or synaptic delay τ , that is independent of excitation, and which measures the time lapse between excitation reaching threshold, and the consequent appearance of action potentials. These potentials are assumed to propagate without decrement along axons, with the velocity v .

All possible types of interconnexion will be permitted: excitatory-excitatory, excitatory-inhibitory, inhibitory-excitatory, and inhibitory-inhibitory. Each type will be taken to be a function only of the distance between points on the pial surface. Thus the tissue is assumed to be isotropic as well as homogeneous. A representation of this arrangement is shown in Fig. 1. It will be seen that all interconnexions are recurrent (Ratcliff, 1965) and either spatially local, or lateral. There is ample anatomical evidence for such circuitry in many cortical areas (Colonnier, 1965;

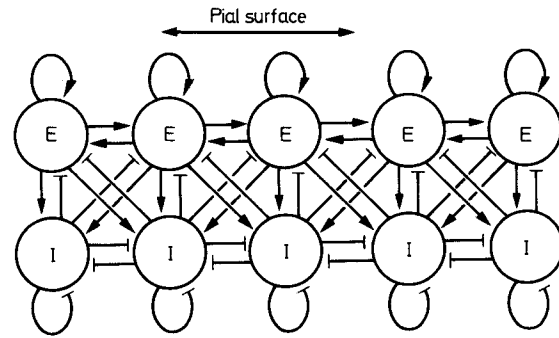


Fig. 1. Diagram of spatial interactions in neural tissue model. Lettered circles represent spatially localised excitatory and inhibitory neural aggregates; while the symbols \rightarrow and \dashrightarrow represent excitatory and inhibitory interactions, respectively. Note that interactions occur both locally and laterally relative to the pial surface. In the actual model, excitatory and inhibitory neurons are assumed to be interspersed and uniformly distributed relative to the pial surface

Szentagothai, 1967; Scheibel and Scheibel, 1970) (see § 5.2.4). The various types of interconnexion are represented by the functions $\beta_{jj'}(x)$ that measure the probability that post-synaptic neurons of type j' are connected with presynaptic neurons of type j located at a distance x . Following Uttley (1955) and Sholl (1956), $\beta_{jj'}(x)$ will be taken to be of the form $b_{jj'} \exp(-|x|/\sigma_{jj'})$ where both j and j' run over excitatory and inhibitory classes.

1.1. Neuronal Activity Variables

To describe the successive states of neuronal tissue composed of both excitatory and inhibitory neurons, two variables must be introduced, one for excitatory neuronal states and the other for inhibitory neuronal states. These variables will naturally be functions of position and time. By assumption, neuronal tissue is highly redundant especially in the radial direction, a fortiori, it is not single unit activity which is functionally important, but rather the net activity of all neurons of a given type located under a small element of the pial surface. This suggests the following representation of neuronal activity. Let excitatory and inhibitory neurons be homogeneously and isotropically distributed throughout the tissue with surface densities ϱ_e and ϱ_i respectively. Let $E(x, t)$ be the proportion of excitatory neurons becoming active per unit time at the instant t , at the point x in the tissue, and $I(x, t)$ the corresponding proportion of inhibitory neurons. Then $\varrho_e E(x, t) \delta x \delta t$ and $\varrho_i I(x, t) \delta x \delta t$ will give the numbers of excitatory and inhibitory neurons, respectively, becoming active under the surface element δx at x , during the interval between t and $t + \delta t$. Following

Beurle (1956), the functions E and I will be referred to as neuronal activities, or simply as activities.

In general the activities E and I will be random variables. However several factors suggest that the probability distribution associated with these variables will be sharply peaked about their mean values. Firstly, if the inter-connexions between neurons, although locally random, are sufficiently dense to justify the postulate of redundancy, then one can expect the variance in connectivity from neuron to neuron to be small. Therefore, within proximate cells of the same type, post-synaptic potentials should be highly correlated. Secondly, neuronal thresholds, plus the relatively rapid decay of post-synaptic potentials, combine in their effects to provide an effective temporal noise filter, so that only the co-operative activity of a large number of afferents will be sufficient to fire or inhibit neuronal activities. It is therefore legitimate to treat E and I as deterministic variables representing the mean values of sharply peaked probability distributions.

Before turning to the derivation of equations for E and I , two conventions are to be noted. Firstly, the state $E=0$, $I=0$, the resting state, is assumed to reflect low level background activity, so that small negative values of the activities are allowed, and serve to represent a depression or inhibition of the resting discharge. Secondly, only afferent activities (stimuli or forcing functions) that vary in one spatial dimension are considered. Because of the homogeneity and isotropy of the model, the neuronal response to such patterns will also vary only in the same dimension. Consequently the mathematical treatment will be limited to that of a one-dimensional tissue. The generalisation to two-dimensions is obvious.

1.2. Derivation of Equations for Neuronal Activity

It follows from the assumption of §1.0. that neurons will become activated only if their post-synaptic potentials exceed threshold, and if they are at the same time sensitive, i.e., non-refractory. Neglecting correlations that may obtain between levels of neuronal excitation and neuronal sensitivities, it follows that at the instant $t + \tau$, the activity located at x is equal to the proportions per unit time of excitatory and inhibitory neurons located at x whose integrated excitations exceeded threshold multiplied by (respectively) the proportions of excitatory and inhibitory neurons which were sensitive at the instant t .

In order to determine the proportions of neurons receiving supra-threshold excitation it is first necessary to derive an expression for the mean integrated

excitation generated in neurons at x by afferent activity. Consider first only excitatory neurons at x . At the instant t ,

$$q_e E \left(X, t - \frac{|x-X|}{v_e} \right) \beta_{ee}(x-X) dX$$

will be the mean rate of arrival of impulses at excitatory neurons located at x , caused by excitatory activity at the instant $t - \frac{|x-X|}{v_e}$ within a segment of length dX at X . Similarly at the instant t

$$q_i I \left(X, t - \frac{|x-X|}{v_i} \right) \beta_{ie}(x-X) dX$$

will be the mean rate of arrival of impulses at excitatory neurons located at x , caused by inhibitory activity at the instant $t - \frac{|x-X|}{v_i}$ within a segment of length dX at X .

The convolutions

$$\int_{-\infty}^{\infty} q_e E \left(X, t - \frac{|x-X|}{v_e} \right) \beta_{ee}(x-X) dX$$

and

$$\int_{-\infty}^{\infty} q_i I \left(X, t - \frac{|x-X|}{v_i} \right) \beta_{ie}(x-X) dX$$

therefore give the mean rates of arrival of impulses at excitatory neurons located at x , from all excitatory and inhibitory neurons (respectively) in the tissue.

In similar fashion because of the assumed linear time-invariant nature of temporal summation in neurons, the difference

$$\begin{aligned} \bar{N}_e(x, t) = & \int_{-\infty}^t \left[\int_{-\infty}^{\infty} q_e E \left(X, T - \frac{|x-X|}{v_e} \right) \beta_{ee}(x-X) dX \right. \\ & - \int_{-\infty}^{\infty} q_i I \left(X, T - \frac{|x-X|}{v_i} \right) \beta_{ie}(x-X) dX \\ & \left. \pm P(x, T) \right] \alpha(t-T) dT \end{aligned}$$

gives the mean value of integrated excitation for excitatory neurons at x (see 5.2.5). Here $P(x, t)$ has been introduced to represent afferent activities. As such it may be either excitatory or inhibitory in nature.

Let $\mathcal{S}_e(\bar{N}_e)$ be the expected proportion of excitatory neurons that receive at least threshold excitation per unit time, as a function of the mean integrated excitation \bar{N}_e . $\mathcal{S}_e(\bar{N}_e)$ is evidently a monotone-increasing function of \bar{N}_e , whose shape depends upon the actual distribution of neuronal thresholds at x . Let $G(\mathfrak{g}_e)$ be the distribution function of excitatory neuronal

thresholds in the tissue. Then by definition

$$\mathcal{S}_e(\bar{N}_e) = \int_0^{\bar{N}_e} G(\vartheta_e) d\vartheta_e.$$

Consequently \mathcal{S}_e will be a monotone function of \bar{N}_e as required. In addition it is bounded by the asymptotes 0 and 1, and if $G(\vartheta_e)$ is unimodal, \mathcal{S}_e will have only one point of inflexion, and will therefore be *sigmoidal* in shape (see 5.2.6). The value of \bar{N}_e at which the inflexion occurs will be referred to as the aggregate threshold for the excitatory neurons in the tissue. Some experimental support for the sigmoidal form of \mathcal{S}_e has been found in recordings from motoneuron pool discharges (Rall, 1955) (see 5.2.7).

To compute the number R_e of excitatory neurons which are sensitive at the instant t , note that it comprises all those neurons which have not been activated in the interval from $t - r_e$ to t . Evidently

$$R_e(x, t) = \left[1 - \int_{t-r_e}^t E(x, T) dT \right] \varrho_e \delta x.$$

If the expected number of excitatory neurons receiving at least threshold excitation during an interval δt at t is statistically independent of the proportion which is sensitive, then the expected number activated during the interval δt at $t + \tau$ is

$$E(x, t + \tau) \varrho_e \delta x \delta t = R_e \cdot \mathcal{S}_e(\bar{N}_e) \delta t.$$

In general there will be some correlation between the level of excitation in a neuron, and the probability that it is sensitive. However the correlation will be small in the densely inter-connected and redundant net considered (Wilson and Cowan, 1972) (see § 5.2.8).

Similar considerations obtain for the inhibitory interneurons in the sheet. Thus the expected number of such neurons activated during the interval δt at $t + \tau$ is

$$I(x, t + \tau) \varrho_i \delta x \delta t = R_i \cdot \mathcal{S}_i(\bar{N}_i) \delta t$$

where $R_i(x, t)$ and $\mathcal{S}_i(\bar{N}_i)$ are appropriately defined.

Introduction of the explicit forms of the expressions contained in these two equations results in the following equations for the neuronal activities generated in the sheet:

$$\begin{aligned} E(x, t + \tau) \varrho_e \delta x \delta t = & \left[1 - \int_{t-r_e}^t E(x, T) dT \right] \cdot \varrho_e \delta x \\ & \cdot \mathcal{S}_e \left[\int_{-\infty}^t \left[\int_{-\infty}^{\infty} \varrho_e E \left(X, T - \frac{|x-X|}{v_e} \right) \beta_{ee}(x-X) dX \right. \right. \\ & \left. \left. - \int_{-\infty}^{\infty} \varrho_i I \left(X, T - \frac{|x-X|}{v_i} \right) \beta_{ie}(x-X) dX \pm P(z, T) \right] \right. \\ & \left. \cdot \alpha(t-T) dT \right] \delta t, \end{aligned} \quad (1.2.1)$$

$$\begin{aligned} I(x, t + \tau) \varrho_i \delta x \delta t = & \left[1 - \int_{t-r_i}^t I(x, T) dT \right] \cdot \varrho_i \delta x \\ & \cdot \mathcal{S}_i \left[\int_{-\infty}^t \left[\int_{-\infty}^{\infty} \varrho_e E \left(X, T - \frac{|x-X|}{v_e} \right) \beta_{ei}(x-X) dX \right. \right. \\ & \left. \left. - \int_{-\infty}^{\infty} \varrho_i I \left(X, T - \frac{|x-X|}{v_i} \right) \beta_{ii}(x-X) dX \pm Q(x, T) \right] \right. \\ & \left. \cdot \alpha(t-T) dT \right] \delta t. \end{aligned} \quad (1.2.2)$$

The expressions $P(x, t)$ and $Q(x, t)$ represent afferent impulses or stimuli from other regions of the brain, or from sensory end organs. It will be seen that the surface densities ϱ_e and ϱ_i cancel out in Eqs. (1.2.1) and (1.2.2), to leave equations for the activities E and I that are valid in the limit $\delta x \rightarrow 0$, $\delta t \rightarrow 0$.

1.3. Simplification of the Neuronal Tissue Equations

Equations (1.2.1) and (1.2.2) provide a precise a representation of the phenomena involved in neuronal activation. These equations however are of a complicated non-linear character, and require considerable simplification to make them useful. Fortunately there are two ways in which the complexity may be significantly reduced without detracting from the physical plausibility of the equations. First, it is evident that if the average range of interaction between neurons is small compared to the velocities of impulse propagation v_j , then the time lags $|x-X|/v_j$ in Eqs. (1.2.1) and (1.2.2) will be negligible. Typical ranges of lateral interaction in Cat are of the order of 50μ (Sholl, 1956), whereas axonal conduction velocities are typically of the order of several mm per msec (Tasaki, 1959). Thus the time lags involved will be less than 0.1 msec (see § 5.2.9). This is very short compared to the membrane time constant μ of several msec, so that it is legitimate to treat the conduction velocity v_j as effectively infinite, and therefore to drop the time lag terms. Further support for this approximation will be provided later when it will be shown that the velocity of propagation of activity waves in the tissue is much slower than any reasonable estimate of axonal conduction velocity.

