Pergamon studies in neuroscience No 11

Series editor Dr William Winlow, *Dept. of Physiology, University of Leeds, LS2 9NQ, UK*

Neuroscience is one of the major growth areas in the biological sciences and draws both techniques and ideas from many other scientific disciplines. *Pergamon studies in neuroscience* presents both monographs and multi-author volumes drawn from the whole range of the subject and brings together the subdisciplines that have arisen from the recent explosive development of the neurosciences.

Pergamon studies in neuroscience includes contributions from molecular and cellular neurobiology, developmental neuroscience (including tissue culture), neural networks and systems research (both clinical and basic) and behavioural neuroscience (including ethology). The series is designed to appeal to research workers in clinical and basic neuroscience, their graduate students and advanced undergraduates with an interest in the subject.

6. **Neurophysiology of ingestion** *ed.* D. A. Booth

7. **A theory of the striatum** *ed.* J. Wickens

8. **Neuroregulatory mechanisms in ageing** *ed.* Maynard H. Makman *and* George B. Stefano

9. **Thalamic networks for relay and modulation** *ed.* Diego Miniacchi

10. **Glycobiology and the brain** *ed.* M. Nicolini *and* P. F. Zatta

Neural Modeling and Neural Networks

edited by

F. Ventriglia

Pergamon Press
Oxford . New York . Seoul . Tokyo

U.K. Pergamon Press Ltd, Headington Hill Hall, Oxford OX3 0BW, England

U.S.A. Pergamon Press, Inc., 660 White Plains Road, Tarrytown, New York 10591-5153, U.S.A.

KOREA Pergamon Press Korea, K.P.O. Box 315, Seoul 110-603, Korea

JAPAN Pergamon Press Japan, Tsunashima Building Annex, 3-20-12 Yushima, Bunkyo-ku, Tokyo 113, Japan

---

Copyright © 1994 Pergamon Press Ltd

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from the publishers.

First edition 1994

Library of Congress Cataloging in Publication Data
Neural modeling and neural networks/ edited by F. Ventriglia.--1st ed.
p. cm.--(pergamon studies in neuroscience)
Includes bibliographical refernces and indexes.
1. Neural networks (Neurobiology) 2. Nervous system--Computer simulation. 3. Neurons--Computer simulation. I. Ventriglia, F. (Francesco) II. Series.
[DNLM: 1. Neural Networks (Computer) 2. Nerve Net. 3. Models, Neurological. WL 102.5 N49236 1993]
QP363.3.N43 1993
612.8'01'13--dc20
DNLM/DLC

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN 0 08 042277 2

Printed in Great Britain by Galliard (Printers) Ltd, Great Yarmouth

Preface

In the last ten years massive research efforts have been made in neural modeling and neural networks. Terms and concepts have been acquired in that time such as learning, memory, perception, recognition etc., which are the basis of neuropsychology. But a great number of neurophysiologists retain the view that much of neural modeling has only a metaphorical value, having no connections with brain activity. Most difficulties in neural modeling arise from an ignorance of what is fundamental and what is not for the expression (and description) of superior neural activity. Still other difficulties result from the fact that the elements used in the modeling could undergo changes, since a complete knowledge of neural structures and functions, both at cellular and at population levels is lacking. Mathematical models of neural systems suffer from these restrictions and, hence, must be able to admit knowledge derived from new experimental results. Models that are too open, however, could be subject to criticism of their scientific value.

The purpose of this book is to describe some important markers in this growing area of research. Towards this aim, different topics on neural structures and functions, techniques of spike train analysis, dynamics of single neurons, oscillations in groups of neurons, randomness and chaos in neural activity, (statistical) dynamics of neural networks, learning, memory and pattern recognition are developed in 12 chapters.

This book is an outgrowth of a two-week-long summer school on "Neural Modeling and Neural Networks" held at Capri (Italy), during September–October 1992, under the Editor's direction. *Neural Modeling and Neural Networks* is a suitable textbook for courses in biological information processing, biological cybernetics and neural networks and could form a source book for brain researchers and modelers. Students need preliminary knowledge of neurophysiology and notions of probability theory, statistics, differential equations and stability analysis.

I thank the Institute of Cybernetics (CNR) [Italian National Research Council], the National Committee for Physics (CNR), the National Group of Cybernetics and Biophysics (CNR), the Naples Research Area (CNR) and the American Society for Mathematical Biology (SMB) for the funding that made the summer school possible.

I would like to thank Dr Antonella Raddi for editorial assistance.

Francesco Ventriglia, Arco Felice (Italy)

List of Contributors

L. F. ABBOTT
Department of Physics and Center for Complex Systems, Brandeis University, Waltham, MA 02254, U.S.A.

M. ABELES
Department of Physiology School of Medicine and The Neural Computation Center, The Hebrew University of Jerusalem, P.O. Box 1172, Jerusalem 91-010, Israel

D. J. AMIT
INFN, Sezione di Roma, Istituto di Fisica, Universita di Roma "La Sapienza", P.le Aldo Moro, Roma, Italy

M. A. ARBIB
Center for Neural Engineering, University of Southern California, Los Angeles, CA 90089-2520, U.S.A.

P. ERDI
Central Research Institute for Physics, Group of Biophysics, PO Box 49, H-1525 Budapest 114, Hungary

B. ERMENTROUT
Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

K. FUKUSHIMA
Department of Biophysical Engineering, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan

W. J. FREEMAN
Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, CA 94720, U.S.A.

J. G. R. JEFFERYS
Department of Physiology and Biophysics, St Mary's Hospital Medical School, Imperial College, London W2 1PG, U.K.

J. SZENTAGOTHAI First Department of Anatomy, Semmelweis University Medical School, Tuzolto u. 58, Budapest IX, H-1450, Hungary

R. D. TRAUB
IBM Research Division, T. J. Watson Research Center, Yorktown Heights, NY 10598 and Department of Neurology, Columbia University, New York, NY 10032, U.S.A.

F. VENTRIGLIA
Neurodynamics Department, Istituto di Cibernetica, C.N.R., Via Toiano 6, 80072 - Arco Felice (NA), Italy

1

Anatomical Bases of Neural Network Modeling

JANOS SZENTÁGOTHAI

First Department of Anatomy, Semmelweis University Medical School, Tuzolto u. 58, Budapest IX, H-1450, Hungary

General introduction

Modern neuroanatomical methods have made it possible to define practically any neuron (1) *Anatomically* (shape and location of the cell body, dendritic arborizations, a considerable part of axonal arborizations, synapses received from and given to other neurons [or other tissue elements]); (2) *Physiologically* (especially spiking activity under suitable experimental conditions, biophysical conditions, prevailing on the outside or the inside of the neuronal membrane, number and location as well as distribution of several ionic channels, etc; the situation is somewhat more difficult in neurons that lack spiking activity and have to be assumed to convey changes of their state electrotonically; the sign of synaptic action, whether excitatory or inhibitory can be defined usually); (3) *Biochemically* (by identification of the mediator, [mediators] through which one neuron can influence other neurons with which they are connected synaptically [in addition to mediators other substances called modulators can also be defined in many cases]). Beyond these classical properties investigation into the nature of many neurons is well on its way to uncover the molecular and genetic mechanisms that are at the bases of the anatomical, physiological and biochemical properties of the neurons.

Considering the extreme complexity of structure and functioning of any piece of neural tissue we have to start with a clear objective, in order not to

loose our way already at the very beginning. The classical approach of the neuroanatomist was and will probably always remain, to give reasonable descriptions of neuron networks. Highly abstract *black box* type models of neural centers have their —rather limited — fields of application, however, this is certainly not the territory of the neuroanatomist. If we were able to build realistic — albeit even radically simplified — models of neuron networks, we might enable students of neural models to try their hand at giving life to such models either by mathematical, general purpose or special hard-wired simulation experiments.

Neuronal connectivity in most neural centers is sufficiently specific to permit disassembly of the whole network into distinct pieces (or units) of characteristic internal connectivity that are arranged into larger structures by repetition of similar architectural units. These units have been termed *neuronal modules* and this architectural principle is referred to by speaking of the *modular architectonic principle* of neural centers. In the following chapters an attempt will be made to illustrate this *modular architectonic principle* as it can be recognized in various parts of the central nervous system (CNS) of the vertebrates, starting with the spinal cord; and the changes that occur in the lower brainstem; subsequently we shall move to the example of the cerebellum, then to the neocortex and eventually — as a special case of the cortex — to its archaic part, the limbic lobe and its highly specialized structure, the hippocampus.

Spinal cord and brainstem

The spinal cord of the vertebrates has two major functions; (a) to integrate neural (reflex and other) functions in one or few neighbouring segments of the body, (b) to serve as a conducting medium along the entire body axis. These two basic functions cannot be separated completely, because functions bridging only minor or even major parts of the body axis have to be separated from the true connections between the higher parts of the CNS and the spinal segmental level as *intersegmental mechanismus*. The brainstem listed in caudo-cranial direction: the *medulla oblongata pons* and *mesencephalon* while preserving the essential architectural principle of the cord, are gradually widened literally by being *blown up* — by new neuron systems, that besides acting as centers for higher coordination of body posture and movement, are serving the increased demand for central structures by the concentration of specific sense organs (equilibrium, audition, vision) on the head and act also as structures connecting with the CNS axis highly specialized centers, like the cerebellum in all vertebrates and *lateral lobes* subserving specific electric senses and lateral line organs that were developed in the fishes, but gradually disappeared in the upper vertebral classes.

The architectonic principle of the segmental apparatus

It has to be born in mind that the segmental arrangement in the vertebrates differs fundamentally from the segmentation in invertebrate phyla where segmentation belongs to the essence of the building plan, whereas in the vertebrates segmentation is superimposed upon an originally continuous neuraxis by the mesodermal segmentation of the body. In spite of this, the secondarily imposed segmentation serves as an important system of landmarks for the description of spinal cord architecture.

Although it may seem difficult to reduce the strange form of the spinal gray matter: two dorsal and two ventral gray columns, connected before and behind the central canal by two gray commissures — into some simplified geometry for a more schematic representation of the spaces occupied by the cell and neuropil material, such a subdivision of the spinal gray matter is indeed possible. The central core of the gray matter can be easily reduced into a *double barrel* structure — resembling the shape of a hunting rifle — and to consider the remaining parts of the dorsal and ventral (in some parts of the cord also into a lateral) appendages attached to the dorsal, lateral and ventrolateral perimeters of the barrels. This architecture was not apparent for the early students of the CNS, because the classical Golgi procedures could be applied with success mainly to very young animals. In these early stages of life the *double barrel* structure of the central gray core does not become apparent. It is only during postnatal development that the neuropil is expanded by ingrowing arborizations of preterminal and terminal arborizations. This became visible only when the application of Golgi type methods by perfusion made the study of adult material possible (Réthelyi, 1976).

The spatial orientation of dendrites and terminal axonal arborizations have been understood already by the authors of the early classical period of neuroanatomy in the years 1888 until shortly after 1900 — notably by Ramón y Cajal (see Fig. 1) — as the most important cue for understanding the synaptic relations in any part of the neuropil. Based upon these criteria the two barrels of the spinal central core can be reduced to a relatively simple geometry of two columns of stacked coins as illustrated in Fig. 2. Both dendrites of the interneurons and axonal arborizations appear to be compressed into flat cylinders from which it is easy to deduce that synapses are established with the largest probability between terminal axons that enter any given flat cylinder and the interneurons whose bodies are lying in and whose dendritic arborizations are confined to the same cylinder. The neuropil of the ventral horn is also restricted — although with much less rigour — to halfmoon-shape extensions in ventrolateral direction of these discs. The motoneurons are not only much larger than the interneurons, but are oriented with the longer axes of their bodies and their dendrites parallel to the spinal cord axis, so that they transgress — and hence may receive synapses from the preterminal axon arborizations within 4–5 neighbouring

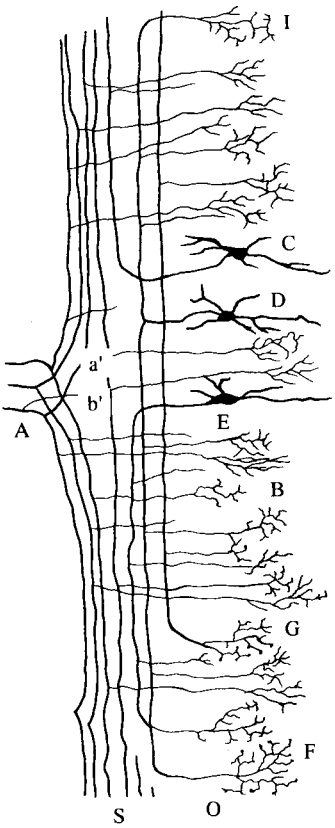


Fig. 1. Drawing of Ramón y Cajal from 1894 illustrating the entrance of dorsal root (A) and its branching into an ascending and descending main branch in the dorsal funicules. Collaterals (a′,b′) are transversally oriented, interneurons C, D, E are giving similarly oriented transversal collaterals. Sagittal longitudinal section.

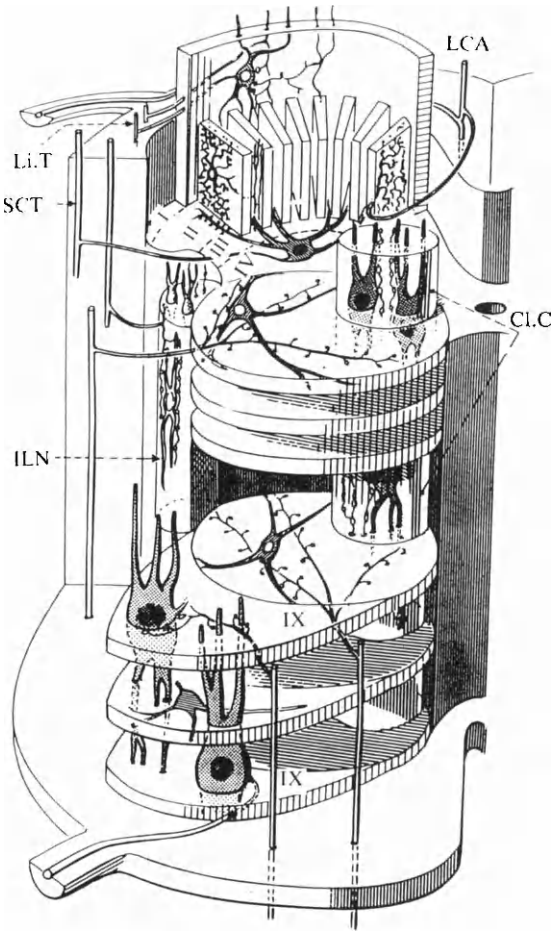


Fig. 2. The spatial orientation of neuropil architecture in the spinal cord in quasi-discrete tissue blocks illustrated in more detail. The neuropil — both dendritic and axonal —is strictly longitudinally oriented in Clarke's column and in the intermedio-lateral nucleus. Li.T., Lissauers tract; SCT, spino-cervical tract; ILN, intermedio-lateral nucleus; LCA, large cutaneous afferents; Cl.C, Clarke's column. Roman numbers indicate laminae of Rexed (Szentágothai and Ráthelyi, 1973).

discs. The neuropil architecture of the dorsal gray column differs very considerably from that of the ventral gray column in being very clearly layered into cell sheets lying parallel to the dorsal surface of the dorsal horn. On this basis the entire spinal gray matter was subdivided by Rexed (1954) into ten layers, generally labeled in dorso-ventral direction by Roman numerals I–X. For practical purposes, especially in physiological experiments, this subdivision is quite useful, although it is neuroanatomically realistic only for laminae I–IV. Lamina I has its neuropil (both dendritic arborizations and terminal axon ramifications) arranged parallel to the dorsal surface of the dorsal column gray matter. Consequently, this layer is crucial for receiving and transmitting the impulses entering the cord via the medium size A_δ fibers. These fibers serve a variety of sensory functions, among else also of fast conducted pain. Conversely the neuropil of laminae II and III is arranged in *lobuli* in radiate direction, i.e. vertical to the dorsal surface of the gray matter. The lobuli are narrower and more numerous in lamina II and somewhat wider (in medio-lateral direction) in lamina III, but the relation between the neuropil of the two layers is not clear, the lobuli become intermixed at the border of the two laminae. There are two specific territories of the spinal gray matter where the neuropil is oriented longitudinally: in Clarke's column and in the so called *intermedio-lateral nucleus*, containing the majority of the sympathetic preganglionic neurons.

Many important functional conclusions can be drawn from this curious combination in spatial orientation of the neuropil, however, going into any detail would require more than the whole space available for the functional anatomy of the entire CNS.

Let me, therefore, enlarge upon a single point here that I believe to be the most important for understanding the essence of the segmental apparatus. Imagine the *double barrel* central core of the spinal cord as two columns of coins one stacked upon the other and the tracts of the white matter —especially the short intersegmental connection located close to the border of the gray matter in the so called lateral and ventral fundamental fasciculi. A pattern of cell connectivity, recognized and illustrated first by Ramón y Cajal (1894, in Fig. 1.) shows the principle — here for the case of the dorsal root fibers that after entering the cord bifurcate into an ascending and a descending branch from which only the main ascending branch serves direct forward conduction toward the dorsal column nuclei of the medulla oblongata (also indirectly to the cerebellum and the ventrobasal nuclear complex of the thalamus). The descending branch and the initial part of the ascending branch gives rise to horizontally (transversally to the spinal cord axis) oriented collaterals that arborize in the flat transversally oriented discs of the central core and contact either interneurons or also directly motoneurons. These are mainly serving local segmental reflexes. This principle is illustrated for the central core and the ventral horn in Fig. 3 (from Szentagothái, 1981), which neglects — for the sake of simplicity — the dorsal root collaterals entirely. As seen from the lower — longitu-

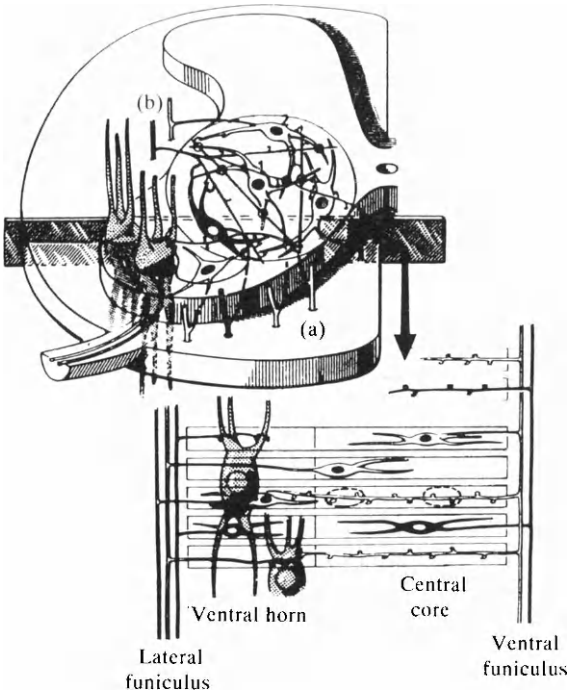


Fig. 3. *Stacked chips* architecture principle of the intermediate zone (center core) of the spinal cord, in the form of flat discs (represented as a circle in the upper transverse section diagram and as brick-shape compartments in the lower longitudinal section diagram). Assumed excitatory interneurons are indicated in outlines, inhibitory ones in full black; motoneurons are stippled. Note the straight courses of interneuron axon collaterals penetrating through the flat neuropil discs, by which they may establish synaptic contacts with any element encountered on their way. Certain potential contacts are considered *forbidden* by some *mismatch* between the respective elements (indicated by small circles), e.g. between excitatory interneuron axon *a* and cell numbered 1, or inhibitory interneuron axon *b* with cell numbered 2.

dinally cut — part of the diagram the principle shown by Ramón y Cajal for the dorsal root fiber holds also for the axons of local interneurons that enter into and bifurcate in the lateral and ventral (white) funiculi and give transversally oriented preterminal branches confined to the mentioned flat discs. Although this diagram is radically and selectively simplified, it may convey the basic principle of the segmental design. It can be compared — very roughly — to the *gear shift* of an old fashioned mechanical calculator, or a lathe (or an early automobile model) in which cogwheels of various sizes are mounted on a couple of parallel axes and the gear is shifted by moving to and from the axes and bringing together cogwheels of different sizes. (For a model builder of today the architectonic principle may be more aptly compared with *stacked electronic chips*, essentially with the same principle of functions on a level more sophisticated by orders of magnitude.)

Looking at the architecture of the spinal segmental apparatus is by no means idle speculation. It has been shown in embryonic transplantation and tissue recombination experiments (Székely and Szentágothai, 1962; Székely and Czéh, 1971) that fragments of the early medullary tube, lacking the dorsal horn and any sensory input, may generate spontaneously (i.e. lacking in any input from without) *walking-like* movements in supernumerary limbs implanted nearby. Provided, of course, that the structure illustrated in Fig. 3 — and only that — remained intact and that the medullary tube fragment was taken from spinal cord segments designed for innervating limbs.

This is, of course, beside the point for our present considerations, however, such experiments do show that nature-like functions (for example walking-type movements of a limb) can be generated by the CNS spontaneously (i.e. without any input) if the connectivity and character (excitatory and/or inhibitory) of the interneurons was preserved in the form as it would develop under natural circumstances. Such experiments show additionally, that neural functions at their very bases (and earliest stages) are spontaneously generated — i.e. *autopoietic* — probably by chaotic (random) activities of the neurons (known to exist in all neural tissues) and self-organized into activity cycles (Szentágothai, 1987; Szentágothai and Érdi, 1989; Szentágothai, in press), provided again that the connectivity and character of neurons remained as occurring in normal embryonic development.

The lower brainstem

The lower brainstem (medulla oblongata, pons and midbrain) is essentially arranged in the same quasi-segmental (transversally oriented disc) pattern, as first recognized and most elegantly illustrated by Scheibel and Scheibel (1958); in fact it were the Golgi studies of these authors in which this fundamental (stacked discs) principle of the entire neuraxis was first recognized, although the consequences could then not be stated explicitly.

An entirely new principle — this time of chemo-architectonic nature — enters here the segmental principle. Groups of specific monoaminergic neurons are built into the brainstem, however, only the serotoninergic neurons do really fit into the segmental *stacked chips* architecture, while catecholaminergic and dopaminergic neurons more often than not are concentrated in separate cell groups (nuclei; for example in the locus ceruleus) that belong to the new structures appearing in the brainstem. The specific sensory and motor (also vegetative) nuclei of the cranial nerves are subdivided into separate entities in the brainstem, but the continuity of sensory and motor *columns* can still be recognized and traced upwards until the mesencephalon (see a relatively recent description by Székely, 1989).

A fundamental change of the design occurred with the appearance, in the lower brainstem, of larger nuclear masses connecting the longitudinal arrangement of the neuraxis with higher centers that have developed — already beginning with very early stages of the vertebrate phylogeny — for auxiliary, although very highly specialized and essential functions, notably: the cerebellum and the tectum. (The mesencephalic tectum has become reduced in the latter stages of phylogeny so that in the mammalian brain practically only remnants of the optic [in amphibia, reptiles and birds] and the acoustic lobes [in some genera of all classes of vertebrates with highly organized sense of hearing] in the earlier phylogenic history of the brain.) The two main nuclear formations for the connection with these higher (auxiliary) centers are the inferior olive of the medulla oblongata and the nuclei of the pons. This is a crude oversimplification, because the pathways feeding into the cerebellum arise also from the spinal cord itself (spino-cerebellar tracts, spino-cervico-cerebellar tracts and from the medulla oblongata (dorsal column nuclei)).

From these nuclei the inferior olive deserves our special attention, because it is this nucleus that supplies the most specific afferent system of the cerebellum, the so called climbing fibers. While the termination of all other afferent systems of the cerebellar cortex (and nuclei) is effected by the mossy fibers of the cerebellar cortex (and nuclei) that terminate in the cerebellar cortex in a more diffuse manner. The climbing fibers are highly selective in contacting locally only one single Purkinje cell and with synaptic systems that, if activated, are insured for an effective transmission. In addition, the olivary nuclei receive massive afferent input both from ascending spinal and from descending pathways. This does not usually occur with the pontine nuclei although it occurs with the dorsal column nuclei.

The upper brainstem

The upper diencephalic and telencephalic (striatum and putamen) part of the brainstem does not retain anything resembling the quasi-segmental arrangement of the lower neuraxis. There is, however, a small part of the

diencephalon — the so called hypothalamus (i.e. the ventralmost part of the diencephalon) — in which elements of the basic architectural principle of the neuraxis are preserved longitudinally oriented fiber tracts and transversally oriented (coronal) quasi-discs of neuropil (Makara *et al.*, 1980; see also a more recent summary by Szentágothai, 1983). The longitudinal fiber systems are concentrated in a medial periventricular and a more lateral (but still medial in its name) so called *medial forebrain bundle.* Although the transversal (here in the brain already *coronal*) orientation of the neuropil can still be detected — but this applies less to the cells that are generally organized in nuclei, than to the terminal axon arborizations. But these very vaguely recognizable architectural properties are of little use for the model-builder, because connectivity — although important —ranges behind a very sophisticated system of chemical information processing. Transmission is effected in this part of the brain by a system of over twenty (as known today) neuropeptide transmitters and modulators that, in addition, do not act exclusively or even predominantly synaptically, so that the classical neuroanatomical approach becomes secondary to considerations based on peptide biochemistry and its background of molecular biology and genetics.

Complex synaptic systems

Space (and time) limitations forbid us to discuss the thalamic system proper (and the sensory relay nuclei showing similar peculiarities) in any detail, apart from mentioning an entirely new somatotopic architecture principle discovered only recently by Kievit and Kuypers (1977). The overall somatotopic relation of gross frontal (coronal) discs of the cerebral cortex and the striatum to the sagittal, in the anterior direction diverging discs of the thalamus, cannot be discussed mainly because its consequences have not yet been elaborated upon, or has any general functional consideration been attached to this curious, and probably very meaningful relation.

There is, however, one aspect of the thalamic system that has not received, so far, any legitimate physiological explanation, in spite of gallant efforts first by Rakic (1975) and later in more elaborate form by Schmitt, Dev and Smith (1976) under the name of *local circuit neurons.* Local neurons, having only relatively short axonal arborizations have been known already to the early classical neurohistologists and have been labeled as Golgi II type neurons (or simply Golgi type neurons in contrast to the classical neuron type having one relatively long axon, first described by Deiters in the 1850s and hence called [in the German literature] Deiters type neurons or more generally Golgi type I neurons), because the long axon of most neurons became clearly recognizable only in the classical studies of Camillo Golgi (1873).

The existence of complex synaptic systems (so called glomeruli) were described already by Ramón y Cajal in his classical studies, but it was not before analysis under the electron microscope became possible that the real structural significance (because a convincing physiological explanation is still lacking) and the wide distribution over the entire nervous system (especially in lower sensory mechanisms) could be appreciated.

Under the term synaptic glomeruli the classical neurohistologists understood the combination of several synaptic structures in what appeared to be *islands* of the neuropil (in the cerebellar cortex, the olfactory bulb, some of the anterior thalamic nuclei). Electron microscopic studies did show much later — beginning with the mid-1960s — that such neuropil islands may be surrounded and almost completely enclosed by glial envelopes, in which case they were labeled *encapsulated zones* , or may lie virtually free in the general neuropil. Such complex synaptic arrangements may not even be recognizable as specific islands. What is standard in the complex synaptic arrangements are so called *synaptic triads*, i.e. at least three synaptically connected (terminal) elements combined in characteristic arrangements. Although found in an infinite member of combinations, the most general (or basic) type of these triads contains (i) a main input element, excitatory by structural and transmitter biochemistry standards, that is presynaptic to (ii) a dendrite (appendage) of a forward conducting (projective) cell and also presynaptic to (iii) a *presynaptic* dendrite element (usually appendage) of a local — generally Golgi type II — cell that is inhibitory both by transmitter biochemical (GABAergic) and structural standards (flattened or at least not clearly spherical) synaptic vesicles and giving symmetric synapses (i.e. lacking in postynaptic densities). Element (iii) is always presynaptic (again based on structural criteria) to (ii). These three synaptic contacts are rarely more distant from one another than a couple of microns. It should be emphasized though that this are only the most common combination of synapses, there are many exceptions and additional complications that will not be discussed further at this point.

Although the role attributed to these synaptic arrangements led to vastly more ambitious hypotheses by Rakic (1975) and by Schmitt, Dev and Smith (1976), this author favors more parsimonious explanations offered by the theoretical studies based on logic networks by Lábos (1977, 1981; see also an attempt to express them in realistic anatomical terms by Szentágothai, 1983). Most recently, it appears that the possible functional significance of the triadic synaptic arrangements has become more meaningful in a new approach by Roska and Hámori (in press).

The cerebellar system

The cerebellar cortex has been the most favored anatomical structure since the classical study of Ramón y Cajal (1888) of the bird cerebellar

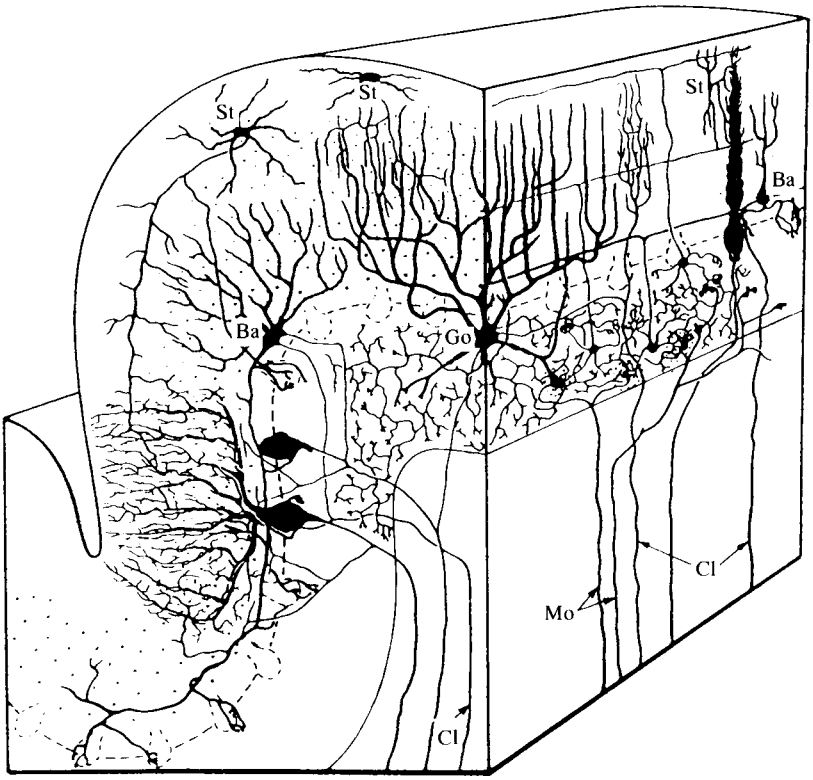


Fig. 4. Stereodiagrammatic illustration of a tissue block taken from a cerebellar folium. The section surface at left is a transversal, at right a longitudinal section through the folium axis. B, basket cells; St, stellate cells; Go, Golgi cells; Cl, climbing fibers; Mo, mossy fibers contacting granule cells in the granular layer below the row of Purkinje cells, which in turn give rise to the parallel axons running parallel to the folium axis (from Eccles *et al.*, 1967).

cortex. This was due primarily to the rectangular lattice — quasi crystalline — structure of the most important synaptic stratum, the molecular layer and later in electron microscopy due to the clear stratification of the cortex into very distinct layers of the gray cortex: the (outermost) molecular layer, a single layer of the main output cells the (ganglionic) layer of the Purkinje cells and an inner layer of the granule cells. This enabled the investigator to see immediately exactly where he was in any electron microscopical visual field.

The characteristic rectangular lattice structure of the cerebellar cortex is shown in Fig. 4 reproducing one of the conventional text-book diagrams, that needs no further comments beyond the figure explanation. Once the existence of specific inhibitory interneurons became known, it was inevitable that students of the cerebellar cortex began speculating about possible operational mechanisms of such neuron networks. One of our early specula-

tions (Szentágothai, 1963, 1965) caught the interest of Sir John Eccles and since the origin of the climbing fibers in the inferior olive was already known at that time, Sir John Eccles tested the basic assumption of our speculations and found it to work in physiological experiments, we soon were able to develop a coherent hypothesis of the operational features of the cerebellar cortex neuron network (Eccles *et al.*, 1967). It soon turned out that this speculation, although essentially correct for the cerebellar cortex itself, did not give us any real insight into the functions of the cerebellum simply because the cerebellar reflex arc into which the longitudinal paths both descending from the cerebral cortex and ascending from the spinal cord and the more complex sense organs are derouted (into a quasi *set-aside* center) is closed not in the cerebellar cortex, but in the cerebellar nuclei. The cerebellar cortex serves only as a very sophisticated neuron network — extended over an extremely large two-dimensional, but intricately folded sheet of neural tissue, that is, as if it were superimposed upon the basic nuclear reflex level in order to supply it with a fine grain (somewhat delayed) control system. This cortical control system is peculiar (1) by having little, if any association in space (i.e. along the two dimensions of the enormous surface) (2) by containing an enormous number of nerve cells (the granule cells, contributing probably to the vast majority of the nerve cells in the body of the higher mammals * [and cetaceans] and hence giving the system an extremely high degree of redundancy) (3) by exercising its influence upon the nuclear level exclusively by inhibition: the Purkinje cells being inhibitory. Although spatial association is insignificant, this lack may be well compensated by association using relatively high speed (reverberating) temporal mechanisms (Pellionisz, 1984).
1

However this may be, the cerebellar system can be understood more completely by two diagrams of Ito (1984), shown with some modifications in Fig. 5. The simpler diagram A shows the elementary circuit, while B tries to convey an impression of the more complex real neuron circuit. The neuron network of the cerebellum is a Garden of Eden for the model builder working on a basic level — from molecular neuroanatomy, over neuron circuit models, well into the realm of more ambitious holistic models attempting to simulate the acquisition of motor skills as it occurs in the postnatal development of many animals (the few exceptions, like for example in the guinea-pig, which is born as miniature adult, are suited to caution against rush generalizations) and finally the almost mysterious ability of the human to develop non-biological motor skills (as seen beginning with the survival technologies of so called primitive societies, over various forms of artistic expression, well into professionally overbred sports and acrobatics and — most importantly — in articulated speech).

[1] *In humans one may calculate with roughly 10^{10} cells in the cerebral cortex, perhaps 1.5×10^{10} for everything else (including spinal cord, brainstem as well as sensory and vegetative ganglia) and at least $3–5 \times 10^{10}$ for cerebellar granule cells.

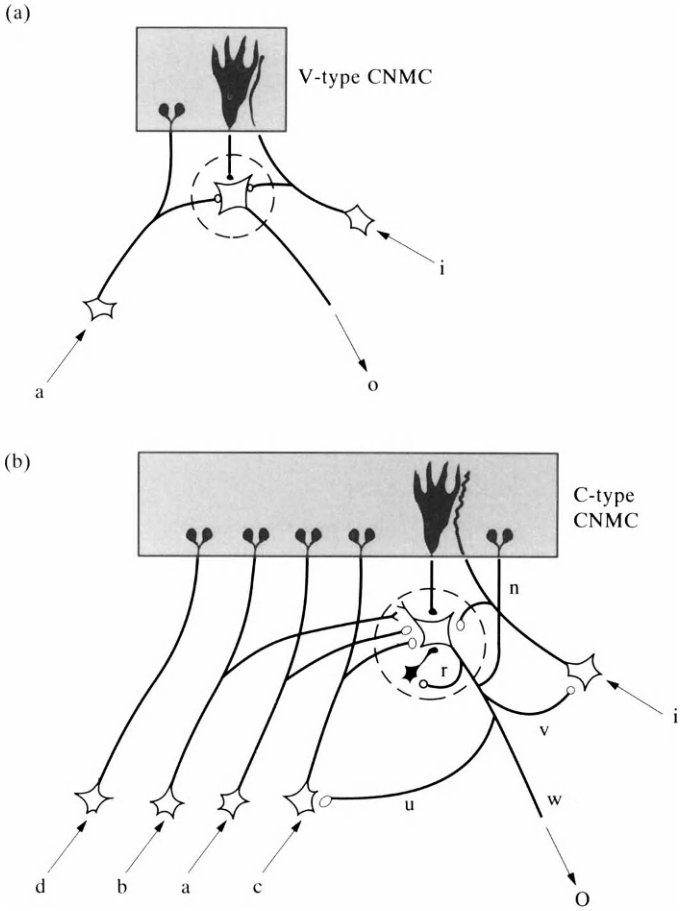


Fig. 5. (A) Simple representation. (B) More realistic complex relations between cerebellar nuclear cells (in dashed circles) and the cerebellar cortex (stippled). This diagram is preliminary.

Cerebral cortex

For a quick look through the light microscope at a histological specimen stained for cells, the only apparent —but undoubtedly striking — feature of the cerebral cortex is its clear separation, in most brain regions, into six distinct surface parallel layers some of which in certain regions of the brain can be subdivided again into sublayers, already in the *cytoarchitectonic* picture. Corresponding, but different layers can be seen in stains for myelin sheaths of the intracortical fibers this is called conventionally *myeloarchitectonics.* A corresponding architectonic picture can be observed in cortex specimens stained for lipids, called *lipidarchitectonics.* For the time being, only the cytoarchitectonic picture is really useful for practical work. The Golgi picture, showing various cell types with their very characteristic dendritic arborizations and the arborization of incoming or local axons seems to make things more obscure rather than contributing to our understanding. With the help of modern neural tracing methods and particularly with the reconstruction of the local arborizations (both dendritic and axonal) of single cells impaled with microelectrodes and studied in physiological experiments and eventually labeling the neurons with the injection of either horse-radish-peroxydase or phaseolus lectin we have been able recently — over the last 15 years — to gain a real insight into the nature of many of the hitherto known neuron types, their anatomical and synaptic connections, both locally and at longer distances and have come closer to a real understanding of the several mediators (modulators, multiple mediators) involved in synaptic transmission.

Figure 6 conveys the state of knowledge at the end of the 1970s, when the concept of the cortical module, an architectural unit, a cylinder of some 200–300 μm diameter and about 3 mm height — i.e. extending vertically through the entire depth of the cortex — became first solidified, although the present level of characterization of various cell-types was still in its beginnings. This figure deals with the module based primarily on cortico-cortical connectivity, probably the soundest basis yet available. The major cylinder is therefore built around an arborization of a cortico-cortical afferent. The concept of the modular architectonics principle of the cerebral cortex arose originally from the early physiological observation of Mountcastle (1957) of a vertical columnar organization principle of the somatosensory cortex, soon followed by an analogous architectural principle in the visual cortex found by Hubel and Wiesel (1959). Anatomical studies — in continuation of the classical observations of Ramón y Cajal (1899) and the then revolutionary insight of Lorente de Nó (1938) on the predominantly vertical orientation of neuron chains — were taken up by this author (Szentágothai, 1961) by using an entirely new approach of chronically isolated cortical tissue slabs and were made available to the Western readership only in 1965 by Colonnier. The first anatomically based neuron circuit model (Szentágothai, 1969) was still very much under the influence of the

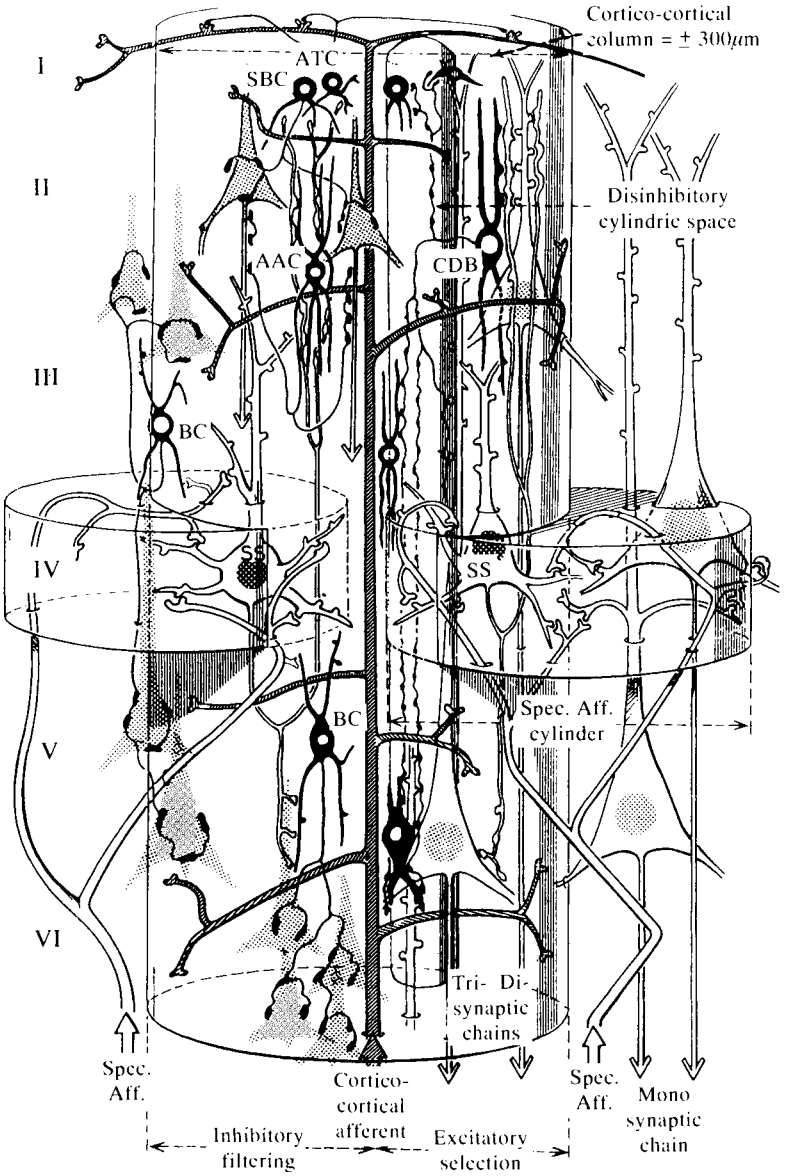


Fig. 6. Relatively early stereodiagrammatic view from Szentágothai (1978): Layering in Roman numerals at left margin. BC, basket cells; SS, spiny stellate cells; ATC, axonal tuft cell; SBC, small basket cell; AAC, axo-axonic cell; CDB, cellule double bouquet. Excitatory cells and fibers in outlines, inhibitory cells in full black.

earlier circuit model of the cerebellar cortex. It is not as if the early specific sensory cortex model had to be abandoned. This is incorporated by the two flat cylinders, located at both sides of Fig. 6. in lamina IV so that the cortico-cortical module model is simply a more realistic application of the model based on specific sensory afferents to a more general architectonic principle holding true in the several regions of the neocortex.

Recently the whole concept of the module principle has been questioned by Swindale (1989), arguing that continuity in surface parallel direction would be more in accordance with reality. However, no sooner than this attack on the concept of modularity became known, it was shown by Peters and Sethares (1991) that instead of the several earlier forms of relatively simple modules as could be built around the arborizations of the specific sensory afferents or cortico-cortical connections, there is an entire hierarchy of my highly regular modularities within the primary visual cortex of the monkey. The architecture of the visual cortex of the primates is more sophisticated in reality than could have ever been imagined earlier. A hierarchy of at least three different types of regularities (modules) built into each other in a real *crystalline structure,* no less regular in structure than that of naturally grown crystals. Figure 7 makes the duality of continuity and of modularity (i.e. discontinuity) intelligible. This diagram tries to show the various ranges within which neurons are connected with each other over distances of a couple of millimeters, i.e. still locally. Simultaneously, the local arborizations of various types of individual neurons that have been drawn to scale, to indicate how the several types of architecture units (modules) fit together and that there is no conflict whatever between continuity and discontinuity.

The dentate gyrus–hippocampus system

The hippocampal system, including its *propylea* the dentate gyrus (or fascia dentata) are the most important formations of the archicortex both from the functional and anatomic point of view. It seems virtually certain by now that its main physiological function is in being instrumental in memory functions. Not in the sense that it would directly contain the traces — if memory were based on gathering and encoding traces (earlier called engrammes) of past experience which is probably not the case — but in issuing the order for any functional event to be imprinted upon (or fixed in) some other part or parts of the brain. However important, these questions are definitely beyond the frame of this chapter. Although clearly a cortical structure, the hippocampal formation shares with the cerebellar cortex the crucial advantages that its layering separates cells and neuropil, with an additional advantage that even the neuropil and fiber strata radiating into or out from the hippocampal formations are separated into various distinct layers or sheets of fibers. This makes the entire hippocampal formation an

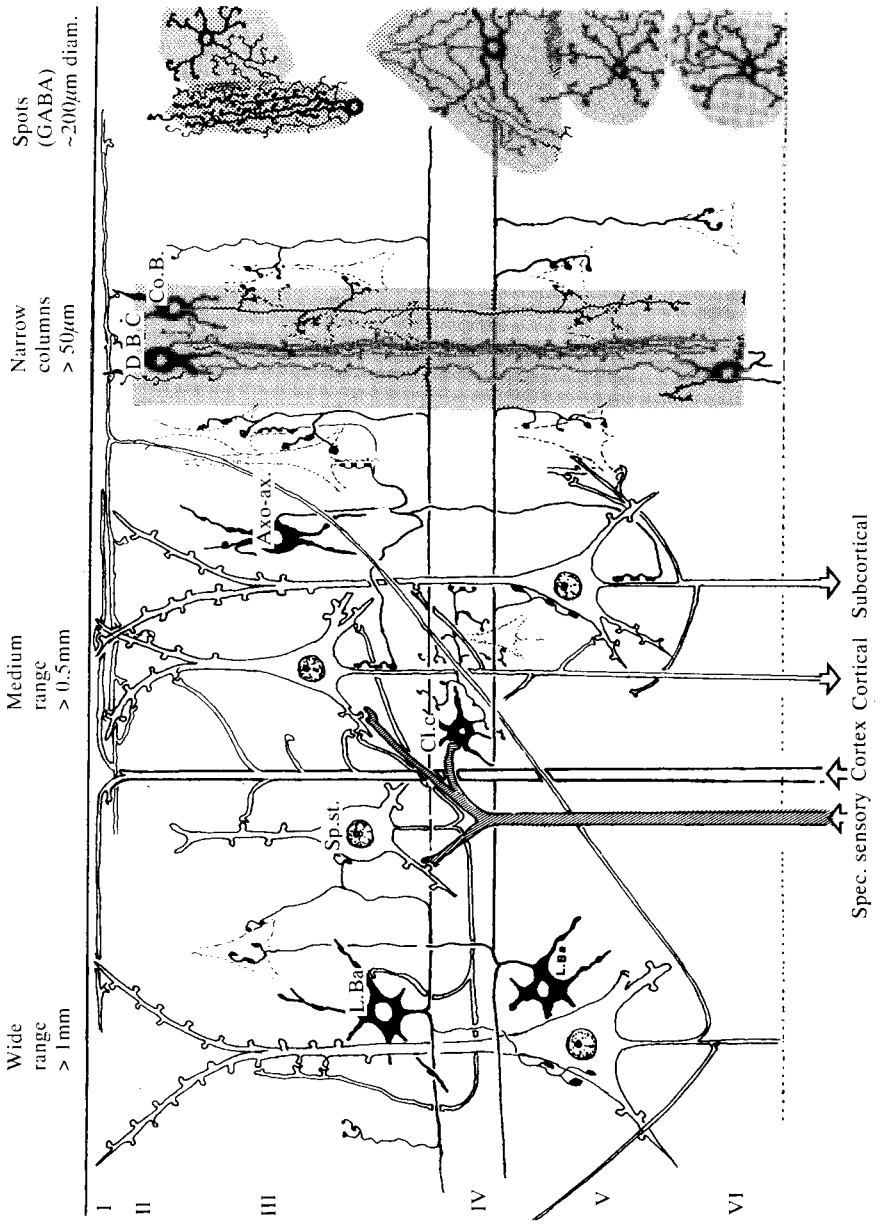


Fig. 7. Recent composite diagram of Szentágothai (in press) of the ranges of local intracortical connections. LB, large basket cells; Cl, Clutch cell; CDB, cellule double bouquet of Cajal; CoB, columnar basket cell; Mar, Martinotti type cell. Stippling (at right) shows the extension of the axonal arbors of small inhibitory interneurons. All elements are drawn approximately to scale.

exceptionally favored ground of combined anatomical, physiological and immunohistological analysis.

The essential features of the hippocampal formations were already understood from the Golgi studies of the classical period as magnificently summarized by the beautiful drawing of Ramón y Cajal (1911) and reproduced here in Fig. 8. The figure and its legend are largely self-explanatory. The input channels were not sufficiently understood at the time, although the two main fiber systems, the so called perforant path, arising in the entorhinal cortex — entering the dentate gyrus through its convexity — is correctly indicated. What was thought to be and indeed is the main output line of the hippocampal formation, the fornix, supplies the other main input path both of the dentate gyrus and the hippocampus proper with strong afferent systems from both sides of the septum. Hence, coming from the hippocampal commissure, the path is also labeled commissural path. What was not known was the sophisticated system of short GABAergic (hence inhibitory) interneurons. Only the basket cells, with their axonal terminations directed preferentially to the cell bodies of the pyramidal cells of the hippocampus proper and of the dentate gyrus granule cells were known to early authors. Even so the neuron chains: entorhinal cortex — perforant path — dentate gyrus granule cells — mossy fibers —pyramidal cells of the proximal C_{A3} (adjacent to the dentate gyrus) — Schaffer collaterals — pyramidal cells of the distal C_{A2} and C_{A1} regions of the hippocampus proper — and, eventually, the pyramidal cell main axons entering the fornix were already well known and correctly illustrated in the classical figure (Fig. 8) of Ramón y Cajal.

It was much later that the axo-axonic interneuron was discovered first in the hippocampus proper and later also in the dentate gyrus. But it took still more time and a considerable refinement of the techniques, before additional interneuron types were recognized and it took new immunocytochemical methods before the crucial importance of the inhibitory interneurons in the regulation of hippocampal neuronal operations began to dawn upon us.

Most of this material from Péter Somogyi's and Tamás Freund's laboratories have not yet been published, so only a few preliminary results can be discussed in more detail. Some of the newly discovered interneurons are stimulated or inhibited directly over longer fiber tracts of the perforant path and of the commissural (septal) system. In this case inhibitions of the granule cells of the dentate gyrus or of the pyramidal cells of the hippocampus is a *feed-forward* inhibition (or disinhibition if, as is the case in several systems, the long afferents [input fibers] are themselves of inhibitory function). Conversely, if the inhibitory interneurons are primarily stimulated by the collaterals of dentate gyrus granule cell collaterals, or by the Schaffer collaterals of the hippocampal pyramid cells (both being excitatory, disinhibition cannot be envisaged) the inhibition over the interneurons is *feedback* in nature.

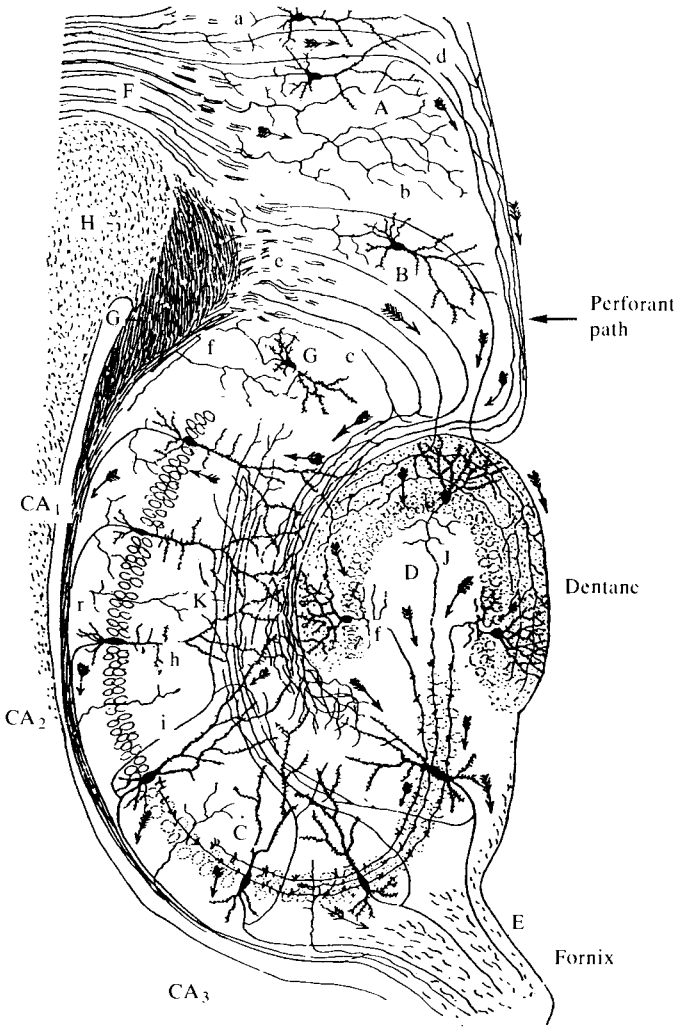


Fig. 8. Hippocampal-dentate formation from a drawing of Ramón y Cajal (1911). A, Entorhinal cortex; B, cells of the subiculum giving rise to perforant path fibers; D, hylus of the dentate gyrus with the origin of the mossy fibers; (j), arising from a single but multicellular layer of the dentate granule cells; i, Schaffer collateral; K, stratum of the Schaffer collaterals; C, row of hippocampus pyramidal cells (the several parts of which [C_{A1}, C_{A2}, C_{A3ab} and C_{A3c}]) are indicated around the perimeter of the hippocampus.

The great advantage in studying the hippocampal formation, apart from gaining deeper insight into the possible neural mechanisms involved in memory and hopefully also into the pathological mechanisms of epileptic seizures, is often the possibility to understand neuron coupling mechanisms in the neocortex. The analyses possible in the hippocampal formations need further development leading to an improved insight into the intricacies of the complex network of the many types of inhibitory interneurons in the neocortex.

2

Models of Visuomotor Coordination in Frog and Monkey

MICHAEL A. ARBIB

Center for Neural Engineering, University of Southern California, Los Angeles, CA 90089-2520, U.S.A.

Approach and avoidance in frog and toad

To simplify rather drastically, we may say that the ability to find food and escape enemies in the frog can be reduced to the ability to tell small moving objects from large moving objects.

A frog surrounded by dead flies will starve to death, but the frog will snap with equal "enthusiasm" at a moving fly or a pencil tip wiggled in a fly-like way. On the other hand, a larger moving object can trigger an escape reaction. Thus, at a very simple level, we might imagine that the brain of the toad has two basic pattern-recognition routines (what we shall call *perceptual schema*), one for recognizing small moving objects (food-like stimuli) and one for recognizing large moving objects (enemy-like stimuli). We could then come up with a very simple model in which we have the signals from the eye routed to these two perceptual schemata. If the small-*moving-object* schema is activated, it will in turn trigger the *motor schema* (our term for an automaton or control system for controlling action) to get the animal to *approach* what is apparently its prey; while if the perceptual schema for large-moving-object is activated, it will trigger the motor schema for *avoidance*, causing the animal to escape an apparent enemy.

We now consider how such a model can become a *biological model* as distinct from a purely functional model. Our assertion will be that a schema-based model (we turn to neural network models in a later section) becomes

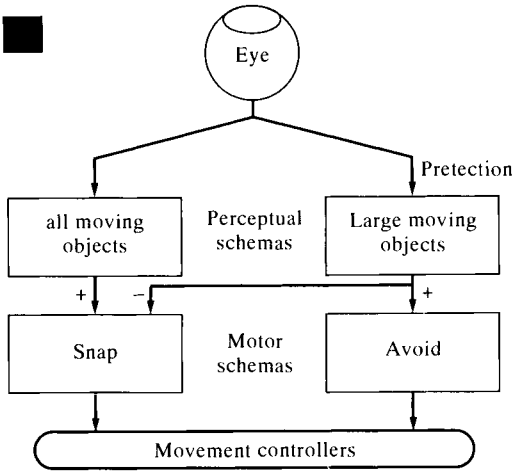


Fig. 1. A schema program for approach and avoidance that takes into account data on the effect of lesioning the pretectum.

a biological model when explicit hypotheses are offered as to how the constituent schemata are played over particular regions of the brain. Each eye of the frog projects to the opposite half of the brain, especially to the important visual midbrain region called the *tectum.* This projection from the *retina* to a layered structure of the brain preserves the neighborhood relationships of the visual field — it is called a *retinotopic map.* Another retinotopic map goes to the *pretectum* (so called because it is in front of the tectum).

Returning to the schemata mentioned above, experimentalists had suggested that the small-moving-object schema was in the tectum, while the large moving objects schema was in the pretectum. Such hypotheses about neural localization may be tested by *lesion experiments.* The above model predicts that animals with a pretectal lesion would continue to approach small moving objects just as the normal animal would, but would not respond at all to large moving objects. However, the model is false! Peter Ewert studied toads (Ewert, 1987) in which the pretectum had been lesioned. He found that not only did the toads respond to small moving objects with approach behavior, but they also responded to large moving objects with approach behavior! This observation leads to the new schema-level model shown in Fig. 1. The new data tell us that in the absence of the pretectum, the animal must be able to respond to all moving objects with approach. Thus, we replace the perceptual schema for *small* moving objects by a perceptual schema for *all* moving objects. On the other hand, in the normal animal, we have that *recognition* of large moving objects triggers avoidance and so we leave the right hand column the way it was. How-

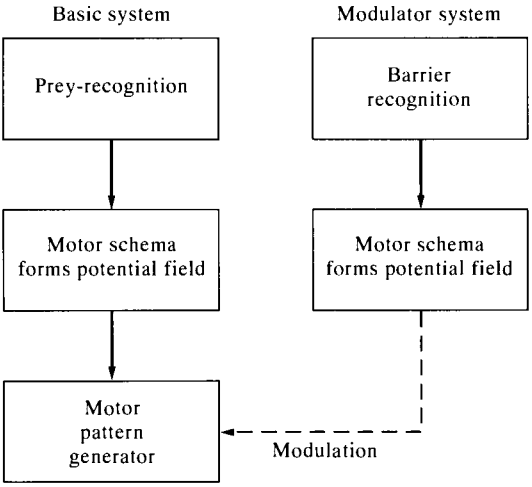


Fig. 2. The evolutionary refinement of adding detour behavior to prey-approach by modulating the basic system for approaching prey.

ever, although we have now explained the response of the lesioned animal to all moving objects and the response of the normal animal to large moving objects, it remains to tune the model so that the normal animal will respond to small moving objects with approach but not avoidance. This we can achieve by having an inhibitory pathway running from the perceptual schema for large moving objects (in the pretectum) to the approach schema — or, equivalently, to the schema for all moving objects. With this model we can now explain our small database on the behavior of both normal animals and those with a lesion of the pretectum. We have thus established that hypothesis about neural localization of subschemata may be tested by lesion experiments. As we shall see below, we may then proceed further to model a brain region to see if its neural circuitry will indeed implement the positive schemata. The important point is that biological models can be expressed at the level of a network of interacting "automata", irrespective of whether or not data are available about the fine-grain implementation of those automata in neural networks.

Detours and path planning

Our task now is to see how brain function can be analyzed in a process of evolutionary refinement in which basic systems serve as the substrate for the evolution of more refined systems. The motivation for this work is the behavior (Ingle, 1976) of a frog observing a worm through a semi-transparent grating or barrier. Instead of launching directly at its prey as would occur if no barrier is present, the frog reacts appropriately, detouring

around the barrier (as we see in progress in the fourth frame) to get its prey. This behavior already extends the schema picture of Fig. 1. Now, the perceptual schema for recognizing prey must be augmented by the perceptual schema for recognizing a barrier and there can no longer be a simple direct path from prey recognition to the triggering of approach behavior. Rather, there must be some way for this path to be modulated by the recognition of the barrier to yield an indirect detour, rather than the direct response.

In the particular situation under consideration here, the animal must not only recognize *prey and barrier* , but must locate them in space. If it can recognize that the prey is in front of the barrier or at most a tongue's length behind, then the animal will indeed snap directly. But if the prey is further behind the barrier, then the animal must use its recognition of where the prey is and where the barrier is to come up with a path which will carry it around the barrier towards the prey. Arbib and House (1987) offered a model of this in which perceptual schemata for prey and barrier can drive motor schemata which compete and cooperate to yield the overall behavior. We postulate that the ground plane in front of the animal is represented in the brain, with the worm being represented as a *global attractor*: each point in the "arena" has a vector pointing towards the worm, with the vectors decreasing in length with distance from the prey, but not vanishing. On the other hand, each fence post is represented as a *local repellor*, such that animal will be repelled either to left or right if it comes close to the fence post, but will not be affected if it is further away. When we combine the action of the individual fence posts, we get a strong vector field to the left of the post and a strong vector field to the right. If we combine all this activity we get a summed field which can be integrated to yield trajectories which either pass to the left or to the right of the fence and then continue *en route* to the prey.

It is not our claim here that the *brainstem* of the frog implements the above potential field algorithm in its neural circuitry. Rather, the crucial point is that we have an evolutionary account of how such a system might arise (Fig. 2). First, the elements of the prey-recognition system — perceptual schema, motor schema and *motor pattern generator* (MPG) — co-evolve so that activity in the prey-schema can represent a goal in such a way that the approach-schema provides the right control signals for the MPG to determine a path to the prey; then the detour system evolves by combining a perceptual schema for stationary objects with a motor schema that *modulates* the effect that the approach-schema has on the motor pattern generator.

To close this section, we emphasize not so much the particular implementation of the schemata, but rather the evolving subtlety of the schema interactions. We start with two basic systems for response to small and large moving objects, respectively, and then find that if we are to match the biological strategy we must come to the more subtle interactions shown in

Fig. 1, in which recognition of small moving objects is not in fact localized in any one region, but is rather a system property involving the *modulation* of the tectum by the pretectum. We then extend the complexity of the environment to which the animal responds — it no longer contains a single prey or a single predator to which the animal may respond with the most basic forms of the "survival behaviors" of feeding or fleeing, but now contains the more subtle structure of obstacles which can block the animal's path and around which the animal must now be equipped to detour. In evolutionary terms, this corresponds to expansion of the ecological niche in which the animal is well suited to survive.

New schemata often arise as "modulators" of existing schemata, rather than as new systems with independent functional roles. Further examples of this may be seen in the lesion-based analysis of schemata for approach and avoidance behavior (Cobas and Arbib, 1991). Here it simply suffices to note that the strategy shown here for the analysis for biological systems may be traced back to the work of Braitenberg (1984) and Walter (1953) (see also the evolutionary approach to the design of an optic flow algorithm (Arbib, 1981)) and has a number of points in common with the approach to the design of robot controllers offered by Brooks (1986).

Neural mechanism of avoidance behavior

We now turn to neural modeling, "filling in the schemata", using as our example the escape behaviors of frogs (Ingle, 1976). The escape direction in response to the approach of a large moving object may be briefly characterized as a compromise between the forward direction of the animal and the direction immediately away from the *looming stimulus*. Barriers can modify *avoidance* behavior, just as they modify approach behavior. If we interpose a barrier to block the preferred direction of escape for a stimulus coming from a particular direction, then the behavior of the animal changes and it tends to jump just to the left or just to the right of the barrier. In the rest of the section, we will not return to the issue of how barriers modify escape behavior, but will instead look in more detail at how the looming stimulus is recognized and how this recognition is transformed into action.

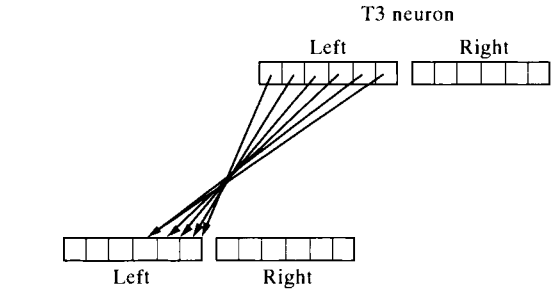
Before doing so, we look at more recent experimental data (Ingle and Hoff, 1990) which shows that the behavior is actually more subtle. We have seen that, if a stimulus is looming directly at the frog, its escape direction will be a compromise between the direction away from the stimulus and the forward direction of the frog. However, if, instead of moving directly towards the animal, the stimulus is on a trajectory which will carry it in front of the animal, the animal will not respond based solely on the position of the looming stimulus, for in this case the animal would jump on a collision path with the *predator*, thus making it highly likely to be captured. Instead, the animal exhibits a highly adaptive "cutback" behavior which carries it on a

path that is well away from that being pursued by the looming object. Here again we note an evolutionary refinement going from simple recognition of a predator to recognition of the trajectory that the predator is following to find an escape direction based on this extra information. Clearly, it is an interesting question to understand how evolutionary pressure could yield an animal able to exhibit this more subtle repertoire. However, in this section our concern is with the neural mechanisms which provide the frog with this functional repertoire. Unlike the previous sections in which we have simply sought a specification of schemata (whether or not analyzed in terms of specific brain regions) to provide a functional analysis of the behavior, we now look at the availability of neurophysiological data.

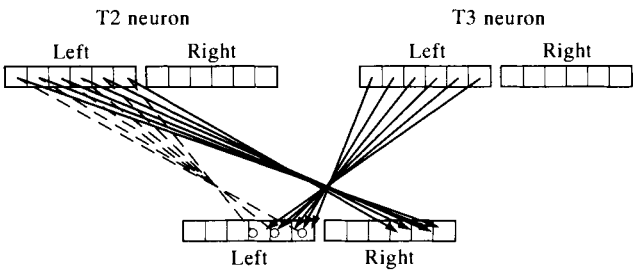
Back in 1959 (Lettvin *et al.*, 1959), the *ganglion cells* of the frog retina, those which send their axons back to the brain, were classified into four different classes and these four classes of cells were found to project to different depths in the *tectum*, forming four specific *retinotopic maps*. Even more excitingly, these maps appeared to be tied to the behavioral repertoire of the animal. If we call the four classes R1, R2, R3 and R4 (remember each one is a spatially arrayed population of cells in the "output layer" of the retina) we find that the R3 cells seem to respond best to small moving objects — as if they were bug detectors — whereas the R4 cells seemed to respond best to the large moving objects — as if they were enemy detectors.

Fortunately or unfortunately, research over the subsequent years has shown that the story is not so simple and that the frog really does need a brain (!) and not just a retina to determine whether to feed or flee. For example, it has been shown that the activity of the R3 cells is quite complicated.

To a first approximation it can be viewed as responding to the *leading edge* of a moving object entering its *receptive field* (the region of the visual field in which activity can affect the activity of the cell), rather than the presence of a prey-like object within the receptive field. Again (and this was already recognized by Lettvin *et al.*, 1959) the activity of the R4 cell can be better interpreted as just a measure of *dimming* — but clearly a predator by casting a larger shadow will dim the receptive field more than prey and thus excite the R4 cell more strongly. Many studies of visual processing or visually guided behavior ignore the particular transformations conducted by the retina, or simply reduce them to a contrast enhancement difference-of-Gaussians lateral inhibition mask. However, we (Liaw and Arbib, 1991) have taken as our starting point the properties of the R3 and R4 neurons. Recall that the R3 and R4 neurons form an array which stretches across the output layer of the retina. If we consider the effect of a dark looming stimulus, we will then see greatest activity amongst those R3 cells whose receptive fields include the leading edge or expanding boundary of that looming stimulus, whereas the R4 cells which will respond most strongly are those contained within the interior of that expanding pattern. We can then combine these cells to provide a model of the T3 cells (so called because they are type 3 among cells of the tectum as characterized by their



(b) Motor heading map



(b) Motor heading map

Fig. 3. Gating of the tectal projection onto motor heading map. Only half of the projections are shown here for simplicity. The T3 neurons project to the heading map to indicate the escape trajectory for a looming object on a colliding trajectory. When the stimulus is crossing the visual field, the T2 signal blocks the T3 signal while exciting the contralateral heading map, thus resulting in a "cut-back" jump.

physiological response). The T3 cell will respond more and more strongly as a stimulus looms to a position at the center of its receptor field if the connections from R3 cells to the T3 cell are radially symmetric but with a sort of inverted Gaussian form in which activity towards the periphery is more effective than activity at the center. The larger the looming stimulus, the further out the pattern of R3 activity and thus the stronger the input to the T3 cell. Since the T3 cells also form a retinotopic array, many cells will be activated by this looming stimulus but the T3 cell whose receptive field is centered on the center of the looming stimulus will have the strongest response. To complete the design of the T3 cell and to make it responsive to a large looming stimulus, but not to several small objects flying apart from each other or to an expanding ring, we give the R4 cells a standard Gaussian projection to the T3 cells so that darkness at the center of the receptive field of the T3 cell will increase its response.

With this we have given an analysis that (although we have not shown all the details here) does not simply explain how the presence of a *looming stimulus* can be represented by a peak of activity in an array of neurons, but also uses circuitry with cells whose firing rates provide a good model of firing rates actually observed neurophysiologically.

With this we have the neural network implementation of the perceptual schema required to explain the avoidance of objects coming directly at the animal, where the current *position* of the looming stimulus on the retina implies the preferred direction of escape. However, to provide the necessary perceptual schema for the cutback response, we must come up with cells that recognize the *direction of motion* across the *retina.* It turns out that the T2 neurons, also in *tectum,* do have this sensitivity. However, we have no information as to how these cells might actually be wired up and so we resort to the standard model of directional selectivity to ensure that the passage of the stimulus from left to right will increase the likelihood of a T2 neuron firing, whereas a pattern moving in the opposite direction on the T3s will yield little or no T2 response.

This provides all the perceptual information we need to complete our model. In Fig. 3(a) we see the *retinotopic map* of T3 neurons (simplified from a 2-dimensional to a 1-dimensional retina for ease of comprehension here) with an appropriate projection pathway from neurons in the T3 array to neurons in what we call the motor heading map which will cause the animal to turn towards the retinotopically corresponding location. Unlike the projection from prey recognition neurons to the motor heading map, in which retinotopically corresponding points are linked by the projection, we now make the connections so that each peak of activity on the T3 layer will have a peak of activity centered at that point which is a compromise between the forward direction and the direction away from the looming stimulus to control the motor pattern generators to escape a directly looming threat.

We show how to model the cutback data Fig. 3(b). First, we add the

T2 array which again is retinotopic but is now signalling not simply that there is a looming stimulus at that point of the retina but there is a stimulus moving from left to right. (We will not look at the corresponding connections for a stimulus on the right half of the visual field moving to the left — this will simply be the mirror image.) This activity must block the normal escape response and so we see a projection from cells of the T2 array to the motor heading map which is a replica of the projection from the T3 array to the motor heading map — but this new projection is *inhibitory.* Thus, if there is no T2 activity the system will respond as in the Fig. 3(a), but not if there is T2 activity. To complete this extended model, we must now add a new excitatory pathway from the T2 neurons which project to the cutback direction and thus can trigger the appropriate motor behavior in this case, while avoiding the normal escape behavior. What is perhaps worth stressing here is that the T3 neurons serve as the basis for the "evolution" a more "sophisticated" set of neurons, the T2 neurons which not only signal the presence of a looming stimulus but the presence of a looming stimulus moving on a particular type of trajectory. This new system then *modulates* T3 activity by projecting to the same *motor heading map* and competing there with the T3 input, rather than acting upon the T3 array itself. This is another instance of schema evolution and modulation. We refer the reader to Liaw and Arbib (1991) for further details of the neurophysiological data which grounds the model, for further details of the circuitry involved in the model and for a number of analyses of the simulations that we have conducted. Here, we close by noting a general property of vertebrate neural control: we see that there may be no place prior to the motor pattern generator where the different representation of the animal's situation are brought together. In general, a multiplicity of different representations must be linked into an integrated whole. However, this may be mediated by distributed processes of competition and cooperation. There is no one place in the brain where an integrated representation of space play the sole executive role in linking perception of the current environment to action.

Control of eye movements

When a limb moves, it needs a burst of agonist contraction to accelerate the limb in the desired direction, followed by an appropriately timed antagonist burst to decelerate it to rest at the desired position (with a possible small agonist correction thereafter); a new resting level of muscle contraction holds the limb in its new position. By contrast, the eye has little inertia and so no antagonist burst is required and the eye has no changing load to require feedback.

Robinson (1964) showed that a *saccade* involved an initial pulse of force to drive the eye to its new position (of fixed amplitude, but increasing in duration with the size of the saccade), followed by a maintained force

to hold the eye in its new position. It was later found (Robinson, 1970) that, when the gaze was fixed, a typical oculomotor neuron would fire increasingly for greater deviations of the eye in the direction controlled by contraction of the muscle it innervated (its "on direction"), with a cut-off at some angle θ_o about 15° into the off-direction, providing co-contraction in forward gaze. The relation is approximately linear, with firing rate

$MN(t) = K[\theta - \theta_o]^+$, where $[x]^+ = \max(x, 0)$, i.e. the motoneuron fires only if the angle θ of gaze exceeds the cut-off θ_o and the firing rate $MN(t)$ increases linearly thereafter. K varies from neuron to neuron. In contrast to this position-dependent resting level, about 5 ms before a saccade in its on-direction, $MN(t)$ jumps to a high frequency which is maintained until about 10 ms before the end of the saccade, when it drops to the resting level appropriate to its new θ. During a saccade in the off-direction, $MN(t)$ is 0. According to Optican and Robinson (1980), the *cerebellum* contains the adaptive circuitry required to adjust the components of $MN(t)$ to ensure proper coordination.

With this, we now turn to a brief review of the neurophysiology of the circuits controlling the *ocular motoneurons* for eye movements in the horizontal plane (Fuchs *et al.*, 1985). The excitatory burst neurons (EBNs), located in the PPRF (*paramedian pontine reticular formation*), are thought to drive the burst of activity in ipsilateral motoneurons. The inhibitory burst neurons (IBNs), located in the *dorsomedial medullary reticular formation*, inhibit the contralateral motor neurons, thus seeming to control the pause in motoneuron firing during movements in the off direction. Where the burst neurons seem involved in the control of the saccade, the tonic neurons (TNs) exhibit regular firing at a rate that is related (in a nonlinear way) to eye position during and after saccades. TN firing does vary with eye velocity in smooth pursuit of a target, as distinct from a rapid saccade to a target.

The gain of burst cells is very large, to keep the duration of saccades as small as possible by generating a high velocity movement when any appreciable motor error exists. In fact, the gain is so high that the model would be unstable — save that the pulse-generators are suppressed during fixation. This is done by the *omnipause neurons* (OPNs) which fire continuously except that they pause during saccades. There is strong evidence that pause cells inhibit burst cells and prevent them from generating a saccade until released. Another group of burst cells exhibits roughly the same discharge for all but the smallest ipsilateral saccades. These are thought to play a role in triggering a saccade and are thus called trigger cells (TRIGs).

These data led to the following hypotheses about *saccade* control: When the omnipause neurons fire (i.e. between saccades), they inhibit the burster neurons (both EBNs and IBNs), thus blocking the burst drive to the motoneurons. For a saccade, the trigger cells silence the omnipause neurons; while the excitatory burst neurons receive a signal to fire and thus raise the firing rate of ipsilateral motoneurons — and via the IBNs, silence the

contralateral motor neurons — for the duration appropriate to the length of the saccade.

Robinson (1975) suggested that the tonic neurons (TN) integrate the firing of the EBNs, $TN(t) = k \int t\ EBN(t)\ dt$, to obtain the desired eye position to maintain the gaze after the saccade has occurred.

It is possible to trigger saccades by stimulation of the *superior colliculus* (the mammalian homologous of the tectum), with the length and direction of the saccade encoded retinotopically on the collicular surface (with further subtleties to be explored below). Based on the anatomy of collicular connections, including intracellular fills in the cat, it appears that colliculus synapses on EBNs as well as long-lead burst neurons (LLBNs), i.e. burst neurons which fire for some time before the actual saccade-correlated burst, in the PPRF. This leads to the hypothesis that these cells may mediate between this topographic map and the temporal encoding of saccades in the EBNs. These neurons may also mediate the ability of frontal eye fields to control saccades, even in an animal whose superior colliculus has been lesioned. But how is the transformation from a spatial to a temporal code brought about? It might be that saccade-related neurons that lie more peripherally and thus encode a larger movement, synapse more strongly on the LLBNs — but one must still explain how greater excitation yields increase in the duration, rather than firing rate, of burster response.

A crucial idea in neurophysiology is efference copy or corollary discharge. When we move our eyes normally the world appears stationary, but when we press against the eyeball with a finger and jiggle it, the world appears to move up and down. Helmholtz suggested that in the former case, but not the latter, the neural circuitry that commands the eye movement also sends a copy of this efference (thus, efference copy) to the perceptual system, which can thus distinguish the self-induced movements from world-induced movements. van Gisbergen *et al.* (1981) model the saccade generator by a servomechanism in which desired eye position (E_d) has subtracted an efference copy/corollary discharge signal (E') replicating eye position (supplied by a neural integrator NI — cf. the role of the tonic neurons as hypothesized above — and not available in the anesthetized animal) to yield a motor error signal e_m which serves as the input to the burst generator B. It should be noted that this tonic neuron recurrent collateral is only a diagrammatic way of indicating that integration occurs. There is no anatomical or physiological evidence for such a construct. Cannon and Robinson (1985) offer a plausible neural-network model of the neural integrator, but the model has not yet been verified by experimental data.

Burst neurons B send their burst directly to the motoneurons (MN) and also indirectly through an integrator path (T, the neural integrator NI) which has a recurrent collateral for the "positive feedback" which provides the integration. B is also coupled in a loop of mutual inhibition with the omnipause neurons (P). Between saccades, firing of P inhibits activity in B absolutely.

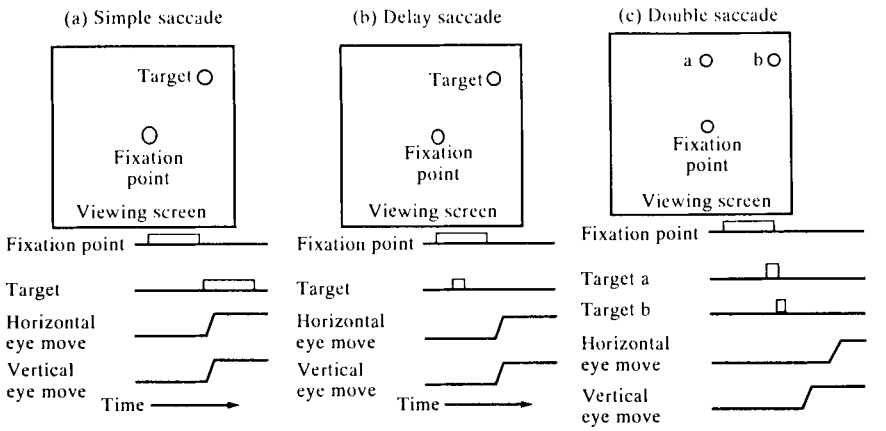


Fig. 4. Saccade task protocols. Time runs from left to right. Traces indicate time course of presentation of fixation point, target(s) and eye movements.

However, trigger input can inhibit P, thus freeing B from the tonic inhibition. And once B is activated, it can "latch" P off for the duration of the saccade. More formally, $P > 0$ unless either trig > 0 (this is the input that triggers a saccade) or latch > 0 (i.e. a saccade is "in progress" and e_m has not yet reached 0). Although this is not quite the formulation in their model, let us introduce the function h such that $h(x) = k$ (the burst frequency) for x greater than some threshold e, while $h(x)$ is otherwise 0. Then the firing of the burster B is described by $B(t) = h(e_m) = h(E_d - E')$ if $P = 0$ but $B(t) = 0$ if $P > 0$.

If e_m differs from 0, B drives MN, charges up NI and "latches" P off, via the latch neuron B1. The saccade ends when e_m is reduced to 0 and the inhibitory bias from the pause neurons is restored. E_d (and thus e_m) plays no further role until another saccade is triggered. During a saccade, the MN output exhibits a pulse step in its discharge, $R(t)$, caused by the transient burst $B(t)$ riding atop the eye position signal E'. The eye is held in position by the output from the neural integrator T to the motor neuron MN.

Schemata for the analysis of saccadic eye movements

The *superior colliculus* (SC) is the primary recipient of the projection from the retina to the *midbrain* and it has long been known that this region plays an important role in the control of saccadic eye movements. SC is a layered structure, with activity of cells in the superficial layers of the SC indicating visual activity but with cells in the deep layers activated only for the target of the upcoming saccade.

Both maps are *retinotopic* and in register — i.e. a hypothetical coordinate grid drawn on the retina may be related to a grid on SC such that, given a point on the retina, a vertical penetration at the corresponding point of the SC will yield cells in the upper layers which are activated by stimulation of the retinal locus and cells in deep SC whose firing correlates with shift of gaze to a target in that retinal direction. Note that the SC commands an eye movement *relative to the current direction of foveal gaze*, rather than giving commands in a head-centered or body-centered frame.

Deeper layers of monkey superior colliculus contain cells that fire before saccades. Each such neuron fires before a range of saccades, so a large population fires before a particular *saccade*. Identical discharges by a given cell can precede a variety of saccades, but Sparks and Jay (1985) state that "it is the location of the active neurons within the topographical map of movement fields, not their frequency of firing, that specifies the trajectory of the saccade." However, since they note that "neurons in the center of the [active] population fire earlier and more vigorously than surrounding cells," it seems appropriate to qualify their assertion, suggesting that some form of distributed coding may apply. Specifically, it is probably only the "center of gravity" of active neurons which is important in coding saccadic size and direction, though current evidence suggests that the burst frequency of the ensemble of active collicular neurons affects saccade velocity.

There are superior colliculus neurons that burst prior to saccades in response to visual targets that also burst before saccades in response to auditory targets. Moreover, with each movement of the eyes in the orbit, the population of neurons responsive to a stationary auditory stimulus changes to a new location within the superior colliculus — the auditory input gets remapped into retinotopic coordinates for the current eye position, i.e. coordinates which represent the eye movement required for fixation. Somatosensory input is also mapped into superior colliculus. Note that these are very different inputs, since the somatosensory input depends on the position of body and limbs, while position in auditory input is time-coded via the difference of arrival of signals at the two ears. Mapping the input into visual coordinates thus seems complex in the former case, comparatively simple in the latter.

The colliculus is intimately related with the pulse generators: the projection from superior colliculus to long lead bursters is monosynaptic, while the latency from deep superior colliculus activity to motoneurons is polysynaptic.

An especially interesting property of the maps in superior colliculus is that existing visual targets in superficial layers get remapped to deep layers when the eye moves. Sparks and Mays (1980), using trials in which an intervening saccade changed the position of the eyes after a brief visual target had been extinguished, discovered quasi-visual (QV) cells the location of whose activity, even if the eyes had moved after the target disappeared, represented the current retinotopic position of the remembered target.

Scudder (1988) developed an alternative model for the *saccade* generator which addresses the issue of spatial coding in superior colliculus. (The model may be seen as a refinement of that of Jürgens *et al.*, 1981.) The output of the superior colliculus is integrated by a neural integrator composed of long lead burster neurons (LLBN) — the integration is shown schematically by self-excitation with unit weight. The retinotopic code of saccade size in superior colliculus is converted to an intensity code by differentially weighting the output of collicular neurons according to the eccentricity of each superior colliculus neuron's movement field. The first event is the onset of the superior colliculus discharge which simultaneously begins to charge the LLBN/integrator and to inhibit the OPN by a long latency indirect pathway. The firing rate envelope of the superior colliculus neurons is a simple Gaussian having a standard deviation of 15 ms. As excitation from the LLBN/integrator discharge increases and the inhibition from the OPN decreases, the net EBN input goes from inhibitory to excitatory. As the EBN fires, the IBN fires and further inhibits the OPN, causing a regenerative increase in EBN and IBN discharge rate. The firing of the IBN also inhibits the LLBN/integrator, causing it to begin to be discharged. Progressively, the LLBN/integrator firing rate decreases until the inhibition of the OPN by the IBN no longer exceeds the excitatory bias of the OPN. The OPN begins firing and the burst ends. The model matches the physiological data for saccades up to 15°. Scudder suggests a number of ways the model might be extended to better fit the data.

We now turn to a *multi-schema* view of the saccadic eye control system. Once I have established what those schemata are, I will provide a brief pointer to work which begins to replace the constituent schemata by neurophysiologically plausible networks localized in different regions of the brain.

In the *simple saccade* experiment, the monkey views a fixation point which is turned off at the same moment that a new target light comes on and the monkey is trained to saccade immediately to the target when it appears. Fig. 4 contrasts this type of experiment with two new experiments. In the *delay memory saccade* paradigm (Hikosaka and Wurtz, 1983) the target is flashed briefly while the fixation light remains on and the monkey is instructed to saccade at offset of the fixation point. Clearly, this task requires a spatial memory to maintain the target position after it has been turned off, providing a *working memory* which will remain active until the saccade occurs. In the third experiment, the *double saccade* task (Mays and Sparks, 1980), two targets are briefly flashed at the offset of the fixation point, first A and then B, but so quickly that both had been turned off before the first saccade begins. Nonetheless, the monkey will successfully make a double saccade, first to A and then to B. What makes the situation so interesting is that Mays and Sparks were able to monitor what they call *quasi-visual (QV) cells* in the deep SC which were located not in the region of the SC corresponding to where stimulus B *had* appeared, but rather to where it

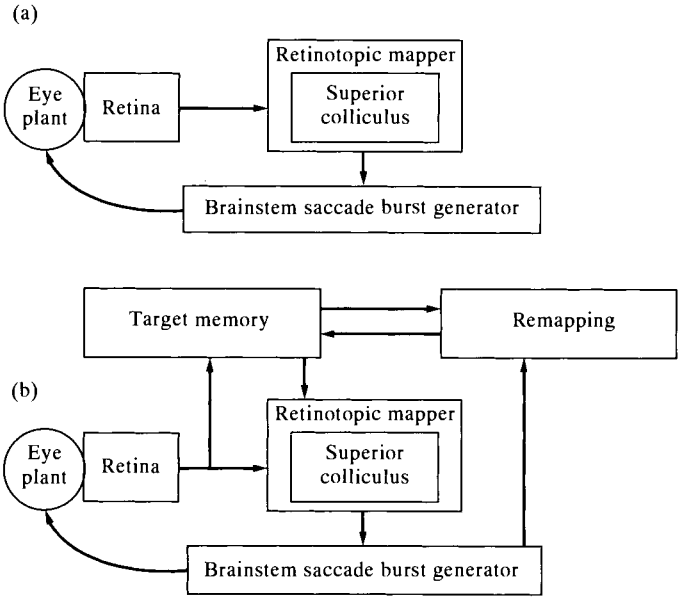


Fig. 5. (a) The basic model for control of simple saccades has two schemata, the Retinotopic Mapper and the Saccade Burst Generator (SG), but it also embodies the structural hypotheses that the Retinotopic Mapper is implemented by the superior colliculus, while the Saccade Burst Generator is implemented by the brainstem. The input to SG codes target position by the position of a peak of activity in a neural map, rather than by a single neural variable. (b) An expanded schema model adds two new schemata, the Target Memory schema and the Remapping schema, to accommodate data on delay saccades and memory saccades.

would appear were it still visible after the saccade to A had occurred. In other words, not only do we need a working memory for B for the second saccade, but we require a *dynamic remapping* to shift the representation of B in a direction equal and opposite to the intervening saccade so that it will now be able to control the appropriate eye movement via the QV cells of the SC.

With this we come to the schema level view of the saccadic system provided in Fig. 5. Figure 5(a) shows the original model in which the brainstem saccade burst generator (essentially the control system modeled by van Gisbergen *et al.*, 1981) is controlled by the retinotopic map maintained in the SC (cf. Scudder, 1988). In Fig. 5(b) we have offered one possible schematization of how this model must be extended. In this case, the retinotopic map of the SC obtains its input not only from the retina but also from target memory. Further, the brainstem saccade burst generator not only controls the eye movements, but also sends an efference copy (corollary discharge) to the Remapping schema which can adjust the target memory appropri-

ately each time an eye movement takes place.

We have devoted considerable efforts to replacing these schemata with neural networks which embody a great deal of neurophysiological and neuroanatomical data about the monkey brain and its performance of saccadic eye movements (Dominey and Arbib, 1992). It would burden the present paper unduly to go into details, but it is perhaps worth a brief tour of the structure of the model, as shown in Fig. 6, to indicate something of what happens in the give-and-take between the general functional understanding provided by schema analysis and the demands and the lacunae of the neurophysiological data. Our first question will be where in the brain can neurons be found whose firing could be seen to represent the different signals involved in the schema theoretic model. We find that QV cells occur not only in SC but also in *posterior parietal cortex* (PP) and in the *frontal eye-fields* (FEF). We also find that the repertoire of saccadic eye movements we have described can occur after the lesion of the SC, mediated instead via FEF control of the brainstem saccade generator. We are thus led to see the QV activity in SC as derivative rather than central.

For the moment, we have chosen to implement a neural network for dynamic remapping in PP and have the results of this remapping imposed on FEF and from there to SC. The interaction between FEF and *thalamus* then provides a reverberatory implementation of the target memory. In the neural network, as distinct from the schema level analysis, we have been led to introduce a subsystem whose role was not explicitly called for in the schema level analysis, namely the *caudate nucleus* (CD) and the *substantia nigra pas reticulata* (SNr) of the basal ganglia. This is because data on the monkey (Hikosaka and Wurtz, 1983a,b) and the rat (Chevalier *et al.*, 1985) have shown that these systems can provide an important inhibitory role, blocking the release of saccades by the superior colliculus. The detailed analysis of the neurophysiology leads to a further refinement of the schema model, as we seek to understand this "unexpected" role. The claim that it is "unexpected" is somewhat disingenuous, since the same papers that reported this important role of the basal ganglia were the experiments which led us to analyze the memory saccade.

Thus, the involvement of the basal ganglia in the multiregion model of the saccades was always an important consideration. Nonetheless, I have attempted to provide a conceptual path, rather than an historically accurate chronology of the research. The real point that emerges from this admission is that the analysis of the brain in fact is neither purely top–down (from function to neural network) nor purely bottom–up (from neural or synaptic components up towards function), but rather requires a continual "middle-out" dialogue.

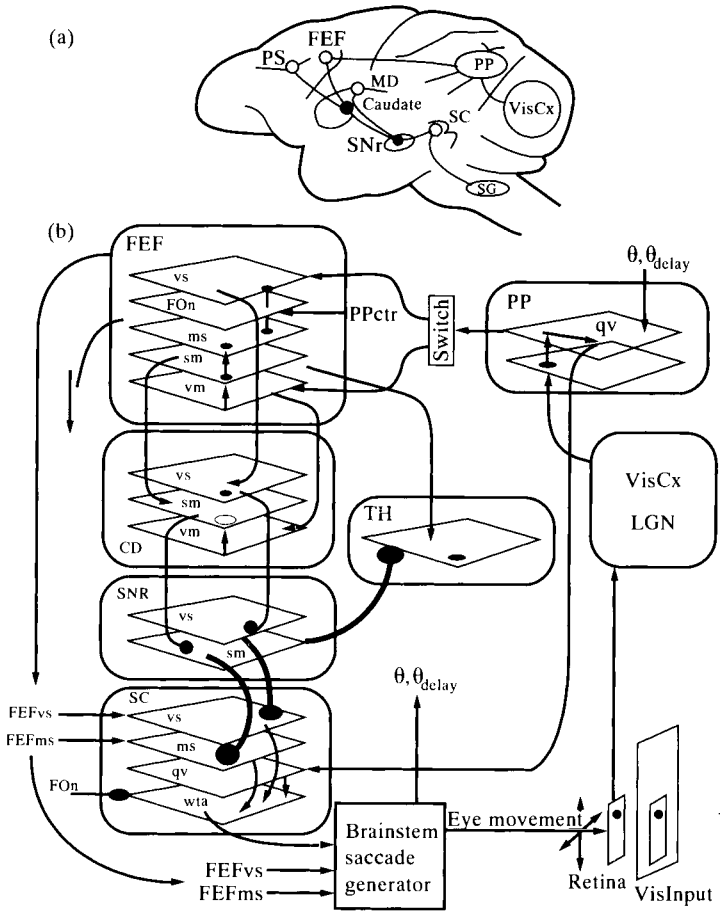


Fig. 6. An expansion of the schema model of Fig. 5(b) to show the posited role of frontal eye fields (FED), mediodorsolateral (MD) thalamus and posterior parietal cortex (PP). In neural network simulations, we expands the model to include the role of basal ganglia and then represents each brain region by one or more layers of physiologically identified cell types. The whole network is implemented in our Neural Simulation Language NSL (Weitzenfeld, 1991) and yields physiologically plausible traces of cell activity and saccadic response. See Dominey and Arbib (1992) for details.

The emergence of space and time

To conclude this short exposition, I want to point out the neural representations of Space and Time that our studies have revealed. In *Rana computatrix* , we studied pattern recognition for Prey, Predators and Barriers. The models revealed the following representations:

Space: Not only did we see retinotopic mapping to multiple brain regions, but we also saw that the animal's pattern of action may be based on trajectory of objects ("normal" vs "cutback" avoidance behavior) and on the relative position of objects (the "maps" for prey and barrier must be brought into register so that the animal will snap directly or detour, as is appropriate).

Time: Once behavior is triggered by perception it proceeds to completion in the specified and appropriate order.

The Representation Of The World is then provided by the distributed interactions of a set of partial (incomplete, possibly inaccurate) representations.

In studying the neural mechanisms in Parietal Cortex, Frontal Eye Fields, Thalamus and Basal Ganglia underlying saccadic eye movements, we find that the Brainstem Saccade Generator converts SPACE (*retinotopic code*) to TIME (firing of the *excitatory burst neurons*). Further aspects of the model provide simple examples of capabilities crucial to the development of intelligent behavior. The Working Memory holds a "plan" (retinotopic position of target) until it is executed while Dynamic Remapping updates the "plan" as action proceeds.

Such schemata allow an overall functional analysis of the system but each brain region cooperates in the implementation of many schemata and each schema may involve the competition and cooperation of many brain regions — and find its "true meaning" only through its embedding within the overall schema network. Our brains and our science proceed by successive approximations to "reality".

Acknowledgments

The research described in this chapter was supported in part by grant no. 1RO1 NS 24926 from the National Institutes of Health (M. A. Arbib principal investigator).

3

Analysis of Single-Unit Activity in the Cerebral Cortex

MOSHE ABELES

Department of Physiology School of Medicine and The Neural Computation Center, The Hebrew University of Jerusalem, P.O. Box 1172, Jerusalem 91-010, Israel

Recording technique

Extracellular spiking activity of single units is usually recorded by inserting a very fine needle (micro-electrode) into the cortex while recording the voltage at the tip of the electrode. Living cells are characterized by having a potential difference across their membranes, with the inside negative relative to the outside. Active nerve cells are characterized by having different membrane potential at different sites. Such differences result in currents flowing through the salt solutions both inside and outside the cell. The currents, in the extracellular fluid, set potential differences which can be recorded by a micro-electrode. Analysis of current flow and potential differences in the extracellular fluid is difficult. Therefore it is customary to analyze the extracellular fields as if they were recordings of potentials generated by dipoles in dielectric space. The electric field induced by a surface dipole depends on the dipole moment, the area and orientation of the surface, and the distance from the dipole. Standard physics says that:

$$V(P) \approx \frac{\Omega \cdot \Delta Q \cdot \delta.}{\epsilon} \tag{1}$$

where $V(P)$ is the potential at point P, Ω is the solid angle at which the dipole is observed from P, ΔQ is the density of charge separation across

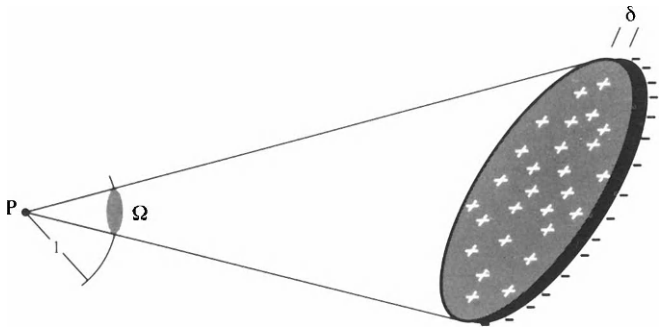


Fig. 1. A surface dipole as observed from *P*.

the dipole (i.e. charge separation per unit area), δ is the dipole thickness, and ϵ is the dielectric constant. Figure 1 illustrates this concept.

Since the density of charge separation depends on the membrane capacity per unit area *(C)* and voltage *(Vm)* difference, while the capacity depends on the membrane thickness (δ), and dielectric constant (ϵ), we get:

$$\Delta Q = Vm \cdot C$$

$$C = \frac{\epsilon}{\delta}$$

$$\frac{\Delta Q \cdot \delta}{\epsilon} = Vm.$$

Substituting into equation (1) we get:

$$V(P) = k \cdot \Omega \cdot Vm \tag{2}$$

where k is the proportionality constant. When measuring the potential at the center of a spheric cell, the observed potential *(V(P))* is equal to the membrane potential *(Vm)*. This yields a proportionality constant of $\frac{1}{4\pi}$. And therefore we end up with:

$$V(P) = \frac{\Omega}{4\pi} Vm. \tag{3}$$

Two properties become immediately obvious.

(1) When the electrode is outside a resting neuron we always measure 0 potential, because a cone tangent to the cell (Fig. 2A) divides the membrane into two surface dipoles one with the positive side towards the measuring point, and the other with the negative side towards the measuring point. Both parts have the same potential magnitude *(Vm)* and are observed with the same solid angle Ω.

(2) When measuring from the inside of a resting cell (Fig. 2B) one always measure the true membrane potential *(Vm)*, no matter where in the cell

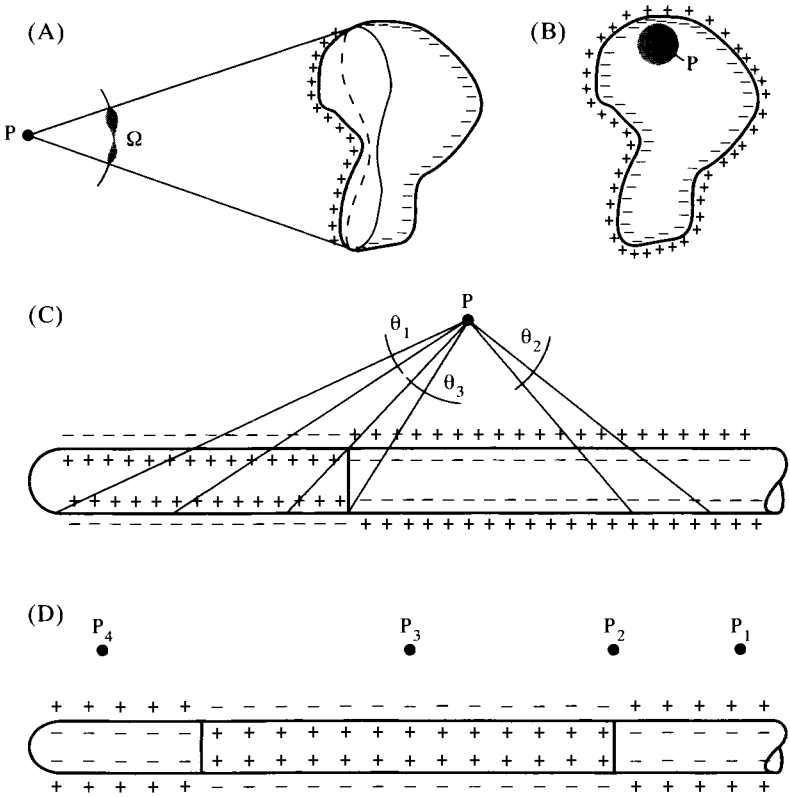


Fig. 2. Measuring voltages at various configurations (a) From any point outside a resting cell. The measured potential is 0. (b) From any point inside the cell the measured potential is *Vm*. (c) Measuring a depolarization front from outside. The measured potential depends on the solid angle spanned by the front alone. (d) What would be the potential measured at points *P1, P2, P3,* and *P4* along an axon conducting a box-car shaped action potential?

the measuring electrode is, because a membrane with the potential Vm is seen in all directions (i.e with a solid angle of 4π). For simplicity, assume a step-wise transition from resting to active zones in a nerve axon (Fig. 2C). What would one measure from the outside. Suppose we measure from point P at Fig. 2C. Looking on an active zone through the angle θ_1 would contribute no potential, because the point sees two dipoles of equal magnitude and opposite signs spanning the same solid angle. Similarly the contribution of the resting zone through an angle like θ_2 would contribute nothing. Only the transition front, observed through θ_3 would contribute a non zero value. In the configuration shown at Fig. 2C a positive potential would be recorded. If the front was propagating from left to right, then as it passes exactly under P a zero potential would be measured, and then (as the front is at the right of P) a negative potential would be observed. If activity, looking like a box-car, propagates along an axon a succession of positive–negative–positive waves would be observed. The reader can verify this by graphic analysis of the contribution of the depolarization and repolarization fronts to recordings at points *P1*, *P2*, *P3*, and *P4* in Fig. 2D. (See Woodbury, 1960, for a more detailed analysis of recording potentials in a volume conductor.) In a realistic situation the voltage profiles do not look like step functions. Then, the contribution of the membrane potential at position x, to the recorded voltage at P, depends on the second derivative of the membrane potential:

$$V(P,x) \approx \frac{\partial^2 V(x)}{\partial x^2}.$$

This relation tells us that the magnitude of extracellularly recorded field will critically depend on the spread of intracellular processes along the cell. A local change in membrane conductivity (due to synaptic activity for instance) causes a local current, which then spreads along the cell. The extent of spread depends on the amount of shunting, through the membrane, along the neuron. Generally, this spread is measured by the *length constant* (λ) of the structure. For stationary potentials, the length constant is given by:

$$\lambda = \sqrt{\frac{r_m}{r_i + r_o}} \tag{4}$$

where r_m is the resistance of the membrane per unit length, r_i and r_o are the resistance (per unit length) inside and outside the cell. When dealing with transients, this equation is insufficient, as transient voltage changes cause leakage also through the membrane capacitance. The membrane presents a poor insulator for fast transients, such as action potentials. Therefore the equivalent length constant for such transients is much shorter then for slowly varying potentials such as postsynaptic potentials. The resulting situation is that synaptic potentials spread well along the dendrites, while the action potential does not. Therefore the second derivatives of the synaptic

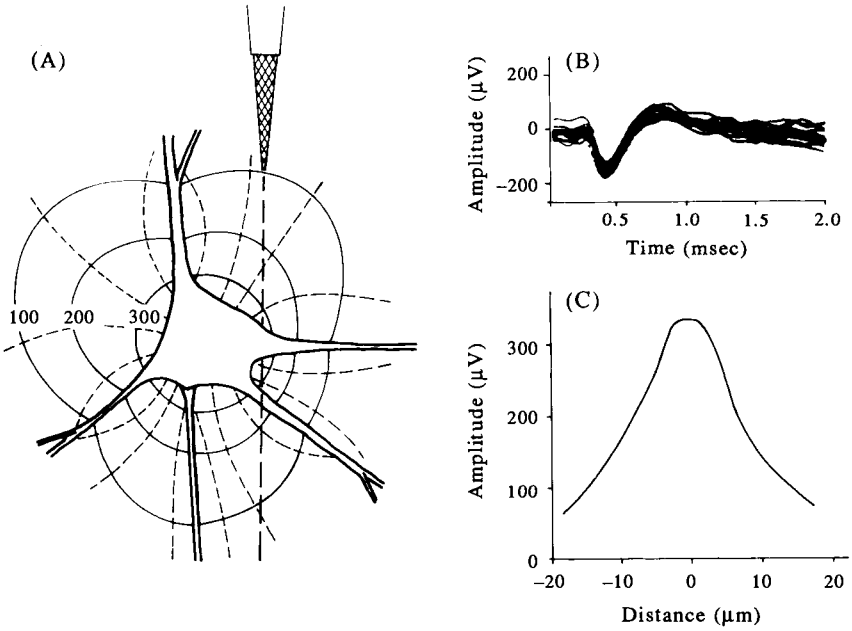


Fig. 3. Recording of extra-cellular action potentials in the cortex. (a) A hypothetical recording situation. (b) Superposition of many extracellularly recorded action potentials. (c) The variation of the peak to peak amplitude with recording depth (from Abeles, 1982a).

potentials are very small when compared to that of the action potential. On the other hand, synaptic potentials affect wide areas in the extracellular space while action potentials can be recorded extracellularly only over a limited range. Fig. 3 illustrates a schematic drawing of the cell body, the currents and potentials around it during an action potential, and an approximate size of a metal micro-electrode, through which the potential is recorded. It also shows an experimental measurement of the size of an action potential as recorded while moving such an electrode in small steps through the cortex. The accepted concept is that when a given electrode, at a given position around a given neuron, records action potentials repetitively, these (extracellularly) recorded action potentials will have the same shape.

Figure 3B illustrates a case where this prediction was met. Consequently, when in the experiment one records repetitively the same spike shapes, one assumes that the activity of a single nerve cell is recorded. To be on the safe side, physiologists call this type of recording *single unit recording* (and not single nerve cell recording), and the extracellularly recorded action potentials are referred to as *spikes*. The relevant parameter in such single

unit recordings are the times at which the action potentials occurred. Thus the result of such recordings is a time series:

$$\{t_1, t_2, ..., t_n\}$$

describing the times at which spikes were recorded. Typically time is measured with 1 ms accuracy.

The following two sections describe the standard methods for analyzing such time series, as used by neurophysiologists.

Stationary time processes

The auto-correlation function

The statistics of the activity of a single unit is most commonly studied by means of the *renewal density function* (Cox, 1970). This function describes the rate of firing of a single unit as a function of time which elapsed since it fired. The renewal-density function is estimated from recorded data by treating every spike as if it were a *stimulus* and measuring the rate of occurrence of spikes following that stimulus (see Perkel et al., 1967a and Abeles, 1982b for more details).

Figure 4 illustrates three examples of such renewal density functions. In Fig. 4A we see a firing patterns which is flat, except for a brief refractoriness at short times. Figure 4B shows a situation in which following a brief refractoriness, an increased firing rate is observed, which then subsides to the constant average firing rate of the single unit. Finally, Fig. 4C shows a case of periodically oscillating firing rate. Shapes such as in Figs 4A and B are found most commonly in the cortex of unanesthetized cats and monkeys (Abeles, 1982a, 1991). However, in some specific instances, periodic oscillation were found too (Eckhorn *et al.*, 1988; Gray and Singer, 1989; Ahissar and Vaadia 1990; Murthy and Fetz, 1992).

The classical view of the mammalian neuron has been that of an *integrate and fire* neuron, which resets its membrane potential after each spike, to some value well below threshold. This view is based on the pioneering studies of Sir J. Eccles (1957) on the motor-neuron of the spinal cord. Such neurons show prolonged depressions in their renewal density functions, and tend to fire periodically when driven to fire at elevated rates. Most cortical neurons do not have this property! This is evident from the absence of prolonged depressions in their renewal density functions, from the non periodic nature of their responses, and from illustrations of intracellularly recorded cortical neurons, where the absence of a deep hyperpolarization after the action potential is evident (e.g. Douglas *et al.*, 1991). A more detailed discussion of this issue can be found in Abeles (1991, 1993).

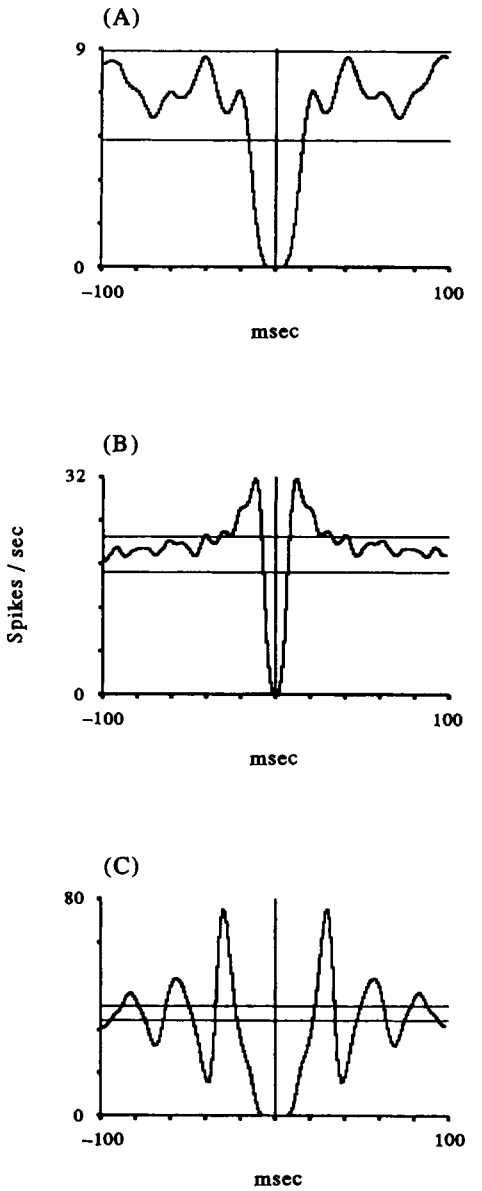


Fig. 4. Schematic representations of renewal density functions. Data obtained from 200 seconds of simulated spike trains. (a) Poissonian with refractoriness. (b) Bursting. (c) Periodic oscillations. The limited amount of data resulted in randomly fluctuating curves. The band marked by the thin lines limit the area which should contains 99% of the random fluctuations (see Abeles, 1982b, for more details). (A) was smoothed with a 10 ms wide bin. (B) and (C) were smoothed with 5 ms wide bins.

Finally, one should mention that in physiological literature the *renewal density* is usually called the *auto-correlation* function. If the spike train $\{t_1, t_2, ...\}$ is idealized by a series of δ functions at the appropriate times:

$$f(t) = \sum_i \delta(t - t_i). \tag{5}$$

Then the auto-correlation function:

$$\varphi(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f(t) \cdot f(t + \tau)\, dt \tag{6}$$

is related to the renewal density function $\lambda_{1,1}(\tau)$ by:

$$\varphi(\tau) = \lambda \cdot (\lambda_{1,1}(\tau) - \lambda) + \lambda \cdot \delta(\tau) \tag{7}$$

where λ is the average firing rate of the single unit. Thus, the two functions have essentially the same shape, except for a scaling factor λ, the subtraction of the average, and the addition of an impulse at the origin. In the text that follows the misnomer *auto-correlation* will be used for *renewal density*.

The cross-correlation function

Analysis of interactions between two spike trains by means of cross-correlations was introduced to neurophysiology by Perkel *et al.* (1967b) and is widely used nowadays. In practice what is computed should more properly be called *cross-renewal density* or *cross-intensity function* (Cox and Isham, 1980). However, the misnomer cross-correlation is deeply entrenched in the neurophysiological literature, and will be also used here.

The cross-correlation is computed just like the auto-correlation, by treating the spikes of one single unit as "stimuli", and estimating the firing rate of the other single-unit around it. Figure 5 illustrate some commonly found shapes of cross-correlations. Figure 5A shows an essentially flat cross-correlogram, meaning that there were no relations between the firing times of the two neurons. This is the most commonly found shape. A surprising finding if one considers the fact that the two single units were recorded through the same micro-electrode and therefore are presumably next neighbors.

Cross-correlograms such as seen in Fig. 5B have the next commonly observed shape. They show that one single unit has the tendency to elevate its firing rate around the time that the other single unit fired. A standard interpretation of such shapes in neurophysiology is that both units share some common drive which elevates, or depresses, the firing rates of both units depending on whether it is an excitatory, or inhibitory, in nature. Two single units having such a cross-correlogram are said, in the neurophysiological jargon, to have *shared inputs* or *common inputs*. The shape of the

cross-correlogram does not permit one to say whether the shared input is excitatory or inhibitory. The time-width of the positive correlation can vary from 1 ms to several seconds. The distribution of width's shows several peaks. Therefore it is probably correct to assume that there are several types of shared inputs (e.g. Nelson *et al.*, 1992).

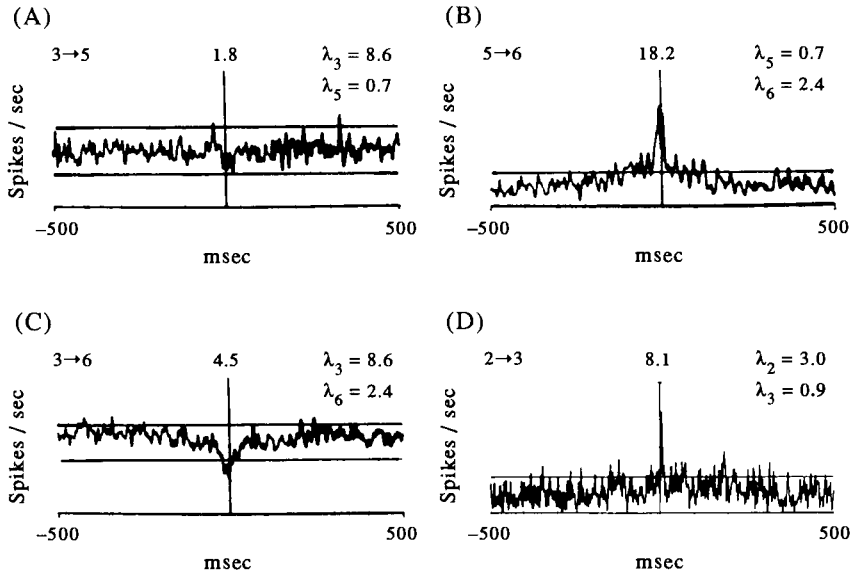


Fig. 5. Cross-correlations for pairs of single-units in the auditory cortex of unanesthetized cat. (A)–(C) For single units 3, 5, and 6, which were recorded through a single micro-electrode. (D) For another pair of single-units. Had the data been taken from two independent, homogeneous Poisson processes, then 99% of the data is expected to be included within the band which is marked by the thin horizontal lines (adapted form Abeles, 1982a).

Another type of cross-correlogram that indicates a shared input, is the one shown in Fig. 5C, where a trough straddling the 0 time-delay is seen. In such a case one says that the two single units share inputs of opposite signs, excitatory to one and inhibitory to the other. Again, it is impossible to state which of the single units is excited by this hypothetical shared input, and which is inhibited. The correlograms of Figs 5B and C indicate the existence of some unspecified relations between the firing of the two single-units. A correlogram which shows a sharp peak (or trough) on one side of the time axis is said to indicate direct interaction between the two single-units. A peak, such as in Fig. 5D indicate excitatory synaptic interaction, while a trough (which is found extremely rarely), indicates direct inhibitory interaction. One should stress that correlations cannot tell cause and effect, and therefore such shapes only indicate the type of synaptic connection, but do not prove it. We note that the cross-correlogram can be normalized

in various ways to estimate their strength. First, the choice of which single-unit is considered as a "stimulus" and which is the response is arbitrary. Conversion between the two cells is easily achieved by:

$$\lambda_{2,1}(\tau) = \frac{\lambda_1}{\lambda_2}\lambda_{1,2}(-\tau) \tag{8}$$

where, $\lambda_{i,j}$ is the cross-correlation between single-unit i (serving as a "stimulus") and single unit j, and λ_i is the average firing rate of single-unit i.

For correlograms as in Figs 5A, B, or C, there is no particular reason to choose one single-unit or the other as the "stimulus". However, when synaptic relations are implicated (as in Fig. 5D) it is desired to use the suspected presynaptic unit as the "stimulus". Then, the hump (or trough) would appear at positive time delays.

When causal relationships are implicated by assuming that one of the units is driving the other through an excitatory (or inhibitory) synapse, then the size of the hump (trough) has a very simple interpretation. The area between the hill (trough) and the baseline is the average number of spikes which is added (subtracted) from the postsynaptic train, following each spike in the presynaptic unit. This area is called the *gain* of the synapse, having positive values for excitatory synapses, a negative values for inhibitory synapses. Typical values for cortical excitatory connections are in the range of 0.01–0.2, i.e. the cortical synapses never transmit at a 1:1 ratio. While these values seem small, they are in fact large when considering the number ($\sim 20,000$) of excitatory synapses that a cortical neuron receives on the average. It is plausible that the cross-correlation technique is inefficient for resolving weaker synapses. Indeed, Cope *et al.* (1987) reported that in the spinal motoneurons, when the amplitude of the EPSP is less then 80 μV its effect on cross-correlations is not significant. The gain of the inhibitory synapses is not known, since it is extremely rare to find one-sided trough in cross-correlations. This situation probably stems from the difficulty to detect suppression in units that fire at low rates. See Aertsen and Gerstein, (1985), for further elaboration of the difficulty to detect inhibition by cross-correlations, and Fetz *et al.* (1991) for comprehensive review of cross-correlations in the cortex.

The use of the concept of synaptic gain, may be illustrated by the work of Ahissar *et al.* (1992) which examined the validity of Hebb's learning rule in the auditory cortex of behaving monkeys. They measured several single units simultaneously, and whenever the cross-correlation suggested a synaptic relation, they modified the amount of correlated activity between the units by generating the appropriate sounds. When the sound increased the correlation between the units the synaptic gain was increased. This phenomenon is illustrated in Fig. 6, where the synaptic gain as measured by the cross correlation between the two cells is increased from 0.039 to 0.175 by playing sound bleeps for 9 min, and then falls back to 0.056 after 7 min without sound. These experiments revealed that Hebb's rule is only part of

the requirements for synaptic strength modulations. Inducing correlations by sound was effective only when the sounds were part of a behavioral scheme in which the monkey had to press, or release a button according to the sound's nature. When the monkey did not respond behaviorally to the sounds (because he was not rewarded for making the correct movement), induction of correlations by sound had only negligible effect on the synaptic strength. Note that the idea about the increased efficiency of learning when there are behavioral implications was put forward by Thorendike (1911).

When the sound reduced the background correlation (by inducing activity in one unit, while suppressing it in the other) and the monkey was attending, the strength of the synapse was reduced.

The shape of the hump (trough) can be used to compute the modulation of firing rates which occur as activity traverse from the pre- through the postsynaptic neuron. For moderate to low firing rates, the relation between the firing rate of the presynaptic unit $\lambda_i(t)$, and the firing rate of the postsynaptic unit $\lambda_j(t)$, is given by:

$$\lambda_j(t) - \lambda_j = (\lambda_i(t) - \lambda_i) * (\lambda_{i,j}(\tau) - \lambda_j). \tag{9}$$

That is, the modulations of firing rate of the postsynaptic unit *(j)*, is given by the convolution between the modulations of firing rates of the presynaptic unit *(i)* and the modulations of the cross-correlogram between the two units (see Abeles, 1991 for more details).

What was called up to here a correlogram, was in fact an intensity (rate) function. It can be converted to a covariance function by:

$$\kappa_{i,j}(\tau) = \lambda_i \cdot \lambda_{i,j}(\tau) - \lambda_i \cdot \lambda_j \tag{10}$$

where, $\kappa_{i,j}$ is the covariance between cells i and j. While rates are easily interpretable, covariances are less interpretable. For continuous stochastic processes, a relevant question is how much of the variance of one process can be explained by the other? This is estimated by the correlation coefficient function which is computed by:

$$\rho_{i,j}(\tau) = \frac{\kappa_{i,j}(\tau)}{\sigma_i \sigma_j} \tag{11}$$

where σ^2 is the variance of each process. For point processes this approach does not hold. As the point process is represented by an infinite series of δ functions, its variance is infinity, and therefore the correlation coefficient is 0. It has been suggested by H. Sompolinsky (private communication), that this problem can be solved if one assumes that the observed spike trains is the realization of a Poisson process with varying rate functions and wishes to deal with the correlations between the underlying rate processes, and not between the spike trains.

Finally, interactions between more then two units can be studied too. This can be achieved by means of many pair-wise correlations, as is done

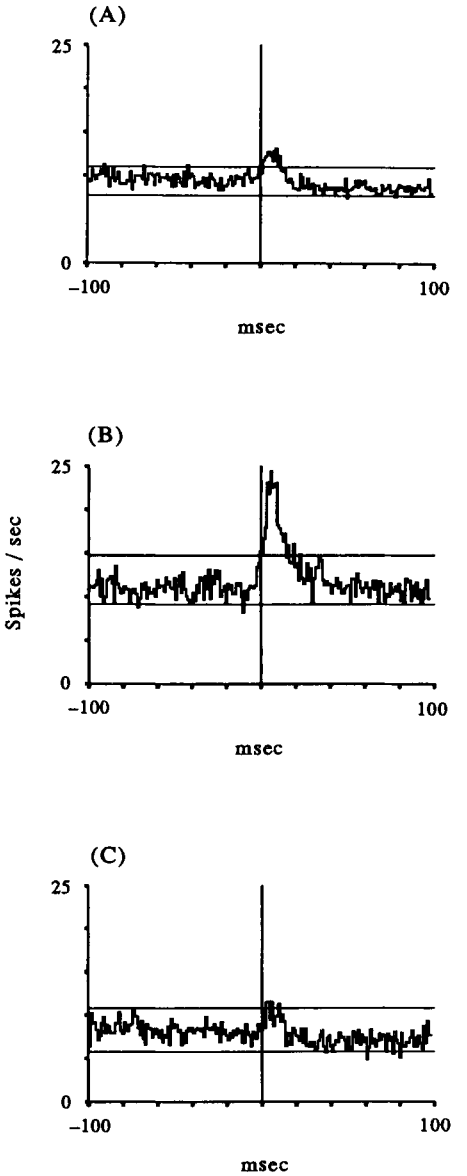


Fig. 6. Modulation of synaptic strength by Hebb's rule. (a) Cross-correlogram between two units in auditory cortex of a behaving monkey. (b) After 9 min of playing sounds which had behavioral meaning to the monkey. (c) After 7 min in a quiet environment.

in the *gravity* method (Gerstein *et al.*, 1985), or by higher order correlations. Cross-correlations of order higher then two were introduced by Perkel *et al.* (1975) under the nick-name *snow-flakes*. Their normalization to intensity functions was introduced by Abeles (1983), and their use was demonstrated in Abeles (1982a), and Abeles *et al.* (1983). Despite the rich information that can be derived from high order correlations they are used infrequently.

Non-stationary processes

In the experimental reality, neurophysiologists are usually interested in relations between single-unit activity and some external events. This external event might be a stimulus delivered by the experimenter, or a movement performed by the subject. Typically the firing rate of a single unit is modulated by such an event and one speaks about *event related activity*. In such a situation, it is not appropriate to treat the spike train as a stationary stochastic process.

The instantaneous firing rate

Most typically, event related activity is studied by means of *Peri Event Time Histograms* (PETH), in which the event is considered to happen in time 0 and the firing rate of the single unit around this time is estimated (see Abeles, 1982b for details). In these cases one cannot speak anymore on stationary point processes, and therefore, averaging properties along time has little meaning. In the following sections the procedures which are usually used by neurophysiologists, are described by using the concepts of non-stationary stochastic processes.

Let ¡Ω, P,**B**¿ be the following probability space: Ω is the collection of nerve cells with similar activity properties. Our "sampling experiment" is defined by the procedure of inserting a micro-electrode into the cortical tissue, and moving the electrode until the spiking activity of one of the neurons in Ω is sampled. **B** is the field of events that one can generate from enumerable collections of neurons in Ω. For instance an event in **B** may be {the sampled neuron is in layer IV}. The sampling procedure defines the probability of sampling any given neuron in Ω, and therefore, also of obtaining the various events of **B**. The spiking activity of every neuron (ω) in Ω, defines a point process $S_\omega(t)$, which is attached to that neuron. The probability of observing a spike at an infinitesimally small time window $(t, t+dt)$, is that of sampling one of the neurons in Ω which fired at that time window. Let us define the instantaneous firing rate $\lambda(t)$ of the process by:

$$Prob\{\ a\ spike\ at\ (t, t+dt)\} = \lambda(t). \tag{12}$$

To estimate the firing rate $\lambda(t)$, the following experiment (Fig. 7) has to be performed: measure simultaneously many *(N)* neurons from Ω, place a

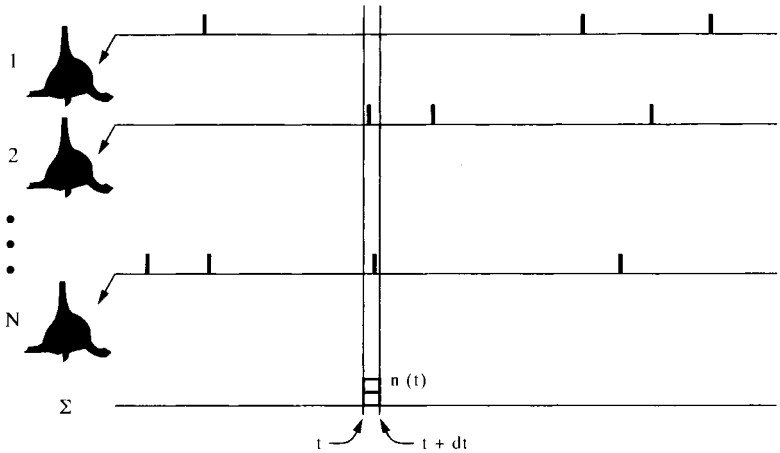


Fig. 7. Estimating the instantaneous firing rate by averaging across the ensemble of realizations.

narrow bin at $(t, t + dt)$ and count the number of neurons *(n(t))* for which a spike was recorded at that bin. As N goes to infinity, we expect to have:

$$\frac{n(t)}{N dt} \rightarrow \lambda(t). \tag{13}$$

While this evaluation method is mathematically simple, its experimental realization is not possible.

To overcome this difficulty the following experiment is done: the activity of a single unit is recorded around many repetitions of the same event. (For example, if the event is a stimulus, the experimenter delivers the same stimulus many times.) Care is taken to space the events along time, so that the intervals between successive occurrences is long, and it is safe to assume that the effect of one event on the firing rate faded before the other event occurred. The activity of the unit around each of the events is treated as if it was a different realization of the point process. Thus instead of measuring the activity of N different neurons (in parallel) around a single event, the activity of a single unit is measured around N repetitions of the event. Then, the instantaneous firing rate is computed just as for the parallel recording. The main price paid is that the function $\lambda(t)$ can be estimated only over a limited time range (less then the interval between successive events).

It is assumed that time (on the grand scale) does not affect the underlying mechanisms, so that there is no systematic differences between the activities around the events that happened early in the experiment and those that happened later on. In physiology when experimenters speak about stationarity they usually mean the stability of the properties along the experiment and not the absence of modulations in firing rates which are related to some external event.

Non-stationary correlations

The correlation functions (both auto and cross-correlations), of non-stationary processes are functions of two times (t_1, t_2). For the auto-correlation, let us define $prob_{1,1}\{t_1, t_2\}$ as the probability of seeing a spike at time $(t_1, t_1 + dt_1)$, and also at time $(t_2, t_2 + dt_2)$, then the basic equation is:

$$Prob_{1,1}\{t_1, t_2\} = \lambda_{1,1}(t_1, t_2)\ dt_1\ dt_2. \tag{14}$$

Similarly, for non-stationary cross-correlation, let us define $prob_{1,2}\{t_1, t_2\}$ as the probability of seeing a spike of unit no. 1 at time $(t_1, t_1 + dt_1)$, and of unit no. 2 at time $(t_2, t_2 + dt_2)$, then the basic equation is:

$$Prob_{1,2}\{t_1, t_2\} = \lambda_{1,2}(t_1, t_2)\ dt_1\ dt_2. \tag{15}$$

Estimating these correlations experimentally is done as for the PETH. That is, instead of measuring the activity of many different units around a single event at time 0, and then averaging across all the units, the activity of one (or two) units is measured while the event repeats many times, and then the activities around these repetitions are considered as an ensemble of independently sampled realizations of a stochastic process. The procedure of estimating these joint correlations is known as JPSTH (*Joint Peri Stimulus Time Histogram*) (see Aertsen *et al.*, 1989, for a more detailed description of how this histogram is constructed and interpreted).

The neurophysiologist is often interested in evaluating which part of the correlation is due to direct effect of the stimulus (or event) on the single unit(s), and which part is due to local interactions in the network. For that end it is assumed that the correlation can be decomposed into two processes:

$$\lambda_{1,2}\{t_1, t_2\} = \lambda_1(t_1) \cdot \lambda_2(t_2) + I_{1,2}(t_1, t_2) \tag{16}$$

where λ_i is the firing rate (as estimated by the PETH) of unit i, and $I_{i,j}$ is the cross-correlation contributed by the internal processes in the network. A question which is often asked is whether the internal correlation is stationary so that:

$$I_{1,2}(t_1, t_2) = I_{1,2}(t_2 - t_1). \tag{17}$$

That is, whether the internal cross-correlation depends only on the delay $(t_2 - t_1)$ between the two spikes. Very often one finds that the internal cross-correlations are not stationary, and show large modulations of size, and even of sign, within a fraction of a second (e.g. Vaadia and Aertsen, 1992).

Dynamic modulations of internal correlations suggest that the relations amongst the activities of single units, is not merely a reflection of the anatomical connections amongst the neurons, but reflect the dynamics of organization of activity in the network. Thus the correlations are an emergent property of the cooperative effects of large number of neurons on each other, and not a reflection of few, simple, connections. It is hoped that the

neural network mechanisms operating in the brain can be deciphered by analysis of the shape, magnitude and dynamics of the correlations. However, at present time, the theoretical relations between the network behavior and the properties of correlations are only partially understood.

Acknowledgments

The author is indebted to Mrs. V. Sharkansky for help in the artwork. This research was supported in part by a grant from the United-States Israeli Binational Science Foundation (BSF).

4

Single Neuron Dynamics: an Introduction

LARRY F. ABBOTT

Department of Physics and Center for Complex Systems, Brandeis University, Waltham, MA 02254, U.S.A.

Introduction

The development and growth of computational neuroscience is attracting a large number of newcomers to neurobiology from nonbiological fields like physics and computer science. This chapter is intended as a crash course in single-neuron neurophysiology and modeling for new arrivals to the field. I have tried to gather together here some of the basic information about neurons that will make it easier for people with a background in the physical and mathematical side of science to model neurons and to communicate with neurobiologists. In contrast with broader issues of brain function, a great deal of precise information is available about how single neurons behave. This material forms the foundation on which neuroscience is built.

Throughout this introduction I have tried to give representative values of the many constants and parameters that arise. The reader must be warned that there is great deal of variation in neurons and these values should not be taken as anything more than order of magnitude indications. In addition, I will give some examples of membrane currents and synaptic transmitters but my lists will be far from complete.

This chapter is devoted to developing mathematical descriptions of neuronal behavior. Much of the material is available in standard texts. Useful general texts on neuroscience are Levitan and Kaczmarek (1991), Nicholls *et al.* (1992) and Kandel *et al.* (1991). Suggested texts on theoretical and modeling aspects are Tuckwell (1988) and Koch and Segev (1989) (see

also Bower 1992). Ionic conductances are covered in detail by Hille (1992) and cable theory is presented in Rall (1977), Tuckwell (1988) and Jack *et al.* (1975). Other reviews are quoted in the text and where I discuss less standard material I have included references to research articles.

The neuronal cable equation

To understand the electrical characteristics of a neuron, we must identify all of the factors that affect its membrane potential. The potential in the bath surrounding a neuron is relatively constant under normal circumstances, and it is conventionally taken to be zero. The potential inside the cell is then, by definition, the membrane potential. It can be measured by inserting a hollow glass electrode filled with an electrolytic solution through the cell membrane. At rest, the membrane potential is typically around -70 mV and during the normal course of cell activity it may vary from about -80 mV to $+50$ mV. These variations are caused by changes in the conductances of channels that allow ions to cross the cell membrane. Because of its negative membrane potential, the neuron is polarized at rest. Making the membrane potential more negative is termed hyperpolarization and moving it in the positive direction is depolarization.

Figure 1 shows a rough schematic of a bipolar neuron. Since the tubular cables that form the dendritic and axonal structures of a neuron are long and narrow, variations in the potential in the radial or axial directions are negligible compared to longitudinal variations. Therefore, neuronal cables are usually treated as one-dimensional structures parameterized by a single longitudinal coordinate x. The lower part of Fig. 1 is an enlargement of a small segment of the neuron that we will use to derive the equation that governs the behavior of the membrane potential. This segment has a radius a and a short length Δx. The basic equation we will derive relates the rate of change of the membrane potential to currents passing into the cable segment. The cell membrane is a lipid bilayer that is an excellent insulator. When the membrane potential is nonzero, charges build up on either side of the membrane which thus acts like a capacitor. Changing the membrane potential V requires a charging current proportional to $\partial V/\partial t$. The capacitance of the cell membrane is around $C = 1\ \mu\text{F/cm}^2$. The small cylinder of membrane shown in Fig. 1 has a surface area of $2\pi a\Delta x$ so it will take a current equal to $2\pi a\Delta x C\partial V/\partial t$ to change the membrane potential at a rate $\partial V/\partial t$.

We derive the basic cable equation for the neuron by identifying all of the sources of current capable of charging up the cell membrane. First, current can flow longitudinally down the cable segment shown in Fig. 1 from neighboring segments. The longitudinal current flow depends on the resistivity of the intracellular medium, a quantity I will call r. Typical values of the longitudinal resistivity lie between $r = 70\Omega\text{cm}$ and $r = 300\Omega\text{cm}$.

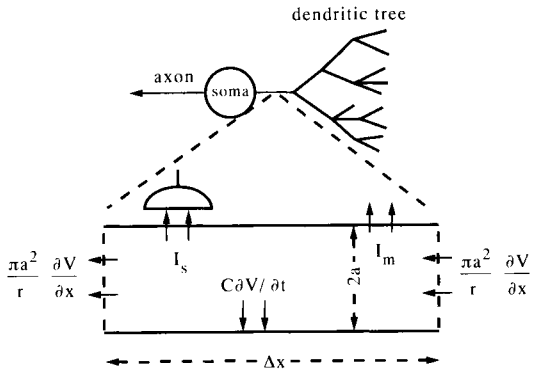


Fig. 1. A schematic of a typical neuronal structure with a portion of the cable blown up to identify specific currents. These are used in the text to derive the cable equation.

The longitudinal resistance of a given cylinder is obtained by multiplying the resistivity by the length of the cylinder and dividing by the cross-sectional area. This explains the rather unintuitive units for the resistivity. The longitudinal flow of current at any point along a cylinder is then given by Ohm's law as

$$I_L = \frac{\pi a^2}{r}\frac{\partial V}{\partial x}. \tag{1}$$

This is a useful result for many applications.

Although the membrane itself is a good insulator, numerous protein channels imbedded in the membrane are capable of conducting ions into and out of the cell. We will discuss these more fully in the next section. For now, we will merely call the current per unit area of membrane due to these channels I_m. By convention, the membrane current is defined as positive outward. Channels are also located at the sites of synaptic contacts with other neurons. The current conducted by synaptic channels, again per unit area of membrane, is labeled I_s. For the small cylinder in Fig. 1, the total membrane current is the area of the membrane times I_m or $2\pi a\Delta x I_m$ and the total synaptic current is $2\pi a\Delta x I_s$.

All of the terms we have discussed are shown in Fig. 1. The cable equation is derived by setting the sum of all the currents equal to the current needed to charge the membrane. The total longitudinal current entering the cylinder is the difference between the current flowing in on the right and that flowing out on the left. We label this different with a Δ so that

$$2\pi a\Delta x C\frac{\partial V}{\partial t} = \Delta\left(\frac{\pi a^2}{r}\frac{\partial V}{\partial x}\right) - 2\pi a\Delta x(I_m + I_s). \tag{2}$$

Dividing through by $2\pi a\Delta x$ and taking the limit $\Delta x \to 0$ we get the cable equation

$$C\frac{\partial V}{\partial t} = \frac{a}{2r}\frac{\partial^2 V}{\partial x^2} - I_m - I_s. \tag{3}$$

Note that we have allowed the derivative to pass through the factor a in this derivation. If the cable has a taper and the radius a depends on x then the second spatial derivative term in equation (3) would be

$$\frac{1}{2a}\frac{\partial}{\partial x}\left(\frac{a^2}{r}\frac{\partial V}{\partial x}\right) \tag{4}$$

instead.

The cable equation is often written in another equivalent form. The total resistance of a piece of membrane is inversely proportional to its surface area. The constant of proportionality has units of resistance times area and we will label it with the letter R. R is often called the membrane resistance although it is really the inverse of the membrane conductance per unit area. Its value may depend on the membrane potential so it is conventionally defined when this is equal to the resting potential of the neuron. If we multiply equation (3) by R we get an equivalent but more transparent form of the cable equation

$$\tau\frac{\partial V}{\partial t} = \lambda^2\frac{\partial^2 V}{\partial x^2} - R(I_m + I_s) \tag{5}$$

where $\tau = RC$ is the membrane time constant and $\lambda = (aR/2r)^{1/2}$ is the cable length constant. The importance of the two quantities should be clear, τ sets the basic time scale for temporal variations and λ the length scale for spatial variations in the potential.

A pyramidal neuron in the cortex might have a total surface area of around $50,000\mu\text{m}^2$, a capacitance of 0.5 nF and a total input resistance of about 100 MΩ. R typically varies between about 10 kΩ cm^2 and 100 kΩ cm^2 which means that τ = 10–100 ms. Typical cable radii are in the range 0.1–10 μm and length constants range from hundreds of microns to millimeters. These numbers are important for characterizing the properties of a neuron but they should not be used too rigidly. Although I have written them as currents, the terms I_m and I_s represent conductance changes which affect both the time constant and the length constant of the cable. Under operating conditions, the relevant time and length scales can be quite different from those applicable to a resting neuron.

The membrane potential everywhere on a neuron can be obtained from equation (5) if appropriate boundary conditions are imposed. The boundary conditions specify what happens to the membrane potential at a node where the neuronal cable branches or at a terminal where it ends. Specifically, at a branching node, the potential V must be continuous and the longitudinal current given by equation (1) must be conserved, so that the current entering along one branch adds up to that leaving along the other branches. At a terminal, different boundary conditions can be imposed, but

a reasonable condition is that no current flows out the end of the cable. By equation (1) this means that the spatial derivative of potential must vanish at a terminal point.

Equation (3) or (5) and the boundary conditions determine the behavior of the membrane potential everywhere on a neuron. However, before we can proceed with this calculation we must specify the membrane current I_m and the synaptic current I_s appearing in these equations. The real complexity and richness of neuronal dynamics arises from these terms. They are discussed in the next two sections.

Membrane conductances

The electrical currents that affect the membrane potential of a neuron arise from the flow of ions, principally potassium, sodium, calcium and chloride, across the cell membrane. Ions enter or leave the cell through pores formed by protein channels located in the cell membrane. A typical neuron may have around a dozen different types of channels. There are anywhere from a few to hundreds of channels in a square micron of membrane. Channels are often highly selective allowing just a single ionic species to pass through them, although some channels allow more than one type of ion through. Ion channels can exist in either open or closed states and recordings of the current flowing through a single channel indicate that channels fluctuate rapidly between open and closed states. In order to compute the total current passing through a large population of channels, we need to know the probability P that a channel is in the open state.

Before discussing the open probability P, we will derive a general expression for the total current per unit area passing through a given type of channel. For a large population of channels, the open probability P determines the fraction of channels that are open at any given time. These open channels correspond to a membrane conductance. The value of this conductance depends on the density of channels and the conductance of an individual channel. The product of these two factors is called the maximal conductance and is denoted by $\overline{g}$. It is the total conductance per unit area if all of the channels are in the open state. The actual conductance is the maximal conductance $\overline{g}$ times P the fraction of channels that are open. The total current is the product of this conductance times the driving force that pushes ions through the channels.

Two forces are responsible for driving ions through the channel pores. First, any potential difference between the exterior and interior of the cell will produce an electric force on the ions. By Ohm's law, this results in a current proportional to the potential difference V. When a neuron is at rest, the negative resting potential attracts positive ions into the neuron and repels negative ions. In addition, ions diffuse through channels because the ionic concentrations differ inside and outside the neuron. These differences

are maintained by ion pumps within the cell membrane. The concentrations of sodium, calcium and chloride are higher outside the cell than inside so these are driven into the neuron by diffusion. Potassium is more concentrated inside the neuron than outside so it tends to diffuse out of the cell. It is convenient to characterize the current flow due to diffusion in terms of an equilibrium or reversal potential E. This is the membrane potential for which the electric force exactly cancels the diffusive flow. For a single ion, the equilibrium potential E can be computed easily. The flow of ions into the cell is proportional to the concentration outside the cell while the flow out is proportional to the concentration inside. The energy difference between an ion of charge q outside and inside the neuron is qE when the membrane potential is E. Equilibrium of ion flow will occur when the ratio of concentrations is equal to the Boltzmann factor corresponding to this energy difference,

$$\frac{[\text{inside}]}{[\text{outside}]} = e^{-qE/kT}. \tag{6}$$

The notation [ ] stands for the concentration of a particular ion, k is Boltzmann's constant and T is the temperature. This determines the equilibrium potential for each ion

$$E = \frac{kT}{q} \ln\left(\frac{[\text{outside}]}{[\text{inside}]}\right) \tag{7}$$

which is the Nernst equation. Typical values based on normal inside and outside concentrations are $E_{Na} = 50$ mV, $E_{Ca} = 150$ mV, $E_{Cl} = -70$ mV and $E_K = -80$ mV. Some channels pass a combination of ions and have equilibrium potentials between these values.

When the membrane potential V is equal to the equilibrium potential E for a particular ion, the net flow of that ion through an open channel is zero. In general, the driving force for a particular conductance is just the difference $V - E$ between the membrane potential and the equilibrium potential. The total membrane current is the sum of currents corresponding to various different types of channels. In discussing the form of the different terms, I will not bother to introduce an extra index to label the particular conductance being discussed so keep in mind that the following equations refer to individual components of the total membrane current and that each component has its own parameters and variables labeled by a suppressed index.

The total current flowing through a set of membrane channels is determined by multiplying the conductance discussed above by the driving force so

$$I_m = \overline{g}P(V - E). \tag{8}$$

Maximal conductances $\overline{g}$ tend to be in the range of μS/cm^2 to mS/cm^2 where $S = 1/\Omega$.

The fraction of open channels P is of course between zero and one and its description plays the most important role in determining the properties of a given membrane conductance. When a particular conductance is active ($P \neq 0$) it tends to push the membrane potential toward the equilibrium potential E characterizing that conductance. Sodium and calcium conductances which have positive equilibrium potentials will depolarize the neuron while potassium conductances hyperpolarize it. When a neuron is near its resting potential, a chloride conductance may have little direct affect on the membrane potential, rather its primary impact is to change the membrane resistance of the cell.

If a pore is open at all times so that $P = 1$, the corresponding conductance is termed passive. Such conductances produce what are called leakage currents. Leakage conductances are often included in neuronal models as fudge factors correcting for conductances that have not been properly measured and leakage caused by the measuring electrode. Active conductances have open probabilities P that depend on a number of factors. For voltage-gated channels, P depends on the membrane potential. For example, in one class of conductances P increases steadily with voltage. These are called persistent or noninactivating. Although P increases with increased V it does not do so instantaneously. It takes a certain amount of time for the conductance to respond to a voltage change. To describe this behavior, P is written in the form

$$P = n^p \tag{9}$$

where p is an integer chosen to best fit the data and n is determined by first-order kinetics with voltage-dependent opening and closing rates $\alpha_n(V)$ and $\beta_n(V)$ (Hodgkin and Huxley, 1952)

$$\frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n. \tag{10}$$

This equation can be written in the equivalent form

$$\tau_n(V)\frac{dn}{dt} = \overline{n}(V) - n \tag{11}$$

if we identify

$$\tau_n(V) = \frac{1}{\alpha_n(V) + \beta_n(V)} \tag{12}$$

and

$$\overline{n}(V) = \frac{\alpha_n(V)}{\alpha_n(V) + \beta_n(V)}. \tag{13}$$

$\tau_n(V)$ is a voltage-dependent time constant reflecting the rate at which the conductance responds at a particular voltage. $\overline{n}(V)$ is the value that n reaches asymptotically at a given voltage (it is sometimes called n_∞). In

other words, if the membrane potential jumps to a value V, n will approach $\overline{n}(V)$ exponentially with time constant $\tau_n(V)$. $\overline{n}(V)$ is typically a sigmoidal function which vanishes for extremely negative potentials and approaches one for positive potentials. A well-known example of a conductance of this type is the delayed-rectifier potassium conductance responsible for repolarizing neurons after an action potential has fired.

Some conductances have a more complicated dynamics than those just discussed. We can think of the conductances described by equation (9) as having a single gate that opens with depolarization, and closes with hyperpolarization. Other conductances act as if they are controlled by two gates. One gate, described by an activation variable m, behaves like the one discussed above. The other, described by an inactivation variable h, behaves in the opposite way, it closes when the membrane is depolarized and opens when it is hyperpolarized. The probability of a channel being open is the product of two factors accounting for the state of each gate,

$$P = m^p h^q \tag{14}$$

where p and q are integers. Since the same process (for example depolarization) that opens one gate closes the other, these conductances are often referred to as transient and they are also called inactivating. The activation m and inactivation h are described by equations just like (11),

$$\tau_m(V)\frac{dm}{dt} = \overline{m}(V) - m \tag{15}$$

and

$$\tau_h(V)\frac{dh}{dt} = \overline{h}(V) - h. \tag{16}$$

The form of the asymptotic value functions $\overline{m}(V)$ and $\overline{h}(V)$ distinguishes an activation from an inactivation gating variable. $\overline{m}(V)$ increases as a function of V while $\overline{h}(V)$ decreases. Both have a sigmoidal shape. The description in terms of independent activation and inactivation gates is useful for understanding how these conductances are modeled but it should not be taken too literally. It is known from single channel studies that activation and inactivation are not independent of each other (Hille, 1992).

The presence of the two factors in equation (14) gives this type of conductance some interesting properties. In order to turn on such a conductance, it may be necessary to first hyperpolarize the neuron and then to depolarize it. Hyperpolarizing the membrane, raises the value of the inactivation h, a process called deinactivation. The second step, depolarization, increases the value of the activation m which is known as activation of the conductance. Only when m and h are both nonzero is the conductance turned on. Note that the conductance can be reduced in magnitude either by decreasing m or h. Decreasing m is called deactivation to distinguish it from decreasing h which is called inactivation. This terminology takes a

bit of getting used to. The best known conductance in this class is the fast sodium conductance responsible for the sharp rise in potential that forms the leading edge of an action potential. There are also transient calcium (T-type) and potassium (the A current) currents in this class.

Some conductances act as if they are gated by a single gate, like the persistent conductances described first, but this gate opens when the membrane is hyperpolarized not depolarized. They are described just like (9) with (11) except that the asymptotic value function $\overline{n}(V)$ is a decreasing sigmoidal function of V as is typical of an inactivation variable. Strictly speaking these conductances deinactivate when they turn on. However, most people cannot bring themselves to say deinactivate all the time so instead they say these conductances are activated by hyperpolarization.

The persistent, transient and hyperpolarization-activated conductances are all the result of voltage-dependent channels. In addition, the state of some channels can be modified by intracellular factors, like calcium. For example, calcium-dependent potassium conductances open if the intracellular calcium concentration near the channel rises. This provides an important addition mechanism for activating membrane conductances. These conductances are described as above but their activation variables depend on the intracellular calcium concentration as well as on the membrane potential (Yamada *et al.*, 1989; Buchholtz *et al.*, 1992).

Synaptic conductances

The synaptic current I_s can receive contributions from two types of synapses, electrical and chemical. Synaptic current at a chemical synapse is due to membrane channels similar to those that determine I_m except that they are gated not by voltage but by an extracellular neurotransmitter. The most straightforward type of synaptic channel is activated directly by binding to a neurotransmitter. As in the last section, we can write the total synaptic current as a sum of terms arising from different synapses and we will discuss each term separately without introducing extra indices to label a particular synapse. The synaptic current due to a particular class of receptor channels can be expressed in terms of a maximal synaptic conductance $\overline{g}_s$, an open probability which we will denote by s, and a driving force factor that is the difference between the membrane potential of the postsynaptic neuron V_{post} and the equilibrium or reversal potential for the synaptic conductance E_s,

$$I_s = \overline{g}_s s(V_{post} - E_s). \tag{17}$$

E_s depends on the type of ion carrying the current just as it did for a membrane conductance. A simple equation can be derived for s by assuming first-order kinetics for the transmitter-channel binding process. Transmitter binds to a closed channel and opens it. The rate for this binding process

is proportional to the number of closed channels and the concentration of transmitter in the synaptic cleft. Since s is the fraction of open channels, the number of closed channels is proportional to $1 - s$. If T is the transmitter concentration and r transmitter molecules must bind to the channel to open it, the rate of transmitter binding and channel opening can be written as $k_1(1 - s)T^r$ where k_1 is a constant. Release of the transmitter from the open channel is proportional to the number of open channels and so can be written as $k_2 s$ where k_2 is another rate constant. Combining these two expressions we find

$$\frac{ds}{dt} = k_1(1 - s)T^r - k_2 s. \tag{18}$$

If we define the asymptotic value

$$\bar{s} = \frac{k_1 T^r}{k_2 + k_1 T^r} \tag{19}$$

and the synaptic time constant $\tau_s = 1/k_2$ we can write the equation for s in an equivalent form

$$\tau_s(1 - \bar{s})\frac{ds}{dt} = \bar{s} - s. \tag{20}$$

The time constant and the function $\bar{s}$ determine the behavior of the synapse. Synaptic time constants range from milliseconds to hundreds of milliseconds. Since $\bar{s}$ depends on the transmitter concentration it is determined by the presynaptic neuron. It is often written as a function of the membrane potential of the presynaptic neuron, V_{pre}. Typically, there is a threshold for transmitter release; $\bar{s}$ is zero below this voltage and it rises from zero above the threshold. Ultimately, this rise must stop because $\bar{s}$ cannot get bigger than one. For example, we can use half of a hyperbolic tangent function for $\bar{s}$, $\bar{s} = \tan h[(V_{pre} - V_{th})/V_s]$ for $V_{pre} > V_{th}$ and $\bar{s} = 0$ for $V_{pre} < V_{th}$. The parameter V_s determines the slope of the rise in the synaptic response with V above the threshold.

Equation (20) reveals one common property of synaptic responses. When a large amount of the transmitter is released, $\bar{s}$ approaches its maximal value of one. The effective time constant for the resulting rise to activation of the synaptic conductance is, from equation (20), $(1 - \bar{s})\tau_s$ which is significantly less than τ_s if $\bar{s}$ is near one. However, when the transmitter is removed and the synapse deactivates, $\bar{s} \to 0$ so this process will occur with a time constant τ_s. This reproduces the effect that synaptic rise times tend to be significantly shorter than their fall times.

Although $\bar{s}$ is often written as a function of the potential of the presynaptic neuron it really depends on the calcium concentration inside the presynaptic terminal. This is because the release of vesicles containing transmitter molecules from the presynaptic neuron is mediated by intracellular calcium. Transmitter release varies like the third or fourth power of the calcium concentration. The presynaptic calcium concentration is related to

the presynaptic potential because calcium enters the presynaptic terminal through voltage-dependent calcium channels. Thus, it is not unreasonable to write $\bar{s}$ in terms of the presynaptic voltage, though in some cases this may leave out important effects (Calabrese and De Schutter, 1992).

Some synaptic channels are not directly gated by a neurotransmitter but rather are activated indirectly by a second messenger system. In this case, the binding of the transmitter to the postsynaptic receptor sets off a biochemical cascade that ultimately leads to the production of a factor that opens the channel. Modellers should be aware that such indirect processes exist because they provide a pathway by which synaptic transmission can, in addition to opening channels, cause fundamental and long-lasting changes inside a neuron. This relates issues of synaptic transmission to general mechanisms of molecular signaling. In addition, the second messenger cascade can act as a chemical amplifier.

As examples of chemical synapses, I will mention four important synaptic conductances. Glutamate is the major excitatory neurotransmitter in the brain and it activates two different types of conductances that are mentioned frequently in connection with long-term potentiation (long-lasting strengthening of synapses) and learning. Both are mixed-ion conductances with equilibrium potentials around 0 mV. They are named for chemicals (agonists) that simulate the effects of the natural transmitter glutamate. The AMPA current is fast activating and deactivating. NMDA receptors also activate quickly but they deactivate more slowly. In addition, NMDA receptors have a novel voltage dependence that makes them interesting in connection with the Hebb rule for synaptic modification. The Hebb rule requires that potentiation at a synapse depend strongly on both the pre and postsynaptic potentials. In our discussion thus far, the postsynaptic potential enters only linearly in equation (17). This is hardly a strong dependence. How then can something like the Hebb rule arise? For the NMDA receptor, the factor $\bar{s}$ depends on both the pre and postsynaptic potentials (Jahr and Stevens, 1990). The presynaptic potential dependence is just the usual one reflecting the amount of transmitter released. The unusual postsynaptic dependence is due to the fact that NMDA receptors are normally blocked by magnesium ions. In order to unblock the conductance the postsynaptic neuron must be depolarized to knock out the blocking ions.

Two important inhibitory synaptic conductances are both activated by the neurotransmitter GABA (γ-aminobutyric acid) (see Connors *et al.*, 1988 for example). $GABA_A$ receptors form relatively fast synapses with a reversal potential near the resting membrane potential of many neurons. $GABA_B$ synapses are slower and longer lasting potassium conductances with a more negative reversal potential. They are an example of an indirectly (G-protein) coupled synaptic channel.

In addition to chemical synapses, neurons can be coupled through electrical synapses (gap junctions) that produce a synaptic current proportional to the difference between the pre and postsynaptic membrane potentials,

$$I_s = G(V_{post} - V_{pre}) \tag{21}$$

where G is a constant. Some gap junctions rectify so that positive and negative current flow is not equal for potential differences of the same magnitude. Even though they are described quite simply, these couplings can have interesting and complex effects on the behavior of neural networks (Abbott *et al.*, 1991).

Single compartment models

The simplest way to model a neuron is to assume that the entire cell can be characterized by a single membrane potential. This is only valid if the neuron is electrotonically compact, i.e. smaller than its electrotonic length constant, or at least if the behavior being studied is not strongly affected by potential differences between different parts of the neuron. Some care should be taken in evaluating this condition. For example, when large numbers of membrane and synaptic conductances are active, shunting will cause the electrotonic length to be shorter than it is for a silent isolated neuron (see for example, Abbott, 1991a; Bernander *et al.*, 1991; Rapp *et al.*, 1991). However, if these conditions are met, the second spatial derivative in equations (3) can be ignored and we can use the simpler, single compartment equation

$$C\frac{dV}{dt} = -I_m - I_s. \tag{22}$$

I would like to discuss the model of the action potential introduced by Hodgkin and Huxley (1952), and show how it can be reduced and understood as a typical nonlinear oscillator model. In this example, there will be no synaptic current, $I_s = 0$, but an external current I_{ext} will be added to the right side of equation (22) to probe the behavior of the model. The Hodgkin-Huxley model has two active conductances, one for sodium and one for potassium, described as discussed above. There are four dynamical variables V, n, m and h and

$$I_m = I_m(V, m, h, n) = \overline{g}_L(V - E_L) + \overline{g}_K n^4(V - E_K) + \overline{g}_{Na} h m^3(V - E_{Na}) \tag{23}$$

where the maximal conductances are $\overline{g}_L = 0.3mS/cm^2$, $\overline{g}_K = 36mS/cm^2$ and $\overline{g}_{Na} = 120mS/cm^2$ and the equilibrium potentials are $E_L = -54.402mV$, $E_K = -77mV$ and $E_{Na} = 50mV$. The sodium activation m, sodium inactivation h and potassium activation n are determined by equations (11), (15) and (16). The six functions $\tau_m(V)$, $\tau_h(V)$, $\tau_n(V)$, $\overline{m}(V)$, $\overline{h}(V)$ and $\overline{n}(V)$ are expressed in terms of α and β functions as in equations (12) and (13) and these are given by fairly complicated formulas which are the result of detailed fits to experimental data,

$$\alpha_m = \frac{.1(V + 40)}{1 - \exp[-.1(V + 40)]} \qquad \beta_m = 4\exp[-.0556(V + 65)] \tag{24}$$

$$\alpha_n = \frac{.01(V+55)}{1-\exp[-.1(V+55)]} \quad \beta_n = .125\exp[-.0125(V+65)]$$
$$\alpha_h = .07\exp[-.05(V+65)] \quad \beta_h = (1+\exp[-.1(V+35)])^{-1}.$$

The units in these formulas have been suppressed with the understanding that all potentials are in mV , all times in ms and all currents in μA/cm^2. With these units, $C = 1$.

When $I_{ext} = 0$, V remains at a resting potential of -65 mV for this model. A positive I_{ext} of sufficient strength induce a sudden depolarization to about 50 mV followed rapidly by a hyperpolarization and then a slower recovery back to the resting level. This is an action potential. The behavior of V, m, h and n during an action potential is shown in Fig. 2A. The sudden rise of the action potential is caused by a rapid increase of the variable m which activates an inward sodium current. The action potential is terminated when the variables h and n, which adjust more slowly to the change in membrane potential, shut off the sodium current and activate an outward potassium current respectively. The final recovery involves the readjustment of m, h and n back to their original values.

The Hodgkin–Huxley equations involve four time-dependent dynamical variables and this makes it difficult to visualize and intuitively understand the workings of the model. There is a conceptual advantage in reducing the model to two dynamic variables so that the phase space can be depicted in a straightforward manner (FitzHugh, 1961; Krinskii and Kokoz, 1973; Rose and Hindmarsh, 1982; Rinzel, 1985; Abbott and Kepler, 1990; Kepler *et al.*, 1992). The first step in one such reduction (Abbott and Kepler, 1990; Kepler *et al.*, 1992) is to describe the three gating variables V, m, h and n in terms of potentials so they are in the same units as V. The functions $\overline{m}(V)$, $\overline{h}(V)$ and $\overline{n}(V)$ appearing in equations (15), (16) and (11) are monotonic in V. They can thus be inverted to define 'equivalent potentials' corresponding to the gating variables through the relations

$$m = \overline{m}(V_m) \qquad h = \overline{h}(V_h) \qquad n = \overline{n}(V_n). \tag{25}$$

Figure 2B shows the behavior of these equivalent potentials during a series of action potentials. It is clear that V and V_m follow very similar trajectories as do V_h and V_n. Because of this, the model is approximately two dimensional and we can form combinations of these potentials to build a fairly accurate reduced model. There are several ways of forming combinations of V and V_m and V_h and V_n to produce a two variable version of the Hodgkin–Huxley model. I will present one method that results in a particularly accurate reduction. More complex neuron models can also be treated using these methods (Kepler *et al.*, 1992; Golomb *et al.*, 1992).

We begin by re-expressing the membrane current as a function of the equivalent potentials by writing $F(V, V_m, V_h, V_n) = I_m(V, \overline{m}(V_m), \overline{h}(V_h), \overline{n}(V_n))$. The reduction of the Hodgkin–Huxley model to a two variable

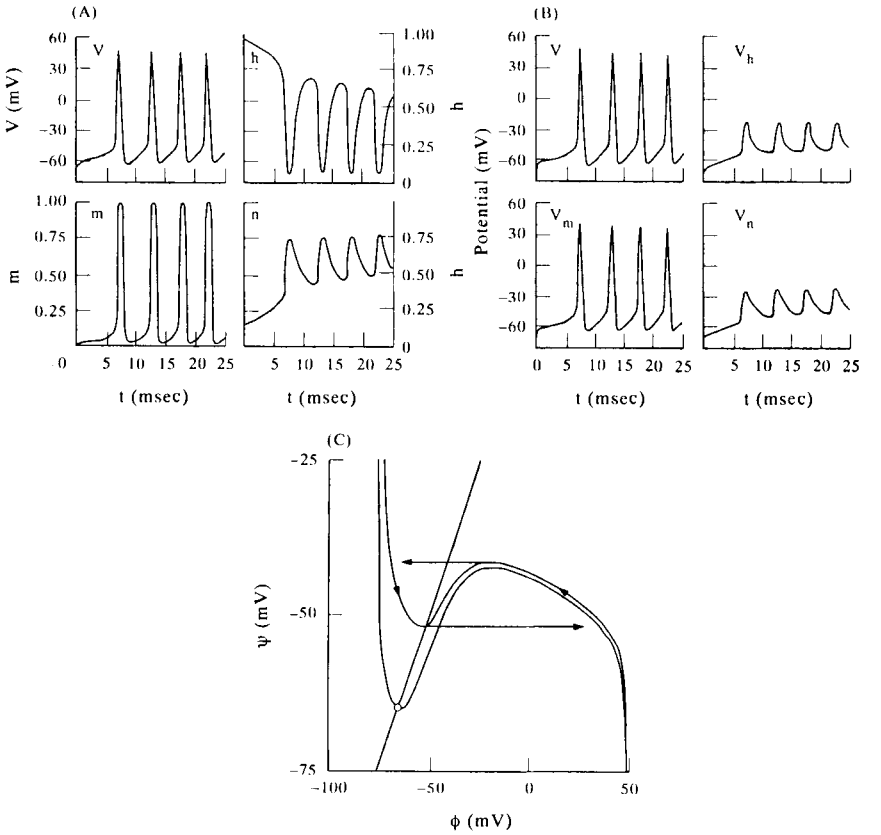


Fig. 2. Reduction of the Hodgkin–Huxley model. (A) V, m, h and n plotted against time while the model is firing a sequence of action potentials. (B) V and the equivalent potenials V_m, V_h and V_n plotted against time while the model is firing a sequence of action potentials. Note that V and V_m and V_h and V_n are similar. (C) Nullclines for the reduced, two-variable system. The lower curved nullcline refers to the zero current case where there is a stable equilibrium point shown by the open circle where the curved and the straight nullclines cross. The upper curved nullcline is for nonzero current when there is a stable limit cycle (indictaed by arrows) representing repetitive firing.

system involves replacing the four variables V, V_m, V_h and V_n with two variables defined by $\phi = yV + (1-y)V_m$ and $\psi = zV_h + (1-z)V_n$. I will not derive the procedure by which we choose the best y and z values to use, I will just state the result. The reader interested in the full procedure can consult Kepler *et al.* (1992). I will define $F_x = \partial F/\partial x$, where x can represent V, V_m, V_h or V_n. In the following equations all expressions involving the function F or its derivatives are to be evaluated with $V = V_m = \phi$ and $V_h = V_n = \psi$. An excellent approximation to the full model can be found if we set

$$z = \frac{F_h}{F_h + F_n} \tag{26}$$

and

$$y = \left[F_V + \frac{C}{\tau_m(\phi)} - \left(\left(F_V - \frac{C}{\tau_m(\phi)}\right)^2 - \frac{4CF_m}{\tau_m}\right)^{1/2}\right]\left(\frac{1}{2(F_V + F_m)}\right). \tag{27}$$

With these choices, the two-variable model is given by the equations

$$C\frac{d\phi}{dt} = y(I_{ext} - F) \tag{28}$$

and

$$\frac{d\psi}{dt} = \frac{z[\overline{h}(\phi) - \overline{h}(\psi)]}{\tau_h(\phi)\overline{h}'(\psi)} + \frac{(1-z)[\overline{n}(\phi) - \overline{n}(\psi)]}{\tau_n(\phi)\overline{n}'(\psi)}. \tag{29}$$

The primes in the last equation refer to first derivatives.

Simulations show that this reduced model is a good approximation of the full model. The easiest way to envision the dynamics given by equations (28) and (29) is to plot the curves (nullclines) $d\phi/dt = 0$ and $d\psi/dt = 0$ in the $\phi - \psi$ phase plane as in Fig. 2C. The straight line in this figure corresponds to $\phi = \psi$ which makes $d\psi/dt = 0$. Along the curved lines $d\phi/dt = 0$. There are two curved lines because two different conditions are plotted. The lower curved nullcline corresponds to $I_{ext} = 0$. The circular point $\phi = \psi = -65mV$ where this curve crosses the straight line is the resting equilibrium point. The higher curved nullcline corresponds to adding an external current sufficient to make the neuron fire. The system now goes into the limit cycle shown by the arrows as it fires a sequence of action potentials. The nullclines and behavior exhibited in Fig. 2C are very typical of nonlinear oscillator models.

Most neurons have many more membrane conductances than the two measured in the squid giant axon by Hodgkin and Huxley (McCormick, 1990; Connors *et al.*, 1982; McCormick *et al.*, 1985; Adams and Gavin, 1986; Buchholtz *et al.*, 1992; Yamada *et al.*, 1989). It is common for neurons to have, for example, one or two sodium conductances, two or three different types of calcium conductances and many different potassium conductances. As a result single neurons can exhibit a wide variety of intrinsic

behaviors even when they are not embedded in a network. The issue of how a neuron regulates such a large number of conductances to produce a particular pattern of activity has recently been the subject of modeling studies (LeMasson *et al.*, 1992; Abbott and LeMasson, 1992).

When many conductances are present, computer simulation is the appropriate method to use for analyzing behavior. The single compartment equations are actually quite easy to integrate numerically due to their minimal coupling. The membrane potential V is affected by all of the gating variables, but these do not directly couple to each other, they only interact through V. Furthermore, the equation for the membrane potential involves nonlinear functions of the gating variables (like m, h and n) but it is linear in V. It can thus be written in the form

$$C\frac{dV}{dt} = D - BV \tag{30}$$

where all the dependence on the gating variables has been absorbed into D and B. For example, in the Hodgkin–Huxley model $D = g_L E_L + g_K n^4 E_K + g_{Na} h m^3 E_{Na}$ and $B = g_L + g_K n^4 + g_{Na} h m^3$. One method for integrating this equation that is stable and second-order accurate is to solve equation (30) exactly over one integration time step Δt treating the gating variables as fixed. Thus, we write

$$V(t + \Delta t) = \frac{D}{B}\left(1 - e^{-B\Delta t/C}\right) + V(t)e^{-B\Delta t/C}. \tag{31}$$

Of course, the gating variables are not fixed, but once V has been updated by the above rule the gating variables can be updated similarly. All of them are treated in the same manner, so take for example m which satisfies

$$\tau_m \frac{dm}{dt} = \overline{m} - m. \tag{32}$$

Again, the method is to integrate this equation exactly treating V as fixed so that

$$m(t + \Delta t) = \overline{m}\left(1 - e^{-\Delta t/\tau_m}\right) + m(t)e^{-\Delta t/\tau_m}. \tag{33}$$

An efficient integration scheme (see Mascagni, 1989) is to alternate using rule (31) to update the membrane potential and rule (33) to update all the gating variables. It is important to alternate the updating of V with that of the gating variables (rather than doing them all simultaneously) as this keeps the method accurate to second-order in Δt. If calcium-dependent conductances are included, the intracellular calcium concentration should be computed simultaneously with the membrane potential. This requires solving the coupled equations for the membrane potential and the calcium concentration, but usually these are linear equations in these two variable so they can be solved easily.

Solving the cable equation

In cases where a single compartment model is insufficient, the full cable equation (5) must be solved (Rall, 1977; Tuckwell, 1988; Jack *et al.*, 1975). The basic problem of cable theory is to compute the membrane potential everywhere on a complex neuron as a function of the external and synaptic currents entering the cell. This is done either analytically or by computer. A commonly used simplification for the analytic approach is to assume that the membrane conductance is constant over the voltage range being considered. In addition, synapses are typically treated as sites of current injection, ignoring the accompanying synaptic conductance changes. This is valid if the synaptic contribution to the total membrane conductance is small. When these assumptions are made, several powerful and practical techniques exist to solve the cable equation for neurons with dendritic trees of arbitrary geometry (Butz and Cowan, 1974; Koch and Poggio, 1985; Holmes, 1986; Horwitz, 1981; Abbott *et al.*, 1991). Among the methods available are diagrammatic rules for computing the Laplace transform of the cable potential (Butz and Cowan, 1974; Koch and Poggio, 1986) and a new method for solving cable theory problems recently developed on the basis of a path integral approach (Abbott *et al.*, 1991; Abbott, 1992; Cao and Abbott, 1992). In addition to these general methods, much work has been done for neurons with restricted dendritic structures (Rinzel and Rall, 1974; Jack and Redman, 1971).

The only method for solving the cable equation that allows a full treatment of the nonlinear voltage-dependent membrane conductances and time-dependent synaptic conductances found in real neurons is compartmentalization (Rall, 1964; Segev *et al.*, 1989) and computer integration (Hines, 1984). Figure 3A shows a sample compartmentalization of a dendritic tree. The compartment sizes must be chosen so that each individual compartment is electrotonically compact and can be described by a single value of the membrane potential. Each compartment has its own membrane potential V_i (where i labels compartments) and its own gating variables producing the membrane current within compartment i, $I_m^{(i)}$. Each membrane potential V_i satisfies an equation like (22) except that compartments couple to their nearest neighbors in the tree. For simplicity, I will treat a single linear cable of constant radius a. After seeing how this case is integrated it should be clear how to treat the more general case of branched and/or tapered cables (Hines, 1984). For a single cable, each compartment is coupled to two neighbors so the equations for the membrane potentials of the compartments have the form

$$C\frac{dV_i}{dt} = -I_m^{(i)} - I_s^{(i)} + g_x(V_{i+1} - V_i) + g_x(V_{i-1} - V_i) \tag{34}$$

except at the ends of the cable where there is only one neighboring compartment. The coupling constant g_x is determined by computing the cur-

rent that flows between two neighboring compartment due to Ohm's law. Recall that r is the resistivity of the intracellular fluid. If the length of the compartments is L, then the resistance between any two compartments (measured from their centers) is the resistivity times the length divided by the cross-sectional area, $rL/\pi a^2$. The total current flowing between compartment $i+1$ and compartment i is then $\pi a^2(V_{i+1} - V_i)/rL$. Recall that in the derivation of the cable equation we divided by the surface area of the compartment $2\pi aL$. Thus, we find that $g_x = a/2rL^2$.

To solve the compartmental equations numerically we must integrate a set of equations, one for each compartment. The gating variables like m can be handled as in the single compartment case using equation (33). The equations for the membrane potentials can be written in the form

$$\frac{dV_i}{dt} = A_i V_{i-1} + B_i V_i + C_i V_{i+1} + D_i. \tag{35}$$

As in equation (30), the gating variable and other dependences have been absorbed into the parameters A, B, C and D. Due to the coupling between compartments, these equations are more difficult to integrate than the single compartment equation (30) and we cannot use the rule (31). Again we want a method that is stable and accurate. The finite difference form of equation (35) would say that the change in V_i at time t is the right side of equation (35) times Δt. Stable (implicit) methods make use of this but evaluate the right side of (35) not at the time t but at a later time $t + \rho\Delta t$. Two methods are predominantly used, the reverse Euler method for which $\rho = 1$ and the Crank-Nicholson method with $\rho = 0.5$. The reverse Euler method is the more stable of the two and the Crank-Nicholson is the more accurate (see Mascagni, 1989). In either case, we write

$$V_i(t + \Delta t) = V_i(t) + \Delta V_i \tag{36}$$

and solve for the ΔV_i to do one step of the integration. To determine ΔV_i we use the finite difference version of (35) with the right side evaluate at time $t + \rho\Delta t$,

$$\Delta V_i = [A_i V_{i-1}(t + \rho\Delta t) + B_i V_i(t + \rho\Delta t) + C_i V_{i+1}(t + \rho\Delta t) + D_i]\Delta t. \tag{37}$$

Now $V_i(t + \rho\Delta t) \approx V_i(t) + \rho\Delta V_i$ so this becomes

$$\Delta V_i = a_i \Delta V_{i-1} + b_i \Delta V_i + c_i \Delta V_{i+1} + d_i \tag{38}$$

where $a_i = A_i \rho\Delta t$, $b_i = B_i \rho\Delta t$ and $c_i = C_i \rho\Delta t$

$$d_i = [D_i + A_i V_{i-1}(t) + B_i V_i(t) + C_i V_{i+1}(t)]\Delta t. \tag{39}$$

This is a set of coupled linear equations for the needed quantities ΔV_i. An efficient method exists for solving these equations (Hines, 1984; Tuckwell

1988). Note that if the cable begins with the compartment $i = 1$, then $a_1 = 0$ and if it ends at compartment N, then $c_N = 0$. The method consists of solving (38) for ΔV_i in terms of ΔV_{i+1} sequentially starting at one end of the cable and proceeding to the other end. For example, if we start the procedure at compartment one, ΔV_1 can be expressed as

$$\Delta V_1 = \frac{c_1 \Delta V_2 + d_1}{1 - b_1}. \tag{40}$$

Substituting this into the equation (38) for ΔV_2 gives

$$\Delta V_2 = \tilde{b}_2 \Delta V_2 + c_i \Delta V_3 + \tilde{d}_2 \tag{41}$$

where $\tilde{b}_2 = b_2 + a_2 c_1/(1 - b_1)$ and $\tilde{d}_2 = d_2 + a_2 d_1/(1 - b_1)$. We now repeat the procedure going down the cable. At each stage we solve for ΔV_{i-1} in terms of ΔV_i finding

$$\Delta V_{i-1} = \frac{c_{i-1} \Delta V_i + \tilde{d}_{i-1}}{1 - \tilde{b}_{i-1}}. \tag{42}$$

and define

$$\tilde{b}_{i+1} = b_{i+1} + \frac{a_{i+1} c_i}{1 - \tilde{b}_i} \tag{43}$$

and

$$\tilde{d}_{i+1} = d_{i+1} + \frac{a_{i+1} \tilde{d}_i}{1 - \tilde{b}_i}. \tag{44}$$

Finally, when we get to the end of the cable we can solve for

$$\Delta V_N = \frac{\tilde{d}_N}{1 - \tilde{b}_N} \tag{45}$$

because $c_N = 0$.

The procedure for finding the ΔV_i is the following. Define $\tilde{b}_1 = b_1$ and $\tilde{d}_1 = d_1$ and iterate equations (43) and (44) down the length of the cable to define all the $\tilde{b}$'s and $\tilde{d}$'s. Then solve for ΔV_N from equation (45) and iterate back up the cable solving for the ΔV's using equation (42). This process takes only $2N$ steps. Figure 3B shows an action potential propagating down the axon of a multi-compartment model neuron computed using this method.

One of the most useful concepts for treating cable problems is the equivalent cylinder concept introduced by Rall (1959, 1977). Suppose we wanted to replace an entire dendritic tree with a single equivalent cylinder, that is, by a single uniform cable as in Fig. 3C. This is, in general, a drastic approximation but if we hope to make it as accurate as possible we need to match the two most important elements in the cable. These are the electrotonic length which determines the amount of attenuation, and the total surface

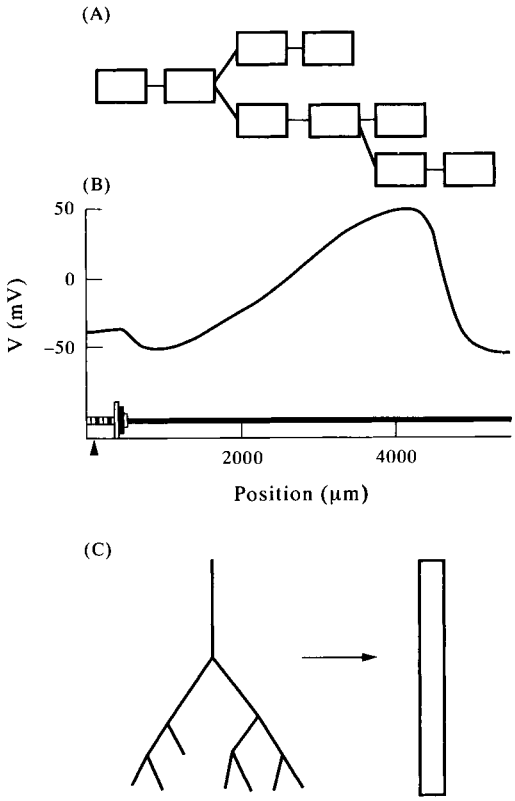


Fig. 3. (A) Compartmentalization of a sample tree. (B) A compartmental simulation of an action potential travelling down the axon of a model neuron. A schematic of the model neuron is shown at the bottom of the plot. The triangle marks the site of simulated synaptic input. An equivalent cable has been used for the dendritic tree. (C) The equivalent cable.

area which determines the total membrane current. The electrotonic length of a cable is the length measured in units of the length constant λ. Suppose a dendritic tree has a total surface area A and an electrotonic length w. For w to be well defined the electrotonic length along every path down the tree must be approximately the same, and indeed this is one of the conditions for the approximation to be valid. We must choose the length L and radius a of the equivalent cable so that $2\pi aL = A$ and $L(aR/2r)^{-1/2} = w$. Thus, we find that $a = (2r/R)^{1/3}(A/2\pi w)^{2/3}$ and $L = (R/2r)^{1/3}(Aw^2/2\pi)^{1/3}$. Under some restrictive circumstances the equivalent cable reproduces the effects of a full tree exactly (Rall, 1959, 1977). However, even when it is not exact, the equivalent cable is an extremely useful and often fairly accurate simplification. Note that in Fig. 3B an equivalent cable was used for the dendritic part of the neuron.