The second simplification follows from the observation that E and I appear on the right hand side of Eqs. (1.2.1) and (1.2.2) only as the time averages

$$\int_{t-r_e}^t E(x, T) dT, \quad \int_{t-r_i}^t I(x, T) dT, \quad \int_{-\infty}^t E(x, T) \alpha(t-T) dT$$

and

$$\int_{-\infty}^t I(x, T) \alpha(t-T) dT.$$

Consider first the integrals involving the membrane impulse response $\alpha(t)$. By assumption $\alpha(t)$ is of the form $\alpha \exp(-t/\mu)$ where μ , the membrane time constant, is of the order of several milliseconds. Thus both E and I are time averaged with a weighting function $\alpha(t)$ that damps out high frequencies. This suggests the introduction of the expressions

$$\langle E(x, t) \rangle = \frac{1}{\mu} \int_{-\infty}^t E(x, T) \exp(-(t-T)/\mu) dT$$

and

$$\langle I(x, t) \rangle = \frac{1}{\mu} \int_{-\infty}^t I(x, T) \exp(-(t-T)/\mu) dT$$

in place of E and I . It follows from the equation for $\langle E \rangle$ that $E(x, t) = \mu \frac{\partial}{\partial t} \langle E(x, t) \rangle + \langle E(x, t) \rangle$ and therefore that

$$\int_{t-r_e}^t E(x, T) dT = \int_{t-r_e}^t \left[\mu \frac{\partial}{\partial T} \langle E(x, T) \rangle + \langle E(x, T) \rangle \right] dT.$$

In most cases the refractory period r_e of cortical excitatory neurons is much less than the membrane time constant μ , Oshima (1969). It follows that $\langle E(x, t) \rangle \sim \langle E(x, t-r_e) \rangle$ and therefore that

$$\int_{t-r_e}^t \left[\mu \frac{\partial}{\partial T} \langle E(x, T) \rangle + \langle E(x, T) \rangle \right] dT \sim r_e \langle E(x, t) \rangle.$$

Furthermore the synaptic delay τ is usually an order of magnitude smaller than μ , so that

$$E(x, t + \tau) \delta t \sim \left[\mu \frac{\partial}{\partial t} \langle E(x, t) \rangle + \langle E(x, t) \rangle \right] \delta t.$$

Similar considerations obtain for I so that Eqs. (1.2.1) and (1.2.2) can be rewritten in terms of the time averaged activities $\langle E \rangle$ and $\langle I \rangle$, and take the forms

$$\mu \frac{\partial}{\partial t} \langle E(x, t) \rangle = -\langle E(x, t) \rangle + [1 - r_e \langle E(x, t) \rangle]$$

$$\cdot \mathcal{S}_e \left[\alpha \mu \left[\int_{-\infty}^{\infty} \varrho_e \langle E(X, t) \rangle \beta_{ee}(x-X) dX \right. \right. \\ \left. \left. - \int_{-\infty}^{\infty} \varrho_i \langle I(X, t) \rangle \beta_{ie}(x-X) dX \pm \langle P(x, t) \rangle \right] \right]$$

$$\mu \frac{\partial}{\partial t} \langle I(x, t) \rangle = -\langle I(x, t) \rangle + [1 - r_i \langle I(x, t) \rangle]$$

$$\cdot \mathcal{S}_i \left[\alpha \mu \left[\int_{-\infty}^{\infty} \varrho_e \langle E(X, t) \rangle \beta_{ei}(x-X) dX \right. \right. \\ \left. \left. - \int_{-\infty}^{\infty} \varrho_i \langle I(X, t) \rangle \beta_{ii}(x-X) dX \pm \langle Q(x, t) \rangle \right] \right].$$

These equations are valid in the limiting case when $\delta t \rightarrow 0$. It will be noted that they are partial differential-integral equations in which the integrals are spatial convolutions. As such they may be written in a more compact form. Let the convolution operation be denoted by the symbol \otimes . Then the simplified equations can be rewritten as

$$\mu \frac{\partial}{\partial t} \langle E(x, t) \rangle = -\langle E(x, t) \rangle + [1 - r_e \langle E(x, t) \rangle] \\ \cdot \mathcal{S}_e [\alpha \mu [\varrho_e \langle E(x, t) \rangle \otimes \beta_{ee}(x) \\ - \varrho_i \langle I(x, t) \rangle \otimes \beta_{ie}(x) \pm \langle P(x, t) \rangle]] \quad (1.3.1)$$

and

$$\mu \frac{\partial}{\partial t} \langle I(x, t) \rangle = -\langle I(x, t) \rangle + [1 - r_i \langle I(x, t) \rangle] \\ \cdot \mathcal{S}_i [\alpha \mu [\varrho_e \langle E(x, t) \rangle \otimes \beta_{ei}(x) \\ - \varrho_i \langle I(x, t) \rangle \otimes \beta_{ii}(x) \pm \langle Q(x, t) \rangle]] \quad (1.3.2)$$

In effect Eqs. (1.3.1) and (1.3.2) are time coarse-grained versions of Eqs. (1.2.1) and (1.2.2) respectively. Time coarse-graining is a technique which was first applied to problems of statistical physics (Kirkwood, 1946). Its validity in the present case depends of course on whether functionally relevant aspects of the activities $E(x, t)$ and $I(x, t)$ occur on a time scale that is sufficiently long compared to the membrane time constant μ . A more detailed analysis (Wilson and Cowan, 1972) has already shown that the technique can in principle be applied to the activity of many cortical regions. Thus the oscillations observed in evoked potential studies typically have periods of 40 msec or more, which is much longer than any reasonable value for the neuronal membrane time constant (Andersen and Eccles, 1962; Freeman, 1967, 1968a, 1968b; MacKay, 1970). Furthermore, the maximum rate at which a flickering visual stimulus is just perceptible, lies between 0.25 and 0.1 times the rate of discharge of retinal ganglion cells during the transient ON response (Brindley, 1971). This indicates that the limit of temporal resolution of the visual system is several times the length of the refractory period, i.e., in the range of the membrane time constant. Equations (1.3.1) and (1.3.2) thus provide a valid approximation for the types of dynamical phenomena to be analysed.

1.4. The Functional Form of \mathcal{S}_j and Constraints on Parameter Values

An earlier study has shown that the qualitative properties of solutions of Eqs. (1.3.1) and (1.3.2) are independent of the particular analytical form of the

sigmoidal functions \mathcal{S}_j (Wilson and Cowan, 1972). This is a valuable property of the equations, because of the difficulty in obtaining experimental determinations of the threshold distribution functions $G(\mathcal{S}_j)$, and also because of the likelihood that these functions will differ in different cortical regions. It is therefore convenient to introduce the logistic curve, translated downward so that $\mathcal{S}_j(0) = 0$, to serve as the representative for \mathcal{S}_j :

$$\mathcal{S}_j(\bar{N}_j) = [1 + \exp(-v(\bar{N}_j - \theta_j))]^{-1} - [1 + \exp(v\theta_j)]^{-1}.$$

The parameter θ_j , the aggregate threshold, determines the position of maximum slope of the function \mathcal{S}_j , and the parameter v determines the value of the maximum slope through the relation

$$\max \left\{ \frac{d\mathcal{S}_j}{d\bar{N}_j} \right\} = \frac{d\mathcal{S}_j}{d\bar{N}_j} \Big|_{\bar{N}_j = \theta_j} = \frac{v}{4}.$$

\mathcal{S}_e will be specified by the parameter values (v_e, θ_e) and \mathcal{S}_i by the parameter values (v_i, θ_i) . Figure 2 shows the graph of $\mathcal{S}_j(\bar{N}_j)$ for $v_j = 1$, $\theta_j = 5$.

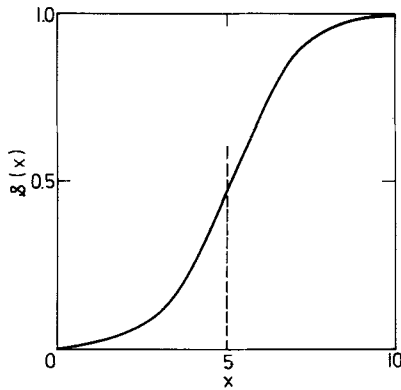


Fig. 2. A typical sigmoid function. The particular function shown here is the logistic curve: $\mathcal{S}_j(\bar{N}_j) = [1 + \exp(-v_j(\bar{N}_j - \theta_j))]^{-1}$ with $\theta_j = 5$ and $v_j = 1$

As noted in § 1.0, exponentials have been chosen for the connectivity functions $\beta_{jj'}(x)$:

$$\beta_{jj'}(x) = b_{jj'} \exp(-|x|/\sigma_{jj'}).$$

It is possible that Gaussians would have been a more propitious choice, but here again the qualitative results are independent of the form of $\beta_{jj'}(x)$, so long as it is monotone-decreasing with $|x|$ (see § 5.2.10). With the exception of Sholl's data (Sholl, 1956) suggesting that typical length constants for neuronal fields in Cat striate cortex are of the order of 50 μm , there is at present very little data from which to infer

values for the parameters $b_{jj'}$ and $\sigma_{jj'}$. Physiological considerations have therefore been used to obtain constraints on these parameters. The first consideration is that the resting state, $\langle E \rangle = \langle I \rangle = 0$, be stable for small perturbations. This is an important form of noise-insensitivity, essential to the normal functioning of cortex. A sufficient condition for such stability is that θ_e be sufficiently large compared to the product $q_e v_e b_{ee} \sigma_{ee}$, and that θ_i be sufficiently large compared to $q_e v_i b_{ii} \sigma_{ii}$ (see § 5.1).

The second consideration is that no uniformly excited state should be stable in the absence of a maintained stimulus. Such an activity state is not found in normal physiological preparations, and could only correspond to conditions of seizure. Since the functions \mathcal{S}_j have unity as their maximum value, the maximum attainable values of $\langle E \rangle$ and $\langle I \rangle$ are $[1 + r_e]^{-1}$ and $[1 + r_i]^{-1}$ respectively. Therefore the following inequalities are sufficient to ensure that such an activity state will not occur (assuming that $q_e = q_i = 1$):

$$[1 + r_e]^{-1} \cdot 2b_{ee} \sigma_{ee} - [1 + r_i]^{-1} \cdot 2b_{ie} \sigma_{ie} < \theta_e \quad (1.4.1)$$

and

$$[1 + r_e]^{-1} \cdot 2b_{ei} \sigma_{ei} - [1 + r_i]^{-1} \cdot 2b_{ii} \sigma_{ii} > \theta_i. \quad (1.4.2)$$

These inequalities are obtained by integrating the functions $\beta_{jj'}(x)$ over the interval $(-\infty, \infty)$. The second inequality ensures that uniform high level excitatory will generate uniform high level inhibitory activity, while the first ensures that the inhibitory activity will in turn be strong enough to extinguish the excitatory activity. The subsequent decline of excitatory activity is necessarily followed by a decline of the resultant inhibitory activity (see § 5.2.11).

The third and final constraint is that excitatory to inhibitory interactions must be of longer range than excitatory to excitatory interactions. That is:

$$\sigma_{ei} > \sigma_{ee}. \quad (1.4.3)$$

This assumption is supported by strong anatomical and physiological evidence (Colonnier, 1965; Szentagothai, 1967; Scheibel and Scheibel, 1970; Bishop *et al.*, 1971). As will be seen, the net effect of this inequality is to spatially localise excitatory activity, and to prevent the propagation of activity waves as obtained by Beurle (1956).

To these various constraints must be added several conventions that have been followed throughout this study. The first is that the aggregate threshold for excitatory neurons θ_e is always about 10 (in units of excitation) in keeping with observations by Eccles

(1965) on the thresholds of motoneurons and pyramidal cells. Secondly, the neuronal membrane time constant μ has been assigned a value of 10 msec. Such a value gives reasonable values for the frequency of oscillatory evoked potentials in spatially localised aggregates (Wilson and Cowan, 1972). Thirdly, the refractory periods r_e and r_i are both equal to 1 msec, a value that is reasonable for the duration of the absolute refractory period (Eccles, 1965). The effect of this is simply to fix the maximum attainable values of $\langle E \rangle$ and $\langle I \rangle$ at 0.5. Lastly, the values of the length constants $\sigma_{jj'}$, are all within a range around 50 μm , as suggested by Sholl's data (Sholl, 1956).

Although all these parameter values are definitely of the correct order of magnitude, no great significance need be attached to the particular values chosen, since the qualitative properties of the model are invariant to changes in such values over any reasonable range (see § 5.2.12).

2. Elementary Properties of the Model

2.0. Spatially Localised Aggregates

Prior to the analysis of Eqs. (1.3.1) and (1.3.2) it will be useful to review briefly the qualitative aspects of the activity generated within localised neuronal aggregates (Wilson and Cowan, 1972). These aggregates comprise interconnected excitatory and inhibitory neurons which are so closely packed together that all activity is spatially constant throughout the aggregate. In § 0.1 aggregates of this type were identified with cortical columns. The activity generated within such aggregates is represented mathematically by a simplified form of Eqs. (1.3.1) and (1.3.2) in which the connectivity functions $\beta_{ji}(x)$ are assumed to be constant over the aggregate. Thus the convolutions $\langle E(x, t) \rangle \otimes \beta_{je}(x)$ and $\langle I(x, t) \rangle \otimes \beta_{ji}(x)$ are replaced by $\beta'_{je} \langle E(t) \rangle$ and $\beta'_{ji} \langle I(t) \rangle$ respectively, where $\beta'_{jj'}$ are appropriately defined. The equations therefore become

$$\mu \frac{d}{dt} \langle E(t) \rangle = - \langle E(t) \rangle + [1 - r_e \langle E(t) \rangle] \cdot \mathcal{S}_e [\alpha \mu [\beta'_{ee} \langle E(t) \rangle - \beta'_{ie} \langle I(t) \rangle] \pm \langle P(t) \rangle] \quad (2.0.1)$$

and

$$\mu \frac{d}{dt} \langle I(t) \rangle = - \langle I(t) \rangle + [1 - r_i \langle I(t) \rangle] \cdot \mathcal{S}_i [\alpha \mu [\beta'_{ei} \langle E(t) \rangle - \beta'_{ii} \langle I(t) \rangle] \pm \langle Q(t) \rangle] \quad (2.0.2)$$

The major features of the solutions of these equations are listed below.

2.1. Local Properties of the Activity

2.1.1. Hysteresis. Localised aggregates, as represented by Eqs. (2.0.1) and (2.0.2), may be excited to a high level of activity, which subsequently is sustained by a much lower excitation than was necessary to trigger the change. In other words, multiple stable steady states may occur with the transitions between them forming hysteresis loops. There is a *threshold* for the excitation or stimulus intensity required to trigger this change of state within a given aggregate. There is a strength-duration law for threshold excitations, so that localised aggregates exhibit temporal summation.

2.1.2. Limit Cycles. Limit cycle oscillations of the activity may be obtained in response to constant stimulation. There is a threshold value of stimulus intensity for the generation of such oscillations. For super-threshold stimuli, the frequency of limit cycle oscillations is a monotone-increasing function of stimulus intensity. For extremely intense stimuli, limit cycle activity extinguished as the aggregate becomes saturated.