From neurons to networks

How can information on the behavior of individual neurons be used in constructing a model of a large neural network? Most neural network models use the firing rates f_i of the neurons $i = 1, 2, \ldots, N$ to describe the activity of the network (Wilson and Cowan, 1972, 1973; Hopfield, 1984; Frolov and Medvedev, 1986; Abbott, 1991b). The equations for f_i must relate the firing rate of neuron i to the firing rates of other neurons in the network. To construct these equations, we must know how the firing rate of a neuron depends on the firing rates of its presynaptic neurons.

One approach to this problem (Abbott, 1991b) is to start by determining how the firing rate of neuron i depends on the amount of current injected into its soma or cell body. This firing rate function $F_i(I_{ext})$ can be determined easily from a single neuron model and, more importantly, it can be measure experimentally by injecting current into the soma of a neuron through a microelectrode. Typical firing rate functions are zero below some threshold value of the current and then rise like a square root or a logarithm of I_{ext}. If the neuron in question exhibits spike-rate adaptation, F_i may have to be described by one or more differential equations rather than as a static function of the external current I_{ext}. Once the functions F_i are known, the next step is to relate the dynamic firing-rate variables f_i to the firing rate functions F_i. How this is done is a fairly subtle issue (Frolov and Medvedev, 1986; Abbott and Van Vreeswijk, 1992). An approach commonly used (Wilson and Cowan, 1972; Hopfield, 1984; Abbott, 1991b) is to write a set of first-order differential equation with time constants τ,

$$\tau \frac{df_i}{dt} = -f_i + F_i(I_s^{(i)}). \tag{46}$$

In this equation, the argument of the firing rate function is not the external current I_{ext} but instead the amount of synaptic current entering the soma

of neuron i. This is, of course, because in a natural setting neurons receive synaptic currents not currents passing through electrodes. The limitations of equation (46) are quite severe and the reader should not hesitate to improve upon it. For example, many neurons respond more rapidly to a rise in the synaptic current than a fall in it whereas equation (46) prescribes the same time constant τ for both processes.

The final step in building a firing-rate model is to describe how the synaptic current of neuron i, $I_s^{(i)}$, depends on the firing rates of other neurons in the network that synapse onto neuron i. Techniques of cable theory are useful for this step (Abbott, 1991a). This is because the current I_s appearing in (46) is the amount of synaptic current that reaches the soma from synapses located all over the dendritic tree of the neuron as well as those located on the soma itself. This current is not just a simple sum of the currents entering at various synapses because of shunting and other losses through the dendritic tree. The dependence of the synaptic current on the presynaptic firing rates reflects how the geometry of the neuron and the placement of its synapses affect network behavior.

The approach outlined here allows the details of single neuron behavior and geometric structure to be incorporated into network models. In addition to producing more realistic descriptions of large neural populations, information on single neuron dynamics is likely to provide important new mechanisms for solving fundamental network problems.

Acknowledgments

I thank Tom Kepler and Eve Marder for both leading and guiding the way. Research supported by National Science Foundation grant DMS-9208206 and NIMH grant MH46742.

5

An Introduction to Neural Oscillators

BARD ERMENTROUT

Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

Preamble

Oscillations abound in neural systems as hallmarks of behavior. They range from the ubiquitous stereotypical patterns of invertebrate motor patterns, the synchronous flashing of thousands of fireflies, to the 40 Hz oscillations in visual cortex and the large amplitude synchronized activity in epilepsy. In this chapter, I will describe some of our current research on mathematical methods by which one can understand and analyze these behaviors. I will begin by defining what is meant by an oscillation and then discuss the principle mechanisms by which such rhythms are generated. I will primarily be concerned with neuronally generated rhythms although much of what is described holds for many other physiological phenomena such as hormonal rhythms, smooth muscle oscillations and biochemical periodicity. Next I will describe how such oscillators interact with the external world and each other. Concepts such as relative phase and phase response curves play an important role in experimental and mathematical analyses of these interactions. Once interactions have been introduced, I will then describe a number of mathematical techniques by which one can reduce the intractable system of many coupled biophysical oscillators to a more mathematically approachable system. This will constitute a large part of this chapter. In the remaining space, I will describe a particular instance of the general theory outlined above and how it has been applied to understanding averaged oscillators.

What is an oscillator and how can we get one?

In order to avoid a number of complicated issues dealing with stochastic behavior (although they certainly are important, particularly in biology) I will restrict the types of models that I will describe to systems of deterministic autonomous differential equations of the form:

$$dX/dt = F(X) \tag{1}$$

where $X \in R^n$ and F is a continuously differentiable function from R^n to R^n. An *oscillator* will be defined to be any such system that admits as a solution an orbitally asymptotically stable solution, $X_0(t)$ such that $X_0(t+T) = X_0(t)$ for all t and some minimal positive number T which is called the *period* of the oscillator. For completeness, I will define orbital asymptotic stability. We say a solution to eq. (1) is *orbitally asymptotically stable* if for each $\epsilon > 0$ there is a $\delta > 0$ such that if $|X(0) - X_0(0)| < \delta$ then $|X(t) - X_0(t)| < \epsilon$ for all $t > 0$ and that furthermore there exists a ϕ such that,

$$\lim_{t\to\infty} |X(t) - X_0(t+\phi)| \to 0. \tag{2}$$

This value ϕ is called the asymptotic phase shift and will prove to be important later when phase response curves are described. From now on oscillator means a system of differential equations that admits an orbitally asymptotically stable periodic solution which will often be referred to as simply asymptotically stable or limit cycle. As is the case with fixed points of differential equations, we can infer stability of periodic solutions by looking at the linearized equations:

$$\mathcal{L}(t)w(t) \equiv dw(t)/dt - A(t)w(t) = 0 \tag{3}$$

where $A(t)$ is the Jacobi matrix of F evaluated at the solution $X_0(t)$. The matrix $A(t)$ is $T-$periodic so that Floquet theory tells us that the solutions to eq. (3) are of the form:

$$w(t) = e^{\nu t}P(t) \tag{4}$$

where ν are complex scalars and $P(t)$ is a $T-$periodic function. There are n linearly independent solutions to eq. (3) and these have associated with them n values of ν. Since ν are not uniquely determined (the factor $2\pi ik/T$ can be added while still maintaining the periodicity of $P(t)$) one usually considers $\mu = e^{\nu T}$ which are unique and are called the *Floquet multipliers.* Clearly one solution to eq. (3) is

$$w(t) = X_0'(t) \tag{5}$$

the time derivative of $X_0(t)$ which is periodic so that for any periodic solution to an autonomous differential equation, there is always a Floquet multiplier of 1. If the remaining multipliers are inside the unit circle (i.e. $\Re\nu < 0$)

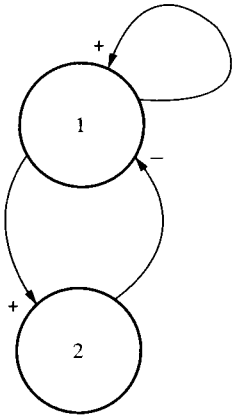


Fig. 1. Schematic showing self-activation and delayed inhibition leading to oscillations.

then the periodic solution is asymptotically stable. The single multiplier of 1 corresponds to the fact that if $X_0(t)$ solves eq. (1) then so does $X_0(t+\theta)$ where θ is any constant. This translational invariance plays a critical role in our later analysis of coupled oscillators. Now that we have defined oscillators, we can ask how they arise in biology, particularly neurobiology. The basic variable for measuring neural activity is the membrane potential. Suppose that at a certain level of depolarization, the potential has positive feedback so that it is further depolarized (the sodium channel and many calcium channels have precisely this property). Suppose furthermore that there is a delayed negative feedback term that forces the potential back down. These assumptions are illustrated in Fig. 1.

Such a qualitative picture is often sufficient for oscillations to occur. I will produce several examples. The first example which you may have already encountered is a model for the membrane potential that is a simplification of the Hodgkin–Huxley equations and has the following form:

$$CdV/dt = g_L(V_L - V) + g_K w(V_K - V) + g_{Ca} m_\infty(V)(V_{Ca} - V) + I \quad (6)$$

$$\tau dw/dt = \phi(V)(w_\infty(V) - w) \quad (7)$$

Here, m_∞ and w_∞ are positive monotone increasing functions of V with minimum and maximum values of 0 and 1 and represent the fraction of open channels as a function of voltage. $\phi(V)$ is some strictly positive function and τ is a time constant for the variable w. In absence of the recovery variable w, the equation above acts as a switch with a high and low potential possible. Since $V_K < 0$ when w is close to 1, the potential is brought down to a low state and when w is zero it is in a higher state. Thus, if τ is large enough then once the potential reaches a maximum, w slowly builds

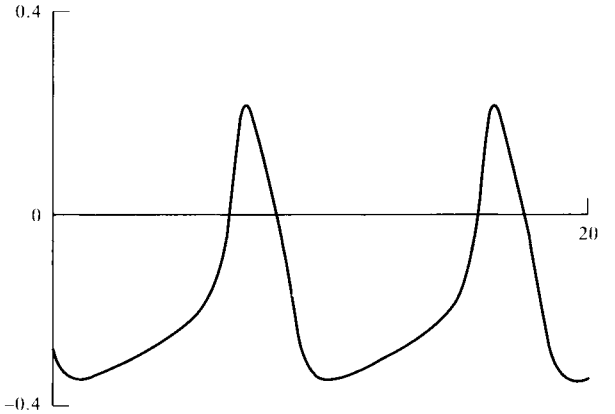


Fig. 2. Solution to the Morris–Lecar model membrane oscillator.

which shuts down V and thus shuts down w. This can lead to oscillatory behavior if the parameters are chosen correctly. Figure 2 shows the potential of a typical oscillation. The key point to emphasize is that the negative feedback is delayed relative to the positive feedback and thus oscillations are possible. This basic mechanism appears to be responsible for almost all oscillations seen in neurobiology. A complete analysis of this simple model and all of its interesting behavior can be found in Rinzel and Ermentrout (1989).

A more direct example where the relevant variable is firing rate is easy to construct. Let $x(t)$ be the number of excitatory neurons firing at any given time and suppose that inhibition directly proportional to this excitation occurs with a time delay τ. Then, the following equation holds:

$$dx(t)/dt = -x(t) + f(ax(t) - bx(t-\tau)) \tag{8}$$

where f is some positive monotone increasing function, a, b are positive parameters, and τ is the delay. In absence of the feedback delay term, this system acts like a switch with a high state near $x = 1$ and a low state near $x = 0$. The presence of the delay allows for the possibility of oscillations. Let $\bar{x}$ be an equilibrium solution satisfying

$$\bar{x} = f((a-b)\bar{x}). \tag{9}$$

The stability is determined by linearizing about this rest state:

$$dy/dt = \alpha y - \beta y(t-\tau) \tag{10}$$

where $\alpha = -1 + af'((a-b)\bar{x})$ and $\beta = bf'((a-b)\bar{x})$. The solutions to the above equation are $y(t) = e^{\lambda t}$ where

$$\lambda = \alpha - \beta e^{-\lambda\tau}. \tag{11}$$

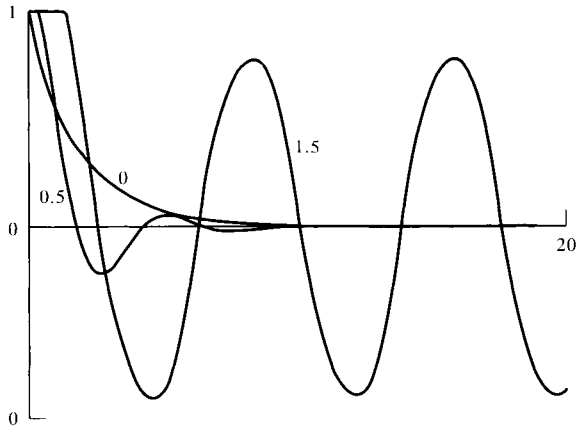


Fig. 3. Behavior for various delays of the delayed feedback model.

If I assume that $\beta > \alpha > 0$ then in absence of delay, the rest is stable so no oscillations occur. Let me now increase the delay. An elementary calculation shows that for $\tau > \tau^*$ where

$$\tau^* = \frac{\tan^{-1}\sqrt{r-1}}{\sqrt{1-1/r}} \tag{12}$$

and $r = \beta^2/\alpha^2$ the equilibrium is unstable and oscillations arise with period close to $2\pi/\sqrt{1-1/r}$. This direct example shows that if the inhibition is sufficiently delayed then there are oscillations. Figure 3 shows the behavior for small and increasingly large delays.

A host of complex behavior is possible in delayed feedback systems (I refer the interested reader to the book by Glass and Mackey). As a final example of a neural oscillations, I will extend the second example to distinct populations of neurons, $x(t)$ representing the excitatory cells and $y(t)$ representing the inhibitory cells. The circuit for these two populations is shown in Fig. 1, where cell 1 is x and cell 2, y. This type of circuit is ubiquitous in cortical structures (see Shepherd 1990; White 1990) and has a vast array of interesting behavior particularly when the network is spatially distributed. The equations governing this little network are:

$$dx/dt = -x + f_e(a_{ee}x - a_{ie}y) \tag{13}$$

$$\tau dy/dt = -y + f_i(a_{ei}x - a_{ii}y) \tag{14}$$

where a_{jk} are non-negative weights and f_j are non-negative monotone nondecreasing functions. In Ermentrout and Cowan (1979) this network was analyzed and it was shown that under some circumstances it acts as an oscillator. Essentially, the main requirements are that the self feedback, a_{ee} be sufficiently strong, the cross feedback $a_{ie}a_{ei}$ be large enough and finally

that the inhibition be sufficiently slow, i.e. τ must be large enough. I will illustrate this last point as it is the point I have been emphasizing throughout this chapter. We will obtain oscillations by destabilizing an equilibrium state. This change in behavior is a well known path to oscillations and is called the *Hopf bifurcation.* Let $(\bar{x}, \bar{y})$ be a rest state for eq. (13). Since this state is independent of τ the inhibitory time constant is a convenient parameter to vary. The linearized system can be written as:

$$du/dt = \alpha_{ee}u - \alpha_{ie}v \tag{15}$$

$$dv/dt = (\alpha_{ei}u - \alpha_{ii}v)/\tau \tag{16}$$

where $\alpha_{ee} = -1 + a_{ee}\bar{f}_e$, $\alpha_{ei} = a_{ei}\bar{f}_i$, $\alpha_{ie} = a_{ie}\bar{f}_e$, and $\alpha_{ii} = 1 + a_{ii}\bar{f}_i$, and $\bar{f}_e = f_e'(a_{ee}\bar{x} - a_{ie}\bar{y})$, $\bar{f}_i = f_i'(a_{ei}\bar{x} - a_{ii}\bar{y})$. We will assume that the self excitation is large enough so that $\alpha_{ee} > 0$ and that the cross feedback is large enough so that $\alpha_{ei}\alpha_{ie} > \alpha_{ii}\alpha_{ee}$. This assumption implies that the determinant of the matrix corresponding to eq. (15) is positive so that the equilibrium is either a node or a focus. Stability is assured if the trace,

$$\alpha_{ee} - \frac{\alpha_{ii}}{\tau} \tag{17}$$

is negative. Clearly since $\alpha_{ii} > 0$ if the inhibition is fast enough (τ sufficiently small) then the trace will be negative and the equilibrium state will be stable. As the inhibition slows down (τ increases) it is also clear that when $\tau > \tau^*$ the rest state will lose stability and oscillations will generally occur. The critical time constant is given by $\tau^* = \alpha_{ee}/\alpha_{ii}$. Figure 4 illustrates the behavior of eq. (13) for various values of τ.

We have seen that self excitation and delayed inhibition will often lead to oscillators. Furthermore, in all of the most common models of neural rhythms, this is precisely the mechanism that leads to their repetitive behavior. A short description of the behavior of such an oscillator when subjected to some perturbation follows. The advantage of the technique described below is that it can be applied experimentally to arbitrary preparations. Let $X_0(t)$ denote the trajectory of an orbitally stable oscillator. Suppose that at time t_0 a brief stimulus is applied to the system (e.g. a brief current pulse). Since the oscillator is stable, if the pulse is not too large, eventually the system will return to its stable oscillating state but there will likely be a phase shift, ϕ. This shift will depend on the strength, duration, timing, and other properties of the stimulus. Assuming the stimulus is fixed, then the only aspect that can vary is the timing of the stimulus, t_0. The result of a series of such experiments is the *phase response curve* (PRC), $\phi(t_0)$ for an oscillator. Figure 5a shows a typical PRC for the flash response of the Southeast Asian firely *Pteroptyx malaccae* which regularly flashes about once per second.

A numerically determined PRC for the model in eqs (6) and (7) is shown in Fig. 5b. Notice a particular property of the PRC. For excitatory stimuli

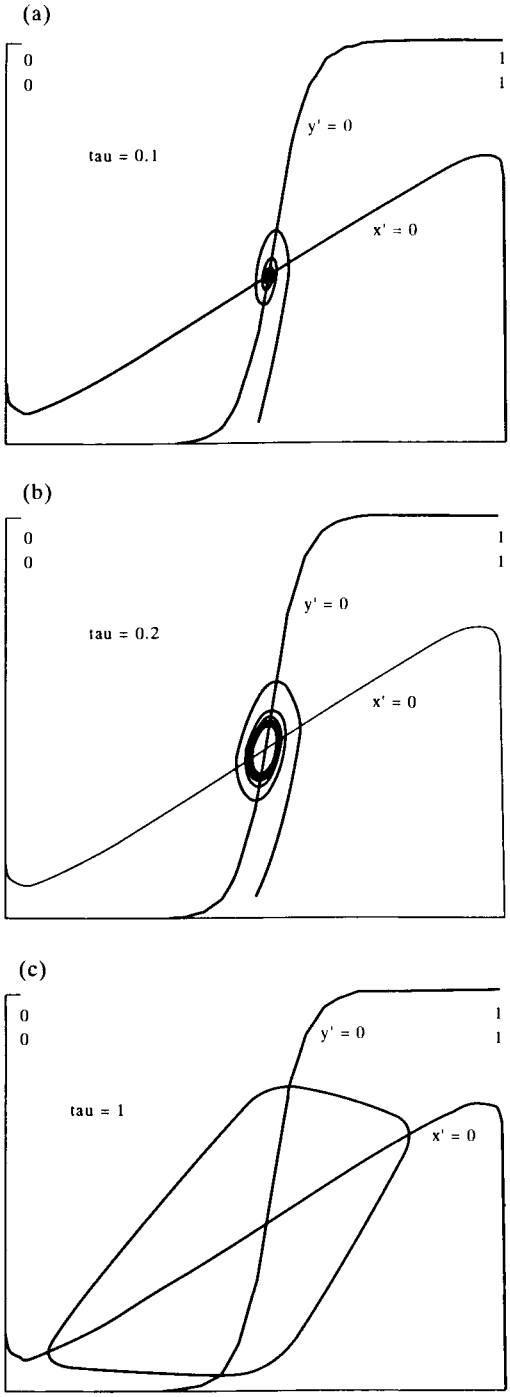


Fig. 4. Hopf bifurcation in phase-plane for Wilson–Cowan model.

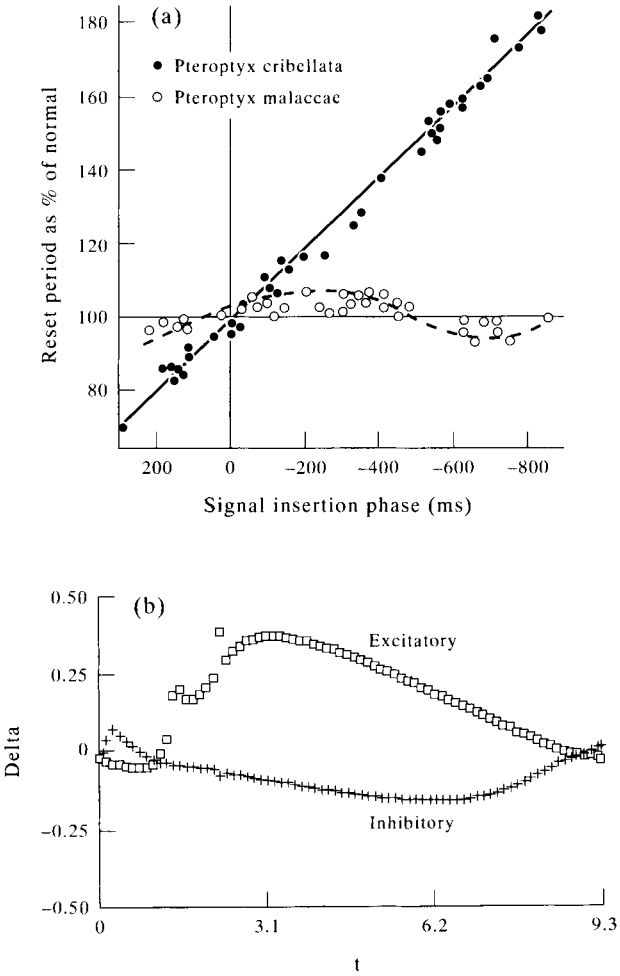


Fig. 5. (a) PRC for firefly; (b) PRC for Morris–Lecar.

that occur right after the cell fires, the phase is retarded while for inhibitory stimuli it is advanced. Another way to view this is that for excitatory stimuli, if the stimulus comes right after the cell has fired the cell will delay its next firing in order to "try" to fire coincidently with the stimulus. This is typical of excitatory stimuli; they try to synchronize with the stimulus. In fact, an "excitatory" stimulus will be defined as one which induces the cell to try to synchronize with the stimulus. On the other hand, inhibitory stimuli attempt to push the cell away from synchrony toward a state in which the cell fires half a cycle away from the stimulus in what is called "antiphase". For this particular model, it is instructive to note that it is very hard to phase delay with excitatory stimuli and hard to phase advance with inhibitory pulses. Strogatz and Mirollo (1991) have used a strict phase advance model for the PRC of fireflies. While this is not a bad approximation for the Morris–Lecar model, Fig. 5a shows that it is not a particularly good model for the firefly. The PRC holds a clue to the behavior of the oscillator when it is periodically driven by an external pulsatile stimulus of the same type as generated the PRC. That is, suppose the PRC of the firefly is obtained via a short pulse of light. Then the effect of repeated pulses can be determined as long as they are sufficiently far apart so that the oscillator can return to its periodic orbit. This last point is not a minor one, for the use of PRCs in forcing experiments assumes that the relaxation back to the oscillation is fast for otherwise when the next stimulus comes, the variables comprising the oscillator will not be on their limit cycle and the results of the PRC experiment will be invalid. For many biological oscillators, the time to relax back to the limit cycle is fast so that PRC methods are often good predictors of the periodically driven preparation. In terms of the previous discussion, the Floquet multipliers are very small in magnitude for this implies that perturbations of the oscillator quickly damp out. Suppose that the PRC has been computed and it is given by the function $\phi(t)$ where t is the time of the stimulus relative to some initial phase in the oscillation which has period T. We can define the phase of the oscillator to be $t/T \equiv \theta$ so that it lies between 0 and 1. Let P be the times between repeated stimuli. Suppose that the oscillator is at phase θ when the first stimulus comes. Then its new phase after the stimulus is $\theta + \phi(\theta)$ since ϕ measures the shift in phase due to a stimulus. P units of time later the stimulus comes again so that the oscillator will have move by an amount P and its current phase right before the next stimulus is $\theta + \phi(\theta) + P$. We thus have defined a map:

$$\theta' = \theta + P + \phi(\theta) \qquad \text{mod } 1. \tag{18}$$

The behavior that we are most interested in is called *phaselocking* by which we mean that the forced (or later, coupled) system has a periodic solutions. In particular, we are interested in so called 1:1 locking in which the oscillator traverses one cycle for each cycle of the periodic stimulus (or, in the case of coupled oscillators; each oscillator traverses one cycle). To see if the oscillator will 1:1 phaselock to the stimulus, θ should advance by one cycle

for every stimulus. That is θ' should equal $\theta + 1$. Substituting this into eq. (18) we see that the condition for 1:1 locking is that

$$1 - P = \phi(\theta) \tag{19}$$

for some value of θ. Since the PRC is a bounded periodic function, eq. (19) puts constraints on the periods to which locking can occur: $1 - \max(\phi) < P < 1 - \min(\phi)$. In order for the solution to be stable, we must have additionally that $|1+\phi'(\theta)| < 1$. An example will illustrate both the stability and the constraints on 1:1 locking. Suppose that $\phi(\theta) = -\alpha \sin(2\pi\theta)$. Then as long as $P \in (1-|\alpha|, 1+|\alpha|)$ a 1:1 locked solution will exist. Note that for P within this range, there are two possible roots to eq. (11) corresponding to values of θ for which $\cos(\theta)$ is positive or negative. Stability is assured if $|1 - 2\pi\alpha\cos(2\pi\theta)| < 1$. Consider entrainment at the same period of the oscillator $P = 1$. Then the two roots of eq. (19) are 0 and 1/2. If $|\alpha| < 1/(2\pi)$ and $\alpha > 0$ then the in phase or synchronized solution is stable. Recall that our "definition" of excitation is that stimuli that precede the peak of the oscillator speed it up. If the peak of the oscillator is $\theta = 0$ then if the stimulus precede the peak, $\theta < 0$ and $\phi(\theta) > 0$ so that $\alpha > 0$ corresponds to excitatory stimulus. In contrast if $\alpha < 0$ which corresponds to inhibitory stimulus, then as expected the out-of-phase solutions, $\theta = 1/2$ is stable. This verifies our intuition that stimuli that are "excitatory" lead to stable in phase solutions and those that are "inhibitory" lead to stable antiphase solutions. For stronger stimuli, more complex effects such as chaotic behavior and Arnold tongues occur (Winfree, 1980; Glass and Mackey, 1988). The use of PRCs is pretty much limited to periodic forcing of nonlinear oscillators and is generally hard to apply when one is interested in the behavior of coupled systems. The reason for this is that a PRC is intrinsically a map and as such requires that there be a "Zeitgeber" or timekeeper. When there are many coupled oscillators, all of whom are influencing each other, there is no obvious choice for the one that determines the timing. In a periodically driven system, this is obvious. Thus, we now turn our attention to the second part of this chapter; coupling of neural oscillators.

Coupling of oscillators

Neurons and neural regions communicate to other cells and regions through a variety of means. The simplest and the means that is seen in early development is through gap junctions. Transport of ions or potential is done via diffusion resulting in a coupling term that depends on the difference between the two coupled entities. The most common communication is via chemical synapses and it is this mechanism that is described in detail. Other ways are via pooled substances that can accumulate in the extracellular space such as ions (Ca^{2+}, K^+, etc) and second messengers (e.g. NO, cAMP).

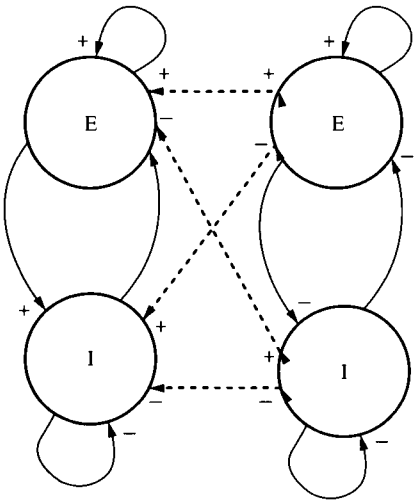


Fig. 6. Coupled Wilson–Cowan net.

Rinzel and Stokes (1992) describe the role of potassium accumulation in synchronizing β cells in the pancreas. If we can understand how a pair of oscillators is coupled, then the extension to many such units in a variety of connection topologies is clear.

Connectionist coupling

I will first consider a connectionist or neural-net type of oscillator such as described by eq. (13) since one need not discuss in detail the mechanisms underlying chemical synapses at this level. Let x_j, y_j denote the activities of two such circuits where $j = 1, 2$ and allow them to be coupled as in Fig. 6. Then the equations are:

$$dx_j/dt = -x_j + f_e(a_{ee}x_j - a_{ie}y_j + \beta_{ee}^{kj}x_k - \beta_{ie}^{kj}y_k) \quad (20)$$

$$\tau dy_j/dt = -y_j + f_i(a_{ei}x_j - a_{ii}y_j\beta_{ei}^{kj}x_k - \beta_{ii}^{kj}y_k) \quad (21)$$

where $k \neq j$. The coupling between the two oscillators is determined by the parameters β^{kj} which are all non-negative. This type of coupling is direct and represent the interaction of the activities between two groups of neurons. Equation (20) has many complex behaviors which are beyond the scope of this chapter (see Khibnik *et al.*, 1991, for a numerically determined bifurcation picture of this model). Obviously when many such circuits are coupled the complexity of the resulting system is manifold.

Generally, additional coupling terms are *added* within the nonlinearity with a positive sign if the presynaptic cell is excitatory and negative if it is inhibitory. Later in this section, several techniques will be introduced

which will allow us to better understand the behavior of systems such as eq. (20) under certain circumstances.

Gap junctions

Let us next consider the membrane models of the form described by eqs (6) and (7). If there are two such oscillatory cells coupled via a gap junction, then the result is equivalent to a resistive junction between the two potentials. This adds a term to eq. (6) of the form:

$$\frac{\Delta V}{R_{gap}} \tag{22}$$

where ΔV is the potential drop between the two cells ($V_{pre} - V_{post}$) and R_{gap} is the resistance of the gap junction. We call such coupling "diffusive" as it is related to the discrete approximation to the diffusion equation. The first important point to note is that if the two cells are equipotential than they are effectively decoupled; no current flows between them. This contrasts with the direct coupling in eq. (20) where even if the cells are synchronized, the effect of the coupling is felt. The consequences of "diffusive" coupling are that one possible solution to a pair of so coupled identical cells is the *synchronous* solutions, $V_1(t) = V_2(t), w_1(t) = w_2(t)$. Synchrony is always a possible solution to two symmetrically coupled identical cells whether or not the coupling is diffusive, however with diffusive coupling the synchronous solution is the same as the solution to the uncoupled cells. (For direct and synaptic coupling, it is possible to destroy the oscillation if the coupling is strong enough — this is called oscillator death — and cannot happen in diffusively coupled identical cells.) The obvious question to ask is whether the synchronous solution is a stable solution to the coupled system. The intuitive notion that diffusion somehow is able to smooth things out would indicate that the answer to this question is yes. But of course as so often happens, intuition fails us here and the answer is an equivocal "sometimes it is stable." Sherman and Rinzel (1992) have recently published a short article showing that two neuron-like oscillators coupled with a gap junction do not have a stable synchronous solution under some circumstances and instead have an out-of-phase solution which is stable. For a different class of oscillators Aronson *et al.* (1990) show instability of the in-phase oscillation for two diffusively coupled identical cells. In both of these examples, it is shown that as the coupling gets stronger the synchronous state is stable. A general but for neurons very unphysical result is the following theorem. theorem Let $X'(t) = F(X(t)), X \in R^n$ have as a solution an asymptotically stable periodic solution $X_0(t)$. Let $d > 0$ be a real number. Consider the coupled system:

$$X'(t) = F(X(t)) + d(Y - X) \tag{23}$$

$$Y'(t) = F(Y(t)) + d(X - Y) \tag{24}$$

Then $X(t) = Y(t) = X_0(t)$ is an orbitally asymptotically stable solution. This is unphysiological because the coupling is scalar indicating (in our neural example) an interaction between the two recovery variables w_1, w_2 that is through diffusion. Since these represent fractions of open channels, it makes no sense to coupled them by diffusion. Thus, the reason that the synchronous solution is not stable in some circumstances is that the diffusion is not scalar. For models of the type in equations (**??**) and (**??**) the following holds (D. Terman, personal communication.) theorem Suppose that two systems of the form shown in eqs (6) and (7) are coupled via a term of the form shown in eq. (22) and suppose that $\phi(V) > \phi_0 > 0$ for all V. Then if R_{gap} is sufficiently small, the synchronous state is stable.
The proof is straightforward. Let $g = 1/R_{gap}$. First one writes down the linearization of the 4 dimensional system

$$d/dt \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} a(t) - g & -b(t) & g & 0 \\ c(t) & -d(t) & 0 & 0 \\ g & 0 & a(t) - g & -b(t) \\ 0 & 0 & c(t) & -d(t) \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix}. \tag{25}$$

Then, split this into a symmetric and an antisymmetric subsystem. Since we assume that the uncoupled oscillator is stable, this implies that solutions to the symmetric subsystem are stable (as its linearization is the same as that of a single uncoupled cell). The antisymmetric subsystem is two-dimensional and has the form:

$$d/dt \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} a(t) - 2g & -b(t) \\ c(t) & -d(t) \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}. \tag{26}$$

We must show that for g sufficiently large, all solutions tend to zero. Consider the function $V = (x_1^2 + y_1^2)/2$. This is positive definite and

$$dV/dt = (a(t) - 2g)x_1^2 - d(t)y_1^2 + (c(t) - b(t))xy. \tag{27}$$

For g sufficiently large, this function is strictly negative for all t thus, $V(t)$ decreases and must tend to zero.

Chemical synapses

Now we turn to models for synaptic interactions. There is no generally accepted model for synapses, the simplest model is described. A synapse can be regarded as a time varying conductance that is added to the right-hand side of the equation for the membrane potential (4). If V_{rev} is the reversal potential for the synapse, then the synaptic current is given by:

$$I_{syn}(t) = \alpha(t)(V_{rev} - V_{post}(t)) \tag{28}$$

where $\alpha(t)$ is the conductance of the channel and $V_{post}(t)$ is the postsynaptic potential. The major differences in models regard the function $\alpha(t)$. The following model is used for this conductance:

$$\tau_1\tau_2 d^2\alpha/dt^2 + (\tau_1 + \tau_2)d\alpha/dt + \alpha = \bar{g}_{syn} m_{syn}(V_{pre}(t)) \tag{29}$$

where τ_1, τ_2 are time constants, $\bar{g}_{syn}$ is the maximum value of the conductance and $m_{syn}(V)$ is a non-negative increasing function of the presynaptic potential, V_{pre} with maximum 1. For "instantaneous" synapses, $\tau_1 = \tau_2 = 0$, and for a simple exponential, $\tau_2 = 0$. The synapses are said to be *slow* if τ is very large, they are *fast* if τ is small, and they are *weak* if $\bar{g}_{syn}$ is small. These extremes can be mathematically exploited in order to *reduce* a coupled system to a more tractable model. I will finally assume that synapses do not interact with each other, thus a neuron receiving multiple synapses will simply add the resulting synaptic currents to its potential. The importance of spatial organization of the synaptic inputs and the role of dendrites has been neglected. This can have a major effect; a distal excitatory synapse can act as effective inhibition if both the pre- and postsynaptic cells are oscillating. Summarizing the discussion so far, we note that coupled membrane oscillators have terms of the form:

$$\sum_k g_k(t)(V_k - V(t)) \tag{30}$$

added to the potential equation where $g_k(t)$ depends on the time course of the presynaptic voltage and $V(t)$ is the postsynaptic potential; or if the coupling is through gap junctions, of the form:

$$(V_{pre}(t) - V_{post}(t))/R_{gap}. \tag{31}$$

Connectionist or neural network models are coupled through the direct addition of weighted activities as in eq. (20). Equation (13) can be rewritten in terms of the potential as:

$$dE/dt = -E + a_{ee}f_e(E) - a_{ie}f_i(I) \tag{32}$$

$$\tau dI/dt = -I + a_{ei}f_e(E) - a_{ii}f_i(I) \tag{33}$$

which means that coupling terms are just added to the right hand side and thus this form of equation parallels that of the membrane models.

Reduction of oscillators

I want to turn to the reduction of models. The following sections discuss four simple reductions and the assumptions inherent in them (i) weak coupling, (ii) strong attraction, (iii) slow coupling, and (iv) coupling near a

Hopf bifurcation. The first two reductions lead to so called phase-models in which the only relevant variable is phase and can be applied to membrane models and neural network models. The third reduction is applied to membrane models to yield a connectionist type model to a first approximation. The last reduction yields a system of equations involving phase and magnitude.

Weak coupling

Consider a pair of oscillators described by vectors in R^n, $X_1(t)$ and $X_2(t)$ satisfying:

$$dX_1/dt = F_1(X_1, X_2, \epsilon) \tag{34}$$
$$dX_2/dt = F_2(X_2, X_1, \epsilon) \tag{35}$$

where $0 < \epsilon << 1$,

$$F_1(X, Y, 0) = F_2(X, Y, 0) \equiv F(X) \tag{36}$$

and

$$dX/dt = F(X) \tag{37}$$

has a limit cycle solution $X_0(t)$. This says that when ϵ is zero, the two systems are uncoupled and behave as independent oscillators. Equations for their behavior for small ϵ will be derived using a series method and the Fredholm alternative. The technique used is equivalent to the method of averaging but it is computationally (and conceptually) simpler. In the appendix of Ermentrout and Kopell (1991) we show this equivalence. Before plunging on, I want to introduce a simple technical point. Let $\mathcal{L}(t)$ be a linear T-periodic operator with a one-dimensional nullspace spanned by, say, $\eta(t)$ and consider the following equation:

$$\mathcal{L}(t)y(t) = f(t) \tag{38}$$

where $f(t)$ is T-periodic. Then there is a $T-$periodic solution, $y(t)$ if and only if:

$$< f(t) \cdot \eta^*(t) > \equiv \int_0^T f(t) \cdot \eta^*(t)dt = 0 \tag{39}$$

where $\eta^*(t)$ is the T-periodic solution to $\mathcal{L}^*(t)\eta^*(t) = 0$, the adjoint equation. (Recall that for a linear operator of the form

$$\mathcal{L}(t) \equiv d/dt - A(t), \tag{40}$$

the adjoint is given by

$$\mathcal{L}^*(t) = -d/dt - A^T(t) \tag{41}$$

where A^T is the transpose of A.) Without loss of generality, η^* can be specified by requiring that

$$< \eta(t) \cdot \eta^*(t) >= 1. \tag{42}$$

With this technical point (called the Fredholm alternative) we can proceed. I will look for solutions to (15) of the form:

$$X_j(t) = X_0(t + \theta_j(s)) + \epsilon w_j \tag{43}$$

where $s = \epsilon t$ is a slowly varying phase and w_j is perpendicular to the derivative of X_0. Substitution of eq. (43) into eq. (34) and setting ϵ to zero results in:

$$\mathcal{L}(t) \equiv (d/dt - A(t+\theta_j))w_j = -X_0'(t+\theta_j)d\theta_j/ds + G_j(X_0(t+\theta_j), X_0(t+\theta_k)) \tag{44}$$

where $A(t)$ is the Jacobian of F evaluated at $X_0(t)$, G_j is the derivative of F_j with respect to ϵ at $\epsilon = 0$ and $k \neq j$. Since w_j is required to be periodic, we must have that the right hand side is orthogonal to the adjoint eigenfunction since the linear operator, $\mathcal{L}(t)$ has a one-dimensional nullspace spanned by $\eta(t) = X_0'(t)$. Application of the Fredholm alternative to equation (44) results in the equations

$$d\theta_j/ds = H_j(\theta_k - \theta_j) \tag{45}$$

where

$$H_j(\phi) = \int_0^T \eta^*(t) \cdot G_j(X_0(t), X_0(t+\phi))dt. \tag{46}$$

The dynamics of the large coupled system reduces to a pair of differential equations on the two-dimensional torus. The crucial point to note is that all of these functions are easily computed numerically and so one can get a representation of these coupling functions for a variety oscillators. Let me illustrate some examples for synaptic and gap junctional coupling of the Morris–Lecar model. The membrane satisfies eqs (6) and (7) and the synaptic current has the form:

$$I_{syn}(t) = \epsilon(\alpha(t)(V_{rev} - V_{post}(t)) + d(V_{pre}(t) - V_{post}(t))) \tag{47}$$

where

$$\tau d\alpha(t)/dt + \alpha(t) = g_{syn}(V_{pre}(t)). \tag{48}$$

This does not fall precisely into the mold of eq. (34) however, we can integrate eq. (48) to obtain:

$$\alpha(t) = \frac{1}{\tau}\int_0^\infty \exp(-s/\tau)g_{syn}(V_{pre}(t-s))ds \tag{49}$$

so that the coupling is of the required form. We need only multiply the coupling by the adjoint and integrate over t to obtain:

$$H(\phi) = H_{syn}(\phi) + H_{gap}(\phi) \tag{50}$$

where

$$H_{gap}(\phi) = d\int_0^T \eta_1^*(t)(V_0(t+\phi) - V_0(t))dt \tag{51}$$

and

$$H_{syn}(\phi) = \int_0^T \eta_1^*(t)(1 - V_0(t))\frac{1}{\tau}\int_0^\infty \exp(-s/\tau)g_{syn}(V_0(t-s+\phi))ds\ dt. \tag{52}$$

Here V_0 is the voltage component of the uncoupled oscillator and η_1^* is the first component of the adjoint. We can rewrite eq. (52) as:

$$H_{syn}(\phi) = \frac{1}{\tau}\int_0^\infty \exp(-s/\tau)H_0(\phi - s)ds \tag{53}$$

where

$$H_0(\phi) = \int_0^T \eta_1^*(t)(1 - V_0(t))g_{syn}(V_0(t+\phi))dt \tag{54}$$

is the coupling function for an instantaneous synaptic connection. In fact, this trick is very nice since it tells us that any linear functional (such as a finite rise time for the synapse or a synaptic delay) commutes with the averaging operator so that if we know the coupling function for instantaneous coupling, we can obtain it for coupling with temporal dynamics by convolving it with the kernel for the dynamics. Thus, an implicit synaptic delay does nothing more than phase shift the coupling function. These coupling functions for the Morris–Lecar model are illustrated in Fig. 7.

Gap junctional coupling, excitatory instantaneous synaptic coupling and inhibitory instantaneous synaptic coupling are shown. Note that both excitatory and gap junctional coupling have positive slopes at the origin while inhibitory coupling has a negative slope. A synaptic delay will translate these in such a way that it is possible for the slopes to change sign. The sign of the slope of the "H" functions is important for stability. For the neural network oscillators with general coupling of the form in eq. (20) and such that each coupling term is premultiplied by ϵ then the coupling function is a sum of 4 terms:

$$H(\phi) = \beta_{ee}H_{ee}(\phi) + \beta_{ie}H_{ie}(\phi) + \beta_{ei}H_{ei}(\phi) + \beta_{ii}H_{ii}(\phi) \tag{55}$$

where

$$H_{ee}(\phi) \quad = \quad \int_0^T \eta_1^*(t)F_e(t)X_0(t+\phi)dt \tag{56}$$

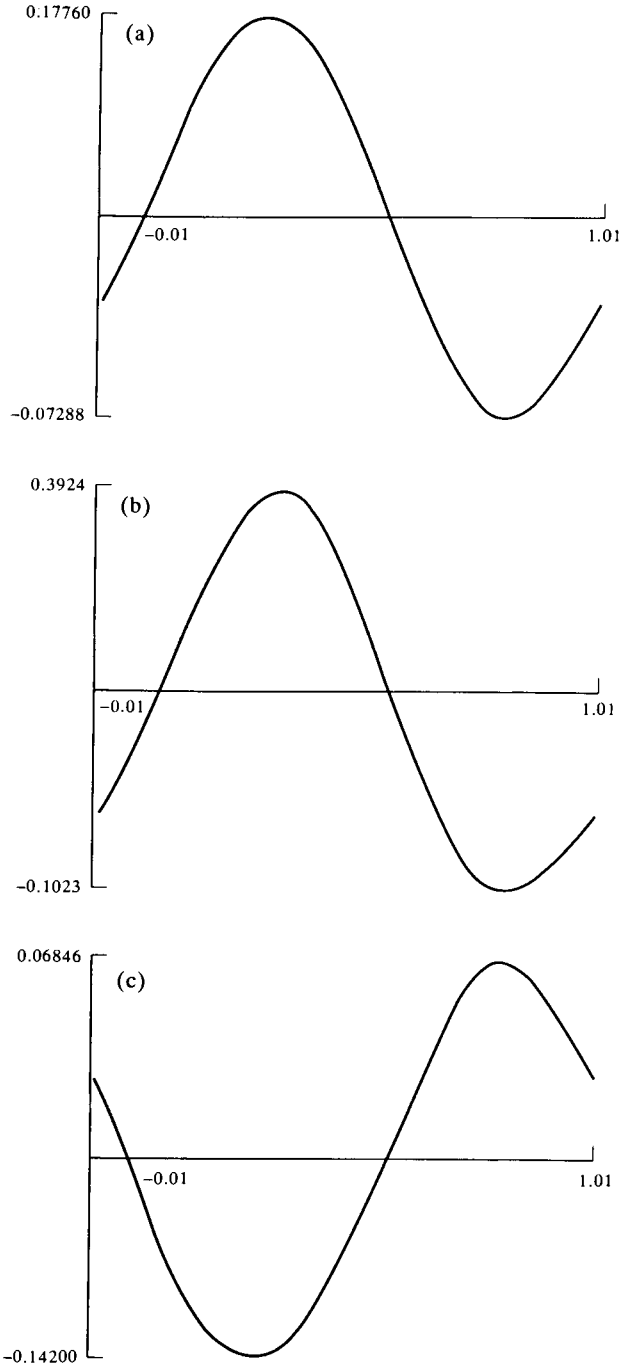


Fig. 7. Coupling functions for Morris–Lecar. (a) Weak gap junctions; (b) weak excitatory coupling; (c) weak inhibitory coupling.

$$H_{ie}(\phi) = -\int_0^T \eta_1^*(t)F_e(t)Y_0(t+\phi)dt \tag{57}$$

$$H_{ei}(\phi) = \int_0^T \eta_2^*(t)F_i(t)X_0(t+\phi)dt \tag{58}$$

$$H_{ii}(\phi) = -\int_0^T \eta_2^*(t)F_i(t)Y_0(t+\phi)dt. \tag{59}$$

Here, X_0, Y_0 are the components of the limit cycle to (13), η_1^*, η_2^* are the components of the adjoint, and

$$F_e(t) = f_e'(a_{ee}X_0(t) - a_{ie}Y_0(t)) \tag{60}$$

$$F_i(t) = f_i'(a_{ei}X_0(t) - a_{ii}Y_0(t)). \tag{61}$$

It is curious to note that for the parameters chosen for this figure, the excitatory–excitatory coupling has a negative slope at the origin. I will return to this point later.

Strong attraction

The attraction of the limit cycle is now supposed to be very strong; that is the Floquet multipliers are all very close to zero. It is easiest to work with additive coupling and since this is reasonable for the membrane and neural network models, we will assume that the pair of coupled oscillators satisfies:

$$dX_j/dt = F(X_j) + G_j(X_k, X_j) \tag{62}$$

where for simplicity, we assume that the oscillators are identical except for added coupling terms. (Note that the additional terms could be applied currents or other conductances so that, in effect, the oscillators could be quite different.) In Ermentrout and Kopell (1991b) we show that as long as the coupling is such that the limit cycle is not *severely* distorted, then there is a change of coordinates

$$X_j(t) = U(\theta_j(t)) + M(\theta_j(t))y_j(t) + O(|y_j|^2) \tag{63}$$

so that in the limit of strong attraction eq. (62) becomes:

$$d\theta_j/dt = 1 + \frac{U'(\theta_j)\cdot G_j(\theta_k, \theta_j)}{|U'(\theta_j)|^2}. \tag{64}$$

Thus, for a strongly attracting system, there is a reduction to a different type of phase model. I want to emphasize that *strong attraction* and *weak coupling* do not lead to the same types of phase models. The latter leads to a model in which the interactions depend only on the phase-difference whereas strong coupling yields far more complex interactions. Let me now

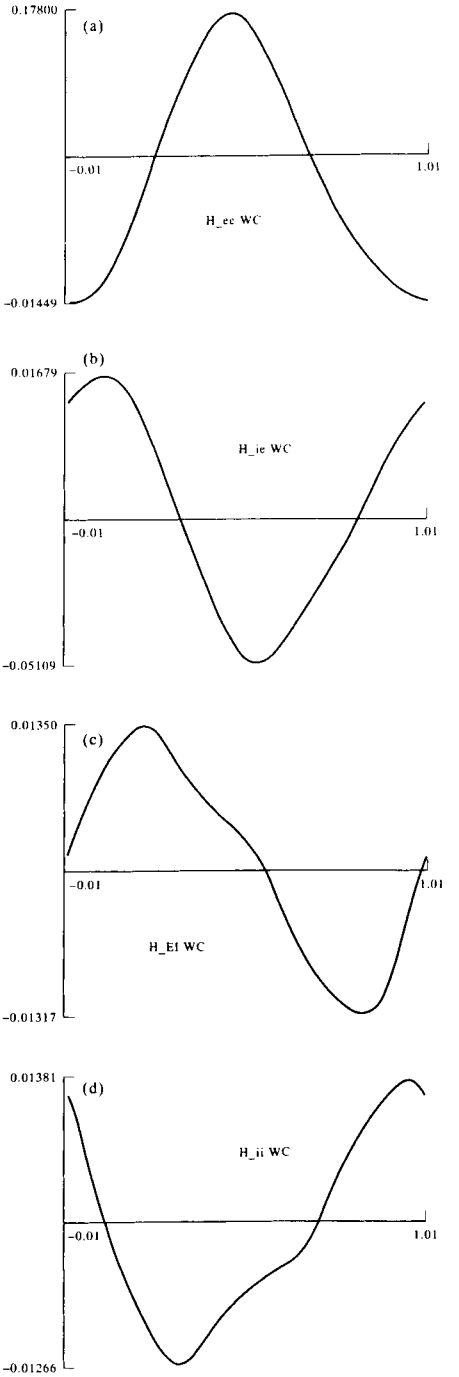


Fig. 8. Coupling functions for the Wilson–Cowan model. (a) EE; (b) IE; (c) EI; (d) II.

apply eq. (64) to a pair of membrane models coupled via one of the two types of synaptic interactions (inhibitory coupling, excitatory coupling). I will let there be an additional current that is different for the two oscillators thus allowing for some intrinsic differences. I will also assume that the synaptic interactions are instantaneous. Using eq. (47) as my interaction, and applying eq. (64) to the Morris–Lecar system, we obtain:

$$\begin{aligned} d\theta_j/dt &= 1+\omega_j(\theta_j) \\ &+ \frac{V_0'(\theta_j)}{V_0'^2(\theta_j)+w_0'^2(\theta_j)} d(V_0(\theta_k)-V_0(\theta_j)) \\ &+ (V_{rev}-V_0(\theta_j))g_{syn}(V_0(\theta_k)). \end{aligned} \tag{65}$$

Here $\omega_j(\theta_j))$ is a term that arises from the additional conductance distingusihing the two oscillators. The synaptic interactions can be rewritten as:

$$\frac{V_0'(\theta_j)}{V_0'^2(\theta_j)+w_0'^2(\theta_j)}(V_{rev}-V_0(\theta)) \times g_{syn}(V_0(\theta_k)) \equiv R(\theta_j)P(\theta_k). \tag{66}$$

It is clear that for multiple synapses, terms of the eq. (66) will simply be added to eq. (65). A single synaptic interaction yields a model that was poularized by Winfree (1967) referred to as a *pulse-response phase model.*

Figure 9 illustrates the pulses and response functions, $P(\theta)$ and $R(\theta)$ for excitatory and inhibitory coupling. Note the similarity of the response function to the phase response curve in Fig. 5b.

Slow synapses

Now I want to consider the consequences of slow synapses. I will only present a cursory analysis of this situation. Rinzel and Frankel (1992) have recently analyzed this case for an excitable/oscillatory system that is of Type II (the onset of oscillations from the rest is a Hopf bifurcation) I will give a short overview of the consequences when the membrane is Type I (onset of periodic orbits of infinite period). The following system of 6 equations will be the subject of my analysis:

$$dV_e/dt = f(V_e,w_e)+a_{ee}g_e(t)(V_{rev}^e-V_e)+a_{ie}g_i(t)(V_{rev}^i-V_e) \tag{67}$$

$$dw_e/dt = g(V_e,w_e) \tag{68}$$

$$dV_i/dt = f(V_i,w_i)+a_{ei}g_e(t)(V_{rev}^e-V_i)+a_{ii}g_i(t)(V_{rev}^i-V_i) \tag{69}$$

$$dw_i/dt = g(V_i,w_i) \tag{70}$$

$$\tau_e dg_e/dt = \epsilon(\bar{g}_e(V_e)-g_e) \tag{71}$$

$$\tau_i dg_i/dt = \epsilon(\bar{g}_i(V_i)-g_i). \tag{72}$$

There is no reason why w could not be a vector of variables and thus that this approach could be applied to more complex membrane models. Similarly, the effects of the synapses on the membranes need not be of the form

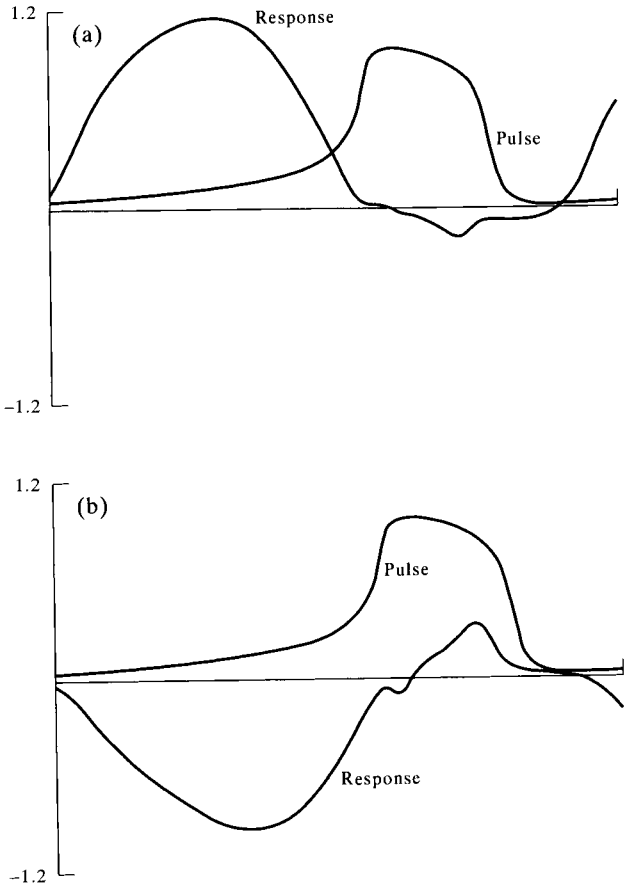


Fig. 9. Pulse-response functions for Morris–Lecar. (a) Excitatory; (b) inhibitory.

above although this is typical. The parameter ϵ is small so that the interactions between the two neurons are quite slow compared to the dynamics of the actual oscillations. The functions f, g represent the uncoupled dynamics of the system. The functions $\bar{g}_e$ and $\bar{g}_i$ are monotone increasing functions of the potential. The quantities V^e_{rev}, V^i_{rev} are the reversal potentials of the excitatory and inhibitory synapses. These equations represent the same network as is shown in Fig. 1 with V_e the potential of the excitatory neuron and V_i the potential of the inhibitory cell. I will now assume that the following holds. There are intervals $J_e = [0, G_e)$ and $J_i = [0, G_i)$ such that if g_e, g_i are held constant and in J_e, J_i respectively, then eqs (67) and (69) have periodic solutions, $V_E(t; g_e, g_i)$ and $V_I(t; g_e, g_i)$ respectively. If I assume that ϵ is very small, then it is possible to substitute V_E and V_I (the fast variables) into eqs (71) and (72) and then average the resultant equations over the periods of their respective oscillations. The result of this calculation is that the slow synapses satisfy:

$$\tau_e dg_e/dt = \epsilon(\Gamma_e(g_e, g_i) - g_e) \tag{73}$$

$$\tau_i dg_i/dt = \epsilon(\Gamma_i(g_e, g_i) - g_i) \tag{74}$$

where

$$\Gamma_e(g_e, g_i) = \frac{1}{T_e}\int_0^{T_e} \bar{g}_e(V_E(t; g_e, g_i))dt \tag{75}$$

$$\Gamma_i(g_e, g_i) = \frac{1}{T_i}\int_0^{T_i} \bar{g}_i(V_I(t; g_e, g_i))dt \tag{76}$$

and T_e, T_i are the respective periods of the oscillators. I have reduced the 6 dimensional system (and even higher dimensions if w is not scalar) to a pair of nonlinear differential equations for the local synaptic gate activity of the excitatory and inhibitory cells. Thus, I have provided a theoretical mechanism for reducing Hodgkin–Huxley type models to neural net models. In order to illustrate a more striking parallel, the interactions have been simplified. Suppose that the reversal potentials of the synapses are bounded away for the potential so that $1 - V(t)$ and $V_k - V(t)$ are nearly constant. Then, we can replace the synaptic terms of eqs (67) and (69) by

$$\mathcal{I}_e \equiv V^e_{rev} a_{ee} g_e + V^i_{rev} a_{ie} g_i \tag{77}$$

$$\mathcal{I}_i \equiv V^e_{rev} a_{ei} g_e + V^i_{rev} a_{ii} g_i. \tag{78}$$

Now for a fixed current, I, let $V_0(t; I)$ be the potential for eqs (67) or (69). For I is some interval, there is a periodic solution. Let

$$f_e(I) = \frac{1}{T(I)}\int_0^{T(I)} \bar{g}_e(V_0(t; I))dt \tag{79}$$

$$f_i(I) = \frac{1}{T(I)}\int_0^{T(I)} \bar{g}_i(V_0(t; I))dt \tag{80}$$

Then, eqs (71) and (72) are transformed into:

$$\tau_e dg_e/ds = -g_e + f_e(c_{ee}g_e - c_{ie}g_i) \quad (81)$$

$$\tau_i dg_i/ds = -g_i + f_i(c_{ei}g_e - c_{ii}g_i) \quad (82)$$

where $s = \epsilon t$ and $c_{ej} = a_{ej}V^e_{rev}$, $c_{ij} = -a_{ij}V^i_{rev}$. These equations are almost exactly the same as eq. (13) and in fact as we will demonstrate below, the functions f_e and f_i are monotone and non-negative. To compare the simplified and the full equations, I must compute the functions f_e, f_i and then approximate them by some simple function. Theoretical calculations (see Rinzel and Ermentrout, 1989) for type I neurons show that the frequency and the magnitude of the average value of the oscillation goes as $\sqrt{(I - I^*)}$ where I^* is the critical frequency at which the repetitive firing begins. Thus, I will attempt to fit f_e, f_i by a function of the form:

$$f(z) = \sqrt{az + bz^2} \quad (83)$$

for $z > 0$ and zero otherwise. a and b will be parameters that are chosen as the best fit. The numerical parameter values of f_e, f_i are computed by integrating the membrane equations for a fixed applied current and then taking the time average over a period or several periods.

Figure 10 shows the behavior of the full system and a comparison of the averaged two-dimensional slow subsystem. Figure 11, sketches the functions f_e and the approximate curve-fit. It is clear that they are monotone and behave like the model described by eq. (13).

It is also clear that the present results can be extended to populations of cells and that there is no reason why the fast oscillators must be identical. Rinzel and Frankel have derived a similar set of equations basing their ideas on a simplified model for bursting suggested by Baer *et al.* The present simulations violate my assumption that there is periodicity for all values of the slow synapses. However, the slow passage from excitable to oscillatory via a saddle-node bifurcation on a circle is well-defined (see Ermentrout and Kopell, 1986) whereas for the Hopf bifurcation case, ramp effects for supercritical bifurcations and hysteresis for sub-critical bifurcations introduce many subtleties. Recently, Kiemel and Frankel (1992) have explored the relative phase of the spikes in two slowly coupled cells. The simple phase-difference models obtained by averaging do not give a complete picture of slow coupling; additional amplitude variables must be introduced.

Coupled Hopf bifurcations

The final reduction that is possible is in the following situation. Suppose that the individual cell is right near the onset of a supercritical Hopf bifurcation. Then, I can ask what the effect of coupling these systems has on the relative phases of the oscillators. Consider the case of several oscillators

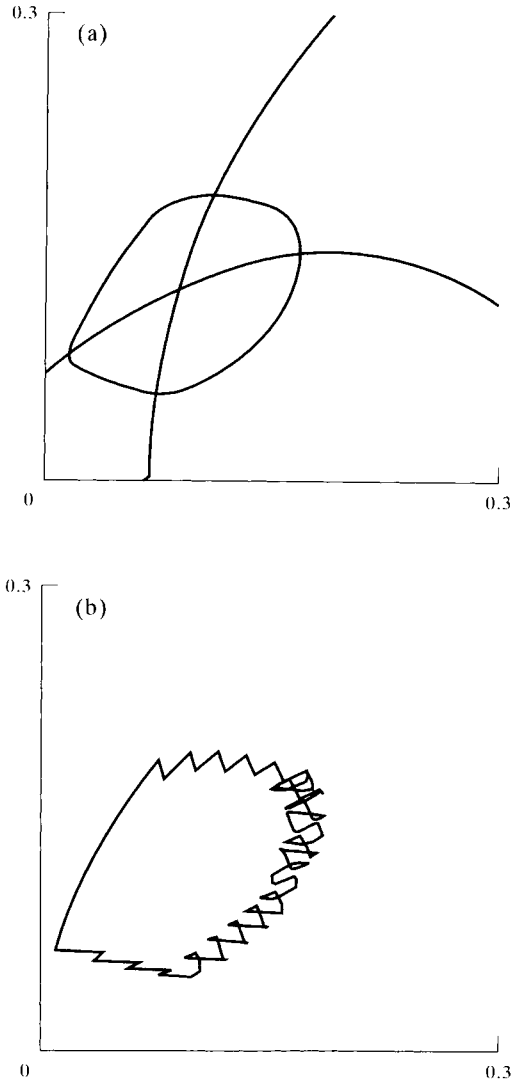


Fig. 10. Comparison of reduced and full models for slow synapses.

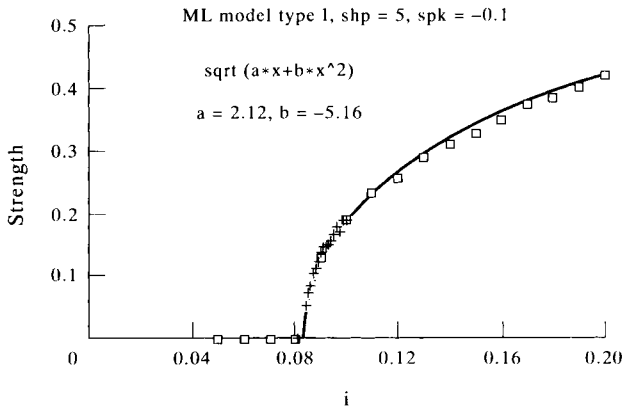


Fig. 11. The numerically computed curves, f and their approximations.

$$dX_j/dt = F(X_j, \lambda) + \eta \sum_k G_{jk}(X_j, X_k). \tag{84}$$

We are interested in when the uncoupled ($\eta = 0$) system undergoes a Hopf bifurcation. Thus we assume with no loss in generality that $F(0, \lambda) = 0$ for all λ and that at, say $\lambda = 0$, the Jacobian of F at $X = 0$, (denoted by A) has a pair of pure imaginary eigenvalues, $\pm i\omega$. Let Φ denote the corresponding eigenvector and let Ψ be the left eigenvector normalized so that $< \Psi, \Phi > \equiv \bar{\Psi}^T \Phi = 1$. We can write $F(X, \lambda)$ as

$$F(X, \lambda) = AX + \lambda BX + Q(X, X) + C(X, X, X) + \cdots \tag{85}$$

where Q is a quadratic form and C a cubic form. Let $A_n = -A + in\omega$. Note that $A_1\Phi = 0$ by definition and A_n is invertible for all $n \neq \pm 1$. Let $\epsilon > 0$ be small and suppose that $\lambda = \epsilon^2\gamma, \eta = \epsilon^2\delta, s = \epsilon^2 t$ so that the coupling is weak and λ is near the bifurcation point. Then, a variety of methods can be used to show that $X_j = \epsilon z_j(s)\Phi \exp(i\omega t) + O(\epsilon^2)$ and that z_j satisfies:

$$dz_j/ds = z_j(\gamma b + q z_j \bar{z}_j) + \delta \sum_k \left((d_1^{jk} + i d_2^{jk}) z_k + (c_1^{jk} + i c_2^{jk}) z_j \right) \tag{86}$$

where

$$\begin{aligned}
b &= < \Psi, B\Phi > && (87)\\
q &= 3 < \Psi, C(\Phi, \Phi, \bar{\Phi}) > + 2 < \Psi, [Q(\Phi, w_0) + Q(\bar{\Phi}, w_2)] > && (88)\\
d^{jk} &= < \Psi, G^2_{jk}\Phi > && (89)\\
c^{jk} &= < \Psi, [G^1_{jk}\Phi + 2Q(\Phi, w^{jk})] > && (90)\\
w_0 &= A_0^{-1} 2Q(\bar{\Phi}, \Phi) && (91)\\
w^{jk} &= A_0^{-1} G_{jk}(0, 0) && (92)\\
w_2 &= A_2^{-1} Q(\Phi, \Phi). && (93)
\end{aligned}$$

Here, G^1, G^2 are the derivatives of the coupling with respect to its first and second arguments. Aronson *et al.*, derived this equation for the special case of diffusive coupling, $G(X,Y) = D(Y - X)$. To connect this model with the phase model approach, we need only convert to polar coordinates, $z_j = r_j \exp(i\theta_j)$ hence, eq. (86) becomes

$$dr_j/ds = r_j(\gamma b_1 + \delta c_j^1 + q_1 r^2) + \delta \sum_k \left(d_1^{jk} r_k \cos(\theta_k - \theta_j) - d_2^{jk} r_k \sin(\theta_k - \theta_j) \right)$$

$$\mathrm{d}\theta_j/ds = \gamma b_2 + \delta c_j^2 + q_2 r^2$$

$$+\delta \textstyle\sum_k \left(d_1^{jk} \frac{r_k}{r_j} \sin(\theta_k - \theta_j) + d_2^{jk} \frac{r_k}{r_j} \cos(\theta_k - \theta_j) \right) \tag{94}$$

The subscripts 1 and 2 denote the real and imaginary parts of the given quantities and $c_j \equiv \sum_k c^{jk}$. A phaselocked solution to eq. (94) is one for which $r_j(t) = \bar{r}_j$, and $\theta_j(t) = \Omega t + \bar{\theta}_j$ similar to the phase models. The advantage of models of this type is that periodic solutions are fixed points of the form $r_j = \bar{r}_j$ and $\theta_j = \Omega t + \bar{\theta}_j$. Thus, the existence of phaselocked solutions is reduced to an algebraic equation. This problem remains to be analyzed in all but the simplest cases. Even the case of two identical diffusively coupled oscillators is capable of complex behavior. Some results on "all to all" coupling have been found however they are for very specific cases (Ermentrout, 1990). The terms q_2 and d_2^{jk} are responsible for much of the difficulty. Kuramoto (1984) has numerically explored this system in two-dimensions and shown complex behavior and turbulence.

Phase difference models

A number of properties of chains of oscillators that have local connectivity and depend only on phase-differences have been described (see Kopell and Ermentrout, 1986, 1990; Kopell *et al.*, 1990, 1991). These arise from the averaging reduction described in the previous section and have been used to understand the central pattern generator of the lamprey. To set the stage for the more general results, let me first describe the behavior of two coupled oscillators. I will then move to chains with localized coupling and finally to completely general coupling topologies. My main interest will be in 1:1 phaselocked solutions.

Two oscillators

Consider a pair of oscillators:

$$d\theta_1/dt = H_1(\theta_2 - \theta_1) \tag{95}$$

$$d\theta_2/dt = H_2(\theta_1 - \theta_2). \tag{96}$$

(Often it is useful to break the functions H_j into two portions representing intrinsic differences in the oscillators and the effects of coupling. Thus, I will often write $H_j(u) = \omega_j + h_j(u)$ where ω_j might arise from slightly different currents in one oscillator from the other.) Let $\phi = \theta_2 - \theta_1$ denote the phase difference between the two oscillators so that ϕ obeys the differential equation:

$$d\phi/dt = H_2(-\phi) - H_1(\phi) \equiv g(\phi). \tag{97}$$

In particular, if $H_1 = H_2$, then $g(\phi)$ is an odd function of ϕ and thus vanishes at 0. If the phase difference between the two oscillators is constant, then they must be phaselocked for they both must traverse their cycles at the same rate, $\Omega = H_1(\phi) = H_2(-\phi)$. The frequency Ω is called the *ensemble* frequency of the coupled system. If this is positive then coupling speeds up the oscillators while if it is negative, coupling slows them down. We see that if the oscillators are identical, then $\phi = 0$ is a rest state for eq. (97) so that as we remarked earlier, the synchronous solution solves the equation. If $H_j(0) \neq 0$ the coupling has an effect on the oscillator when synchronized acting to either slow it down or speed it up. In fact, it is exactly this distinction that is important in contrasting diffusive coupling ($H_j(0) = 0$) with "synaptic" coupling ($H_j(0) \neq 0$.) Returning to eq. (97), we suppose that there is a $\bar{\phi}$ such that $g(\bar{\phi}) = 0$ so that eqs (95) and (96) have a phaselocked solution. Since g is periodic and continuous, this means that there will generically be at least one other root to $g(\phi) = 0$. (The exceptional case occurs when $\bar{\phi}$ lies on a global extremum of g.) We can now ask whether such a root is asymptotically stable. If so, then it can be shown that a phaselocked solution to the full unreduced system exists and is orbitally asymptotically stable. The theory of scalar autonomous differential equations tells us that a sufficient condition for asymptotic stability is that $g'(\bar{\phi}) < 0$ and for instability that $g'(\bar{\phi}) > 0$. For two identical oscillators, the synchronous solution is stable if the slope of H_j is positive at the origin for $g'(0) = -H_1'(0) - H_2'(0)$. By continuity, if there are precisely two distinct roots, then one will be stable and the other unstable. In particular, for two identical oscillators, the other root is a half a cycle out of phase from synchrony and represents the anti-phase solution. From the figures in the previous section, we see that diffusive and excitatory synaptic coupling of membrane oscillators have stable synchronous solutions while for inhibitorily coupled systems, the anti-phase solution is stable. There is nothing in the analysis that forces there to even be a rest state for eq. (97); if no such state exists, then the two oscillators simply move apart from each other at a variable rate and the resulting state is called *drift*. With these preliminaries in mind, I can next address the behavior of groups of oscillators.

Rings of identical oscillators

We briefly consider a ring of N identical oscillators coupled in a translation invariant fashion:

$$d\theta_j/dt = \omega + \sum_{k=0}^{N-1} h_k(\theta_{j+k} - \theta_j). \tag{98}$$

The fact that this is a ring of oscillators implies that $h_{k\pm N}(\phi) = h_k(\phi)$ k. One class of solutions to this problem is the travelling wave solution

$$\theta_j = \Omega_k t + 2\pi jk/N \tag{99}$$

where $k = 0, \ldots, N-1$ is the wave number and Ω_k is the ensemble frequency for that wave number. The plot of Ω_k versus k is called the dispersion relation. The wave number $k = 0$ corresponds to the synchronous solution in which all oscillators are in phase. From eq. (98) it is clear that

$$\Omega_k = \sum_{l=0}^{N-1} h_i(2\pi lk/N). \tag{100}$$

The stability of these solutions is found by linearizing about the equilibrium solution. The result is a translation invariant linear system whose eigenvectors are $\exp(i2\pi jk/N)$ so that the eigenvalues are easily found to be of the form

$$\lambda = \sum_{l=0}^{N-1} h_l'(2\pi lk/N)(\exp(-2\pi isl) - 1). \tag{101}$$

For each wave number k, the real part of λ must be nonpositive for all $s = 0, 1, \ldots, N-1$. There must also be at most one zero eigenvalue corresponding to $s = 0$ and the time-translational invariance of any periodic solution. More details on ring coupling can be found in Ermentrout (1985).

Chains of oscillators

For simplicity of exposition, I will consider only nearest neighbor coupling of a chain of N oscillators although much of what I describe here has been extended to multiply coupled chains (Kopell *et al.*, 1990). I will also separate out the intrinsic frequency differences between oscillators and assume that the coupling is otherwise identical between segments. (This is also unnecessary for the mathematics to work.) Finally, allowing for anisotropy in the coupling, letting h^- denote the effect of oscillators above and h^+ denote the effect of oscillators below. The equations are:

$$d\theta_1/dt = \omega_1 + h^+(\theta_2 - \theta_1) \tag{102}$$

$$d\theta_j/dt = \omega_j + h^-(\theta_{j-1} - \theta_j) + h^+(\theta_{j+1} - \theta_j) \quad (103)$$

$$d\theta_N/dt = \omega_N + h^-(\theta_{N-1} - \theta_N). \quad (104)$$

A *phaselocked* solution to this system is one for which $\theta_j = \Omega t + \bar{\theta}_j$ where $\bar{\theta}_j$ is a constant. Thus, I am interested in solutions to a set of N equations in the $N+1$ unknowns, $\bar{\theta}_j$ and Ω, the ensemble frequency. This system may appear to be underdetermined, however, one can easily see that the replacement of $\bar{\theta}_j$ by $\bar{\theta}_j + c$ where c is any constant is also a solution. Thus, I can assume that $\theta - 1$, e.g. is fixed to be 0. Thus, there are N equations in N unknowns. It is not obvious that there will even be a solution and if so what this solution looks like. However, with very few assumptions, it is possible to not only show that a solution exists, but that it is stable and has a particular form as long as N is large. Suppose that the frequencies ω_j vary regularly, i.e. there is a function $\omega(x)$ such that $\omega_j = \omega(j/N)$. Suppose also that there is an interval J for which $h^{+\prime}(\phi) > 0$ and $h^{-\prime}(-\phi) < 0$. Let $2f(u) = h^+(u) + h^-(-u)$ and let $2g(u) = h^+(u) - h^-(-u)$. In Kopell and Ermentrout (1986, 1990), we have shown that for large N, phaselocked solutions are approximated to within $O(1/N)$ (except possibly at the boundaries) as $\theta_j = \Theta(j/N)$. If I let $\phi(x) = d\Theta(x)/dx$ then ϕ satisfies

$$\Omega = \omega(0) + f(\phi(0)) + g(\phi(0)) \quad (105)$$

$$\Omega = \omega(x) + 2f(\phi(x)) + \frac{1}{N} g(\phi(x))_x \text{ for } 0 < x < 1 \quad (106)$$

$$\Omega = \omega(x) + f(\phi(1)) - g(\phi(0)). \quad (107)$$

The usefulness of the continuum model is that it allows one to easily compute the phaselocked solutions in a variety of interesting cases. For example, a frequency gradient is modelled by letting $\omega(x)$ be some monotone function, a "pacemaker" is modelled by a local increase or decrease in $\omega(x)$ and so on. Suppose that N is large and that the solution has no large gradients. Then the "outer solution" to eq. (106) satisfies:

$$\Omega = \omega(x) + 2f(\phi(x)). \quad (108)$$

Thus, if we know the ensemble frequency, we need only invert f to get the phase gradient and integrate to get the actual phases. However, the ensemble frequency remains unknown and one must look to the boundary conditions. To satisfy equation (105) and setting $x = 0$ in equation (106) and subtracting, I find that,

$$f(\phi(0)) - g(\phi(0)) \equiv h^-(-\phi(0)) = 0. \quad (109)$$

Thus, I choose $\phi(0)$ to be a root of $h^-(-u)$ whence, $\Omega = \omega(0) + 2f(\phi(0))$. Generally, this will not lead to a solution that is compatible with the condition at $x = 1$ so that there will be a boundary layer and the inner equa-

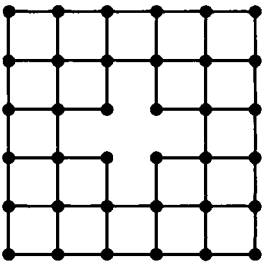


Fig. 12. Graph of the matrix for the coupled 2d array showing connectedness.

tion must be examined. Letting $\xi = N(x-1)$ and $\Phi(\xi)$ denote the inner solution, it is easy to show that it must satisfy

$$\Omega = \omega(1) + 2f(\Phi(\xi)) + g(\Phi(x))_x \tag{110}$$

$$\Omega = \omega(1) + f(\Phi(0)) - g(\Phi(0)) \tag{111}$$

$$\lim_{\xi \to -\infty} \Phi(\xi) = \phi(1). \tag{112}$$

The last condition is the matching condition. Wheter it is solvable or not depends on the details of the problem. Indeed, there may be a boundary layer at the other end instead in which case the outer equation would be different. There are many other possibilities, but all are solved in a manner similar to what I have sketched above. Kopell and Ermentrout (1990) tabulated all the possibilities in terms of the zeros of $h^{\pm}$ and the frequency gradients. A solution of particular interest as far as the central pattern generator of the lamprey assumes that $h^{\pm}(0) \neq 0, \omega(x) = 0$, and that the coupling is *anisotropic;* it is stronger from tail to head than the other way around. Under these conditions, it can be shown that the phase gradient is a nonzero constant except at the boundaries. This corresponds to a uniform travelling wave which controls the muscle activity of the swimming lamprey. The assumption that $h^{\pm}(0) \neq 0$ says that the coupling between the oscillators controlling the swim pattern is mediated by chemical synapses rather than gap junctions. In a number of other papers, we have shown evidence for this coupling anisotropy. Further infomation on this model as well as the behavior of the coupled cahin can be found in these papers. Kopell *et al.* (1990), extended the present results to a chain for which there is coupling beyond nearest neighbors. There is little qualitative difference other than the fact that the travelling wave solutions are less affected by local variations in frequency. Periodically forced chains have been exploited to provide information on the structure of the lamprey spinal cord central pattern generator (Kopell *et al.*, 1991; Williams *et al.*, 1990).

General connection topologies

The most general case is one in which there are no restrictions on the form of the coupling between the oscillators and each oscillator can communicate to any other one. If there are N coupled oscillators, then the phases satisfy:

$$d\theta_j/ds = H_j(\theta_1 - \theta_j, \ldots, \theta_N - \theta_j). \tag{113}$$

A solution to eq. (113) is phase locked if there are constants $\bar{\theta}_1, \ldots, \bar{\theta}_N$ and a constant Ω such that:

$$\Omega = H_j(\bar{\theta}_1 - \bar{\theta}_j, \ldots, \bar{\theta}_N - \bar{\theta}_j) \tag{114}$$

for all $j = 1, \ldots, N$. Any number can be added to each of the $\bar{\theta}_j$ and still provide a solution that is just translated in phase form the first solution. This is a consequence of the fact that a phase-locked solution is a periodic solution to an autonomous system of differential equations so there is always translation symmetry. Thus, I assume that $\bar{\theta}_1 = 0$ with no loss in generality and the remaining phases are then relative to the fist phase. The following theorem (Ermentrout, 1992) is quite useful in assessing stability. theorem Let $\alpha_{jk} = \partial H_j/\partial\theta_k$ evaluated at a phaselocked solution to eq. (113). If $a_{jk} \geq 0$ and the graph of the matrix (a_{jk}) is complete, then the phase-locked solution is orbitally asymptotically stable.
This theorem also enables one to extend solutions that depend on a parameters since it says that the eigenvalues are bounded away from zero and thus the implicit function theorem can be applied. As an application of this theorem, I will show that there are stable rotating waves in a two-dimensional $2N \times 2N$ array of oscillators. Consider the following system of equations:

$$d\theta_{jk}/dt = \sum_{j',k'} (H(\theta_{j',k'} - \theta_{j,k} + \alpha) - H(\alpha)) \tag{115}$$

where α is a parameter, (j', k') are nearest neighbors to (j, k) and $H(u)$ is an odd periodic function with positive derivative as long $-\pi/2 < u < \pi/2$. Recently (Paullet and Ermentrout) we have shown that there is a nonzero solution to eq. (115) for $\alpha = 0$ such that the phase-differences are all in the interval, $[-\pi/2, \pi/2]$. Only in the very center 4 squares do the phase lags equal $\pm\pi/2$. The graph of the matrix corresponding to this coupling and this solution is shown in Fig. 12. Clearly it is connected, so that the computed solution is asymptotically stable and can be continued for small α. In fact, we have numerically shown that at a critical value of α the solution disappears.

I want to point out that the synchronous solution to eq. (115), $\theta_{jk} = 0$ is always stable by the theorem as long as $|\alpha| < \pi/2$. An analogue of this theorem is true for continuum models of phase difference oscillators which arise when one tries to describe continuum neural network models.

6

Mechanisms Responsible for Epilepsy in Hippocampal Slices Predispose the Brain to Collective Oscillations

ROGER D. TRAUB* and JOHN G. R. JEFFERYS†

*IBM Research Division, T. J. Watson Research Center, Yorktown Heights, NY 10598 and Department of Neurology, Columbia University, New York, NY 10032, U.S.A.
† Department of Physiology and Biophysics, St Mary's Hospital Medical School, Imperial College, London W2 1PG, U.K.

Introduction

Clinical observation has provided a source of understanding of the brain for hundreds if not thousands of years (Clarke and Dewhurst, 1972). In the last hundred years or so, advances in experimental techniques have allowed observation of patients and post mortem tissues to be supplemented by data aimed at elucidating cellular mechanisms of diseases of the nervous system. This evolution has been particularly fruitful in the case of *epilepsy*. We have progressed from the clinician Hughlings Jackson's ideas on cortical localization (from direct study of his patients) in the last century to study of induced seizures *in vitro* systems such as brain slices and isolated whole brains (Paré *et al.*, 1992). The analysis of cellular mechanisms of epilepsy has in turn provided considerable information on basic brain mechanisms, some of which has been quite provocative (e.g. Miles and Wong, 1987a; Schwartzkroin and Prince, 1978; Johnston and Brown, 1981; Michelson and Wong, 1991; see Jefferys, 1990, and Traub and Miles, 1991 for review). In this chapter we shall consider one particular type of *in vitro* epilepsy called *synchronized afterdischarges* or "ADs" using it as a means to review the

an important role in memory (Squire, 1986) and it is therefore interesting that the results to be discussed raise some intricate questions concerning synaptic modification.

Neuronal Cell Types in the Hippocampus

About 90% of the neurons in the *CA3* region are *pyramidal* cells, the principal neurons that supply CA3 output to other parts of the brain (including in particular the *CA1* hippocampal region and the opposite CA3 region — there is one hippocampus in each cerebral hemisphere). Of particular relevance to epilepsy and probably to normal brain function, is the fact that CA3 neurons synaptically excite one another. These *connections* are called *recurrent* , and are described in more detail below. The remaining neurons, generally non-pyramidal in shape, are probably all inhibitory and most or all of them use *GABA* (γ-aminobutyric acid) as their *transmitter.* Some of these neurons stain for various peptides known to have transmitter or *modulatory* roles, so that if these *peptides* are released under physiological conditions, then the inhibitory neurons might produce effects on short (milliseconds), intermediate (hundreds of milliseconds) and long time scales. The physiological properties of inhibitory neurons (also called "interneurons") are considered elsewhere in Miles (1990a,b), the series of papers by Lacaille and colleagues (Lacaille, 1991) and are reviewed in the book of Traub and Miles (1991).

Pyramidal neurons have two sets of *dendrites,* called *basilar* and *apical,* with recurrent synaptic connections into both sets of dendrites. Presynaptic neurons release transmitters onto a given postsynaptic neuron, causing *membrane synaptic currents* to flow and possibly initiating *intracellular "secondary" metabolic processes* that in turn influence the membrane. There are also synaptic receptors whose primary function is to couple to intracellular biochemical processes. Because metabolic processes are, in general, slower than voltage-gated membrane processes, the different receptor coupling mechanisms introduce a range of time scales into neuronal function. Located in the membrane are diverse populations of *channels* , each specific for a particular ion or set of ions (Na^+, Ca^{2+}, K^+, Cl^-) and each regulated in its own characteristic fashion by transmembrane potential, intracellular [Ca^{2+}], intracellular enzymes, modulatory substances and so forth. There is a classification of channel types (Llinás, 1988), growing constantly and hippocampal neurons are blessed with a considerable variety, particularly K channels. To add to the complexity, membrane channels, even of a given type, do not appear to be distributed in uniform (i.e. spatially homogeneous) concentration across the cell membrane (Traub *et al.,* 1991 and *loc. cit.*).

CA3 pyramidal neurons are able to generate two kinds of *action potentials, fast* (Na-mediated, 1–2 ms) and *slow* (Ca-mediated, 5 or more ms). Ca

action potentials tend to be larger in the apical dendrites than in the soma, while the reverse appears to be true for Na action potentials this probably reflects differences in channel density between the two membrane regions. Very crudely, the roles of the different K channels are to help repolarize action potentials (with time constants of a few milliseconds) and to control the time intervals between either fast or slow action potentials. The time constant for relaxation of all kind of K channel, the AHP channel (AHP standing for "afterhyperpolarization") is about 1s (Numan *et al.*, 1987), while the time course for closure of the D type of K channel is even slower (Storm, 1988).

Membrane currents of hippocampal neurons interact with one another to generate different patterns of cellular output. Thus, *in vivo* cells may fire isolated action potentials or *spike doublets* , or they may fire *"intrinsic bursts "* (*"complex spikes "* in the *in vivo* literature), consisting of a brief train of *action potentials* at typically 5–10 ms intervals. Intracellular recordings indicate that the action potentials in a burst ride on a slower depolarizing wave mediated by Ca^{2+} ions (Wong and Prince, 1981) and that bursts can be generated in the dendrites as well as in the cell body (Wong *et al.*, 1979; R. Miles, personal communication; Benardo *et al.*, 1982). Bursts are usually followed by a long AHP. Bursts in hippocampal neurons are called intrinsic because of the demonstration that isolated neurons can produce them (Wong *et al.*, 1984; H. Michelson and R. K. S. Wong, personal communication).

Intrinsic bursting is not, by itself, abnormal, although it occurs rarely in awake animals that are moving about (G. Buzsáki, personal communication). In awake immobile animals, many neurons (but still a minority of them) fire brief trains of approximately simultaneous action potentials, producing *"physiological sharp waves"* (Buzsáki, 1986; Buzsáki *et al.*, 1992). Intrinsic bursts occur under resting conditions in healthy hippocampal slices, apparently uncorrelated in time between different neurons. When most or all of the neurons in a local population burst simultaneously, however, it is considered abnormal ("epileptogenic", or *"epileptic"* for short).

Of particular importance for brain function, we feel and too little appreciated by non-specialists is the fact that nerve cell membranes and synaptic receptors involve physical processes that operate on many overlapping time scales. Examples for membrane processes have been given. We shall give some specific examples for synaptic interactions that are relevant to the hippocampus. Thus, pyramidal neurons probably use glutamate as their (excitatory) *neurotransmitter. Glutamate* binds to at least three chemically and functionally distinct types of receptor on pyramidal cell membranes, designated *"AMPA "* (an acronym for a compound that binds to this receptor type), *"NMDA "* (for *N* -methyl-D-aspartate, a ligand for the second receptor type) and *"metabotropic "* (Desai and Conn, 1991). AMPA responses peak within a few milliseconds and have decay time constant also of a few milliseconds (Williams and Johnston, 1991). NMDA-gated currents[1] peak

more slowly than for AMPA and relax with a time constant of the order of 100 ms (Forsythe and Westbrook, 1988; Hestrin *et al.*, 1990a,b). The action of metabotropic receptors is slower still. Analogously, the inhibitory transmitter GABA can interact with at least two types of receptors in the hippocampus, called, $GABA_A$ and $GABA_B$ (Miles and Wong, 1984; Miles, 1990; Hablitz and Thalmann, 1987; Ashwood *et al.*, 1987; Traub and Miles, 1991). $GABA_A$ IPSPs have very rapid onset and the underlying conductance decays with a time constant of the order 10 ms or so. In contrast, the onset of $GABA_B$ can be delayed as long as about 40 ms (R. Miles, personal communication) and decays with a time constant of 100 ms or more.
[1]

The hippocampal slice

The hippocampus of a rodent such as the rat contains many hundreds of thousands of neurons. Furthermore, experiments on cellular mechanisms *in vivo* are difficult because in the living animal the brain does not sit still but pulsates with heart beat and respiration. In addition when drugs are applied *in vivo* it is difficult to attain a uniform concentration throughout the tissue. To some extent these problems are circumvented by using slices which have only some thousands of cells (and can be microdissected still further). Slices are maintained in a motionless state in a chamber and drugs can be applied uniformly. It is fortunate that properly prepared slices of hippocampals and neocortex retain sufficient synaptic circuitry to generate a wide variety of interesting neuronal population behaviors. Population activity can be modified in several ways, including altering the ionic composition of the bathing medium (e.g. by removing Mg^{2+}), application of drugs that block specific receptors or ion channels and repetitive electrical stimulation. Population activity can be monitored by *intracellular recording* from one (sometimes two or more) neurons, by *field potential recording* (a means of monitoring synaptic currents and action potentials from localized clusters of cells), or by more esoteric means (e.g. optical recording with voltage-sensitive dyes or with intracellular compounds sensitive to Ca^{2+} or Na^{2+} concentrations).

[1] AMPA and NMDA receptors differ not only in time course but in other important properties. Thus, in the presence of neurotransmitter, NMDA channels exhibit a voltage-dependent conductance, with conductance increasing in the range from resting potential (say −60 mV or −70 mV) to 0 mV. AMPA-gated channels are not voltage-dependent in this potential range, although rectification can occur outside the range. NMDA channels are regulated not only by membrane potential but also by the concentration of extracellular small neutral amino acids, including glycine and by Mg concentration. Reducing extracellular $[Mg^{2+}]$ diminishes the voltage-dependence of the channel and increases the conductance at any given transmembrane voltage. A quantitative model of the voltage- and Mg-dependence of the NMDA channel has been presented by Jahr and Stevens (1990) and it is this formulation that we use in our simulations.

How can "epilepsy" be defined in brain slices?

Epilepsy is a clinical syndrome, of diverse causes, in which recurrent spontaneous seizures occur. (We shall not try to define *seizure* further, but instead refer the reader to standard neurological texts, such as Bradley *et al.*, 1991.) In order to apply the term to brain slices, one must consider a sequence of logical transformations. Clinically apparent seizures usually are accompanied by characteristic electrical signals, recorded in the *EEG*, representing some sort of average behavior of a large neuronal population. It turns out that similar signals can be recorded at times without clinical phenomenology apparent either to the patient or to a trained observer. Such signals are called electrographic seizures; their occurrence allows one to think of epilepsy in terms of electrical signals as well as clinically. The next transformation is the production of signals similar to electrographic seizures by experimental manipulation of the brains of animals, first involving large parts of the brain and later only involving confined regions. In this way, some of the cellular correlates of electrographic seizures were recognized (Matsumoto and Ajmone Marsan, 1964). Specifically, the local field potential during an electrographic seizure usually consisted of a series of waves; in phase with each wave, nearby neurons tended either to fire a *burst of action potentials* or to exhibit an IPSP (inhibitory postsynaptic potential). Slower ("DC") potentials, either depolarizing or hyperpolarizing, might be superimposed. This finally set the stage for the study of epilepsy *in vitro*: namely, the production of population events in which cellular activities (usually the firing of one or more action potentials) were highly correlated with extracellularly recorded waves. Thus, the cycle was closed onto one of the original 19th century definitions of epilepsy, in which excessive neuronal firing was assumed to be critical: "occasional, abrupt and excessive discharges of parts of the cerebral hemispheres" (Jackson, 1931).

Methods of inducing epileptic events in vitro

Effective techniques include, but are not limited to, the following: altering ionic composition of the bathing medium (lowering $[Ca^{2+}]$, $[Mg^{2+}]$ or $[Cl^{-}]$; elevating $[K^{+}]$); blockade of K currents, affecting both somadendritic membranes and probably presynaptic terminals (e.g. with TEA (tetraethyl ammonium) or 4AP (4-aminopyridine)) blockade of $GABA_A$ receptors (e.g. with picrotoxin, bicuculline or penicillin); repetitive electrical stimulation; or prior *in vivo* injection of a biological toxin (cholera toxin, tetanus toxin). The latter two techniques induce multiple secondary changes, now being further characterized. In this chapter, we shall consider only the conceptually simplest method, the blockade of $GABA_A$ receptors. It is complicated enough. A major challenge remains in identifying the mechanisms that the different forms of epileptogenesis share in common.

What can be learned from the study of epilepsy in vitro

One long-term goal is to understand better the clinical condition. This will take time. We may note the difficulties of reproducing *in vitro* the conditions operating when a patient has a seizure, or even knowing what these conditions are. Then there is the near-impossibility of studying control human tissue. In the meantime, there is much to be learned from analysis of the detailed structure of epileptic events: how do they arise, what determines the frequency of bursting within an event and what controls the frequency with which events occur? These questions lead directly into membrane and synaptic physiology.

Modeling approach

Our modeling approach has been reviewed in detail elsewhere (Traub and Miles, 1991; Traub *et al.*, in press), but some comments are in order so that the figures will be intelligible. Each neuron is modeled as a nonlinear circuit, intended to capture reasonably accurately those membrane processes taking place in the 0.5–1000 ms time range. Each model neuron consists of 19 compartments, so that there is some representation of the dendritic systems, but no attempt to capture details of dendritic branching or dendritic spines, indeed, others have modeled individual neurons with many thousands of compartments, but such models do not lend themselves to *network simulations.* Each compartment contains any of a repertoire of 5 or 6 different intrinsic ionic conductances, one for Na, one for Ca and three or four for K. The conductances are not distributed uniformly across the membrane. The basic single-cell model is described in Traub *et al.* (1991), but in order to speed up computation and allow simulations with 1000 pyramidal neurons, we have in this chapter used a variant in which only 7 compartments (the soma and proximal and middle apical dendrites) are "active" and one of the K currents (the A current) is removed. Synaptic connections are assumed to be random for "small" networks (typically 1000 cells or fewer) and locally random for larger networks (Miles *et al.*, 1988). The density of synaptic connections and the properties of individual synapses, are constrained by experimental data (particularly dual intracellular recordings) whenever possible (Miles, 1990a; Miles and Wong, 1984, 1986, 1987). The aim is to elucidate unexpected and interesting aspects of population behaviors, and, needless to say, to make novel experimental predictions.

Example of in vitro *epileptiform activity*

Figure 1 illustrates an experimental afterdischarge (left), recorded from the CA3 region of a *guinea pig hippocampal slice* , bathed in picrotoxin to block

$GABA_A$ receptors. The upper trace is from an intracellular electrode in the soma of a pyramidal cell and the lower trace is the extracellular potential in the pyramidal cell body layer. The tight correlation of these signals demonstrates the synchrony of any particular cell with the population as a whole. *Afterdischarges* occur spontaneously at intervals of several seconds, or they can be evoked by simulation of a few neurons or even one neuron (Miles and Wong, 1983). As described previously (Miles *et al.*, 1984; Hablitz, 1984), the event consists of an initial (F) burst that is longer than the later (S) bursts. There can be 1 to perhaps 10 2S bursts, at frequencies of about 15 Hz (slower in ADs produced by tetanus toxin). In slices from immature rats, afterdischarges can last a few seconds or tens of seconds (Brady and Swann, 1986, 1988). The entire event is followed by a prolonged (seconds) *afterhyperpolarization*. Figure 1 also illustrates how a *network model* with 1000 pyramidal neurons (and 100 inhibitory neurons, with the output of interneurons activating $GABA_A$ "receptors" blocked) is able to reproduce the essential features of the afterdischarge. A detailed account of this model and some associated experiments, is given in Traub *et al.* (in press); here, we shall outline the key elements.

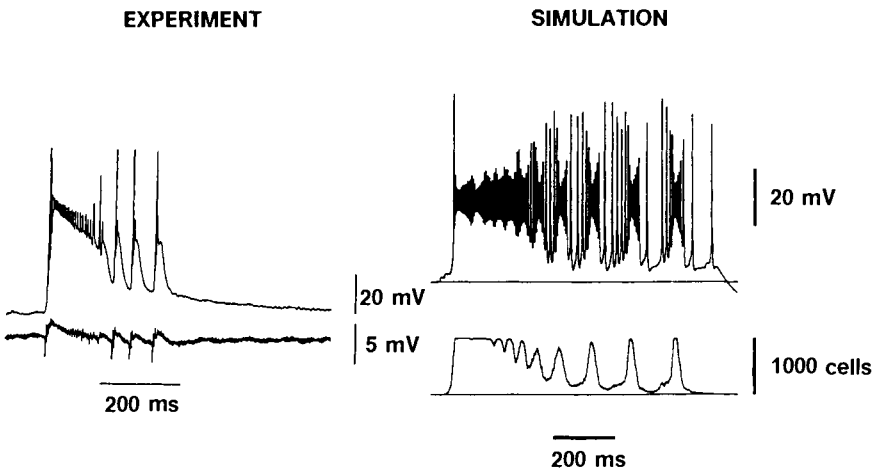


Fig. 1. Epileptic afterdischarges *in vitro*. Experiment shows a spontaneous event from a guinea pig hippocampal slice bathed in the $GABA_A$-blocker picrotoxin (methods as in Miles *et al.*, 1984). Upper trace is intracellular, lower trace extracellular. Note the synchrony between the two traces. Simulation is from a randomly connected model with 1000 pyramidal neurons and 100 inhibitory cells, half producing $GABA_A$ IPSPs — blocked in this case — and half producing $GABA_B$ IPSPs (methods as in Traub *et al.*, in press, except that these neurons have 7 active compartments and AMPA and NMDA inputs are into both apical and basilar dendrites). Upper trace is somatic potential of one model neuron and lower trace is the number of pyramidal cells depolarized more than 20 mV. Experimental data courtesy of R. Miles.

The pyramidal neurons are known to be recurrently interconnected, with a probability of interaction between two nearby cells being a few percent (Traub and Miles, 1991). We make the specific assumption that when two neurons are functionally connected, then both AMPA and NMDA receptors are activated at the respective synapse (Bekkers and Stevens, 1989; McBain and Dingledine, 1992). The assumption is reasonable because AMPA blockers generally abolish synchrony (Lee and Hablitz, 1989) suggesting that AMPA receptors are present in recurrent connections; this is further supported because NMDA blockers do not abolish synchronized firing induced by picrotoxin, although its pattern is altered (Lee and Hablitz, 1990). On the other hand, NMDA receptors *also* appear to be present in the recurrent connections, because synchronized firing can be elicited by lowering Mg concentration (increasing NMDA-gated conductances) and this synchronization survives AMPA-blockade, but not NMDA-blockade (Jefferys and Traub, 1992).

Given this background, we now need to account for several aspects of the AD. First, is the *growth phase* : how does stimulation of a single neuron lead to such a complicated event involving the whole population? Second is the *steady-state* phase: once all the neurons are firing together, why is the event organized into a series of synchronized bursts? The growth phase appears to arise as a chain reaction. What makes this possible is the fact the synapses between CA3 pyramidal neurons are *strong* enough to allow a burst in a presynaptic cell to elicit, with probability at least 0.3, a burst in a postsynaptic cell. Because each pyramidal cell excites *in vitro* at least 20 other pyramidal cells on average, a burst in one cell elicits bursts in perhaps 6 others. This is enough to excite a number of interneurons which, under normal conditions, would suppress or damp the spread of excitation, but with $GABA_A$ blocked, the chain reaction can proceed. Why does $GABA_B$ not damp it? The latency for burst spread from one cell to another can be as short as about 10 ms. Once many cells are firing, new cells are recruited even more rapidly by the firing of many presynaptic cells together. On the other hand, $GABA_B$ IPSPs are delayed for perhaps 40 ms. Thus, the chain reaction can *outrun* slow inhibition. Somewhat analogous considerations arose in the design of the uranium bomb: how to ensure sufficiently many steps in the fission → neutron release → induced fission (analogous to spread of bursting) before thermal expansion ("slow inhibition") of the uranium rendered most of the released neutrons unable to induce new fissions (inhibition preventing the spread of bursting). This chain reaction has been described previously (Traub and Miles, 1982, 1991; Traub *et al.*, 1987); here, we shall consider more fully the steady-state phase of the AD. We note in passing that in this earlier work on the growth phase of synchronization, the pharmacology of the recurrent synapses was not discussed, but the way in which these synapses was simulated corresponds to the properties of AMPA receptors. Indeed, we would guess that AMPA receptors alone can support the spread of bursting from one CA3 cell to another

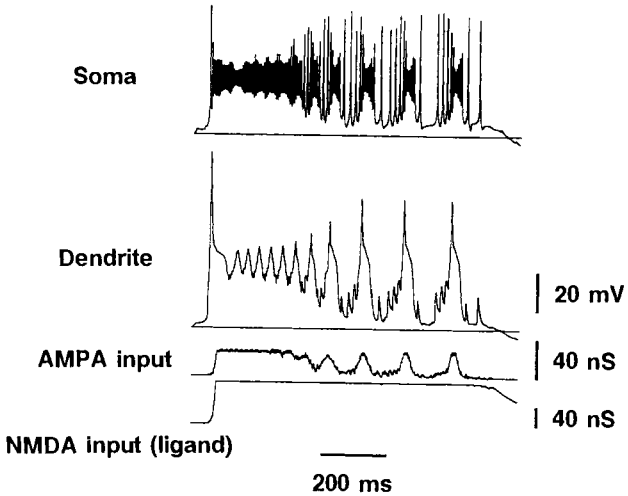


Fig. 2. Synaptic inputs during simulated afterdischarge (same simulation as in Fig. 1). The AMPA input (consisting of summated rapidly decaying signals) follows the cellular potential (and hence the population output). In contrast, the NMDA input rises quickly and saturates. It is this input that, in the model, elicits the series of calcium spikes (2nd trace) and somatic bursts.

at least under optimal conditions, but further experiments are needed on this issue.

How does the model account for afterdischarges?

This is illustrated in Fig. 2, recapitulating in a larger model the basic ideas discussed in Traub *et al.* (in press). Because the afterdischarge is synchronized between neurons, we may pick a representative neuron and examine both its membrane potential (in different parts of the cell) and its synaptic inputs, the latter in turn sampling some very highly correlated subsets of neurons in the population. The dendritic region plotted is from a compartment containing a calcium conductance but no sodium conductance, so that slow action potentials can be produced there but not fast ones. What is immediately apparent is that the AMPA input to this neuron follows the somatic potential of the neuron (and hence the somatic potential of all the neurons) closely. This is expected because AMPA-mediated EPSPs (excitatory postsynaptic potentials) have a rapid time course and because axon conduction delays between neurons are small. On the other hand, the NMDA input quickly saturates (because of the long time course of individual EPSPs, together with the finite number of receptors) and then remains almost constant for the duration of the afterdischarge. Blocking the

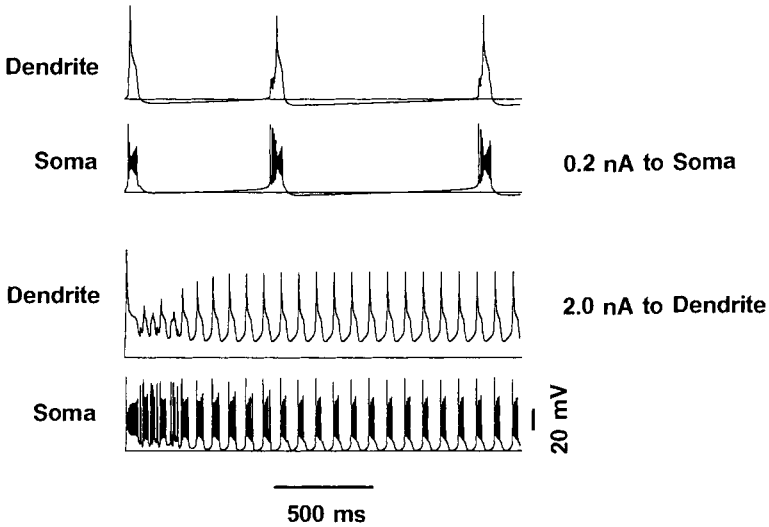


Fig. 3. Two modes of firing in single cell model. Top: small current injected into soma elicits low frequency bursting (c.f. Wong and Prince, 1981). Each somatic burst is correlated with a dendritic calcium spike. Bottom: injection of a large current into the apical dendrite leads to high frequency (about 10 Hz) bursting (c.f. Traub *et al.*, 1991, in press). It is this pattern of firing that is elicited in our model of afterdischarges by NMDA synaptic input.

AMPA input to one neuron in the model population does not change the response of that neuron markedly (not shown). What we are seeing, therefore, is the response of a neuron to a constant large dendritic input — the NMDA input — with the AMPA input acting to maintain synchrony of the population. Is this conclusion reasonable, given the known experimental data? And why is the initial burst longer than subsequent bursts?

Repeating dendritic bursts in hippocampal neurons

Our *model of a pyramidal neuron* (Traub *et al.*, 1991) was constructed in an attempt to understand intrinsic bursts in terms of the measured kinetics of identified types of membrane channels data available from the isolated neuron preparation (Kay and Wong, 1987). This model worked most successfully when there was a relative segregation of different channel types, with Na channels closer to the soma and Ca channels more predominant in the dendrites — an assumption that does appear to be consistent, so far, with experimental data. The model then would generate *intrinsic bursts*, for example when the soma was depolarized by a small current (Fig. 3; compare Wong and Prince, 1981). It was subsequently discovered that *den-*

dritic stimulation, with a large enough current, would elicit a quite different activity, consisting of a transient long burst (or series of bursts) followed by steady repeating short bursts. The initial transient is relatively long because it takes more than 100 ms for the AHP to develop completely (Lancaster and Adams, 1986); succeeding bursts take place when the membrane is shunted and are therefore briefer in duration than the initial burst. This type of firing behavior upon dendritic stimulation has been confirmed experimentally (Traub *et al.*, in press). Therefore, the idea that a large sustained NMDA input elicits the afterdischarges appears reasonable. This idea is supported by an additional experimental observation: NMDA blockers tend to abolish S bursts, leaving an attenuated initial burst (Lee and Hablitz, 1990; Traub *et al.*, in press). This observation on the effects of NMDA blockers is also consistent with our guess that bursting can spread from cell to cell supported by AMPA receptors alone.

Our ideas on afterdischarges were developed for the hippocampus but they may apply to the neocortex as well. Thus, pyramidal neurons in the neocortex also can generate repeating calcium spikes (Amitai *et al.*, 1991), the neocortex contains NMDA receptors (Thomson *et al.*, 1989) and the neocortex can generate afterdischarges.

Scaling the excitatory synaptic conductances

The pattern of simulated ADs depends on the *strength* of the *AMPA* and of the *NMDA conductances* (i.e. the peak conductance elicited by a single presynaptic action potential). Roughly the strength of the AMPA synapses determines the duration of the initial burst and the strength of the NMDA synapses determines the number of secondary bursts (Traub *et al.*, in press). In the examples illustrated here, AMPA and NMDA conductances were scaled to give an initial burst lasting about 200 ms and 4 secondary bursts. Are these conductance values consistent with experimental data on unitary synaptic interactions, the effect one cell has on another? The pharmacological data are incomplete but we can make some comparisons. These comparisons are also useful for trying to predict the effect of NMDA receptors on non-epileptic activity in the CA3 region. We shall present simulated unitary interactions for a single presynaptic action potential and for a presynaptic burst. In each case, the conductances are scaled as they are in Figs 1 and 2 and we shall examine the AMPA-mediated EPSP, the EPSP elicited by AMPA and NMDA together in *normal* (1.0 mM) [Mg^+] and the AMPA + NMDA in 0.0 mM Mg (this augmenting the NMDA component, as described quantitatively by Jahr and Stevens (1990)).

Figure 4 (left) shows that the simulated unitary AMPA-mediated EPSP is brief and followed by an undershoot, a result of *recruiting* Ca-mediated outward K currents. The undershoot, of course, can not happen in a purely passive membrane, but is observed experimentally in some EPSPs (Miles

and Wong, 1986). The EPSP with both receptors active (Fig. 4 middle) has a prolonged *depolarizing hump*, even though the NMDA input (bottom trace) is small. This hump, as is the case experimentally (Miles and Wong, 1986), is voltage-dependent, decreasing with cellular hyperpolarization: if the model soma is held 9 mV hyperpolarized to rest, the AMPA + NMDA EPSP is quite similar to the AMPA EPSP at rest (not shown). In the model, the voltage-dependence arises not only because of the properties of the NMDA receptor but because of intrinsic conductances as well (c.f. Thomson *et al.*, 1988). Thus blocking the intrinsic Ca conductance reduces the EPSP hump (at resting potential) but does not abolish it. What we conclude from this, is that with NMDA receptors in recurrent synaptic connections, the autonomous behavior of the CA3 region (the behavior determined by interactions between CA3 pyramidal neurons) might be strongly regulated simply by regulating the resting potentials of the cells, something that might be accomplished by neuromodulators or by changes in the ionic composition of the extracellular medium.

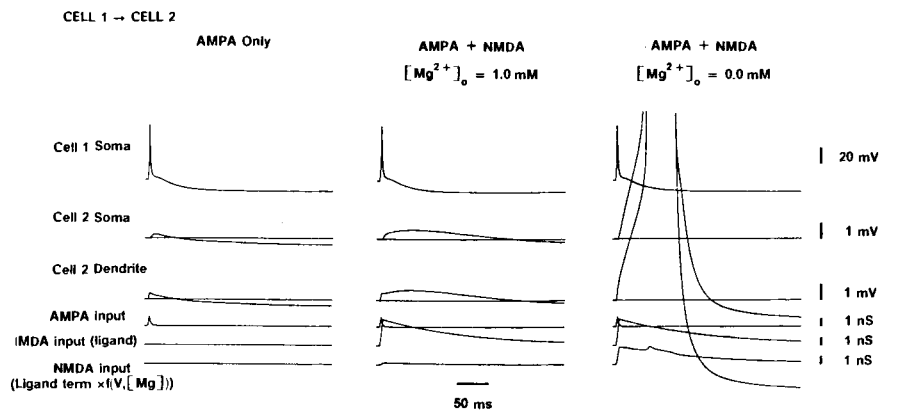


Fig. 4. Simulated unitary synaptic interactions. Cell 1 is connected to cell 2 and cell 1 is stimulated to fire a single action potential. This action potential activates either AMPA receptors alone (left), using the same unitary conductance as in Figs 1 and 2; AMPA and NMDA receptors at 1.0 mM Mg (middle), again scaled as in Figs 1 and 2; or AMPA and NMDA receptors in 0 Mg (right). The NMDA conductance is the product of a ligand-gated term, $h(t)$ and a function of the local dendritic potential and extracellular Mg concentration, $f(V,[Mg])$, the latter simulated as described in Jahr and Stevens (1990). The unitary AMPA input is simply a function of time and does not depend on voltage or Mg in this model. The simulations of Figs 1 and 2 were done with [Mg] = 1.0 mM. See text for further details. Lowering Mg increases the value of f and in the illustrated case allows a single spike to evoke a postsynaptic burst.

Next, it is instructive to predict what might happen in a medium with

no Mg ion (Fig. 4, right). In comparing the middle and right parts of Fig. 4, note that the ligand-gated NMDA input to the postsynaptic neuron is the same, because the cell is stimulated in the same way in each case, but the actual NMDA conductance (bottom trace) is much larger on the right, owing to an increase in f(V,[Mg]). The result is the transduction of a single presynaptic action potential into a postsynaptic burst, the action potentials of which are truncated at this gain. Such transduction is not observed in normal media indeed, it is rare for a single action potential to evoke an action potential at all in a postsynaptic CA3 pyramidal cell (Traub and Miles, 1991). The low Mg case has not yet been tested experimentally to our knowledge.

Now let us examine the simulated effects of a *presynaptic burst.* (While cell-to-cell transmission of bursting has been demonstrated experimentally (Miles and Wong, 1987a) we are not aware of an analysis of this in terms of AMPA-mediated and NMDA-mediated components.) With the conductances scaled as in Figs 1 and 2, the AMPA input alone from a presynaptic burst can elicit a postsynaptic burst (Fig. 5 left). This burst transmission, however, is not robust in the sense that small hyperpolarizations of the postsynaptic cell abolish it (not shown). With AMPA and NMDA acting together (Fig. 5, middle) burst transmission occurs much more robustly, although the form of the burst is little altered compared with AMPA acting alone: the postsynaptic burst now has an extra action potential and occurs at slightly reduced latency. The fact that burst morphology is so little affected by NMDA is interesting, given the large NMDA conductance (bottom trace of the middle column). Why should this be? The reason is that the burst is dominated by the dendritic Ca spike (middle trace), with the synaptic input from a single presynaptic neuron having little additional influence. Of course, simultaneous NMDA input from *many* presynaptic neurons does exert a significant influence, tending to produce a train of *dendritic Ca spikes* — that is how we believe afterdischarges arise (Fig. 2). Consistent with this, we predict that in lowered Mg media, a single presynaptic neuron, firing a burst, might be able to elicit more than one dendritic Ca spike in a postsynaptic cell (Fig. 5, right).

How might NMDA receptors influence population activity that is not epileptic?

If the effect of NMDA receptors on recurrent synapses is subtle at the level of one cell communicating with one other cell (Fig. 5, middle), we might expect minimal influence of NMDA receptors on population activity under conditions where there is little overall firing. On the other hand, if there is enough firing to produce simultaneous inputs to neurons from more than one presynaptic neuron (as happens in the extreme during an afterdischarge), then NMDA receptors could show their presence in the population output — even without the synaptic plasticity that NMDA receptors might

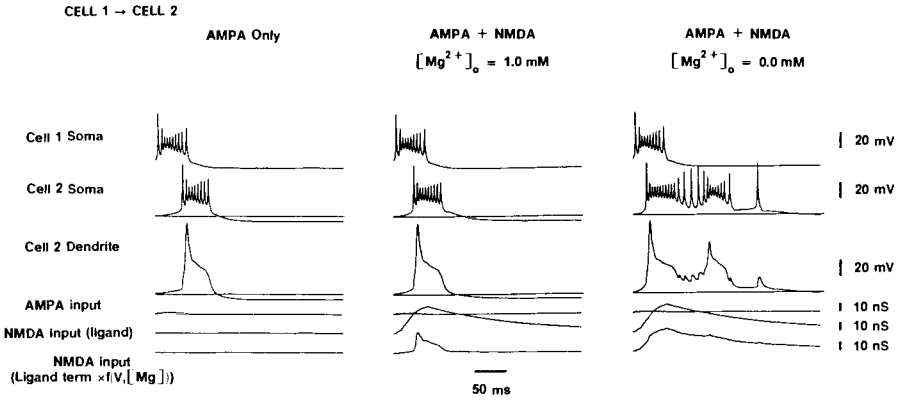


Fig. 5. Simulated unitary synaptic interactions when the presynaptic cell (cell 1) fires a burst. The arrangement of the figure is as for Fig. 4, but note the different voltage and conductance calibrations.

mediate. To explore this notion, we simulated several seconds of spontaneous activity in a *CA3 network model* (1000 pyramidal cells, 100 inhibitory cells), at different strengths of the $GABA_A$ synapses (Fig. 6). With inhibition blocked completely (upper traces), we see what has already been described: AMPA receptors alone (left) can mediate synchronization to give simultaneous bursts, but do not give ADs; while AMPA and NMDA receptors together do give ADs (right). On the other hand, at powerful levels of inhibition (bottom), NMDA receptors only increase the overall level of firing; while the two cells shown do exhibit simultaneous bursts in this epoch, it is not reflected in a *blip* in the population activity. The most interesting case is shown in the middle of Fig. 6: here, at an intermediate level of inhibition, the presence of NMDA receptors — at least in this model — predisposes the system to a *population oscillation* in which synchrony during any one wave is incomplete, i.e. not all of the cells fire during each wave. The activity shows up in pairs of neurons, held hyperpolarized, as correlated synaptic potentials. While similar activity has been observed during partial blockade of inhibition (Schneiderman, 1986; Miles and Wong, 1987a), or even in control conditions (Schwartzkroin and Haglund, 1986; Traub *et al.*, 1989), the role of NMDA receptors in the activity remains to be explored. It is interesting that the frequency of partially synchronized activity in low concentrations of penicillin is about 2 Hz, as in the simulation (Traub *et al.*, 1987). A role for synaptic inhibition in regulating NMDA effects has been proposed for other neuronal systems (Davies and Collingridge, 1989;

Luhmann and Prince, 1990).

Concluding experimental and theoretical issues

There is a fashionable general *model of memory* and neural development that involves *Hebbian learning* , i.e. the strengthening of excitatory synapses that are active when the postsynaptic cell is active. It is no longer easy to cite references for this idea comprehensively or fairly, as it seems to be almost taken for granted. It plays a part in the thinking behind the Hopfield or spin-glass model (Hopfield and Tank, 1986), as well as in models of neural development (Constantine-Paton *et al.*, 1990; Bear *et al.*, 1990; Linsker, 1990; Singer, 1990). Many have pointed out that *long-term potentiation* (LTP) (Gustafsson and Wigstrom, 1988) is "Hebbian" because of the requirement for postsynaptic depolarization coincident with presynaptic activity. NMDA receptors enter into this picture as well: NMDA gated currents contribute to postsynaptic depolarization and postsynaptic Ca entry (the NMDA channel being permeable to some extent to Ca ions) seems to be a specific trigger necessary to initiate LTP. Blockade of *NMDA receptors* also prevents LTP at many (but not all) synapses in which LTP has been described (Muller *et al.*, 1988). In order to explore these ideas in the specific context of the CA3 region, many quantitative data are needed, particularly relating to the precise postsynaptic dendritic "state" required for synaptic modification to occur and the extent to which cooperative or correlated firing of presynaptic neurons must take place (Friedlander *et al.*, 1990). Our results, however, suggest some interesting possibilities. Let us suppose that excitatory synaptic plasticity in recurrent synapses does take place, is important and depends on the presence of NMDA receptors. Such ideas have been considered by Muller *et al.* (1991), who suggest that recurrent CA3 synapses are critical in the formation of a computationally useful *"spatial map"*. We may consider two alternatives: (1) if plasticity in recurrent synapses can take place at low levels of activity, perhaps even when single presynaptic and postsynaptic neurons fire together (Ahissar *et al.*, 1992), then the presence of NMDA receptors should have only subtle effects on population input–output, that is, on a short time scale compared to the time for plasticity to take place. (2) On the other hand, when plasticity in recurrent synapses requires the correlated firing of many presynaptic neurons, the presence of NMDA receptors will themselves alter population output by facilitating the spread of activity within the network. Perhaps an influence of NMDA receptors on population dynamics (by enhancing the spread of firing along recurrent pathways) is functionally useful. On the other hand, if such an influence is not useful or is frankly disruptive, then the brain may have mechanisms for suppressing the spread of recurrent firing. The theta EEG rhythm has been proposed to be one such mechanism (Traub *et al.*, 1992, in press).

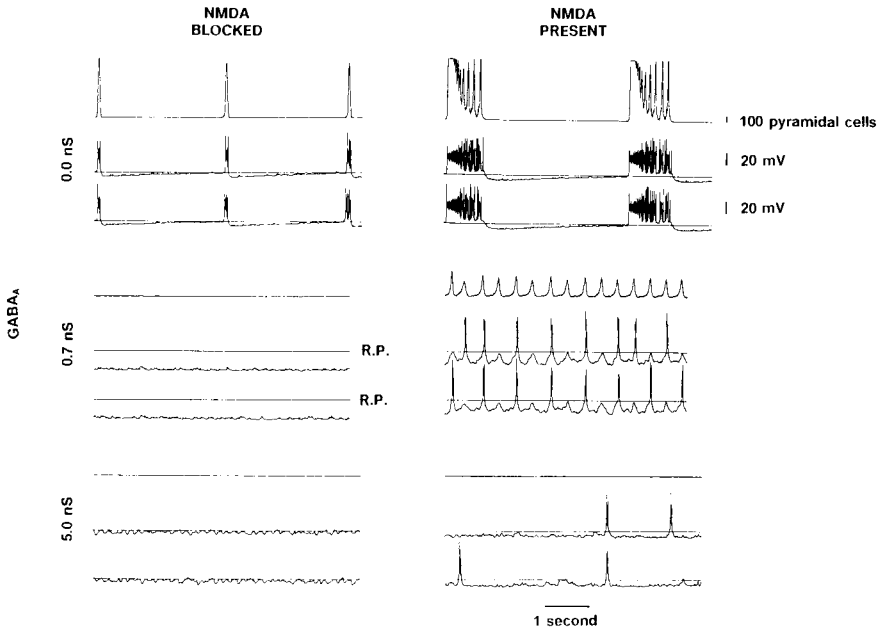


Fig. 6. Autonomous behavior of the 1000 pyramidal cell model, with and without NMDA blockade and at different levels of $GABA_A$ synaptic efficacy. Each simulation shows the number of pyramidal cells depolarized more than 20 mV at the soma (upper trace) and the somatic membrane potential of two arbitrarily chosen (and not monosynaptically connected) pyramidal cells (middle and lower traces). Every pyramidal neuron receives independent Poisson-distributed synaptic noise, with AMPA component only, into the proximal apical compartment, mean interval 100 ms; this was intended to represent random mossy fiber input. The peak conductance of the mossy input was 2.94 nS. With $GABA_A$ blocked, AMPA receptors, in the recurrent synapses, mediate synchronized bursts analogous to interictal EEG spikes (upper left), while AMPA and NMDA receptors in combination, in the recurrent synapses, can produce synchronized after discharges (upper right). When $GABA_A$ is sufficiently large (bottom), cellular activity is, for the most part, uncorrelated. There is an intermediate range of $GABA_A$ (middle) for which AMPA receptors alone mediate disorganized activity, whereas both receptors together can produce (in the simulation) a population oscillation at about 2 Hz. In the middle simulations, the 2 cells whose potentials are plotted have been held hyperpolarized to emphasize EPSPs. R.P. is resting potential.

Acknowledgements

We wish to acknowledge helpful discussions with Drs Richard Miles, William N. Ross, Robert K. S. Wong, Hillary Michelson and Gyorgy Buzsáki. We especially thank RM for sharing unpublished data (as in Fig. 1). This work was supported by IBM and the Wellcome Trust. JGRJ is a Wellcome Trust Senior Lecturer.

7

Diffusion Models of Single Neurones' Activity

LUIGI M. RICCIARDI

Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Universita, degli Studi di Napoli "Federico II", Via Cinzia, 80126-Napoli, Italy

Introduction

The literature existing on the subject single neurones' dynamics within the context of the theory of stochastic processes is much too wide to be reviewed in this chapter, even on a perfunctory base. On the other hand, a detailed discussion of any particular neuronal model would turn up to be of interest only to a restricted audience and miss to provide a feeling of the current state of the researches in progress in this field and of the numerous still open related mathematical questions both of an abstract and of a computational nature. I have thus come to the conclusion that an acceptable compromise could be achieved by organising the contents of this chapter in a way to aim at the following target: To provide a bird's eye view of a whole class of neuronal models as well as on the mathematical machinery involved by these models. Thus doing, anyone interested in a deeper understanding of the topic that I shall discuss will possess the basic knowledge and the indispensable tools to be able to dig autonomously into the literature or to undertake some challenging extensions or further derivations on his own.

More specifically, I shall offer a survey of one-dimensional stochastic diffusion models for the membrane potential of a single neurone, with emphasis on the probabilistic properties of the models and on the related first-passage-time (FPT) problems, namely on the determination of the neuronal output. While I make use of the formalism of stochastic differential equations (SDE's), I shall also show in a couple of paradigmatic instances how diffusion models can be constructed from first principles, that is without resorting to SDE's. I shall thus indicate how the much celebrated Ornstein–Uhlenbeck (OU) neuronal model can be obtained as the limit of a Markov

process with discrete state space in continuous time. Similarly, a straightforward limit procedure can be seen to lead to a neuronal model which embodies the notion of reversal potentials. Some more references and numerous complementary considerations can be found in review article by Lánský *et al.* (1990), that I have taken as the basis for the present chapter.

It is interesting to point out that the task of constructing neuronal models by means of diffusion processes has led some of us to face a variety of new and truly challenging analytic and computational problems about which I shall say something in the sequel. Hence, once again it emerges the fruitful feedback between fields as diversified as mathematics and biology. This is an important point that, in my view, deserves some considerations. Indeed, while everyone would agree that use of mathematical tools are beneficial, if not indispensable, to biologists, the converse statement may sound groundless. Still, there exist numerous examples showing that biology has proven useful to mathematics as suggestive of truly new and fecund investigations of an exquisite mathematical nature. One outstanding example is the mathematical theory of reaction–diffusion systems; a second example, probably unfamiliar to non specialists, is offered by Feller's theory of parabolic differential equations (Feller, 1952) that has successively led this author to construct the rigorous theory of diffusion processes in one dimension (Feller, 1954). This second example is particularly enlightening and thus deserves to be briefly outlined hereafter.

To formulate a mathematical model for the growth of a population of micro-organisms, Feller was led to make use of the well-known stochastic process that bears the name of " branching process". Hence, he considered (Feller, 1951a) a population consisting of similar individuals and assumed that each of them could generate, independently of the others, a number j of offspring with probability π_j $(j = 0, 1, ...)$, with $\pi_0 > 0$ and $\pi_0 + \pi_1 < 1$ to make extinction possible and to avoid the case where each generation consists at most of one individuals. This offspring constitutes the population after one generation. The process is then iterated: Each individual of the first generation produces j offspring with the same probabilities as above, thus leading to the second generation, and so on. Denoting by X_n the stochastic process representing the number of individuals forming generation n, Feller was able to construct a limit process by suitably making n go to infinity while simultaneously shrinking to zero the time elapsing between successive generations. He thus obtained a continuous process $X(t)$ of the diffusion time, whose transition pdf

$$f(x, t|x_0) = \frac{\partial}{\partial x} P\{X(t) \leq x|X(0) = x_0\} \tag{1}$$

satisfies a Fokker–Planck (FP) equation which is obtained for $q = 0$ from the more general FP equation (Feller, 1951b)

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}\left[\left(p\,x + q\right) f\right] + \frac{\partial^2}{\partial x^2}\left(r\,x\,f\right) \qquad 0 \leq x < \infty \tag{2}$$

with $p, q \in \Re$ and $r \in \Re^+$. Equation (2) is a very peculiar one. Indeed, Feller discovered that the behaviour of its solutions near the boundary $x = 0$ critically depends upon the numerical values of the parameters q and r. More precisely, if $0 < q < r$ the boundary $x = 0$ behaves as a " regular boundary", namely equation (2) can be solved together with an arbitrary boundary condition at $x = 0$, in general expressing absorption, reflection or a mixture of the two. If $q \leq 0$, $x = 0$ is an " exit boundary" in the sense that attainment of the boundary is possible from any starting point with a positive abscissa whereas from the boundary $x = 0$ no transition to the positive half-line can take place. Finally, if $q \geq r$ the point $x = 0$ is an " entrance boundary", meaning that particles can enter the positive half-line starting from the boundary $x = 0$ whereas the converse cannot occur. This is a very peculiar behaviour of the solutions of singular diffusion equations, such as equation (2), which after all emerged from the consideration of an instance of biological interest, i.e. the afore-mentioned diffusion approximation for a population subject to a branching process-like growth scheme.

The present chapter will primarily deals with neuronal models based on Markov stochastic processes of the diffusion type. However, I shall also say a few words on general Gaussian processes. In this case analytical results are scarce and fragmentary so that one is forced to rely on simulation algorithms. Here the way is full of " traps": Indeed, simulation of stochastic processes must be performed with extreme care to avoid to come to wrong conclusions or mere artefacts.

One more comment is in order. For a long time I have been deeply convinced that digital computers should be used as little as possible and, in any case, never as an investigation tool but only as the last resort when one really needs to come up with numerical answers to questions that cannot be answered by other means. However, during the past few years I had to admit to myself that I was wrong. Indeed, in more than one instance my co-workers and I have been led to discover some rather general and unsuspected theorems by a careful inspection of a large number of results obtained by purely numerical computations. I shall be more specific about this when talking about the asymptotics of certain neuronal diffusion models both in the absence and in the presence of oscillatory output behaviour.

Diffusion processes and neuronal modelling

Let us look at a neurone as a black box characterised by an input and by an output. We may then pose two distinct problems: (i) to determine the output for a given input, or (ii) to guess the input by analysing the output. In the present framework problem (i) is an *FPT problem*, while (ii) is an inverse problem, that is the *inverse of an FPT problem*. In both cases the class of input functions must be preliminarily specified to make these problems mathematically sound. As preannounced, in the sequel I shall

primarily assume that the neurone's membrane potential is modelled by a one-dimensional diffusion process.

It should be stressed that stochastic models of neuronal activity rank among the most advanced applications of the theory of stochastic processes to biology. There are several reasons why many of these neuronal models stem out of the theory of diffusion processes. First of all, the theory of diffusion processes is well developed; hence, it allows the neural modelers to apply many general results to this specific application area. Secondly, for neurones with many synaptic inputs there is a rather good correspondence between the models and the biological reality. Thirdly, even when some of the assumptions underlying the model may be somewhat questionable, the output behaviour of the neurone can often be well approximated by that of a suitably chosen diffusion process. This is, for instance, strikely true for the well-known Gerstein–Mandelbrot diffusion model. Finally, from a biophysical point of view neuronal models of single cells reflect the electrical properties of the membrane via electric circuit models that contain energy storage elements. Such circuit models can be described by means of differential equation for the membrane potential. When the input to the cell includes a random component, then the differential equation includes a " noisy" term and hence becomes itself a stochastic differential equation whose solution can sometimes be approximated by a diffusion process. In other cases, for instance when the neurone has very few synaptic inputs near the trigger zone, a Poisson driven differential equation may be a biologically more appropriate model (see Tuckwell, 1988, 1989 for a review).

An action potential (spike) is produced when the membrane voltage exceeds a threshold voltage; this situation corresponds to an FPT problem for the associated diffusion process.

Let us now turn to the mechanisms responsible for the production of stationary spike trains. In this case we are interested in modelling the "spontaneous" or " resting" activity of a neurone, or the steady-state response to a constant stimulus. Such an analysis is indeed usually required before attempting to model the neurone's response to time varying stimuli. The interested reader can refer to the large body of existing literature on the derivation of other models. Several reviews on stochastic neuronal models are indeed available (see, for instance, Fienberg, 1974; Holden, 1976; Sampath and Srinivasan, 1977; Yang and Chen, 1978; Lee, 1979; Tuckwell, 1988, 1989). The construction of some of the diffusion approximations considered in the sequel can be found in the above reviews or in a number of research papers including this topic (cf. Gerstein and Mandelbrot, 1964; Gluss, 1967; Johannesma, 1968; Roy and Smith, 1969; Capocelli and Ricciardi, 1971; Ricciardi, 1976; Ricciardi and Sacerdote, 1979; Tuckwell and Cope, 1980; Hanson and Tuckwell, 1983; Kallianpur, 1983; Ricciardi and Sato, 1983, Lánský, 1984; Lánský and Lánská, 1987; Giorno *et al.*, 1988a). Finally, many results can be found in the monographs by Goel and Richter–Dyn, (1974) and Ricciardi (1977).

It should be mentioned that at a microscopic level, the biophysical properties of nerve membrane are based on the opening and closing of ionic channels. Models based on FPT problems are important here as well. The movement of an ion through a pore can indeed be viewed as the random movement of a particle in a multistable potential well. (For a recent review see Cooper *et al.*, 1988).

Let us now discuss briefly the ways in which it is customary to construct and to describe a scalar diffusion process. To this purpose I have to recall some basic notation and background. Hence, I shall briefly talk about stochastic differential equations, Fokker–Planck equations and FPT problems.

A scalar diffusion process $\{X(t);\ t \geq t_0,\ t_0 \in \Re\}$ can be described by the stochastic differential equation (SDE)

$$dX(t) = \mu\,[X(t), t]\ dt + \sigma\,[X(t), t]\ dW(t), \tag{3}$$

where $W = \{W(t);\ t \geq t_0\}$ is a standard Wiener process and μ and σ are real-valued functions of their arguments satisfying certain regularity conditions (Karlin and Taylor, 1981). Here the process $X(t)$ represents changes in the membrane potential between two consecutive neuronal firings (spikes). The reference level for the membrane potential is usually taken to be equal to the resting potential. The initial voltage, namely the reset value following a spike, is often assumed to be equal to the resting potential, viz. $X(t_0) = x_0$. The threshold potential, denoted by $S = S(t)$ $[S(t_0) > x_0]$, will be assumed to be a deterministic function of time. In the sequel we shall mainly be interested in the properties of the FPT random variable (r.v.) T, defined as

$$T = \inf_{t \geq t_0} \{t : X(t) > S(t) | X(t_0) = x_0\}. \tag{4}$$

Hence, T is the theoretical counterpart of the interspike interval. The importance of interspike intervals is due to the generally accepted hypothesis that the information transferred within the nervous system is usually encoded by the timing of occurrence of neuronal spikes. Therefore the reciprocal relationship between the firing frequency and the interspike interval naturally leads to the problem of determining the probability density function (pdf) of T, namely the function

$$g[S(t), t | x_0, t_0] = \frac{\partial}{\partial t}\ P\{T \leq t\}. \tag{5}$$

When this density cannot be obtained analytically (which is practically the rule) or when it is too difficult to give sufficiently precise estimations of it, the analysis is restricted to its moments, primarily mean and variance. The coefficient of variation is also used as it measures the relative spread of the distribution and its deviation from exponentiality. In some cases it can finally be useful to work out asymptotic estimates of the FPT pdf for large times or for large thresholds, as it will be briefly indicated in the sequel.

An alternative description of the process $X(t)$ is sometimes obtained via the so called "diffusion equations approach". First of all, in analogy with (1), one defines the pdf of $X(t)$ conditional on $X(t_0) = x_0$:

$$f(x,t|x_0,t_0) = \frac{\partial}{\partial x} P\{X(t) \leq x | X(t_0) = x_0\}\,. \tag{6}$$

It can then be seen that f satisfies the Fokker–Planck (FP) equation

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}\left[A_1(x,t)\,f\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[A_2(x,t)\,f\right] \tag{7}$$

and the Kolmogorov (K) equation

$$\frac{\partial f}{\partial t_0} + A_1(x_0,t_0)\frac{\partial f}{\partial x_0} + \frac{1}{2}A_2(x_0,t_0)\frac{\partial^2 f}{\partial x_0^2} = 0\,, \tag{8}$$

where the coefficients $A_1(x,t)$ and $A_2(x,t)$ are the "drift" and "infinitesimal variance" of the process defined as

$$A_i(x,t) = \lim_{\Delta t \downarrow 0}\frac{1}{\Delta t}\int (y-x)^i\, f(y,t+\Delta t|x,t)\,dy$$

$$\equiv \lim_{\Delta t \downarrow 0}\frac{1}{\Delta t}\, E\{[X(t+\Delta t) - X(t)]^i | X(t) = x\} \qquad (i = 1,2). \tag{9}$$

It is essential to mention that the quantities $A_1(x,t)$ and $A_2(x,t)$ just defined are related to the functions μ and σ earlier considered when talking about the differential equations approach. However, the literature here is at times misleading, to the point to have generated the so called "Ito–Stratonovich" controversy. Without embarking here in a discussion of this point (see, for instance, Ricciardi 1977), we limit ourselves to mentioning that within the so called Ito approach the relation is the following:

$$\begin{cases} A_1(x,t) = \mu(x,t) \\ A_2(x,t) = \sigma^2(x,t). \end{cases} \tag{10}$$

The second approach, known as the " Stratonovich approach" relates these functions as follows:

$$\begin{cases} A_1(x,t) = \mu(x,t) + \dfrac{1}{4}\dfrac{d}{dx}A_2(x,t) \\ \qquad\qquad A_2(x,t) = \sigma^2(x,t). \end{cases} \tag{11}$$

(Note that the two approaches yield the same result only if the infinitesimal variance of the process is a constant. The meaning is that in general

they actually model different processes!). To determine f via the diffusion equations the following initial conditions must be considered:

$$\lim_{t \downarrow t_0} f(x, t|x_0, t_0) = \delta(x - x_0) \tag{12a}$$

$$\lim_{t_0 \uparrow t} f(x, t|x_0, t_0) = \delta(x_0 - x) \tag{12b}$$

expressing the circumstance that initially the whole probability mass is concentrated at the initial value x_0. However, these conditions are not always sufficient to determine uniquely the transition pdf and suitable boundary conditions have to be imposed (Feller, 1952, 1954; Karlin and Taylor, 1981).

In view of the considerations to come, it must be mentioned that equations (7) and (8) characterize the class of diffusion processes. These are special Markov processes. We recall that a process with continuous state space in continuous time is Markov if the function (6) satisfies the following integral relation (Chapman–Kolmogorov or Smolukowski equation):

$$f(x, t|x_0, t_0) = \int_{\Re} f(x, t|y, \tau)\, f(y, \tau|x_0, t_0)\, dy \tag{13}$$

for all $x, x_0, y \in \Re$ and for all t_0, τ, t such that $t_0 < \tau < t$. Hence, the denomination of "transition" pdf given to the conditional density (6) in the case of Markov processes. From (13) the so called "forward" and "backward" differential expansions can be derived (cf., for instance, Ricciardi, 1977):

$$\frac{\partial f}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \left[A_n(x, t)\, f\right] \qquad (forward) \tag{14}$$

and

$$\frac{\partial f}{\partial t_0} + \sum_{n=1}^{\infty} \frac{1}{n!} A_n(x_0, t_0) \frac{\partial^n f}{\partial x_0^n} = 0 \qquad (backward), \tag{15}$$

where the functions $A_n(x, t)$ are the so called "infinitesimal moments" defined as follows:

$$A_n(x, t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \int (y - x)^n\, f(y, t + \Delta t|x, t)\, dy$$

$$\equiv \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} E\{[X(t + \Delta t) - X(t)]^n | X(t) = x\} \qquad (n = 1, 2, \ldots). \tag{16}$$

Note that (16) contain the definitions (9) of drift and infinitesimal variance as a special case. The role of the two descriptions, equation (3) vs. equations (7) and (8), of the membrane potential $X(t)$ is now briefly outlined. As (3) describes the time evolution of the trajectories of the membrane potential,

a discretized version of it can be obtained by a mesh of points t_i such that $t_0 < t_1 < \ldots < t_n = t$. Equation (3) can then be re-written as

$$x(t_{i+1}) = x(t_i) + \mu\left[x(t_i), t_i\right](t_{i+1} - t_i) + \sigma\left[x(t_i), t_i\right]\left[W(t_{i+1}) - W(t_i)\right]. \quad (17)$$

Hence, an approximation procedure for solving (3) consists of computing $x(t_{i+1})$ in terms of $x(t_i)$ realizing that the increments of the standard Wiener process appearing in (17) are mutually independent, normally distributed r.v.s with zero mean and variances $t_{i+1} - t_i$, independent of $x(t_i)$. However, such a procedure is too crude, and thus often not immune from the risk of generating too poor an approximation to the diffusion process $X(t)$ or of yielding unacceptable approximation errors. More sophisticated methods for the simulation of (17) can be found, for instance, in Rümelin (1982), Pardoux and Talay (1985).

The neuronal models based on the applications of diffusion processes are predominantly time homogeneous and thus the functions μ and σ appearing in (3), (7) and (8) do not depend explicitly on t, i.e., $\mu(x,t) = \mu(x)$, $\sigma(x,t) = \sigma(x)$. Consequently, one has $f(x,t|x_0,t_0) = f(x,t-t_0|x_0,0) = f(x,t-t_0|x_0)$. Time nonhomogeneous diffusion models have been mentioned by Matsuyama *et al.* (1974), Matsuyama (1976), Ricciardi (1982) and Lánský (1984). However, since very few analytical results are available, simulation methods are required.

The firing pdf and its stochastic counterpart

Hereafter we shall outline some basic ideas underlying the determination of the firing pdf for a neurone modeled by a diffusion process. As already mentioned, this is an FTP problem in which the unknown is the FPT pdf through a preassigned boundary $S(t)$. Let us now define the transition pdf $\alpha(x,t|x_0,t_0)$ of $X(t)$ in the presence of an absorbing boundary $S(t)$. This means that the process is terminated when the membrane potential reaches the threshold. Mathematically, we set

$$\alpha(x,t|x_0,t_0) = \frac{\partial}{\partial x} P\{X(t) \le x;\ X(\vartheta) < S(\vartheta)\ \ \forall\vartheta \in [t_0,t]|X(t_0) = x_0\}. \quad (18)$$

It can be shown that such a function can be obtained by solving eqs (7) and (8) subject not only to initial conditions (12) but also to the boundary condition $f[S(t),t|x_0,t_0] = 0$. A second boundary condition, expressing the behaviour of $X(t)$ at the other hand of the interval where diffusion takes place, must be added. This may be a regularity condition, a reflection condition or other kinds of conditions depending upon the model one is dealing with.

The functions (5) and (18) are intimately related. Indeed, denoting by ν the lower bound of the state space ($-\infty$ if we assume that the membrane

potential can be hyperpolarized without bounds), the following identity clearly holds:

$$P\{\nu < X(t) < S(t);\ X(\vartheta) < S(\vartheta)\ \ \forall\vartheta \in [t_0, t] | X(t_0) = x_0\} = P\{T > t\}. \tag{19}$$

Indeed, the left-hand-side is the probability that the membrane potential has never attained the firing threshold up to time t, namely that the FPT of $X(t)$ through $S(t)$ is greater than t. In terms of definitions (5) and (18) identity (19) can be re-written as

$$\int_{\nu}^{S(t)} \alpha(x, t|x_0, t_0)\, dx = \int_{t}^{\infty} g[S(\tau), \tau|x_0, t_0]\, d\tau. \tag{20}$$

Hence,

$$g[S(t), t|x_0, t_0] = -\frac{\partial}{\partial t}\int_{\nu}^{S(t)} \alpha(x, t|x_0, t_0)\, dx. \tag{21}$$

The left-hand-side is the firing pdf. Its determination thus requires the determination of the transition pdf of $X(t)$ in the presence of an absorbing boundary at $S(t)$.

Let us now make the assumption (to be removed later) that the neuronal threshold is a constant, i.e. $S(t) = S$, with $S > x_0$ a constant. Also, let us assume that $X(t)$ is time-homogeneous. From (21) we then obtain:

$$g(S, t|x_0) = -\int_{\nu}^{S} \frac{\partial}{\partial t}\, \alpha(x, t|x_0)\, dx = \left\{A_1(x)\, \alpha - \frac{1}{2}\frac{\partial}{\partial x}\left[A_2(x)\, \alpha\right]\right\}\Big|_{\nu}^{S} \tag{22}$$

where we have used the fact that α satisfies Eq. (7) with A_1 and A_2 independent of t.

An alternative approach is the following. Since the sample paths of $X(t)$ are continuous functions, any sample path that reaches a state $x > S$ at time zero must necessarily cross S for the first time at some intermediate time τ $(0 < \tau < t)$. We can then write

$$f(x, t|x_0) = \int_0^t f(x, t-\tau|S)\, g(S, \tau|x_0)\, d\tau \qquad (x > S). \tag{23}$$

This is a first-kind Volterra integral equation in the unknown firing pdf $g(S, t|x_0)$. Its solution is made complicated by the circumstance that the kernel $f(x, t-\tau|S)$ exhibits a singularity of the type $1/\sqrt{t-\tau}$ as $\tau \uparrow t$. Hence, the problem of determining $g(S, t|x_0)$ from Eq. (23) via numerical methods is by no means trivial (cf. Buonocore *et al.*, 1987 and references therein). An analytic approach to the solution of Eq. (23) is based on the Laplace transform (LT). Set

$$\begin{cases} g_\lambda(S|x_0) = \mathcal{L}_t\{g(S, t|x_0)\} \\ f_\lambda(x|y) = \mathcal{L}_t\{f(x, t|y)\} \\ f_\lambda(x|x_0) = \mathcal{L}_t\{f(x, t|x_0)\}, \end{cases} \tag{24}$$

where $\mathcal{L}_t$ denotes the LT with respect to t of the function in curly brackets. Henceforth, we shall safely assume $\lambda > 0$. Since on the right-hand-side of (23) appears a convolution integral, passing to the LT we obtain:

$$g_\lambda(S|x_0) = \frac{f_\lambda(x|x_0)}{f_\lambda(x|S)} \qquad \forall x > S. \tag{25}$$

Note that the right-hand-side of (25) is actually independent of the arbitrarily chosen state x. Hence, if the transition pdf of $X(t)$ is known and if its LT can be calculated, the right-hand-side of (25) can be written down. The function $g(S,t|x_0)$ can then be obtained as an inverse LT. Alternatively, we can obtain $g_\lambda(S|x_0)$ as a solution of an ordinary differential equation. Indeed, the Kolmogorov equation (8) for a time-homogeneous process reads

$$\frac{\partial f}{\partial t} = A_1(x_0)\,\frac{\partial f}{\partial x_0} + \frac{1}{2}\,A_2(x_0)\,\frac{\partial^2 f}{\partial x_0^2}. \tag{26}$$

Taking the LT of both sides we obtain:

$$\mathcal{L}_t\Big\{\frac{\partial f(x,t|x_0)}{\partial t}\Big\} = A_1(x_0)\,\frac{\partial f_\lambda(x|x_0)}{\partial x_0} + \frac{1}{2}\,A_2(x_0)\,\frac{\partial^2 f_\lambda(x|x_0)}{\partial x_0^2}. \tag{27}$$

We now have:

$$\mathcal{L}_t\Big\{\frac{\partial f(x,t|x_0)}{\partial t}\Big\} = \int_0^\infty e^{-\lambda t}\,\frac{\partial f(x,t|x_0)}{\partial t}\,dt$$

$$= \lim_{t\downarrow 0}\,\big[e^{-\lambda t}\,f(x,t|x_0)\big] + \lambda\int_0^\infty e^{-\lambda t}\,f(x,t|x_0)\,dt. \tag{28}$$

Since the limit in (28) is zero by virtue of the initial condition (12*a*) and of the assumption $x > S > x_0$, from (25), (27) and (28) we finally obtain:

$$\frac{A_2(x_0)}{2}\,\frac{d^2 g_\lambda(S|x_0)}{dx_0^2} + A_1(x_0)\,\frac{dg_\lambda(S|x_0)}{dx_0} - \lambda\,g_\lambda(S|x_0) = 0. \tag{29}$$

This is a second order ordinary differential equation, that must be solved with the conditions

$$g_\lambda(S|S) = 1 \tag{30a}$$

$$g_\lambda(S|x_0) < \infty, \qquad \forall x_0 < S. \tag{30b}$$

Equation (30*a*) expresses the circumstance that the FPT through S starting from S itself is a degenerate random variable that takes the value 0 with probability 1 (instantaneous first-passage). Hence, $g(S,t|S) = \delta(t)$ and hence its LT is unit. Equation (30b) is instead a regularity condition whose role is to discard solutions of (29) that blow up and thus cannot be pdf's.

It is worth to point out that even though the inverse LT of the function $g_\lambda(S|x_0)$, obtained either by means of (25) or by solving Eq. (29) with condition (30), cannot be calculated it can nevertheless provide useful information on the FPT. Indeed, $g(S,t|x_0)$ is the moment generating function of T, so that

$$t_n(S|x_0) = E(T^n) = \frac{d^n g_\lambda(S|x_0)}{d\lambda^n}\Big|_{\lambda=0} \qquad (n = 1, 2, \ldots). \tag{31}$$

We conclude this section by noting that in the above considerations we have made the assumption that the initial state $X(t_0) = x_0$ is fixed. However, some extensions are possible in which the initial state is viewed as a random variable with an assigned distribution function (cf. Lánský and Smith, 1989; Giorno *et al.*, 1992; Ricciardi *et al.*, 1992)

Neuronal models

The year 1964 marks the beginning of the history of neuronal models based on diffusion processes. Indeed, in a much celebrated article Gerstein and Mandelbrot (1964) postulated that for a number of experimentally monitored neurones subject to spontaneous activity the firing pdf could be modelled by the FPT pdf of a Wiener process. Actually, these authors were able to show that, by suitably choosing the parameters of the model, the experimentally recorded interspike intervals histograms of many units could be fitted to an excellent degree of approximation by means of the FPT pdf of a diffusion process characterised by the constant infinitesimal moments

$$A_1(x) = \mu, \qquad \mu \in \Re \tag{32a}$$

$$A_2(x) = \sigma^2, \qquad \sigma \in \Re^+. \tag{32b}$$

As is well known, in this case the transition pdf (6) is normal, with mean $\mu(t-t_0)+x_0$ and variance $\sigma^2(t-t_0)$. Since this is a temporally homogeneous diffusion process, we can take $t_0 = 0$. By means of the methods outlined earlier, one can then prove that the FPT pdf of such a process is given by

$$g(S,t|x_0) = \frac{S - x_0}{\sigma\sqrt{2\pi t^3}} \exp\Big[-\frac{(S - x_0 - \mu t)^2}{2\sigma^2 t}\Big], \qquad x_0 < S. \tag{33}$$

This function is often appropriately called the "inverse Gaussian" (IG) pdf because its cumulant generating function is the inverse of the cumulant generating function of a Gaussian random variable (cf. Chhikara and Folks, 1989, for a survey of its properties). For $\mu \geq 0$, neuronal firing is a sure event as (33) is normalised to unity. If one takes $\mu < 0$, the FPT pdf (33) can be interpreted as the firing pdf conditional upon the event "firing occurs".

The case $\mu = 0$ is also of interest since (33) is a stable law (Feller, 1966). This case provides an interpretation of numerous experimental results indicating that the shapes of histograms are sometimes preserved when the adjacent interspike intervals are summed (see, for instance, Rodieck *et al.* 1962; Gerstein and Mandelbrot, 1964; Holden, 1975).

To provide an interpretation of the neuronal model based on assumptions (32), let us imagine that the neurone's membrane potential undergoes a simple random walk under the effect of excitatory and inhibitory synaptic actions. For simplicity, let us assume that the neuronal dynamics develops on a discrete time scale, viz. that time is quantized as a sequence of instants $0, \tau, 2\tau, \ldots$, with $\tau > 0$ an arbitrary time unit. We shall then denote by $X_{n\tau}$ $(n = 0, 1, \ldots)$ the membrane potential at time $n\tau$ and assume that the random walk depicting the neuronal dynamics is of the following type:

$$X_{(n+1)\tau} = X_{n\tau} + Z_{n\tau} \qquad (n = 0, 1, \ldots) , \tag{34}$$

where $\{Z_{n\tau}\}_n$ is a sequence of independent identically distributed random variables. More precisely, we assume that

$$P\{Z_{n\tau} = \epsilon\} = p , \qquad P\{Z_{n\tau} = -\epsilon\} = q \equiv 1 - p \qquad (n = 0, 1, \ldots) \tag{35}$$

where ϵ is a positive real number. These assumptions model the time course of the neuronal membrane potential as a random walk characterised by jumps of magnitudes ϵ and $-\epsilon$ that, in turn, are representative of the excitatory and inhibitory post-synaptic potentials (PSPs). These are equally likely to occur if $p = q$, whereas $p > q$ $(p < q)$ implies the existence of a trend leading the random walk towards (away from) the firing threshold S. The quantities τ and ϵ can now be chosen in such a way that the random walk (33) converges to a continuous-time stochastic process. To this purpose, let us impose that mean and second order moment of the PSPs are both proportional to the time unit τ:

$$E(Z_{n\tau}) = \mu\tau , \qquad \mu \in \Re \tag{36a}$$

$$E(Z_{n\tau}^2) = \sigma^2\tau \qquad \sigma \in \Re^+ . \tag{36b}$$

Making use of (35), relations (36) yield:

$$\epsilon p - \epsilon q = \mu\tau , \qquad \mu \in \Re \tag{37a}$$

$$\epsilon^2 = \sigma^2\tau , \qquad \sigma \in \Re^+ . \tag{37b}$$

Recalling that $p + q = 1$, from eq. (37) we obtain the specification of the PSPs under constraints (36):

$$P\{Z_{n\tau} = \sigma\sqrt{\tau}\} = \frac{1}{2} + \mu\frac{\sqrt{\tau}}{2\sigma} ,$$

$$P\{Z_{n\,\tau} = -\sigma\sqrt{\tau}\} = \frac{1}{2} - \mu\,\frac{\sqrt{\tau}}{2\,\sigma} \longleftrightarrow (\mu \in \Re,\ \sigma \in \Re^{+}). \tag{38}$$

By the above procedure we have been led to model the time course of the neurone's membrane potential as a random walk described by equations (34) involving the random PSPs (38). It is now easy to see from (34) and (38) that the following relations hold:

$$\frac{1}{\tau}\,E\{X_{(n+1)\,\tau} - X_{n\,\tau}|X_{n\,\tau} = x\} = \mu \tag{39a}$$

$$\frac{1}{\tau}\,E\{[X_{(n+1)\,\tau} - X_{n\,\tau}]^2|X_{n\,\tau} = x\} = \sigma^2 \tag{39b}$$

$$\frac{1}{\tau}\,E\{[X_{(n+1)\,\tau} - X_{n\,\tau}]^{2+r}|X_{n\,\tau} = x\} = o(\tau) \qquad (r = 1, 2, \ldots). \tag{39c}$$

Passing to the limit in eq. (39) as $\tau \to 0$ and recalling (9), we conclude that the random walk $X_{n\,\tau}$ converges to the diffusion process having drift μ and infinitesimal variance σ^2, viz. to a Wiener process.

The above procedure provides the simplest example in which the Wiener model for neuronal activity can be constructed. Although the underlying assumptions are undoubtedly oversimplified and some well-known electrophysiological properties of neuronal membrane are not taken into account, it must be stressed that the fitting of some experimental data by the FPT pdf (33) is truly remarkable (Gerstein and Mandelbrot, 1964). This is probably the reason why some authors have attempted to generalise the Wiener process in various respects. For instance, Lánský and Smith (1989) have analysed the case when the Wiener model is characterised by a random initial value X_0 such that

$$P\{X_0 = x_{0,1} < S\} = p_1\,, \qquad P\{X_0 = x_{0,2} < S\} = p_2 \equiv 1 - p_1. \tag{40}$$

The firing pdf then becomes the mixture of IG pdf's

$$g(S, t|X_0) = \frac{1}{\sigma\sqrt{2\,\pi\,t^3}}\sum_{i=1}^{2} p_i\,(S - x_{0,i})\,\exp\Big\{-\frac{(S - x_{0,i} - \mu\,t)^2}{2\,\sigma^2\,t}\Big\}\,, \tag{41}$$

for which Amoh (1984) has raised some parameter estimation problems. Another generalization has been proposed by Wise (1981) who has substituted (33) with a generalised IG pdf that Barndorff–Nielsen *et al.* (1978) have obtained as the FPT pdf for a time-homogeneous diffusion process (cf. Lánský *et al.*, 1990).

In (33) the firing threshold is assumed to be a constant. If, more realistically, a time-varying threshold is introduced to account for relative refractoriness — which is relevant in high firing rate conditions — the determination of the firing pdf for the Wiener neuronal model cannot be accomplished analytically unless the firing threshold is unrealistically assumed to

be linearly time-dependent. For other types of threshold functions, *ad hoc* numerical methods had to be devised (cf., for instance, Durbin, 1971; Smith, 1972; Anderssen *et al.*, 1973; Park and Shuurmann, 1976, 1980; Buonocore *et al.*, 1987 and references therein).

In conclusion, despite the excellent fitting of some data, Gerstein and Mandelbrot's neuronal model based on the Wiener process has been the object of various criticisms (cf. Tuckwell, 1988). We limit ourselves to pointing out that this model does not include the well-known spontaneous exponential decay of the neurone's membrane potential that occurs between successive PSPs. Hereafter we shall thus discuss another diffusion model for neuronal activity that includes this specific feature. This is customarily denoted as the Ornstein–Uhlenbeck (OU) neuronal model because of its analogy with the model used by these two authors to describe Brownian motion. Indeed, the OU neuronal model is defined as the diffusion process characterised by the following drift and infinitesimal variance:

$$A_1(x) = -\frac{x}{\vartheta} + \mu \tag{42a}$$

$$A_2(x) = \sigma^2 \tag{42b}$$

where $\mu \in \Re$ while ϑ and σ are positive constants. Comparing (42) with (32) we see that now the drift is state-dependent. However, in the limit as $\vartheta \to \infty$ moments (42) tend to moments (32), meaning that the OU model yields the Wiener model for infinitely large ϑ. Recalling (3), (10) and (11) we can interpret the OU model as generated by the following SDE:

$$dX(t) = \left(-\frac{X}{\vartheta} + \mu\right) dt + \sigma\, dW(t)\,, \tag{43}$$

where $W(t)$ is a standard Wiener process. Note that since A_2 is a constant, both Ito and Stratonovich's approaches lead to the same diffusion process. Equation (43) is taken as representative of the time course of the membrane potential. In the absence of randomness ($\sigma = 0$), equation (43) yields the spontaneous exponential decay of the membrane potential towards the resting potential ρ, with $\rho = \mu\vartheta$. Hence, ϑ can be viewed as the time-constant of the neurone's membrane (approximately 5 ms). The above mentioned limit as $\vartheta \to \infty$ would deprive the membrane model of its spontaneous exponential decay and thus change equation (43) into

$$dX(t) = \mu\, dt + \sigma\, dW(t)\,, \tag{44}$$

viz. into the SDE for the Wiener process with drift μ and infinitesimal variance σ^2.

Although equation (43) can be taken as definitory of the OU model, we shall now again resort to a simple procedure that is effective to show how such a model can be obtained from first principles, in particular by

using the formalism of diffusion equations. To this purpose, let us initially assume that the neurone is subject to a sequence of excitatory and inhibitory PSPs of constant magnitudes $e > 0$ and $i < 0$ occurring in time with Poisson laws of parameters α_e and α_i, respectively. Hence, the membrane potential is viewed as a stochastic process $X(t)$ in continuous time with a discrete space consisting of the lattice $x_0 + k\,i + h\,e \quad (k, h = 0, \pm 1, \ldots)$ with the points of discontinuity randomly occurring in time. Denoting by x_0 ($x_0 < S$) the resting potential, we assume that all sample paths start at x_0 at the fixed initial time $t_0 = 0$. The firing pdf is the pdf of the instants at which the sample paths for the first time reach, or cross in the upward direction, the barrier S. The process $X(t)$ is Markov, and thus its transition pdf satisfies equation (13). Let us first assume that $\vartheta = \infty$ and let us make the substitutions $t \to t + \Delta t$, $\tau \to t$, $t_0 \to 0$ in equation (13). Because of the assumed Poisson nature of PSPs, apart from infinitesimal quantities $o(\Delta t)$, equation (13) now reads:

$$f(x, t + \Delta t | x_0) = \left[1 - (\alpha_e + \alpha_i)\,\Delta t\right] f(x, t | x_0) + \alpha_e\,\Delta t\, f(x - e, t | x_0)$$

$$+\alpha_i\,\Delta t\, f(x - i, t | x_0). \tag{45}$$

Intuitively, this equation says that a sample path is in the neighbourhood of x at time $t + \Delta t$ if at time t it is already in the neighbourhood of x *and* no PSPs occur in $(t, t + \Delta t)$, or if it is around $x - e$ $(x - i)$ at time t and one excitatory (inhibitory) PSP occurs during the time interval $(t, t + \Delta t)$. From (45), in the limit as $\Delta t \downarrow 0$ one obtains:

$$\frac{\partial f}{\partial t} = \alpha_e \left[f(x - e, t | x_0) - f(x, t | x_0)\right] + \alpha_i \left[f(x - i, t | x_0) - f(x, t | x_0)\right]. \tag{46}$$

Expanding the functions in square brackets as Taylor series about x and setting

$$\mu_n = \alpha_e\, e^n + \alpha_i\, i^n \qquad (n = 1, 2, \ldots), \tag{47}$$

equation (46) becomes:

$$\frac{\partial f}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\, \mu_n\, \frac{\partial^n f}{\partial x^n}. \tag{48}$$

Since by making use of (16) the quantities (47) can be seen to be the infinitesimal moments of $X(t)$, equation (48) is the forward differential expansion (14) of the Smolukowski equation. Note that if we set $\mu_n = 0$ for $n > 2$ in (48), we re-obtain the Wiener process with drift $\mu = \alpha_e\, e + \alpha_i\, i$ and infinitesimal variance $\sigma^2 = \alpha_e\, e^2 + \alpha_i\, i^2$. Let us now assume that $0 < \vartheta < \infty$. Then, $X(t)$ is a Markov process with a continuum of states because now

the membrane potential exponentially decays to the zero resting potential in between successive PSPs. Equation (13) thus yields:

$$f(x,t+\Delta t|x_0) = \int_{-\infty}^{+\infty} f(x,t+\Delta t|y,t)\, f(y,t|x_0)\, dy\,. \tag{49}$$

Let us now remark that the first transition pdf under the integral sign can be expressed, apart from quantities $o(\Delta t)$, as follows:

$$f(x,t+\Delta t|y,t) = \left[1-(\alpha_e+\alpha_i)\,\Delta t\right]\,\delta(y_0-x)$$

$$+\alpha_e\,\Delta t\,\delta(y_1-x)+\alpha_i\,\Delta t\,\delta(y_2-x) \tag{50}$$

where we have set:

$$\begin{cases} y_0 = y - y\,\frac{\Delta t}{\vartheta} \\ y_1 = y - y\,\frac{\Delta t_1}{\vartheta} + e - \left(y - y\,\frac{\Delta t_1}{\vartheta} + e\right)\frac{\Delta t_2}{\vartheta} \\ y_2 = y - y\,\frac{\Delta t_1}{\vartheta} + i - \left(y - y\,\frac{\Delta t_1}{\vartheta} + i\right)\frac{\Delta t_2}{\vartheta} \end{cases} \tag{51}$$

with $\Delta t_1 + \Delta t_2 = \Delta t$. We now substitute (51) in (50) and use the identity

$$\delta[\phi(y)] = \frac{\delta(y-\bar{y})}{|\phi'(\bar{y})|}\,, \tag{52}$$

holding if $\phi(y)$ is a monotonic function vanishing for $y=\bar{y}$, to obtain:

$$f(x,t+\Delta t|y,t) = (1-\Delta t/\vartheta)^{-1}\left\{\left[1-(\alpha_e+\alpha_i)\,\Delta t\right]\,\delta\left(y-\frac{x}{1-\Delta t/\vartheta}\right)\right.$$

$$+\alpha_e\,\Delta t\,\delta\left(y-\frac{x-e+e\,\Delta t_2/\vartheta}{1-\Delta t/\vartheta}\right)$$

$$\left.+\alpha_i\,\Delta t\,\delta\left(y-\frac{x-i+i\,\Delta t_2/\vartheta}{1-\Delta t/\vartheta}\right)\right\}. \tag{53}$$

Making use of (53) and writing $(1-\Delta t/\vartheta)^{-1} \simeq 1+\Delta t/\vartheta$ we then have:

$$f(x,t+\Delta t|x_0) = \left(1+\frac{\Delta t}{\vartheta}\right)\left\{\left[1-(\alpha_e+\alpha_i)\Delta t\right]\, f\left(x+x\,\frac{\Delta t}{\vartheta},t\middle|x_0\right)\right.$$

$$+\alpha_e\,\Delta t\, f\left[x-e+(x-e)\,\frac{\Delta t}{\vartheta},t\middle|x_0\right]$$

$$\left.+\alpha_i\,\Delta t\, f\left[x-i+(x-i)\,\frac{\Delta t}{\vartheta},t\middle|x_0\right]\right\}. \tag{54}$$

Hence, in the limit as $\Delta t \to 0$ we find that $f(x,t|x_0)$ satisfies the following equation:

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial x}\left(\frac{x}{\vartheta} f\right) + \alpha_e \left[f(x-e,t|x_0) - f(x,t|x_0)\right]$$

$$+\alpha_i \left[f(x-i,t|x_0) - f(x,t|x_0)\right]. \tag{55}$$

This is clearly a generalization of equation (46) to the case of finite time constant ϑ. Similarly to the previous case we can then prove that (55) can be written as

$$\frac{\partial f}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n}(A_n f), \tag{56}$$

with

$$A_1(x) = -x/\vartheta + \alpha_e e + \alpha_i i$$

$$A_n = \alpha_e e^n + \alpha_i i^n \qquad (n = 2,3,\ldots). \tag{57}$$

The form of the expansion (56) indicates that the quantities (57) are the infinitesimal moments of the process. This can, however, be directly proved by using (16) with the "infinitesimal" transition pdf given by (50). To obtain a differential equation out of the infinite expansion (56) one might be tempted to assume that the PSPs are small enough to allow the A_ns (57) to be disregarded from a certain n on. However, such a procedure is incorrect. While referring to Capocelli and Ricciardi (1971) and to Ricciardi (1977) for further details, here we limit ourselves to indicating how a diffusion approximation can be obtained provided one takes the magnitudes of the PSPs infinitely small and the rates of arrival of synaptic inputs infinitely large, with proper precautions. Let us denote by $\delta > 0$, σ_i^2, σ_e^2 and $k_i < 0$ otherwise arbitrary quantities and set:

$$\alpha_e = \left(\frac{\sigma_e}{\delta}\right)^2, \quad \alpha_i = \left(\frac{\sigma_i}{\delta}\right)^2, \quad i = k_i\,\delta, \quad e = -k_i\left(\frac{\sigma_i}{\sigma_e}\right)^2 \delta. \tag{58}$$

From (57) we then obtain:

$$\begin{cases} A_1 = -x/\vartheta + \sigma_i^2(k_i - k_i) = -x/\vartheta \\ A_n = \delta^{n-2} k_i^n \sigma_i^2 \left[(-1)^n \left(\frac{\sigma_i}{\sigma_e}\right)^{2n-2} + 1\right] \qquad (n = 2,3,\ldots) \end{cases} \tag{59}$$

which, in the limit as $\delta \to 0$, yield:

$$\begin{cases} A_1 = -x/\vartheta \\ A_2 = k_i^2 \sigma_i^2 \left[(\sigma_i/\sigma_e)^2 + 1\right] \equiv \sigma^2 \\ A_n = 0 \qquad (n = 3,4,\ldots). \end{cases} \tag{60}$$

The above sketched procedure has led us to conclude that if the input rates are taken larger and larger while simultaneously taking smaller and

smaller PSPs with the constraint (58), the membrane potential "converges" to the diffusion process having infinitesimal moments (60). These identify with moments (42) if $\mu = 0$. The case $\mu \neq 0$ can be obtained by a slightly more complicated model in which multiple Poisson-distributed excitatory and inhibitory inputs impinge on the neuronal membrane. For brevity, here we skip the description (see Ricciardi, 1976, 1977).

Let us now return to the OU model with infinitesimal moments (42). Its transition pdf is obtained by solving either the Fokker–Planck equation (7) or the Kolmogorov equation (8) with the initial conditions (12). One thus finds (cf., for instance, Ricciardi, 1977):

$$f(x,t|x_0) = \left[2\,\pi\,V(t)\right]^{-1/2} \exp\left\{-\frac{\left[x - M(t|x_0)\right]^2}{2\,V(t)}\right\}, \tag{61}$$

where

$$M(t|x_0) = \mu\,\vartheta\left(1 - e^{-t/\vartheta}\right) + x_0\,e^{-t/\vartheta}$$

$$V(t) = \frac{\sigma^2\,\vartheta}{2}\left(1 - e^{-2t/\vartheta}\right). \tag{62}$$

Hence, at each time t the transition pdf is normal with mean $M(t|x_0)$ and variance $V(t)$. It must be pointed out that the OU model differs from the Wiener model in some relevant respects. For instance, an equilibrium regime exists since in the limit as $t \to \infty$ the pdf (61) becomes normal with mean $\mu\,\vartheta$ and variance $\sigma^2\,\vartheta/2$. Furthermore, the crossing of the firing threshold is a sure event. However, differently from the Wiener model, the FPT problem is in general very complicated, even in the case of constant thresholds. Indeed, analytical solutions can only be determined in very particular cases, namely for thresholds of the type

$$S(t) = \mu\,\vartheta + A\,e^{-t/\vartheta} + Be^{t/\vartheta}, \tag{63}$$

with A and B arbitrary constants (Siebert, 1969; Sugiyama *et al.*, 1970; Ricciardi, 1977; Buonocore *et al.*, 1987). However, the thresholds (63) are of no interest within the neuronal modelling context.

To give an idea of the difficulties one meets when trying to solve the FPT problem for the OU neuronal model for threshold other than (63), let us make use of the approach outlined earlier for the case of a constant threshold S. Equation (29) then reads:

$$\frac{\sigma^2}{2}\,\frac{d^2 g_\lambda}{dx_0^2} + \left(\mu - \frac{x_0}{\vartheta}\right)\frac{dg_\lambda}{dx_0} - \lambda\,g_\lambda = 0. \tag{64}$$

As shown in Capocelli and Ricciardi (1971), equation (64) can be solved with conditions (30) to yield:

$$g_\lambda(S|x_0) = \exp\Big[\frac{(x_0-\mu\,\vartheta)^2-(S-\mu\,\vartheta)^2}{2\,\sigma^2\,\vartheta}\Big]\;\frac{D_{-\lambda\,\vartheta}\Big[\sqrt{\dfrac{2}{\sigma^2\,\vartheta}}\,(\mu\,\vartheta-x_0)\Big]}{D_{-\lambda\,\vartheta}\Big[\sqrt{\dfrac{2}{\sigma^2\,\vartheta}}\,(\mu\,\vartheta-S)\Big]}\,, \tag{65}$$

where $D_\nu(z)$ is the Parabolic Cylinder Function, or Weber function:

$$D_\nu(z) = \frac{\exp(-z^2/4)}{\Gamma(z)}\int_0^\infty \exp\Big(\frac{-z-zx^2}{2}\Big)\,x^{-\nu-1}\,dx\,, \qquad \nu<0\,.$$

More explicitly, let us introduce the Kummer function

$$\Phi(a,c;x) = \sum_{n=1}^{\infty}\frac{a\,(a+1)\,\ldots\,(a+n-1)}{c\,(c+1)\,\ldots\,(c+n-1)}\,\frac{x^n}{n!}\,.$$

Then,

$$D_\nu(x) = \sqrt{\pi}\,2^{\nu/2}\,e^{-x^2/4}\Big[\frac{1}{\Gamma[(1-\nu)/2]}\,\Phi\Big(-\frac{\nu}{2},\frac{1}{2};\frac{x^2}{2}\Big)$$
$$-\frac{\sqrt{2}}{\Gamma(-\nu/2)}\,x\,\Phi\Big(\frac{1-\nu}{2},\frac{3}{2};\frac{x^2}{2}\Big)\Big]\,.$$

An alternative expression for $g_\lambda(S|x_0)$ can be found in Nobile *et al.*, 1985a. In any case, the function g_λ is too complicated to lead to any useful approximation for its inverse LT (which, of course, is not known in closed form). Hence, the firing time problem for the OU neuronal model must be investigated by different means.

Before proceeding, one little remark is necessary. In the foregoing we have focused our attention on sketching the genesis of the OU neuronal model for the time-homogeneous case. However, in some cases the necessity exists to include nonstationary effects in the neuronal models (Ricciardi, 1982) in order to be able to account for effects such as neuronal burst activity in which it is conceivable that the firing frequency is subject to some kind of modulation. This is a rather difficult task. A nonstationary version of the OU model can be obtained by slightly modifying the above procedure. Assume, for instance, that the arrival rates α_e and α_i of PSPs are made to be time-dependent, viz. $\alpha_e \to \alpha(t)$, $\alpha_i \to \beta(t)$. This means that two time nonhomogeneous Poisson processes are assumed to describe synaptic excitatory and inhibitory inputs. It is not difficult to see that now the infinitesimal moments (57) change into

$$\begin{cases} A_1(x) &= -\dfrac{x}{\vartheta}+\alpha(t)\,e+\beta(t)\,i \\ A_n(t) &= \alpha_i(t)\,e^n+\beta(t)\,i^n \qquad (n=2,3,\ldots) \end{cases}\,. \tag{66}$$

Hence, if one takes e and $|i|$ infinitely small and $\alpha(t)$ and $\beta(t)$ diverging in a suitable way, one may again end up with a diffusion process. For instance, take

$$\begin{cases} \alpha(t) \sim \dfrac{\gamma(t)}{\omega^2}, & e \sim h(t)\,\omega \\ \beta(t) \sim \dfrac{\eta(t)}{\omega^2}, & i \sim k(t)\,\omega \end{cases} \tag{67}$$

where γ, h, η and $-k$ are positive functions. Then, (66) read:

$$\begin{cases} A_1(x) = -\dfrac{x}{\vartheta} + \dfrac{1}{\omega}\left[\gamma(t)\,h(t) + \eta(t)\,k(t)\right] \\ A_n(t) = \omega^{n-2}\left[\gamma(t)\,h^n(t) + \eta(t)\,k^n(t)\right] \quad \text{(n=2,3,...)}. \end{cases} \tag{68}$$

If we impose that $\gamma(t)\ h(t) + \eta(t)\ k(t) = 0$, in the limit as $\omega \downarrow 0$ (viz. infinitely large rates of PSPs of infinitely small magnitude), (68) yield a time non homogeneous diffusion process, e.g. the diffusion process with the following drift and infinitesimal variance:

$$\begin{cases} A_1(x) = -\dfrac{x}{\vartheta} \\ A_2(t) = \gamma(t)\,h^2(t)\left[1 + \dfrac{\gamma(t)}{\eta(t)}\right]. \end{cases} \tag{69}$$

The above-mentioned modulation effect is then accounted for by suitably choosing the arbitrary functions γ, h and η. Various generalizations of this nonhomogeneous neuronal model are possible. They will be the object of future investigations.

We conclude this section by stressing the fundamental differences existing between the Wiener and the OU neuronal models. The Wiener model is unrealistically lacking the presence of the finite time-constant of the neuronal membrane, but it is susceptible of a closed form solution. Indeed, the firing pdf has the simple expression (33) (from which one can also calculate mean, variance, skewness, etc. of the firing time in the presence of a constant threshold). The OU model, although more realistically depicting the neurophysiological reality, does not allow us to obtain any closed form expression for the firing pdf. Moreover, as we shall see, although the moments of the firing time for the case of a constant threshold can be obtained analytically, their expressions are much too complicated to allow for any intuitive understanding in terms of the values assigned to the involved parameters. Hence, rather cumbersome computations are required to come to any evaluation of the statistics of the firing time.

Numerical evaluation of the firing pdf

We have stressed the lack of a closed form solution to the firing pdf for the OU neuronal model. We shall now briefly indicate an efficient procedure to obtain accurate numerical evaluations for the general case of time-varying thresholds and for arbitrary diffusion models (viz. not necessarily of the OU type).

We have pointed out that the calculation of the FPT pdf is essentially equivalent to the determination of a transition pdf in the presence of an absorbing boundary. This, in turn, is a very complicate task. Hence, efficient numerical algorithms are desirable, especially if one wishes to deal with the case of a time-varying neuronal threshold. It is fair to say that the first fundamental contribution to the solution of this problem is due to Durbin (1971) for the case of a standard Wiener process and a time-dependent boundary. A purely numerical method was successively developed by Anderssen *et al.* (1973). The idea is based on the remark that equation (23) also holds in the limit as $x \to S(t)$, i.e. when the constant boundary S is changed into a continuous function $S(t)$ and the arbitrary state x at time t indefinitely approaches the value that the threshold takes at that time. Hence, equation (23) becomes:

$$f[S(t), t|x_0, t_0] = \int_0^t f[S(t), t|S(\tau), \tau]\, g[S(\tau), \tau|x_0, t_0]\, d\tau, \tag{70}$$

with $x_0 < S(t_0)$. The determination of $g[S(t), t|x_0, t_0]$ for a standard Wiener process was accomplished by Anderssen *et al.* by a method based on the expansion of g in terms of Legendre polynomials and on the solution of the resulting system of linear equation. Such a method was then modified by Favella *et al.* (1982) and applied to the numerical solution of the FPT problem for the OU process. However, all these algorithms necessitate of the use of large computation facilities and of sophisticated library programs. As a consequence, they are expensive to run and not suitable to suggest the modeler in real time how to pick the various parameters to fit the recorded data or to decide whether to reject the postulated model. Some progress along this line was made by adopting an entirely different approach (Buonocore *et al.*, 1987). The guiding idea was to try to remove the singularity of the kernel of equation (70) by trying to modify the equation itself. While referring to the quoted paper for a complete treatment, here we limit ourselves to mentioning the essential results. Let

$$F(x, t|y, \tau) = P\{X(t) \leq x|X(\tau) = y\} \equiv \int_{-\infty}^{x} f(z, t|y, \tau)\, dz \tag{71}$$

be the transition distribution function of the membrane potential $X(t)$ and let $S(t)$ be a sufficiently smooth threshold function. Then, for $x_0 < S(t_0)$ the FPT pdf $g[S(t), t|x_0, t_0]$ of $X(t)$ through $S(t)$ satisfies the following second-kind Volterra integral equation:

$$g[S(t), t|x_0, t_0] = -2\,\psi[S(t), t|x_0, t_0] +$$

$$2\int_{t_0}^{t} g[S(\tau), \tau|x_0, t_0]\, \psi[S(t), t|S(\tau), \tau]\, d\tau, \tag{72}$$

where we have set

$$\left\{ \begin{array}{rcl} \psi[S(t),t|y,\tau] & = & \varphi[S(t),t|y,\tau] + k(t)\ f[S(t),t|y,\tau] \\ \varphi(x,t|y,\tau) & = & \dfrac{d}{dt}\ F(x,t|y,\tau), \end{array} \right. \tag{73}$$

with $k(t)$ an arbitrary continuous function. Apparently, equation (72) is more complicate than equation (70). However, it possesses an extra degree of freedom, viz. the arbitrary function $k(t)$ whose specification can be made in a way to remove the singularity of the kernel or, when possible, to make the integral on the right-hand-side of (72) vanish, thus obtaining the closed form solution $g[S(t),t|x_0,t_0] = -2\ \psi[S(t),t|x_0,t_0]$. To illustrate this method, let us separately consider the Wiener and the OU neuronal models.