2.1.3. Hysteresis \Leftrightarrow Limit Cycles. If stimuli to excitatory and inhibitory neurons within an aggregate are independently variable, then the aggregate may exhibit either limit cycle activity or hysteresis switching between different steady activity states, depending upon the relationship between the two sets of stimuli.

These properties will be seen to extend and generalise with the introduction of spatial interactions.

2.2. The One-Dimensional Tissue

Numerical solutions of Eqs. (1.3.1) and (1.3.2) have been obtained for various sets of parameters satisfying the constraints given in § 1.4. These solutions illustrate three qualitatively different dynamical modes: active (self-generated) transient responses followed by a relaxation to the resting state, spatially localised limit cycles, and spatially inhomogeneous stable steady states that reflect attributes of prior stimuli (see § 5.2.13). Typical values of the parameters that give rise to each of the three forms are shown in Table 2. These values have been used in all the calculations reported in this article. The three dynamical forms, associated as they are with different sets of values of the various parameters, may be thought of as characteristic of anatomically distinct specialisations of sheets of nervous tissue, each serving different functions, in differing regions of cerebral cortex or thalamus.

Table 2. Parameters

Parameters (σ_{ij} in μm)	Active transient mode	Oscillatory mode	Steady-state mode
v_e	0.5	0.5	0.5
θ_e	9.0	9.0	9.0
v_i	0.3	1	0.3
θ_i	17.0	15.0	17.0
b_{ee}	1.5	2.0	2.0
σ_{ee}	40.0	40.0	40.0
b_{ie}	1.35	1.5	1.35
σ_{ie}	60.0	60.0	60.0
b_{ei}	1.35	1.5	1.35
σ_{ei}	60.0	60.0	60.0
b_{ii}	1.8	0.1	1.8
σ_{ii}	30.0	20.0	30.0

2.2.1. Active Transients. The self-generated transient response will be referred to simply as the "active transient". In this mode, the response to a brief localised stimulus of adequate intensity continues to increase even after cessation of the stimulus. The activity therefore reaches a peak value after the stimulus has ceased, and then decays back to the resting state. Figures 3 and 4 show two examples of the active transient. In each case the stimulus is a spatial square-wave of the form

$$\langle P(x, t) \rangle = \begin{cases} 0 & \text{for } x < L_1 \\ P & \text{for } L_1 \leq x \leq L_2 \\ 0 & \text{for } x > L_2 \end{cases}$$

applied for a duration Δt . By assumption $\langle P(x, t) \rangle$ excites only excitatory neurons within the tissue. If the width of the stimulus, $L_2 - L_1$, is sufficiently small, the excitatory response $\langle E(x, t) \rangle$ develops a single spatial maximum located at the centre of the segment originally stimulated. Moreover the excitatory response does not spread to any appreciable extent, but as a consequence of the longer ranged inhibitory effects which it evokes [see Eq. (1.4.3)] remains localised within the stimulated segment.

Figures 3 and 4 indicate that there is a threshold for the generation of active transients. Figure 3 shows that temporal summation is involved in reaching the threshold: a stimulus of supra-threshold intensity P will generate an active transient if and only if it is of sufficiently long duration Δt . It may be concluded that for stimuli of constant width, there is a strength-duration law of Block type (Le Grand, 1957).

At supra-threshold levels of excitation, the *latency* of the peak response decreases with increasing stimulus intensity (see Fig. 3). This phenomenon is characteristic of many sensory systems, notably the

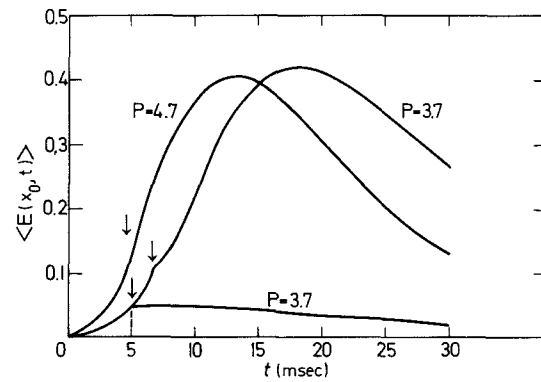


Fig. 3. Temporal summation and latency effects in the active transient mode. $\langle E(x_0, t) \rangle$ is the excitatory activity at the center of the narrow region stimulated. Stimulation begins at $t=0$ and continues until the time indicated by the arrow on each curve. Stimulus width is $80 \mu\text{m}$. Stimulus intensity is indicated by P values. The two curves for $P=3.7$ demonstrates the role of temporal integration in reaching threshold, while the superthreshold curves for $P=3.7$ and $P=4.7$ show the decreased latency to peak response as a function of increasing stimulus intensity

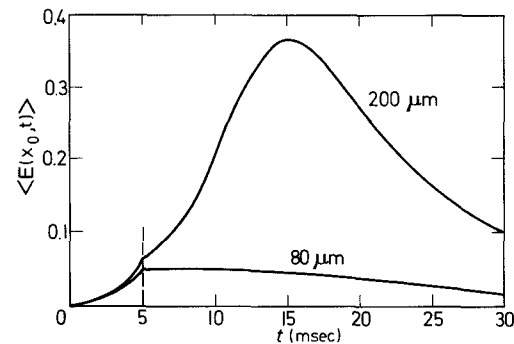


Fig. 4. Spatial summation in the active transient mode. $\langle E(x_0, t) \rangle$ is the excitatory activity at the center of the excitatory activity at the center of the region stimulated. The widths of the stimuli are $80 \mu\text{m}$ and $200 \mu\text{m}$. Both stimuli are of intensity $P=3.7$, and both are presented for 5 msec

visual system (Donchin, 1967; Hartline, 1938). It will be shown later that variable latency is important for the understanding of metacontrast effects.

Figure 4 indicates that spatial as well as temporal summation is involved in reaching the threshold for the generation of active transients. Thus a narrow stimulus applied for 5 msec fails to trigger the active transient, whereas a wider stimulus of the same intensity and duration evokes the transient. Spatial summation is also a characteristic property of the visual system.

It is in fact evident from the form of Eqs. (1.3.1) and (1.3.2) that absolute thresholds exist for stimulus width, intensity and duration. Thus stimuli that are too

narrow can never evoke active transient regardless of their intensity and duration. Similarly, stimuli that are too weak can never evoke active transients regardless of their width and duration, and so forth such threshold properties may be regarded as important forms of noise insensitivity in both cortical and thalamic tissue.

2.2.2. Edge Enhancement. If stimuli even wider than those discussed above are applied to the tissue, an important new phenomenon occurs during the course of the active transient. This is the phenomenon of *edge enhancement*. Figure 5 shows a series of profiles of $\langle E(x, t) \rangle$ at various instants following the cessation of the stimulus. It will be seen that a definite edge enhancement occurs, at about the time the active transient reaches its maximum value. The enhancement is then sustained throughout the decaying phase of the response. It is important to note that a definite latency exists before edge enhancement develops. Such a latency has been observed in psychophysical experiments (Kahneman, 1965), and is of the order of 40 msec (see § 5.2.14). Equations (1.3.1) and (1.3.2) predict latencies that are somewhat smaller than those observed in Kahneman's experiments (some 15 msec in the example shown in Fig. 5). The discrepancy may be accounted for by the observation that the visual pathway from retina to cortical area 18 comprises four distinct sheets of tissue (retina, lateral geniculate

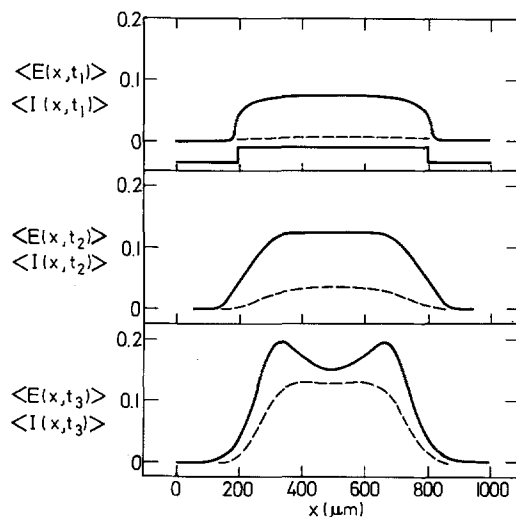


Fig. 5. Sequence of spatial profiles of neural activity showing the latency associated with the edge enhancement effect. The region stimulated is indicated at the bottom of the first panel. Excitatory and Inhibitory activities are indicated by solid and dashed lines respectively. The first panel shows the activities at the end of the period of stimulation, while the second and third show activity at progressively later times

Table 3. Effect of increasing stimulus intensity on maximal neural response

Intensity	$E(75, t_{\max})$	$E_{\max}(\text{edge})$	Effective width
3.4	0.072	0.084	280
3.7	0.151	0.195	320
6.0	0.230	0.284	500

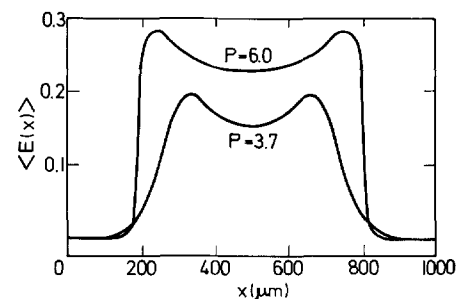


Fig. 6. The size-intensity effect in the active transient mode. $\langle E(x) \rangle$ is plotted at the instant at which the edge response was maximal. The stimuli are of constant width, with differing intensities as indicated

nucleus, cortical area 17, and cortical 18), that the psychophysical experiments almost certainly involve all these sheets, and that their separate latencies add.

Another important property of the edge enhancement effects generated by Eqs. (1.3.1) and (1.3.2) is shown in Table 3 and in Fig. 6. It will be observed that the *disparity* between the two points in the tissue where maximum edge enhancement occurs, increases with increasing stimulus intensity. If such a disparity is utilised in the nervous system to measure stimulus size, then it is predicted that a more intense stimulus will be seen to be somewhat wider than a weaker stimulus of the same actual width. This prediction has been confirmed psychophysically (Zusne, 1971). This size-intensity phenomenon may also provide an explanation of the fact that form perception thresholds are significantly higher than thresholds for detecting the mere presence of a stimulus (Zusne, 1971). This possibility is being investigated in greater detail using a two-dimensional extension of the present model.

2.2.3. Active Transients in Relation to Sensory Information Processing. Edge enhancement in the nervous system is perhaps best known in studies of the Limulus eye (Hartline and Ratliff, 1958 and Ratliff, 1965) has stressed its importance in form perception. The two effects discussed in § 2.2.2, the latency and size-intensity effect for edge enhancement, are both critically dependent on active properties of the neuronal tissue model represented by Eqs. (1.3.1) and

(1.3.2). Such effects cannot be obtained by lateral inhibition alone, but in addition require *recurrent lateral excitation*. This suggests that active transient responses may be characteristic of primary sensory cortex. In the visual system, for example, it is known from studies with stabilised retinal images (Ditchburn and Ginsberg, 1952) and from microelectrode recordings in cortical area 17 (Burns, Heron, and Pritchard, 1962; Burns, 1968) that the burst discharge of ON-centre retinal ganglion cells is of prime importance in stimulating visual cortex. The use of stimuli of very short duration Δt in Eqs. (1.3.1) and (1.3.2) provides a crude representation of this ganglion cell discharge. It follows that active transients are of some importance in providing a means by which sensory processing can continue long after the ON-centre discharge triggered by a stimulus has subsided. These areas are elaborated in § 3 in which it is shown that several more complex psychophysical effects may be simulated by the active transient mode of the present model.

2.2.4. Spatially Localised Limit Cycles. Consider now the second dynamical form of the solutions of Eqs. (1.3.1) and (1.3.2). In this form the response to a constant stimulus consists of limit cycle oscillations of $\langle E \rangle$ and $\langle I \rangle$. Furthermore the response occurs only within the region of stimulation, and does not propagate, except under very special circumstances (see § 2.2.7). Figure 7 shows a sequence of spatial profiles of $\langle E \rangle$ at successive instants during one period of the oscillation. Limit cycle oscillations differ from active transients in that, although they may be similar for stimuli that are of short duration compared to the period of the oscillation, the transients relax for constant maintained stimuli to constant (although spatially inhomogeneous) activity, whereas the oscillations are stable and continue as long as the stimulus is applied.

As one might expect, to generate sustained oscillatory activity, stimuli must exceed an absolute threshold intensity and a minimum width. In addition, Fig. 7 indicates that edge enhancement occurs for sufficiently wide stimuli, whereas narrow stimuli do not show any edge enhancement.

2.2.5. Limit Cycle Encoding of Stimulus Intensity. The salient characteristics of such oscillations are shown in Table 4. It will be seen that the frequency of oscillation increases monotonically with stimulus intensity, and that it is effectively independent of stimulus width. At very high stimulus intensities, however, the net inhibitory activity is insufficient for the generation of oscillations, and consequently the

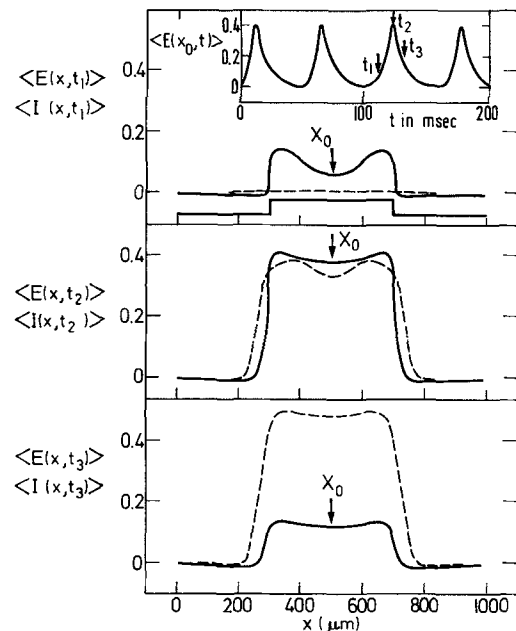


Fig. 7. Three phases in a spatially localized limit cycle oscillation in response to a maintained stimulus with the rectangular profile shown at the bottom of the first diagram. Solid line is $\langle E(x) \rangle$; dashed line is $\langle I(x) \rangle$. Inset at the top shows temporal details of the oscillation at the point marked x_0 in the other diagrams. The times t_1 , t_2 , and t_3 show the points on the cycle at which the spatial profiles are plotted

excitatory activity stabilises at a high (saturated) value. It is doubtful that such a saturated state ever occurs under physiological conditions in cerebral cortex, for its realisation would require an extremely high density of afferent fibre synapses directly with excitatory neurons. It follows that below the saturated state, stimulus *intensity* may be uniquely encoded in terms of limit cycle *frequencies*. In addition since the limit cycles remain localised and exhibit edge enhancement, stimulus *width* may be encoded in terms of the *disparity* between activity peaks, as discussed in § 5.2.2.