We start with the Wiener model. One has (cf. Section on *Neuronal modeling*):

$$f(x,t|y,\tau) = \frac{1}{\sqrt{2\ \pi\sigma^2\ (t-\tau)}}\ \exp\left\{-\frac{[x-y-\mu\ (t-\tau)]^2}{2\ \sigma^2\ (t-\tau)}\right\}. \tag{74}$$

Hence,

$$F(x,t|y,\tau) = \frac{1}{2}\left[1 + Erf\left\{\frac{[x-y-\mu\ (t-\tau)]}{\sqrt{2\ \sigma^2\ (t-\tau)}}\right\}\right]. \tag{75}$$

From (73) and (75) we then have:

$$\psi[S(t),t|y,\tau] = f[S(t),t|y,\tau]\ h(t,\tau,y)\ , \tag{76}$$

where the function $h(t,\tau,y)$ is defined as follows:

$$h(t,\tau,y) = S'(t) - \frac{\mu}{2} - \frac{S(t)-y}{2\ (t-\tau)} + k(t)\ , \tag{77}$$

with $S'(t) = dS(t)/dt$. The following result holds. Let $S(t)$ be twice differentiable with continuous derivatives. Then,

$$\lim_{\tau\uparrow t} \psi[S(t),t|S(\tau),\tau] = 0 \tag{78}$$

if and only if

$$k(t) = \frac{1}{2}\ [\mu - S'(t)]\ . \tag{79}$$

Furthermore, one has:

$$\psi[S(t),t|S(\tau),\tau] = 0 \quad \forall t,\tau\ :\ t_0 \le \tau < t \quad and \quad \lim_{\tau\uparrow t} \psi[S(t),t|S(\tau),\tau] = 0 \tag{80}$$

if and only if

$$\left\{ \begin{array}{rcl} S(t) & = & a\ t + b \\ k(t) & = & \dfrac{\mu - a}{2} \end{array} \right. \tag{81}$$

with a and b arbitrary real numbers. In other words, if a smooth threshold $S(t)$ is assigned, the firing pdf for the Wiener neuronal model satisfies equation (72) in which the kernel $\psi[S(t), t|S(\tau)]$ is made continuous (actually vanishing) at $\tau = t$ with the choice (79) of $k(t)$ in (76). Hence, a straightforward numerical solution of (72) is possible without need to resort to main frame computers or program libraries (cf. Buonocore *et al.*, 1987). We incidentally note that by virtue of (80) and (81) one immediately recovers the well-known result for the FPT pdf for a linearly time-dependent boundary:

$$g(a\,t + b, t|x_0, t_0) = \frac{|a\,t_0 + b - x_0|}{\sqrt{2\,\pi\sigma^2\,(t - t_0)^3}}\,\exp\left\{-\frac{[a\,t + b - x_0 - \mu\,(t - t_0)]^2}{2\,\sigma^2\,(t - t_0)}\right\}$$

$$(x_0 \neq a\,t_0 + b) \tag{82}$$

of which (33) is a particular case.

Similar considerations can be made for the case of a diffusion process with linear drift and constant infinitesimal variance:

$$\begin{cases} A_1(x) &= \alpha\,x + \beta \\ A_2(x) &= \sigma^2, \end{cases} \tag{83}$$

α, β and $\sigma \neq 0$ being arbitrary real constants. Then, under the initial condition $P\{X(t_0) = x_0\} = 1$, for all $y \in \Re$ and $\tau < t$ one has:

$$f(x, t|y, \tau) = \left\{\frac{\exp\,[-2\,\alpha\,(t - \tau)]\,\alpha}{\pi\,\sigma^2\,[1 - \exp\,(-2\,\alpha\,(t - \tau))]}\right\}^{1/2}$$

$$\times \exp\left\{\frac{-\alpha\,[(x + \beta/\alpha)\,\exp\,[-\alpha\,(t - \tau)] - (y + \beta/\alpha)]^2}{\sigma^2\,[1 - \exp\,(-2\,\alpha\,(t - \tau))]}\right\} \tag{84}$$

and

$$f(x, t|y, \tau) = \frac{1}{2}\,\left\{1 + Erf\left[\left\{\frac{\alpha}{\sigma^2\,[1 - \exp\,(-2\,\alpha\,(t - \tau))]}\right\}^{1/2}\right.\right.$$

$$\left.\left.\times\left[(x + \beta/\alpha)\,\exp\,[-\alpha\,(t - \tau)] - (y + \beta/\alpha)\right]\right]\right\}. \tag{85}$$

From (73) and (85) we thus have:

$$\psi[S(t), t|y, \tau] = f[S(t), t|y, \tau]\,H(t, \tau, y)\,, \tag{86}$$

with

$$H(t, \tau, y) = S'(t) - \alpha\left[S(t) + \frac{\beta}{\alpha}\right] - \frac{\alpha\,\exp\,[-\alpha\,(t - \tau)]}{1 - \exp\,[-2\,\alpha\,(t - \tau)]}$$

$$= \left\{\left[S(t) + \frac{\beta}{\alpha}\right]\,\exp\,[-\alpha\,(t - \tau)] - \left(y + \frac{\beta}{\alpha}\right)\right\} + k(t). \tag{87}$$

Let now again $S(t)$ be a twice differentiable function with continuous derivatives. Then (Buonocore *et al.*, 1987),

$$\lim_{\tau \uparrow t} \psi[S(t), t | S(\tau), \tau] = 0 \tag{88}$$

if and only if

$$k(t) = \frac{1}{2} \left[\alpha\, S(t) + \beta - S'(t)\right]. \tag{89}$$

Furthermore:

$$\psi[S(t), t | S(\tau), \tau] = 0 \quad \forall t, \tau \;:\; t_0 \leq \tau < t \quad \textit{and} \quad \lim_{\tau \uparrow t} \psi[S(t), t | S(\tau), \tau] = 0 \tag{90}$$

if and only if

$$\begin{cases} S(t) &= -\dfrac{\beta}{\alpha} + A\, e^{\alpha\, t} + B\, e^{-\alpha\, t} \\ k(t) &= B\, \alpha\, e^{-\alpha\, t}, \end{cases} \tag{91}$$

with A and B arbitrary constants. Therefore, also for the OU neuronal model the firing pdf is a solution of equation (72) in which the kernel is continuous at $\tau = t$ if the choice (89) of $k(t)$ is made. Note that under assumptions (91) we obtain a closed form FPT pdf for the diffusion process with moments (83):

$$g(-\beta/\alpha + A\, e^{\alpha\, t} + B\, e^{-\alpha\, t}, t \,|\, x_0, t_0)$$

$$= 2\, \alpha\, \frac{\left|A\, e^{\alpha\, t_0} + B\, e^{-\alpha\, t_0} - x_0 - \beta/\alpha\right|}{\exp\left[\alpha\, (t - t_0)\right] - \exp\left[-\alpha\, (t - t_0)\right]} \times f[S(t), t | x_0, t_0]$$

$$(x_0 \neq -\beta/\alpha + A\, e^{\alpha\, t_0} + B\, e^{-\alpha\, t_0}). \tag{92}$$

with f given by (84) for $x = S(t)$, $y = x_0$ and $\tau = t_0$ (cf. equation (63)). However, this is not useful for the understanding of the behaviour of the OU neuronal model. Instead, the numerical solution of equation (70) under the continuity conditions (78) and (88), and hence with the choices (79) and (89), respectively, of the originally arbitrary function $k(t)$, becomes a rather straightforward and simple task (Buonocore *et al.*, 1987). Therefore, by using a personal computer and without any need of library programs, the Wiener and the OU neuronal models can be investigated for arbitrary time-varying thresholds and hence parameters identification problems can be effectively approached.

One might wander whether the regularization procedure of equation (70) is limited to the above two neuronal models. Actually, a more general result can be proved. Indeed, under mild assumptions one can prove that for any time-homogeneous diffusion processes characterized by drift $A_1(x)$

and infinitesimal variance $A_2(x)$, the kernel of equation (70) can be made vanishing at $\tau = t$ by selecting (Buonocore *et al.*, 1989)

$$k(t) = \frac{1}{2}\left\{A_1[S(t)] - S'(t) - \frac{A_2'[S(t)]}{4}\right\}, \tag{93}$$

a condition that includes (79) and (89) as particular cases.

The moments of the firing time

In this section we shall refer to the OU neuronal model and, for the sake of simplicity, we shall set $\mu = 0$ in (42*a*).

As we have indicated, the firing pdf for a wide class of time-dependent thresholds can be numerically evaluated to a high degree of precision. Instead, analytical results of neurophysiological interest are scarce, if not completely lacking. Indeed, even in the case of the constant threshold approximation one can determine only the LT of the firing pdf, which is given by (65). Nevertheless, use of this can be made to calculate the moments $t_n(S|x_0)$ of the firing time via (16). For instance, the first three moments have been obtained in such a way by Sato (1978) and by Cerbone *et al.* (1981). A much more compact form can however be derived by an alternative method (Nobile *et al.*, 1985a). By rather cumbersome calculations, the following result can be proved:

$$t_1(S|x_0) = \vartheta\left[\varphi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) - \varphi_1\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right] + \psi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) - \psi_1\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right) \tag{94a}$$

$$\begin{aligned} t_2(S|x_0) = {} & 2\,\vartheta\, t_1(S|x_0)\left[\sqrt{\pi}\,\varphi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) + \psi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right)\right] \\ & + 2\,\vartheta^2\left\{\sqrt{\pi}\,\ln 2\left[\varphi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) - \varphi_1\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right]\right. \\ & \left. - \sqrt{\pi}\left[\varphi_2\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) - \varphi_2\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right] - \psi_2\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) + \psi_2\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right\} \end{aligned} \tag{94b}$$

$$\begin{aligned} t_3(S|x_0) = {} & 3\,\vartheta\, t_2(S|x_0)\left[\sqrt{\pi}\,\varphi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) + \psi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right)\right] \\ & + 6\,\vartheta^2\, t_1(S|x_0)\left[\sqrt{\pi}\,\ln 2\,\varphi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) - \sqrt{\pi}\,\varphi_2\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) - \psi_2\left(\frac{S}{\sigma\sqrt{\vartheta}}\right)\right] \\ & + 3\,\vartheta^3\sqrt{\pi}\left(\ln^2 2 + \frac{\pi^2}{12}\right)\left[\varphi_1\left(\frac{S}{\sigma\sqrt{\vartheta}}\right) - \varphi_1\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right] - 6\,\vartheta^3 \end{aligned}$$

$$\sqrt{\pi}\ \ln 2\left[\varphi_2\left(\frac{S}{\sigma\sqrt{\vartheta}}\right)-\varphi_2\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right]+6\,\vartheta^3\,\sqrt{\pi}\left[\varphi_3\left(\frac{S}{\sigma\sqrt{\vartheta}}\right)\right.$$

$$\left.\varphi_3\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right]+6\,\vartheta^3\left[\psi_3\left(\frac{S}{\sigma\sqrt{\vartheta}}\right)-\psi_3\left(\frac{x_0}{\sigma\sqrt{\vartheta}}\right)\right], \tag{94c}$$

where we have set:

$$\varphi_1(z)=\int_0^z e^{t^2}\,dt\varphi_2(z)=$$

$$\sum_{n=0}^{\infty}\frac{z^{2n+3}}{(n+1)!\,(2n+3)}\sum_{k=0}^{n}\frac{1}{2k+1}\varphi_3(z)=$$

$$\sum_{n=0}^{\infty}\frac{z^{2n+5}}{(n+2)!\,(2n+5)}\sum_{k=0}^{n}\frac{1}{2k+3}\sum_{j=0}^{k}\frac{1}{2j+1}$$

$$\psi_1(z)=\sum_{n=0}^{\infty}\frac{2^n}{(n+1)\,(2n+1)!!}\,z^{2n+2}$$

$$\psi_2(z)=\sum_{n=0}^{\infty}\frac{2^n z^{2n+4}}{(2n+3)!!\,(n+2)}\sum_{k=0}^{n}\frac{1}{k+1}$$

$$\psi_3(z)=\sum_{n=0}^{\infty}\frac{2^n z^{2n+6}}{(2n+5)!!\,(n+3)}\sum_{k=0}^{n}\frac{1}{k+2}\sum_{j=0}^{k}\frac{1}{j+1}. \tag{95}$$

Anyone who has a little familiarity with computational problems would undoubtedly agree that a satisfactory evaluation of quantities (1) require a good deal of work. Equally undoubtedly, the closed-form expressions (1), viz. the analytic expressions of the first three moments of the firing time for the considered OU neuronal model, do not provide the modeler with any more intuition than the complete lack of knowledge on these moments. Differently stated, this is a typical situation where one is obliged to explore a virgin territory groping in the dark of fragmentary computational results. Consequently, we tried to compute systematically the mean t_1, variance $t_2-t_1^2$ and skewness $t_1^3/(t_2-t_1^2)^{3/2}$ of the firing time via (1) and (2) for a variety of thresholds and initial states. As shown in Nobile *et al.* (1985a), a

striking feature emerged, that apparently had passed unnoticed. Namely, for positive boundaries of the order of a couple of units or more, whatever the initial state the variance of the firing time equals the square of its mean value to an excellent degree of approximation. Moreover, the skewness is equal to 2 to a very high precision. Finally, the goodness of these approximations was seen to increase with increasing values of the threshold. Putting all this information together, the conjecture emerged that the firing pdf is susceptible of an excellent exponential approximation for a wide range of thresholds and of initial states. As a matter of fact, these " experimental" results have led us to formulate theorems on the asymptotic exponential behaviour of the FPT pdf not only for the OU process but also for a wider class of diffusion processes, both for constant and for time-varying boundaries (Nobile *et al.*, 1985b; Giorno *et al.*, 1990).

As is to be expected, the convergence to the exponential distribution for increasing threshold is accompanied by a large increase of the mean firing time. Indeed, as the threshold moves away from the reset voltage, spikes become more and more rear and hence the mean interval between successive neuronal firings grows larger and larger. This kind of behaviour is in agreement with the finding that for some neurones the histograms of the recorded interspike intervals become increasingly better fitted by exponential functions as the firing rates decrease (see, for instance, Škvařil *et al.*, 1971).

Some extensions of the OU model

It must be pointed out that the OU neuronal models considered in the foregoing, characterised by linear drift and constant infinitesimal variance or by linear drift and time-varying infinitesimal variance (cf. Eqs. (42) and (69)) can be modified in a way to include a time-dependent drift. For instance, it is conceivable that the model (42) be changed into

$$\begin{cases} A_1(x) &= -\dfrac{x}{\vartheta} + \mu + P(t) \\ A_2(x) &= \sigma^2 \end{cases} \tag{96}$$

if a time-dependent extra-effect is induced by some kind of DC external stimulation acting on the neurone. It is, in particular, interesting to investigate the behaviour of the model (96) when $P(t)$ is a periodic function of period T that reflects some oscillatory action of the environment. This situation quite naturally leads us to considering an FPT problem for the OU process (42) and a periodic boundary. To give a hint of why this is so, let us refer to the OU neuronal model with infinitesimal moments (96), where $P(t)$ is periodic with period T. It is then intuitive that by a suitable transformation an FPT problem through a constant threshold S_0 can be changed into an FPT problem for a time-dependent OU model through a periodic

boundary. Putting it the other way around, an FTP problem through a periodic boundary $S(t)$ for the OU process defined via

$$\begin{cases} dX(t) &= -\frac{1}{\vartheta} X(t)\, dt + \sigma\, dW(t) \\ P\{X(0) &= x_0\} = 1, \end{cases} \tag{97}$$

is equivalent to the FPT problem through the constant boundary $S(0) = S_0$ for the process

$$Y(t) = X(t) - S(t) + S_0 \tag{98}$$

that satisfies

$$\begin{cases} dY(t) &= \left[-\frac{1}{\vartheta} Y(t)\, dt - \frac{S(t)}{\vartheta} + \frac{S_0}{\vartheta} - S'(t)\right] dt + \sigma\, dW(t) \\ P\{Y(0) &= x_0\} = 1. \end{cases} \tag{99}$$

In other words, both $X(t)$ and $Y(T)$ start at x_0 at time 0. However, $Y(t)$ is an OU process containing a periodic drift and the threshold is constant while $X(t)$ is a time-independent linear-drift process but the corresponding threshold is a periodic function. The question then naturally arises: How does the firing pdf look for an OU model and a periodic boundary? This is another paradigmatic case in which a mathematical answer is made possible by a systematic analysis of computational results. We have been able to prove that the periodicity of the threshold reflects in an oscillatory behaviour of the firing pdf which, in turn, approaches a time nonhomogeneous exponential limit as the threshold moves farther and farther away from the reset value of the process. For instance, for the normalised OU process, viz. for $A_1 = -x$ and $A_2 = 2$, as the periodic threshold $S(t)$ of period T moves away from the equilibrium point (the origin, in this case) in the positive direction, the following asymptotic relation holds for the FPT pdf $g(t)$:

$$g(t) \sim Z(t)\, e^{-\lambda t}, \tag{100}$$

where λ is a positive constant and $Z(t)$ is a periodic function of period T. It is quite impressive to see to what degree of accuracy relation (100) emerges (cf. Giorno *et al.*, 1990).

The OU neuronal model has been derived either by means of the SDE (43) or via a limiting procedure on the infinitesimal moments (57). In either case, we can view it as obtained from a deterministic system characterised by a spontaneous exponential decay, towards some resting value, on which suitable random fluctuations have been superimposed. To illustrate this point, let us refer again to the OU model and, for simplicity, let us take $\mu = 0$. Then, from (61) and (62) we see that this is a normal process whose transition pdf we now re-write as

$$f(x, t|x_0) = \frac{1}{\sqrt{2\,\pi\, U(t)}} \exp\left\{ -\frac{1}{2\, U(t)} \left[x - \frac{\gamma(t)}{\gamma(0)}\, x_0 \right] \right\}, \tag{101}$$

with

$$U(t) = \gamma(0)\left[1 - \frac{\gamma^2(t)}{\gamma^2(0)}\right] \tag{102}$$

and

$$\gamma(t) = e^{-t/\vartheta}\,. \tag{103}$$

This way of writing is instructive. Indeed, one can prove that the conditional pdf of a stationary normal process is given by the right-hand-side of (101) in which

$$\gamma(t) = E\big[\gamma(t+\tau)\,\gamma(t)\big] \tag{104}$$

is the correlation function of $X(t)$ and $U(t)$ is given by (102). Therefore, one way of constructing other neuronal models somewhat departing from the OU model is to take stationary normal processes with zero mean and correlation function of the form

$$\gamma(t) = \gamma(0)\,e^{-t/\vartheta} + \varphi(t)\,o[\gamma(0)]\,. \tag{105}$$

Any such process can be thought of as an extension of the OU model in the sense that for small $\gamma(0)$, alike the OU model, it describes some kind of random fluctuations around the spontaneous exponential decay of the membrane potential. Indeed, from (101), (102) and (105) we have:

$$\lim_{\gamma(0)\downarrow 0} P\Big\{X(t) = x_0\,e^{-t/\vartheta}\Big\} = 1\,. \tag{106}$$

In other words, as the variance $\gamma(0)$ of $X(t)$ goes to zero, the conditional pdf (101) of $X(t)$ tends to a delta-function centred around the mean value of the OU process, viz., $x_0\,\exp(-t/\vartheta)$.

The reason for considering a normal process with correlation function of type (105) is two-fold. On the one hand, we have an approximation of the OU neuronal model in the sense just specified; on the other hand, starting from function (105) we can construct an approximation to it that admits of derivatives at $t = 0$. As is well known, such a circumstance implies that the corresponding process $X(t)$ possesses differentiable sample paths, unlikely the OU process whose sample paths are almost nowhere differentiable. The differentiability property of the sample paths play an essential role for the description of the process, which is not of the Markov type. However, we cannot go here into any details. Instead, we stress that, in conclusion, by the above-outlined procedure the OU model can be replaced by a stationary normal process with differentiable sample paths, provided the exponential correlation function is substituted by a function that, while possessing derivatives in the origin, does not depart too much from an exponential.

Here we cannot even sketch the rather subtle and very cumbersome methods that have been devised to deal with FPT problems for normal processes with differentiable sample paths. The interested reader is referred

to Ricciardi and Sato (1983, 1986) for an analytic approach and to Ricciardi *et al.* (1991, 1992) for a description of a simulation procedure and for a discussion of a large amount of computational results.

Neuronal diffusion with state-dependent infinitesimal variance

In order to embody some other physiological features of real neurones, several alternative models have been proposed. However, the properties of such models are generally obscure, because the corresponding FPT problems still remain unsolved. The papers by Capocelli and Ricciardi (1973), Hanson and Tuckwell (1983) and Giorno *et al.* (1988) on the FPT problem for the considered neuronal models are still the only exceptions. Hereafter, following Lánský *et al.* (1990), we shall briefly mention some of the diffusion models with state-dependent infinitesimal variance.

Capocelli and Ricciardi (1973) proposed a model in which the effect of the noise, that influences the membrane potential, decreases as the membrane potential gets closer to the threshold. Their model is defined on $(-\infty, S)$ by

$$\begin{cases} A_1(x) &= -\alpha\, x + \beta \\ A_2(x) &= \sigma\sqrt{S - x} \end{cases} \tag{107}$$

where α, β and $\sigma > 0$ are constants. A thorough analytical treatment of this model is provided in this paper, including a closed-form expression for the firing pdf. However, the role of this model seems to be limited by the biological assumptions leading to the form of the infinitesimal variance in (107). Successively, Matsuyama (1976) examined shunting inhibition in neural models by studying the stochastic diffusion model defined on (ν, S) by the following infinitesimal moments:

$$\begin{cases} \mu(x) &= -a\, x + \mu \\ \sigma(x) &= \sqrt{\sigma_1^2 + \sigma_2^2\,(S - x)^2} \end{cases} \tag{108}$$

where $a, \mu, \sigma_1 > 0$ and $\sigma_2 > 0$ are constants and where $\nu < x_0$ is a reflecting boundary. For this model numerical methods for the solution of the FPT problem via Fokker–Planck equation where used, and a comparison of the firing times of model (108) and of the Ornstein–Uhlenbeck model was performed.

The remaining models reviewed in this section are known as the models with "reversal potentials". In this class of models it is assumed that the input-related changes of the membrane potentials do depend on the actual value of the membrane potential. This is suggested by some experimental findings (cf. for instance, Schmidt, 1985). It is enlightening to quote the following sentences by Schmidt: "It can be seen that an IPSP (inhibitory PSP) triggered at a membrane potential of −65 mV (the resting potential of this cell) has an amplitude of 5 mV in the hyperpolarizing direction. If the

membrane potential is lowered by a current applied through a microelectrode, the amplitude of IPSP (with the same peripheral stimulus) increases greatly. This means that the driving force for the IPSP becomes greater during depolarization of the membrane potential. Conversely, when the membrane potential is increased, the amplitude of the IPSPs initially becomes smaller, and finally drops to, zero at approximately -80 mV". A similar behaviour should hold also for EPSPs (excitatory PSPs). However, the value at which EPSPs would vanish is above the firing threshold S, and thus it cannot be measured. Hence, the reason for the introduction of models embodying the concept of reversal potentials was to account for the biological restriction that the state space of the process $X(t)$ must be finite and for the assumption that a nonlinear summation of synaptic input takes place. Although the need for such models was first pointed out by Johannesma (1968), their analysis has appeared much later.

Hanson and Tuckwell (1983) considered two reversal potential diffusion models; the first is given by

$$\begin{cases} A_1(x) &= -\dfrac{x}{\tau} + \mu\,(V_E - x) \\ A_2(x) &= \sigma\,(V_E - x) \end{cases} \tag{109}$$

where $\tau > 0$, μ and $\sigma > 0$ are constant and where V_E is a constant, called the " excitatory reversal potential" for which $V_E > S$ holds. The model (109) is studied in the cited paper either without any additional boundary or being restricted from below by means of an artificial reflecting boundary condition at the resting potential. The second model proposed and analysed by these authors was defined by the infinitesimal moments

$$\begin{cases} A_1(x) &= -\dfrac{x}{\tau} + \mu_1\,(V_E - x) + \mu_2\,(x - V_I) \\ A_2(x) &= \left[\sigma_1^2\,(V_E - x)^2 + \sigma_2^2\,(x - V_I)^2\right]^{1/2}, \end{cases} \tag{110}$$

where τ, μ_1, μ_2, σ_1 and σ_2 are constants and where $V_I < x_0$ is an additional constant called the " inhibitory reversal potential". The role of V_E is the same as in model (109). As before, the lower boundary was taken to be either $-\infty$ or a constant. Here V_I was a reflecting boundary. For both the models (109) and (110) the first two moments of the free process $X(t)$ (viz., in the absence of the threshold S) and the stationary pdf were derived. The first two moments of the FPT through S were then computed.

The diffusion processes (109) and (110) are characterised by a linear drift $A_1(x)$, like as in the case of the OU model. However, the form of the infinitesimal variance $A_2(x)$ does not permit a simple interpretation of the model in the sense that there is no value of the membrane potential for which the infinitesimal variance is zero. Therefore an artificial additional boundary condition has to be imposed. This issue has motivated the development of additional diffusion neuronal models. Giorno *et al.* (1988) thor-

oughly analysed the diffusion neuronal model defined by the following infinitesimal moments:

$$\begin{cases} A_1(x) &= -a\,x + \mu \\ A_2(x) &= \sigma^2\,(x - V_I) \end{cases} \tag{111}$$

where $\sigma > 0$, $\mu > 0$ and a are constants. Now the boundary V_I is either regular, for $\mu < \sigma^2$, or entrance for $x \geq \sigma^2$. When the process is restricted from below at V_I, the transition pdf of $X(t)$, its limiting form and the first two moments have been determined. For the FPT problem, the first three moments have been obtained in an analytical form. The computational method mentioned above was then used for the numerical evaluation of the FPT pdf. Moreover, the probability distribution of the firing frequency, i.e. the probability distribution and mean and variance of having n firings in the time interval $(0, t)$ under the condition that there is a spike at zero, was also calculated. The exponential behaviour of the pdf for large values of the threshold in this model follows as a special case of the result by Nobile *et al.* (1985b). This implies that the frequency description of the model behaves in a Poisson-like manner. Finally the joint distribution was derived for the number of firings together with the probability density for the value of the membrane potential. An interesting feature of this model is that it can be constructed by starting with a discrete model of the birth and death type and then passing to the limit in much the same way as we constructed the OU diffusion model. The advantage of this procedure is that the hypotheses underlying the final diffusion model are pinpointed and the role of the inhibitory potential clearly emerges without need to introduce artificial boundary conditions.

Two other diffusion models with reversal potential are due to Lánský and Lánská (1987). They are characterised by the following infinitesimal moments:

$$\begin{cases} A_1(x) &= -a\,x + \mu \\ A_2(x) &= \sigma\,(x - V_I) \end{cases} \tag{112}$$

and

$$\begin{cases} A_1(x) &= -a\,x + \mu \\ A_2(x) &= \sigma\,(V_E - x)\,(x - V_I) \end{cases} \tag{113}$$

respectively, where the constants and the reversal potentials are defined and denoted as in the previous cases. The properties of the diffusion processes with infinitesimal moments (112) and (113) have not been studied previously in the neuronal modelling context, but they have been discovered in other applications of stochastic diffusion processes. For example, in the case of (113) by transforming the interval (V_I, V_E) into the interval $(0, 1)$ we obtain a process that is well known in genetic applications (see, e.g. Karlin and Taylor, 1981). However, differently from the genetic applications, now the transformed threshold S lies somewhere inside the interval $(0, 1)$ and not at its boundaries.

As stressed by Lánský *et al.* (1990), the variety of forms for the infinitesimal variance and the linear form of the drift of the above sketched models are to be expected. In particular, the linear drift models the passive electrical circuit properties of the membrane at the trigger zone and also the mean effect of the noisy input. The infinitesimal variance, on the other hand, must not only take into account the diversity of spatial configurations for different neurones, but the location and type of synaptic input on that neurone as well. Hence, a variety of forms for this term in the diffusion equation appear to be appropriate.

Concluding remarks

An important topic, not extensively studied as yet, consists of model verification. For diffusion models some results are available (Pernier, 1972; Nilsson, 1977; Tuckwell and Richter, 1978; Lánský, 1983; Lánský et al., 1988). Noteworthy are also Segundo's (1986) neurophysiologically oriented critical review of the integrate-and-fire models and the thorough comparison between experimental and simulation results in the stochastic after-hyperpolarization model of Smith and Goldberg (1986).

It must be stressed that the neuronal behaviour described by diffusion models ultimately assume that the output is a renewal process, viz. that intervals between firings are independent and identically distributed. However, Frigessi and Den Hollander (1989) introduced a model aimed at simulating the clustering effect in spike generation. This is based on the Ornstein–Uhlenbeck process with a threshold of two different values and the dynamics of the resultant output can be expressed in the terms of a semi-Markov process. Serial dependence among interspike intervals has also been modelled in other ways, e.g. by adjusting the reset value after each spike in the Ornstein Uhlenbeck process (Geisler and Goldberg, 1966).

Stochastic neuronal models are also used for the description of more complex neuronal structures. For instance, Holden (1978) used the Wiener and Ornstein-Uhlenbeck processes for the study of information transfer in a chain of neurones. Furthermore, the papers by Brannan and Boyce (1981), Brannan (1982), Kryukov (1984), Buhman and Schulten (1987) are among the few examples where stochastic diffusion models are used to study the dynamics of neural nets made of neuronal units which are modelled by an Ornstein-Uhlenbeck process with a constant or time-varying threshold. A critical review of deterministic and stochastic approaches to the analysis of neural networks can be found in Cowan and Sharp (1988). A review of several specific neuronal systems and networks, together with simulation programs and an extensive bibliography can be found in MacGregor (1987).

In the foregoing we have striven to outline a variety of stochastic models for single neurones' activity based on the theory of diffusion processes. As it has been shown, the prediction of the firing pdf is a difficult problem

whose solution can usually only be approached by numerical or simulation procedures. To conclude this birds' eye view of this topic, I would like to mention that an alternative approach to the construction of diffusion models that are able to fit experimental data consists of reversing the problem. Namely, instead of formulating a neuronal model $X(t)$ based on some reasonable assumptions and then trying to compute the firing pdf as a FPT pdf through a preassigned threshold, one can proceed as follows. Firstly, construct the histogram of the experimentally recorded spike train and try to fit it by a function $g(S, t|x_0)$, with S and x_0 parameters to be determined by the standard methods. Once this task has been accomplished, ask the following question: Can $g(S, t|x_0)$ be viewed as the FPT pdf, through the threshold S, of a diffusion process conditioned on $X(0) = x_0$ and, if answer is yes, can such a process be uniquely determined? This is, so to say, the "the inverse" of the FPT problem. Quite surprisingly, at least in principle a precise answer to this question can be provided (Capocelli and Ricciardi, 1972).

Acknowledgements

I wish to mention that on April 8, 1992 my colleague and dear friend Professor Renato M. Capocelli passed away. Some of the work that I reviewed in this chapter has been the outgrowth of a happy and long-lasting collaboration between us that started at the Institute of Cybernetics of the Italian National Research Council in the late sixties and that ceaselessly continued for many years in various research areas at the University of Chicago and at the University of Salerno. I dedicate this chapter to his memory.

8

Noise and Chaos in Neural Systems

PETER ÉRDI

Central Research Institute for Physics, Group of Biophysics, PO Box 49, H-1525 Budapest 114, Hungary

Neural structures: between determinism and randomness

Introductory remarks

"Whenever he is looking at any piece of neural tissue, the investigator becomes immediately confronted with the choice between two conflicting issues: the question of how intricate wiring of the neuropil is strictly predetermined by some genetically prescribed blueprint and how much freedom is left to chance within some framework of statistical probabilities or some secondary mechanism of trial and error, or selecting connections according to necessities or the individual history of the animal. Even on brief reflection one has to arrive at the conclusion that the case may not rest on either extreme" (Szentágothai, 1978, 1990).

The nervous system is considered as a prototype of hierarchical structures and can be investigated at molecular, cellular, synaptic, network and systemic level. Both elementary and more complex neural structures are dynamic structures in that sense that they are maintained by the material, energetic and informational interaction with the environment. Dynamic structures in general and regular spatial patterns in biological systems, in more particular, are the results of symmetry-breaking morphogenetic mechanisms. Two questions will be studied in this section. First, how to characterize the dendritic arborization patterns and what could be the mechanism of the generation of dendritic branching trees? Second, what can be the role of environmental noise in the ontogenetic formation of ordered cortical structures?

Dendritic patterns

Though dendrites were originally defined as processes specialized to be the receptive part of the neuron, it turned out that sometimes they can be presynaptic elements. The *dendritic arborization* can be classified based on (1) "randomness" versus "regularity" in regard to branching of the various dendritic segments, resulting in a continuous spectrum ranging from the "radiate" to the "tufted" types of arborization and (2) the degree of deviation of the individual branches from a radiated and rectilinear course (Szentágothai and Arbib, 1974). The main types of patterns were illustrated by Ramon-Moliner and Nauta, 1966.

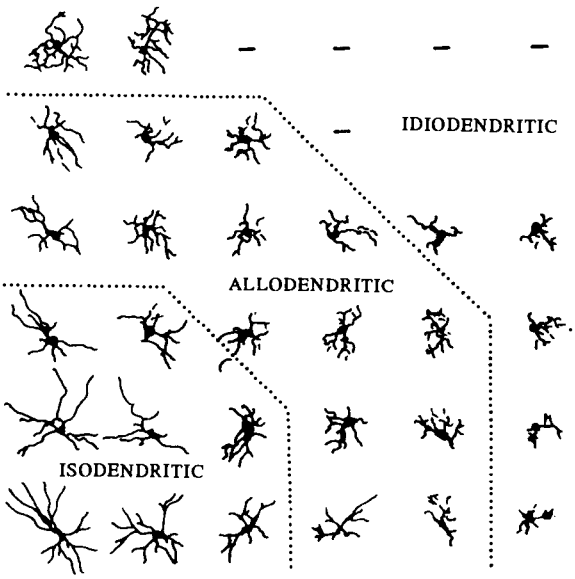


Fig. 1. Varieties of dendritic patterns found in the brainstem of the cat.

What kinds of *developmental mechanisms* might generate the complex branching patterns of neurons? Schierwagen (1987) and later Pellionisz (1991) suggested that neurons might have approximately fractal structure. Schierwagen even analyzed the most important possible developmental models, leading to *fractals*, as well. A main problem of developmental biology is whether elementary constituents in different positions become programmed to follow different developmental pathways and thus to produce the characteristic anatomical pattern of the organism? Basically now there is a consensus that development is not genetically programmed, but only controlled. A phenomenological approach based on the concept of *positional information* was suggested by Wolpert (1969), emphasizing the role of *morphogenetic gradients*. Therefore reaction — diffusion systems seemed

to be a proper general framework to describe biological pattern formation (see e.g. Slack, 1987; Tsonis, 1987).

A rather general class of dynamic models are:

$$\frac{\partial c_i}{\partial t}(x,t) = f_i(c(x,t)) + D_i \tag{1}$$

where $c_i(x,t)$ is the concentration of the ith component, the function f prescribes the kinetics due to the chemical reactions and D_i is the redistribution operator due to diffusion, convection etc.

A *reaction–diffusion model* of pattern formation was first proposed by Turing (1952) and has extensively been elaborated by Gierer-Meinhardt (1972), (see also Meinhardt, 1982; Harrison, 1987; Murray, 1989 — Chapter 14). The Turing–Gierer–Meinhardt (TGM) model is a specific example of equation (1) to describe the (a) activator and (b) inhibitor interactions:

$$\begin{aligned} \frac{\partial a}{\partial t}(x,t) &= f(a(x,t), b(x,t)) + D_a(x,t) \\ \frac{\partial b}{\partial t}(x,t) &= g(a(x,t), b(x,t)) + D_b(x,t) \end{aligned} \tag{2}$$

A three-variables extension of the TGM model has also been given to describe network formation based on the diffusible morphogen concept. This model led to uniform behavior.

It was argued by Schierwagen (1987) that the TGM model explains the spatial distribution of the morphogens only and not the shape of the anatomical structure. Pattern formation of neurons can be better described by other models. Furthermore, it is not clear, under what conditions (if at all) the TGM model might lead to fractals (Meakin, 1986). Both purely stochastic and so-called random growth models leading to fractals were suggested to describe the generation of dendritic branching.

The branching patterns of dendrites were modeled according to the Galton–Watson process in varying environments (Kliemann, 1987). Using this stochastic approach, the geometric structure of the neuron is characterized by statistical means.

Accepting the view that neural dendritic branching might have a *self-similar* character, its development can be modeled by fractal generating growth models (Stanley and Ostrowsky, 1986). The main important property of these models is that a growth process having random character may lead to structure with inherent symmetry properties. Specifically, the diffusion limited growth model is proper for generating fractal dendritic patterns. In this model the rate of growth depends on the local concentration of some substance, which diffuses from a distant source and is consumed by the growing object. The main assumptions of the algorithm are: (i) the diffusion process is fast compared to the growth process, a steady state is reached before each growth event; (ii) for the growth probability p, $p = a$

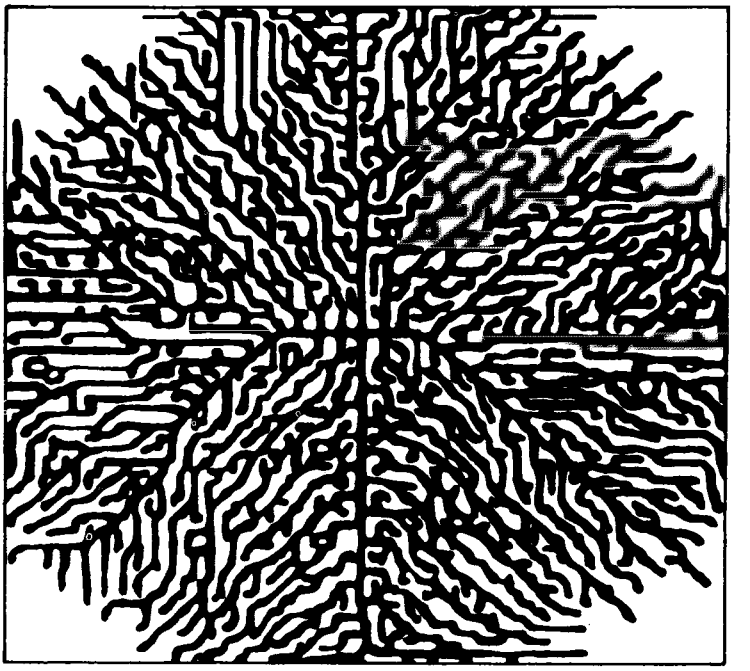


Fig. 2. A uniform branching pattern, generated by the Turing–Gierer–Meinhardt model.

c^n, where the value of the exponent n strongly influences the nature of the growing form; (iii) potential growth point are picked at random. A structure generated under these conditions is shown in Fig. 3.

It should be an open research program to decide (i) what kinds of neural dendritic patterns can be considered as fractal structures and (ii) how variations of the parameters of the growth process influence the geometry of the branching.

Cortical patterns

Some remarks on the modular architectonic principle

Experimental facts resulting from anatomy, physiology, embryology and psychophysics give evidence that the vertebrate nervous system is a highly ordered structure composed of *building blocks* of repetitive structures. The building blocks or *modular architectonic principle* is rather common in the nervous system (Szentágothai, 1983). After the anatomical demonstration of the so-called cortico-cortical columns (Goldman-Nauta, 1977) it was suggested (Szentágothai, 1978) that the cerebral cortex is considered on a large

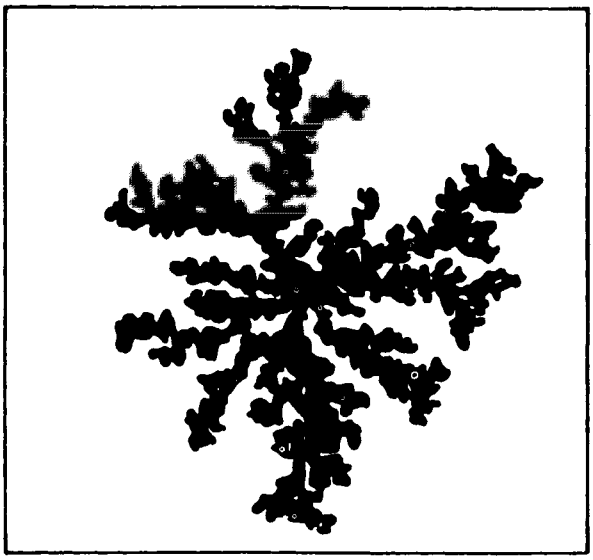


Fig. 3. A fractal branching pattern, generated by limited growth.

scale as a mosaic of vertical columns interconnected according to a pattern strictly specific to the species. Having been motivated by the pioneering work of Katchalsky (1974) on the dynamic patterns of neural assemblies Szentágothai offered to interpret the cortical order in terms of *dynamic structures* instead of applying a crystal-like approach. Claims regarding the modularity of the cortex have recently been raised (Swindale, 1990; for partial answer see Chapter 1, this volume).

Ocular dominance columns

The visual cortex is considered as the best paradigm of modular organization of the neural centers. *Ocular dominance* (OD) columns or *stripes* are characteristic examples of this kind of organization. In normal cats and monkeys, the afferents from lateral geniculate nucleus (LGN) laminae corresponding to the two eyes innervate common target structures, e.g. layer IV of the visual cortex and form partially overlapping bands. This alternating termination of LGN afferents showing about 300–400 μm wide periodicity is thought to be the anatomical substrate for the physiologically defined ODs (Hubel and Wiesel, 1972; Hubel, 1982; Wiesel 1982). Normal development leads to near-periodic spatial patterns (Hubel and Wiesel, 1977; see Fig. 4).

Activity patterns heavily influence the formation of ocularity domains. Treatment with tetrodotoxin of one eye of a kitten from 2 to 6 weeks of age

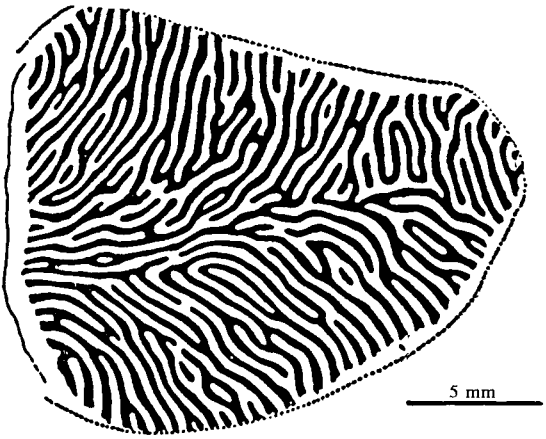


Fig. 4. Reconstruction of ocular dominance columns (layer IVc) of the visual cortex of a normally developed macaque monkey (from Hubel and Wiesel, 1977).

prevent the normal development of OD columns. It was demonstrated that activity plays a role in the columnar segregation and spontaneous activity (having random character) must be sufficient (Stryker and Harris, 1986).

A well-written comparative study on models of ontogenetic formation of OD was given recently (Goodhill, 1992). Our approach is slightly different from the models reviewed there, since we have been arguing for the beneficial (even indispensable) role of environmental noise during the developemnt of ordered neural structures (Érdi and Barna, 1984, 1985, 1987, 1991; Barna and Érdi, 1986; Érdi and Szentágothai, 1985).

Neurodynamic models are often two level dynamic systems which describe the activity change of single neurons and the modifiability of the synaptic efficacy. Schematically:

$$\begin{aligned} \Delta \mathbf{c}(t) &= \mathbf{e} f(\mathbf{c}(t), S(t))\Delta t) \\ \Delta S(t) &= I g(\mathbf{c}(t), S(t)) u \Delta t + R(t) \end{aligned} \tag{3}$$

Here $\mathbf{c}(t)$ is the activity vector, $S(t)$ is the matrix of the synaptic efficacy, $R(t)$ is an additive noise term to simulate environmental noise, u scales the time, $\mathbf{e}$ is the unit vector and I is the identity matrix. The function f is determined by the velocity of the transfer of the presynaptic information and the spontaneous activity decay. The function g contains the modification due to the local cooperation among neighboring synapses and global competition of ingrowing axons for receptor molecules. *Simulation experiments* showed that environmental noise is necessary in establishing globally ordered structures. Noise superimposed (i) might change the qualitative behavior of the deterministic system by destabilizing temporary, metastable

states; (ii) help to accelerate the transient processes. Figure 5 (Érdi and Barna, 1984) visualizes the values of the matrix S of the formed topographic connections established by an (over)simplified model.

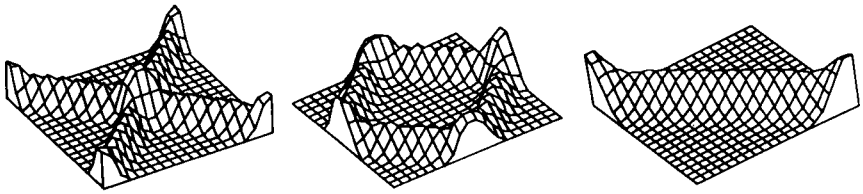


Fig. 5. Visualization of the qualitative behavior of a self-organizing algorithm. Values of the synaptic strength are plotted. *Double diagonal* and *cratern-like* solutions are produced by deterministic learning rules. The *diagonal* attractor associated to topographic order emerged by including a small additive random term.

Similarly, local and global order was generated for OD between one-dimensional pre- and postsynaptic sheets (Fig. 6).

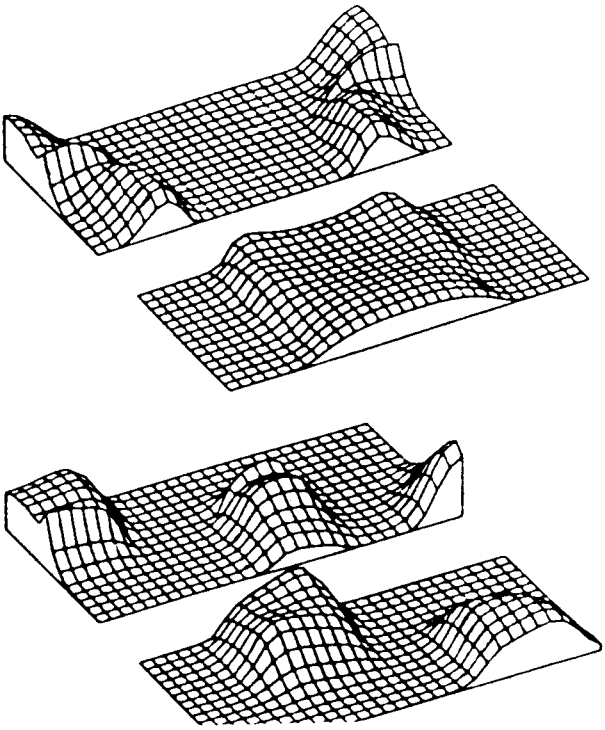


Fig. 6. Locally and globally ordered structures with pure deterministic and superimposed learning rules.

Simulation experiments with two-dimensional pre- and postsynaptic sheets consume much more time. However, the formation of alternating segregation of afferents was demonstrated even in this case (Fig. 7). To reduce computational time this experiment started from connectivity patterns corresponding to topographic order.

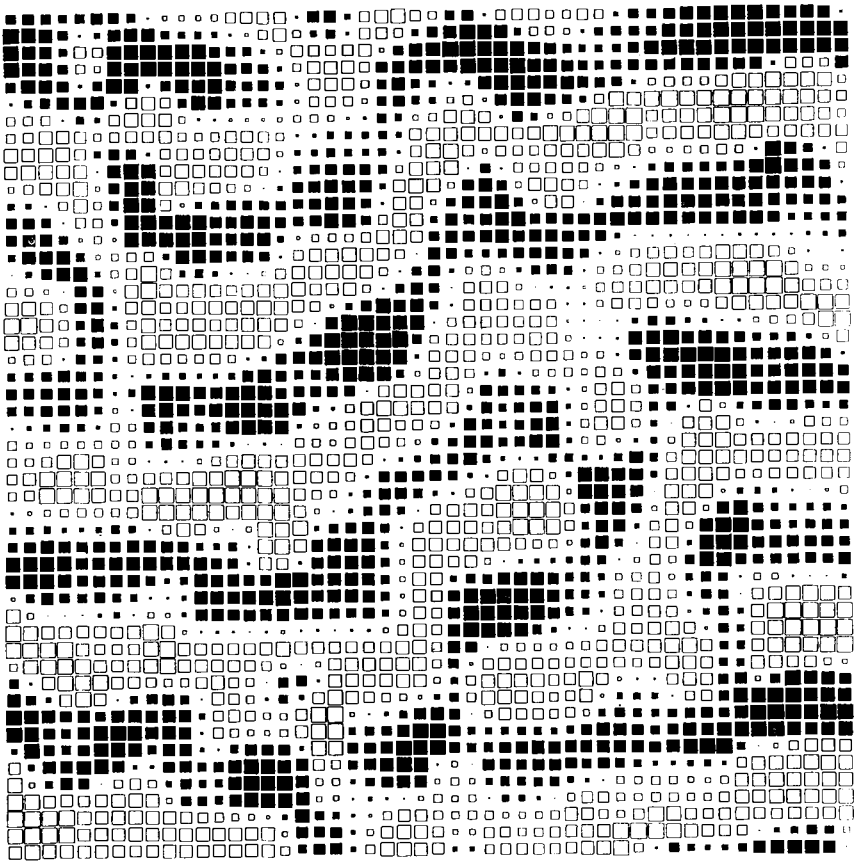


Fig. 7. Ocular dominance strips on the postsynaptic array. 50 x 50 cell matrix is shown with periodic boundary conditions.

Noise and chaos in neural systems: some facts and models

Noise versus chaos: some mathematical remarks

There are a lot of philosophical debates whether or not deterministic mechanisms leading to chaos are the ultimate and only sources of randomness. It is well known now that simple deterministic systems may show statistical mechanical behavior. The simplest of these deterministic systems is the Bernoulli shift. The Bernoulli shift is defined by:

$$x_n = \begin{cases} Dx_n; & 0 \leq x_n < 1/2 \\ Dx_n - 1; & 1/2 \leq x_n \rfloor 1 \end{cases} \tag{4}$$

To investigate the properties of this mapping the arithmetic is made by using binary expansion: an arbitrary initial value is represented by the dyadic form.

$$x_0 = \sum_{j=1}^{N} \frac{\epsilon_j}{2^j} = .\epsilon_1 \epsilon_2 ... \epsilon_N 00 \tag{5}$$

The mapping is called "shift", because each iteration of the map corresponds to a shift of the base point one place to the right. If x_0 is rational than the orbit is periodic, if it is irrational than a true chaotic orbit is obtained. The periodic orbits of shifts are (1) dense in the unit interval (since rational numbers are dense), (2) unstable (because of the exponential magnification of errors). For irrational initial values the closure of the orbit in the infinite time limit is the entire interval. The stationary probability distribution is just the uniform distribution. It can be shown that every smooth initial density evolves irreversibly to $p(x) = 1$. The ordinary household mixing is the proper analogy for the approach of a conservative dynamical system to statistical equilibrium. The most important properties of the Bernoulli shift (and a large class of others mappings) are:

(1) Sensitivity to initial conditions — If x_0 and x'_0 are different only at the nth digit, after n step they are different at the first digit.

(2) Randomness — The Bernoulli shift has the same statistical characteristics as the coin tosses defined by

$$x_n = \begin{cases} L; & if \quad x_n < 1/2 \\ R; & if \quad x_n \geq 1/2 \end{cases} \tag{6}$$

(3) Ergodicity — Almost all irrational numbers contain any given finite sequences of digits an infinite number of times. So, starting from any irrational x_0 we can approximate x with any precision for an unlimited number of times by the successive application of B shift.

Transmitter release: random or deterministic mechanism?

Miniature end-plate potentials (MEPPs) occurring spontaneously were first described by Fatt and Katz (1952). Statistical analyses of MEPPs seemed to be in accordance with the vesicular hypothesis based on the morphological demonstration of the synaptic vesicles (De Roberties and Bennett, 1954). According to the classical, vesicular hypothesis (Del Castillo and Katz, 1954) packets of transmitter molecules ("quanta") contained synaptic vesicles are released by exocytosis. The random character of the amplitude- and time-interval distributions of the MEPPS was explained by variable content of the quanta.

Recently Kriebel *et al.* (1990) made an alternative hypothesis on the nature of the transmitter release. Comparing the statistical behavior of the two processes, the *dripping faucet* was suggested to be a proper model of the release process. Dripping faucet is a common example of the transition from periodicity to chaos (Crutchfield *et al.*, 1986). The main points of the suggestion by Kriebel *et al.* are (i) that the packets are not preformed but dynamically generated during the release process, (ii) MEPPs are not random but chaotic signals.

Noise

Stochastic neurodynamics

Stochastic neurodynamics offers a compromise between models based on detailed single cell level description and oversimplified networks of binary units: large-scale neural activities are calculated (Cowan, 1991).

A neuron is supposed to have three possible states: *quiescent* (q), *activated* (a) and *refractory* (r). Transition probabilities for elementary changes between subpopulations have to be specified; e.g. an elementary activation process q → a implies the transition

$$(n_q, n_a, n_r) \rightarrow ((n-1)_q, (n+1)_a, n_r). \tag{7}$$

Assuming that the system is a time-homogenous continuous time discrete state space Markovian jump process, the evolution equation for the probability distribution of the states is

$$dP_n(t)/dt = AP_n(t), \tag{8}$$

where A is the (time independent) interaction matrix.

What can be done if the assumptions behind the model are violated?

(i) The transition probabilities can be time-dependent, i.e., the Markovian process is inhomogeneous in time and the corresponding evolution equation is

$$dP_n(t)/dt = A(t)P_n(t). \tag{9}$$

(ii) While Markovian jump processes are characterized by exponential waiting time distribution, the relaxation of this condition leads to semi-Markov processes governed by the integral equation

$$dP(t)/dt = \int_{t_0}^{t} A_{ts} P(s) ds \tag{10}$$

where A_{ts} is the *kernel function.*

(iii) The violation of locality in state space leads to "Levy processes", (see e.g. West, 1986). The characteristic function of the Levy process is

$$f(k,t) = exp(-bt \mid k \mid^{\alpha}), b > 0, \alpha \leq 2.$$

For $\alpha = 1$ the Cauchy process, for $\alpha = 2$ the diffusion process with Gaussian distribution is specified. For $0 < \alpha < 2$ and $\alpha \neq 1$ the Levy process is associated to random walk on fractals. The $P_{\alpha}(x,t)$ distribution function (generated from the characteristic function) do not posses finite moments. The divergence of the variance implies that conventional statistical methods do not work. The r.h.s. of the evolution equation can not be constructed by a differential operator.

Noise, as source of information

The fluctuation–dissipation theorem of thermodynamics suggests that noises identified with fluctuations can be analyzed to obtain information for the parameters of the process tending to equilibrium. Ion channels mechanisms, e.g. kinetic parameters of conformational transitions of ionic channels can be estimated using noise analysis (e.g. Holden, 1976; Sakmann and Neher, 1983). *Biological* contrary to *physical* noise phenomena are explained by Abeles (see Chapter 3 in this volume).

Noise-induced transitions

A possible resolution of the determinism–randomness dilemma was based on the principle described as *"randomness in the small and structure in the large"* (Anninos *et al.*, 1970; Harth *et al.*, 1970). The spirit of this principle accepts the existence of stochastic effects in the brain, but allows the *averaging out* of the fluctuations. Equations for single neuron level can have random character (Chapter 7 this volume), but the evolution equations at the population levels are deterministic ones.

In the theory of dynamic systems, it turned out that in certain situations (mostly in the neighborhood of instability points) the effect of the fluctuations can not be averaged out, on the contrary, they may cause drastic effects at the macroscopic level. The theory of *noise-induced transitions* (Horsthemke and Lefever, 1984) shows that fluctuations may operate as "organizing forces". Stated in a more technical way, noise may cause the instability of the deterministic attractors and a stochastic model might exhibit properties qualitatively different from those of the deterministic model.

Chaos in neural systems

Chaotic temporal patterns

Both periodic and chaotic temporal patterns can be generated at single neuron level, at multicellular network level due to the interactions among neurons and at global level in consequence of spatiotemporal integration.

Chaos in neural membranes

Chaos can be generated endogenously and by periodic forcing. A few chaotic phenomena have experimentally been demonstrated at the single neuron level. Molluscan neurons have shown chaotic activity induced by K^+ blocking agents (Holden *et al.*, 1982). Rapp *et al.* (1985) reported that spontaneous activity in the precentral and postcentral gyri of the *squirrel monkey* may be chaotic. The *Onchidium pacemaker neuron* can show regular beating, regular bursting and irregular bursting discharges depending on the current applied (Hayashi *et al.*, 1982, 1992). Membrane potential of the *squid giant axons* show periodic and non periodic responses (Matsumoto *et al.*, 1984). *Hippocampal pyramidal cells* may also be a candidate for the chaotic pattern generator (Hayashi *et al.*, 1990). Holden (1984, 1991) remarked that chaotic behavior only occurs in neurons (and model neurons) under rather restrictive conditions and for small ranges of parameters.

Global cortical dynamics

The technique of recording electroencephalograms (EEG) has opened a window for measuring and understanding brain activity. Though the underlying biophysical mechanisms of EEG signal generation is not completely clear, still the processing of EEG curves gives information about the structure–function relationship of the brain. According to the nowadays already classical approach, the records, as time series, are considered as stationary stochastic processes and their characteristic frequencies are

determined by spectral analysis. Nonlinear dynamic system theory offers a different conceptual approach to EEG signal processing (see e.g. Dvorak and Holden, 1991). Time series, even irregular ones, are considered as deterministic phenomena generated by nonlinear differential equations. First, the *phase space* of the dynamic system is reconstructed, then, its (embedding and correlation) dimensions are calculated (Grassberger and Procaccia, 1983). Low (but larger than two) *fractal dimensions* are associated with *deterministic chaos*, while it is not easy (if at all) to discriminate between high-dimensional chaos and random process. Lyapunov exponents measure the exponential divergence of nearby trajectories (for a new and efficient algorithm, see Barna, 1992). The presence of at least one positive *Lyapunov exponent* is the signature of chaotic dynamics. The criteria of chaos were reported to be fulfilled for the *sleep cycle* (Babloyantz *et al.*, 1985), *epileptic petit-mal seizure* (Babloyantz and Destexhe, 1986), alpha rhythm (Soong and Stuart, 1989) and Creutzfeld–Jacob disease (Gallez and Babloyantz, 1991).

Dynamical disease and controlling chaos

A *dynamical disease* is defined as a disease that occurs in an intact physiological system that leads to abnormal dynamics. Chaotic behavior was found in a model of the central dopaminergic neuronal system and was associated with schizophrenics (King *et al.*, 1984). Synaptic level rhythmic generator of cholinergic system requires a fine-tuned neurochemical control system. Even mild impairment of the metabolism might imply "abnormal" dynamic synaptic activity. (Simulation experiments of a model based of the transmitter–recycle hypothesis: Barna and Érdi, 1988.) Not only the appearance of chaos but other qualitative changes (emergence or loss of rhythms, appearance of altered or new periodicities) can be related to dynamic diseases (about the temporal disorders in the human oscillatory systems see Rensing *et al.*, 1987).

Controlling chaos is a concept initiated by the following question: how to stabilize unstable periodic orbits embedded within a strange attractor by making only small time-dependent perturbations in accessible system parameter (Ott *et al.*, 1990)? While control algorithms (e.g. proportional-feedback algorithm, see Petrov *et al.*, 1992) was adopted for controlling systems "from outside", as in the case of open chemical systems, in neural context properly chosen "learning rules" can play the role of "internal" control algorithms.

Possible functional role of chaos

In spite of a lot of effort, we still do not have a coherent view about the eventual functional role of chaos in information processing (see Freeman, this volume). A list is given below based on Tsuda (1992):

- novelty filter (Skarda and Freeman, 1987),
- a nonlinear pattern classifier (Freeman *et al.*, 1988),
- a catalyst for learning (Skarda and Freeman, 1987),
- a short-term memory generator (Nicolis and Tsuda, 1985),
- a stimulus interpreter (Tsuda, 1984),
- a memory searcher (Tsuda *et al.*, 1987),
- a copy machine of a periodic oscillation (Tsuda and Shimizu, 1985),
- a dynamic storeroom of long-term memory (Tsuda, 1991),
- a self-referential logic generator (Nicolis and Tsuda, 1985).

Network structure and stability

General remarks

Studies on connections between the structure of neural networks and the quantitative and/or qualitative properties of their dynamic behavior gave the possibility to reinterpret the structure–function relationship. Synthesis and analysis of *neural networks* relies on the eventual existence of general relationship between the excitatory–inhibitory patterns of the network and the different types of temporal behavior, as convergence to equilibrium point, oscillation and chaos. As it is well known, dynamic behavior of *autonomous systems* is determined by three different factors, as network structure, parameter values and initial conditions. The quantitative and qualitative effects of parameter changes are studied by *sensitivity analysis and bifurcation theory* , respectively. The set of initial conditions are classified by the basins of the *attractors.* Here the relationship between network structure and dynamic behavior is discussed.

Searching the structural conditions of stability, periodicity, chaos etc. we have to realise that the overwhelming majority of the statements is conditional, i.e., system properties also depend on the numerical values of parameters and initial conditions. Some properties (e.g. qualitative stability) depend only on the sign of parameters characterizing the strengths of connections. The question to be studied here is what kinds of network architecture lead to stable equilibrium and which are the structural conditions for a network to be candidate for a rhythmicity or chaos generator (Érdi, 1991)?

The connectivity–stability dilemma

In the cybernetic literature there was a statement asserting that a system loses its stability as the connectivity exceeds a certain threshold. In other words, in order to have a *stable equilibrium point* a large system should

not be strongly connected. Simulation experiments of randomly assembled deterministic systems supported this statement (Gardner and Ashby, 1970). Theoretical results for the connectivity–stability problem have been given within the framework of the *Wigner–May theorem* (Wigner, 1959; May, 1972) by analyzing the eigenvalue of certain matrices. Let B a random $n \times n$ matrix of connectivity c_n (B has $c_n n^2$ non-zero elements). The non-zero elements are chosen independently from a fixed symmetric distribution of mean 0 and variance α^2. Then the trivial solution of the system of differential equations:

$$x = A\,x \qquad\qquad \text{with } A := B - I,$$

is for large n almost surely stable if $\alpha^2 n\, c_n < 1$ and almost surely unstable if $\alpha^2 n\, c_n > 1$. (α can be interpreted in neural network context as the synaptic strength.) This theorem is usually considered as a general statement on randomly assembled large deterministic systems, connecting structural complexity and dynamic stability. Many misinterpretations of the theorem have appeared in the biomathematical literature due to the neglect of its very specific assumptions; e.g. the immediate consequence of assuming symmetrically distributed matrix elements with zero mean is that the expected number of positive and negative elements (i.e. of the excitatory and inhibitory interactions) are the same. (This condition is also assumed by Hopfield, 1982.) Some remarks on the applications and some hints on the extensions and limits of the Wigner–May theorem have been made (Érdi and Tóth, 1990). In order to apply the theorem it is necessary to precisely check *all* its assumptions, or to clearly delineate its domain of validity. The ratio of positive and negative values in the Jacobian is the assumption most often violated. Intuitively it is not surprising that non-balanced excitatory–inhibitory patterns might lead to instability. Furthermore, in certain situations, increase in connectivity may enhance stability. We can make assumptions for the sign but not for the numerical values of the strength of interactions. Given a sign matrix formed by the J_{ij} Jacobian matrix by writing +, if $\partial F_i/\partial x_j > 0$ and −, if $\delta F_i/\delta x_j < 0$, (and assuming that the J preserves its sign stability) we might ask whether or not there is an admissible system of differential equations $dx(t)/dt = F(x(t))$ that admits at least one stable equilibrium point. The answer is yes, if J is qualitatively stable. A J matrix (derived from F) is qualitatively (un)stable if every sign equivalent is (un)stable. The requirements of the qualitative stability (not listed here, see e.g. Jeffries, 1974) are very restrictive. Very loosely speaking, qualitative stability requires many zeros in the Jacobian. (Of course, the lack of qualitative stability does not imply qualitative instability. A qualitative instability test has been given by Eisenfeld and DeLisi (1985).) Qualitative stability implies convergence to a stable equilibrium point (regular behavior). The violation of the criteria of qualitative stability can (conditionally) lead to periodic and chaotic (i.e. exotic) behavior.

Regular and exotic behavior: some remarks

It was demonstrated by examples in the context of chemical instabilities (Tyson, 1975), that two-dimensional models of chemical reactions might exhibit oscillatory behavior when some condition of qualitative stability does not hold. The existence of self-excitation ($J_{ii} > 0$), the occurrence of competition ($J_{ij} < 0$ and $J_{ji} < 0$) and cooperation ($J_{ij} > 0$ and $J_{ji} > 0$), the existence of positive feedback loop ($J_{ij}\ J_{jk} \ldots J_{qi} > 0$) and of negative feedback loop ($J_{ij}\ J_{jk} \ldots J_{qi} < 0$) are possible sources of instabilities. Recent results in the theory of differential equations (Hirsch, 1991 and references cited therein) made relationships between structures with given properties and dynamics of the system. The structure of the Jacobian matrix of five (not disjoint) important groups are shown (definitions and statements are not restricted to 3D systems).

Cooperative (1) systems (supplemented with proper negative elements) show convergence. Competitive systems can be transformed to cooperative (2) ones, but the original system can have attracting period points. Non-cooperative, but sign-symmetric (3) systems show convergence, when the system has a feedback loop (4) with the even-loop property (i.e. when the number of edges with negative sign is even). Self-excitation (5), (positive feedback or autocatalysis, using different terminologies) is being thought to be an important source of periodicity.

Network structure and chaos: some designing principles

It is well-known, that the method of local analysis, as linear stability analysis, is not an appropriate method for detecting chaos. Still, for 3D systems without self-activation King (1983) gave conditions for Jacobian matrix structures to be prone to chaos. Only those Jacobians are candidates for chaos, which graphs are characterized with three conditions: (i) the graph is connected; (ii) each node has at least one edge directed towards it and one edge directed away from it; (iii) there is a node with at least two edges directed towards it. The structures of the Jacobian satisfying these criteria are shown below.

Oscillation and chaos in a model of the olfactory bulb: a case study

On the dynamics of the olfactory bulb

The olfactory system and in particular, the *olfactory bulb* (OB) is a "working horse" to investigate the structural conditions and functional significance of *chaotic information processing* in *neural networks* (Freeman, 1987). Oscillatory activities were studied by Li and Hopfield (1989) in the framework of a model based on the connections among oscillators formed by pairs of an *excitatory mitral* and an *inhibitory granule cells.* The extended version of this model, taking into account even the *lateral connections* in the layer of the mitral cells, is systematically studied (Érdi *et al.*, 1991). To study the inherent dynamics of the bulbar activity due to the interaction among the two most important cell populations, the explicit time dependence of the sensory input ("sniff cycle") has been neglected. By taking into account the time-dependence of inputs the model would become a system of non autonomous differential equations. Non autonomous systems do not have attractors in the general case and, obviously, the change of state depends not only on the state itself, but also on the external effects. A series of bifurcation phenomena in terms of the strength and sign of these connections is demonstrated. Chaos occurs at relatively large values of excitatory connections only.

The model

The model consists of two layers of neurons: (i) the layer of excitatory *mitral cells* and (ii) the layer of inhibitory *granule cells.* Mitral cells receive inputs from the olfactory receptor cells. Each mitral cell is connected to the three nearest granule cells. These mitral–granule synapses are dendrodendritic and excitatory. Granule cells receive modulatory input from cortical regions and other higher centers. The granule–mitral synapses are inhibitory. No lateral connection is assumed in the granule cell layer. The physiological nature of lateral connections in the mitral layer is not specified: both excitatory and inhibitory connections are investigated. Membrane potentials of mitral and granule cells at time t are expressed by the real vectors m (t) and g (t), respectively. Their transfer functions are denoted by ϕ_m and ϕ_g. The underlying differential equations for the mitral and granule cell activities are as follows:

$$
\begin{aligned}
m_i(t)/dt &= -am_i(t) - b\textstyle\sum_{j=1}^{N} G_{ij}\phi_g(g_j(t)) + \\
&\quad c\textstyle\sum_{j=1}^{N} L_{ij}\phi_m(m_j(t)) + I_i^{rec} \\
dg_i(t)/dt &= -qg_i(t) + e\sum_{j=1}^{N} M_{ij}\phi_m(m_j(t)) + I_i^{cort}
\end{aligned}
\tag{11}
$$

where a, b, q, e are positive constants and c can be positive or negative. The synaptic matrices M, G, L correspond to the mitral–granule, granule–mitral and lateral (mitral–mitral) connections, respectively. I^{cort} and I^{rec} denote the cortical and receptor inputs, respectively and are constant in time. While all elements of the modulatory cortical input vector I^{cort} have the same values, odor patterns are taken into account by assigning random values to the different elements I_i^{rec}.

Given a model structure and the set of ODEs to be solved, the first question arising is what qualitative dynamical behavior can be expected with different sets of parameter values. One way of answering this question is to find the attractors to which the time evolution of the system leads. The most striking difference between the above equation and the model set up by Li and Hopfield is the existence of lateral connections. It is therefore particularly interesting to study the different — positive and negative — values of the parameter c which controls the sign and strength of lateral connections while the synaptic matrix L (and also M and G) are fixed. The main question is which values of the parameter c imply stationary, oscillatory or chaotic behavior and what kinds of bifurcation phenomena can be observed between the different parameter regions. To complete the analysis of the model, tests with the other parameters listed above have also been done exhaustively.

Simulation results

Simulations have shown that the parameters listed above can be classified into two groups: c, q and I^{rec} produce the same sequence of *attractors* together with the corresponding bifurcation points although the size of the parameter windows may be different. This can be explained by the fact that all these parameters act in the direction of mitral cell excitation (or disinhibition). The other group of the parameters, i.e. a, b, e and I^{cort}, have the opposite effect (mitral cell inhibition or granule cell excitation) thus they produce the same sequence of attractors in reversed order. These observations apply for both positive and negative values of c. With this in mind, results with only the parameter c (lateral connection strength) are presented. Non-zero elements of the matrices M, G and L are all set to 1. Values of the other parameters are fixed and these values are used in all subsequent simulations: $N = 11$, $a = 0.1$, $b = 1.0$, $f = 0.2$, $e = 1.2$, $I_i^{cort} = 0.1$. The mean value of I_i^{rec} is $\langle I^{rec} \rangle = 1.0$.

At $c = 0$ the trajectories converge into a stable focus. When lateral excitation is considered ($c > 0$), the following attractor regions can be observed as c increases:

I. stable focus
II. limit cycle
III. strange attractor

IV. coexistence of limit cycle and strange attractor

V. coexistence of stable node and strange attractor.

For negative values of c (lateral inhibition), the attractor regions are (as c increases) as follows:

VI. stable focus

VII. limit cycle

VIII. stable focus.

These results show that chaos can only be found when $c > 0$ so, in our model, lateral excitation is necessary to obtain chaotic activity. The Poincaré plot of chaotic activity in region III is shown in Fig. 8.

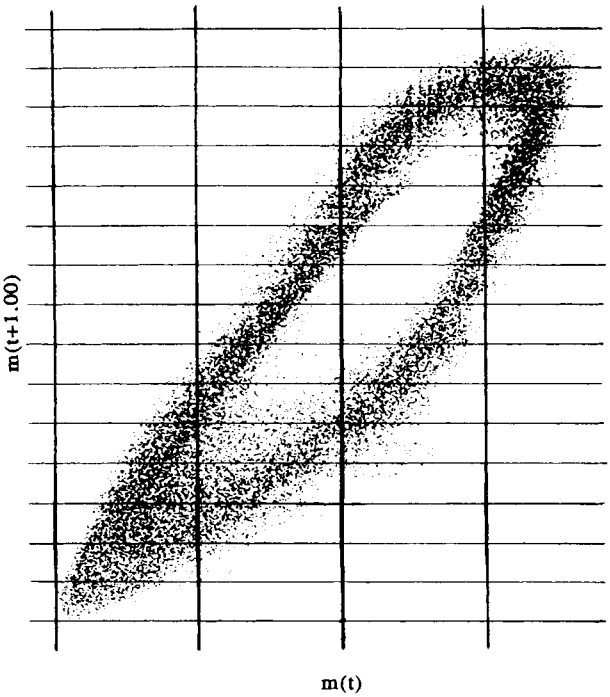


Fig. 8. Poincaré plot of chaotic activity.

For a complete analysis, not only the attractor regions but also the different types of bifurcations between them should be identified. A proper way of representing the results is the *phase diagram.* The activity of an arbitrary mitral–granule pair is chosen for two-dimensional representation. First, it can be observed that weak excitatory, weak inhibitory connections and the state of non-existing lateral connections lie in the same region, there is no qualitative difference between them. The first bifurcation from this region with one stable focus is a *Hopf bifurcation* in the case of both lateral excitation and inhibition (between regions I–II and VI–VII). The bifurcation between regions II and III is a *period-doubling bifurcation.* The *limit cycle* ,

the *double-period limit cycle* and the *strange attractor* are shown in Figs 9a, b and c, respectively. While the strange attractor remains stable over the regions III, IV and V, a *stable limit cycle* also appears in region IV (Fig. 10). Further increasing c, the period of this limit cycle grows unboundedly and, as it becomes infinitely large, the limit cycle disappears and a stable node emerges. This is characteristic of the so-called saddle-node infinite period bifurcation.

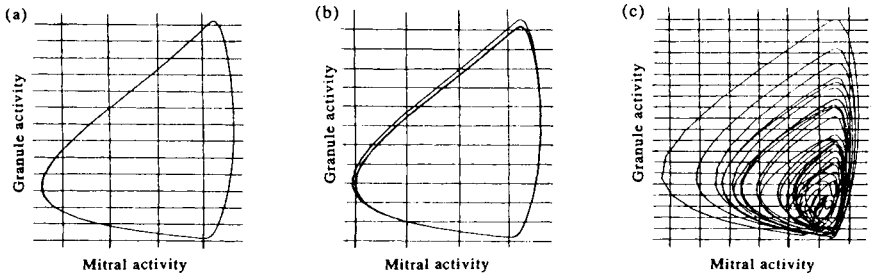


Fig. 9. Period doubling bifurcation. (a) Periodic attractor; (b) biperiodic attractor; (c) strange attractor.

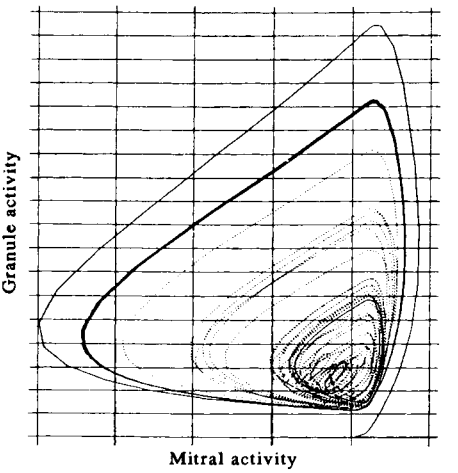


Fig. 10. Coexistence of periodic and strange attractor. Solid line: periodic attractor with the initial transient. Broken line: strange attractor.

For negative values of c, the bifurcation between the regions VI and VII has already been mentioned. The bifurcation between the regions VII and VIII is different from this in that the limit cycle has a finite amplitude at the bifurcation point. The effects of both lateral inhibition and excitation on the dynamic behavior of a model of the OB have been studied. The connection strengths are taken to be constant over time but lateral coupling can be varied through a control parameter (c). The sign of c determines whether the lateral connections are considered as excitatory or inhibitory. The results of the computer simulations show that *chaotic activity* can only

be found if lateral excitation is assumed. Lateral interactions in the mitral layer seem to have special importance in the model. A transition from oscillatory to chaotic region by the continuous time modification of lateral synaptic strengths can also be demonstrated (Fig. 11).

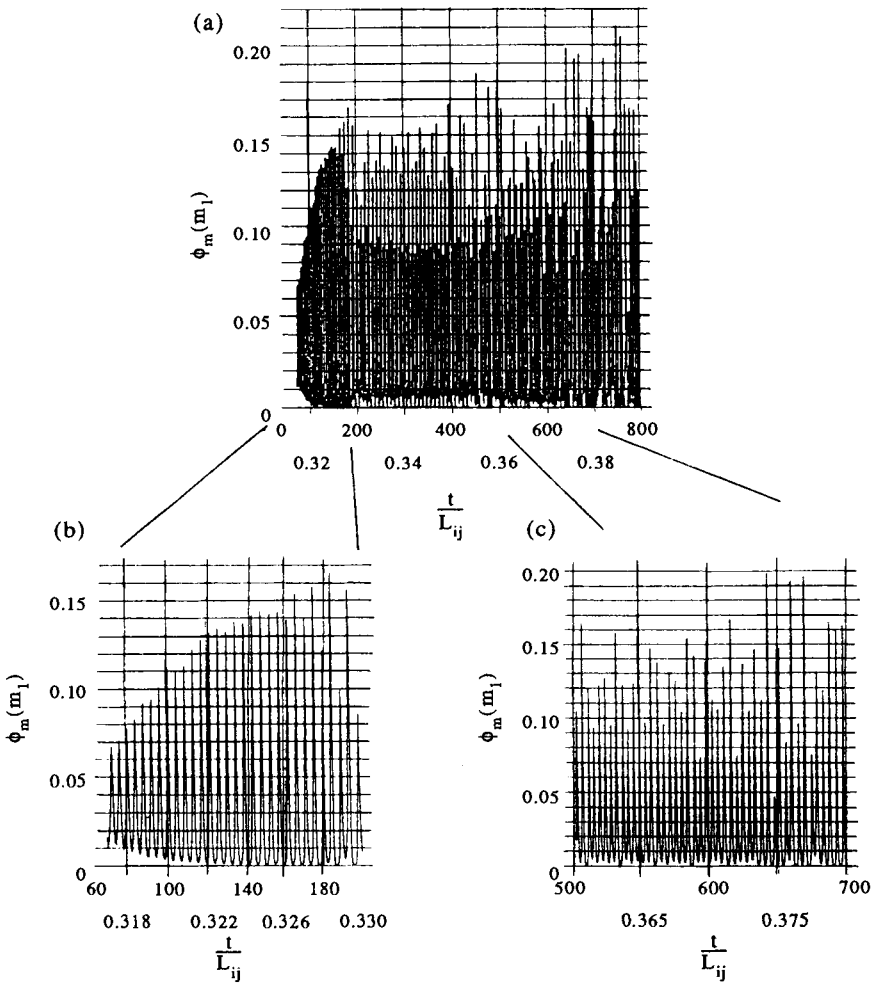


Fig. 11. Synaptic modification induced bifurcation. Both the time scale and the average lateral synaptic strength are indicated at the horizontal axes. (a) Full learning process; (b) early learning phase; (c) emergence of chaos.

Olfactory bulb can exhibit a large set of dynamic behaviors and bifurcations. The role of dynamics and particularly of chaos in the odor information processing is discussed by Freeman (see Chapter 9 in this volume).

Acknowledgements

This chapter is dedicated to 80th birthday of János Szentágothai.

9

Qualitative Overview of Population Neurodynamics

WALTER J. FREEMAN

Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, CA 94720, U.S.A.

Introduction

Two salient facts about neurons are that each functional part of the brain contains immeasurably large numbers of them and that each neuron forms synaptic connections to and from immeasurably large numbers of other neurons (Braitenberg and Schutz, 1991). Statistical sampling is used. Estimates of the number of neurons in a cubic millimeter of cortex range from 10^4 to 10^6 and for the number of synaptic connections on each dendritic tree from 10^3 to 10^5. The typical sensory input for a conditioned stimulus is carried by a parallel array of an immeasurably large number of receptor axons and the motor outflow for a typical conditioned response a fraction of a second later is carried in parallel by an immeasurably large number of motor axons. Every goal-directed response involves stupendous numbers of neurons in multiple brain parts.

Much of what is understood about the functions of these massive numbers of neurons comes from recording the activity of neurons one at a time with microelectrodes. Networks of model neurons are constructed by induction that represent the actions of a small subset of neurons, on the implicit assumption that what the other (unobserved) neurons are doing during a perceptual act or a conditioned reflex is not important for the activity of the observed subset. This approach might be called statistical me-

chanics (Wilson and Cowan, 1972; Amari, 1974; Amari and Maginu, 1988; Hansel and Sompolinsky, 1990) of brain function, comparable to the study of molecules in an ideal gas in terms of their kinetic energy, position, collision rates, etc.

A different approach is to conceive not individual neurons but interactive populations of neurons. The weak interactions among immeasurable numbers can be predicted to give rise to cooperative properties, which can only be accessed by measuring local mean field quantities constituting observable macroscopic variables for the masses. This alternative approach might be called a thermodynamics of brain function, by which the equivalents are sought for the temperature, pressure, viscosity, etc. of an ideal gas. However, physical analogs on their face are inappropriate and have led many physicists astray. Neurons are not molecules in a gas, so a fresh beginning must be made for a new brain science (Skarda and Freeman, 1987).

Neural populations come into play significantly during animal behavior and their properties cannot be understood from studies of isolated neural preparations *in vitro* or in brains under anesthesia. Therefore this new approach is introduced in the context of acts of animal perception and in particular the *perception* of odors through the *olfactory system* (Fig. 1), because it is the simplest and best known at present and because there is good reason to propose (Herrick, 1948) that it is the phylogenetic prototype for the algorithms by which all sensory systems operate in perception and indeed for all instances of behaviorally related interactions of cortex within itself and with subcortical nuclei and the far reaches of the brainstem and spinal cord.

Consideration is restricted here to acts of pre-attentive perception that do not require inspection or combining successive stimuli over time. Such acts occur in each sensory system independently of other senses on first arrival, but in the time to the next heartbeat, eye movement, or inhalation, all the senses are involved. This takes place by convergence from several sensory cortices into the *limbic* and *motor systems* of the brain (Fig. 1). These acts of perception occur in all vertebrates and perhaps in all animals. The acts are rapid in onset, short in duration, reliable and reproducible. In animals they are easily controlled by standard techniques for conditioning. For these reasons a *rabbit* that has been trained to sniff an odorant is an optimal subject for the study of the neurobiology of perception in the olfactory system, with the hope that what is learned about perceptual coding will hold for all other senses as well (Freeman and van Dijk, 1987; Tsuda, 1991).

It is known from lesion studies, in which the damage is caused by disease or experimental surgery, that the neural activities sustaining these acts of pre-attentive perception take place in the outer shell of the brain, the cerebral cortex. From behavioral experiments with trained animals it is known that each act takes place just after some sensory input has been transmitted to each of the sensory cortices in the brain and that it then takes

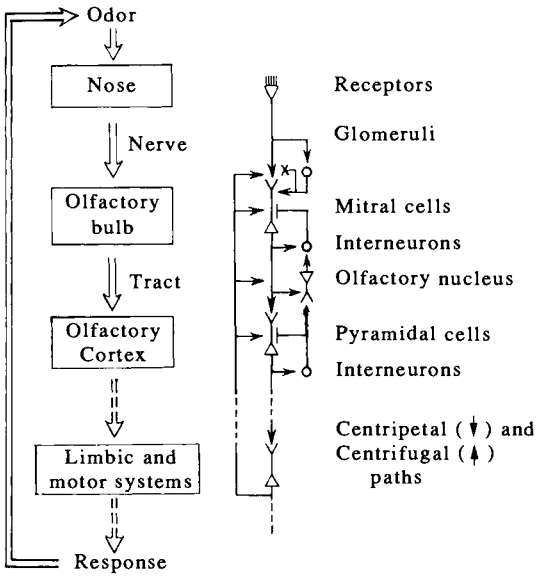


Fig. 1. The main flow of neural activity in the olfactory system is forward (centripetal) through three stages. Feedback occurs within, between and around the stages. Feedback from the brain into the bulb and cortex is denoted centrifugal. There are 10–12 centrifugal paths with various functions (from Freeman, 1983).

place within a few tenths of a second of stimulus arrival, before any decision is taken on what to do next. *Response latency* measurements show that it happens in the time needed for a *sniff* or a *saccade* ("a twinkling of the eye"). These findings show exactly where and when to look for the behaviorally related physiological process in the brain that sustains a percept.

Previous studies in neuronal physiology (e.g. Shepherd, 1983) have shown how sensory stimuli are analyzed, not how percepts are synthesized in the brain. From studies combining theory and experiment it is known how sensory stimuli are transduced by receptor neurons in a two-stage process. First the stimulus is converted to a loop current with its energy source in the membrane at or near the site of action of the stimulus and its effective action at the initial segment of the sensory axon, where its amplitude is re-expressed by the frequency of a train of action potentials. The loop current flows inside the cell in one direction and outside the cell in the other direction across the tissue resistance, giving rise to a "generator potential", which for the olfactory receptors in the nose is called (Fig. 2) the *electro-olfactogram*. This two-stage operation also serves as a model for the function of cortical neurons.

They receive action potentials ("units") at the *synapses* on their dendrites and convert them to dendritic loop currents. The currents are summed at the initial segments of their axons (the *"trigger zones"*) where the net

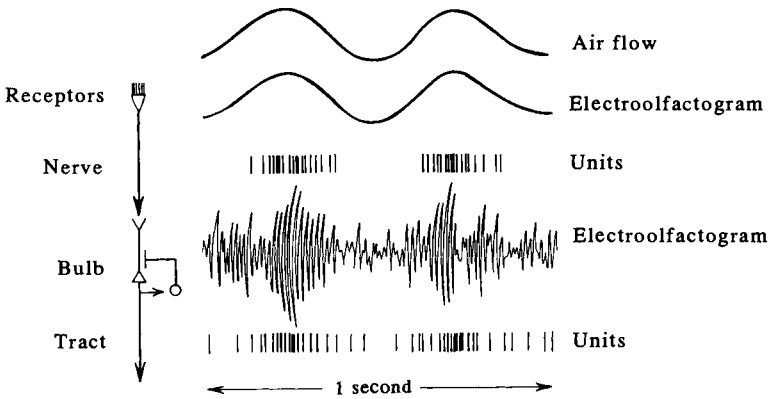


Fig. 2. For each neuron the dendrites receive input at synapses, convert the axon pulse input to loop current and sum it at the cell body. The axon converts the sum to a pulse train for transmission to other neurons. The loop current in the wave mode for olfactory receptors is the basis for the receptor potential known as the "electro-olfactogram" (EOG) and for cortical neurons it is the main contributor to the electroencephalogram (EEG). The air flow over the receptors varies with respiration, leading to the low frequency waves in the EOG and EEG, the correlated volleys of nerve action potentials ("units") and the high frequency oscillations in the "bursts" in the EEG.

amplitude determines the frequency of axonal firing. The *dendritic currents* cause oscillating waves of electrical potentials to appear in and around cortex (Basar, 1980) called the *electroencephalogram* (EEG). Thus there are two main state variables by which neural activity is carried: pulses on axons and waves on dendrites.

These sensory stimuli are pre-analyzed in both the pulse and wave modes at intermediate stations. In the visual system which is the best known in regard to pre-processing, there are complex operations in the retina and the lower brain called "adaptation", "range compression", "contrast enhancement", "motion detection" and so on. Sensory stimuli in the cortex selectively activate neurons that are called "feature extractors", such as "line" or "bug" detectors, "face" or "hand" cells, or the archetypal "grandmother cells" that are conceived to fire whenever that person appears. Comparable operations have been described for the auditory and somatic cortices as well. Typically the neurons in these cortices receive input axons that are organized in parallel arrays to provide the anatomical basis for *topographic mapping* onto the brain from receptor arrays in a body surface such as the retina, skin and ear. The alignment of input axons side by side provides the anatomical basis for extracting spatial and temporal derivatives of input for edge and motion detection and for contrast enhancement.

Similar pre-analysis also takes place in the *olfactory system*, but the processing is far less complicated and largely consists of dynamic range com-

pression and signal normalization (Freeman, 1975), though some degree of contrast enhancement may take place (Shepherd, 1983). The simplicity makes easier the task of finding *percepts*. The olfactory system has a rough *topographic map* from its receptor neurons in the nose into the receiving sheet of cortical tissue, the *olfactory bulb* (Fig. 3).