Table 4. Dependence of spatially localised limit cycle frequency upon stimulus intensity and stimulus width

Stimulus intensity	Stimulus width	Frequency of limit cycle response (sec^{-1})
2.5	80	14
2.5	600	14
5	80	18
5	400	18
5	600	18
10	400	22
10	600	23

Physiological evidence exists that certain sheets of tissue, thalamic rather than cortical, may indeed encode stimulus intensities in this fashion. In micro-electrode studies of thalamic somatic sensory neurons in Monkey, Poggio and Viernstein (1964) found that renewal density functions of the firing patterns of many units driven by joint angle receptors exhibited undamped oscillations. The frequencies of these oscillations were found to be monotone-increasing functions of joint angle. Although this evidence does not imply that aggregates of such units are firing in a cooperative oscillatory mode, it is what one would expect of the single unit properties deducible from Eqs. (1.3.1) and (1.3.2) given the constraints that generate limit cycles. Freeman (1968b) has observed similar relationships in evoked potentials obtained by macroelectrode recordings from the olfactory bulb and pre-pyriform cortex of Cat. The frequencies of oscillation of such potentials are again monotone-increasing functions of stimulus intensity, and are consistent with the predictions of the model.

2.2.6. Thalamic Oscillators. It is now well established that various thalamic nuclei are incapable of sustained oscillatory activity, and that these oscillations do not require feedback from cerebral cortex (Eccles, 1965; Andersen and Andersson, 1968; Purpura, 1970). Various mechanisms have been postulated to account for such oscillations, including rebound excitation from long lasting hyperpolarisation (Eccles, 1965; Andersen and Andersson, 1968). However Eccles (1965) has admitted that feedback excitation could also account for the observations and has postulated recurrent interactions within the ventrobasal complex of the thalamus that are essentially equivalent to those of Fig. 1. It is therefore suggested that Eqs. (1.3.1) and (1.3.2), suitably constrained to produce limit cycles, may serve to represent the specialisation of neuronal tissue to carry out appropriate thalamic functions.

To test the hypothesis that limit cycle activity in the tissue model is related to thalamic activity, an attempt was made to reproduce recent observations by Purpura (1970) of *frequency demultiplication* in the responses of certain regions of the ventro-laminar nucleus of the thalamus. The observations are that under certain conditions VL neurons respond only to every second pulse in a stimulus pulse train. A stimulus consisting of a train of pulses each of duration 5 msec at a frequency of 25 per sec was therefore applied to the tissue model. The pulse width was small enough so that no edge enhancement occurred. The excitatory activity generated by such a stimulus is shown in Fig. 8, where it is seen that the neuronal response does indeed occur at one-half the frequency

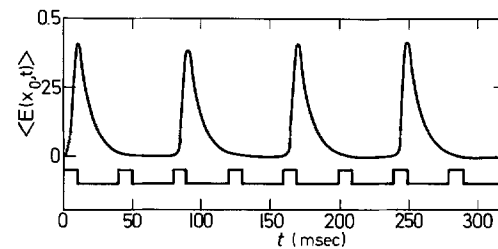


Fig. 8. Frequency demultiplication in the neural response to a stimulus pulse train. Each stimulus pulse had a narrow, rectangular spatial profile. $\langle E(x_0, t) \rangle$ is the response at the center of the region stimulated. Duration of each stimulus pulse is 10 msec and an intensity of $P = 5$. Pulses are presented at a frequency of 25/sec

of the stimulus train. By varying the frequency and intensity of the stimulus train, the neuronal response may be made to exhibit frequency demultiplication by $1/3$, $1/4$, and so on.

These results indicate that Eqs. (1.3.1) and (1.3.2) in the limit cycle mode may prove to be of value in theoretical studies of thalamic function, and of the consequent generation of some of the spatio-temporally organised rhythms seen in the EEG. The suggestion that the synchronised activities seen in the EEG (e.g., alpha, beta, delta, and theta rhythms) are generated by limit cycle oscillators, was first made by Dewan (1964).

2.2.7. Travelling Wave Responses. It is well known that most regions of cortical and thalamic tissue receive both specific and non-specific afferents (Lorente de No, 1949). In view of the fact that non-specific afferents affect very extensive regions in a diffuse fashion, it seems evident that their function is to provide a tonic biasing or threshold-setting effect. Furthermore it is known that cortico-fugal feedback, e.g., from cortical area 18 to the lateral geniculate nucleus (Kalil and Chase, 1970) serves to inhibit inhibitory interneurons there, i.e., tonic biasing operates by way of dis-inhibition.

Consider therefore Eqs. (1.3.1) and (1.3.2) with specific afferent excitation

$$\langle P(x, t) \rangle = \begin{cases} 0 & \text{for } x < L_1 \\ P & \text{for } L_1 \leq x \leq L_2 \\ 0 & \text{for } x > L_2 \end{cases}$$

and with non-specific afferent inhibition of the form

$$\langle Q(x, t) \rangle = -Q.$$

$\langle P(x, t) \rangle$ represents a spatial square wave whose width $|L_1 - L_2|$ is small enough to prevent edge enhancement from occurring.

For sufficiently small Q , the solutions of Eqs. (1.3.1) and (1.3.2) are characteristic spatially localised limit

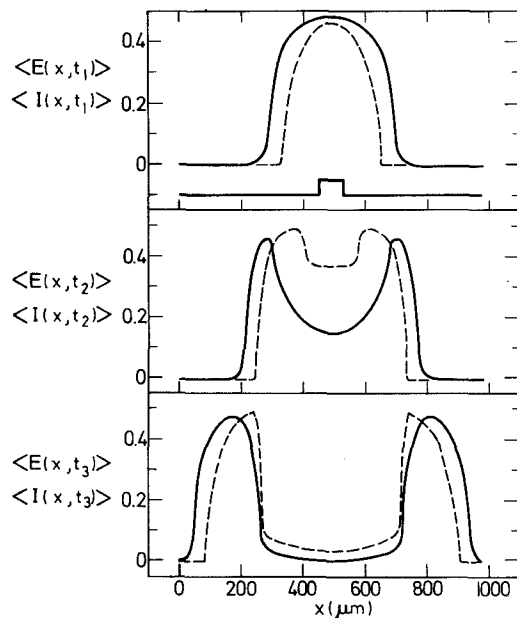


Fig. 9. Generation of traveling wave pair under disinhibitory binging in the oscillatory mode. $\langle E(x, t) \rangle$ is indicated by a solid line, and $\langle I(x, t) \rangle$ by a dashed line. The region initially stimulated is indicated beneath the top graph. $t_1 = 20$ msec; $t_2 = 30$ msec; $t_3 = 50$ msec

cycles, although with a modified frequency-intensity relationship. For sufficiently large values of Q however, a qualitatively new form of solution appears: instead of remaining localised, the peak excitatory phase of each limit cycle gives rise to a pair of waves travelling in opposite directions from the point of excitation. A sequence of spatial profiles detailing the generation of such a wave pair is shown in Fig. 9. Wave pairs are generated once per cycle for as long as the stimulus persists, and each wave travels away from the locus of stimulation without attenuation. Thus, a very brief stimulus $\langle P \rangle$ will generate a single wave pair, whereas a stimulus of longer duration will generate a succession of such pairs. The propagation velocity corresponding to $Q = -30$ is 4 cm/sec. In general this velocity is a function of Q , and of the connectivity parameters $\beta_{jj'}$, and $\sigma_{jj'}$, but not of $\langle P \rangle$, whereas the frequency and therefore the wave-length vary with $\langle P \rangle$ in a non-linear fashion.

2.2.8. Waves and Rhythms in Cortex. Such a generation of propagating wave-pairs as a consequence of non-specific dis-inhibition may be related to several physiological and psychophysical phenomena. For example, it is now known that in addition to the synchronised activities seen in the EEG, various desynchronised activities are also generated within sub-cortical structures (Andersen and Andersson,

1968), and that there are tonic excitatory and inhibitory influences on cortex from such structures (Demetrescu, Demetrescu, and Iosif, 1965). These interactions are extremely subtle and complex, and obviously cannot be accounted for by any simple model such as that discussed in this paper. However, some of the observations suggest that wave propagation and its quenching, of the type discussed here, may be involved in such interactions. For example recent observations by Freeman (personal communication) of spatio-temporal evoked potentials recorded from the surface of the olfactory bulb and pre-pyriform cortex indicate that both spatially localised oscillations and propagating waves obtain under differing experimental conditions. It therefore seems likely that with suitable elaboration, the model described above will prove to be of some utility in the analysis of such complex phenomena.

There is in fact a related but somewhat simpler set of observations on isolated (undercut) cortex (Burns, 1951). In such a preparation, electrical stimulation of one locus on the cortical surface generates waves that propagate without attenuation to the boundaries of the isolated region. On the assumption that under-cut cortex is dis-inhibited, wave generation and propagation, as in the tissue model, is to be expected following localised stimulation of the cortical surface. However, the propagation velocity of 20 cm/sec observed by Burns (1951) is significantly higher than that obtained from the tissue model, suggesting that spatial interactions in cortex are longer ranged than those incorporated in the model.

Burns (1958, 1968) has suggested that such waves may be related to the activity waves discussed by Beurle (1956). However, the neuronal medium considered by Beurle contained only excitatory neurons whose refractory periods were of sufficient duration so that each could be activated only once during the passage of a wave. In the tissue model presented here, the mechanism of wave propagation is quite different. Each neuron may fire many times during the passage of a wave front, but such activity generates enough inhibition in its wake to quench the excitation as the wave passes. The inordinately long refractory periods required in Beurle's model make it seem much less plausible than the excitatory-inhibitory model suggested here.

2.2.9. Spatially Inhomogeneous Steady States. The third form of the solution of Eqs. (1.3.1) and (1.3.2) consists of spatially inhomogeneous stable steady states of activity. [The existence of spatially homogeneous steady states is precluded by the inequalities

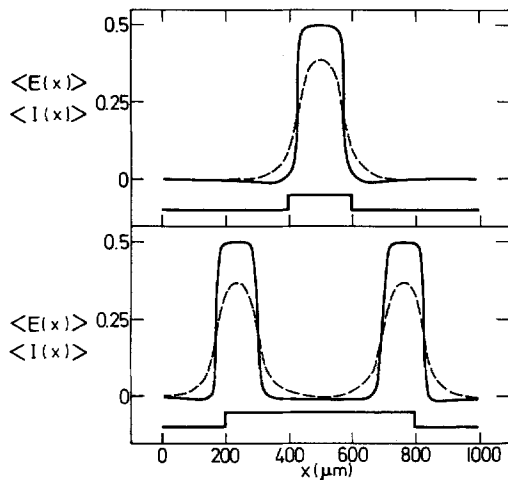


Fig. 10. Spatially inhomogeneous stable steady states of neural activity generated in response to two different stimuli. $\langle E(x) \rangle$ is represented by a solid line and $\langle I(x) \rangle$ by a dashed line. The rectangular inset below each graph shows the spatial configuration of the stimulus used to excite the tissue to each of the steady states. Once established by brief stimulation, the neural activity is self-maintained

(1.4.1) and (1.4.2).] Once reached such states are maintained by neuronal activity within the tissue without need of any further afferent stimulation. Two examples of these activity states are shown in Fig. 10 together with their generative stimuli. It is apparent that broad spatial square waves generate two peaks of activity approximately located at their edges, whereas sufficiently narrow stimuli generate only one peak. As one might expect, there is a threshold for the triggering of such stable peaks, related to the width, intensity, and duration of the applied stimuli in the conventional fashion.

2.2.10. A Possible Mechanism for Short Term Memory. The functional significance of these stable steady states becomes manifest with the observation that the two-dimensional generalisation of Eqs. (1.3.1) and (1.3.2) also generates stable steady states that reproduce the contours of prior two-dimensional stimulus patterns. (Preliminary studies of the two-dimensional case have confirmed this.) Such states may provide a mechanism that could account for *short-term memory*, and may be thought of as a complex spatial generalisation of dynamical hysteresis, a phenomenon first suggested as the physiological basis for short term memory by Cragg and Temperley (1955). Thus the contours of any brief but supra-threshold stimulus will be actively stored as a pattern of neuronal activity until erased by sufficiently strong afferent inhibition of excitatory neurons in the sheet, or by an equivalent afferent excitation of inhibitory

interneurons. Spatially localised activity of this type may also be thought of as resulting from localised *reverberation*: activity may circulate among neurons in the excited region in such a manner that the total activity at each point remains constant. The interpretation of spatially inhomogeneous stable steady states as a basis for short term memory is therefore compatible with Hebb's theory (Hebb, 1949).

Although this discussion of short term memory has been largely hypothetical, recent experiments by Fuster and Alexander (1971) provide some physiological support for the postulate. In these experiments Monkeys were trained to perform delayed response discriminations. During the interval between cue presentation and execution, marked increases were observed in the firing rates of neurons in prefrontal cortex and in the medial dorsal nucleus of the thalamus. Increases of up to ten times the resting discharge rate were observed which persisted for as long as one minute. These microelectrode recordings are what one would predict from the solutions of Eqs. (1.3.1) and (1.3.2) in the steady state mode.