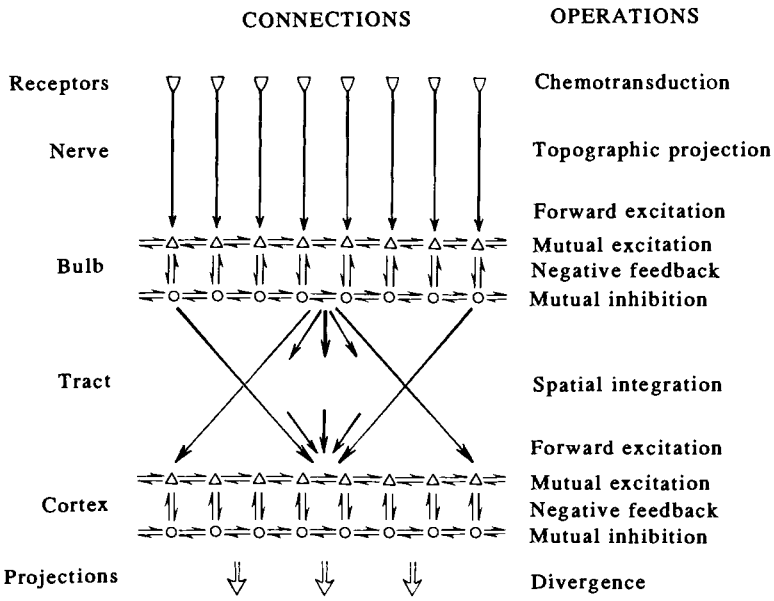


Fig. 3. Each stage of the olfactory system is formed by sheets of neurons. The receptors form a single noninteractive layer and their axons form an array that projects in parallel by the primary olfactory nerve with a degree of topographic order that maps the receptors into the bulb. The bulb and cortex contain both excitatory and inhibitory neurons, whose negative feedback interactions sustain the oscillations seen in the EEGs of these structures (Fig. 2). The output of the bulb to the cortex by the lateral olfactory tract has strong divergence as the basis for spatial integration and does not have topographic order (from Freeman, 1983).

Roughly 50 million *olfactory receptors* in the rabbit transmit their output by unbranched axons to half a million neurons in the bulb. As is typical for sensory systems a large number of receptors converges to a small number of bulbar neurons, the convergence ratio here being about a thousand to one.

The axons of the bulbar neurons that carry pulses to the olfactory cortex do not form a topographic map. Each output axon has many branches that diverge widely over the cortex. Conversely, each cortical neuron receives input from neurons that are widely dispersed throughout the bulb, thereby performing not only temporal integration but spatial integration as well. This divergence is an important property which will repeatedly referred to in this review. This type of divergent connection may in fact be much more common in cortical connections than is the topographic map.

The studies in olfaction will now be used to answer the question of how, within a few tenths of a second, an act of pre-attentive perception is accomplished, that is, how relevant sensory input is (a) extracted from the environment by a sniff, (b) globally integrated by arrays of cortical neurons and (c) combined with pertinent past experience and expectancy of future events into a space-time pattern of cortical activity. That pattern is called a *"wave packet"* in its physical aspect and a *"percept"* in its behavioral or functional aspect (Freeman, 1975). The answer will be expressed in the language of nonlinear dynamics of neural populations.

Populations and ensemble averages

The basic anatomy and physiology of neurons is familiar in terms of the *Neuron Doctrine*, which holds that the element of neural function is the individual neuron. Complementary to this doctrine is the concept of the cooperative neural mass, that exists through synaptic connections among participating neurons (Fig. 3). In this hierarchical view the *population activity* co-exists with the trains of action potentials from individual neurons in the form of shared spatially distributed patterns of activity occupying large areas of cortex, having sudden onsets and offsets and lasting small fractions of a second. These *wave packets* have only recently become available for measurement with large arrays of electrodes and amplifiers. They form the foundation of this new approach.

The concept of coexisting "microscopic" and "macroscopic" activity is familiar. Each person as an individual speaks with one voice, but each participates in families, committees and economic units. Those larger participations become apparent only in the aggregate, such as committee reports, vote tallies or economic indices. Similarly, each animal is also part of a species, a food chain and an ecological system. A water molecule is simultaneously part of a hailstone, a cloud and a weather front. The challenge of applying this concept to brain function is not theoretical. It is experimental. How can one observe the activity of a neural mass? How can one determine that what one sees relates to behavioral brain function? The collective activity must be observed as a sum or an average of the activity of members of the population. Like a census taker one must collect data from individuals, make averages over samples and then find the behavioral correlates of the results, which have no meaning unless and until they are related to what the brain is doing.

One must choose between two methods of averaging. A single amplifier suffices for recording a trace of activity from one neuron in a mass or from a local group of neurons acting in concert (Fig. 2). The recorded trace typically lasts merely a few seconds or even a fraction of a second. On repeated stimulation one can collect and store an ensemble of traces over time. If one aligns these stored traces using the stimulus time marker, one can

compute an average trace called an *ensemble time average*, also known as an *averaged evoked potential* or event related response (Basar, 1980). With many electrodes in a spatial matrix one can simultaneously obtain an ensemble of traces each time the stimulus is presented (Fig. 4).

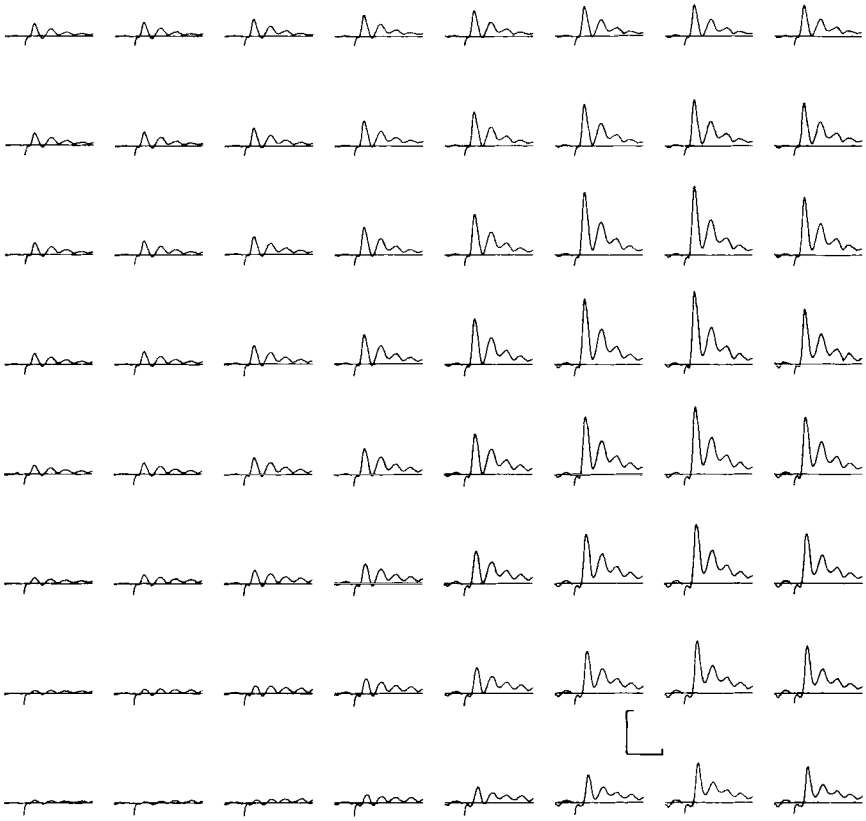


Fig. 4. Each trace is a time ensemble average (N = 100) from an electrode in an 8 x 8 array (4 x 4 mm) on the olfactory bulb, with repetitive electrical stimulation of the primary olfactory nerve (Fig. 3) at 1/s and with simultaneous recording from the 64 electrodes. Time: 100 ms; amplitude: 200 μV (from Freeman, 1975).

However, by using the stimulus time marker to align successive evoked potentials from serial stimuli, one can also compute a spatial ensemble average for the single stimulus presentation on every channel (Fig. 5). In this example it is computed from the collection of averaged evoked potentials instead of from the unaveraged time traces.

Profound differences exist between these two averages. For example, suppose that a chorus sings a musical phrase under the baton of a conductor. The listener hears a spatial ensemble average from the singers, who also hear and respond to each other interactively. Now suppose that the conductor were to ask each member to sing the phrase solo while the sound

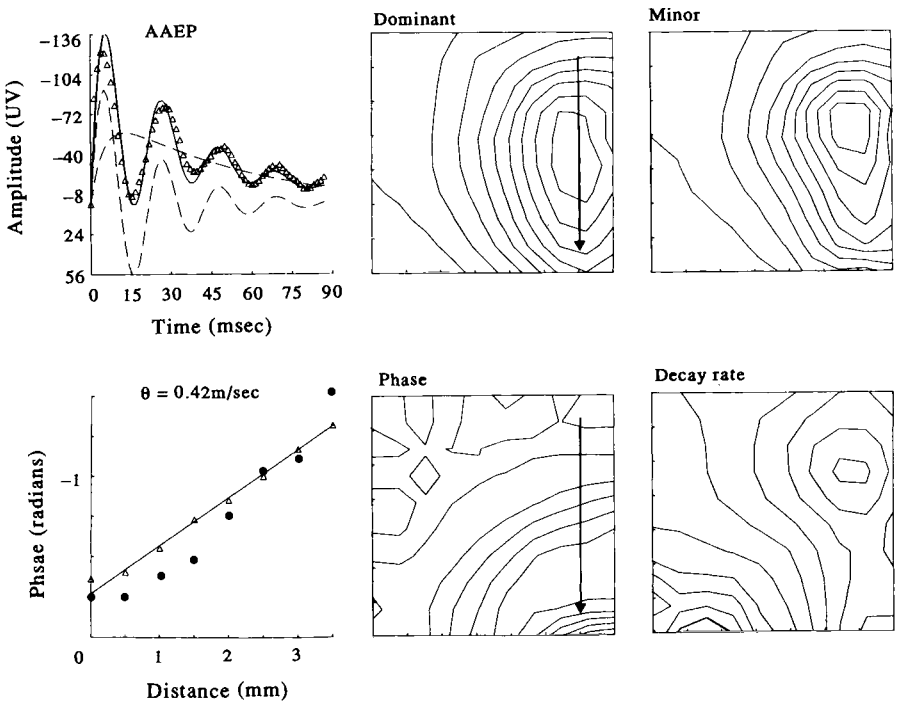


Fig. 5. (a) A spatial ensemble average is formed across the 64 simultaneously recorded and averaged traces (Fig. 4). The fitted curve is the sum of a damped cosine and a baseline shift (dashed curves). (b) The contour plot of the amplitude of the cosine is derived by second order extrapolation from the 64 values. The arrow shows the direction of propagation of the action potentials in the primary olfactory nerve. (c) The spatial pattern of the amplitude of the baseline shift is shown. The locations of both components are dependent on the location of the stimulus, owing to the topographic mapping by the nerve. (d) The frequency of the cosine is the same for all 64 traces. The phase is monotonically related to distance along the nerve as shown by the arrow in (b). (e) The contours show the spatial pattern of phase along the direction of input propagation. (f) The decay rate of the cosine is proportional to the initial amplitude (from Freeman, 1975).

is taped, using a tap of the baton to start each singer. The conductor could make a time ensemble average by aligning the tape segments at the tap and summing the sounds to reconstitute the chorale. The reason this method cannot give the same result as simultaneous singing is that the singers do not have the necessary precision of pitch and timing unless they can interact to form a population as they sing.

Time ensemble averaging works well for neuronal events that are solidly time-locked to stimuli, but it does not work on population events, because each collective cortical event is internally generated. The onset time of each event is related but not bound to the stimulus and variable in latency. Endogenous events typically consist of brief fluctuations that vary in their frequency both within particular events and over successive events (Fig. 2). These variations cause destructive time ensemble averaging. A *spatial ensemble average* of a collection of traces that are recorded simultaneously, in contrast, tends to wash out local detail but to emphasize the common wave form of the population.

Time ensemble averaging can be used to display the impulse response of cortex to its sensory input and spatial ensemble averaging can be used to derive an estimate of its perceptual output. This is shown by comparing two forms of electrical field potentials recorded from the olfactory bulb. First, an electrical stimulus delivered to the afferent pathway leads to a characteristic damped oscillation. This can be recorded with an *array of 64 electrodes.* When the stimulus is low intensity, the response of the bulb is an oscillation at 40–60 Hz like the ringing of a bell when it is struck (Fig. 4). The response is restricted to a part of the bulb to which the stimulated axons carry the input. It has the same frequency of oscillation everywhere over its duration and its amplitude pattern reflects accurately the spatial pattern of the input (Freeman, 1991). In other words, this is a stimulus-bound response of the bulb. Because the oscillation occurs only when the input intensity is low, it is obscured by the ongoing EEG, so it must be extracted by repeated stimulation and time ensemble averaging at each of the 64 channels. Measurement of the spatial pattern of phase of the 64 damped cosines reveals a time lag of the oscillation corresponding to the direction and speed of propagation of the afferent axons that carry the electrically evoked action potentials (Fig. 5). The contributions of the EEG are removed by this approach, along with the amplitude and phase patterns of the endogenous bursts, because the times of onset of the bursts and their frequencies of oscillation vary unpredictably from one burst to the next without precise relation to the times of onset of the electrical stimulus.

In contrast, the normal input is from olfactory receptors that are activated by inhalation, which induces a burst of oscillation in the EEG (Fig. 2). The burst involves the entire bulb (Bressler, 1984). It is observed through the same array of 64 electrodes placed on the bulbar surface, serving as a 4 × 4 mm window. A *typical burst* with one inhalation has a duration of about 100 ms with 4–6 cycles of oscillation at typical frequencies of 40–60

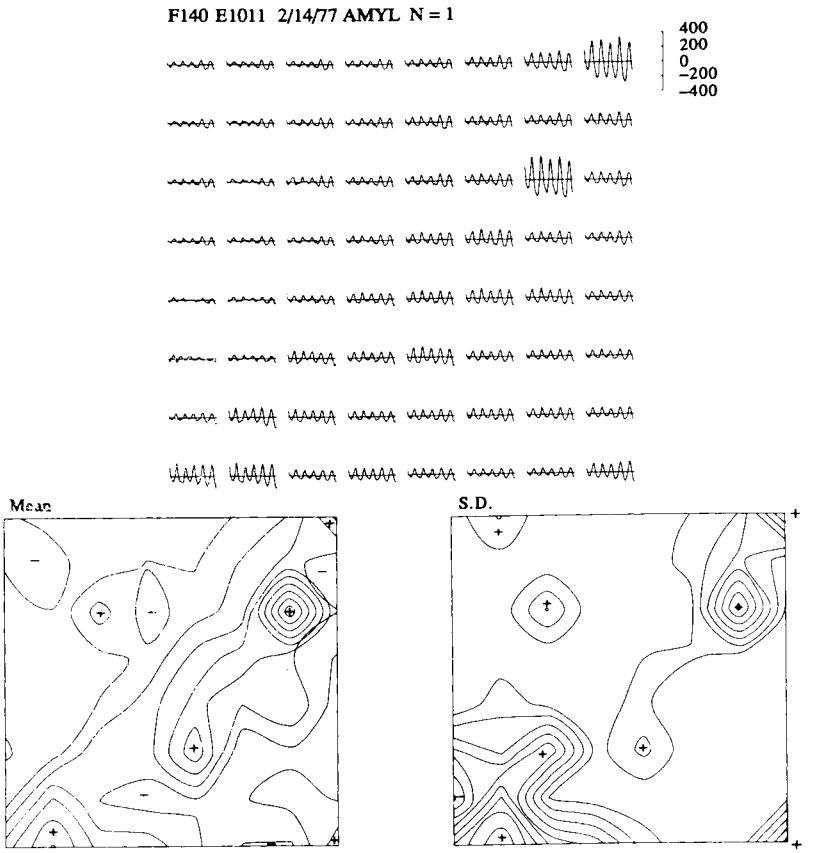


Fig. 6. A single burst as seen in Fig. 2 was recorded simultaneously with 64 electrodes as in Fig. 4. The common wave form at the same instantaneous frequency as determined by spatial ensemble averaging is apparent on all channels in this example without ensemble averaging. The differences in amplitude are displayed in a contour plot for the mean of 10 bursts, giving the spatial amplitude modulation (AM) pattern. The variations between the spatial AM patterns of successive bursts are summarized by the plot of the standard deviations (S.D.) for the 64 channels. The scale is in microvolts. (From Freeman and Schneider, 1982.)

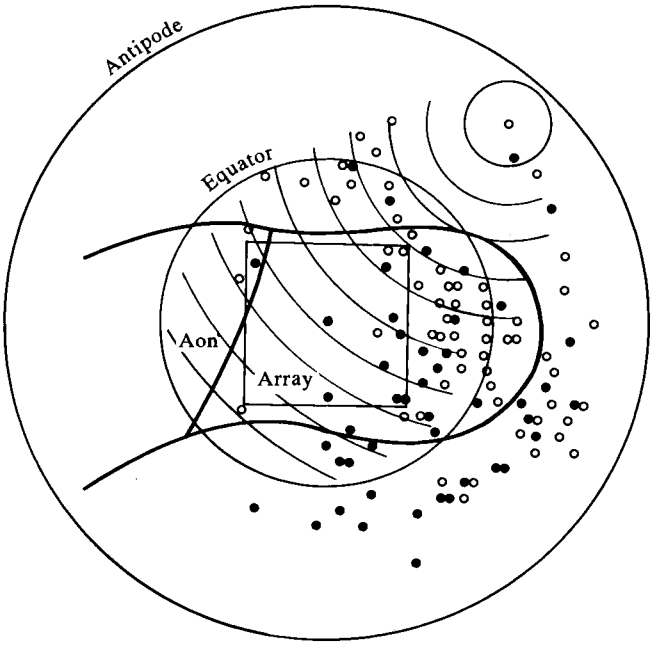


Fig. 7. The spatial pattern of phase for each bulbar burst takes the form of a cone in spherical coordinates. The square shows the outline of the array superimposed on the lateral view of the rabbit olfactory bulb, the surface of which has been flattened into a circle. The location of the apex of the cone for each burst is shown by a dot (maximum phase lead) or small circle (minimum phase lag). The concentric circles show an example of the phase portrait for one burst with a contour interval of 0.1 radians/mm. The location and sign of the apex were not related to behavior or to the spatial AM pattern and the phase was not related to the directions of propagation of the nerve (Fig. 5) or the tract (Fig. 3). These results show that temporal ensemble averaging (Fig. 4) can be used to retrieve the driven response of cortex to input and that spatial ensemble averaging (Fig. 6) can be used to retrieve the endogenously generated output of cortex as its self-organized response to the stimulus. (From Freeman and Baird, 1987.)

Hz. There is a common wave form distributed over the entire bulb, which is derived by statistical analysis and it is found to have a *spatial pattern of amplitude modulation* (AM) (Fig. 6). Furthermore, perceptual information is carried in this spatial AM pattern.

This burst is not evoked by the input. It is induced, because the input makes the bulbar populations unstable. The instability is manifested when the bulb jumps to a new state and creates a new spatial AM pattern of activity. The evidence for this state change is found in the spatial pattern of the *phase* of the common oscillation. This has the form of a cone, which implies that each transition starts at one point, called a site of nucleation and spreads uniformly in all directions at rates near 2 m/s, the velocity of axon collaterals in the bulb (Freeman and Baird, 1987). The location of the apex of the cone is a random variable between successive bursts and can be found anywhere on the bulbar surface unrelated to the odorant stimuli that select the spatial AM pattern (Fig. 7).

The polarity of the apex (maximal phase lead or lag) is also a random variable, so that the point cannot serve to "locate" a "pacemaker" neuron. Hence the phase pattern proves that the bursts are endogenous or "self-organized".

Coding by neuronal populations

There is a yet more profound aspect of spatial ensemble averaging to be considered. Spatial integration is the mechanism by which neurons in a local neighborhood corresponding to a *cortical "column"* make their own spatial ensemble average and thereby establish the basis for *cooperative collective action.* Spatial integration occurs continually in each neuron as its dendrites converge and sum the waves of current that are generated by its synapses upon axonal pulse input. It also occurs as each neuron transmits to many others in the surround and then receives their transmissions back again. In many if not all areas of cortex the excitatory neurons excite each other (but not themselves) and the cooperative interaction is the basis for the emergence of the local mean field and the population neighborhood. Receptors cannot form a population in this sense because they are not synaptically connected and they are not globally interactive in other ways. Spatial integration also occurs between cortical areas, as for example in the divergence of the axon projection from the bulb to the *olfactory cortex,* which is important for the readout of the population, but divergence alone and parallel action alone are not sufficient. There must be feedback, as indicated by feedback within the transmitting population and by the "centrifugal" connections (Fig. 1) from the target populations back to the transmitter population.

The macroscopic activity of the population is difficult to observe in the pulse activity of the individual neurons, because a large amount of

random variation exists in the intervals between pulses in the *"spontaneous" background activity* of each cell. From some indirect measures it is estimated that in the cortex only one part in 1000 to 10,000 of the total variance in the activity of the individual neurons in the population is covariant with the activity of the population. Because time ensemble averaging destroys the endogenous components, the minimum for detecting population activity would require simultaneous recording from 1000 to 10,000 cells. At present the experimental limit for recording "units" simultaneously is from about 100 individual cells.

Nature, however, provides easy access to cortical spatial ensemble averages. This comes about because the same dendritic currents that determine the firing rates of single neurons at trigger zones flow across the tissue resistance and their ohmic potentials add as the currents pass through the tissue (Fig. 8). When the extracellular field potential is properly recorded, it provides a basis to estimate the strength of the local mean field activity in the neighborhood of the neuronal population from the amplitude of its EEG.

Multiple simultaneously recorded EEGs from an *array of electrodes* placed on the cortical surface at suitable spacings provide a basis for observing the spatial AM patterns of cortical activity (Fig. 6). Other methods currently under development include measurement of cortical electric fields with *voltage-sensitive dyes* and *optical recorders,* (T'so *et al.*, 1990) or with *magnetic sensors* to measure the magnetic components of the fields of dendritic current (Williamson and Kaufmann, 1989; Llinas *et al.*, 1991). At present their instrumental noise levels are so high that they cannot be used without time ensemble averaging or narrow band pass filtering, so the use of these methods is restricted to measuring time ensemble averaged evoked potentials and not the unaveraged traces.

Simultaneous recordings of EEGs from 60 to 64 electrodes in rectangular arrays placed surgically on the olfactory bulb or cortex reveal a substantial degree of *spatial coherence* of the activity at all times. In contrast, the time series recorded at each electrode is highly irregular. The wave form from each electrode is as nonreproducible as a freehand scrawl, but the simultaneous recordings always contain the same scrawl (Fig. 6). To be sure, there are local differences in the amplitudes and peak latencies (Fig. 7) of the scrawl, but the same peaks and troughs occur everywhere and the instantaneous frequency of the EEG traces tends to be the same.

The commonality is often difficult to see in the visual, auditory and olfactory cortices, because the EEG reflects the sum of electric currents from multiple sources, many of them unknown and the common wave form comprises only half to three quarters of the total variance of the activity in the traces (Freeman and Viana di Prisco, 1986; Freeman and Grajski, 1987). The common wave form is extracted by computing a spatial ensemble average from each set of EEG traces, in much the same way that the cortex extracts the bulbar output. This commonality of wave form extends over the

entire bulb. Similarly other common wave forms cover the entire *olfactory cortex* (Bressler, 1987) and much if not all of the *primary visual* (Freeman and van Dijk, 1987) and *auditory* (Pantev *et al.*, 1991) *cortices.* In contrast to lower frequencies such as the α (8–12 Hz) and θ (3–7 Hz) *rhythms* these time series appear to be self-organized within the cortices and not imposed by pacemakers lying in the basal ganglia or *brainstem.*

What makes these spatial AM patterns interesting is the fact that they contain behavioral information as demonstrated experimentally by conditioning. After an animal is trained to discriminate and respond to an odorant (Fig. 9), a unique spatial AM pattern reappears in the EEG of the bulb whenever the animal inhales that odorant.

If it inhales another odorant that it can discriminate, a different reproducible pattern emerges. The common oscillatory activity constitutes a *"carrier" wave* and the information about an odorant is expressed by the spatial AM pattern of the carrier in the *wave packet* independently of its wave forms, frequencies, or latency patterns.

These findings suggest that *olfactory coding* resembles the frames in a movie film. The light is a fluctuating energy or carrier (wave form) that bursts on and off with each frame. The information in each frame (sniff) is given by the intensity of the light (amplitude of the wave) at each pixel (local neighborhood) and it is held for the duration of the frame. If the same scene is shown in sequential frames, the spatial pattern of the carrier amplitude is reproduced in each frame. If a new scene appears, a new spatial pattern is seen. This analogy may suggest the intermittent flow of perceptual information in the olfactory system by a process known as *chaotic itinerancy* (Tsuda, 1991).

To recapitulate, microscopic coding is found in its clearest form in the peripheral nervous system and in the sensory and motor relays to and from the cerebral cortex. The information is carried in the pulse trains of a selected subset of neurons. Observers extract its statistical time series analysis of point processes using time ensemble averaging on the recorded data, because each pulse sharply defines the time and place of its piece of information in terms of the time interval from the last preceding pulse at that trigger zone. The pulse frequency gives the intensity of the stimulus and the origin of the axon determines the quality of the stimulus. Stimulus parameters are modified by cellular mechanisms that control receptor sensitivity, but they are not changed by associative learning. Macroscopic coding is found most clearly in cerebral cortex, co-existing with the microscopic activity that is injected into the cortex by afferent axons. This information is spatially distributed over the entire population comprising an area of cortex, but it is resolved into time segments. Within those segments it is stationary in the form of an amplitude modulation of a common carrier. The patterns are self-organized, not topographically related to input and strongly dependent on modifications by near and remote experience that has been incorporated into the cortex by learning.

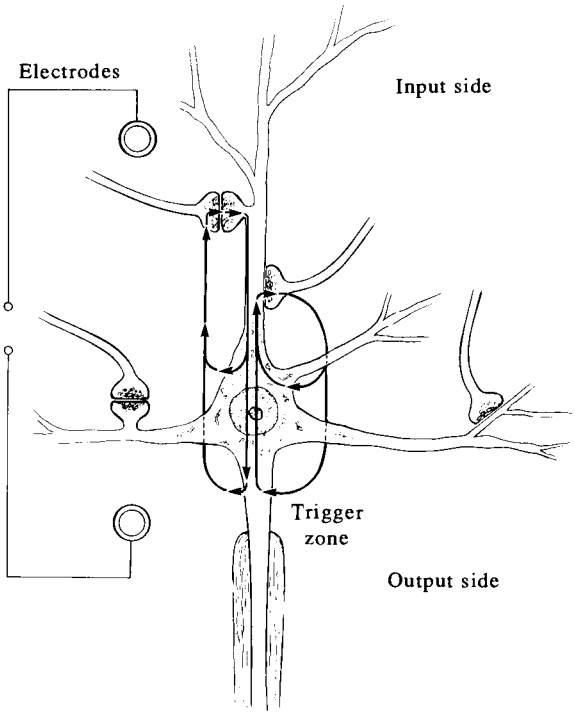


Fig. 8. A typical neuron is shown in outline form with its branching arbor of apical dendrites above, its cell body and nucleus in the middle and its trigger zone and myelinated axon below. The basal dendritic arbors extend to both sides. Two examples of excitatory synapses ending on spines are shown on the left, one active and giving rise to loop current. Two examples of inhibitory synapses are shown on the left, one active and also giving rise to loop current. The directions of the currents are opposed, that for excitation crossing the trigger zone membrane outwardly, increasing pulse probability and that for inhibition crossing inwardly, which reduces pulse probability. The sum of the inner limb of the loop currents from synapses to trigger zone determines the synaptic potential of the single neuron, a microscopic variable. The outer loop currents sum in the extracellular space over the neurons in the neighborhood of the population, giving the local mean field potential and EEG, a macroscopic variable. (See Freeman, 1991.)

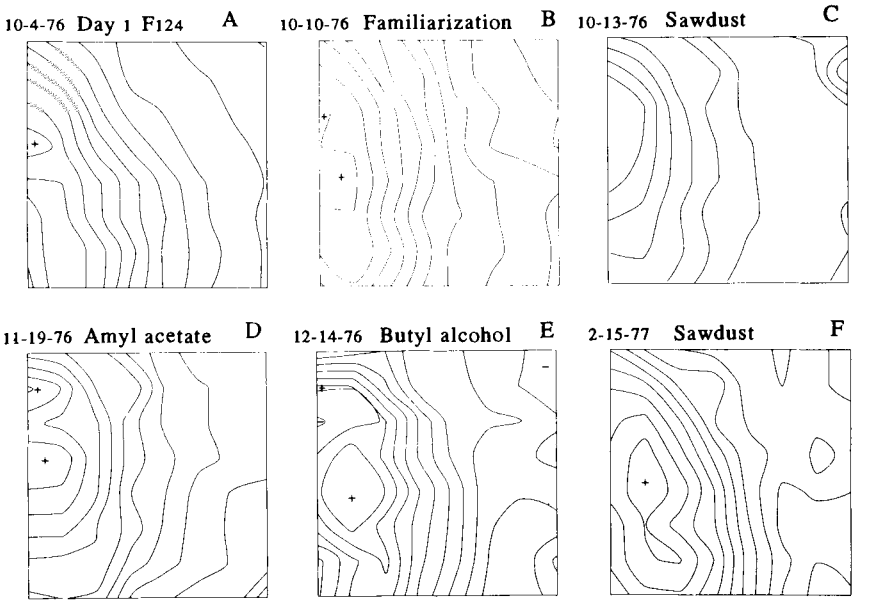


Fig. 9. The bulbar spatial AM pattern changes with familiarization to the odorant delivery device and thereafter with each new discriminated odorant or change in the reinforcement contingency (Viana di Prisco and Freeman, 1985; Freeman and Grajski, 1987). After serial conditioning (Freeman and Davis, 1990) the return to a previously learned odorant gives a new pattern, not the old one ("sawdust"). These and related results (Freeman and Skarda, 1985) show that the spatial AM patterns manifest the meaning of the learned stimuli, not the details of the stimuli *per se*, suggesting that meaning is created by the neural events that constitute burst formation. (From Freeman and Schneider, 1982.)

Learning and pattern recognition by populations

Continuing with olfaction, the EEG recordings reveal changes in the spatial AM patterns of the carrier wave. The changes imply that synaptic strengths are modified with *associative learning*. How and where do the changes with learning take place? The answer is found by training an animal to discriminate an electric stimulus given to an olfactory pathway and measuring the monosynaptically initiated impulse responses of the *bulb* and *olfactory cortex* (Freeman, 1975). The averaged evoked potentials (similar to those in Fig. 4) do not increase in amplitude but show greater durations in the first upward peak and in the subsequent oscillations. They are less stable in the oscillatory mode. This pattern can only occur when some of the synapses that interconnect excitatory neurons within the olfactory bulb and cortex are selectively strengthened (Fig. 3). The increase in synaptic strength with learning does not, as is commonly supposed, take place at the synapses between the incoming axons and the excitatory neurons, because the initial amplitude does not increase. It appears instead that, in accordance with a *variant* of the *Hebb rule*, which holds that synapses between neurons that fire together are strengthened, the synapses that are augmented are only those between bulbar neurons that are simultaneously excited by the training stimulus. The synaptic changes with behavior only occur when the training stimulus is an odorant and when it is accompanied by a reinforcing stimulus that is rewarding or punishing. The data show that over a sequence of a few dozen sniffs on as few as four or five trials a *"nerve cell assembly"* forms for a particular odorant as the animal learns to discriminate it. The assembly consists of a small subset of cells that were co-excited on one or more learning trials on which reinforcement was given. This combination of conditions strengthens the synaptic connections between them (Gray *et al.*, 1986). When any subset of neurons in this assembly thereafter receives input, the entire assembly is excited by the enhanced mutual excitation. These principles have been embodied in sets of differential equations, the solutions for which give spatial AM patterns of chaotic carrier wave forms after modification of connection strengths by learning in accordance with a form of the Hebb rule (Freeman, 1987; Yao and Freeman, 1990).

Studies of the EEGs from these animals and in particular the *phase patterns* (Fig. 7) show that with each onset and offset of a burst there is a dramatic change or *state transition* of the bulbar population, consisting of an abrupt change from one global activity pattern of the neural population to another one and from one set of properties to another set. The change is induced by the input and the new pattern is held or "clamped" until the input subsides.

The concept of a global state transition in the brain is familiar from behavior. Both sleeping and waking are global brain states with abrupt transitions between them. During standing, walking and running the bones,

muscles and neurons are the same, but their global patterns of periodic activity are susceptible to sudden changes. The analysis of EEGs is complicated by the fact that the carrier is irregular in its wave form and it never repeats itself, whereas those state transitions that are best understood are typically from one steady state to another, like water on a pond freezing or melting, or a muscle from relaxation to contraction. Another well known transition is from a steady state to rhythmic oscillation known as the *Hopf bifurcation.* This concept can be used to describe the change in an animal at rest that starts to walk or to swim.

These dynamic modes form a well known hierarchy of stable states. The stability of a state of a system is determined by perturbing the system by brief input such as a pulse to determine whether it returns to the initial state. That state is characterized as an *"attractor"* owing to its perceived tendency to draw the system to itself and the set of inputs for which this tendency holds is called the *"basin" of attraction,* in analogy to the tendency of an object to move downhill to rest. The simplest attractor is a point, which represents zero change. This holds for single neurons at rest, for isolated slices of cortex and for brains under very deep anesthesia or in legally defined "brain death". The next is the *limit cycle attractor* characterized by a single frequency to which the system returns in periodic oscillation after perturbation. This holds for single neurons that are firing pulses at a constant frequency, for populations in rhythmic activity such as in locomotion and to an adequate approximation for cortex in respect to its capacity for generating regular EEG oscillations in the alpha and theta ranges. The more complex *quasi-periodic attractor* has two or more discrete frequencies, for which the spectrum consists of a spike at each frequency.

The steady states that aptly characterize normal cortex differ from these three attractors in that the activity is not periodic. There is no one frequency or discrete number of frequencies into which the activity can be decomposed. The single neuron has a pulse train for which the interval histogram conforms to the Poisson distribution. The *autocorrelation function* has no sustained periodic oscillation (Fig. 10, upper frame).

Yet the statistics of the pulse train are time-invariant whether or not the neuron has been transiently perturbed by external input and is observed after the transient (the evoked activity) has died out and the conditional probability of firing pulses on the local mean field of the population in which the neuron is embedded oscillates in the same manner as the EEG of the population (Fig. 10, lower frame). The *EEG* may often seem to be close to periodic, but its *spectrum* is broadly distributed about certain preferred peaks and its amplitude histogram is Gaussian (Freeman, 1975). Activity that is broad-spectrum and locally unpredictable at first glance might be characterized as "noise", but the population activity that is manifested in the EEG has spatiotemporal structure and global predictability. The structure can be observed in the common aperiodic carrier wave that is shared by many millions of neurons comprising an area of cortex, by the *phase gra-*

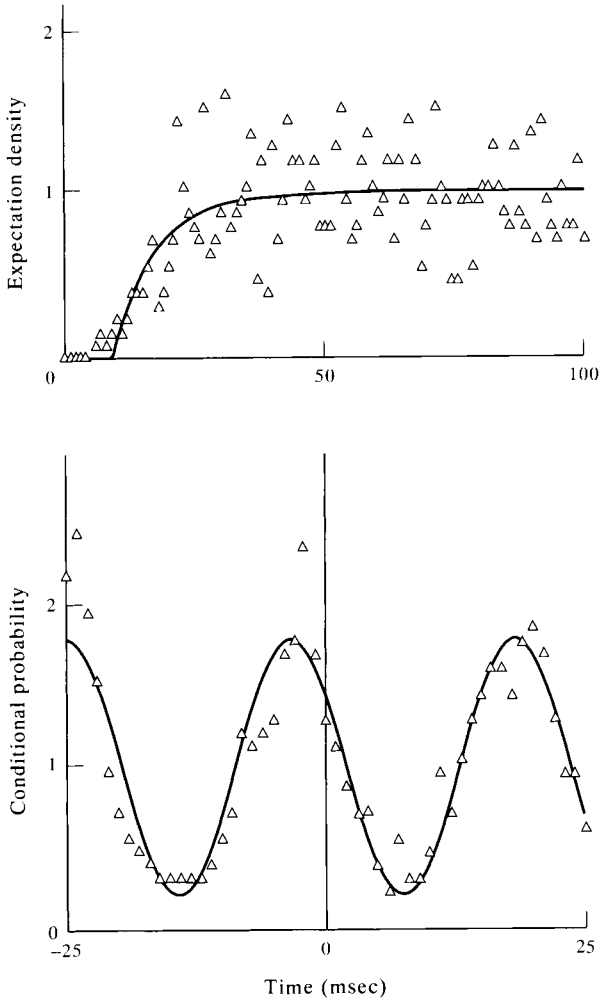


Fig. 10. The microscopic variables of single neurons do not reflect the spatial AM patterns or the common carrier wave forms when viewed singly. The upper frame shows the expectation density (autocorrelation function) for a single neuron spike train, revealing the refractory period and the mean frequency but little else. The lower frame shows the same train when it is reinterpreted as the probability of firing conditional on the amplitude of the simultaneously recorded EEG in the neighborhood. The use of this macroscopic variable as a key serves to make explicit the structure that is implicit in the pulse train. (From Freeman, 1975.)

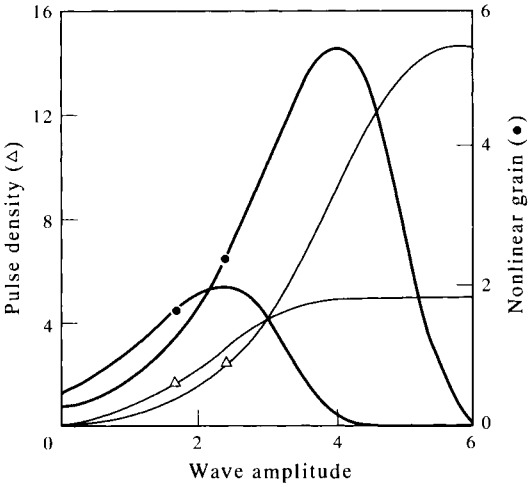


Fig. 11. The conversion of loop current amplitude to pulse density takes place in populations in accordance with a sigmoid function, which is derived from the Hodkgin–Huxley system. In brief, the exponential increase above threshold in firing probability for a single neuron is governed by the voltage dependent sodium conductance and the upper limit on firing rate is determined in the main by the potassium conductance. When expressed at the population level the combined nonlinearity is static. Two curves are shown, one corresponding to a resting or anesthetized state in which bursts do not occur and the other to a highly motivated state that is characterized by bursts with inhalation. The dark curves are the slopes or nonlinear gains of the sigmoid curves. The small symbols denote the values of wave and pulse densities without input. The asymmetry of the sigmoid curve is shown by the fact that the maximal gain is to the excited side of the basal state. It is this property that underlies input-dependent gain increases and state transitions with input into bursts. (From Freeman, 1979.)

dient that precludes the existence of a "pacemaker" for the common activity and by the ability of the cortex to return to a definite spatial pattern in circumstances of receiving an appropriate learned input. These properties clearly indicate that global cortical activity does not consist of unstructured noise, but it is a manifestation of *deterministic chaos* (Stewart and Thompson, 1986; Basar, 1990).

The stability of a chaotic brain state is easily shown by use of perturbation. When an electrical stimulus is given that drives a sensory cortex away from its pre-stimulus basal state, it returns to that state by a characteristic trajectory that is revealed by an *evoked potential* (Fig. 4). The wave form of the evoked potential changes with the parameters of the stimulus (location and intensity, for example) but the state to which the system returns is the same by a variety of statistical assays for a specifiable domain of the parameter values. Therefore the basal state can be said to manifest an attractor and the domain of the input parameters specifies a basin of attraction for that attractor.

More generally, the form of the activity of a sensory cortex as revealed

by its EEG shows striking changes with modification of the behavioral state, as from sleeping to waking or from resting to aroused in hunger, fear, etc. Each identifiable state that can be shown to be stable under perturbation can be said to manifest an attractor for that cortex and the change in dynamics from one attractor to another can be described as a bifurcation. The change implies that at least one parameter within the system has changed and this is called a bifurcation parameter.

An example of a bifurcation parameter is shown in Fig. 11 by the height of the sigmoid curve. Each neuron sums its dendritic currents and converts the instantaneous amplitude to a probability of firing. This probability is governed by the properties of the membrane, mainly the *voltage-dependent sodium conductance*, which determines the concave-upward part of the sigmoid and the *voltage-dependent potassium conductance*, which determines the convex-upward part of the curve. In a local population the output is a continuous pulse density as a function of the *local mean field potential* (Freeman, 1979; Eeckman and Freeman, 1991).

The lower of the two sigmoid curves shows the slope and maximum for an animal at rest. The mean level of dendritic potential and of pulse density is shown by the small triangle. In this state there is basal chaotic activity, but no bursts occur. When the animal is *aroused*, the mean values, the slope and the upper asymptote increase with a single parameter. In this state the basal activity persists and a burst occurs with each inhalation, showing that the system has a different attractor. Bifurcation has occurred. When the animal is given a large dose of an anesthetic, the basal chaotic activity vanishes and the system bifurcates to yet another attractor, a point that is stable under perturbation.

The burst is dependent on the presence of input during inhalation and not on a parameter change in the system, so it does not manifest a bifurcation. It is due to the asymmetry of the sigmoid curve, such that its steepest slope is not at the rest point for zero input (the triangle) but to the right (excitatory) side of the sigmoid. The two heavy curves show the derivatives of the sigmoids, which constitute the nonlinear gains. They show that excitatory input not only increases the output of a population. It also increases the sensitivity and the input–output gain. This feature depends upon the property of the individual neuron, that as it is brought closer to its firing threshold, its probability of firing increases exponentially. Automatically this increases the feedback gain between coupled populations, so that they tend to be destabilized. If the input has an odorant that activates a nerve cell assembly and if the sigmoid curve is sufficiently steep, then the population undergoes a state transition and breaks into an endogenous oscillation, a burst. The spatial *AM pattern* reflects the previous experience embedded in the strengthened synaptic connections of the assembly. When *learning* takes place and a new assembly forms, the change in synaptic strengths constitutes a parametric modification of the system and a bifurcation can be said to have taken place, but accessing the new spatial AM pattern follows and

does not constitute the bifurcation.

The structure of chaotic attractors in perception

The combination of continual and unpredictable local variation with long term constancy is shown by the two sets of *EEG records* in Fig. 12 lasting 5 s and taken 15 s apart. The upper trace (A) is from the *olfactory bulb*, B is from the *anterior nucleus* and C–E are from the *olfactory cortex*.

The bursts occur with inhalations that bring the system out of its low amplitude basal chaotic state. The successive traces are more similar to each other than the traces from the different parts and the three traces from the cortex are more similar to each other than to those from the other parts. The constancy is such that an experienced observer can tell at a glance the source of a trace, but the variability is such that no prediction can be made as to the detailed future time course of any of the traces.

The same combination of variability and constancy is found in the spatial AM patterns. Each pattern is described by the 64 amplitudes of the common wave form and therefore by a 64 × 1 column vector or a point in 64-space. For display purposes a set of points from a number of bursts can be projected into 2-space while preserving the relative distances between them (Sammon, 1969). The variation in pattern is shown by the dispersions of the clusters from two animal subjects and the constancy is shown by the clear separation of the two subjects (Fig. 13, upper frame).

Each animal has its own unique pattern, but the expression of it varies randomly from each burst to the next in the absence of deliberate odorant stimulation, as shown by the nearly symmetric projection of the points for the burst from each animal alone (Fig. 13, lower frame). In Fig. 14 the dimensions have been reduced to 2 by means of step-wise discriminant analysis (Freeman and Grajski, 1987) for a data set from one subject containing bursts in 3 conditions: control, a conditioned stimulus odorant paired with a reward (CS+) and an unreinforced odorant (CS–).

The centroids of the three groups are denoted by the large (0, +, –). The inset table shows the classification of the individual spatial AM patterns with respect to the centroids by a Euclidean distance measure. The variability in pattern is shown by the spread of the individual points and the constancy is reflected in the grouping around the centroids. The rate of correct classification (72%) is typical for these data and is only slightly below the rate of correct performance of the conditioned response by the animal.

Whenever a new odorant is introduced and a new pattern forms for that odorant, all of the other patterns change as well. Global AM pattern changes accompany the switching of reinforcement contingencies (CS+ and CS– reversal), even though the odorants and the responses are the same. When an old CS+ is reintroduced, its spatial AM pattern changes to a new

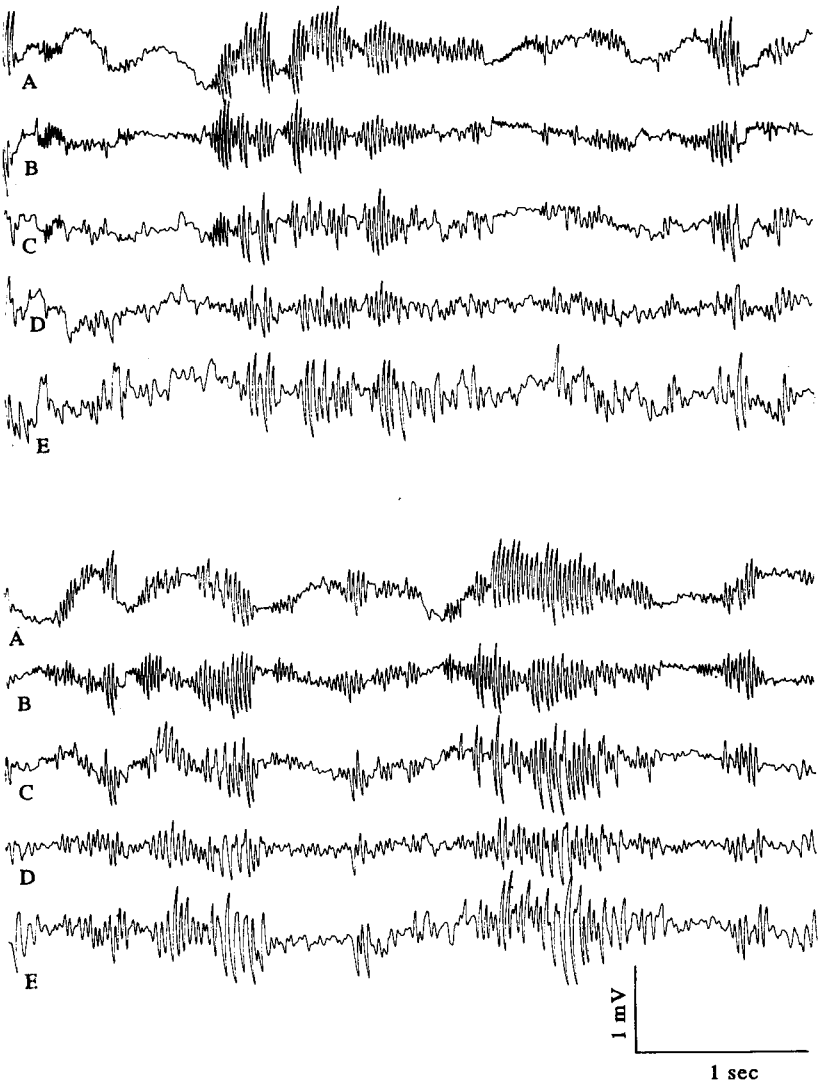


Fig. 12. Electroencephalographic (EEG) traces from the olfactory bulb (A), nucleus (B) and olfactory cortex (C–E) show a low frequency wave driven by respiration and high frequency oscillations called "bursts". The irregularity and local unpredictability of the activity has led many observers to dismiss it as random "background noise". Close study shows that it is the activity of deterministic chaos that carries perceptual information. The long term constancy is suggested by comparing the segments A–E with A′–E′ taken about 15 s later. Such constancy can hold for weeks to months if substantial learning does not take place in the interim. (From Freeman, 1960.)

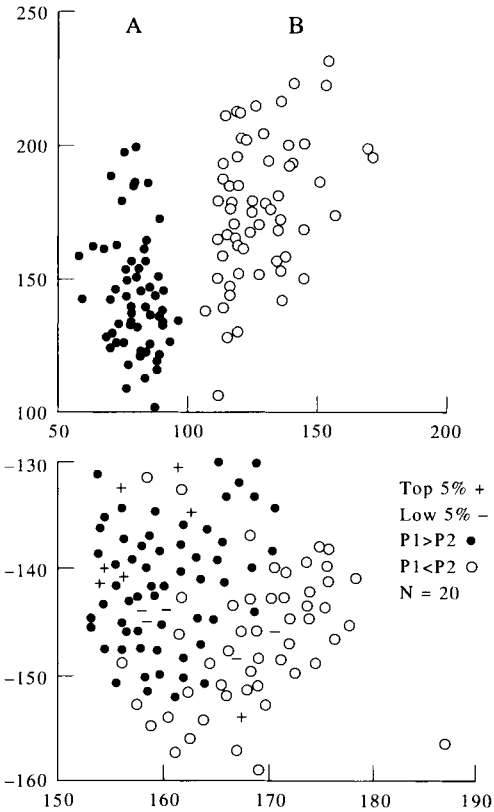


Fig. 13. Each burst has a spatial AM pattern that is expressed by the 64 amplitudes of the common wave form, or as a 64 x 1 column vector, or as a point in 64 space. A collection of bursts can be viewed as a cluster of points by projection into 2-space by a nonlinear mapping that preserves the relative distances within the cluster but optimizes the separation of the clusters of points in the projection (Sammon, 1969). The variation on spatial AM pattern among bursts for each of two rabbits is shown in the upper frame; the constancy is shown by the separation between the clusters for the two rabbits. Each subject has its own characteristic spatial AM pattern. The lower frame shows that the distribution of points conforms to a multivariate Gaussian density, in which local features of bursts (e.g. highest or lowest amplitude) can be readily distinguished. (From Freeman and Viana di Prisco, 1986.)

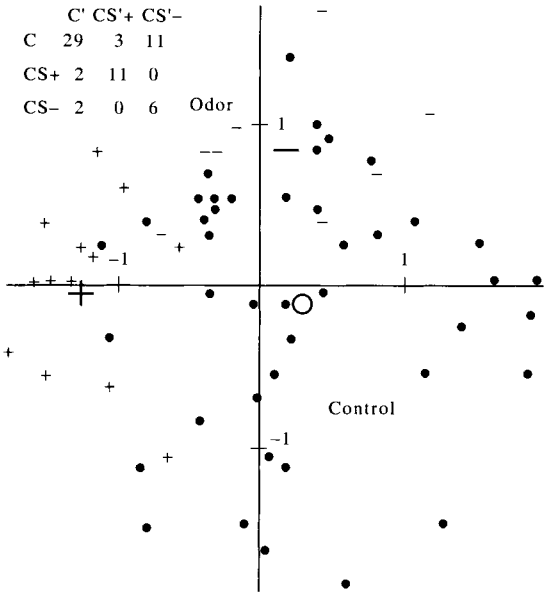


Fig. 14. An alternative method for classification of bursts is to decompose the temporal wave forms as basis functions, measure the loading on each channel of the dominant basis function and by stepwise discriminant analysis determine the projection of the points from 64 space down to the plane determined by two discriminant axes for three clusters. Those three here are the control (C) state prior to an odorant (o), the rewarded odorant (CS+) and an unrewarded odorant (CS–). The table of classification by a Euclidean distance measure is shown in the inset. (From Freeman and Grajski, 1987.)

form, not the old one (Fig. 9). These findings lead to conclusions that the macroscopic AM patterns are not *invariant with respect to stimuli* and that they reflect instead the meaning of the stimuli for the subject. Perhaps most significantly, the patterns cannot be derived from the stimuli by filtering. They are endogenously created by the bulb after its synaptic connections have been modified during learning.

One interpretation of these data is that the neurodynamic machinery of the olfactory system develops and maintains a *global chaotic attractor*, which resembles a *Lorenz attractor* in having *wings*, but instead of two it has multiple wings, one for each class of odorant that it can discriminate. In this view the system when left without input settles into a basal chaotic activity that can be maintained for indefinitely long time periods. When receptors are stimulated by a learned odorant, the system is forced out of the basin of the interburst attractor by the amplitude-dependent nonlinear gain (Fig. 11) of the populations (Freeman, 1979) and is constrained in one wing of the global attractor (Freeman, 1992).

The system does not learn single episodes. It learns to assign inputs into classes. This property reflects the nature of the operation that the olfactory system performs, which is to generalize over its specific inputs. There is a very large number of receptor neurons in the nose, on the order of 10^8 and about 10^5 of these may be capable of responding to any one odorant. On any one inhalation only hundreds or perhaps thousands of receptors actually receive the odorant, but because of turbulence in the nasal passage, the selection of the 10^3 or 10^4 from the 10^5 receptors available is different on every trial. A nerve *cell assembly* can form during training by strengthening synapses among the olfactory bulb excitatory cells that are co-excited by input and it thereafter responds in a stereotypic way when any of its members are excited in any combination. The system is forced out of the basal state by receptor input with a known odorant and is constrained to oscillate in one of the wings of the global attractor. For this reason the oscillation can be visualized for only one wing with each inhalation and then only briefly (Figs 3 and 12).

The geometry in state space of EEG attractors can be visualized by plotting two or more traces of maintained activity against each other or by plotting one trace against itself lagged twice in time as shown by example (Fig. 15). An analogy is to plot on a map of a city those routes that are taken by a salesperson making routine rounds to customers and from home to office and back again.

After several weeks certain preferred trajectories become apparent, in which there is a prominent home-office loop and a loop for each preferred customer. The collection of loops forms a global attractor and the appearance of a certain loop on any given day discloses an input from a preferred customer. The introduction of a new loop (customer) corresponds to a bifurcation with learning a new odor in olfactory dynamics, instituting a structural change in the global system under associative learning.

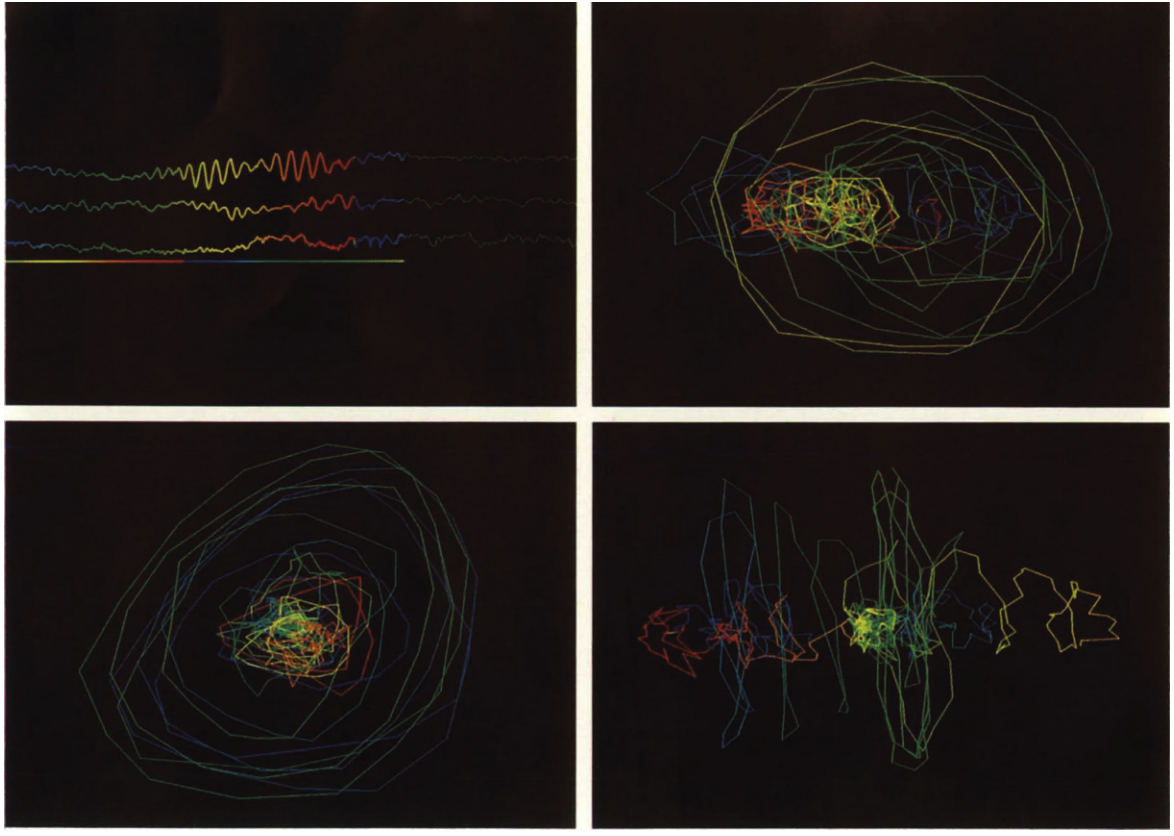


Fig. 15. A chaotic attractor in the olfactory system can be conceived as an object in a dynamic space that is traced or ``filled in'' by recordings of EEGs as a subject explores its environment. These traces were recorded from the bulb after the subject was trained to its olfactory environment by classical conditioning. The brain waves were amplified $\times$ 10,000, digitized at 2 ms intervals, visually edited and transferred to an IRIS workstation for display in three dimensions, using color to display elapsed time, red to blue. The bulbar EEG from a representative site was twice lagged against itself to create a colored object, which was rotated to give perspective views that optimally displayed the characteristics of the attractor, as inferred from behavioral studies, physiological analysis and simulations with solutions of networks of nonlinear ordinary differential equations (Freeman, 1987). This object has a central core of low amplitude activity between bursts and it has a set of wings that have been formed by previous training to identify odorants, one for each learned class. Only one wing can be observed at a time and the occupancy of the wing is revealed in the spatial AM pattern over the duration of the burst plotted in time (upper left). During exhalation the system reverts to the basal core. The core of this high dimensional object can be seen to resemble a ball of yarn when it is projected down into 3-D. The push of the input from odor receptors causes a wing to appear (upper right). The burst can also be viewed from one end, so that the core lies in the center of spiral rings (lower left), or it can be seen from one side, so the rings appear as a coil with the return path in the core (lower right). This example shows a double burst, in which the rings are spread by an input bias in two planes.

This analogy breaks down for brain dynamics, because there is no equivalent of a city map and olfactory dynamic space not restricted to 2 dimensions. At least 3 and preferably 4 dimensions can be used for display of EEG phase portraits with computer graphics resembling a flight simulator to manipulate the structures in 3-D with color to display EEG amplitudes in a fourth dimension. Extended *raw EEG records* give phase portraits that are too complicated to reveal meaningful structure. There are several reasons. One is that the bursts and interbursts are too brief to trace out the forms of chaotic attractors. Successive bursts trace different parts of the wing for the same odorant and different wings for different odorants. Other parts of the brain participate with the olfactory system during perception of an odor, which adds to the complexity of the trajectory. Yet another is that raw EEGs have contributions from multiple sources, many being still undefined, that perturb the trajectories and distort the phase portraits. Finally, the defining properties of a wing are its spatial AM pattern and its basin of attraction with respect to the domain of input and these properties are not readily apparent in displays of 64 simultaneously recorded time series.

In summary, the olfactory system comprising the bulb, nucleus and cortex maintains a global attractor with many wings observed one at a time can be observed. A wing of the attractor in state space looks like a loose coil of wire. By hypothesis an act of perception consists in the explosive leap of the dynamic system from a chaotic basal state to a wing or from one wing to another under the selection of a nerve cell assembly by a stimulus. An act of *associative learning* is a bifurcation, an irreversible structural change in the global attractor that adds a new wing and modifies existing wings.

Readout of percepts

Given the synthesis of a *wave packet* corresponding to a *percept*, the next question is, how is it conveyed to those parts of the brain that will respond with it? The AM patterns are observed in dendritic potentials, but they are conveyed by action potentials of the many neurons from the bulb and widely disseminated over the olfactory cortex. While the commonality of the carrier of the population output is limited to a large fraction of the total macroscopic activity manifested in the EEGs, it consists of a very small fraction of the microscopic neuronal activity in the bulb, perhaps only one part in 1000 to 10,000. How in the next stage is the population output extracted by the receiving neurons under the conditions of an extremely low signal to noise ratio?

One explanation is that because the output of each bulbar neuron diverges widely over the cortex (Fig. 3), each cortical neuron sums input from neurons scattered widely over the bulb, thereby effecting a spatial integration over the bulbar population. The only activity that survives this integration is that which has a common instantaneous frequency, provided fur-

ther that the phase dispersion of the carrier is under a quarter of a cycle of the dominant frequency of oscillation. All other activity averages toward zero. But the common activity is the *carrier.* It is the cooperative output of the population that creates and transports the perceptual information and the measurements of the phase gradient of the carrier show that its spread is under the specified limit to avoid degradation by phase dispersion.

In a word, divergence in the bulbar output pathway "launders" the bulbar message and it can only do so if the carrier oscillates (Freeman, 1991). The divergence will do this optimally if the undesired components are not spatially coherent, that is, at different frequencies at different points in the transmitting cortex. This answer explains the utility of the divergence in the output pathway of the bulb and also of the oscillatory carrier. However, this answer cannot account for the utility of the chaotic carrier, because a limit cycle carrier or a noise carrier would serve as well if it were spatially coherent. Then why is brain activity chaotic? This fundamental question remains unanswered, but there are some directions in which to search for answers. One suggestion is that *chaos* is inevitable in so large and complex a system as the simplest brain with its many interconnections. Engineers and computer scientists already find undesired chaos emergent in their more advanced systems. If this is the source and significance of brain chaos, perhaps the study of brain dynamics can show how to function in chaos.

Yet the evidence suggests that brain chaos is the result of design by evolution and is not an epiphenomenon of hypercomplexity. Brains are intrinsically unstable. Even when animals are at rest and immobile their brains ceaselessly generate temporally unpatterned activity reflected in their EEGs. The chaotic olfactory EEG is not, as is commonly said, "a noise like the roar of a crowd at a ball game", because the EEG is not the smoothed and degraded sum of single cell activities. It is a deterministic output like a complex modern symphony that is orchestrated by the system. It is not a product of the bulb or cortex alone but of these and other parts of the olfactory system in feedback interaction. If they are surgically separated from each other, the activity vanishes (Freeman, 1975).

One role for this mechanism might be to provide unpatterned activity as a form of exercise for neurons, which must keep active or die. In contrast to this view the *"grandmother cell"* hypothesis requires that neurons wait in silence through days or years for their trigger stimuli to appear and then fire faultlessly on feature presentation. This is biologically unrealistic. But a more important postulated role of chaos for brain function is predicated on the fact that chaotic systems not only destroy information; they create it. The essence of a perceptual act may consist of the neurodynamics which creates a neural activity pattern from the raw materials of stimulus, experience and expectancy. The greatest achievement by an act of perception is the immediate creation of the meaning for the subject of some relevant information from a continuous environmental inflow that is complex beyond all measure.

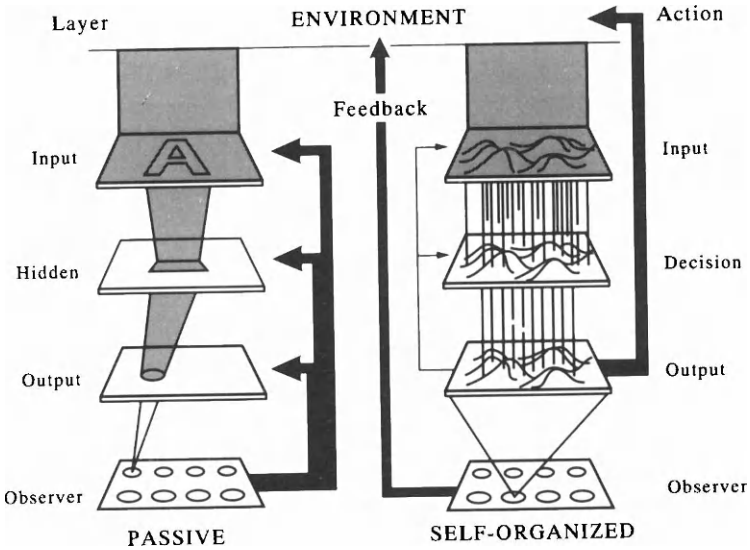


Fig. 16. Man-made systems for extracting and categorizing information from the environment reduce the infinite flow by using filters that are designed by an operator. The brain with its self-organizing chaotic activity makes its own dynamic filters by relying on the intrinsic creative action of its chaotic generators to replace environmental input with low dimensional flows of information shaped by experience and expectation. The activity patterns that are formed by the dynamics are spread out over large areas of cortex, not concentrated at points. Motor outflow is likewise globally distributed, even if the observed movement is only the twitch of a finger or an eyelid. In other words, the pattern categorization does not correspond to the selection of a key on a computer keyboard but to an induction of a global activity pattern.

Man-made information processing systems face this same infinite rate of information flow and they handle the task by use of filters (Fig. 16). The filters define what is "signal" and remove everything else as "noise".

In all cases it is the engineer or observer who decides what is signal and what is noise, either by direct construction of the filter or by use of a reference standard with which to compute an error, so as to shape a filter through a "teaching" procedure as in *back propagation*. Such an agent in a brain would be an homunculus and is not admissible. Brains do not allow direct intervention by outside observers to determine what is feature or signal and what is to be ground or noise. Brains act to organize themselves and to shape the environment and the relations of their sensors to the environment, as in sniffing or moving the eyes. The raw act of perception by hypothesis may be replacement of an infinite environmental inflow with a finite dimensioned chaotic activity pattern. The *fractal dimension* may be determined within the system and may vary according to prevailing circumstances and constraints. The requisite perceptual meaning may be freshly minted under the stimulus conditions and the state of the brain holding at the instant of the state transition of the system away from the basal chaotic attractor into a wing.

Acknowledgements

Thanks for computer graphics are owed to Leslie Kay, Brandon Savage, Bill Baird, Frank Eeckman, Michael Greenspon, Sven Jakubith and Koji Shimoide. The work was supported by grants MH06686 from the National Institute of Mental Health, the Office of Naval Research N63373 and AFOSR 88-0268 from the Air Force Office of Scientific Research.

10

Towards a Kinetic Theory of Cortical-like Neural Fields

FRANCESCO VENTRIGLIA

Neurodynamics Department, Istituto di Cibernetica, C.N.R., Via Toiano 6, 80072 - Arco Felice (NA), Italy

Introduction

From the origin of the mathematical modeling of brain activities *statistical theories of neural fields* had always been highly active. Their principal aim is to explain the dynamics of brain neural processes and global cerebral activities. Continuity and randomness are the main characters of these theories and statistical methods and differential/difference equations are their base. These characteristics differentiate clearly these theories from others, namely those based on the logic approach to neural networks (McCulloch and Pitts, 1943 and later workers), or those which do not consider the effects of the low conduction velocity of spikes in brain, such as the spin glass neural theory of Hopfield (1982). Rashevky (1938) and Rapoport (1948, 1952), who considered random elements in their models, made the early approaches towards a statistical theory of neural fields. Later, Beurle (1957) introduced spatial and continuity aspects into that. In his model the neurons are diffused in space in a continuous way and the description of the neural activity is accomplished by using the spike frequency. An interesting result of Beurle's work is the demonstration of the ability of the model to explain the propagation of activity waves. Subsequently, Griffith (1963) proposed a neural field theory in which he defines the sources of the neural field and a linear operator on it. A field equation, relating the strength of the sources

to the excitation influencing the neurons in distant zones, is formulated. This approach was continued by Fisher (1973). Wilson and Cowan (1973) introduced a more complete spatial model, based on the spike frequencies of excitatory and inhibitory neurons. The time evolution of the spike frequency was described by a couple of integro-differential equations whose numeric solution showed some interesting results concerning the propagation of excitatory and inhibitory waves. Amari (1974) proposed a statistical theory of neural fields in which he studied the phenomenon of the formation of topographic maps. Recently, he introduced a statistical neurodynamic method for the analysis of nonequilibrium behaviors in auto associative memory models (Amari, 1987). In 1974 a statistical theory of neural fields based on a kinetic approach was proposed (Ventriglia, 1974). A similar proposal was later made by Sbitnev (1975). More recent advances in this field can be found in the works of Ingber (1982) and Peretto and Niez (1986); albeit some unrealistic assumptions about the neural activity have been made by these authors.

The *kinetic theory of neural fields* is based fundamentally on the description of electrical properties of single neurons and on some statistical hypotheses about their connecting network. The basis of this theory and its main results are exposed in this chapter. Initially a basilar, minimal set of biological neural features, except those linked to neural plasticity, have been defined. Such characteristics are utilized for the construction, through statistical-physics methods, of a general mathematical model, which does not refer to particular brain regions. In the subsequent sections, some results of the application of the theory to cortical-like neural fields and to attention and learning are discussed. I have further exposed an attempt to extend my theory with elements related to the dynamics of neurotransmitters' activity. There, a set of kinetic equations, describing the activity of complex neural systems having several interacting sub-fields have been formulated.

Neurobiological base

The following two sub-sections contain items that should constitute a necessary base for neural modeling. They state a minimal set of basilar characteristics that all the neurons share independent of their shape (granular, pyramidal, basket, chandelier and so on) and their character (excitatory or inhibitory). These characteristics are related both to properties of single neurons and to connections among them. They can be used in a natural way in a modeling based on statistical-physics. The model, which sustains the kinetic theory is exposed in the following sections. After some preliminary elaboration this model adheres better to the below minimal set (Ventriglia, 1974, 1978, 1980).

Characters of single neurons

Each neuron can be characterized, from an electrical point of view, by the following elements:

• A *membrane potential* as recorded at the axon hillock, whose value changes in time according to neural influences impinging on the neuron or due to spontaneous fluctuations (but consider the warning of Abbott, Chapter 4, this volume).

• A *threshold* level for the *firing* (i.e., the generation of action potentials — hereafter called *impulses*).

• A decay constant for the membrane potential.

• A resting potential, which is taken by the membrane potential in absence of neural influences.

• A *maximum hyperpolarization level* (MHL), that limits the effects of the incoming inhibitions on the membrane potential.

• *Periods of absolute and relative refractoriness,* during which the neuron cannot fire or has great difficulty in firing, respectively.

• Rising times and decay constants for Excitatory Postsynaptic Potentials (EPSPs) and for fast Inhibitory Postsynaptic Potentials (fast IPSPs).

• A rising time and a decay constant for slow IPSPs, whose existence has been revealed in studies on the hippocampus (Thalmann, 1984).

• A significant delay between the arrival of impulses and their effects on the membrane potential at hillock, in neurons with extended dendritic branching.

• Occurrence of not rare *spontaneous EPSPs* in some families of neurons, that produce a background noise.

Elements characterizing the connections among neurons

To a very large extent, the common features of the interactions among neurons in brain can be expressed by the following elements (details of connections in or among specific cerebral regions are not considered).

• The *impulse conduction velocity* is significantly finite, in some cases even it is very low (tens of centimeters per second).

• Also significant is the dispersion of the velocity, that in some cases is very high.[1]

• As a consequence of the previous assumption, topological properties such as the nearness cannot be disregarded.

[1] The necessity of considering such elements in theories of neural systems has been already stressed by Pellionisz and Llinas (1982) with particular reference to velocity dispersion. These assumptions contribute to sharply differentiate theories close to biology from theories which do not consider the effects of assuming a finite conduction velocity.

• Occurrence of not rare *failures of transmission at axonic branches* produces a random lack of effects on neurons connected to those branches (Martin, 1984; Scharfman *et al.*, 1990).

• The axonic branch terminals can be classified in two groups, namely terminals with *short range, random effects* — related to axons (or collateral of axons) which branch near (range of some hundreds of microns) the neuron's soma and do connections in an apparently random way and terminals with *long distance, deterministic effect* —related to axons which abandon the zone of origin (in some cases travel within the white matter) and reach distant zones (separation even by some centimeters) where they branch.[2]

• Existence of a synaptic delay time should also be considered.

• The quantum of excitation (or inhibition) transferred at each *synapse* has a very low value, 60 to 150 μV. (It has been computed that several hundreds of simultaneously active excitatory synapses are necessary for the production of the action potential in a postsynaptic neuron (Martin, 1984).)

• The *divergence* is very high, i.e. each neuron makes a very low (1–10) number of synapses with each connected neuron (Martin, 1984). In general, a low level of re-entry exists in the brain. For instance, Amaral *et al.* (1990) computed that in the hippocampal field CA3, which is considered to be the brain structure with the higher level of *re-entry* , each pyramidal neuron is innervated only by 1.9% of the total pyramidal population.

Kinetic model

The neural kinetic theory is based on the following intuitive premises.

Neural systems are considered as finite regions of the three-dimensional space, filled with two different types of *particles.*

Neurons, characterized by an inner excitation corresponding to the subthreshold membrane potential, constitute a kind of motionless, solid background. They, under some conditions, emit impulses in space.

Impulses, quasi-particles without mass (neuron-interacting), move freely in the neural region. They are absorbed by neurons.

The neuronic characteristics which are explicitly used in the following model are:

(1) Neurons are characterized by their excitatory and inhibitory functions. Their inner excitation e assumes a continuous range of values.

(2) The impulses convey a quantum of excitation i (excitatory or inhibitory) which produces an increasing concentration of related neurotransmitters and hence, an excitation (or inhibition) on the postsynaptic neurons.

[2] Long distance connections are generally conveyed by excitatory (pyramidal) neurons, whereas the inhibition is of local type (short-range, random) (Jones, 1986); the excitation related to collaterals of pyramidal neurons or to axons of not-pyramidal, excitatory neurons is of local type.

The base of the kinetic theory of neural fields rests on the following statistical assumptions, the first of which corresponds to the hypothesis of the *grand canonical ensemble* in Statistical Mechanics.

• *The linkage between the impulses and the axonic structure*, which sustains the movement of the action potentials, *is of a statistical type*. The impulses have velocities distributed according to a *probability distribution function* (PDF) which can depend on the position at which the impulses are emitted.

• The neurons are distributed in space and in inner excitation according to probabilistic distribution functions.

The neurons are characterized by a threshold level for firing (S), a resting level for the inactivity state (e_r), a maximum hyperpolarization level (0), an absolute refractoriness period (τ) and a synaptic delay time (t_0). These values can depend on the type of neuron. Both EPSPs and IPSPs have a null rising time. This last assumption imposes an instantaneous jump of the inner excitation in the postsynaptic neuron reacting to incoming stimuli.

Simple model

In a first, simple model the impulses have the following characteristics. They move along straight constant paths from the origin (a neuron) to the point of absorption (another neuron). They are not self-reacting and are split in two classes, namely impulses coming from external sources and impulses which are *random short-range* — and hence only a local propagation of activity is considered.

In the description of the dynamics of *neurons and impulses interaction* only *unitary absorptions* are taken into account and they obey the following rules.

During the encounters between impulses and neurons some impulses are absorbed by neurons in a random way. As a consequence, since impulse convey a quantum i of excitation, the inner excitation e of the neuron (assumed here to be normalized in the range (0, 1)) suffers a jump of height i towards higher values, for excitatory impulses or a dip of depth i towards lower values, for inhibitory ones.

Then:

if the new value of e is below the threshold S (= 1) and above the level 0, the excitation decays towards the resting level e_r, with decay constant μ;

otherwise,

if $e \geq S$, the emission of λ new impulses by neurons occurs and the inner variable assumes the value of the resting level e_r, after an absolute refractory period (ARP) τ;

or,

if the value of e is below 0, it is fixed at e_h ($0 < e_h < e_r$) (assuming a reflecting boundary for the hyperpolarization).

To proceed formally some notations and definitions are presented.

Def. 1 —$f(\mathbf{r}, \mathbf{v}, t)\, d\mathbf{r}\, d\mathbf{v}$ = Probable number of impulses in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$, whose velocites lie in the element of velocity space $d\mathbf{v}$ around $\mathbf{v}$, at time t.

Def. 1′ —$f^+(\mathbf{r}, \mathbf{v}, t)\, d\mathbf{r}\, d\mathbf{v}$ = As above but the *impulses are excitatory.*

Def. 1″ —$f^-(\mathbf{r}, \mathbf{v}, t)\, d\mathbf{r}\, d\mathbf{v}$ = As above but the *impulses are inhibitory.*

By definition:

$$f(\mathbf{r}, \mathbf{v}, t) = f^+(\mathbf{r}, \mathbf{v}, t) + f^-(\mathbf{r}, \mathbf{v}, t);$$

and

$$a = \frac{f^+(\mathbf{r}, \mathbf{v}, t)}{f(\mathbf{r}, \mathbf{v}, t)}$$

denotes the percentage of excitatory impulses.

Def. 2 — $g(\mathbf{r}, e, t)\, de\, d\mathbf{r}$ = Probable number of neurons in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$, whose excitation lies in the element de around e, at time t.

Def. 3 —$g_d(\mathbf{r}, t)\, d\mathbf{r}$ = Probable number of neurons in the volume element $d\mathbf{r}$ around point $\mathbf{r}$ in refractory state, at time t.

Related to the PDF f are the current and the impulse density:

Def. 4 — $\mathbf{j}(\mathbf{r}, \mathbf{v}, t) = \mathbf{v} f(\mathbf{r}, \mathbf{v}, t)$.

It can be easily shown that:

$\mathbf{j}(\mathbf{r}, \mathbf{v}, t).\ \mathbf{n}\ dS\ d\mathbf{v}\ dt$ = Probable number of impulses in $d\mathbf{v}$ around $\mathbf{v}$, which cross a small area dS with unit normal $\mathbf{n}$ in time dt.

Def. 5 — $\rho(\mathbf{r}, t) = \int d\mathbf{v} f(\mathbf{r}, \mathbf{v}, t)$.

Clearly, $\rho(\mathbf{r}, t)$ represents the total number of impulses in $d\mathbf{r}$ at $\mathbf{r}$ independent of their velocity.

Related to the PDF g are:

the neuron density:

Def. 6

$$\psi(\mathbf{r}) = \int_0^1 g(\mathbf{r}, e, 0)de;$$

the pre firing neuron density:

Def. 7

$$\psi_f(\mathbf{r}, t) = \int_\gamma^1 g(\mathbf{r}, e, t)de$$

where $\gamma = 1 - i$ and $\psi_f(\mathbf{r}, t)$ represents the total number of neurons in $d\mathbf{r}$ at $\mathbf{r}$ in pre firing conditions;

the mean neuronic excitation:

Def. 8

$$\varepsilon(\mathbf{r}, t) = \int_0^1 eg(\mathbf{r}, e, t)de.$$

ϵ ($\mathbf{r}$, t) can be related to one of the parameters which is used to describe the observable neurophysiological activity that is the *EEG*.

The following *principle of conservation of probabilistic mass* must be observed:

$$\forall \mathbf{r}, \forall t \geq 0, \qquad \psi(\mathbf{r}) = \int_0^1 g(\mathbf{r}, e, t)de + g_d(\mathbf{r}t)$$

Impulses inserted into the neural mass by external sources, which are independent of the impulse population present, can be described by the following source distribution function:

Def. 9 — Z($\mathbf{r}$, $\mathbf{v}$, t) d$\mathbf{r}$ d$\mathbf{v}$ dt = Probable number of impulses inserted in d$\mathbf{r}$ at $\mathbf{r}$ and d$\mathbf{v}$ at $\mathbf{v}$ between t and t + dt.

The impulse-neuron encounters can be formally described by using the following definitions.

Def. 10 — σ ($\mathbf{r}$, $\mathbf{v}$) denotes the *macroscopic cross-section* (or the *absorption coefficient*) of neurons in d$\mathbf{r}$ at $\mathbf{r}$ for impulses with velocity in d$\mathbf{v}$ at $\mathbf{v}$.

Accordingly, the rate of impulse-neuron encounters for impulse of velocity $\mathbf{v}$ at position $\mathbf{r}$ is given by:

Encounters rate = $v\ \sigma$ ($\mathbf{r}$, $\mathbf{v}$) ψ ($\mathbf{r}$) f($\mathbf{r}$, $\mathbf{v}$, t) d$\mathbf{r}$ d$\mathbf{v}$

Encounters rate for each neuron = $v\ \sigma$ ($\mathbf{r}$, $\mathbf{v}$) f($\mathbf{r}$, $\mathbf{v}$, t) d$\mathbf{r}$ d$\mathbf{v}$

Encounters rate for each neuron independently from the impulse velocity

$$h(\mathbf{r}, t)d\mathbf{r} = d\mathbf{r} \int v\sigma(\mathbf{r}, \mathbf{v})f(\mathbf{r}, \mathbf{v}, t - t_0)d\mathbf{v}$$

where v denotes the module of the velocity, i.e. $v = |\mathbf{v}|$.

Hence, the probable number of firing neurons in d$\mathbf{r}$ around the point $\mathbf{r}$, at time t is given by:

$$N_f(\mathbf{r}, t) = a \int_\gamma^1 g(\mathbf{r}, e, t)de \int dv'\sigma(\mathbf{r}, \mathbf{v}')v'f(\mathbf{r}, \mathbf{v}', t - t_0) \qquad (1)$$

It should be noted the term $t - t_0$ is due to the synaptic delay.

Def. 11 — λ ($\mathbf{r}$) = the average number of impulses produced at $\mathbf{r}$ by a neuron in firing conditions or emission coefficient.

Def. 12 — f*($\mathbf{r}$,$\mathbf{v}$) d$\mathbf{v}$ = Normalized distribution of impulses with velocity in d$\mathbf{v}$ at $\mathbf{v}$, produced in d$\mathbf{r}$ at $\mathbf{r}$ by neurons in the firing state.

Def. 13 — N_f ($\mathbf{r}$, t) λ ($\mathbf{r}$) f*($\mathbf{r}$, $\mathbf{v}$) d$\mathbf{r}$ d$\mathbf{v}$ dt = Probable number of impulses emitted into d$\mathbf{r}$ at $\mathbf{r}$ and d$\mathbf{v}$ at $\mathbf{v}$ in time dt by neurons in the pre firing state at time t , due to absorptions (encounters) suffered by impulses of all velocities.

Let us note that whereas, the excitatory or inhibitory character of neurons has not been considered in the model description, the character of impulses has been dealt with by stating the percentage of excitatory (and inhibitory) impulses. In such a way, I have also implicitly assumed that the excitatory (and inhibitory) neurons are homogeneously diffused within the neural population and have the same percentage of impulses.

Kinetic equations

A *kinetic* (evolution) equation for the impulse distribution can be formulated according to the following procedure.

The equation:

$$dI = d\mathbf{v}dt \int_V \frac{\partial f(\mathbf{r}, \mathbf{v}, t}{\partial t} dr \tag{2}$$

denotes the change in time dt of the number of impulses which are located in a small volume V with surface S around the point **r** and have velocity in d**v** at **v**.

We also have, by a balance relation:

dI = – (i) net number of impulses flowing out of S in dt

– (ii) number of impulses suffering absorptions in V in dt

+ (iii) number of impulses of velocity **v** produced in V in dt by firing neurons

+ (iv) number of impulses of velocity produced in V in dt by external sources.

In mathematical terms we have (from Definition 4):

$$(i) d\mathbf{v}dt \int_S \mathbf{j}(\mathbf{r}, \mathbf{v}, t).\mathbf{n}dS \tag{3}$$

where **n** is the outward normal to S. By application of Gauss theorem, the surface integral can be transformed to a volume integral:

$$(i) = d\mathbf{v}dt \int_V \frac{\partial \mathbf{j}(\mathbf{r}, \mathbf{v}, t)}{\partial \mathbf{r}} d\mathbf{r} \tag{4}$$

Moreover:

$$(ii) = d\mathbf{v}dt \int_V v\sigma(\mathbf{r}, \mathbf{v}), \psi(\mathbf{r}) f(\mathbf{r}, \mathbf{v}, t) d\mathbf{r} \tag{5}$$

and from Definition (13) we obtain:

$$(iii) = d\mathbf{v}dt \int_V \lambda(\mathbf{r}) f^*(\mathbf{r}, \mathbf{v}) a \int_\gamma^1 g(\mathbf{r}, e, t) de \int dv' \sigma(\mathbf{r}, \mathbf{v}) v' f(\mathbf{r}, \mathbf{v}', t - t_0) d\mathbf{r}. \tag{6}$$

Finally, Definition (9) gives:

$$(iv) = d\mathbf{v}dt \int_V \mathbf{Z}(\mathbf{r}, \mathbf{v}, t)d\mathbf{r} \tag{7}$$

Then we can obtain:

$$\begin{aligned} d\mathbf{v}dt \int_V \Bigg[\frac{\partial f(\mathbf{r}, \mathbf{v}, t)}{\partial t} + \frac{\partial \mathbf{j}(\mathbf{r}, \mathbf{v}, t)}{\partial \mathbf{r}} + v\sigma(\mathbf{r}, \mathbf{v})\psi(\mathbf{r})f(\mathbf{r}, \mathbf{v}, t) - \\ \int v'\sigma(\mathbf{r}, \mathbf{v}')\psi_f(\mathbf{r}, t)f(\mathbf{r}, \mathbf{v}, t - t_0)\lambda(\mathbf{r})f^*(\mathbf{r}, \mathbf{v})d\mathbf{v}' - \mathbf{Z}(\mathbf{r}, \mathbf{v}, t) \Bigg] d\mathbf{r} = 0 \end{aligned} \tag{8}$$

Noting that

$$\frac{\partial \mathbf{j}(\mathbf{r}, \mathbf{v}, t)}{\partial \mathbf{r}} = \frac{\partial \mathbf{v}.f(\mathbf{r}, \mathbf{v}, t)}{\partial \mathbf{r}}$$

$$\frac{\partial \mathbf{v}.f(\mathbf{r}, \mathbf{v}, t)}{\partial \mathbf{r}} = \mathbf{v}.\frac{\partial f(\mathbf{r}, \mathbf{v}, t)}{\partial \mathbf{r}}$$

and that the volume is arbitrary, the searched kinetic equation can be formulated as below:

$$\begin{aligned} \frac{\partial f(\mathbf{r}, \mathbf{v}, t)}{\partial t} + \mathbf{v}.\frac{\partial f(\mathbf{r}, \mathbf{v}, t)}{\partial \mathbf{r}} + v\sigma(\mathbf{r}, \mathbf{v}), \Psi(r)f(r, v, t) = \\ \lambda(\mathbf{r})f^*(\mathbf{r}, \mathbf{v})a \int_\gamma^1 g(\mathbf{r}, e, t)de \\ \int d\mathbf{v}'\sigma(\mathbf{r}, \mathbf{v}')v'f(\mathbf{r}, \mathbf{v}', t - t_0) + Z(\mathbf{r}, \mathbf{v}, t) \end{aligned} \tag{9}$$

The *kinetic equation* related to the neuronic excitation distribution g(**r**, e, t) can be formulated in the following way.

At first, we must note that:

$$[g(\mathbf{r}, e + \mu(e_r - e)dt, t + dt) - g(\mathbf{r}, e, t)]d\mathbf{r}de = \left[\frac{\partial g}{\partial t}\right]_{ext} d\mathbf{r}dedt \tag{10}$$

where the term $\left[\frac{\partial g}{\partial t}\right]_{ext}$ denotes the variation of distribution due to impulse absorption.