2.2.11. Specifying the Different Solution Modes. The fundamental role played by the various parameters of Eqs. (1.3.1) and (1.3.2) in determining the dynamical mode of the solutions is obvious. Unfortunately the equations are so complex as to preclude the analytic determination of either necessary or sufficient conditions for the specification of each of the three forms of solution. Nevertheless, qualitative relationships between the sets of parameters that generate the different forms are evident in Table 2 (see § 2.2). Certain predictions based on these relationships have been tested and verified by numerical integration of the equations.

It will be seen from Table 2 that the parameter sets which lead to the active transient and steady state modes differ only in the value of the parameter b_{ee} . Since b_{ee} measures the strength of lateral excitatory-excitatory interactions, it follows that mode switching from steady states to active transients may be effected by decreasing the strength of excitatory (positive) feedback within the model tissue. Conversely, by increasing the strength of excitatory feedback, stable spatially inhomogeneous steady states can be created. This suggests the possibility that one of the effects of learning may be the creation of new large-scale stable states of activity in cortical tissue by way of the modification of excitatory-excitatory synapses. Such possibilities require further investigation.

The relationship between the parameter set determining the limit cycle mode and those determining the other two modes is much more complicated. The

major differences are in the values of the parameters b_{ie} , b_{ei} and b_{ii} . In the limit cycle mode b_{ie} and b_{ei} are larger, and b_{ii} is smaller than in the other modes. In effect, there is stronger negative feedback within the model tissue as a consequence of the increased strength of excitatory-inhibitory-excitatory circuits and a decreased strength of inhibitory-inhibitory (or dis-inhibitory) circuits. In addition, the parameter v_i is also greater in the limit cycle mode. This implies an increase in the maximum slope or sensitivity of the aggregate response function $\mathcal{P}_i(\bar{N}_i)$ (see § 1.4), and therefore a decrease of the range between resting and saturated activity of the inhibitory interneurons in the tissue. It seems that by sharpening the inhibitory response function, the instability necessary for limit cycle oscillations is created. The slightly lower value of the aggregate threshold θ_i is not critical (see § 1.4).

There are clearly other possibilities in the relationships among the various parameter sets. For example, increasing σ_{ie} or σ_{ei} should produce similar effects to those obtained by increasing b_{ie} or b_{ei} . Further insight into such possibilities may perhaps be obtained from the analysis of spatially localised aggregates developed in a previous paper (Wilson and Cowan, 1972).

3. Three Sensory Experiments

3.0. Introduction

In § 2 it was suggested that neuronal tissue in the active transient mode might be involved in sensory information processing. In this section it will be shown that the active transient response can serve to provide plausible explanations for a variety of psychophysical observations associated with the visual system. In discussing the various experiments, it is to be emphasised that the same parameters are used throughout (those listed in Table 2 that generate active transients): only stimulus parameters are varied for the different experiments.

3.1. Spatial Modulation Transfer Functions

The spatial modulation transfer function is a measure of the response of the visual system to different spatial frequencies. In the first sensory experiment this function was determined for the model tissue in the active transient mode. To accomplish this the excitatory neurons were stimulated for a duration $\Delta t = 5$ msec with spatially varying afferent excitation of the form

$$\langle P(x, t) \rangle = \frac{k}{2} \left[1 + \cos\left(\frac{2\pi n x}{L}\right) \right]$$

where n is an integer. Subsequent to this stimulation excitatory activity in the tissue continued to increase until a maximum amplitude was reached at a latency of approximately $\Delta t = 10$ msec, after which the activity decayed to the resting state. In all cases the response was spatially periodic, with a frequency equal to that of the stimulus, but because of the non-linearities in Eqs. (1.3.1) and (1.3.2), the actual waveform was not sinusoidal. The maximum peak to trough amplitude of the excitatory response $\langle E(x, t) \rangle$, was chosen as a measure of the sensitivity of the model tissue to stimulation at each spatial frequency. Figure 11 shows this amplitude plotted as a function of spatial frequency, for stimuli of constant amplitude.

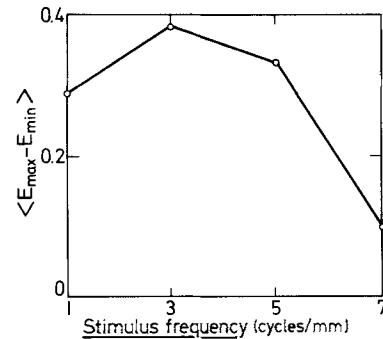


Fig. 11. Spatial modulation transfer function of the tissue model. Stimulation is brief and of the form:

$$\langle P(x) \rangle = 1 + \cos\left(\frac{2\pi n x}{L}\right)$$

$\langle E_{\max} - E_{\min} \rangle$ gives the maximum peak to trough amplitude of the excitatory response

The attenuation of both low and high spatial frequencies exhibited by the model is consistent with the properties of the modulation transfer function for the Human visual system, as measured psychophysically (Campbell and Green, 1965). Such measurements involve the determination of contrast thresholds for distinguishing spatially varying gratings from a *Ganzfeld*. Since the model exhibits both threshold phenomena and spatial summation, it is clear that results similar to those shown in Fig. 12 can be obtained from the determination of threshold sensitivities to spatially varying gratings. However such threshold determinations require large amounts of computer time and have not been carried out.

It has been suggested that the spatial modulation transfer function for the visual system is formed by the optical attenuation of high frequencies by the lens,

and the neuronal attenuation of low frequencies by lateral inhibitory nets (Ratliff, 1965). Campbell and Green (1965) used laser generated interference f fringes to show that the optical properties of the Human eye can contribute little to the high frequency attenuation, and therefore that it must be mainly neuronal in origin. The computations reported here show that the neuronal tissue model incorporating *recurrent lateral excitation* in addition to recurrent lateral inhibition [subject to the inequality of (1.4.3)] can account for the spatial filtering properties of the Human visual system. The available evidence does not permit one to speculate however, on the precise anatomical location where this filter occurs.

3.2. Metacontrast Phenomena

3.2.0. Introduction. In a second series of model experiments, various spatio-temporal interactions of the *metacontrast* type have been investigated. Metacontrast is defined as the perceptual masking of a brief target stimulus by a second stimulus presented subsequently at a location in the visual field that does not overlap that of the target (Kahneman, 1968). All psychophysical experiments on metacontrast involve two-dimensional targets, such as discs or alphanumeric characters, and a two-dimensional mask that is either an annulus encircling the area of target presentation, or else a pair of bars flanking it (Kahneman, 1968; Kohlers, 1962; Kohlers and Rosner, 1960). Since only one dimensional spatial patterns are considered in this paper, it is obviously impossible therefore to replicate the actual psychophysical experiments. Instead, one dimensional bars have been used as target stimuli, and pairs of such bars as masks. The problem thus reduces to a consideration of one-dimensional interactions. The results obtained in this fashion may be expected to hold for rectangular bars that are much longer than they are wide. Whether or not qualitative differences are likely to arise in more complex two-dimensional cases will be discussed later.

3.2.1. Type A Masking. The procedure adopted was to stimulate the tissue model with a target bar of specified width and supra-threshold intensity for an interval of duration $\Delta t = 5$ msec. The response was allowed to form unimpeded for a variable interval, at the termination of which masking bars of the same width but greater intensity than the target were presented, at locations flanking but not overlapping the locus of target presentation. The spatial relationships are shown in the inset of Fig. 12. To measure the degree of masking, responses to the target alone were compared with responses to the target followed by the

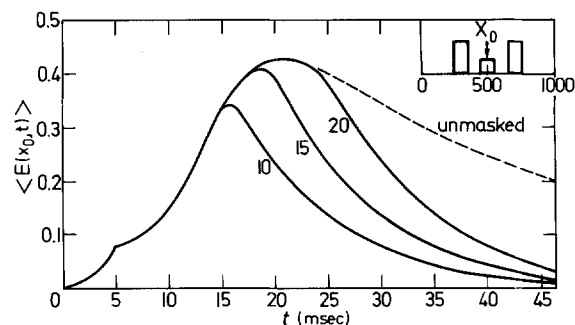


Fig. 12. Type A metacontrast effects in the active transient mode. $\langle E(x_0, t) \rangle$ is the excitatory response at the point of target presentation, as indicated in the inset. Figures on the curves are times at which masking bars are presented at the locations shown by cross-hatching in the inset. The dashed line indicates the unmasked result. Target presentation is for 5 msec, and the relative intensities of target and mask are given by heights of bars in the inset

mask. The results are shown in Fig. 12 for various intervals between target and mask presentation. In each case the width of the target was chosen to be sufficiently narrow so as to preclude edge enhancement, so that the target response might be accurately represented by a temporal graph of the activity generated at the centre of the target location (see inset, Fig. 12).

It will be noticed that unless the interval between target and mask is short, the mask has little effect on the *maximum* target response, but affects mainly the decaying phase of the target response. However it will be recalled that the tissue model exhibits temporal summation (Fig. 3, § 2.2.1) so that an appropriate measure of the target response is not the maximum value attained, but the integral of the response over its entire duration, i.e., the area under each curve shown in Fig. 12, or $\int_0^t \langle E(x, T) \rangle dT$. Using this integral to

measure masking, it will be seen that masking is a monotone-decreasing function of the interval between target and mask presentations. Other computations have shown that the degree of masking is also a monotone-decreasing function of the distance between target edges and the inside edges of the masking bars.

The monotonic masking shown in Fig. 12 has been designated type A by Kohlers (1962). Typically, type A masking functions measured psychophysically show strong effects even if masks are presented as long as 50 msec after targets. One might anticipate that such long intervals would not lead to masking in the model. In type A masking however, the mask is typically of much higher contrast than the target. It was shown in § 2.2.1 that high intensity stimulation generates

active transients of very short peak latency; therefore the neuronal response to a high contrast mask may be expected to propagate through retina, lateral geniculate, and several cortical areas at a much higher velocity than the target response. The mask response can therefore reach cortex concurrently with the target response, even if it is initiated some 50 msec after the target response. This implies that the location of the neuronal interactions leading to type A masking must be cortical. The fact that type A masking may be obtained in dichoptic stimulus presentations (i.e., with the target presented to one eye and the mask to the other at the appropriate retinal location) (Kohlers, 1962; Kahneman, 1968) supports the presumption that it is a cortical phenomenon.

3.2.2. Type B Masking. A second non-monotone masking phenomenon exists, first observed by Kohlers and Rosner (1960) and designated type B by Kohlers (1962). In this type of masking the optimum effect is obtained if the mask is presented approximately 50 msec after the stimulus has terminated. If the mask follows the stimulus with a very short delay, both stimulus and mask are perceived as simultaneous. If the interval between target and mask is very long however, the target and mask are perceived as consecutive. The appropriate stimulus arrangements for realising type B masking are that target and mask be virtually contiguous and of equal intensities (Kahneman, 1968).

These stimulus conditions have been used in an effort to obtain type B masking within the tissue model, so far without success. Since the mask used in type B experiments is usually always an annulus, it is possible that the two-dimensional model will reproduce the phenomenon, but this has yet to be attempted. Whatever the outcome, there is one detail which suggests that there may be a fundamental difference between type A and type B masking. In type B masking, both target and mask are of equal intensity and therefore will generate responses of equal latency. There can therefore be no question of mask responses arriving at cortex concurrently with target responses, as postulated for type A phenomena, so that propagation through successive sheets of neuronal tissue cannot account for the occurrence of optimal type B masking effects at a delay of around 50 msec. There is another interesting possibility. It is now known that corticofugal fibres to the lateral geniculate nucleus originate almost entirely from cortical area 18 in Cat (Hollander, 1971). The proper latency for optimal type B masking would obtain if such effects were the result of interactions within the lateral geniculate nucleus of retinally

generated mask responses and cortical feedback of target responses. If this were the case, all aspects of type B masking could be readily explained: simultaneous presentation of target and mask would lead to the simultaneous perception of both stimuli; since both stimuli are of equal intensity, neither could mask the other. [Actually, masking is mutual but weak for both stimuli in this case, in both model and psychophysical experiments (Kahneman, 1968).] The optimal delay of 50 msec would then be accounted for as the delay necessary to produce maximum interactions of retinal and cortical responses in the lateral geniculate nucleus. This possibility is currently being examined in greater detail.

3.3. Spatial Hysteresis in Binocular Vision

3.3.0. The Fender-Julesz Experiment. In the final series of model experiments reported in this paper an important phenomenon based on binocular fusion has been investigated. This phenomenon was discovered by Fender and Julesz (1967) in the course of experiments with stabilised retinal images. Binocular stabilisation techniques were used so that the locations on the retinae of identical patterns simultaneously presented to the two eyes could be independently controlled. These patterns are initially presented with zero binocular disparity, so that subjects report seeing the patterns as a single fused percept. The patterns are then slowly and symmetrically moved apart on the two retinae until the subject perceives a double image. At this point the direction of pattern motion is reversed, and the patterns are slowly moved together until the subject again perceives a single fused image. The surprising result is that the actual retinal disparities necessary for refusion are always significantly less than those required to maintain fusion once it has been established. Thus a form of *spatial hysteresis* is operative in stereopsis (see Fig. 3 and 6 of Fender and Julesz, 1967).

Since there are topographic projections from retina to striate cortex, it can be supposed that retinal stimulus disparities produce equivalent disparities in the resultant cortical responses, and that spatial hysteresis is a cortical phenomenon associated with stereopsis. In view of the demonstrable existence of hysteretic effects in spatially localised neuronal aggregates (§ 2.1.1 and Wilson and Cowan, 1972), it is appropriate to look for spatial hysteretic effects directly analogous to those discussed above in the one-dimensional tissue model.

3.3.1. Spatial Hysteresis in the One-Dimensional Tissue Model. In the one-dimensional model all stimuli

are equivalent to vertical bars. However one of the stimulus patterns used by Fender and Julesz (1967) was a single vertical bar, so that no difficulties arise in considering the experiment as a one-dimensional spatial problem.

It was assumed that the retinal stimuli could be reasonably approximated by a pair of sharply peaked Gaussian distribution functions, one for each eye. These stimuli were equal in amplitude and shape throughout the experiment, and were assumed to represent the maintained retinal discharge rather than the transient ON response, since it must be supposed that under the high contrast stabilised image conditions employed by Fender and Julesz, the maintained retinal discharge, rather than the transient ON response, provided the effective stimulus to cortex. Initially the two Gaussian stimuli were placed in exact register on the tissue model and summed point by point to produce the net (binocular) stimulus. The active transient response was then allowed to develop and reach equilibrium (because of the intense and sustained stimulation, the active transient does not in this case decay to the resting state). Following this the two stimuli were symmetrically moved apart by a small increment, re-summed, and the neuronal response was again allowed to reach equilibrium. This procedure was repeated to produce increasing stimulus disparity.

Mathematically this stimulation routine may be written as

$$\langle P(x, t) \rangle = k[\exp(-(x - vt)^2/\sigma^2) + \exp(-(x + vt)^2/\sigma^2)],$$

where v is the (small) velocity at which the two "retinal" stimuli are moved apart, and σ measures the width of each Gaussian stimulus. In the model experiment the continuous velocity v was approximated by shifting the two stimuli apart in small steps at a constant (slow) rate.

Initially with the two stimuli in exact register, the response consisted of a single peak of sustained high level activity as shown in Fig. 13a. As the disparity between the two stimuli increased, the response continued as a single peak located at the mid-point of the binocular stimulus until a critical disparity was reached. At this stimulus disparity, the single peak response decayed rapidly to zero, and twin response peaks formed at the locations of the now rather widely separated stimuli. A spatial profile of this response is shown in Fig. 13b. Further increases in stimulus disparity merely produced further corresponding increases in the disparity of these response peaks.

Following this the stimuli were gradually moved together again in the same fashion. Again a critical disparity was reached at which the response pattern

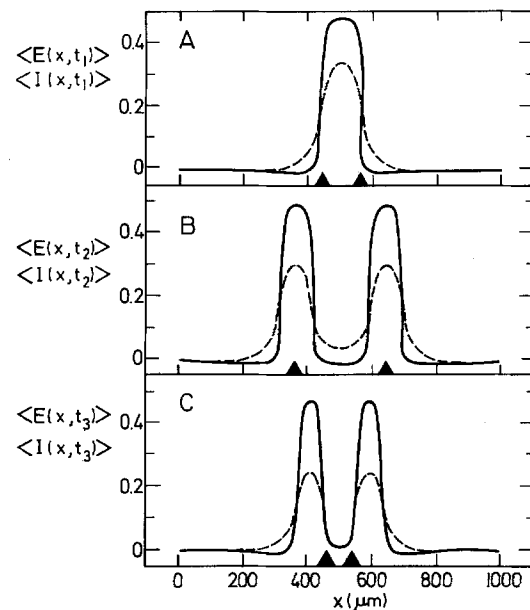


Fig. 13. Neural activity at three stages of the Fender-Julesz hysteresis loop. $\langle E(x, t) \rangle$ is plotted as a solid line and $\langle I(x, t) \rangle$ as a dashed line. A, B, and C indicate the points on the hysteresis loop in Fig. 14 at which the activities are plotted. Positions of the maxima of the Gaussian stimuli are indicated under each graph by the solid triangles. Note that although the stimuli are further apart in A than in C, the response in A is a single peak, whereas the response in C exhibits twin peaks

switched from two peaks of activity to a single peak. However, this critical disparity was much smaller than the first critical disparity described above. The difference may be seen by contrasting Fig. 13c with 13a. Although the stimulus disparity shown in Fig. 13a is greater than that shown in Fig. 13c, the response profile shows a single peak in the first case and twin peaks in the second case.

A more effective way of representing the results of this model experiment is shown in Fig. 14 in which the disparity between peaks of excitatory neuronal activity is plotted as a function of the disparity between peak amplitudes of the two stimuli. The disparity is taken to be zero when the response consists of a single peak of activity. Starting at zero stimulus disparity, the arrows in Fig. 14 indicate that as the stimuli are moved apart, the response remains unimodal until the disparity increases beyond approximately 180 μm . Thereafter the bimodal response is rapidly established. As the stimuli are again brought back into register, the arrows indicate that the response peaks do not move together as quickly as do the stimulus peaks (because of the retarding effects of recurrent lateral inhibition). Furthermore refusion does not occur at 180 μm and the bimodal response is maintained until a disparity

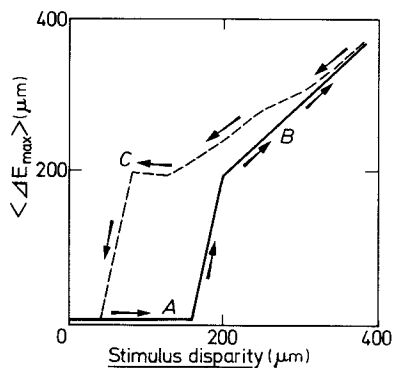


Fig. 14. Simulation of the Fender-Julesz experiment on hysteresis in binocular vision. Stimulus disparity is the distance between maxima of the two displaced Gaussian stimuli, while $\langle \Delta E_{\max} \rangle$ is the separation between maxima in the neural response. Arrows show the hysteresis loop generated by first pulling the stimuli apart and then bringing them back into register

of about $80 \mu\text{m}$ is reached, after which refusion is rapidly established. The letters *a* – *c* in Fig. 14 indicate the points on the hysteresis loop from which the response profiles shown in Fig. 13 were taken.

The circuit indicated by the arrows in Fig. 14 is clearly hysteretic, and is generated by spatio-temporal interactions within the model tissue. If unimodal responses are identified with binocular single vision and bimodal responses with double vision, then the hysteresis loop shown in Fig. 14 reproduces the results of the Fender-Julesz experiment. The reader is referred to Fig. 3 of the original article for comparison (Fender and Julesz, 1967).

3.3.2. Random Dot Stereograms. Two other aspects of the Fender-Julesz experiments are also of interest. In addition to bar targets, experiments were also performed using highly textured two-dimensional patterns [random dot stereograms (Julesz, 1971)]. Spatial hysteresis was again observed, but much greater disparities were found to be necessary for breakaway and much smaller ones for refusion. In effect, the range of stimulus disparities over which spatial hysteresis was operative increased in both directions relative to the results obtained using vertical bar targets. Since the model problem is now two-dimensional, attempts have not yet been made to reproduce the results. It seems certain however, that the increased effectiveness of highly textured patterns in producing spatial hysteretic effects is directly related to the fact that highly textured chess board patterns generate much stronger cortical evoked potentials than do simpler line targets (Spehlmann, 1965). It has been found that a "chessboard" stimulus will similarly produce a stronger response than a bar in the two-

dimensional tissue model, so that it is very likely that highly textured target will also generate stronger hysteretic effects in the two-dimensional model.

3.3.3. Occlusion Effects. The other interesting aspect of the Fender-Julesz experiments is contained in the determination of the duration of occlusion of the two retinal images that is necessary to cause the breakaway of a fused image at various disparities. As one might expect, the critical duration of occlusion was found to be a monotone-decreasing function of increasing retinal stimulus disparity. It has already been shown that strength-duration curves of Block type obtain for threshold phenomena in the neuronal tissue model (§ 2.2.1). This fact, together with the observation that the strength of the stimulus actually delivered to the tissue model at the site of single (fused) response peak, decreases with increasing stimulus disparity, indicates that a qualitatively similar occlusion effect can be expected from the model.

3.3.4. Relation to Two Theories of Binocular Vision. Spatial hysteresis is a key feature of the model for stereopsis and depth perception recently developed by Julesz (1971). The Julesz model employs two interacting arrays of spring-coupled magnetic dipoles, and is based upon an analogy to magnetic hysteresis, since no neuronal mechanisms for spatial hysteresis were known at the time it was formulated (see § 5.2.15). The close qualitative similarity between the Julesz model, and the neuronal tissue model described in this paper, becomes apparent when it is realised that the locking of magnetic dipoles effectively provides local excitatory feedback, whereas the springs that couple adjacent dipoles provide the equivalent of longer ranged recurrent lateral inhibition.

Sperling (1970) has also developed a rather detailed though qualitative physical and neuronal model of binocular vision. This model is similar to that one described here, in that short range recurrent excitation is coupled to longer ranged recurrent lateral inhibition. However, the implications of these assumptions for large-scale spatio-temporal phenomena are not developed in a quantitative fashion. Sperling's postulated neuronal binocular field could readily be constructed from several interconnected model tissue sheets, and it seems clear that the spatial hysteresis effects previously discussed would provide a quantitative foundation for the predictions of Sperling's model. The Sperling model is actually more closely related to the difference-field model, also developed by Julesz (1971), than it is to the spring-coupled magnetic dipole model described above. Furthermore, the two models do differ in their representation of depth information, and

in the dynamics of the process of fusion. The single sheet neuronal tissue model presented in this paper is not by itself a complete model for depth perception and stereopsis. However it does provide a plausible neuronal interpretation of visual hysteretic phenomena. It is expected that further extensions of the tissue model will provide more insights into the differing neuronal mechanisms underlying the Julesz and Sperling models.

3.4. Concluding Remarks

In concluding this section it is re-emphasised that the spatial modulation transfer function, metacontrast and hysteretic effects were all obtained using the same choice of parameters in Eqs. (1.3.1) and (1.3.2), i.e., those for the active transient mode. The different results represent the responses of one model to differing stimuli. This is consistent with the hypothesis advanced in § 2.2.3 that the active transient mode may represent qualitatively the responses of sensory cortical tissue. The known anatomy of striate cortex is certainly much more complicated than is reflected in Eqs. (1.3.1) and (1.3.2), but it is hoped that the relative simplicity of the model may serve as a basis for a better understanding of the functional significance of cortical complexity.

4. Discussion

Throughout this paper it has been assumed that the appropriate description of the dynamics of nervous tissue is in terms of the interactions between aggregates of neurons distinguish according to cellular type. Given this assumption it is natural to choose as dynamical variables the proportions of neurons in the various aggregates becoming activated per unit time in a small region of the tissue. With this choice of variables, and a few simplifications such as time coarse-graining and the neglect of relative refractoriness, the model embodied in Eqs. (1.3.1) and (1.3.2) is readily derived.

The anatomy represented in the model is a reasonable approximation to actual neuronal tissue, for the basic circuitry has been proposed for many neuronal regions. For example, the model anatomy is similar to that of the olfactory bulb and pre-pyriform cortex (Allison, 1953). The primitive nature of these structures re-inforces the idea that the present model may serve as a first simple approximation to the more complex structures of cerebral neo-cortex and related thalamic nuclei. However even this simple model exhibits considerable complexity. Three fundamentally distinct dynamical response modes have been found, each associated with a different set of connectivity

parameters. Since these parameters are anatomical in nature, it has been postulated that each mode represents a different anatomical specialisation of neuronal tissue to perform a distinct physiological function.

Of the three anatomies perhaps the most interesting one is that supporting active transient responses. It has been shown that there is a threshold for the generation of active transients, involving both spatial and temporal summation. Further properties include variable latencies, edge enhancement of sufficiently wide stimuli, and a relationship between the width of the neuronal representation of a stimulus and the stimulus intensity. Since these properties are all characteristic of primary sensory systems, it is postulated that the active transient anatomy is characteristic of primary sensory cortex, particularly of primary visual cortex. This postulate is supported by the fact that such a model anatomy reproduces, with one fixed set of parameters, a variety of sensory experiments: spatial modulation transfer determinations, metacontrast effects, and the spatial hysteresis effect discovered by Fender and Julesz (1967).

The second model anatomy supports the generation of long-lasting oscillatory responses, and is assumed to be characteristic of thalamic tissue. The characteristic responses are spatially localised limit cycles for constant spatially inhomogeneous stimuli. The frequency of such oscillations codes for stimulus intensity, and again, edge enhancement occurs for sufficiently wide stimuli. Frequency de-multiplication of stimulus pulse trains also occurs, and with suitable diffuse dis-inhibition, propagating waves are generated. It is suggested that such dis-inhibition is one of the factors contributing to the generation of the various rhythms observed in the EEG.

The third model anatomy supports the generation of spatially inhomogeneous stable steady states of neuronal activity. Excitation to such states by brief stimuli involves threshold effects and spatial and temporal summation. Such spatially inhomogeneous states retain information about the contours of stimuli. It is therefore postulated that these states may be involved in short-term memory and that the associated anatomy, involving strong recurrent excitatory interconnexions, is characteristic of archi- and pre-frontal cortex.

The various properties described above all arise from the interplay between short range recurrent lateral excitation and longer ranged recurrent lateral inhibition. The current model may therefore be thought of as a natural extension of several previous studies of spatial interactions in nervous tissue.

Perhaps the first important study was that of Beurle (1956) who analysed the properties of a medium composed exclusively of excitatory neurons. Inhibitory neurons were not included and the refractory periods of the excitatory neurons were assumed to be so long that each could fire only once in response to a given stimulus, although this assumption was later removed in computer simulation studies by Beurle (1962) and Farley and Clark (1961). The present model may be thought of as an extension of Beurle's to include the multiple firing of neurons and recurrent lateral inhibition as well as excitation. The most elegant and best known studies of a system involving lateral inhibition are those of Hartline and Ratliff on the compound eye of *Limulus* (Hartline and Ratliff, 1958; Ratliff, 1965). The present model may also be thought of as an extension of the Hartline-Ratliff paradigm to incorporate recurrent lateral excitation. An attempt by Griffith (1963, 1965) to develop a neuronal "field theory" may also be cited here, although it was unsuccessful, mainly because of a failure to deal with inhibition as arising from a distinct aggregate of inhibitory neurons. The appropriate two variable description of excitatory-inhibitory interactions in a net was first developed by Cowan (1968, 1970) and has been used in this paper. Finally, the present model is an extension to the spatial domain of the non-linear dynamics of localised neuronal aggregates previously developed by the authors (Wilson and Cowan, 1972).