On dividing by dt and making dt tend to zero, we obtain:

$$\left[\frac{\partial g}{\partial t} + \mu(e_r - e)\frac{\partial g}{\partial e}\right] d\mathbf{r}de = \left[\frac{\partial g}{\partial t}\right]_{ext} d\mathbf{r}de \tag{11}$$

The member on the left hand side denotes the distribution change due to the decay of the neuronic excitation (towards the resting level e_r). By a rule of balancing, we get:

$$\left[\frac{\partial g}{\partial t}\right]_{ext} = \sigma \int d\mathbf{v} v f(\mathbf{r}, \mathbf{v}, t - t_0)(ag(\mathbf{r}, e - i, t) + (1 - a)g(\mathbf{r}, e + i, t) + cg(\mathbf{r}, e, t)) + \delta(e - e_r)N_f(\mathbf{r}, t - \tau) \tag{12}$$

where c , given by:

$$c = \begin{cases} a, & \text{if } e = e_h \\ 1, & \text{otherwise} \end{cases}$$

is related to the effect of reflection that occurs towards the value $e = e_h$ when the inner excitation reaches the MHL. The Dirac function, defined as:

$$\delta(x - x_0) = \begin{cases} 1, & \text{if } x = x_0 \\ 0, & \text{otherwise} \end{cases}$$

is related to a term of gain due to neurons which recover to the state e_r after the ARP. Hence, the neuronic kinetic equation becomes:

$$\left[\frac{\partial g}{\partial t} + \mu(e_r - e)\frac{\partial g}{\partial e}\right] = \sigma \int d\mathbf{v} v f(\mathbf{r}, \mathbf{v}, t - t_0)(ag(\mathbf{r}, e - i, t) + (1 - a)g(\mathbf{r}, e + i, t) + cg(\mathbf{r}, e, t)) + \delta(e - e_r)N_f(\mathbf{r}, t - \tau). \tag{13}$$

The kinetic equation for *refractory neuron distribution* can be obtained easily. Indeed,

$$g_d(\mathbf{r}, t + dt)d\mathbf{r} = g_d(\mathbf{r}, t)d\mathbf{r} + (N_f(\mathbf{r}, t) - N_f(\mathbf{r}, t - \tau))d\mathbf{r}dt. \tag{14}$$

Then, on dividing by dt and making dt tend to zero, we obtain:

$$\frac{\partial g_d}{\partial t}(\mathbf{r}, t) = N_f(\mathbf{r}, t) - N_f(\mathbf{r}, t - \tau). \tag{15}$$

Analytical results and several computations simulating the activity of one- and two-dimensional neural systems, based on the above equations and carried out by appropriate numerical procedure, showed the impossibility to obtain stable and persistent self-sustained activities (Ventriglia, 1978, 1980, 1983). In particular, it was demonstrated that the points (in the space of parameters) in which the emission of impulses balanced the absorption were unstable and the activity either decayed (very rapidly) to zero or increased without limits. Then, for a better description of the neural activity I formulated a more elaborated model in which not only unitary absorptions, but also multiple ones and not only short-range effects, but also long-range ones were considered. The description of this model is the issue of the next section.

Refined kinetic model

In the refined version of the model, the single neurons preserve the previous features, but both their excitatory and inhibitory characters are distinctly described. Moreover, the impulses are split into three *classes* : impulses originating from external sources, *short-range impulses* and *long-range impulses* with assigned target-zone. Thus, the description of the model comes closer to that of cerebral cortex.

The short-range, random impulses are related to collaterals of axons or to axons of interneurons, both excitatory and inhibitory, which remain near the emitting neuron and branch in a random way. They are formally characterized by an emission coefficient λ-related to the density of the axonic branching — which establishes the number of new impulses that each firing neuron must emit. Their velocity is distributed according to a probability distribution function (PDF) f^*.

The long-range, specific impulses are related to axons of excitatory neurons which reach neural regions far from the emitting neuron. They have the following formal characteristics.

(i) A parameter ξ provides for each couple of points $(\mathbf{r}, \mathbf{r}')$ of the neural system an emission coefficient, expressed as an *origin-destination matrix* $\xi(\mathbf{r}, \mathbf{r}', t)$:

$\xi(\mathbf{r}, \mathbf{r}', t) \neq 0$ if the elemental zone $\mathbf{r}$, $d\mathbf{r}$ is coupled to the elemental zone $\mathbf{r}'$, $d\mathbf{r}'$;

$\xi(\mathbf{r}, \mathbf{r}', t) = 0$ otherwise.

(ii) They propagate without absorption in the region between the origin and the destination. In this space the module of the velocity is distributed according a PDF f', to take into account the possible velocity dispersion.

(iii) When they reach the destination zone, they are added to the local impulse population and their velocity is distributed according to a PDF f^o. In this zone they can be absorbed as per the usual rules of local random absorption.

(iv) The parameter ξ can change in time according to suitable *rules of learning*.

To proceed formally, some previous notations and definitions must be modified and yet others must be introduced. In the following, the excitatory or inhibitory character of the impulse/neuron entity is denoted as s and the class is denoted as c.

Def. 14 — $f_s(c, \mathbf{r}, \mathbf{v}, t)\, d\mathbf{r}\, d\mathbf{v}$ = Probable number of impulses of type s and class c in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$, whose velocity lies in the element of velocity space $d\mathbf{v}$ around $\mathbf{v}$, at time t.

Def. 15 — $g_s(\mathbf{r}, e, t)\, de\, d\mathbf{r}$ = Probable number of neurons of type s in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$, whose excitation lies in the element de around e, at time t.

Def. 16 — $g_{ds}(\mathbf{r}, t)\, d\mathbf{r}$ = Probable number of neurons of type s in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$ in refractory state, at time t.

Following are the neuron densities related to the PDF g:
Def. 17

$$\psi_s(\mathbf{r}) = \int_0^S g_s(\mathbf{r}, e, 0)de$$

and

$$\Psi(\mathbf{r}) = \sum_s \psi_s(\mathbf{r})$$

and the mean neuronic excitation:
Def. 18

$$\epsilon_s(\mathbf{r}, t) = \int_0^S e g_s(\mathbf{r}, e, t)de.$$

The principle of conservation of probabilistic mass changes as follows:

$$\forall \mathbf{r}, \forall t \geq 0, \psi_s(\mathbf{r}) = \int_0^S g_s(\mathbf{r}, e, t)de + g_{ds}(\mathbf{r}, t).$$

Now, I shall describe the procedure used to model *high* activity in neural systems. High activity means that in some *infinitesimal* time interval Δt a lot of impulses are flying around in the system and several of them can be absorbed by each neuron (here infinitesimal time interval means a time which is short with respect to some typical neuronal period, such as, for example, the ARP). In such a way, multiple impulse absorptions can occur in these intervals. To describe the dynamics of multiple absorptions, the following two assumptions must be introduced:

(a) In the infinitesimal time interval Δt in which the impulses are absorbed by a neuron, the arrival time of single impulses are not distinguishable.

(b) Each neuron absorbs a whole number of impulses in the time Δt (this is a consequence of the instantaneous jump of the inner excitation).

Assumption (a) allows the following computation of the number of net absorptions for each neuron in an elemental neighbourhood of $\mathbf{r}$ at time t (in a time interval Δt):

$$h_\Delta(\mathbf{r}, t) = \Delta t \sum_c \int d\mathbf{v} \mid \mathbf{v} \mid [\sigma_{+1}(c, \mathbf{r}) f_{+1}(c, \mathbf{r}, \mathbf{v}, t - t_0) - \sigma_{-1}(c, \mathbf{r}) f_{-1}(c, \mathbf{r}, \mathbf{v}, t - t_0)] \tag{16}$$

where $h_\Delta(\mathbf{r}, t)$ is a real number.

According to assumption (b), if the number h_Δ is not an integer, we must redistribute appropriately the absorptions among all the neurons in the elemental neighbourhood of $\mathbf{r}$. I assume that, at time t, all such neurons

absorb n_Δ net impulses (n_Δ is the greatest integer contained in h_Δ), whereas only a percentage d_Δ of neurons absorb an extra impulse.

Here:

$$m_\Delta = int(h_\Delta) \qquad \text{[int(x), denotes the integer of x];}$$

$$n_\Delta = \begin{cases} m_\Delta \text{if} M_\Delta > 0; \\ 0 d_\Delta = \begin{cases} h_\Delta - n_\Delta \text{if} h_\Delta > 0 \\ 0 \text{otherwise;} \end{cases} \end{cases} \quad (40)$$

$$c_\Delta = h_\Delta - int(h_\Delta).$$

Hence, the following probable number of firing neurons can be computed:

$$N_{S\Delta}(\mathbf{r},t) = \int_{S-n_\Delta i}^{S} g_s(\mathbf{r}, e, t)de + d_\Delta \int_{S-(n_\Delta+1)i}^{S-n_\Delta i} g_s(\mathbf{r}, e, t)de \quad (17)$$

As regards the formulation of a kinetic equation for the impulse distribution $f_s(c, \mathbf{r}, \mathbf{v}, t)$, we have:

$$f_s(c, \mathbf{r} + \mathbf{v}\Delta t, \mathbf{v}, t + \Delta t) - f_s(c, \mathbf{v}, \mathbf{r}, t)d\mathbf{r}d\mathbf{v} = \left[\frac{\partial f}{\partial t}\right]_{ext} d\mathbf{r}d\mathbf{v}\Delta t \quad (18)$$

where the left hand side expresses the free streaming of impulses in space and the right hand side is concerned with variations of the distribution function due to absorptions, emissions and external sources. From the usual procedure of balance, the following terms can be obtained:

- a loss term: $\Delta t\ \sigma_s\ \Psi\ |\mathbf{v}|\ f_s$ [impulse absorption]
- three gain terms:

$\lambda_s(\mathbf{r})f^*(\mathbf{v})N_{s\Delta}(\mathbf{r}t)$, [short-range impulse emission]

$f^o(\mathbf{v}) \int_V d\mathbf{r}'\ \xi(\mathbf{r}, \mathbf{r}') \int dv' f'(v' N_{s\Delta}(\mathbf{r}', t\text{-}v^{-1} \mid \mathbf{r}\text{-}\mathbf{r}'|)$, [long-range impulse emission];

$\Delta t Z(\mathbf{r}, \mathbf{v}, t)$, [external source].

Let us note that in the term related to long-range emissions, the integration extends over the volume V occupied by the neural system.

Then, the following impulse kinetic equation can be formulated (the form of finite difference is imposed by difficulties in the formal limit procedure):

$$f_s(c, \mathbf{r} + \Delta t, \mathbf{v}, t + \Delta t) = f_s(c\mathbf{r}, \mathbf{v}, t) - \sigma_s(c, \mathbf{r}) \mid \mathbf{v} \mid \psi f_s(c, \mathbf{v}, \mathbf{r}, t) +$$
$$\Delta t Z(\mathbf{r}, \mathbf{v}, t)\delta(c-1) + \lambda_s(\mathbf{r})f^*(\mathbf{v})N_{s\Delta}(\mathbf{r}, t)\delta(c-2) + f^o(\mathbf{v}) \int_V d\mathbf{r}'\xi(\mathbf{r}, \mathbf{r}') \quad (19)$$
$$\int dv' f'(v')N_{s\Delta}(\mathbf{r}', t - v'^{-1} \mid \mathbf{r}\text{-}\mathbf{r}' \mid)\delta(s-1)\delta(c-3)$$

We must note that the Dirac function $\delta(s-1)$ is a consequence of the assumption that the deterministic long-range effect depends only on excitatory neurons. The other Dirac functions are related to the class c of impulses and select the class for which the affected term is present.

As regards the formulation of the kinetic equation for neuronic inner excitation distribution $g_s(\mathbf{r}, e, t)$, starting from the usual balance equation:

$$[g(\mathbf{r}, e+\mu(e_r-e)dt, t+dt) - g(\mathbf{r}, e, t)]d\mathbf{r}de = \left[\frac{\partial g}{\partial t}\right]_{ext} d\mathbf{r}dedt \qquad (20)$$

we must consider the following loss-gain terms, affecting the right hand side:

- $(\delta(e-e_r)N_{s\Delta}(\mathbf{r}, t-\tau))$, (recovering of refractory neurons to resting potential e_r);
- $[\delta(e-e_h)\int_0^{R(-h_\Delta)} g_s(\mathbf{r}, e, t)de]$, [contributed by neurons whose inner excitation reach the MHL and is reflected to the value e_h]; where R (x) is the ramp function:

$$R(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

- terms related to jumps in the inner excitation occurring as a consequence of the impulse absorption. They are exhausted by the following four cases:

(i) $m_\Delta = 0$ and $c_\Delta = 0$ (no absorptions);
(ii) $m_\Delta = 0$ and $c_\Delta \neq 0$ (partial unit absorptions);
(iii) $m_\Delta \neq 0$ and $c_\Delta = 0$ (multiple integer absorptions);
(iv) $m_\Delta \neq 0$ and $c_\Delta \neq 0$ (multiple partial absorptions).

All these items result as particular cases of the following general equation, which gives the kinetic equation for the neuronic distribution:

$$\begin{aligned} g_s(\mathbf{r}, e+\mu(e-e_r)\Delta t, t+\Delta t) = \delta(e-e_r)N_{s\Delta}(\mathbf{r}, t-\tau) + \\ \delta(e-e_h)\int_0^{R(-h_\Delta)} g_s(\mathbf{r}, e, t)de + (1-\mid c_\Delta \mid)g_s(\mathbf{r}, e-m_\Delta i, t) + \\ \mid c_\Delta \mid g_s(\mathbf{r}e-(m_\Delta + sign(h_\Delta))i, t); \end{aligned} \qquad (21)$$

where *sign* (x) is defined as:

$$sign(x) = \begin{cases} 1, & \text{if, } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

The kinetic equation for refractory neurons is given by:

$$g_{ds}(\mathbf{r}, t+\Delta t) = g_{ds}(\mathbf{r}, t) - N_{s\Delta}(\mathbf{r}, t-\tau) + N_{s\Delta}(\mathbf{r}, t). \qquad (22)$$

Computational experiments

Appropriate parameter definitions, initial and boundary conditions were used in a series of computational *experiments* to investigate the characteristic activities of *two-dimensional* neural models, based on the above kinetic equations. The initial conditions were given by the following formal assignments:

$$\begin{aligned} f_s(\mathbf{r}, \mathbf{v}, 0) &= F_s(\mathbf{r}, \mathbf{v}) \\ g_s(\mathbf{r}, e, 0) &= G_s(\mathbf{r}, e) \\ g_{ds}(\mathbf{r}, 0) &= G_{ds}(\mathbf{r}) \\ Z(\mathbf{r}, \mathbf{v}, 0) &= \zeta(\mathbf{r}, \mathbf{v}). \end{aligned} \tag{23}$$

Short-range connection

In a series of computations, I took into account only short range effects. In such a way the excitability properties of neural systems with connections diffused, more or less homogeneously, throughout the neural mass were studied.

With regards to the impulse distribution, I assumed that an inflow of excitatory impulses was received in a very narrow zone of the neural system at the initial time.

The distribution of the subthreshold neuronic excitation was uniform in space and assumed the form of a Planck function in the inner excitation e. The values attributed to the parameters of the Planck function produced the following neuronic densities: about 3.6 million of excitatory neurons and 0.9 million of inhibitory ones per square centimeter. The ratio between the relative abundance of the two types of neurons agrees with the experimental finding which leads to hypothesize that the 20% of all the neurons in the cortex are inhibitory.

No refractory neurons were present at the initial time and an external source function was assigned. The velocity PDF for the emitted impulses was a Gaussian. An absorbing boundary was assumed for the neural system.

The following values from literature were assigned to the neuronic parameters (Woody, 1982): $\mu = 50\ \mathrm{s}^{-1}$, $\tau = 1.75$ ms, $t_0 = \frac{1}{2}$ ms.

The structural coefficients σ_s, λ_s assumed constant values in velocity; sometimes σ_s depended on the position.

In these computations the activity of a neural system (dimensions: 1.2 $\times$ 1.2 cm^2), having about 5.15 million excitatory neurons and 1.30 million inhibitory neurons grouped in 3600 modules, was simulated. Indeed, the step of the space grid used in the solving numerical procedure was such that each grid point was representative of a small square of neural matter, with the side length of 200 μm, in which about 1400 excitatory neurons and 350 inhibitory neurons are contained. Each module was excited by the

surrounding modules, within the range of a few steps both in excitatory and inhibitory way as well as by external excitatory sources.

The length of the above modules is almost equal to that assumed by some authors (200–300 μm) to be the diameter of a *macrocolumn* (Hubel and Wiesel, 1972; Szentagothai, 1975; Mountcastle, 1978), which forms the hypothetical functional unit of the cerebral cortex (Szentagothai, 1983 and Chapter 1, this volume; Eccles, 1984). The neuron number agrees with the value of 4000 per module (diameter: 300 μm) as hypothesized by Szentagothai (1983).

Several computational experiments have investigated the ability of the simulated system to react to some initial stimuli and to propagate and sustain activity waves (Ventriglia, 1990). A first series of computations demonstrated that the model is capable of responding to an initial stimulation with a space pervasive and stable oscillation in which excitatory and inhibitory effects and effects related to refractoriness are periodically dominant. In a different series of computations, the high inhibition produced the extinction of all the activity in a very short time at the end of the activating stimulation. Other computational experiments have studied the dependence of the propagation on the space variability of structural parameters. In particular, a case was studied in which the threshold S assumed different values in different modules. These values were produced by a random procedure based on a Gaussian distribution. A stable, but random activity was obtained. The ability of the model to generate rotating waves of activity was also studied. A stimulation was applied such that its place of application changed in time in an appropriate way. As a reaction, a phenomenon of spiraling activity (space invasive and stable) was obtained. Several, simultaneous initial stimuli produced multiple spiraling activities, competing for space (the evolution of two spiraling activities is shown in Fig. 1).

Another meaningful feature was the capability of the model to preserve well-developed patterns of activity also in presence of strong disturbances. This ability could be defined as *ultra-stability.* A very strong disturbance was, indeed, applied in a place where an organized collective activity was already settled. The disturbance was unable to modify the space-time characteristics of the developing pattern.

The above results, expressed as space-time maps of some computed macroscopic variables related to the electrical neural activity, could in principle be compared with cortical recordings carried out by multiple microelectrode or with cortical mappings produced by PET (positron emission tomography) and similar equipments.

Kruger *et al.* (1988), for instance, investigated the spiking cortical activity in monkey (striate cortex) by a multimicroelectrode (5 × 6 microelectrode) with an interelectrode distance of 160 μm. This is very close to the space step (200 μm) that I used in my computations. But comparisons were not possible. In fact, the above authors used great care in recording only single cell activities at each microelectrode, whereas my results could be related to

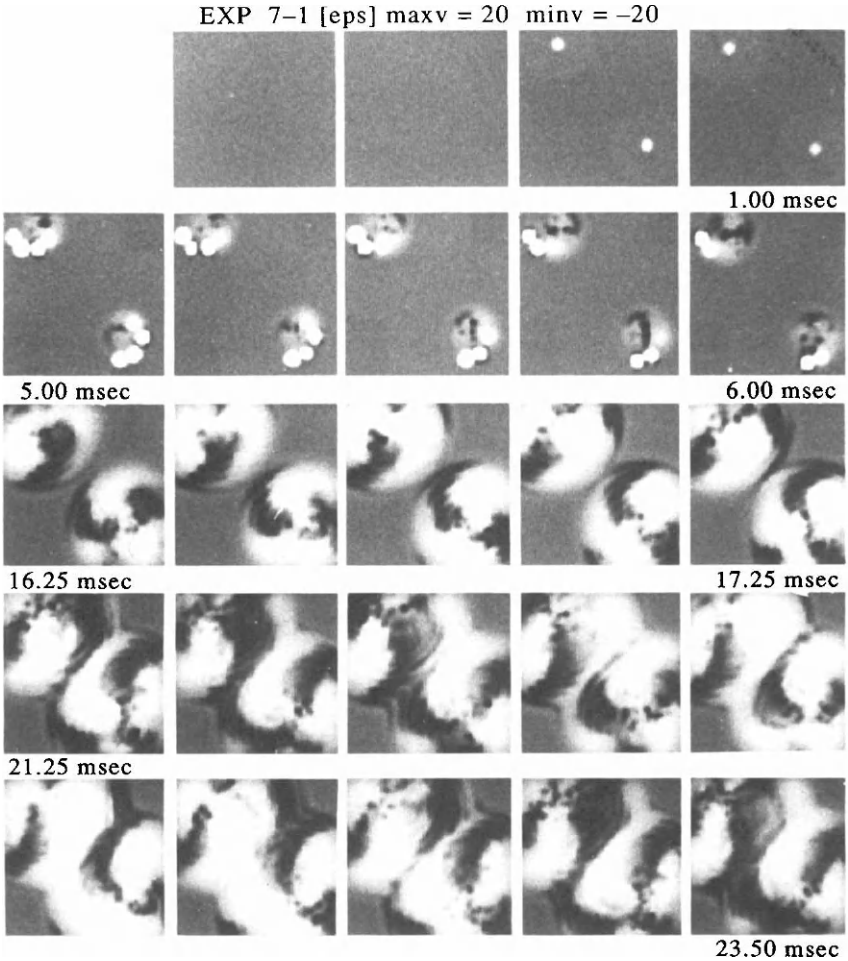


Fig. 1. Two rotating activity waves in a simulated cortical-like field shown by the computed macroscopic variable h_{Δ}. The activating stimuli end in 6 ms. Both rotations are clockwise. Some superimposed oscillatory activity is manifest. (The black is related to dominating inhibition, the white to dominating excitation.)

activities recorded from pools of neurons. Moreover, they deduced that the source of correlations among activities of different neurons was assignable to the retina rather than to the cortex.

PET provides cortical maps of metabolic or neurotransmitter activities, related to rates of neural activity reflecting, in some way, the levels of electrical activity. However, the space and time resolution by this methodology (about few seconds, for time, according to Friedland *et al.*, 1985) is yet too low to describe the rapid changes of the neural (electric) activity as in my model.

Some indications of biological interest are, however, furnished by this set of computations. Indeed, an excessive, oscillating and stable activity, which propagates from the origin focus/foci and tends to invade all the available space with a period bounded only by the ARP, is shown in almost all computational experiments with homogeneous neural systems. The rotating activity waves obtained in the model could be put in connection with similar activities, observed in isolated retina of chickens (Gorelova and Bures, 1983) and in cortical regions (Bures *et al.*, 1983). A relationship between the generation of rotating waves and the epilepsy was hypothesized by the latter authors. Hence, the conditions of the model could be related to those responsible for epilepsy in brain.

Short and long-range connection

In a different series of computations models with both short and long-range effects were used. The activity of a neural system, constituted of about 8 million excitatory neurons and 6.5 million inhibitory neurons (grouped in 900 modules), was simulated. Each module had side length equal to 400 μm and contained about 8800 excitatory neurons and 7200 inhibitory neurons. The activity of each module was affected by surrounding modules, within the range of a few steps (in excitatory and inhibitory way, excitation propagating along local circuits), by distant modules (in excitatory way, excitation propagating through long-distance connections) and by external excitatory sources. Below, I shall report only the values of parameters which differ with respect to those quoted in the previous section.

The distribution function of the subthreshold neuronic excitation was such as to produce about 5.5 million excitatory neurons and 4.5 million inhibitory neurons per square centimeter.

The long-distance impulses in the absorption-free zone assumed a constant velocity $v' = 200$ cm/s.

The boundary of the neural system was reflecting and the ARP was assumed shorter: $\tau = 1.0$ ms.

The structure of the deterministic interzone connections, related to long-distance excitatory effects, determine the collective behavior of the model. In the lack of precise experimental information, some simple geometries

were used in these series of computational experiments such as highly symmetric (crystal-like) and randomly-generated.

The first computations, with crystal-like geometries, showed that an appropriate choice of the structural values σ_s and λ_s, a sufficiently high value of the parameter ξ_0 (related to long-distance effects) and a high number of regions exciting a given zone (trough long-distance effects) were able to produce an excitation whose propagation was sustained mainly by long-distance paths. Also, a periodic and permanent pattern of activity could be obtained by choosing appropriate structures of long-distance connections. Hence, the model was able to sustain complex patterns of excitatory and inhibitory activities for an indefinite time (Ventriglia, 1988).

Some other experiments were carried out to examine the possibility of disturbing or ending the evolution of collective activity patterns. In the first experiment, an excitatory disturbance hit a zone of the neural system shortly before an excitation conveyed by the propagation of a self-sustained pattern reached the zone. The subsequent inhibitory period did not suffice to stop the activity and, after a short period of disturbed evolution (< 15 ms), the pattern recovered its previous space-time features with only a small shift in time ($\frac{1}{2}$ ms). In a second experiment, in which the inhibitory effects were strengthened, the evolution of the self-sustained pattern was stopped completely by the disturbance. Hence, in these systems, a strong inhibitory effect, occurring in a critical time, is able to stop the further evolution of patterns. The above results could be compared with those of Enright and Winfree (1987) on the simulation of excitable systems composed by populations of mutually-coupled relaxation oscillators. It must be stressed that in all the computational experiments *feed-forward* inhibition, i.e. inhibition which reaches a neural zone from distant zones was never used. The principle that long-distance stimulation is conveyed by axons of excitatory neurons and that inhibition has only a local origin was always adhered to. This assumption forbade modification of the activity propagation by direct inhibitory influences. It seems that this principle is not valid in some part of the hippocampus.

The second series of computations (randomly generated long-distance connections) showed that the production of stable, but random activity patterns required larger values of the coefficient λ_s and σ_s and an higher number of interzone connections.

Remarks

As regards the adequacy of the kinetic theory to elucidate long lasting activities in absence of external (sensory) reinforcing stimuli, that also by introspection seems to be a basilar feature of human brain functions, the analysis of the above computational experiments enabled us enlighten the following characteristics.

• In models in which:

(a) inhibitory and excitatory neurons are homogeneously distributed in space;

(b) neurons do not distinguish the character of impulses, i.e. excitatory and inhibitory neurons absorb impulses of both species without selection;

(c) neural parameters do not depend on position;

(d) only short-range effects (propagated through local circuits) are considered;

collective activities, maintained for indefinitely long time periods, can be obtained, but they are high-frequency oscillatory and pervasive (very similar to epileptic effects). Otherwise, the strength of inhibitory effects is such as to produce the extinction of all the activities shortly after the ending of the activating stimulation. Also, the *recruitment* of neurons participating to high frequency activities is so strong that it is hardly possible to change the collective oscillatory pattern. Yet, a random space-variability of the threshold prevents oscillatory recruitments.

• In models in which both short and long-range propagations are taken into account, indefinitely long activities at low frequency are obtained. But, also in this case an abnormal trait is present, that is the extreme difficulty to modify (or stop) the space-time development of periodic patterns by local disturbances. In fact, only accurate definitions of time and place of application of the disturbance and a precise tuning of the excitatory and inhibitory parameters would allow the cessation (or the change) of the collective pattern of activity.

To surmount the difficulties encountered in controlling such nonphysiological behaviors, I have modified my theory following two different approaches. The first is concerned with the introduction of elements related to neural processes of attention which modulate the global brain activity (Ventriglia, 1990, 1991). Such processes are discussed in the next two sections. The second approach involves the description of neurotransmitter-related chemical activities which occur in synapses during the transmission of impulses. These elements ought to help in slowing and smoothing the neuronal activity.

Attention, learning and memory

In this section, some insights on the relationships between attentive behaviors and memory and learning are discussed. (For a more complete review see Ventriglia (1990).)

Vinogradova (1975, 1986) assumed that the memory processes are based on the action of two interconnected neural circuits of the brain: the *informational circuit* and the *regulatory circuit*. The informational circuit, CA1 → mammillary bodies → anterior thalamic nuclei → cingulate cortex → entorhinal cortex → CA1 (also referred to as Papez's circuit in the literature),

manipulates and dynamically stores the information reaching it from different sensory sources and projects into neocortex. The regulatory circuit, CA3 → lateral septal nucleus → hypothalamus → pontobulbar reticular formation → mesencephalic reticular formation → medial septal nucleus → CA3, regulates the global level of the brain activity. Connections between the two circuits in hippocampus are carried out by the intervention of neurons of *entorhinal cortex* and *fascia dentata*. The global system is supposed to act as a *coincidence detector*.

If the activity arriving in field CA3 from cortical input (through entorhinal cortex and fascia dentata) is not balanced by the activity arriving from reticuloseptal input (as it is assumed in the case of a *novel* stimulus), the majority of neurons in CA3 are inhibited and as a result the hippocampal global output to two reticular formations decreases. Since the hippocampus has an inhibitory influence on these formations, its reduced activity causes an increasing activity in reticular formations. In turn, this event increases the global activity of the brain, which seems necessary to fix long-term memory traces. The inclusion in the regulatory circuit of the hypothalamus and the linked amygdala, both related to the modulation of the emotional system, may increase/prolong the excitation arriving at neocortex from the regulatory system.

Mishkin (1982) proposed the following conjecture about the memory formation of visual stimuli. When a new visual stimulus is perceived, information representing visual attributes such as size, colour, etc. converge on single neurons of inferior temporal lobe. This *novel* perception trigger a cortico-limbo-thalamo-cortical circuit, which acts as an automatic rehearsal or imprinting mechanism, strengthening the prestriate-temporal connections that participate in firing the circuit in the first place. Amygdala and hippocampus are considered by Mishkin of crucial importance to this process.

Current hypotheses

Later studies changed in part the above view, in particular the route (CA1, mammillary bodies, anterior thalamic nuclei, cingulate cortex, entorhinal cortex) is substituted by a direct connection between hippocampus and entorhinal cortex via subicular complex.

According Squire *et al.* (1989) "the hippocampal formation receives a great part of its sensory information from *polymodal association cortices* via the entorhinal cortex. This information is then processed in subsequent steps by the neural systems of the hippocampal formation: fascia dentata, hippocampus proper (CA3, CA1) and subicular complex. The processed information is returned to the same polymodal regions of cortex via the entorhinal cortex." Because of extensive convergence and divergence at the stage of the hippocampal formation, the above authors retain improbable

that a point-to-point mapping is maintained from the origin to the destination as the information traverses the hippocampal system. They assume, in accord with Vinogradova, that in correspondence with *relevant events* the hippocampus affects the neocortex in a diffused manner.

As regards the cellular aspects of memory it has been hypothesized that the mechanism of formation of long-term memory traces is based on temporal expression of specific genes (Thompson *et al.*, 1987). Initial experimental evidences seem to support this hypothesis. In fact, several workers have reported that changes in synapses and in their organization, as well as the expression of c-fos protein(s) take place in several neural systems of brain as a consequence of pathologically high stimulation leading to kindling and seizures. These observations and the above results on attention and memory suggest that the physiological mechanism of long-term memory probably requires a very high level of excitation, obtainable by appropriate stimulation from reticular and/or hippocampal formations.

Attention conjecture

Based on the previously discussed computational experiments and on the above attention and memory hypotheses, I proposed a conjecture about the attention process (Ventriglia, 1990, 1991). It takes into account the effects of the activities originating from the hippocampus and from prefrontal cortex on the polymodal areas of neocortex when a novel and relevant event is detected. Also important is the experimental finding that at the prefrontal cortex arrive information derived from the hippocampus and from the amygdala/hypothalamus, both via the MD thalamic nucleus and by direct projections (Fuster, 1985). The action of the general attention model, based on two attention terms, *global* and *focused* (or local), can be described as follows.

Some cortical activity based on ultra-stable patterns is assumed to be present at the beginning. When the *novelty* of the neural activity originating from the sensorial stimulation is detected (in hippocampal formation), a global attention process is started (i.e. enhancement of the cortical activity, via hippocampus and/or reticular formations in brainstem) producing the triggering of all the patterns embedded in the long distance circuits of the association area of the cerebral cortex.

Activities in cortical systems induce further activities in the thalamus and in the sub-cortical circuits of attention, where the single patterns of cortical origin are compared with the pattern related to sensorial stimulation to detect similarity. If such similarity is detected in some pattern, the transition to a new phase is hypothesized to take place.

In this stage, the activity related to the process of global attention decays and a process of local attention is triggered (originating in prefrontal cortex), sustaining only the activity of the selected cortical pattern(s). If a novelty

aspect is detected again, the process of global attention restarts, the local attention process decays and the whole process is repeated. Otherwise, also the local attention process decays (habituation phenomenon) and only some ultra-stable patterns of cortical activity remain active.

An event of *absolute novelty* occurs if no similarity is detected between the pattern(s) related to sensorial stimulation and the patterns triggered by global attention process. Memorized replies or behaviors do not exist for this event. In this case, since the process of global attention is still active, the pattern of activity produced in the cerebral cortex as a consequence of the sensorial stimulation is greatly enhanced and a memorization (learning) process might start. Indeed, the high level of activity needed for memory laying could be reached in the involved neural zone. All these activities are gated from the medial septal input that allows bursts of high activity only in very short periods (a few milliseconds) which occur at the frequency of theta waves (5–12 Hz; see Vertes, 1986; Vinogradova and Brazhnik, 1986).

Attention and learning modeling

Attention

At this stage only some gross features of the influence of attention on the neural activity have been introduced in the kinetic theory (Ventriglia, 1990). Simply two new sources of external impulses have been added to the kinetic equations: one with a widespread effect, related to global attention and the other with a position-dependent effect, related to local attention. These impulses affect only the layer where the external stimulation arrives and are denoted by $A_g(\mathbf{r}, \mathbf{v}, t)$ and $A_f(\mathbf{r}, \mathbf{v}, t)$.

The kinetic equation changes to

$$\begin{aligned} f_s(c, \mathbf{r} + \mathbf{v}\Delta t, \mathbf{v}, t + \Delta t) = f_s(c, \mathbf{r}, \mathbf{v}, t) - \Delta t \sigma_s(c, \mathbf{r}) \mid \mathbf{v} \mid \psi f_s(c, \mathbf{r}, \mathbf{v}, t) \\ + \Delta t (Z(\mathbf{r}, \mathbf{v}, t) + \\ A_g(\mathbf{r}, \mathbf{v}, t) + A_f(\mathbf{r}, \mathbf{v}, t))\delta(c-1) + \lambda_s(\mathbf{r}) f^*(\mathbf{v}) N_{s\Delta}(\mathbf{r}, t)\delta(c-2) \\ + f^o(\mathbf{v}) \int_V d\mathbf{r}' \xi(\mathbf{r}, \mathbf{r}') \int dv' f'(v') N_{s\Delta}(\mathbf{r}', t - v'^{-1} \mid \mathbf{r}\text{-}\mathbf{r}' \mid)\delta(s-1)\delta(c-3) \end{aligned} \tag{24}$$

The effects of the addition of these terms were studied by a series of computational experiments and some interesting results are reported earlier (Ventriglia, 1990).

Learning

The strength of the connection between neurons of the different modules can be characterized by a time-dependent origin-destination matrix $\xi(\mathbf{r}, \mathbf{r}', t)$. A simple rule for updating the connection strength might be formulated according to the hypothesis of Hebb (1949) to describe the associative reinforcement between presynaptic and postsynaptic activities. But at this moment no clear experimental results exist on this widely accepted conjecture. In fact, it does not consider all the implications of the work of Lashley on memory traces (1929, 1950). After a long research experience Lashley, while denying the possibility that memories could be based on localized engrams, attributed some possibilities to memory based on largely distributed traces and stressed the linkage between attention and memory (see Thompson *et al.*, 1980; Thompson, 1986). However, the epigenetic stabilization of modular connections seems less controversial. This phenomenon occurs, during the learning phase of the first days of mammalian life, in some part of the cerebral cortex, as for example, in the visual cortex of kittens (Hubel and Wiesel, 1970). The initial state is taken as the critical stage, where the number of connections is maximal. This phenomenon is described in Ventriglia and Érdi (1987) and Ventriglia (1988) by assuming that the strengths of long-distance specific connections can be modified by following some rules based on selective stabilization or elimination of preexisting connections. Accordingly, an hypothesis more complex than that of Hebb was adopted to update the connection strength . The main differences are listed in the following:

(a) the variation of the connection strength is not linked to the activity of single couple of neurons but it occurs between (and is produced by) populations of neurons;

(b) only long-distance connections, linked to excitatory effects, are subject to time changes.

The assumption (a) is sustained by the conjecture (by Szentagothai) of modularity of the entire cerebral cortex. From this conjecture it follows that the main flows of inter cortical excitation occurs among modules. The assumption (b) is sustained both by the modularity conjecture and by the assumption that memory is not local, but broadly distributed in cortex. In addition

(i) a time delay was introduced to take into account explicitly the finite velocity of impulse propagation. In this case, the firing coincidence, which produces the connection strength growth, takes into account the distance that must be covered by the presynaptic impulses to produce effects on postsynaptic neurons at the appropriate time;

(ii) a selective *weakening rule* has been used. As proposed by Hirai (1980), the lower the efficacy of transmission, the greater is the weakening of the connection; and

(iii) to avoid the unbounded growth of the strength of connections an upper limit is imposed at some predetermined value ξ^*. Additionally, the reversal of the sign of the strength (transition from excitatory synapse to inhibitory one) is implicitly forbidden.

The discrete time version of the learning equation is:

$$\begin{aligned} \xi(\mathbf{r}, \mathbf{r}', t + \Delta t) = \xi(\mathbf{r}, \mathbf{r}', t) + \\ k_1 N_{+1\Delta}(\mathbf{r}, t - t_v) N_{+1\Delta}(\mathbf{r}', t) R(\xi^* - \xi(\mathbf{r}, \mathbf{r}' t))\Delta t + \\ -k_2 R(N_{+1\Delta}(\mathbf{r}', t) - \xi(\mathbf{r}, \mathbf{r}', t) N_{+1\Delta}(\mathbf{r}, t - t_v))\xi(\mathbf{r}, \mathbf{r}', t)\Delta t \end{aligned} \tag{25}$$

where $t_v = |\mathbf{r}\text{-}\mathbf{r}'|/v_0$ and R (x) is the ramp function defined earlier.

The above equation modifies slightly those reported in Ventriglia and Érdi (1987) and Ventriglia (1988). The second term on the right hand side is a modified version of the Hebb's rule which takes into account assumptions (a) and (i) above. The factor $R(\xi^* - \xi(\mathbf{r}, \mathbf{r}', t))$ forbids the unbounded growth of the connection strength. The third term on the right hand side is a version of the Hirai's rule, modified as in the second term. With the present knowledge on learning (and memory) it is not very clear that a similar mechanism could be hypothesized to work for learning processes in the adult.

Attention and learning

Some computational experiments were carried out to study the effect of global attention on learning. In particular, the effects on cortical-like zones of hippocampal activities modulated by a septal input was simulated. Towards this purpose the learning rules and the two-dimensional model were used. The model described the activity of neurons grouped in 3600 modules. Both short and long-range effects were taken into account. The initial configuration of the long-distance connections was randomly generated, with a maximal number of 30 connections per module. Connections between modules which had a distance less than 15 modular diameters were not permitted. External sources (simulating an input by sub-cortical nuclei) stimulated a prefixed zone of the model (a rectangle containing 40 modules) for the first two-thirds of the total duration of the experiment and neurons within a prefixed zone of the model (another rectangle) were preferred targets of this external stimulated zone. $\xi_0(\mathbf{r}, \mathbf{r}')$ (= $\xi(\mathbf{r}, \mathbf{r}', 0)$) and $\xi^*(\mathbf{r}, \mathbf{r}')$ assumed constant values for all couples $(\mathbf{r}, \mathbf{r}')$.

A first computation, without the attention term, was used as a reference for subsequent ones.

In a second series of computations, the modulation of the activity of the neural system by a global attention term $A_g(\mathbf{r}, \mathbf{v}, t)$ was permitted. This was accomplished by inserting additional pulses of impulses, hitting all

the modules of the system at prefixed times (10, 25, 40, 55, 70 ms), each pulse having a duration of 5 ms.

In a third series of computations, the attention term $A_g(\mathbf{r}, \mathbf{v}, t)$ was considered without modulation, by inserting an additional inflow of impulses in all the system (40 ms after the beginning time) which persisted till the end of the computation (75 ms).

The following results were exhibited by the space–time patterns of some relevant macroscopic variables and by the variation of the connections' strength.

1st experiment —The activity died out at the end of the stimulation. The connection strength modification occurred at a very low level.

2nd experiment —The activity persisted even after the end of the stimulation. Residual activity was observed also after the end of the attention term. The frequency of this activity was 166 Hz, yet far from the normal cortical activity (see Fig. 2). We must note that this frequency is not related to EEG but to spiking frequency of neurons of single modules *recruited* by the global activity of the system. The connection strength modification occurred at a level higher than that above.

3rd experiment —The activity persisted after the end of both the stimulation phase and the attention term. The frequency was 333 Hz, far from the normal activity. The connection strength modification occurred at a high level.

Synaptic activity description

The kinetic theory is based fundamentally on the electrical characteristics of single neurons and the modality in which these neurons interact with each other in the cerebral cortex. Herein, an attempt towards its extension with elements related to the dynamics of neurotransmitters is presented. The aim is to slow and smoothen the neural activity by using a model closer to the biological *reality.* To this end, some results on the neurotransmitters' activity at the synapses of single neurons and in particular, the presence of neurotransmitters causing slow IPSPs (Thalmann, 1984) are considered. (The relevance of this finding for neural modeling is emphasized by Traub, 1988; and Chapter 6, this volume.) Such enrichment of the theory leads to the elimination of a very drastic simplifying assumption made earlier and valid until now, assuming an instantaneous rising of the postsynaptic inner excitation subsequent to impulse absorption.

In this section, a complex model of neural systems having several interacting subfields is described. The index c (= 1 ... n) denotes the possible, different neural fields in the model. The index s denotes the different impulse types in each field c , s I^c, $I^c = I_e{}^c + I_i{}^c$, where $I_e{}^c$ and $I_i{}^c$ are sets related to excitatory and inhibitory impulses. The index s' denotes the different neuron types in each field c , s' N^c. The values of the excitation

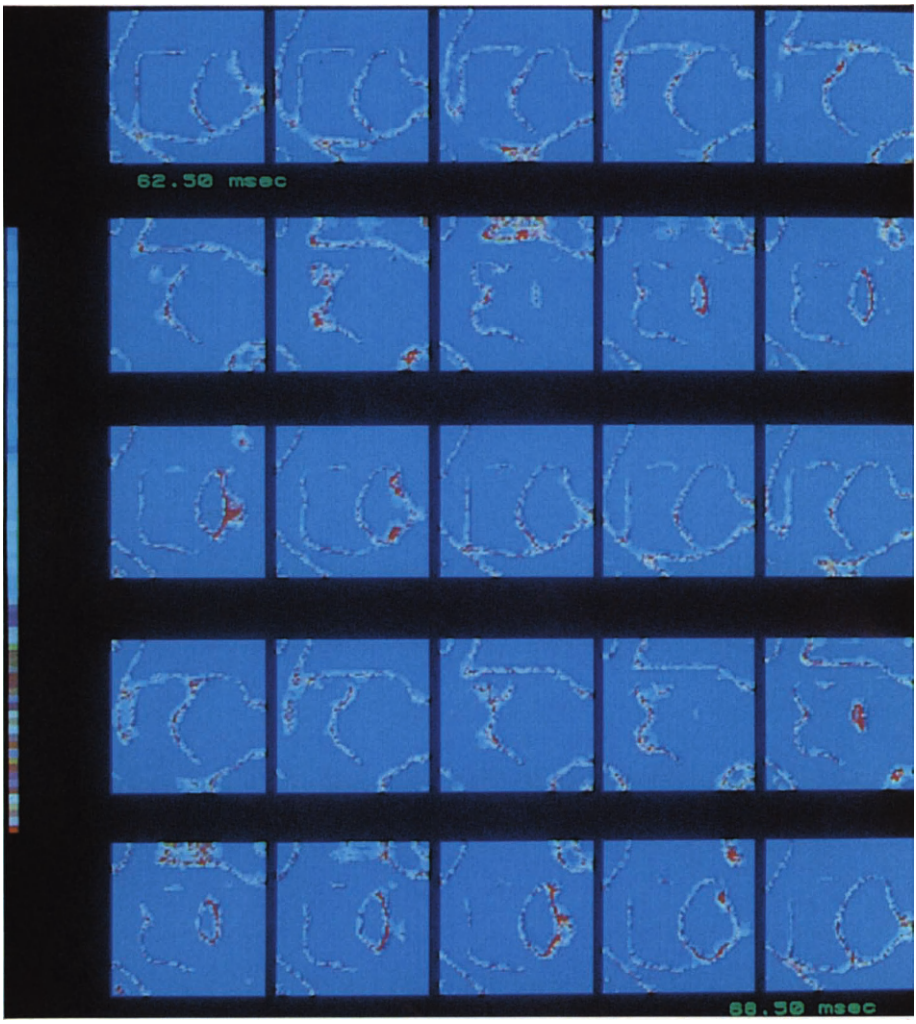


Fig. 2. Oscillatory collective activity at 166 Hz in a simulated cortical-like field shown by the number of firing neurons. (The red is related to high values.)

are normalized to a range between 0 and 1 ($e = 0$ corresponds to the MHL (about −85 mV) and $e = 1$ corresponds to the firing threshold level (about −55 mV). In the following notations the sub-field c is considered explicitly only when necessary. As usual, the mathematical basis for a statistical description of the neural model is given by the following three PDFs: $f_{c,s}(\mathbf{r}, \mathbf{v}, t)$, $g_{c,s'}(\mathbf{r}, e, t)$ and $g^{\circ}{}_{c,s'}(\mathbf{r}, t)$. The first denotes the PDF of the velocity of impulses, the second the PDF of subthreshold neuronic excitation and the third describes neurons which are in a state of absolute refractoriness. They are formally defined in the following way:

Def. 19 — $f_{c,s}(\mathbf{r}, \mathbf{v}, t)\, d\mathbf{r}\, d\mathbf{v}$ = Probable number of impulses of type s in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$ of the field c, whose velocity lies in the element of velocity space $d\mathbf{v}$ around $\mathbf{v}$, at time t.

Def. 20 — $g_{c,s'}(\mathbf{r}, e, t)\, de\, d\mathbf{r}$ = Probable number of neurons of type s' in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$ of the field c, whose excitation lies in the element de around e, at time t.

Def. 21 — $g^{\circ}{}_{c,s'}(\mathbf{r}, t)\, d\mathbf{r}$ = Probable number of neurons of type s' in the volume element $d\mathbf{r}$ around the point $\mathbf{r}$ of the field c in refractory state, at time t.

$$\psi_{c,s'}(\mathbf{r}) = \int_0^1 g_{c,s'}(r, e, 0)de$$

is the probable number of neurons of type s' for unit volume in field c.

The usual principle of conservation of probabilistic mass must be observed.

The description is completed by two other functions $C_{c,s,s'}(\mathbf{r}, t)$ e $C_{c,s,s'}(\mathbf{r}, t)$, connected to free and linked neurotransmitter concentrations, respectively and a law of transformation, coupling linked neurotransmitters concentration $\bar{C}$ to the neuronic inner excitation $E_{c,s'}(\mathbf{r}, t)$.

The rules of interaction between neuron and impulse populations determine the space–time evolution of the functions f, g, g°, C, $\bar{C}$, E. The composition of the related kinetic equations is obtained by the usual balancing method. The formulation procedure requires some new parameters and equations.

The absorption rate of impulses of type s by neurons of type s' in an elemental neighbourhood of $\mathbf{r}$ at time t, $h_{s,s'}(\mathbf{r}, t)$ (linked to the neurotransmitters' release at the related synapses) is given by:

$$h_{s,s'}(\mathbf{r}, t) = \int d\mathbf{v} \mid \mathbf{v} \mid f_s(\mathbf{r}, \mathbf{v}, t - t_0)\sigma_{s,s'}(\mathbf{r}) \tag{26}$$

$C_{s,s'}(\mathbf{r}, t)$ is the concentration of neurotransmitters linked to type s impulses in synaptic clefts of type s' neurons. Let us assume that its time variation is ruled by the following equation:

$$\frac{\partial C_{s,s'}}{\partial t}(\mathbf{r}, t) = -k_d C_{s,s'}(\mathbf{r}, t) - k_1 C_{s,s'}(\mathbf{r}, t) + k_a h_{s,s'}(\mathbf{r}, t) \tag{27}$$

where k_d and k_l are, respectively, the diffusion constant outside the synaptic clefts (without effects on the postsynaptic membrane) and the linkage constant at the postsynaptic membrane (which is responsible for the induction of a variation in the inner neuronic excitation). Let k_a be the elemental contribution to concentration growth due to a single impulse. Moreover, let $C_{s,s'}(\mathbf{r}, t)$ be the concentration of neurotransmitters linked to receptors on the postsynaptic membrane and its time variation be ruled by the equation:

$$\frac{\partial \bar{C}_{s,s'}}{\partial t}(\mathbf{r}, t) = -k_i \bar{C}_{s,s'}(\mathbf{r}, t) + k_1 \bar{C}_{s,s'}(\mathbf{r}, t) \tag{28}$$

where k_i is the deactivation constant, i.e. the rate with which the neurotransmitters lose the control of the ionic channels.

Let us assume that the inner neuronic excitation depends (at a first approximation) in a linear way on the linked neurotransmitter concentration $\bar{C}_{s,s'}(\mathbf{r}, t)$:

$$E_{s,s'}(\mathbf{r}, t) = k_s \bar{C}_{s,s'}(\mathbf{r}, t) \tag{29}$$

where k_s is a constant. Then: $\Delta E_{s,s'}(\mathbf{r}, t) = k_s\ \Delta C_{s,s'}(\mathbf{r}, t)$.

We must note that the excitation obeys the above equation only for values of e in the range (0, 1). In fact, the excitation could assume values in the range (E_{min}, E_{max}), where E_{min} and E_{max} are conditioned only by the maximum available quantity of inhibitory and excitatory neurotransmitters, respectively. Instead, the kinetic rules, imposed by the nonlinearity of the neuronic answer, state that the point $e = 0$ is a natural reflecting boundary and the point $e = 1$ is a natural absorbing boundary. The sum, in an elemental neighbourhood of $\mathbf{r}$ at a time t for a period Δt, of all the contributions, excitatory and inhibitory, induces the following net variation of the excitation:

$$\Delta E_{s'}(\mathbf{r}, t) = \left[\sum_{s \in I_e} \Delta E_{s,s'}(\mathbf{r}, t) - \sum_{s \in I_i} \Delta E_{s,s'}(\mathbf{r}, t)\right]. \tag{30}$$

The computation of the probable number of neurons which fire or are at a MHL requires more attention. However the neurons which are in the state of absolute refractoriness and more specifically those which recover from this state render the computation difficult. Exactly what should be the value of their inner excitation when they exit from the absolute refractoriness? We can assume that during absolute refractoriness, the inner excitation is no more influenced by the concentration variation of different neurotransmitters and hence behaves as if these concentrations were null. When this period ends and the neuron exits from refractoriness, the inner excitation suddenly reacts to the effective values of neurotransmitter

concentrations and suffers the effects of the following concentration variations: $(C_{s,s'}(\mathbf{r}, t) - 0$, for each s. Hence, the value e_t of the inner excitation, corresponding to these concentration variations is given by:

$$e_{ts'}(\mathbf{r}) = \left[\sum_{s \in I_e} k_s \bar{C}_{s,s'}(\mathbf{r}, t) - \sum_{s \in I_i} k_s \bar{C}_{s,s'}(\mathbf{r}, t)\right]. \tag{31}$$

If e_t is greater or equal to 1, a condition in which the related neurons must fire, then they do so and become again refractive. So, we can compute the following probable number of firing neurons in an elemental neighbourhood of $\mathbf{r}$ at a time t :

$$N_{c,s'}(\mathbf{r}, t) = \int_{R(1-R(\Delta E))}^{1} g_{c,s'}(\mathbf{r}, e, t)de + \theta[(e_r + e_t) - 1]N_{c,s'}(\mathbf{r}, t - \tau) \tag{32}$$

where R (x) is the ramp function and $\theta(x)$ is the step function:

$$\theta(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

In the above equation, the first term of the right hand side is related to the firing neurons which were previously in a subthreshold state. The second term is related to neurons which recover from refractoriness, having the inner excitation value greater than the threshold level and fire once again.

Now, let us describe in more detail the process to which the neurons in MHL are subjected. When the subthreshold excitation reaches the MHL due to the increase of the inhibitory neurotransmitters concentrations, the neurons are unable to further react to this increase and so their subthreshold excitation remains fixed at zero, as if the concentration variation were null. At the end of the inhibition period, the inner excitation of inhibited neurons suddenly reacts to the effective values of the neurotransmitter concentrations, suffering the effects of concentration variation. This process is quite similar to that proposed for the refractoriness and the value e_t of the inner excitation, corresponding to these concentration variations. If $M_{c,s'}(\mathbf{r}, t)$ denotes the probable number of neurons in MHL, the following equation can be formulated:

$$\begin{aligned} &M_{c,s'}(\mathbf{r}, t + \Delta t) = \\ &\theta(-\Delta E)[\textstyle\int_0^{R(-\Delta E)} g_{c,s'}(\mathbf{r}, e, t)de + M_{c,s'}(\mathbf{r}, t) + \\ &\theta[-e_r + e_t)]N_{c,s'}(\mathbf{r}, t - \tau)]. \end{aligned} \tag{33}$$

In this equation, the first term in the square bracket is related to neurons which reach the MHL, the third is due to neurons which exit from the state of refractoriness and suffer a high inhibition so as to reach the MHL.

By the usual balancing method, the following set of finite-difference kinetic equations can be formulated for the functions $f_{c,s}(\mathbf{r}, \mathbf{v}, t)$, $g_{c,s'}(\mathbf{r}, e, t)$ and $g^{o}{}_{c,s'}(\mathbf{r}, t)$.

$$f_{c,s}(\mathbf{r}+\mathbf{v}\Delta t, \mathbf{v}, t+\Delta t) = f_{c,s}(\mathbf{r}, \mathbf{v}, t) - \Delta t \mid \mathbf{v} \mid f_{c,s}(\mathbf{r}, \mathbf{v}, t) \sum_{s' \in N^c} \psi_{c,s'}(\mathbf{r})\sigma_{c,s,s'}(\mathbf{r}) + \lambda_{c,s}(\mathbf{r})f^{*}(\mathbf{v})N_{c,s}(\mathbf{r}, t) + f^{o}(\mathbf{v})\sum_{k'}\int_{A^{k'}} d\mathbf{r}'\xi_s(\mathbf{r}, \mathbf{r}')\int dv' f'(v')N_{k',s}(\mathbf{r}', t - v'^{-1} \mid \mathbf{r}\text{-}\mathbf{r} \mid) + \Delta t Z_{c,s}(\mathbf{r}, \mathbf{v}, t) \tag{34}$$

$$g_{c,s}(\mathbf{r}, e, t+\Delta t) = \delta(e-(e_r+e_t))N_{c,s}(\mathbf{r}t-\tau)+ \tag{35}$$

$$g_{c,s}(\mathbf{r}, e-\Delta E, t)+\theta(\Delta E)\delta(e-(e_r+e_t))M_{c,s}(\mathbf{r}, t) \tag{36}$$

In the top equation the second term of the right hand side is related to absorption, the third to local effects, the fourth to distance effects (k′ runs over all the neural field connected to field c by long-distance effects) and the fifth to external sources, respectively. The sum present in fourth term is a consequence of modeling neural systems having multiple, interacting sub-fields. We must note that the first term on the right hand side of the second equation containing a Dirac function, accounts for the probabilistic mass linked to neurons recovering from the state of absolute refractoriness. As discussed above, if $e_r + e_t$ is greater or equal to 1, a condition in which the related neurons must fire, then those neurons go again to a refractory state. In that case, the related probabilistic mass disappears from the second equation and it appears again in the last equation, related to refractory neurons. The third term, also containing a Dirac function, accounts for the neurons which exit from the MHL and takes the inner excitation value of $e_h = e_r - e_t$.

The above equations were tested for a simple case of a neural modulus constituted of excitatory neurons and fast and slow inhibitory neurons (Ventriglia and Raddi, 1993). The above excitatory and inhibitory neurons were given two different values of ARP corresponding to τ_e = 1.5 ms and τ_i = 0.5 ms and two different values of resting level namely e_{r_e} = 0.34, e_{r_i} = 0.67. The latter values affected the firing threshold of related neurons.

From the resulting activity, it was evident that the very strong inhibitory component, linked to slow IPSP, prevented the neuronic sub-populations

from firing until its contribution to the subthreshold excitation was nearly exhausted. Hence, the firing occurred on the tail of the slow IPSP.

An application of this model to the field CA3 of the hippocampus will be presented in a forthcoming paper.

Let me stress that the greater complexity in the description of the neuron activity, obtained by considering features linked to neurotransmitter dynamics, is dropped in the kinetic theory without losing the ability to describe the activity of neural populations by statistical methods. In this respect the above approach differentiates itself from that of other *real* models such as, for example, those proposed by Traub *et al.* (1992) or Ekeberg *et al.* (1991) where neurons having a complex activity have been described; but, when population activities have been studied, networks of singly simulated neurons must be considered.

Conclusion

Let me recapitulate the cardinal features of the evolution of this neural network theory which embodies the greater part of two decades of my research endeavor. I started with a model in which only random effects and neural activity at very low levels were described. In that case it was quite impossible to obtain stable activity surviving the external stimulation phase. Subsequently, random neural networks having high level activities were described, where lasting activities were obtained but they were excessive and similar to epilepsy phenomena. The separation of the connections among neurons in short-range, random connections and long-range, deterministic (or scattered) ones, as it occurs in cerebral cortex, permitted the modeling of low-frequency activities. In this case the frequency was bounded from below by the space dimensions of the simulated neural systems, larger dimensions permitting lower frequencies. The extreme difficulty in those neural models to modify the activity patterns constituted the abnormality appearing at that stage. But two new paths of research, one related to the enrichment of the description of neuronal activity and the other to the description of more global brain activity, seem to provide better opportunities to also eliminate the last anomaly. The first takes into account the dynamics of the neurotransmitters in describing the excitation propagation among neurons and the second is related to the modulation of the global cortical activity by attention processes. A better description of the cortical structures, with the introduction of the total cortical lamination, should complete this modelling procedure. In such a way, neural systems could be described (and simulated) whose activity approaches more closely the normal cortical activity. At this point which is not too far from the current model, by considering appropriate values of the parameters linked to the description of neurons and their connections, we can begin

the description (and the simulation) of the capability of the cerebral cortex to carry out *recognition processes* and *cognitive computations.*

11

Psychology, Neurobiology and Modeling: The Science of Hebbian Reverberations

DANIEL J. AMIT

INFN, Sezione di Roma, Istituto di Fisica, Universita di Roma "La Sapienza", P.le Aldo Moro, Roma, Italy

Introduction

This chapter intends to express a certain unity between cognition, neurobiology and theoretical modeling. The subject is vast and multi-dimensional. I will trace a one-dimensional trajectory which will go from cognitive experience to neurophysiology to toy modeling and back to biology. I find this a very productive trajectory both in terms of generating interesting questions and in terms of proposing answers. The task is far from over, but it is also far from exhaustion.

I shall start by discussing introspective considerations for necessity of having attractors, which neuropsychology has termed reverberations. Then I will show that there exists experimental evidence for reverberations and that attractors exist and are observable in neurophysiology. Attractors that are observed neurophysiologically, which appear to be required for mental processing, are very informative both for describing the functioning of the brain and for measuring information about computation processes.

I will discuss toy models which already reproduce the surprising kind of information contained in the experiments of Miyashita (Miyashita and Chang, 1988; Miyashita, 1988) which have produced some very convincing and beautiful results concerning attractors.

In the last section I shall go beyond toy models to build up a realistic model of processing in cortex from information in anatomy and neurophysiology.

From introspection to neurophysiology

Let me start with a quote from Von Neumann which is a warning about the constraints we impose on our approach to describing computational phenomena in the brain. He said, in 1954, in his book *The computer and the brain*:

> It also ought to be noted that the language here involved may well correspond to a short code in the sense described earlier, rather than to a complete code: when we talk mathematics, we may be discussing a secondary language, built on the primary language truly used by the central nervous system. Thus the outward forms of out mathematics are not absolutely relevant from the point of view of evaluating what the mathematical or logical language truly used by the central nervous system is. However, the above remarks about reliability and logical and arithmetical depth prove that whatever the system is, it cannot fail to differ considerably from what we consciously and explicitly consider as mathematics. [Yale 1954, p.82]

This wisdom we should all keep in mind. One may be losing much potential power by imposing onto a neuronal computing system the kind of computation used when writing on paper a formal proof of a mathematical theorem or when programming a conventional computer.

The gray box

If you try to describe what happens in the brain you can start by drawing a very simplified picture: there are channels of sensory input such as visual, auditory, olfactory systems. Then some preprocessing encodes the incoming stimulus. This is followed by a whole cascade processes which terminates in some computation to decode the "decision", in activating muscles or glands. When you start looking at such a complex problem, one thing is to describe the whole system as a black box without worrying about the detailed description of all the basic modules involved in the computational process. Even in this case it is possible, by reasoning on very general grounds, to characterize the properties of some modules and the main structure of the internal organization of these modules. For instance you can imagine a very simple and homogeneous structure: there are many layers of neural networks and there are synapses connecting the layers.

Therefore, whatever is coming from different sensory channels, is pushed from layer to layer and eventually there is a layer for the expression of the output.

This is one possibility. However, the option that I prefer is that there is really a processing cut which splits the homogeneous feed-forward picture and entails the necessity of computational phenomena which go on and maintain themselves persisting even in the absence of the input stimulus. Obviously these phenomena are selective to the input but they can exist for long times independently of the presence of the stimulus.

Let me justify this picture by a very simple example which I find rather convincing. Suppose that somebody asks you to translate a word in your language. He does that vocally using the auditory system, so, for instance, he says to you "what is the italian for water?". One thing that may happen is that you immediately say "acqua". It may be the case that you recognize the word "water", that you know its expression in italian but you cannot produce the corresponding word. You may spend several minutes going on with the effort of retrieving the required translated word. You may even forget about the entire episode, deal with other matters, and realize hours later that the answer has come up. Note that within a few hundred milliseconds, maybe a second, the stimulus has disappeared. The stimulus consisted of the auditory signal which held the word to be translated and the task to be performed. Both of these somehow have to persist because the stimulus has disappeared, and they must persist for a long time, they may persist for days.

This phenomenon raises a natural question: how do they persist? A possible solution is to suppose that the recognized word and the information concerning the translation task are in memory. However this is not a satisfactory answer because there is a huge number of memorized things (in the synapses) and there is something special about one memory, the one affected by the stimulus. The task must be alive and the internal representation of the stimulus must be in a persisting way active to drive the computational processes of the output system. So somehow there must be a status in which one of the things that you have in passive memory, which is usually described as being stored in the synapses, has been put into a special status so that the computation can consciously or unconsciously go on about a specific task.

The conclusion that you can draw is that there must be somewhere the possibility of keeping one out of large store of items in memory in a special status for relatively long times. This mechanism which maintains one memory out of the stock is not related to short, intermediate or long term memory. In each one of these classes we need an extra mechanism which allows keeping one of the passive memories in a special status. Moreover, you immediately also have to conclude that this kind of extra mechanism has to be local: you cannot employ the whole cortex to keep the word and the task alive in the interim period. Moreover, one can perform

several concurrent tasks such as processing several words essentially at the same time. Therefore, the cortex must be modular and each module has to be able to store many memories and to operate concurrently with other modules. Indeed the classification of the input signal is very fast even for a selection among many different possible memories.

There is a beautiful experiment by Lionel Standing (Standing, 1973) which corroborates this observation. Thousands of uncorrelated slides are collected. Then increasing sets of these images are shown to people in an experiment. Two days later the subjects are asked to recognize previously seen pictures out of pairs of pictures, one of which had been seen and one randomly chosen. He measured the number of correct answers as a function of the number of pictures that the subjects have learned before. The number of correct answers increased linearly (not saturating) until 10,000 pictures, and the recognition time was found to be constant, about half a second, even for a very large number of images.

Experimental evidence for the existence of attractors

In the previous section the logical necessity of a special active state has been discussed. It was argued that strengthening synaptic storage is not enough and that there must be mechanism which allows for tagging of the specific memories. Moreover, this something must be of a kind which can be used by other parts of the brain which eventually perform the task. The simplest carrier seems to be spike activity: indeed it can be communicated over long distances, can be observed by other parts of the brain and furthermore it can be observed by neurophysiologists. So the particular memory, excited by the stimulus, should be flagged by a persistent activity selective to the stimulus. Each pattern should be tagged by a particular spike rate distribution in such a way that it is possible to make available the internal representation for later processing.

I am quite convinced that the picture I have just presented was clearly perceived by Donald Hebb more than 40 years ago. Unfortunately, when one talks about Hebbian learning it is always about how synapses are modified and generally one forgets about the fact that Donald Hebb was not a neurochemist nor a neurophysiologist: he was a psychologist and he was trying to construct an infrastructure of psychology to supplement or replace the mythological one. His words about the dynamic principle of the synaptic modification are:

> Let us assume that the persistence or repetition of a reverberatory activity (or "trace") tends to induce lasting cellular changes that add to its stability... When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes

> place in one or both cells such that A's efficacy, as one of the cells firing B, is increased. [p.62] When one cell repeatedly assists in firing another, the axon of the first cell develops synaptic knobs (or enlarges them if they already exist) in contact with soma of the second cell. [p.63]

The main idea is that the presence of the stimulus provokes a certain activity distribution in a given area of the cortex. Hebb was looking for synaptic dynamics, or, at least, for a dynamical principle, which makes the activity provoked in this module able to persist in the absence of the stimulus. He did not need any synaptic modifications to maintain the activity when the stimulus is on, because stimulus stabilizes it. Instead, he was looking for some synaptic tricks in order to make the activity distribution something similar to the one provoked by the stimulus when the stimulus is on. Namely, the synaptic changes have been supposed to sustain reverberations, providing the local network with a connectivity suited to maintain the activity even in absence of the stimulus.

As was stated before, spike activity is something that can be observed by neurophysiologists. So, if the picture is true, there must be experimental evidence for the existence of reverberations which are attractors. In fact, Miyashita's group and other neurophysiologists have done some very interesting experiments which demonstrate the existence of reverberations of the type discussed above. Moreover, the results provide us with a treasure of additional information which is extremely useful in constructing models.

Before going into a detailed description of Miyashita's experiments, I wish to discuss an apparent paradox concerning single unit recordings. I am claiming that there is a persistent selective activity distribution in a module. In this case it should be possible to actually obtain the full *internal representation* of the stimulus looking at one neuron at time (single unit recordings). This should be possible even if nothing interesting is happening with a single neuron, because a single neuron cannot maintain a selective activity. This means that if one observes a persistent elevated activity distribution, there must be a collective phenomenon which involves a large number of neurons in a certain module. This activity distribution constitutes the internal representation of the stimulus and it is reproducible. Since the internal representation is basically determined by the synaptic structure, every time you show the same stimulus, the same neuron will have the same activity. Therefore, recording neuron by neuron, it is actually possible to reconstruct the full internal representation of a given stimulus by single unit recordings, even if nothing interesting happens in a neuron when considered as an isolated entity.

Let us consider in more detail the chain of experiments performed by Miyashita *et al.* In the basic experiment (Miyashita and Chang, 1988) a set of computer generated images is prepared and two monkeys are trained to perform visual image matching. In order to avoid correlations in the ge-

ometrical structure, the pictures were generated using fractal algorithms which produce presumably uncorrelated images. Then the monkey was trained to do the following task: first there is a short warning signal. Then one of the generated pictures is shown on the screen for a short time (200 ms). Next, the screen is blanked and, after 16 s, another picture is presented and the monkey is trained to respond differently if the second one is the same as the first one or is different. When the monkey begins to perform well and gives a high percentage of correct answers the recording starts.

Miyashita's group was able to find a very small area (1 mm^2) of the anterior ventral temporal cortex (AVT) where selective activity takes place even in the absence of the stimulus, i.e. during the 16 s delay. It is a quite remarkable fact that this type of persistent activity goes on in a very localized area. This fact constitutes experimental evidence for the existence of a localized mechanism which is able to maintain a reverberation for as long as 16 s. Furthermore, when the pictures are presented in a completely random order there is also the evidence of reproducibility of the spike rate distribution: when you show a certain image you get always the same persistent elevated activity distribution. For instance, you can imagine presenting a particular pattern, record the activity level of a given neuron. Then you can show an arbitrary number of different images, and, finally present the same pattern that had been shown first. If you record again the spike activity, you get an activity level which is very similar to the delay activity following the first presentation of this image.

In Fig. 1 there are four figures from Sakai and Miyashita (1991): the rasters above each histogram constitute a representation of spike activity of a certain neuron recorded during twelve different repetitions of the same picture (a dot is drawn at the time a spike is emitted). The reproducibility is self-evident in all the four cases. Furthermore in Fig. 1 there is the experimental manifestation of the effects of the typical collective behavior of physical systems. For example, in case (a) a neuron active during presence of the stimulus maintains high activity in the delay period; in case (b): during the period of 200 ms in which the external stimulus is on, the activity level is very low. However, when the stimulus is removed, the neuron activity goes up and reaches a level that is essentially the same as in the case (a). The interpretation in terms of attractors and reverberations is quite simple: during the period in which the stimulus is on it maintains some activity distribution. But when the stimulus is removed, the high activity level is restored by the reverberations which are governed by the structure of synaptic connections. Actually the system is correcting the "error" on the basis of the acquired past experience which has structured the attractors. In fact it is not necessary to give the full internal representation exactly to obtain the recognition of a word. When you pronounce the word "water", you can pronounce it slowly or quickly, loudly or quietly, with noise etc., and we will consider the same task and the same word for a very wide range of such phenomena.

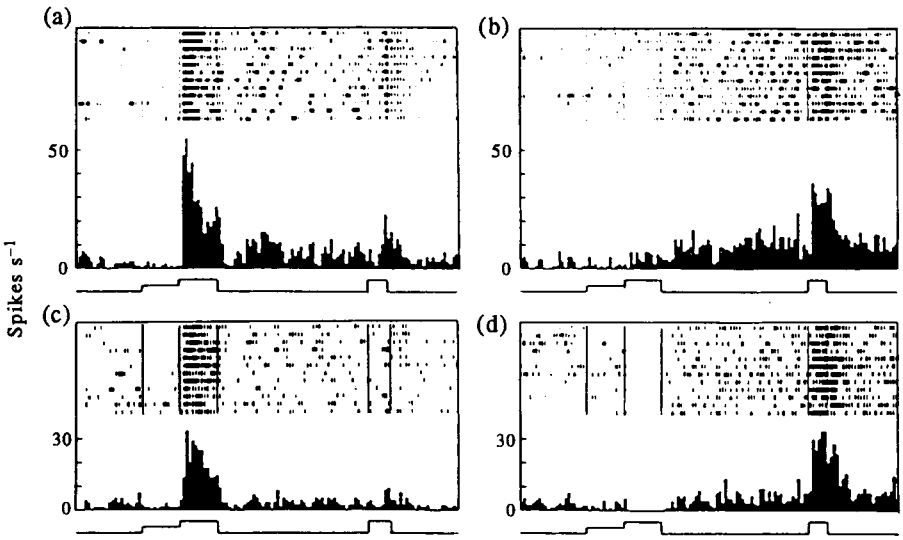


Fig. 1. Four types of neuronal behavior observed with a single unit (in each part of the figure there are 12 rasters and a spike rate histogram below): (a) Active in presence of stimulus and persists in absence; (b) neuron active in the delay period of a different stimulus; (c) active in presence of stimulus, inactive in absence; (d) inactive in presence of stimulus and active in delay period. (From Sakai and Miyashita, 1991.)

There is also another very interesting behavior which characterizes this particular part of anterior ventral temporal cortex (Miyashita and Chang, 1988): varying the size, the angle of rotation or the color of an image, one finds an internal activity distribution which is essentially the same. This means that in this part of the cortex the attractors are already an internal representation of some prototypes and stimuli which differ for size, angular position or color have essentially the same internal representation and all of them flow into the same attractor.

Conversion of temporal to spatial correlations

In the second experiment (Miyashita, 1988) the pictures are shown in a fixed order during the training: each picture is always followed by the same picture. A sampling of the internal representations is reconstructed by recording the activity of single neurons from the AVT cortex. Subsequently, the correlation between internal representations corresponding to different patterns is calculated off-line. It is found that the only thing the correlation between representations depends on, is the distance separating the two patterns in the training sequence.

In order to prevent the spurious effects of some accidental correlations in the internal representation of the images, the two monkeys were trained on the same set of images but the patterns were presented in a different order. For both monkeys the same result was obtained: the correlation between two internal representations becomes insignificant at the separation of about 5 in the training sequence: this is the magic 5 as we shall see below. So we have to conclude that in this part of the cortex the synaptic structure created during learning, has correlated attractors even though the items you learn are presumably uncorrelated. In other words, if you show uncorrelated pictures always in the same temporal order, then the internal representations, the attractors, are correlated and preserve some of the information concerning the temporal order.

Cognitive psychology implications: priming and false alarms

It is not too surprising that when we see two uncorrelated things very often in temporal contiguity, then when you think about one, the other can be easily provoked. There is a very familiar phenomenon in cognitive psychology called priming. People are trained to identify certain distorted patterns, e.g. visual patterns. Then the time for the identification of the patterns is measured and compared to the time required to identify distorted images if preceded by a pattern that is cognitively correlated with the image to be identified. For example showing first the full word "chair" and then the word "table" distorted. The identification time in the second case is significantly shortened.

The interpretation in terms of attractors is quite simple: when the cortical module is in one attractor which is correlated with the internal representation of the incoming stimulus, then the time the system takes for moving into the new attractor is much shorter than the time required for a transition from an uncorrelated state. This interpretation immediately makes a prediction about a neurophysiological experiment. One can measure priming on the level of AVT cortex by observing transition times between internal representations. This is a direct neurophysiological experiment and it could be done with single unit recordings.

There are other phenomena in cognitive psychology whose interpretation in terms of correlated attractors is quite simple. For example there is a well known phenomenon called "false alarm": subjects are shown a set of objects, and then a sublist is defined. Patterns are shown very quickly, at random, and the subject has to respond after each presentation whether the stimulus belongs to the sublist or not. If the pattern shown is correlated with one belonging to the sublist then the number of errors grows. In attractor language this happens because each one of these patterns carries information of the fact that it was seen very often in temporal proximity to a pattern which belongs to the sublist. So their internal representations

are correlated and when a pattern which is correlated with one of stimuli belonging to the sublist is presented, the probability of a wrong answer is higher. This is one more thing that can be verified on the neurophysiological level by single unit recordings.

Lessons from Miyashita's monkey

Let me summarize what we can learn from Miyashita's experiments:

- Long-lived elevated spike rate distributions selective to the stimulus exist. The activity in the module is not moving randomly activating all the neurons in the same way. So ergodicity is broken and selectivity is possible. Moreover the spike rates are self-sustained as they persist even after the disappearance of the stimulus.
- The reverberations can be observed in localized modules ($1mm^2$, 10^5 neurons)
- The reverberations are reproducible on the single unit level. The activity of a single neuron depends only on the stimulus leading to the reverberation and it is always the same when the same pattern is presented. A particular neuron may be shared by several different representations, thus having the same activity is all of them.
- Persistent high rates are not single neuron properties (collective behavior).
- The internal representations are of prototypes: in this area they do not depend on angular position, size, color etc.
- The internal representations are distributed, i.e. many neurons participate in representing a pattern.
- The reverberations are attractors in the sense that some neurons which are activated during the presentation of the stimulus stop during the delay period. So the network is able to perform error correction. A class of similar stimuli leads to the same activity configuration.
- The internal representations code for temporal correlations: patterns which are temporally correlated in the training phase have correlated internal representations.
- The coding level, which is the fraction of neurons in the network that actually are activated in the attractor, is low.
- The elevated rates of the neurons active in the attractors are significantly lower than the maximal rates allowed to a cortical neuron i.e. 500 spikes/s.

Toy models

Miyashita experiments have raised two essential problems: the first one concerns the type of dynamics which can maintain a selective distribution of neuronal activity in absence of the stimulus. The second problem is the way the system creates correlated attractors from patterns that are uncorrelated. In order to solve these two problems I will describe some toy models which reproduce the main features of the Miyashita experiments. Toy models are a very powerful tool for investigating very complicated systems and turn an untractable problem into a tractable one. They represent the essential features of the systems, indicate certain phenomena and make predictions. If the results are satisfactory, their robustness to the removal of the simplifying assumptions can be tested.

The Hopfield model

The Hopfield model (Hopfield, 1982; Amit, 1989) provides a solution to the first problem, that is how to maintain selective activity distributions. The distribution must be sensitive to the stimulus but not too sensitive, in the sense that internal representations of very similar stimuli must lie in the same basin of attraction so that the system is able to perform error correction. On the contrary, if two stimuli are different enough, then the system has to evolve to different attractors enabling the possibility of discriminating between two different representations.

Let us start describing the formal neuron which constitutes one of the basic elements of the Hopfield network. Keeping in mind that the Hopfield model is a toy model one ignores the fact that the biological structure of the neuron is much more complex and has some very different features. In this picture neurons emit spikes when the depolarizing potential reaches threshold and it is impossible to have two spikes running on the axons because a neuron cannot emit a new spike during the absolute refractory period (ARP). So, if time is divided into small enough slices, there are never two or more spikes travelling along the axons of the same neuron in the same time slice. Consequently, states of neurons can be represented, in a toy model, by two state variables. For instance you can conventionally choose 0–1 variables σ_i, with $\sigma_i = 0$ corresponding to no spike emitted and $\sigma_i = 1$ corresponding to one spike emitted by neuron i.

The neurons are put together in order to build up a more complex structure: the network. A state of the network at a given moment is a distribution of 0–1 which specifies which neurons are emitting spikes and which are not. So the state of the N neurons belonging to the network can be completely specified by assigning an N-bit word to each time slice. Then, at each time step, the state of the neuron σ_i is updated on the basis of the activity of the other neurons connected to it. There are some issues con-

cerning the choice of the updating procedure which best imitates biological asynchrony, but we will not worry about this now, as we are studying a toy model.

In the framework of the schematization of the Hopfield model, the question of updating the state of neuron σ_i can be formulated in the following way: first the depolarization of a neuron relative to its threshold is computed. The variable J_{ij} is taken as the depolarization of neuron number-i due to a single spike in neuron number-j. This is the synaptic efficacy which can be positive or negative, excitatory or inhibitory. We will not worry about the fact that neurons are either entirely inhibitory or excitatory. In a toy model we can imagine them to be completely randomly mixed. So in the following we will use positive and negative synapses, and positive and negative values for the two state variables $s_i = \pm 1$ representing the activity of the neurons. The 0–1 representation and the (–1,1) are equivalent (up to a local shift of thresholds) related by the transformation $s_i = 2\sigma_i - 1$.

The difference between the depolarization and the threshold is:

$$h_i(t + \delta t) = \sum_{i=1}^{N} J_{ij} s_j(t) \tag{1}$$

which determines the subsequent state of the neuron, according to:

$$s_i(t + \delta t) = sign[h_i(t + \delta t)]. \tag{2}$$

The new activity will be positive whenever the depolarization is above the threshold ($h_i > 0$) and the neuron emits a spike, and will be negative otherwise. In this case there is no noise in the system, so the new activity of neuron i is determined with certainty and a given state is always followed by the same state. In reality there are many noisy processes which affect the time evolution of the activity, but in a toy model we come back to them when we discuss more realistic situations.

The important thing concerning this simple model (Hopfield, 1982) is that it is possible to construct a specific synaptic matrix J_{ij} which produces an arbitrary number of different attractors. It is possible to choose the states one wants as attractors among the very large sets of all the random uncorrelated patterns (i.e. all the random N-bit words in which each bit is chosen at random, with equal probability, to be either $+1$ or -1). Defining ξ_i^μ as the state neuron i should be in when the network settles in attractor number μ, the connection between neuron i and neuron j will be the sum of the products of activities of neuron i and neuron j in each one of the p patterns to be memorized in the network:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu. \tag{3}$$

It is easy to show that the h_i produced by this synaptic structure has the same sign of ξ_i^ν whenever $s_i = \xi_i^\nu$:

$$h_i = \sum_j J_{ij} s_j = \frac{1}{N} \sum_j \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \xi_j^\nu = \xi_i^\nu [1 + \delta_i] \tag{4}$$

where δ_i is the part of h_i contributed by the $p-1$ patterns ξ^μ with $\mu \neq \nu$:

$$\delta_i = \frac{1}{N} \sum_j \sum_{\mu \neq \nu} \xi_i^\mu \xi_j^\mu \xi_j^\nu . \tag{5}$$

Since all the bits of the patterns are uncorrelated, δ_i is a one-dimensional random walk of Np steps of size N^{-1}. The RMS distance it can reach is:

$$\sqrt{\bar{\delta_i^2}} \simeq \sqrt{p/N}.$$

So, if the number of stored patterns p is much less than N then δ can be ignored and the sign of h_i will be just equal the sign of ξ_i^ν. This means that starting in one of these states ξ_i^μ that are stored in the synaptic structure, the network remains there for an indefinite time, until a new stimulus arrives. Furthermore, these fixed points are true attractors with a certain basin of attraction. Starting from a state which resembles one of the stored pattern, the system is able to perform error correction moving into the nearest attractor. The system works fine as an (auto) associative memory even in the presence of noise and if there are not too many "errors" in the initial state, the system is still attracted to the prescribed set of attractors.

This is expressed in Fig. 2 where the time development of the network is followed by monitoring the overlaps with the stored patterns, defined as:

$$m^\mu = \frac{1}{N} \sum_{i=1}^{N} s_i \xi_i^\mu . \tag{6}$$

These are natural variables which measure the macroscopic distance between the present state of the network and each of the stored patterns. They are the required generalization of the magnetization in a ferromagnet (see e.g. Amit, 1989). The overlaps are plotted for two different initial configurations, which resemble the first stored pattern ξ_i^1. The initial state has been chosen to have a large overlap (m^1 = 0.35–32% errors, m^1 = 0.20–40% errors) only with the first stored pattern. The overlap of the first pattern increases rapidly to 1 and all the errors are corrected dynamically while the overlaps with other patterns fluctuate around 0 and their value is negligibly small when compared to overlap with the first stored pattern. This happens also in the second case in which the number of "errors" in the initial state is significantly bigger than in the first case.

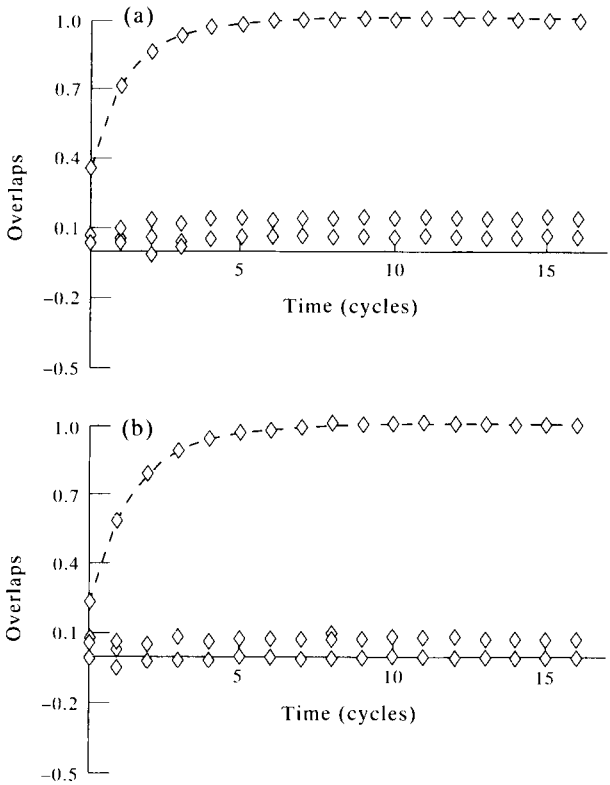


Fig. 2. Time development of overlaps of a network of 400 neurons with 7 stored patterns in two simulations of the Hopfield model with noiseless asynchronous dynamics. (From Amit, 1989.)

Let me stress one important point: what we can see at the end of the time development is essentially a fixed point. This does not mean that the system is very quiet, but exactly the opposite. What happens when the system arrives at a fixed point is that half the neurons are emitting a spike and half the neurons are quiescent. The definition of the model is such that the fixed point implies that the neurons which are 1 emit a spike in every interval of time discretization, i.e. in every 2 ms (the absolute refractory period). Neurons which are fixed at -1 emit no spikes. Thus the fixed point is a very active state of the modeled neural system, in which half the neurons emit 500 spikes/s. The fluctuations in the attractor in the presence of noise imply that the 50% active neurons are active at less than 500 spikes/s. while the other 50% begin to have nonzero spike frequencies.

All the activities are determined by a dynamical process and the fixed point is the result of an attractor mechanism which produces the final firing pattern. Even if there is some noise in the system, the network is still able to increase the overlap with the stored pattern that is the most similar to the initial state (the stimulus) (see Fig. 3). Again one of the overlaps converges rapidly to the vicinity of 1, but it does not stop there. In the presence of noise, one of the overlaps fluctuates in the neighborhood of 1 and the other remains negligibly small. But, again, the system is attracted to one state which depends on the initial stimulus and is able to correct dynamically the "erroneous" bits of the input. When noise increases the system becomes completely ergodic. Eventually all neurons emit 250 spikes/s whatever stimulus you start from. The system mixes all the patterns and produces a state which is uncorrelated with any of the stored patterns.

Correlated attractors from uncorrelated patterns

The second question raised at the beginning of this section concerned the correlations between internal representations of patterns which are presented in a fixed order during the training period. It turns out that it is possible to reproduce Miyashita correlations in a very simple toy model (Griniasty *et al.*, 1993). Again we take the uncorrelated patterns $\xi_i^\mu = \pm 1$ as the distribution of activities in the stimuli to be learned. The dynamics of the activity of the neurons is described again by the equations 1 and 2. However the patterns ξ_i^μ are stored differently by adding to the Hopfield matrix a term that reflects the fact that the patterns in the training period are presented in a fixed order:

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p}[\xi_i^\mu \xi_j^\mu + a(\xi_i^{\mu-1}\xi_j^\mu + \xi_i^\mu \xi_j^{\mu-1})]. \tag{7}$$

The idea is that in the synaptic matrix the information concerning temporal correlation is stored by the second term in the square brackets. If

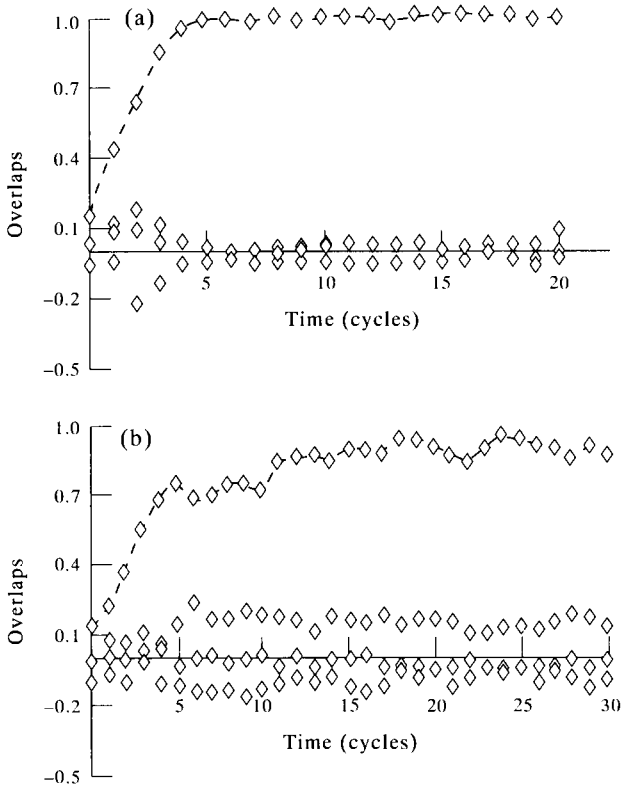


Fig. 3. Same simulations in the presence of noise. Measure of noise "temperature"=0.4, 0.6. At noise level 1.0 the system becomes ergodic. In each case the initial state (the stimulus) has more than 40% errors (recall that 50% errors is a loss of all information in the pattern). (From Amit, 1989).

$a = 0$, then we have the Hopfield matrix. Otherwise there exists a term which connects the structure of one pattern to the preceding one. The last term preserves the symmetry of the Hopfield matrix so that all attractors are fixed points in the absence of noise and therefore the analysis is simple.

The additional term can be constructed by the network even in a realistic learning scenario. During the training period the pattern number μ–1 is always shown before pattern number μ, so when the network is in the attractor number μ–1, and the transition to the next stimulus occurs, the information for the both patterns is available and can be used for learning. This kind of term is possible only in the general framework of attractor dynamics which were discussed in the previous section. When stimulus μ–1 is removed the information about it is stored and maintained active in the corresponding attractor.

It is not necessary to have a further term such as $\xi^{\mu}\xi^{\mu-2}$ to connect the attractor μ–2 to attractor number μ. This follows from the neural dynamics driven by a synaptic matrix like (7). Using the synaptic matrix 7, the difference between the depolarization and the threshold, equation 4, and the definition of the overlaps equation 6, can be written as:

$$h_i = \sum_{\mu=1}^{p} m^{\mu}(t)[\xi_i^{\mu} + a(\xi_i^{\mu+1} + \xi_i^{\mu-1})]. \tag{8}$$

Starting from a state identical to pattern ξ_i^2, i.e. which has overlaps $m^2 = 1$, and $m^{\nu} = 0 \ \forall \nu \neq 2$, the equation for the depolarization, equation 8 becomes:

$$h_i = \xi^2 + a(\xi_i^3 + \xi_i^1).$$

If $a < 1/2$, then the original uncorrelated patterns ξ^{μ} are still attractors because the additional term cannot change the sign of the whole expression. It is always the first term in the square brackets which dominates the sign of h_i. This means that if the monkey has not been trained long enough, then the uncorrelated stimuli produce uncorrelated internal representations. That is a prediction that could be verified neurophysiologically.