It is to be emphasised that all spatial interactions in the model are recurrent. This feature distinguishes the model from retinal models such as Sperling's (1970). It is in fact the incorporation of *recurrent excitatory interactions* that leads to the active properties of the model tissue: thresholds for self-generated excitation, inhomogeneous stable steady states, limit cycle oscillations, spatial hysteresis, and so on. Such active properties seem to be present in cortical and thalamic responses as manifested in such phenomena as short term memory, selective attention, dreaming, and ultimately in cognition. Obviously these active properties must be characteristic of central rather than retinal or other peripheral processes. [Recurrent interactions are present in the retina, but are almost certainly inhibitory, and involved in adaptation (Rushton, 1965).]

The model developed here is only a crude first approximation to the anatomical and physiological complexity of actual nervous tissue. However, now that a framework for dealing with spatio-temporal interactions between aggregates of distinct neuron types has been developed, it will be seen that the model may be readily extended to include spatio-temporal

interactions among three or more aggregates. In general, the model for N aggregates will involve N simultaneous first order integro-differential equations of the same form as Eqs. (1.3.1) and (1.3.2), together with N^2 connectivity functions specifying the spatial interactions among and within the N aggregates. The model can readily incorporate such further complexities as the differential magnification of central and peripheral regions in cortical area 17 (Daniel and Whitteridge, 1961) through the use of spatial connectivity functions with position dependent coefficients. In such a fashion the model may be extended to incorporate much more of the detailed complexity of nervous tissue. To cite but one example, the "simple" and "complex" neurons of Hubel and Wiesel (1963, 1965, 1968) might be modelled as two distinct aggregates of excitatory neurons.

In concluding, one important ambiguity in the anatomical interpretation of the present model should be noted. It has been assumed that all recurrent interactions occur between neurons located in the grey matter of a single anatomical region of cortex or thalamus. However, it is known that many cortical and thalamic regions are reciprocally and topographically connected. If the conduction time for impulse propagation between two such reciprocally connected regions is sufficiently short, then some of the dynamical properties of the cortical tissue model might actually be properties of the two interconnected regions rather than of either one in isolation. The resolution of such ambiguities must await further anatomical and physiological data.

5. Appendix

5.0. Introduction

The fundamental equations discussed in this paper, Eqs. (1.2.1) and (1.2.2) and the time-coarse grained forms (1.3.1) and (1.3.2) are inherently non-linear, and generate complex non-linear phenomena. Linear equations cannot generate such phenomena as active transients that persist after the cessation of stimuli, limit cycle oscillations, or spatially inhomogeneous stable steady states. Furthermore, even in Eqs. (1.3.1) and (1.3.2) it is not possible to solve for the spatial steady states and then determine their stability by linear analysis. Fortunately it is possible to linearise Eqs. (1.3.1) and (1.3.2) about a spatially homogeneous state, $\langle E(x, t) \rangle = \langle E_0(t) \rangle$ and $\langle I(x, t) \rangle = \langle I_0(t) \rangle$ and then study stability properties of the resulting equations. The point is that given $\langle E_0 \rangle$ and $\langle I_0 \rangle$ it is always possible to choose constant values of

$\langle P(x, t) \rangle$ and $\langle Q(x, t) \rangle$ so as to make $(\langle E_0 \rangle, \langle I_0 \rangle)$ a steady state (not necessarily stable) of Eqs. (1.3.1) and (1.3.2).

5.1. Linearised Stability Analysis

It is therefore assumed that appropriate values of $\langle P(x, t) \rangle$ and $\langle Q(x, t) \rangle$ have been chosen in such a fashion, so that the linearised forms of Eqs. (1.3.1) and (1.3.2) are, respectively

$$\mu \frac{\partial}{\partial t} \langle E'(x, t) \rangle = -k_e \langle E'(x, t) \rangle + S_e [\langle E'(x, t) \rangle \otimes \beta_{ee}(x) - \langle I'(x, t) \rangle \otimes \beta_{ie}(x) \pm \langle P'(x, t) \rangle] \quad (5.1.1)$$

and

$$\mu \frac{\partial}{\partial t} \langle I'(x, t) \rangle = -k_i \langle I'(x, t) \rangle + S_i [\langle E'(x, t) \rangle \otimes \beta_{ei}(x) - \langle I'(x, t) \rangle \otimes \beta_{ii}(x) \pm \langle Q'(x, t) \rangle] \quad (5.1.2)$$

where

$$\langle E'(x, t) \rangle = \langle E(x, t) \rangle - \langle E_0(t) \rangle$$

and similarly for $\langle I' \rangle$, $\langle P' \rangle$, and $\langle Q' \rangle$, and where

$$k_e = 1 + r_e \mathcal{S}_e[\langle E_0 \rangle, \langle I_0 \rangle];$$

$$S_e = [1 - r_e \langle E_0 \rangle] \cdot \frac{\partial \mathcal{S}_e(Y)}{\partial Y} \bigg|_{\substack{\langle E' \rangle = 0 \\ \langle I' \rangle = 0}},$$

$$k_i = 1 + r_i \mathcal{S}_i[\langle E_0 \rangle, \langle I_0 \rangle];$$

$$S_i = [1 - r_i \langle I_0 \rangle] \cdot \frac{\partial \mathcal{S}_i(Y)}{\partial Y} \bigg|_{\substack{\langle E' \rangle = 0 \\ \langle I' \rangle = 0}}.$$

In principle Eqs. (5.1.1) and (5.1.2) may be solved exactly by Fourier transforming with respect to x . Unfortunately, the inverse transform involves extraction of quartic roots, so that the method is not practicable. However all the important properties may be obtained by considering only perturbations that are spatially periodic. It will therefore be assumed that $\langle P'(x, t) \rangle$ and $\langle Q'(x, t) \rangle$ have Fourier series representations, so that only solutions need be considered for individual harmonics of $\langle P' \rangle$ and $\langle Q' \rangle$.

The response of the linearised tissue equations to the perturbation $\cos \left[\frac{2n\pi x}{L} \right]$ will be of the form

$$\begin{aligned} \langle E'(x, t) \rangle &= \langle E'(t) \rangle \cos \left[\frac{2n\pi x}{L} \right]; \\ \langle I'(x, t) \rangle &= \langle I'(t) \rangle \cos \left[\frac{2n\pi x}{L} \right]. \end{aligned} \quad (5.1.3)$$

On substituting these expressions into Eqs. (5.1.1) and (5.1.2) one obtains:

$$\begin{aligned} \mu \frac{d}{dt} \langle E'(t) \rangle &= \left[-k_e + \frac{2S_e b_{ee} \sigma_{ee}}{1 + (2n\pi \sigma_{ee}/L)^2} \right] \cdot \langle E'(t) \rangle \\ &\quad - \left[\frac{2S_e b_{ie} \sigma_{ie}}{1 + (2n\pi \sigma_{ie}/L)^2} \right] \cdot \langle I'(t) \rangle + S_e \end{aligned} \quad (5.1.4)$$

and

$$\begin{aligned} \mu \frac{d}{dt} \langle I'(t) \rangle &= \left[\frac{2S_i b_{ei} \sigma_{ei}}{1 + (2n\pi \sigma_{ei}/L)^2} \right] \cdot \langle E'(t) \rangle \\ &\quad - \left[k_i + \frac{2S_i b_{ii} \sigma_{ii}}{1 + (2n\pi \sigma_{ii}/L)^2} \right] \cdot \langle I'(t) \rangle. \end{aligned} \quad (5.1.5)$$

These equations are of the form

$$\mu \frac{d}{dt} \langle E' \rangle = c_1 \langle E' \rangle - c_2 \langle I' \rangle + S_e,$$

$$\mu \frac{d}{dt} \langle I' \rangle = c_3 \langle E' \rangle - c_4 \langle I' \rangle.$$

The eigenvalues are therefore proportional to

$$c_1 - c_4 \pm \sqrt{[(c_1 + c_4)^2 - 4c_2 c_3]}/2. \quad (5.1.6)$$

Thus a necessary condition for stability is that $c_4 > c_1$. It follows from Eqs. (5.1.4) and (5.1.5) that this condition is

$$\frac{2S_i b_{ii} \sigma_{ii}}{1 + (2n\pi \sigma_{ii}/L)^2} + k_i > \frac{2S_e b_{ee} \sigma_{ee}}{1 + (2n\pi \sigma_{ee}/L)^2} - k_e. \quad (5.1.7)$$

If the homogeneous steady state $(\langle E'_0 \rangle, \langle I'_0 \rangle)$ is near the resting state $(0, 0)$, then k_e and k_i will both be approximately unity. If $\sigma_{ee} > \sigma_{ii}$, the conditions (5.1.7) will be satisfied for all n as long as $S_i b_{ii} \sigma_{ii} > S_e b_{ee} \sigma_{ee} - 1$. When this condition is met, a sufficient condition for the stability of the solutions of Eqs. (5.1.4) and (5.1.5) is that the square root in (5.1.6) be imaginary. For this to be the case it is necessary and sufficient that

$$\begin{aligned} &\frac{16S_e S_i b_{ie} \sigma_{ie} b_{ei} \sigma_{ei}}{(1 + (2n\pi \sigma_{ie}/L)^2)(1 + (2n\pi \sigma_{ei}/L)^2)} \\ &> \left[\frac{2S_e b_{ee} \sigma_{ee}}{1 + (2n\pi \sigma_{ee}/L)^2} + \frac{2S_i b_{ii} \sigma_{ii}}{1 + (2n\pi \sigma_{ii}/L)^2} + k_i - k_e \right]^2. \end{aligned} \quad (5.1.8)$$

Suppose that (5.1.8) is satisfied for small n , if k_e is even slightly different from k_i however, the inequality will be violated for sufficiently large n . Thus for sufficiently large values of n the eigenvalues of Eqs. (5.1.4) and (5.1.5) must be real. In such a case the requirement for stability becomes

$$\begin{aligned} &c_1 - c_4 + \sqrt{[(c_1 + c_4)^2 - 4c_2 c_3]} < 0 \\ &\text{or} \\ &c_2 c_3 < c_1 c_4. \end{aligned}$$

In terms of the coefficients in (5.1.4) and (5.1.5) this condition becomes:

$$\frac{4S_e S_i b_{ie} \sigma_{ie} b_{ei} \sigma_{ei}}{(1 + (2n\pi\sigma_{ie}/L)^2)(1 + (2n\pi\sigma_{ei}/L)^2)} > \left[\frac{2S_e b_{ee} \sigma_{ee}}{1 + (2n\pi\sigma_{ee}/L)^2} - k_e \right] \left[\frac{2S_i b_{ii} \sigma_{ii}}{1 + (2n\pi\sigma_{ii}/L)^2} + k_i \right]. \quad (5.1.9)$$

Since k_e and k_i are both positive, it follows that (5.1.9) is automatically satisfied for very large n . Suppose that the coefficients have been chosen so that (5.1.9) is also satisfied for small n . For the remaining intermediate values of n however, the situation may be different. It has already been argued on physiological grounds that $\sigma_{ei} > \sigma_{ee}$ [Eq. (1.4.3)]. If in addition $\sigma_{ie} > \sigma_{ii}$, then the left hand side of (5.1.9) will decrease faster than the right hand side as n increases. It is thus possible that (5.1.9) will be violated for intermediate values of n , thus leading to instability.

These considerations lead to the conclusion that Eqs. (5.1.1) and (5.1.2) may be stable with respect to harmonic perturbations of both low and high spatial frequency, yet unstable to perturbations of intermediate spatial frequency. Even if the equations are not actually unstable with respect to such intermediate frequencies, they will be less stable than at high frequencies, and consequently more sensitive. This property is manifest in the spatial modulation transfer function of the non-linear Eqs. (1.3.1) and (1.3.2). It is also the basis for the edge enhancement effects in the active transient mode, and for the edge extraction effects in the steady state mode.

Unfortunately, linearised analysis tells one little about the oscillatory mode since limit cycle oscillations are functions of the *global* properties of the non-linear equations. Nor does the linearised analysis indicate whether an instability at intermediate frequencies will lead to an active transient response, or to a spatially inhomogeneous steady state.

5.2. Notes and Comments

1. § 0.1 p. 56. Such redundancy is not necessarily a simple increase of the number of neurons involved in any given operation (v. Neumann *op cit*), but rather an increased complexity of the overall operations served by neurons inter-connected into radially organised nets (Winograd and Cowan *op cit*). In such *functionally* redundant nets each neuron is involved in numerous local operations, and each local operation is effected by many different groups of neurons.

2. § 0.1 p. 56. The laminar structure of cortex and thalamus does not vitiate the argument for radial redundancy. Such structure may be no more than the most economical way to specify and realise inter-connexions within and among functional columns.

3. § 1.0 p. 58. The neglect of any relative refractory period during which neurons have higher than normal thresholds leads to a great

simplification of the theory. Equations incorporating relative refractory neurons have been developed in a previous paper (Wilson and Cowan, 1972). All the results obtained in the current paper can be obtained by specialisation from these equations.

4. § 1.0 p. 58. It was first pointed out to the authors by W. J. Freeman Jr., that if excitatory neurons are taken to be mitral cells, and inhibitory neurons to be granule cells, the theory can be seen to incorporate the basic cytoarchitectonic features of the olfactory bulb. Similarly, pre-pyriform cortex may be taken to consist of excitatory pyramidal neurons and inhibitory neurons. The fact that the pre-pyriform or olfactory cortex is phylogenetically one of the most primitive parts of the brain is further support for the investigation of a model such as is presented here as a prelude to the development of more complex models that might more fully incorporate the detailed architecture of cerebral cortex.

5. § 1.2 p. 59. This expression is an extension of that introduced by Beurle (1956) and is based on the assumed *additivity* of post-synaptic effects. Sperling (1970) introduced *divisive* or shunting inhibition to model retinal interactions. As Sperling indicated, shunting inhibition is a convenient means for the introduction of back-ground adaptation. Divisive inhibition could easily be introduced into the present model, but has been excluded, partly to preserve the mathematical simplicity of the model and also because there is as yet very little experimental evidence for the existence of such inhibition in cerebral cortex.

6. § 1.2 p. 60. The effects of multimodal densities $G(\vartheta)$ on $\mathcal{S}_e(\bar{N}_e)$ have been discussed in a previous paper (Wilson and Cowan, 1972).

7. § 1.2 p. 60. In actual fact the analysis given above represents a simplification of the actual details of the interactions that lead to the activation of neurons in the sheet. A more detailed analysis can be sketched out as follows [see Cowan (1972) for more details]. At any instant t , the proportion per unit time of neurons at x reaching a fixed threshold with a mean excitation \bar{N}_e comprises a proportion per unit time reaching the threshold *for the first time* in the interval $(0, t)$, plus a proportion per unit time reaching the threshold for the second time, plus a proportion per unit time reaching the threshold for the third time, ..., and so on. Let $h(\bar{N}_e, \vartheta_1)$ be the proportion per unit time described above. It has been shown for a wide variety of neuronal models (Johannesma, 1967; Roy and Smith, 1969) that h is a sigmoidal function of \bar{N}_e for a fixed threshold ϑ_1 . However, by assumption the threshold is itself a randomly varying parameter with the distribution function $G(\vartheta_e)$. It follows that the distribution function for the proportion per unit time of neurons reaching threshold at t is

$$\int_0^\infty h(\bar{N}_e, \vartheta_e) G(\vartheta_e) d\vartheta_e.$$

This expression is certainly monotone increasing with \bar{N}_e . Indeed it will be sigmoidal if $G(\vartheta_e)$ is unimodal. The simplification in the text amounts to the assumption that h equals the Heaviside step function $\mathbb{1}[\bar{N}_e - \vartheta_e]$. This implies the neglect of neuronal activations caused by *fluctuations* about the mean level \bar{N}_e , hence the description of \mathcal{S}_e as the *expected* proportion of neurons reaching threshold per unit time. Similar considerations would obtain if thresholds were to be fixed at a single value, and fluctuations of connectivity, or in the activity itself, were to be considered.

8. § 1.2 p. 60. This correlation is on a neuron by neuron basis. That is, neurons that are strongly excited at t are likely to have just been activated and therefore used at t . Thus the expected number of neurons activated during the interval δt will tend to be smaller than $R_e \cdot \mathcal{S}_e(\bar{N}_e) \delta t$. As a consequence there will be *fluctuations* about the mean activity $E(x, t) \delta t$. However the correlation-time of such fluctuations cannot be greater than the membrane time-constant μ .

since membrane potential decays with such a time course. Indeed the actual correlation-time will be considerably shorter than mainly because of the smoothing effect of the spatial summation of afferent excitation by individual neurons, the occurrence of multiple firings of neurons, and also because of intrinsic fluctuations of neuronal thresholds and similar noisy events. Thus over an aggregate or tissue, correlations between sensitive neurons and excited neurons may be expected to be small and of short duration. Moreover in § 1.3, the time coarse-grained expressions $\langle E \rangle$ and $\langle I \rangle$ are introduced and used throughout the subsequent analysis. $\langle E \rangle$ and $\langle I \rangle$ are obtained from E and I respectively by filtering with a smoothing time-constant equal to μ . Such a filter automatically eliminates all high frequency oscillations greater than μ^{-1} c.p.s., and consequently, the correlations referred to above.

9. § 1.3 p. 60. Recent evidence suggests that Sholl's estimated may be too small (Scheibel and Scheibel, 1970). However it is often the case that longer axon processes are thicker and have higher conduction velocities. Therefore the argument for neglecting conduction terms for lateral interactions remains valid. However the situation may be very different when projections from one anatomical region to another are considered.

10. § 1.4 p. 62. In a few instances the effects of a non-monotone form for $\beta_{ij}(x)$ have been investigated. The function chosen possessed a local minimum at the origin and symmetric maxima located a specified distance to either side. The attempt was to model Cajal's *cellule à double bouquet dendritique*, which has been supposed to be inhibitory (Szentagothai, 1967). Such a choice did not appear to yield any qualitatively new results, however.

11. § 1.4 p. 62. These inequalities do not strictly preclude the possibility of some intermediate spatially uniform state being stable. However, such a state has never been found.

12. § 1.4 p. 63. The point is that different choices of these parameter values can be compensated for by suitable alterations of the values of the other parameters. For example, an increase in the values of r_e and r_i can be compensated by increasing the values of the coupling coefficients b_{ij} .

13. § 2.2 p. 63. Because of the complexity of Eqs. (1.3.1) and (1.3.2) and the number of parameters involved, one cannot be certain that these modes exhaust the possibilities. Furthermore, it has proved to be impossible to determine the dependence of different modes on parameter values. However, it is shown that certain relationships do exist between the various parameters that serve to partly specify the modes (see § 5.1).

14. § 2.2.2 p. 65. L. M. Glass (personal communication) has observed that the Mach band phenomenon (another edge enhancement effect), also has a definite latency associated with its development.

15. § 3.3.4 p. 74. Hysteresis in model neuronal nets had already been demonstrated (Harth *et al.*, 1970). However the phenomena described were purely temporal, and were not analogous to the spatio-temporal hysteresis discovered by Fender and Julesz.

Acknowledgement. We should like to express our gratitude to our colleagues in the Department of Theoretical Biology, The University of Chicago, for many helpful comments, and to the National Institutes of Health and the Alfred P. Sloan Foundation for partial support of the research work reported in this paper.

References

- Allison, A. C.: *Biol. Rev.* **28**, 195 (1953)
- Andersen, P., Andersson, S. A.: *Physiological basis of the alpha rhythm*. New York: Appleton-Century-Crofts 1968
- Andersen, P., Eccles, J. C.: *Nature (London)* **196**, 645 (1962)
- Beurle, R. L.: *Phil. Trans. Roy. Soc. B*, **240**, 55 (1956)
- Beurle, R. L.: *Phil. Trans. Roy. Soc. B*, **240**, 55 (1956)
- Beurle, R. L.: In: Foerster, H. v., Zopf, G. W. (Eds.): *Principles of self-organization*, 291. Pergamon Press 1962
- Bishop, P. O., Coombs, J. S., Henry, G. H.: *J. Physiol.* **219**, 625, 659 (1971)
- Bishop, P. O., Kozak, W., Levick, W. R., Vakkur, G. J.: *J. Physiol. (London)* **163**, 503 (1962)
- Blinkov, S. M., Glezer, I. I.: *The human brain in figures and tables*. New York: Plenum Press 1968
- Brindley, G. S.: *Physiology of the retina and visual pathway*. London: Edward Arnold 1970
- Burns, B. D.: *J. Physiol.* **112**, 156 (1951)
- Burns, B. D.: *The mammalian cerebral cortex*. London: Edward Arnold 1958
- Burns, B. D.: *The uncertain nervous system*. London: Edward Arnold 1968
- Burns, B. D., Heron, W., Pritchard, R.: *J. Neurophysiol.* **25**, 165 (1962)
- Campbell, F. W., Green, D. G.: *J. Physiol.* **181**, 576 (1965)
- Colonnier, M. L.: In: Eccles, J. C. (Ed.): *Brain and conscious experience*. Berlin-Heidelberg-New York: Springer 1965
- Cowan, J. D.: In: Caianiello, E. R. (Ed.): *Neural networks*. Berlin-Heidelberg-New York: Springer 1968
- Cowan, J. D.: In: Gerstenhaber, M. (Ed.): *Mathematical problems in the life sciences*. I. pp. 1—57. Providence, R.I., American Mathematical Society 1970
- Cowan, J. D.: In: Rice, S. A., Freed, K. F., Light, J. C. (Eds.): *Statistical mechanics — new concepts, new problems, new applications*. Chicago: University of Chicago Press 1972
- Cragg, B. G., Temperley, H. N. V.: *Brain*, **78**, 304 (1955)
- Creuzfeldt, O., Ito, M.: *Exp. Brain Res.* **6**, 324 (1968)
- Daniel, P. M., Whitteridge, D.: *J. Physiol.* **159**, 203 (1961)
- Demetrescu, M., Demetrescu, M., Iosif, G.: *Electroenceph. clin. Neurophysiol.* **18**, 1—24 (1965)
- Dewan, E. M.: *J. Theoretical Biology* **7**, 141 (1964)
- Ditchburn, R. W., Ginsborg, B. L.: *Nature*, **170**, 36 (1952)
- Donchin, E.: *Vision Res.* **7**, 79 (1967)
- Eccles, J. C.: *The physiology of synapses*. New York: Academic Press 1964
- Eccles, J. C.: *Epilepsia*, **6**, 89 (1965)
- Farley, B., Clark, W. A.: In: Cherry, C. (Ed.): *Information theory* (p. 242) (Fourth London Symposium). London: Butterworth and Co. 1961
- Fender, D., Julesz, B.: *J. Opt. Soc. Am.* **57**, 819 (1967)
- Freeman, W. J.: *Logistics Review* **3**, 5 (1967)
- Freeman, W. J.: *Math. Biosci.* **2**, 181 (1968a)
- Freeman, W. J.: *J. Neurophysiol.* **31**, 337 (1968b)
- Fuster, J. M., Alexander, G. E.: *Science* **173**, 652 (1971)
- Griffith, J. S.: *Bull. Math. Biophys.* **25**, 111 (1963)
- Griffith, J. S.: *Bull. Math. Biophys.* **27**, 187 (1965)
- Harth, E. M., Csermely, T. J., Beek, B., Lindsay, R. D.: *J. Theor. Biol.* **26**, 93 (1970)
- Hartline, H. K.: *Am. J. Physiol.* **121**, 400 (1938)
- Hartline, H. K., Ratliff, F.: *J. Gen. Physiol.* **41**, 1049 (1958)
- Hebb, D. O.: *The organization of behavior*. New York: John Wiley 1949
- Hollander, H.: *Exp. Brain Res.* **10**, 219 (1970)
- Hubel, D. H., Wiesel, T. N.: *J. Physiol.* **165**, 559 (1963)
- Hubel, D. H., Wiesel, T. N.: *J. Neurophysiol.* **28**, 229 (1965)
- Hubel, D. H., Wiesel, T. N.: *J. Physiol.* **195**, 215 (1968)
- Johannesma, P. I. M.: In: Caianiello, E. R. (Ed.): *Neural networks*, p. 116. Berlin-Heidelberg-New York: Springer 1968

- Julesz, B.: Foundations of cyclopean perception. Chicago: University of Chicago Press 1971
- Kahneman, D.: Quart. J. exp. Psychol. **17**, 308 (1965)
- Kahneman, D.: Psychol. Bull. **70**, 404 (1968)
- Kalil, R.E., Chase, R.: J. Neurophysiol. **33**, 459 (1970)
- Kirkwood, J.G.: J. Chem. Phys. **14**, 180 (1946)
- Kohlers, P.A.: Vision Res. **2**, 277 (1962)
- Kohlers, P.A., Rosner, B.S.: Am. J. Psychol. **73**, 2 (1960)
- LeGrand, Y.: Light, color, and vision. New York: John Wiley 1957
- Lorente de Nó, R.: In: Fulton, J.F. (Ed.): Physiology of the nervous system, p. 288. New York: Oxford University Press 1949
- MacKay, D.M.: In: Schmitt, F.O. (Ed.): Neurosciences research symposium summaries, p. 397. Cambridge, Massachusetts: M.I.T. Press 1970
- Mountcastle, V.B.: J. Neurophysiol. **20**, 408 (1957)
- Oshima, T.: In: Jasper, H. (Ed.): Basic mechanisms of the epilepsies, p. 253. Boston: Little, Brown & Co. 1969
- Poggio, G.F., Viernstein, L.J.: J. Neurophysiol. **27**, 517 (1964)
- Polyak, S.L.: The vertebrate visual system. Chicago: University of Chicago Press 1957
- Purpura, D.R.: In: Schmitt, F.O. (Ed.): The neurosciences: Second study program. New York: Rockefeller University Press 1970
- Rall, W.: J. Cell. Comp. Physiol. **46**, 413 (1955)
- Ratliff, F.: Mach bands. London: Holden-Day 1965
- Roy, B.K., Smith, D.R.: Bull. Math. Biophys. **31**, 341 (1969)
- Rushton, W.H.: Proc. Roy. Soc. London, B, **162**, 20 (1965)
- Sanderson, K.J.: Exp. Brain Res. **13**, 159 (1971)
- Scheibel, M.E., Scheibel, A.B.: In: Schmitt, F.O. (Ed.): The neurosciences — Second study program. New York: Rockefeller University Press 1970
- Sholl, D.A.: The organization of the cerebral cortex. London: Methuen 1956
- Spehlmann, R.: Electroenceph. clin. Neurophysiol. **19**, 560 (1965)
- Sperling, G.: J. Am. Psychol. **83**, 461 (1970)
- Sperling, G.: Perception and Psychophysics **8**, 143 (1970)
- Szentagothai, J.: In: Lissak, K. (Ed.): Recent development of neurobiology in Hungary **1**, 9 (1967)
- Tasaki, I.: In: Field, J., Magoun, H.W., Hall, V.E. (Eds.): Handbook of physiology, Section 1: Neurophysiology, p. 75, 1950
- Uttley, A.M.: Proc. Roy. Soc. B, **144**, 229 (1955)
- von Neumann, J.: In: Shannon, C., McCarthy, J. (Eds.): Automata studies, p. 43. Princeton: Princeton University Press 1956
- Wilson, H.R., Cowan, J.D.: Biophys. J. **12**, 1 (1972)
- Winograd, S., Cowan, J.D.: Reliable computation in the presence of noise. Cambridge, Mass.: M.I.T. Press 1963
- Zusne, L.: Visual perception of form. New York: Academic Press 1971

Dr. Hugh R. Wilson
Dept. of Theoretical
The University of Chicago
Chicago, Ill., USA