If $0.5 < a < 1$ the uncorrelated patterns ξ_i^{μ} are no more stable points. Indeed, starting from $m^2 = 1$, there is a probability 0.25 that the second term dominate. So, after one step of the dynamics one finds (Griniasty *et al.*, 1993) that there are three non vanishing overlaps: $m^2 = m^1 = m^3 = 0.5$. But now, at the next step, two new terms will contribute to the sum in expression (8) and therefore also the overlaps with patterns at distance two from ξ_i^2 will take non zero values.

The equations which determine the attractors of this system, for low loading levels, are the mean field equations:

$$m^{\mu} = \langle\langle \xi^{\mu} sign \left[\sum_{\nu=1}^{p} \xi_i^{\nu} + a(\xi_i^{\nu+1} + \xi_i^{\nu-1}) \right] \rangle\rangle,$$

in which the double angular brackets indicate an average over the distribution of bits in the stored patterns (see e.g. Amit, 1989, Chapter 4). These equations have been solved numerically (Griniasty *et al.*, 1993). The results are shown in Fig. 4, which indicates that each attractor has finite overlaps with about 5 stored uncorrelated patterns. The solution has been found exactly using **Mathematica** (Cugliandolo, 1992), which propagates rational fractions. It has been found that there are exactly 5 nonzero overlaps in each attractor. Starting from a state identical to one of the stored, uncorrelated patterns, one obtains the following iterations for the overlaps with the stored patterns:

0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0
0	0	0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{5}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	0	0	0
0	0	$\frac{1}{32}$	$\frac{3}{32}$	$\frac{13}{32}$	$\frac{19}{32}$	$\frac{13}{32}$	$\frac{3}{32}$	$\frac{1}{32}$	0	0
0	$\frac{1}{128}$	$\frac{3}{128}$	$\frac{13}{128}$	$\frac{51}{128}$	$\frac{77}{128}$	$\frac{51}{128}$	$\frac{13}{128}$	$\frac{3}{128}$	$\frac{1}{128}$	0
0	$\frac{1}{128}$	$\frac{3}{128}$	$\frac{13}{128}$	$\frac{51}{128}$	$\frac{77}{128}$	$\frac{51}{128}$	$\frac{13}{128}$	$\frac{3}{128}$	$\frac{1}{128}$	0

It is also possible to calculate the correlations between different attractors. The correlations can be calculated exactly from the mean field solution. The correlations between one generic attractor and the attractors corresponding to the patterns neighboring in the training sequence are:

$$\cdots \frac{4}{256} \quad \frac{11}{256} \quad \frac{32}{256} \quad \frac{85}{256} \quad \frac{170}{256} \quad 1 \quad \frac{170}{256} \quad \frac{85}{256} \quad \frac{32}{256} \quad \frac{11}{256} \quad \frac{4}{256} \cdots$$

So the resulting attractors are correlated beyond the first neighbor memorized in the synapses and become small at a separation of 4–5. It is remarkable that this result does not depend on the value of the constant a: the only requirement is that a lie in the interval [0.5–1].

It is quite surprising that this result is so close to the experimental findings of Miyoshita. Of course, in the experiment neither the size of the correlations nor their range correspond exactly to these values. But we must keep in mind that this result has been produced in a toy model.

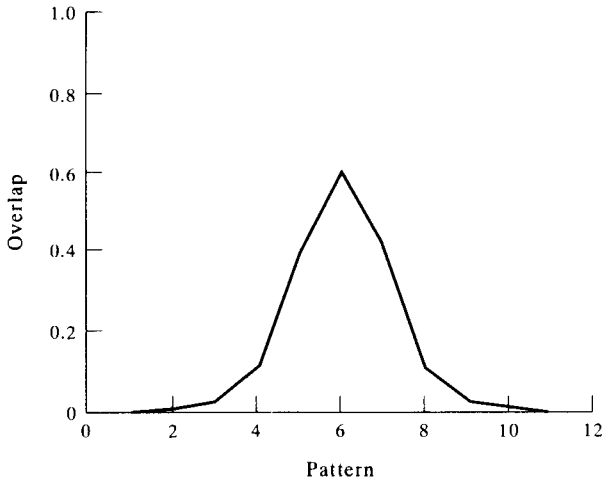


Fig. 4. Overlaps between an attractor and stored patterns as a function of the separation of the pattern in the sequence from the pattern underlying the attractor. (From Griniasty *et al.*, 1993).

Limits and robustness of the Hopfield model

Limits on storage

The analysis of Hopfield model has raised the question of how many patterns can be stored and effectively retrieved. Let us come back and examine in more detail the problem of storage capacity for the Hopfield model itself and for some developments of the Hopfield model. It was shown that the random patterns stored in the synaptic structure of the Hopfield model are stable states as long as the number of patterns grows much more slowly than the number of neurons. This estimate was obtained by imposing that h_i, the difference between the depolarization and the threshold, had the same sign as the corresponding bit in the pattern used as initial condition. If all the neurons of the network are set to states equal to those in one of the stored patterns (for instance ξ_i^1), then at the next step, if the noise is not too large compared to the signal term, one gets the same activity configuration (see equation 4). This is signal to noise analysis and it produces estimates which are surprisingly close to rigorous results obtained by a more elaborate calculation.

The elementary argument (Hopfield, 1982) leads to the naive estimate that storage is safe as long as $p \ll N$. In fact, with a more careful argument (Weisbuch and Fogelman-Soulié, 1985) one finds that for the probability that any bit in a pattern be unstable to be vanishingly small, as $N \to \infty$, one must have:

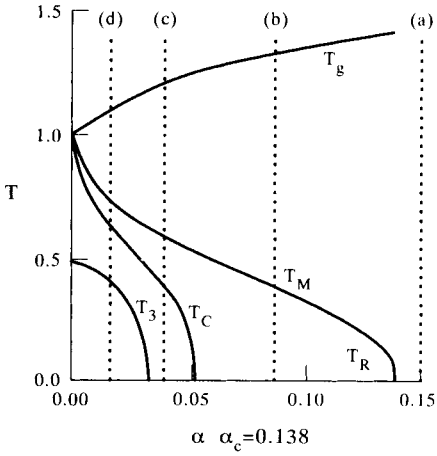


Fig. 5. $T-\alpha$ phase diagram of a network near saturation. T_g is the line of transition into the spin-glass phase (no retrieval); at T_M retrieval appears; at T_c the retrieval phase becomes a global minimum of the free-energy; below T_3 mixtures of three patterns appear; below T_R replica symmetry breaks.

$$p < \frac{N}{2 \ln N}.$$

This estimate is based on the assumption that the noise term can be well approximated by a random variable with gaussian distribution (see e.g. McEliece *et al.*, 1987).

These limitations on the number of stored patterns can be relaxed if one allows a small finite fraction of misaligned bits. In fact, for purposes of storage and retrieval, it is not necessary to require that every one of the attractors be identical to one of the patterns stored in the synaptic structure. The calculation of the storage capacity in these conditions has been solved (Amit *et al.*, 1987, 1985a,b) using tools from statistical mechanics. The results can be summarized in the $T-\alpha$ diagram which is depicted in Fig. 5. T is the temperature which is related to the fast noise in the dynamics of the neuron. At $T = 0$ we have the noiseless case as in the original Hopfield model. α is the storage level i.e. the ratio between p, the number of patterns, and N, the number of neurons. The diagram in Fig. 5 represents the lines in the plane $T-\alpha$ across which the network changes its qualitative behavior.

Let us consider in more detail the behavior in the noiseless case, when $T = 0$: if $\alpha < 0.138$ then it is possible to effectively store and retrieve patterns. This means that if the number of patterns is less than the 14% of the number of neurons, then the attractors are very similar to the configurations of activity stored in the synaptic structure. The percentage of neurons which are not identical in the attractor to their state in the pattern itself, grows as α is approaching the critical value $\alpha_c = 0.138$. When α goes beyond its

critical value, the percentage of errors rapidly reaches the maximum value of 50%. This means that the attractors get completely uncorrelated with the stored patterns. The transition is discontinuous and just before becoming 50% the highest error fraction is less than 1.5% (see e.g. Amit *et al.*, 1987).

Robustness

In the basic model the network is assumed to be fully connected: each neuron sees any other neuron of the network through a symmetric connection. This is clearly too idealized. The robustness of the results has been tested under several types of disruption.

Three types of synaptic dilution have been considered:

- Random symmetric elimination of synaptic efficacies: starting with a network with synapses equation 3, one zeroes synaptic efficacies symmetrically until the average number of nonzero connections per neuron becomes Nc.
- Asymmetric disruption: again one starts from the matrix equation 3, however, in this case the synapses are cut in a completely random way, not preserving the symmetry (a synapse may be cut only in one direction).
- Extreme asymmetric disruption: as in the previous case but with the constraint $c \simeq \ln N/N$.

The first case was treated analytically by Sompolinsky (1987) for a network whose synaptic matrix has the following form:

$$J_{ij} = \frac{c_{ij}}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu .$$

The dilution coefficients c_{ij} take the values 1 and 0 with probabilities c and $1 - c$, respectively. To preserve the symmetry it is also assumed that $c_{ij} = c_{ji}$. The solution makes use of the fact that the interactions remain symmetric and the thermodynamic apparatus is still available.

The properties of such randomly, symmetrically, diluted network are very similar to the properties of the fully connected network. For instance the retrieval quality is very midly impaired by symmetric dilution even for dilution levels up to 0.8 (see Sompolinsky, 1987). The critical storage capacity α_c decreases as the dilution increases and $1 - c$ tends to its maximum value. However, if one computes the storage capacity relative to the number of surviving synapses per neuron (α_c/c), the curve increases with the level of dilution (see e.g. Sompolinsky, 1987; Amit, 1989).

In the second case asymmetry is introduced in the network by cutting the synapses in one direction only. We have no way of studying the system

analytically, but systematic simulations by Gutfreund (see e.g. Amit, 1989, Chapter 7) show that the behavior of the network near attractors is very similar to the symmetric case. For instance, when only 50% of the synapses remain double sided, $J_{ij} = J_{ji}$, there is good retrieval until $\alpha_c = 0.1$ while when all the synapses are cut in at least one direction, α_c decreases to 0.06. If the effective capacity, per remaining synapse is concerned, the network performs as well or better than the intact one.

In the last case, extreme random asymmetric dilution, the connections are randomly diluted to a level where the mean number of remaining synapses per neuron is of order $\ln N$:

$$c \simeq \frac{\ln N}{N}.$$

The interest in this case is not because it is closer to biological reality than the standard Hopfield model. In fact the average cortical connectivity is a few thousand per neuron so the total number of neurons N should be an astronomical number since $\ln N \simeq 10^3$. However, in spite of the fact that the network connectivity is asymmetric, the dynamics and hence the storage capacity can be studied analitically. Derrida *et al.* (1987) have computed the effective critical capacity, α_c/c to be 0.7 instead of 0.14 of the fully connected model.

Toward a more realistic model: 0–1 neurons

In the Hopfield model, when the network is in one of its attractors, half the neurons emit spikes at saturation rates (500 spikes/s) and the remaining neurons are quiet. On the other hand, Miyashita's experiments show that:

- The elevated rates of the neurons active in the attractor (10–20 spikes/s) are very significantly lower than the maximal rates allowed to a cortical neuron.
- The coding rate, which is the fraction of neurons in the network that actually are activated in the attractor, is very low (5–10%).
- No bimodality in the distribution of spike rates has been observed.

The problem of effectively storing and retrieving sparsely coded patterns has been solved (Tsodyks and Feigel'man, 1988; Buhmann *et al.*, 1989) using the 0–1 description of neural states as representing passive and active states, respectively. It was observed that the properties of the new model can be quite different from the ± 1 model, because the random shifts in the threshold, implied by the transformation of variables, are strongly correlated with the elements of the synaptic matrix. Low coding levels of activity entail that memories have to be selected with a probability which is biased

toward the zeroes. So the pN bits $\eta_i^\mu = 0, 1$ of the p patterns ($\mu = 1, ..., p$) are chosen independently, at random, according to the biased distribution:

$$\Pr(\eta_i^\mu = 1) = f \ll 0.5$$
$$\Pr(\eta_i^\mu = 0) = 1 - f.$$

In the noiseless case the dynamics of the activity of the neurons is governed by the following equation:

$$V_i(t + \delta t) = \Theta[h_i(t + \delta t) - \theta] = 0, 1$$

where θ is the threshold and h_i is the depolarization of neuron number i. Note that here h_i is the depolarization proper. The threshold appears explicitly. The depolarization is given in terms of the synaptic weights J_{ij}:

$$h_i(t + \delta t) = \sum_{j=1}^{N} J_{ij} V_j(t)$$

and V_i=1,0 implies that neuron i does (does not) emit a spike in a discretization interval around time t. In this case, however, only active neurons contribute. The synaptic matrix is the sum over all the patterns of the product of the activities that the two neurons should have when getting into the attractors, subtracting from each the value f, the average fraction of ones in a each memory:

$$J_{ij} = \frac{1}{Nf(1-f)} \sum_{\mu=1}^{p} (\eta_i^\mu - f)(\eta_j^\mu - f).$$

It is not hard to prove that this synaptic matrix makes $V_i = \eta_i^\mu$ a fixed point attractor if p is not too large. Indeed, calculating the depolarization, it is possible to express h_i as a linear combination of the similarities (overlaps) of the dynamic state of the network to the stored patterns:

$$h_i = \sum_{\mu=1}^{p} (\eta_i^\mu - f) \left[\frac{1}{Nf(1-f)} \sum_{j=1}^{N} (\eta_j^\mu - f) V_j(t) \right]$$
$$= \sum_{\mu=1}^{p} (\eta_i^\mu - f) m^\mu(t) \qquad (9)$$

where the similarities m^μ are defined as:

$$m^\mu(t) = \frac{1}{Nf(1-f)} \sum_{i=1}^{N} (\eta_j^\mu - f) V_i(t) \in [-1, 1].$$

If the number of patterns p is not too large, it is simple to verify that when starting from a state of the system which is exactly equal to one of

the stored patterns (say for instance $V_i(0) = \eta_i^1 \Rightarrow m^1 = 1; m^{>1} = 0$), at the next step you get the same configuration of activities provided that you choose the uniform threshold θ $(-f < \theta < 1 - f)$:

$$V_i(t + \delta t) = \Theta(\eta_i^1 - f - \theta) = V_i(t).$$

Let us sum up the main properties of the 0–1 model:

- It is possible to design the attractors into the synapses in a learning-like fashion. The η_i^μ are fixed points or noisy attractors if p is not too large.
- The attractors, as in the Hopfield model, have large basins of attraction so that if the network is presented a corrupt version of one of the η^μ, the system is able to perform error correction by producing an attractor which is identical to η^μ.
- Non ergodicity is a collective phenomenon and consequently the process of retrieval is robust to noise and disruption.

In addition there is no active–passive symmetry: the coding rate of the attractors can be very low as in the associative cortex. Moreover the 0–1 representation can be interpreted in terms of spike activity in more natural way than the ± 1 schematization.

Finally, at very low coding rates ($f << 1$) the storage capacity is significantly enhanced and the number of storable patterns can grow as N^2. This storage capacity approaches Gardner's result (Gardner, 1987) for the absolute maximum storage capacity. It is interesting to note that while the result of Gardner has been obtained by searching in the space of all synaptic matrices, here we have an explicit synaptic matrix which performs essentially as well.

Attractor neural network retrieving at low spike rates

The 0–1 model discussed in the previous section does not solve the problems of low spike rates and of bimodality. In this section I will complicate the network to actually be able to describe different levels of spike rates. In the new model, which was developed with Tsodyks (Amit and Tsodyks, 1991) following ideas of Frolov (Frolov and Mendvedev, 1986), the spike rates are represented by analog variables and the dynamics of the network is described as that of integrated synaptic currents driven by spike rates.

Neuron number i is described by two continuous variables: the afferent current I_i and the depolarization potential V_i. I_i is the current that is arriving to the soma of neuron i, it decays with some characteristic time constant τ_I and it is driven by synaptic input which is determined by the contribution from the spikes in the pre-synaptic neurons. To simplify the description, let us assume that:

- The afferent current has no structure inside the neuron.
- The spikes arriving at a synapse are very short events, a point in time (the emission time of the k-th spike by pre-synaptic neuron number j will be indicated as t_j^k)
- The transmission delay on the line from j to i is about 1–2 ms and is negligible compared to the typical time scales of spike rates.

If $\Delta I J_{ij}$ is the amount of current contributed to the post-synaptic neuron i by a single spike on neuron j, then the differential equation for the afferent current is:

$$\frac{dI_i(t)}{dt} = -\frac{I_i(t)}{\tau_I} + \Delta I \sum_{j=1}^{N} J_{ij}\delta(t - t_j^k). \tag{10}$$

The first term on the right hand side represents the decay of the afferent current (τ is the decay constant).

The other variable is the membrane depolarization at the soma V_i, whose dynamics is described by the equation:

$$\tau\frac{dV_i(t)}{dt} = -V_i(t) + I_i(t). \tag{11}$$

When V_i, driven by the afferent current I_i, reaches the threshold then a spike is emitted by neuron i. These spikes then feed back into equation (10). This is the *integrate-and-fire* model for the neuron.

Equations (10) and (11) are a closed set of dynamical equations and so provide essentially a complete description of the model. Given the value of all the parameters and specifying the form of the synaptic matrix J_{ij}, it is possible to simulate the behavior of the network. However, the model can be reduced to a more tractable one (Frolov and Mendvedev, 1986): The detailed spike contributions can be substituted by mean spike rates $\nu_i(t)$ which can be obtained by averaging over a certain time interval δT. Let us consider the transformation from spikes to rates in more detail. Equation 10 can be written in a different form:

$$I_i(t + \Delta t) - I_i(t) = -\frac{I_i(t)}{\tau_I}\Delta t + \Delta I \sum_j J_{ij}\Delta N_j(\Delta t)$$

where $\Delta N_j(\Delta t)$ is the number of spikes arriving through the channel j in the interval Δt. This is a strongly fluctuating random variable: it is mostly zero and rarely more than one if the integration time interval Δt is 10 ms for example (Δt and δt must be longer than the duration of a single spike and of a typical transmission delay). But we have a large number of such random variables corresponding to all the afferent channels and the averaging time δT can be chosen in such a way that $\Delta t > \delta T$ and, that,

in a δT interval the number of afferent spikes per neuron is large enough. If C is the connectivity per neuron, f the coding rate and 10 spikes/s the average activity, then the number of spikes arriving in δt is $10 \cdot C \cdot f \cdot \delta t/1000$. In cortical condition we have that $C \simeq 15000$, and $f \approx 5 - 10\%$ so that the number of spikes arriving in δt is $15 \simeq 75$. In these conditions, the law of large numbers allows one to replace $\Delta N_j(t)/\Delta t$ by the mean frequency $\nu_j(t)$, leading to:

$$\frac{dI_i(t)}{dt} = -\frac{I_i(t)}{\tau_I} + \Delta I \sum_{j=1}^{N} J_{ij}\nu_j(t) \tag{12}$$

with:

$$\nu_j = \lim_{\Delta t \to \infty} \frac{\Delta N_j(\Delta t)}{\Delta t}.$$

Obviously equation 12 constitutes a good approximation but it is not exact, since the law of large numbers entails that the number of random variables tend to infinity. In cortical conditions the finite value of Cf may leave some fluctuations. Those will affect equation (12) as incoming noise.

Transduction of current to rate

To complete the description we need the explicit form of the rates $\nu_j(t)$ in terms of the afferent currents I_j. In order to find the form of the current to rate transduction function, we calculate $\phi(I)$ i.e. the inverse of the time between spikes as a function of the afferent current by integrating equation 11. For instance, if I_i is constant, then the system behaves as an oscillator and the rate $\phi(I)$ is given by:

$$\phi(I) = \begin{cases} 0 & \text{for } I < \theta \\ \left[\tau_0 + \tau \ln \frac{I}{I-\theta}\right]^{-1} & \text{for } I > \theta \end{cases} \tag{13}$$

where θ is the deporalization threshold for the emission of a spike and τ_0 has been introduced to mimic the behavior during the absolute refractory period (ARP). It is important to notice that the ARP essentially establishes the time scale of maximum rates. When I becomes large, the second term in the square brackets in equation (13) goes to zero, and the frequency takes the value $1/\tau_0$. So, for instance, if τ_0 is 2 ms then the maximum frequency will be 500 spikes per second. Choosing τ_0 as the basic time unit, $\tau_0 = 1$, and when a neuron saturates (i.e. when I becomes very large) $\phi(I) \to 1$ the corresponding frequency is $1/\tau$.

In the presence of noise the time interval between spikes is the first passage time (FPT) in a random walk. Assuming that the input noise is a stationary gaussian process, the problem has been discussed in detail in Holden (1976) and the solution is in Amit *et al.* (1991):

$$[\phi(I)]^{-1} = 1 + \tau\frac{\sqrt{\pi}}{\sigma}\int_0^{\theta} e^{(z-I)^2/\sigma^2}\left[erf\left(\frac{z-I}{\sigma}\right)+1\right]dz \qquad (14)$$

where all potential and currents are measured in units of ΔI ($\Delta I = 1$) and I is the mean current while σ is the variance of the gaussian process. I is intended to be a combination of the stationary part of the noise (nonselective mean of a noise input) and the part of the signal which varies selectively depending on the pattern being retrieved.

To proceed further we need to estimates the parameters in the function $\phi(I)$. The ARP τ_0 and τ are measurable and in order to estimate θ we assign n, the number of incoming spikes needed to reach the threshold, which is also measurable. Then $\theta = nJ$ where J is a typical value of the synaptic strength J_{ij}. Here things get less general in the sense that the estimate for θ depends on J which in turn depends on the particular synaptic model. However, at this point, we are ready to test the possibility of having stationary activity states with low spikes rate (Amit *et al.*, 1991).

Stable selective rates

Assume that the matrix J_{ij} is that described in the previous section:

$$\tau_I J_{ij} = \frac{1}{Nf(1-f)}\sum_{\mu=1}^{p}(\eta_i^\mu - f)(\eta_j^\mu - f). \qquad (15)$$

A mean field theory for such a model can be derived, due to the symmetry of the synaptic matrix (see e.g. Kuhn, 1990; Shiino and Fukai 1990). The natural parameters for the description of retrieval are the similarities of the distribution of stationary activities to the stored patterns η^μ:

$$m^\mu = \frac{1}{Nf(1-f)}\sum_i(\eta_i^\mu - f)\langle\nu_i\rangle$$

where $\langle\nu_i\rangle$ is the time average of the activity of neuron i. m^μ can be rewritten as the difference between the mean activity of the neurons in foreground ($\eta_i^\mu = 1$) and the mean activity of the neurons in background (Amit *et al.*, 1991). Given equation (15), the definition of the incoming current equation 12 and the above definition of the overlaps, one can write the expression for the current as:

$$h_i = \sum_{\mu=1}^{p}(\eta_i^\mu - f)m^\mu(t). \qquad (16)$$

Since the currents determine the frequencies, and those in turn determine the overlaps, one derives mean-field equations for the attractors, for low loading:

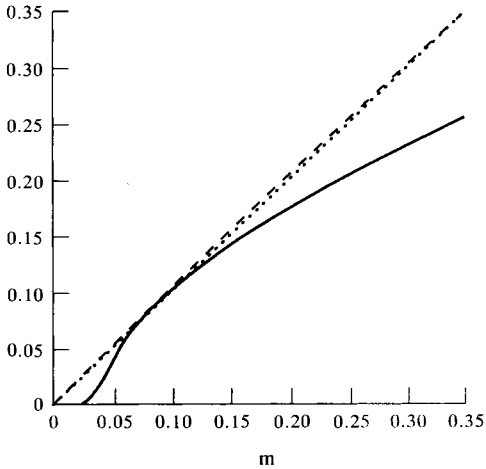


Fig. 6. The two sides of equation 18 for several values of σ. For param: $\theta = 0.25$, $\mu = 0.2$, short-dashed $\sigma = 0.0025$; long-dashed $\sigma = 0.0125$; dot-dashed $\sigma = 0.05$. Both axes measure essentially the spike frequency in units of the saturation. (From Amit *et al.*, 1991.)

$$m^{\mu} = \langle\langle(\eta_i^{\mu} - f)\phi[\sum_{\mu=1}^{p}(\eta_i^{\mu} - f)m^{\mu}]\rangle\rangle. \tag{17}$$

When considering retrieval of a single pattern (for instance the first one: $m^{\mu} = m\delta^{\mu 1}$) at low memory loading level, the mean-field equations reduce to (Amit *et al.*, 1991):

$$m = \phi(m) - \phi\left[-\frac{fm}{(1-f)}\right] \tag{18}$$

Retrieval of memory is achieved when neurons in foreground have an activity distribution around high values, well separated from the activity distribution of the neurons in background, which have low activity rates. In Fig. 6 the two sides of the mean field equation are plotted as functions of m for different values of σ, the level of noise. Any intersection of the two curves constitutes a possible solution of the mean field equation. There are two solutions with $m \neq 0$. The higher one is the stable solution while the lower one is unstable. So the minimum stable rate is about above 50–55 spikes per second and it seems quite difficult to lower these frequencies down to 10–20 spikes per second. This may imply that the neurons are still oversimplified.

Storage capacity and activity distributions of foreground and background neurons

It is possible to calculate, with the tricks of replica method, the structure of the attractors at high loading ($p \approx N$) and in particular the storage

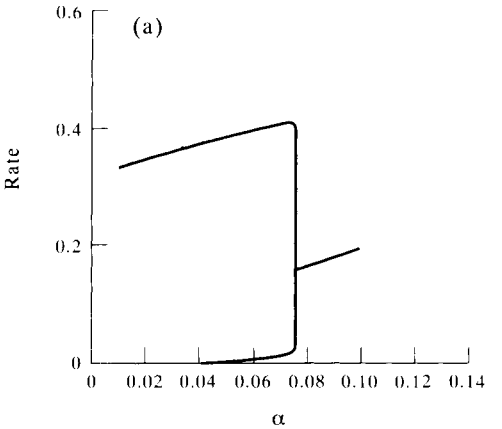


Fig. 7. Mean rates in foreground (upper curves) and background (lower curves), for networks with gain function (14), as a function of the loading level, α. (From Amit *et al.*, 1991.)

capacity. In Fig. 7 one sees, as a function of the loading $\alpha = p/N$, the mean frequencies of the foreground (upper curves) and the background neurons (lower curves) for the gain function $\phi(I)$ of equation (14).

In fact, it is even possible to compute analytically the distributions of frequencies among the neurons in the network. In Fig. 8 there are two activity histograms corresponding to two different cases: in case (a) α is below saturation while in case (b) it is above the critical value. In the first case the background neurons populate very low frequencies and their distribution is very sharp around zero. The distribution of activity among foreground neurons is well separated from that of the background neurons and far below the saturation. In case (b), when α is above its critical value, the two distributions merge and no retrieval is possible. When α reaches its critical value, the two frequency distributions become, discontinuously, indistinguishable. This means that in each class, part of the neurons have high rates, while the others are almost quiescent. The selection of which neurons in each class have high rates and which are quiet is totally uncorrelated with the stimuli, or with the stored patterns (see Fig. 8, case (b)).

Simulations

Numerical simulations are very useful to check the results of the mean field equations and to control the accuracy of the assumption of replica symmetry as well as the effects of the finite size of the network. The dynamics of equation (12) was simulated with the synaptic matrix (15) and the gain function (14) for a network of 1000 neurons. The stimulus was presented as an external afferent current to foreground neurons. Then, after a certain time, this current is removed and the relevant quantities are moni-

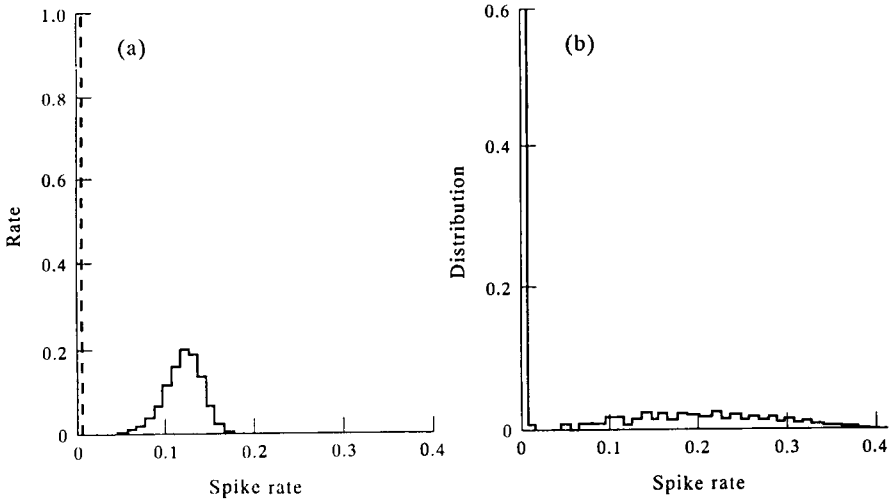


Fig. 8. Activity histograms in attractors, in mean-field theory. (a) Below saturation, retrieval. Full curves represent the distribution of neurons in foreground while dashed lines are neurons which should be quiet. (b) Above saturation: the two distributions merge. The gain function is integrate-and-fire. (From Amit *et al* ., 1991.)

tored as a function of time. The results obtained by simulations are essentially the same as the ones predicted by mean field theory. In Fig. 9 the distributions of foreground and background neurons are plotted in the case of α below saturation. These curves are very similar to the curves of Fig. 8.

As α increases, some general features in which simulations differ from predictions of mean-field theory present themselves. One of the main differences is that the transition to the frozen state is smeared out by fluctuations. Moreover, the critical capacity α_c is greater than in the mean field theory and the distributions of foreground and background neurons are not identical when α is above its critical value. This means that there exists a remnant memory even if the network is overloaded, similar to remnant magnetization in spin-glasses.

It is also possible to study in detail the time development of some relevant quantities. For example, in Fig. 10, we plot the time development of the mean spike rate among the neurons in the foreground. When the stimulus is active the mean rate grows up very quickly. After 50 ms the stimulus is removed and the activity of the neurons decreases to one sixth of the maximum mean rate reached in presence of the input. The convergence to the attractor is very fast: 20–30 ms.

The simulations have also checked the robustness of the network to asymmetric dilution of the synapses. The synaptic matrix was diluted (50% dilution) and 30% errors were introduced in the stimulus. After a few

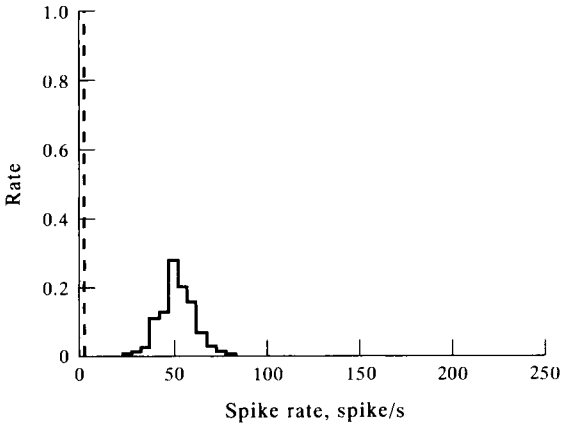


Fig. 9. Simulations: activity distributions in retrieval attractors (α = 0.08), in simulations. Full curve: neurons in foreground; dashed: neurons in background. The gain function is integrate-and-fire. (From Amit *et al.*, 1991.)

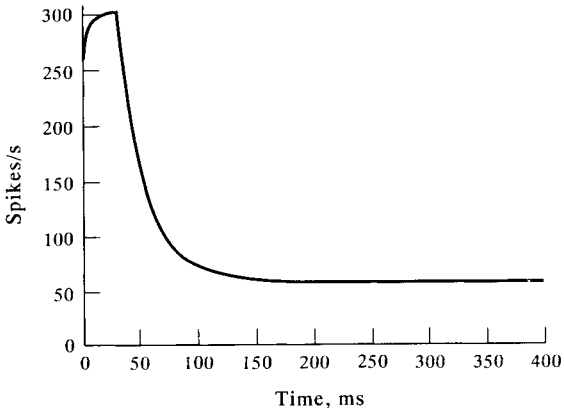


Fig. 10. Time development of mean rate in the network. The stimulus persists until the activity starts to drop ($t \simeq 50$ ms). (From Amit *et al.*, 1991.)

milliseconds the separation into an active class and an inactive class is already attained. When the stimulus is removed the mean frequencies begin to decrease but the two distributions remain well separated. So the network operates very well as auto-associative memory, correcting the errors in the stimulus extremely rapidly and it is robust to the random asymmetric dilution of synaptic matrix.

Concluding remarks

The reader should be left with the feeling that the discussion has stopped midway. There is a large number of interesting investigations which should be continued along this particular path. Some are already under way. For example, questions like storage capacities of networks storing correlated attractors; behavior of realistic networks with correlated attractors; predictions of detailed cognitive and neurophysiological experiments from the models; implications for learning mechanisms etc.

This chapter will leave here with the hope that it has at least communicated the sense that neural network modeling, post Miyashita's monkey, is a productive integrated scientific field with a method and an empirical reference, both in neurobiology as in cognitive psychology.

Acknowledgments

I would like to express my gratitude for the special experience that the Capri International School on Neural Modeling and Neural Networks has provided. These thanks go first to Professor Francesco Ventriglia who made it possible. But the interactions, especially with Drs Roger Traub and Walter Freeman, and not less with the special group of students that attended, have been exceptionally gratifying. We have benefited from the comments of N. Brunel on an earlier version of this manuscript.

The manuscript was prepared by Stefano Fusi (INFN, Sezione di Roma, Istituto di Fisica dell'Università di Roma, La Sapienza).

12

Pattern Recognition with Neural Networks

KUNIHIKO FUKUSHIMA

Department of Biophysical Engineering, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan

Introduction

Even the most advanced electronic computers today exhibit far less ability than human beings when they are required to recognize, say, visual patterns. In order to design an information processing system which has higher functions like the human brain, it is important to understand the mechanism of the brain and learn from the real biological basis of it. One of the largest attracting powers of the artificial neural networks or neurocomputers today resides in their learning ability. They are expected to acquire the ability to do anything if they have been properly trained beforehand. When training an artificial neural network, however, the function of generalization is indispensable. Let us consider the case of training the network to recognize visual patterns such as handwritten characters. If the network lacks the ability to generalize, all of the deformed versions of the patterns have to be taught to the network during the training. It is not guaranteed that patterns which have not been shown before will be recognized correctly. Since an infinite number of deformed patterns can exist, it is almost impossible to teach all of them beforehand. On the other hand, if the network has the ability to generalize, it is sufficient to train the network using typical examples of deformed patterns only. The rest of the deformed patterns can be recognized correctly even though they have not been shown during the training. In order to let the network have the same function of generalization as human beings, it is essential to design the network with

the same architecture as that of the real biological brain. The biological brain, however, has a complicated network architecture and is difficult to understand using only conventional neurophysiological and psychological experiments. Hence, modeling neural networks plays an important role. In the modeling approach, we study how to interconnect neurons to synthesize a brain model, or a network with the same functions and abilities as the brain. When synthesizing a model, we try to follow physiological evidence as faithfully as possible. For parts that are not yet clear, we construct a hypothesis and synthesize a model that follows the hypothesis. We then analyze or simulate the behavior of the model, and compare it with that of the brain. If we find any discrepancy in behavior of the model and the brain, we change the initial hypothesis and modify the model. We then test the behavior of the model again. We repeat this procedure until the model behaves in the same way as the brain. Although we must still verify the validity of the model by physiological experiment, it is probable that the brain uses the same mechanism as the model, because both respond in the same way. The relationship between modeling neural networks and neurophysiology resembles that between theoretical physics and experimental physics. Once we complete a model, we can easily see the essential algorithms of information-processing in the brain, because nonessential factors in the biological brain are discarded in the model. We can use the algorithms directly as design principles for new information-processors. Many models have been constructed along these lines. Among them, some models of visual information processing proposed by the author are introduced in this chapter.

Contrast detection by the retina

Model of on- and off-center cells

Retinal ganglion cells, namely, the output cells of the retina, have concentric on-center or off-center receptive fields. The on-center receptive field consists of an excitatory circular region surrounded by an inhibitory annular region, while the off-center receptive field has the opposite structure. These cells are supposed to extract contrast components from patterns projected onto the retina. Photoreceptors and ganglion cells are not directly connected in an actual biological retina. There are several kinds of cells in between, and information is processed by the interactions of such cells. However, we regard the retina as a "black box", and consider a two-layered neural network (Fukushima, 1971), as in Fig. 1. Let layer U_0 correspond to the layer of photoreceptors, and layer U_1 to the layer of on-center ganglion cells. Analog-threshold elements, whose inputs and outputs take non-negative analog values, are employed as the cells of layer U_1. The spatial distribution of the strength of the input connections of a single cell of

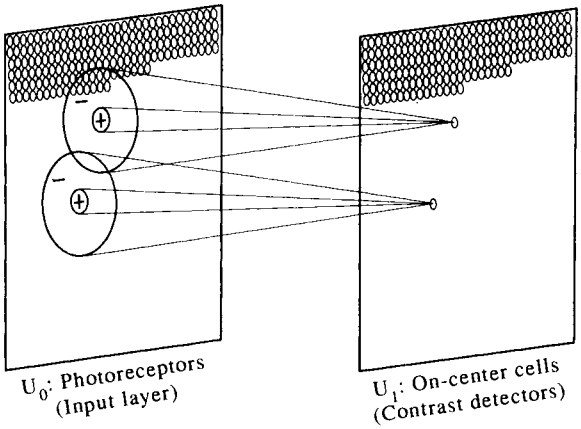
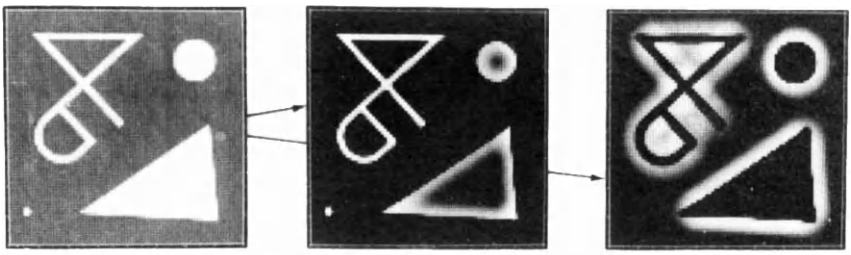


Fig. 1. A two-layered neural network model of the retina (Fukushima, 1976).



U_0: Photoreceptors U_1: On-center cells Off-center cells

Fig. 2. An example of the response of the model of the retina (Fukushima, 1976).

U_1 is positive (or excitatory) at the center, and negative (or inhibitory) in the surrounding area, as shown in Fig. 1. We can simply invert the positive and the negative polarity of the connections in order to make a model for off-center cells. Figure 2 is an example of a computer simulation of this model. As can be seen from this figure, the effect of the uniform background light in the input pattern is completely eliminated at the output layer U_1, and only the contrast component is extracted. The model also shows some psychologically known effects, such as the Mach-band.

Electronic retina

The author and his group constructed a similar model with an electronic network, too (Fukushima, 1970a), (Yasuda *et al.*, 1971). Since it was in 1970 when we did not have the techniques to build VLSI yet, the model was constructed using discrete electronic components. The photoreceptor layer, or input layer, of the model consisted of an array of about 700 silicon solar cells. The output layer, or on-center cell layer, also contained an array of about 700 cells, whose outputs were diplayed by the brightness of the array

of lamps of the same number. If we were to construct a model of Fig. 1 with an electronic network of this scale in a straightforward way, in which each connection between photoreceptors and on-center cells is replaced with one discrete electronic component, a prohibitively large number of connecting elements would be necessary because each on-center cell requires input connections from a huge number of photoreceptors. We developed a new technique to reduce the number of connecting elements. In this method, triangular meshes of resistors are used to make the interconnection. Each node of a resistor mesh is driven by the signal from a photoreceptor. A signal fed to a certain node propagates laterally through the resistor mesh and attenuates with distance. Thus, the response of the photoreceptor layer is blurred. In our model, two such resistor meshes were used: one with a small blurring constant for the excitatory interconnections, and the other with a large blurring constant for the inhibitory interconnections. Each on-center cell, which was actually a differential amplifier, received its inputs antagonistically from these two resistor networks. As the result, on-center receptive fields were created quite economically. A similar technique was used recently by Mead *et al.* to construct a VLSI chip of the retina (Mead and Mahowald, 1988).

Feature extraction in the visual cortex

In early 1960s, Hubel and Wiesel classified the cells of the visual area by the function of their receptive fields into simple cells, complex cells, hypercomplex cells, and so on (Hubel and Wiesel, 1962, 1965). They hypothesized that these classes of cells are connected in a hierarchical manner: retina → LGB (relay cells between the retina and the visual cortex) → simple cells → complex cells → hypercomplex cells. In the 1960s and early 1970s, various kinds of models of the visual system were constructed following this hypothesis (Fukushima, 1976). Figure 3 shows an example of such a model. It is a multi-layered network consisting of a cascade connection of many layers of cells (Fukushima, 1971, 1970). The structure of the connections between the input layer U_0 and the contrast-detection layer U_1 is identical to that of the model shown in Fig. 1. The spatial distribution of the input connections of the cells of individual layers is shown in Fig. 3. The response of the cells of the network to a heart-shaped pattern is shown in the figure. The input pattern is decomposed into line segments of different orientation in layer U_2. The curved components of the input pattern are extracted in layer U_5 in the highest stage.

Learning and self-organization

Engineers in pattern recognition often classify the process of self-organization into "supervised learning" (or learning-with-a-teacher) and

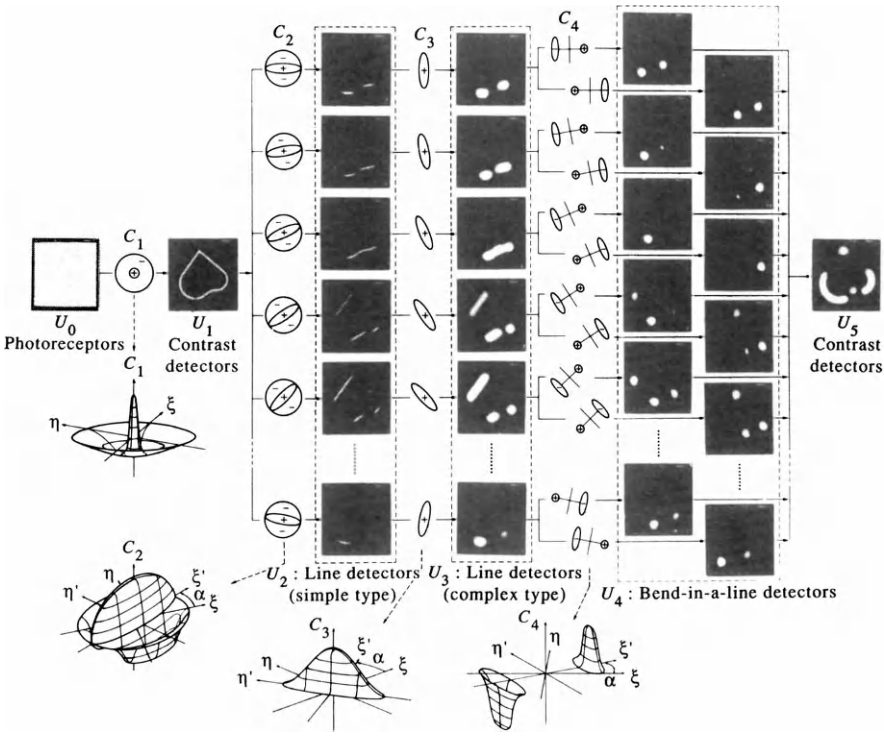


Fig. 3. A multilayered network capable of extracting features from curvilinear patterns (Fukushima, 1976) (modified from (Fukushima, 1971)).

"unsupervised learning" (or learning-without-a-teacher). A famous example of a classical network which can be trained by supervised learning is the three-layered "perceptron" (Rosemblatt, 1962). Although the perceptron has a three-layered hierarchical structure, it is only in the highest stage that the cells have variable input connections. The connections to the cells in other layers are all fixed and unmodifiable. It is well known that the capability of a hierarchical multilayered network to process information is greatly improved by increasing the number of layers in the network. In order to give full play to the capability of a multilayered network, however, the input connections of the cells must be modifiable not only in the highest stage, but also in the intermediate lower stages. An example of supervised learning useful for such a purpose is "back propagation" (Rumelhart, 1986). In the model called the "cognitron" (Fukushima, 1975a), (Fukushima, 1981), which the author proposed earlier, unsupervised learning is used.

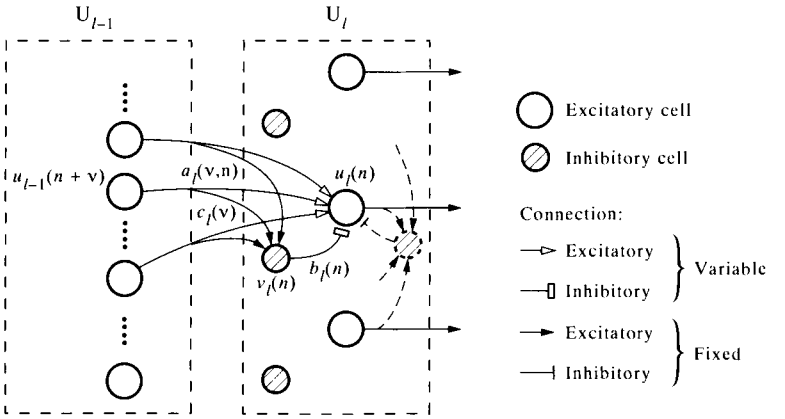


Fig. 4. The structure of the cognitron. The interconnection between two adjoining layers U_{l-1} and U_l is shown, where U_l represents the lth layer in the hierarchical network (Fukushima, 1975a).

Unsupervised learning in the cognitron

The cognitron consists of a number of neural layers of similar structure, connected in a hierarchical manner. Each layer has a number of excitatory cells and the same number of inhibitory cells. These cells are connected with each other as shown in Fig. 4. The figure shows interconnection between two adjoining layers only, because the network has the same structure in all stages. The input connections of excitatory cells are variable, and are reinforced by unsupervised learning. The repeated presentation of a set of training patterns to the input layer is enough for the self-organization of the network. It is not necessary to give any information about the categories in which the training patterns should be classified. During the repeated presentation of training patterns, each cell (strictly speaking, each excitatory cell) in the highest stage of the network gradually comes to respond selectively to only one of the training patterns. Self-organization of the cognitron is performed with a kind of competitive learning procedure: Among the cells situated in a certain small area, only the one responding most strongly has its input connections reinforced. The amount of reinforcement of each input connection to this maximum-output cell is proportional to the intensity of the response of the cell from which the relevant connection leads. This principle is applied to both excitatory and inhibitory connections. Figure 5 illustrates the process of reinforcement, showing only the connections converging on an excitatory cell u_l of the lth layer. The u_l-cell receives variable excitatory connections from a group of excitatory cells u_{l-1} in the preceding layer. The cell also receives a variable inhibitory connection from a subsidiary inhibitory cell v_l. The v_l-cell receives fixed excitatory connections from the same group of u_{l-1}-cells as does the u_l-

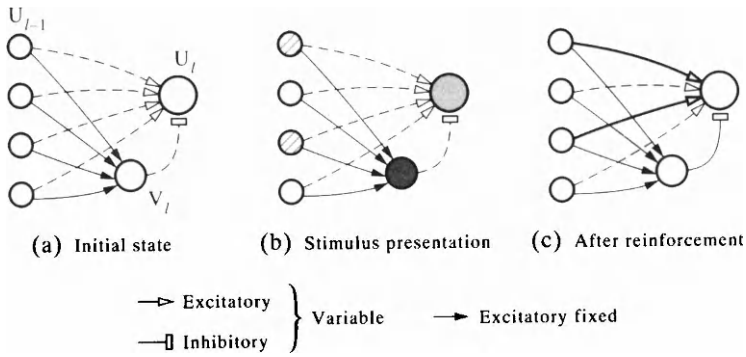


Fig. 5. The process of reinforcement of the input connections converging to an excitatory cell (modified from (Fukushima, 1988)). The density of the shadow in a circle represents the intensity of the response of the cell.

cell, and always responds with the average intensity of the output of the u_{l-1}-cells. The initial strength of the variable excitatory connections is very weak but not completely zero. If it were completely zero, no cell would respond to training stimuli, and the maximum-output cell could not be determined. Suppose the u_l-cell shown in Fig. 5(b) responds the strongest of the u_l-cells in its vicinity when a training stimulus is presented. According to the principle mentioned above, variable connections leading from activated u_{l-1}- and v_l-cells are reinforced as shown in Fig. 5(c). The variable excitatory connections to the u_l-cell grow into a "template" which exactly matches the spatial distribution of the response of the u_{l-1}-cells of the preceding layer. The inhibitory variable connection from the v_l-cell is reinforced at the same time, but not as strongly because the output of the v_l-cell is not as large. After the training (or learning) is complete, the u_l-cell acquires the ability to extract a feature of the stimulus presented during the training period. Through the excitatory connections, the u_l-cell receives signals indicating the existence of the feature to be extracted. If an irrelevant feature is presented, the inhibitory signal from the v_l-cell becomes stronger than the direct excitatory signals from u_{l-1}-cells, and the response of the u_l-cell is suppressed. Because of the "winner-take-all" nature of this principle, the formation of more than one cell in a small area to extract the same feature does not occur, and "division of labor" among the cells occurs automatically. In other words, the formation of a redundant network can be prevented. With this principle, the network also exhibits self-repair. If a cell which has responded strongly to a stimulus is damaged and ceases to respond, another cell, which happens to respond more strongly than the others, starts to develop and substitutes for the damaged cell. Until then, the larger response of the first cell had prevented the development of the second cell.

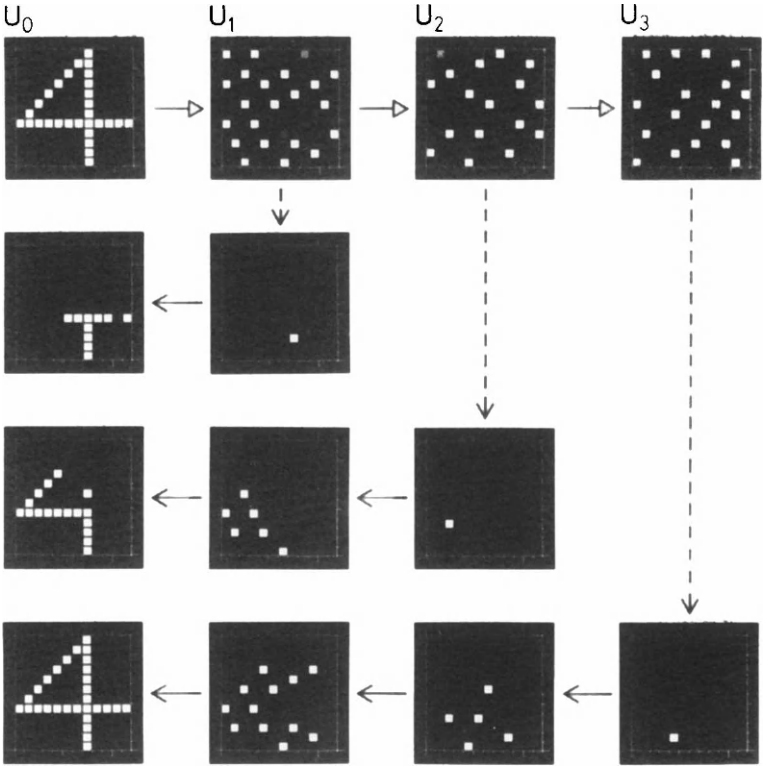


Fig. 6. An example of the response of the cognitron, and the reverse reproduction from single cells (Fukushima, 1975a). The first line shows the normal response to stimulus "4". The second, third and fourth line show reverse reproduction from single cells of layers U_1, U_2 and U_3, respectively.

Computer simulation of the cognitron

The cognitron has been simulated on a digital computer. During the process of self-organization, five stimulus patterns "0", "1", "2", "3" and "4" were presented to the input layer U_0 repeatedly and in a cyclic manner. Twenty cycles of pattern presentations were enough to organize the network. Each stimulus pattern elicits its own response pattern from each layer of the network. Most of the cells, especially the cells in layer U_3 at the highest stage, respond selectively to a single stimulus pattern. The picture at the top line of Fig. 6 shows how the (excitatory) cells in each layer of the network respond to stimulus pattern "4". In order to visualize how the connections have been organized in the network, an experiment called "reverse reproduction" has been made. The reverse reproduction tests how the cells would respond if the direction of all the signal flows in the network were reversed. As mentioned above, the pictures in the top line of

Fig. 6 show the normal response to stimulus "4". At first, we chose from layer U_1 a cell which responds strongly to pattern "4". (It does not matter which cell is actually chosen. We can get a similar results even if another cell is chosen, provided that the cell responds strongly to the stimulus). The pictures in the second line of the figure show the result of reverse reproduction from this u_1-cell. An excitatory response is elicited from the u_0-cells which, in the normal state, have an excitatory effect on this u_1-cell. In other words, the excitatory part of the receptive field of the u_1-cell is visualized. Since the u_1-cell only has a small receptive field at the lower right portion of layer U_0, only the lower right-hand part of the pattern "4" is reproduced. The third and the fourth line of the figure show reverse reproduction from single u_2- and u_3-cells which respond strongly to the pattern "4". We can see that the chosen u_3-cell has converging paths which exactly match the pattern "4" from the input layer U_0.

Neocognitron: deformation-invariant recognition

Outline of the neocognitron

Although the cognitron can acquire the ability to recognize patterns through the process of learning, the cognitron, like many other models, does not have the ability to correctly recognize shifted or distorted patterns . The conventional cognitron usually recognizes the same pattern presented in a different position as a completely different pattern. The author proposed the "neocognitron" (Fukushima, 1980, 1988a; Fukushima and Miyake, 1982) to eliminate this defect. The neocognitron can recognize stimulus patterns correctly, even if the patterns are shifted or distorted. It can acquire this ability through the process of learning. Since the neocognitron has the ability to generalize, it is not necessary to teach all the deformed versions of the training patterns. It even is able to correctly recognize correctly a pattern which has not been presented before, provided it resembles one of the training patterns. The neocognitron, like the cognitron, is a hierarchical multilayered network and consists of a cascade of many layers of cells, as shown in Fig. 7. The initial stage of the network is the input layer U_0, which consists of a two-dimensional array of receptor cells. Each of the succeeding stages has a layer of "S-cells" followed by a layer of "C-cells". Thus, in the entire network, layers of S-cells and C-cells are arranged alternately. The S-cells are feature-extracting cells, and somewhat resemble the simple cells of the visual cortex. They have variable input connections, and can acquire the ability to extract features by learning (or training) in the same way as the excitatory cells of the cognitron. After learning process is completed, an S-cell is activated only when a particular feature is presented in a certain position in the input layer. The features

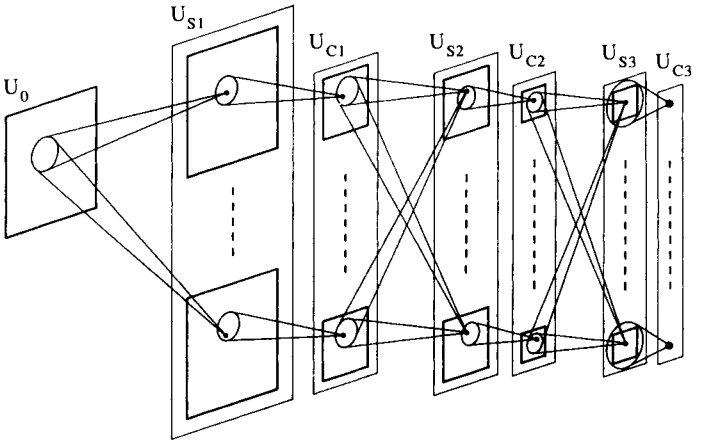


Fig. 7. Hierarchical network architecture of the neocognitron (Fukushima, 1980). Notation U_{Sl}, for example, is used to indicate the layer of S-cells of the lth stage.

extracted by the S-cells are determined during the learning process. Generally speaking, local features, such as a line at a particular orientation, are extracted in the lower stages. More global features, such as part of a training pattern, are extracted in the higher stages. C-cells somewhat resemble complex cells. They are put into the network to allow for positional error in the features extracted by S-cells. The connections from S-cells to C-cells are fixed and invariable. Each C-cell receives signals from a group of S-cells which extract the same feature, but from slightly different positions. The C-cell is activated if at least one of these S-cells is active. Even if the stimulus feature is shifted and another S-cell is activated instead of the first one, the same C-cell keeps responding. Hence, the C-cell's response is less sensitive to shifts of the input pattern. We can also say that the spatial response of a layer of S-cells is blurred at the succeeding layer of C-cells. The density of cells in each layer is designed to decrease with the order of the stage, because each cell in a higher stage usually has a larger receptive field, and neighboring cells receive similar signals owing to the overlapping of their receptive fields. The layer of C-cells at the highest stage works as the recognition layer: the response of the cells in this layer shows the final result of pattern recognition by the neocognitron. Each layer of S-cells or C-cells is divided into subgroups according to the features to which they respond. The connections converging to the cells in a subgroup are homogeneous and topographically ordered: all the cells in a subgroup have input connections of the same spatial distribution, where only the positions of the preceding cells from which the connections lead are shifted in parallel with the positions of the cells in the subgroup. In other words, the connections have a translational symmetry. This condition of homogeneity holds not

only for fixed connections, but also for variable connections. In order to reinforce the variable input connections of S-cells under this condition, another new principle of reinforcement is introduced together with the one previously employed in the cognitron: The maximum-output cell not only grows by itself, but also controls the growth of neighboring cells, working, so to speak, like a "seed" in crystal growth. The neocognitron can also be trained by supervised learning as well as by unsupervised learning. Supervised learning is useful when we want to train a system to recognize, for instance, hand-written characters which should be classified not only on the basis of similarity in shape but also on the basis of certain conventions. In the case of supervised learning, the "teacher" presents training patterns to the network and points out the positions of features which should be extracted. The cells which have receptive fields at the positions are chosen as "seed cells". The other process of reinforcement is identical to that of unsupervised learning, and occurs automatically.

Principles of deformation-invariant recognition

In the entire network, with its alternating layers of S-cells and C-cells, the process of feature-extraction by the S-cells and the toleration of a shift by the C-cells is repeated, as shown in Fig. 8. During this process, local features extracted in lower stages are gradually integrated into more global features. Finally, each C-cell of the recognition layer at the highest stage integrates all the information of the input pattern; each cell responds only to one specific pattern. In other words, only one cell, corresponding to the category of the input pattern, is activated. Other cells respond to patterns of other categories. Tolerating positional errors a little at a time at each stage, rather than all in one step, plays an important role in endowing the network with the ability to recognize even distorted patterns. Let an S-cell in an intermediate stage of the network have already been trained to extract a global feature consisting of three local features of a training pattern 'A', as shown in Fig. 9(a). The cell tolerates a positional error of each local feature if the deviation falls within the dotted circle. Hence, this S-cell responds to any of the deformed patterns shown in Fig. 9(b). The toleration of positional errors should not be too large at this stage. If large errors are tolerated in one step, the network may come to respond erroneously, such as recognizing a stimulus like Fig. 9(c) as an 'A' pattern.

Handwritten character recognition

In order to demonstrate the ability of the neocognitron, the author and his group have designed several systems which recognize hand-written characters: for example, a system recognizing ten numerals (Fukushima, 1988a), and a system recognizing alphanumeric characters (Fukushima and

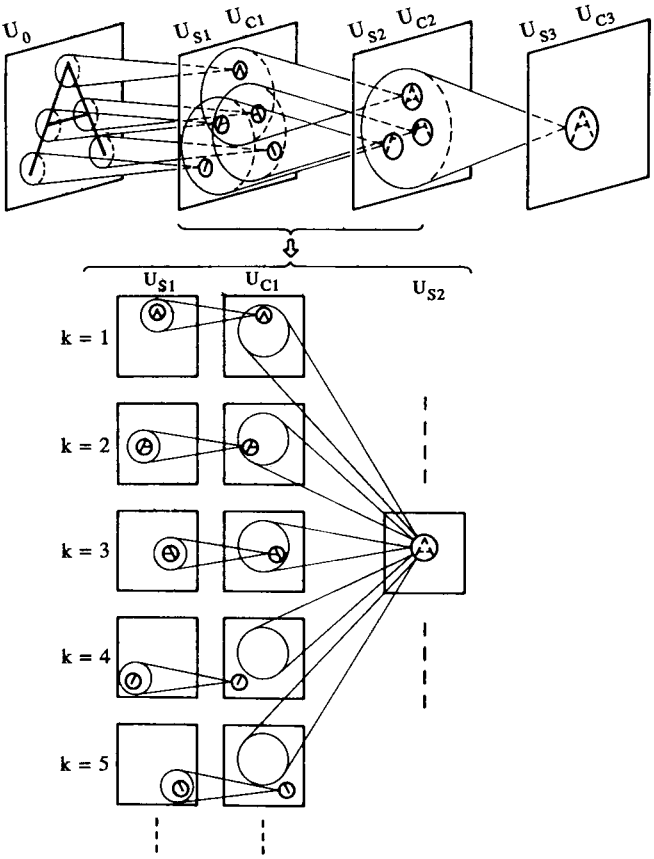


Fig. 8. Illustration of the process of pattern recognition in the neocognitron (Fukushima, 1980).

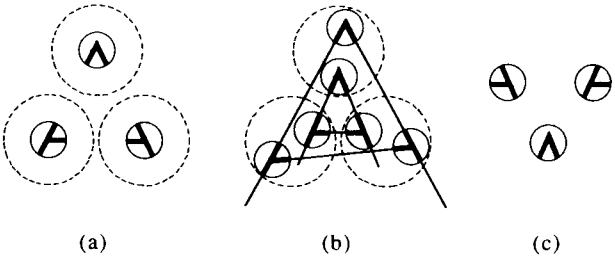


Fig. 9. Illustration of the principle for recognizing deformed patterns (Fukushima, 1988).

Wake, 1991). These systems have been trained by supervised learning. Figure 10 shows how the alphanumeric character recognition system responds

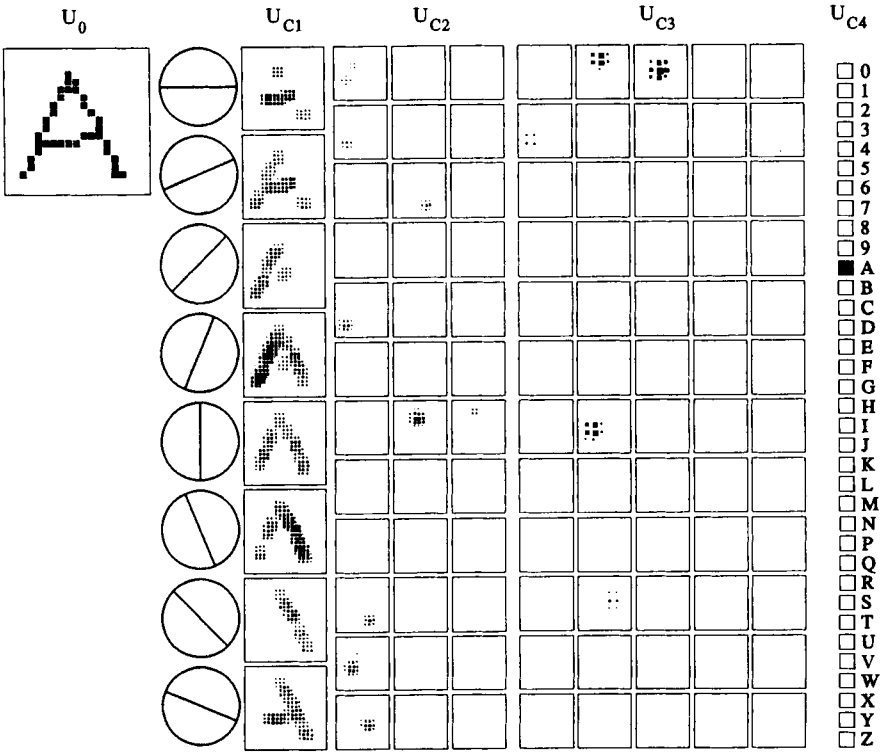


Fig. 10. An example of the response of the network. Only the response of the C-cells are displayed (Fukushima and Wake, 1991).

when pattern 'A' is presented to the input layer U_0. The figure shows only the response of C-cells in the network. In the recognition layer U_{C4}, shown at the extreme right, only one cell corresponding to 'A' is activated. Figure 11 shows some examples of deformed alphanumeric characters which the system has recognized correctly. As can be seen from the figure, the neocognitron recognizes input patterns robustly, with little effect from deformation, changes in size, or shifts in position. Even though the input patterns have some parts missing or are contaminated by noise, the neocognitron recognizes them correctly.

Improvement of the neocognitron

Many approaches are now in progress to improve the neocognitron. This section introduces some of those approaches. As has been discussed above, the conventional supervised learning can endow the neocognitron with the ability of robust pattern recognition, if a good training pattern set

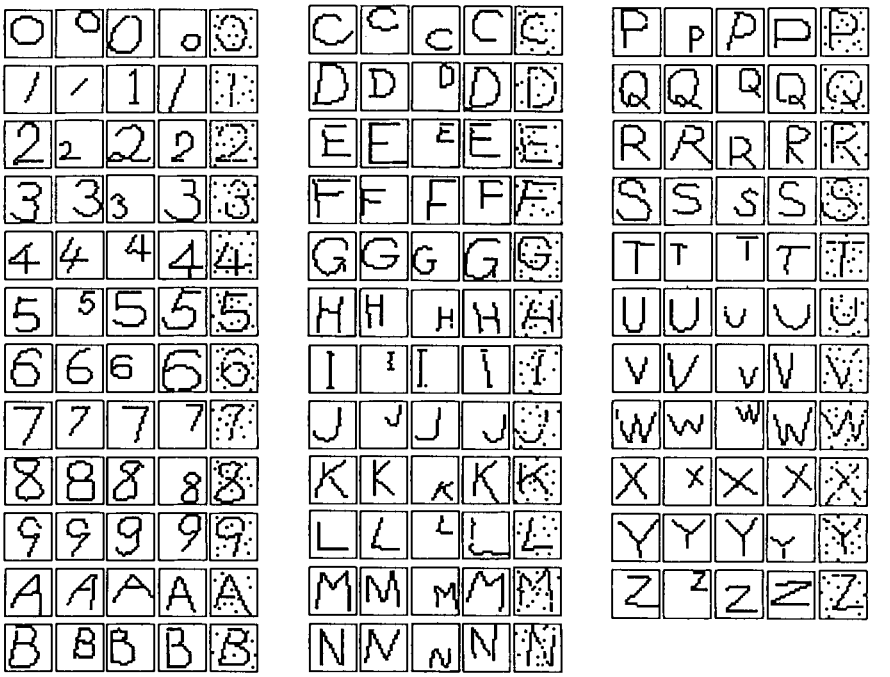


Fig. 11. Some examples of deformed alphanumeric characters which the neocognitron has recognized correctly (Fukushima and Wake, 1991).

is skillfully chosen. However, a great amount of labor is required to construct a good training pattern set, because the "teacher" is required to tell not only the categories of the training patterns but also the positions at which important features exist. On the other hand, the conventional algorithm for unsupervised learning (Fukushima and Miyake, 1982), by which all of the training progresses automatically, produces a somewhat lesser ability to recognize deformed patterns. In order to construct a neocognitron which can be easily trained to have a powerful ability to recognize patterns, improvements have been made from two sides (Wake and Fukushima, 1992). One is the improvement of the network architecture (Fukushima and Wake, 1992a), and the other is the development of a new learning algorithm which does not require a significant amount of labor (Fukushima and Wake, 1992b). The neocognitron with the new architecture (Wake and Fukushima, 1992; Fukushima and Wake, 1992a) has six stages of S- and C-cell layers. It has built-in edge-extracting cells, line-extracting cells, and bend-detecting cells. Layer U_{S1} consists of edge-extracting S-cells. The S-cells of layer U_{S2} extract line components using the edge information extracted in U_{S1}. Although a change in thickness of a line causes a shift in position of both edges, the effect of the positional shift of the edges is absorbed by the blur-

ring operation in the C-cells of U_{C1}. The S-cells of U_{S2} can then extract the line with little effect from the variation in thickness. The bend-detecting cells receive signals from U_{C2}, which responds to line segments, and detect the bend-points and endpoints of the lines. The structure of this part of the network is the same as the one used in the curvature-extracting network. The layers of bend-detecting cells make a branch in the hierarchical network, and the S-cells in the next stage receive input connections from both line-extracting cells and bend-detecting cells. The S-cells in higher stages of the network have variable input connections, which are modified through learning. In contrast to the conventional neocognitron, the network using this architecture shows considerable robustness even when trained by unsupervised learning. The amount of accepted deformation is quite large compared to the conventional neocognitron trained by unsupervised learning. Supervised learning is still required, however, if we want to construct a multi-font recognition system which can recognize not only deformed patterns, but also various patterns of different styles of writing. The new learning algorithm (Wake and Fukushima,1992; Fukushima and Wake, 1992b) combines unsupervised and supervised learning. Variable connections are reinforced by unsupervised learning at first, and are modified afterward by supervised learning if the result of the recognition is wrong. This algorithm does not require a significant amount of labor from the "teacher", who is required to tell only the category of the training pattern he is presenting. Backward signal paths similar to the ones used in the "selective attention model" discussed later have been added to the neocognitron to be used only for the training process. The cells in the backward paths are arranged in a mirror image of the cells in the forward paths. The forward and the backward connections also make a mirror image to each other, but the directions of signal flow through the connections are opposite. After finishing the unsupervised learning, learning by error-correction is begun. Test patterns of various styles of writing are presented to the network, and the response of the recognition layer is observed. If the result of recognition is incorrect, the desired output is fed to the highest stage of the backward paths. In intermediate stages of the network, backward signals from the desired recognition cell are compared with the forward signals from the training pattern, as shown in Fig. 12. If discrepancies between the two signals are detected, new feature-extracting cells are automatically generated in the network. Operations which used to be manually performed can now progress automatically.

Selective attention model

Although the neocognitron has a considerable ability to recognize deformed patterns, it does not always recognize patterns correctly, when two

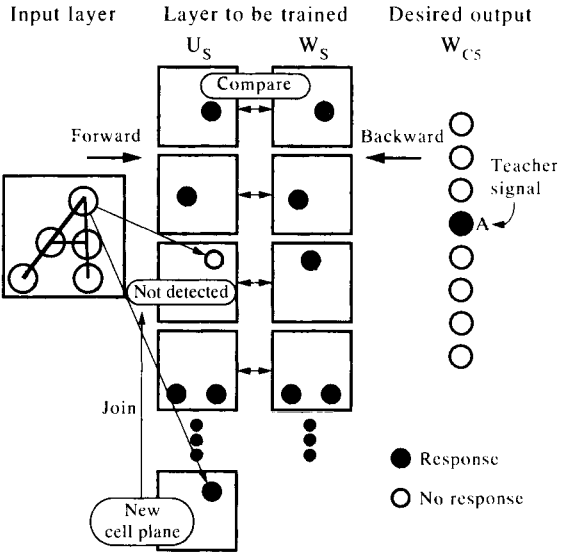


Fig. 12. Supervised learning using backward signals (Fukushima and Wake, 1992b).

or more patterns are presented simultaneously. In order to increase the ability of the neocognitron, backward (i.e. top-down) connections were added to the conventional neocognitron which had only forward (i.e. bottom-up) connections. This new model has the function of selective attention in visual pattern recognition (Fukushima, 1986b, 1987a, 1988). When a composite figure consisting of two or more patterns is presented to the model which has finished learning, the model selectively focuses its attention on one pattern after another, segments each pattern from the others, and recognizes them separately. This model also has the function of associative recall. Even if noise or defects affect the stimulus pattern, the model can recognize it and recall the complete pattern for which the noise has been eliminated and defects corrected. Perfect recall does not require that the stimulus pattern be identical in shape to the training pattern. A deformed pattern, or one shifted in position can be recognized correctly and the missing portions restored.

Network architecture

Figure 13 illustrates schematically the signal flow in the network of the selective attention model. In this diagram, layers of cells in the forward paths and the backward paths have been drawn separately. Figure 14 illustrates how the cells, represented by circles, are connected with each other in the network. Letters u and w represent the cells in the forward paths

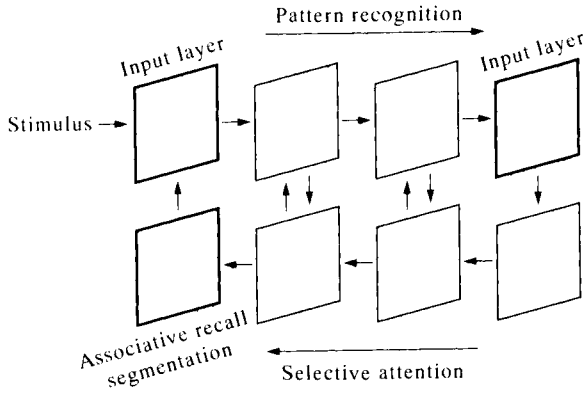


Fig. 13. The forward and backward signals and their interaction in the hierarchical network of the selective attention model (Fukushima, 1988).

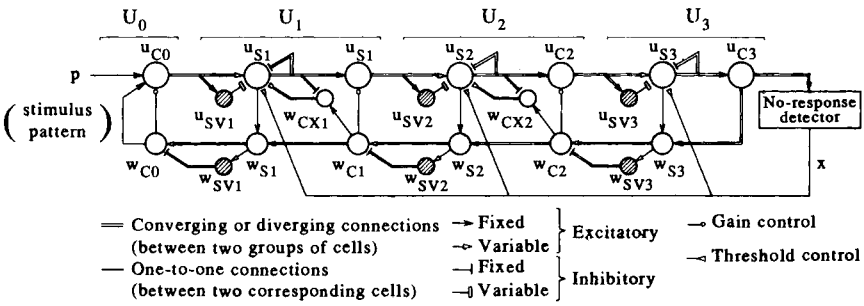


Fig. 14. Hierarchical structure of the interconnections between different kinds of cells (Fukushima, 1987a).

and the backward paths, respectively. Although the figure shows only one of each kind of cell in each stage, there are actually numerous cells in two-dimensional arrays as illustrated in the detailed diagram of Fig. 15. In the hierarchical network, the forward signals manage the function of pattern recognition. The forward path of the network has almost the same structure and function as the neocognitron. Cells u_S and u_C in the forward path correspond respectively to the S- and C-cells of the neocognitron. The lowest stage of the forward path is the input layer U_{C0}, and the highest stage is the recognition layer U_{CL} ($L = 3$ in the case of the figure). The cells in the backward paths are arranged in the network making a mirror image of the cells in the forward paths. The same is true also between the forward and the backward connections. The output of the recognition layer is returned to the lower stages through the backward paths. The forward and backward signals interact. The forward signals gate backward signal flow, and, at the same time, the backward signals facilitate forward signal flow. The backward signals manage the function of selective attention,

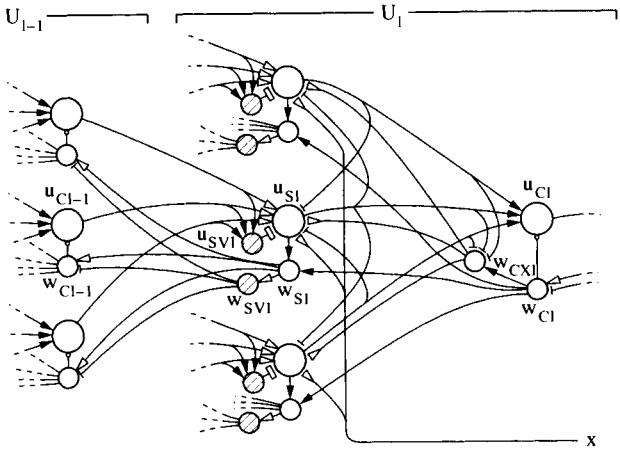


Fig. 15. Detailed diagram illustrating spatial interconnections between neighboring cells (Fukushima, 1987a).

pattern-segmentation and associative recall. The lowest stage of the backward paths is the "recall" layer W_{C0}, in which the result of associative recall appears. We can also interpret the output of the recall layer as the result of segmentation. The response of the recall layer is fed back to the input layer.

Backward flow of signals

The network is so designed that the backward signals, which are sent back from the recognition layer, flow retracing the same route as the forward signals. The backward signal from an arbitrary w_{Sl}-cell of the lth stage is simply transmitted to w_{Cl-1}-cells in the lower stage through direct excitatory backward connections and indirect inhibitory connections via inhibitory w_{SVl}-cell (see Fig. 15). Since the backward connections have been reinforced to make a mirror image of the forward paths from u_{Cl-1}- to u_{Sl}-cells, the backward signals retrace the same route as the forward signals. (This situation resembles the reverse reproduction experiment for the cognitron.) In order to make backward signals from w_{Cl}- to w_{Sl}-cells retrace the same route as the forward signals, however, the mirror-imaged network structure alone is not enough. Corresponding to the fixed forward connections which converge to a u_{Cl}-cell from a number of u_{Sl}-cells, many backward connections diverge from a w_{Cl}-cell towards w_{Sl}-cells (see Fig. 15). However, we do not want all the w_{Sl}-cells receiving excitatory backward signals from the w_{Cl}-cell to be activated for the following reason: To activate a u_{Cl}-cell in the forward path, the activation of at least one preceding u_{Sl}-cell is sufficient. Usually only a small number of preceding u_{Sl}-cells are actually activated. In order to elicit a similar response from the

w_{Sl}-cells in the backward paths, the network is synthesized in such a way that each w_{Sl}-cell receives not only excitatory backward signals from w_{Cl}-cells but also a gate signal from the corresponding u_{Sl}-cell. Guided by the gate signals from the forward paths, the backward signals retrace the same route as the forward signals. Thus, the backward signals finally reach the recall layer W_{C0} at the lowest stage, and at exactly the same positions as the stimulus pattern presented to the input layer. Now let's consider the case in which a stimulus consisting of two or more patterns is presented to the input layer. Sometimes, two or more cells may be activated in the recognition layer. However, all of these cells but one, stop responding because of competition by lateral inhibition between cells in the forward paths, and also because of the process of focusing attention, mentioned later. Since backward signals are sent back only from the activated recognition cell, only signals corresponding to the pattern which is now recognized reach the recall layer W_{C0} at the lowest stage. Even if the stimulus pattern now recognized is a deformed version of a training pattern, it emerges with its deformed shape. Therefore, we can interpret the output of the recall layer as the result of segmentation of a single pattern from the stimulus.

Effect of backward signals on forward signals

Take, for example, a case in which the stimulus contains a number of incomplete patterns which are contaminated with noise and have several parts missing. Even when the pattern recognition in the forward path is successful, and only one cell is activated in the recognition layer U_{CL}, it does not necessarily mean that the segmentation of the patten is also completed in the recall-layer W_{C0}. When some part of the input pattern is missing and the feature which is supposed to exist there fails to be extracted in the forward paths, the backward signal flow is interrupted at that point and cannot proceed any further because no gate signals are received from the forward paths. The w_{CX}-cells find a failure to extract a feature when the cells in the backward paths are active but the corresponding cells in the forward paths are not (Figs 14 and 15). If this situation is detected, the w_{CX}-cells send threshold-control signals to u_S-cells around that area, and decrease the threshold for feature-extraction. Thus, the model is forced to extract even vague traces of the undetected feature. Once a feature is thus extracted in the forward paths, the backward signal can then be further transmitted to lower stages through the path unlocked by the gate signal from the newly activated forward cell. Hence, a complete pattern, in which defective parts are interpolated, emerges in the recall-layer W_{C0}. If all the recognition cells are silent, the no-response detector shown at the far right in Fig. 14 is activated and sends another threshold-control signal to all the u_S-cells of all stages and decreases their threshold for feature extraction. The value of the threshold-control signal increases until at least one recog-

nition cell comes to be activated. Forward cells receive gain-control signals from corresponding backward cells. More specifically, the gain of each u_C-cell is increased by the signal from the corresponding w_C-cell (Figs 14 and 15). Therefore, forward signal flow is facilitated only in paths in which backward signals flow. The gain-control signal plays the role of focusing attention. Let's consider the case in which a stimulus consisting of two or more patterns is presented to the input layer. Let one of the recognition cells be activated, and one of the patterns of the stimulus be recognized. Only the forward paths relevant to this pattern are facilitated by the gain-control signals, because the backward signals flow only from that recognition cell. This means that attention is selectively focused on one of the patterns of the stimulus. In order to switch attention to another pattern, it is sufficient to interrupt momentarily the backward signal-flow and make the gain-control signals disappear. Because of fatigue, each u_c-cell has its gain decreased if it had been earlier kept high by strong gain-control signal. Therefore, signals corresponding to the previous pattern now have difficulty in flowing through the forward paths, and another pattern will be recognized.

Computer simulation

The selective attention model was simulated on a computer. The variable connections were reinforced by unsupervised learning. Figure 16(a) shows the five training patterns, which were repeatedly presented to the network during the learning period. The training patterns were presented in this shape only; deformed versions were not presented at all. Figures 16(b) through (d) show the behavior of a network that has finished the learning process. In these figures, the responses of the cells in the input layer U_{C0} and the recall layer W_{C0} are shown in time sequence. The numeral to the upper left of each pattern represents time t after the start of stimulus presentation. The stimulus pattern presented to this network is identical to the response of the input layer at $t = 0$, shown in the upper left of each figure. (Note that the input pattern p appears directly in layer U_{C0} at $t = 0$, because no response is elicited from layer W_{C0} at $t < 0$.) Figure 16(b) shows the response to a stimulus consisting of two juxtaposed patterns, "2" and "3". In the recognition layer, not shown in this figure, the cell corresponding to pattern "2" happens to be activated first. This signal is fed back to the recall layer through backward paths, but the middle part of the segmented pattern "2" is missing because of interference from the closely adjacent "3". However, the interference soon decreases and the missing part recovered, because the signals for pattern "3", which is not being attended to, are gradually attenuated without receiving facilitation by gain-control signals. At $t = 5$, the backward signal-flow is interrupted for a moment to switch the attention. The mark ∇ denotes this operation.

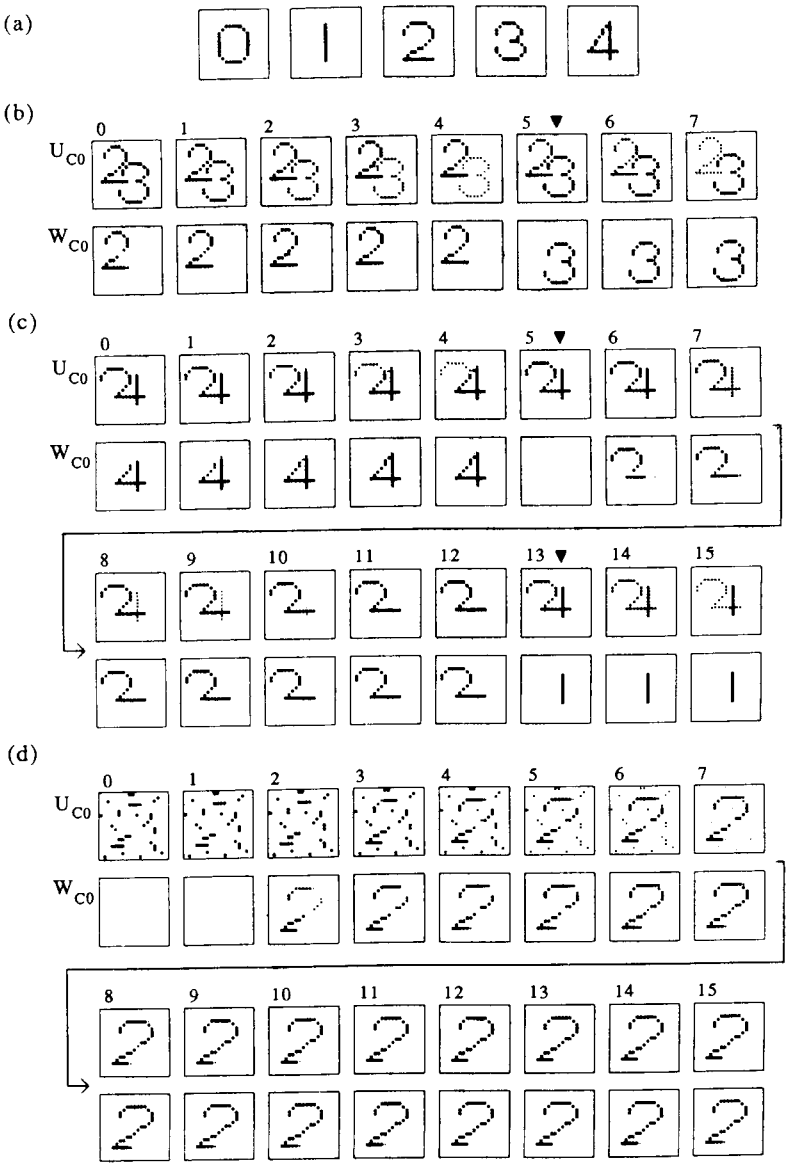


Fig. 16. Some examples of the response of the selective attention model. (a) Five training patterns used for learning. (b) An example of the response to juxtaposed patterns (Fukushima, 1987a). (c) An example of the response to superimposed patterns (Fukushima, 1987a, 1988). (d) An example of the response to an incomplete distorted pattern with noise (Fukushima, 1988).

Since the gain-control signals from the backward cells stop, the forward paths for pattern "2", which have so far been facilitated, now lose their

conductivity. The recognition cell for pattern "3" is now activated. Since backward signals are fed back from this newly activated recognition cell, pattern "3" is segmented and emerges in the recall layer W_{C0}. Recognition and segmentation of individual patterns can thus be successful even if the input patterns are deformed or shifted from the training patterns. Figure 16(c) shows an example of the response to a stimulus consisting of superimposed patterns. The pattern "4" is isolated first, the pattern "2" next, and finally pattern "1" is extracted. The recalled pattern "4" initially has one part missing, but the missing part is soon restored to resemble the training pattern.

Figure 16(d) shows the response to a greatly deformed pattern with several parts missing and contaminated by noise. Because of the large difference between the stimulus and the training pattern, no response is elicited from the recognition layer (not shown in the figure) at first. Accordingly, no feedback signal appears at the recall layer W_{C0}. The no-response detector detects this situation, and a threshold-control signal is sent to all feature-extracting cells in the network, which makes them respond more easily even to incomplete features. Thus, at time $t = 2$, the recognition cell for "2" is activated, and backward signals are fed back from it. Noise has been completely eliminated from the pattern now sent back to the recall-layer W_{C0}, and some missing parts have begun to be interpolated. This partly interpolated signal, namely the output of the recall layer W_{C0}, is again fed back positively to input layer U_{C0}. The interpolation continues gradually while the signal circulates through the feedback loop, and finally the missing part of the stimulus is completely filled in. The missing part is interpolated quite naturally, despite considerable difference in shape between the stimulus and the training pattern. In other words, the style of writing of the stimulus pattern is kept as faithful as possible, and only indispensable missing parts are restored.

Connected character recognition

The principles of this selective attention model can be extended to be used for several applications. One of them is the recognition and segmentation of connected characters in the cursive handwriting of English words. Generally speaking, machine recognition of connected characters is a difficult problem. It cannot be successfully performed by a simple pattern matching method, because each character changes its shape by the effect of the characters before and behind. In other words, the same character can be scripted differently when it appears in different words, in order to be connected smoothly with the characters in front of it and following it. Although the original model of selective attention already had the ability to recognize and segment patterns, it did not always work well when too many patterns were presented simultaneously. The model was then mod-

ified and extended to be able to recognize connected characters in cursive handwriting (Imagawa and Fukushima, 1991; Wake and Fukushima, 1992). In the model for connected character recognition, a search controller was added to the original model, in order to restrict the number of patterns to be processed simultaneously. The model mainly processes the patterns contained in a small "search area", which is moved by a search controller. To be more specific, in all stages except U_{CL}, the u_C-cells receive gain control signals not only from the w_C-cells but also from the search controller. The gain control signal from the search controller produces the search area by decreasing the gain of the u_C-cells situated outside of the search area. The position of the search area is shifted to the place in which a larger number of line extracting cells u_{C1} are activated. The search area has a size somewhat larger than the size of one character. The boundary of the search area is not sharply restricted: the gain of the u_C-cells is controlled to decrease gradually around the boundary. It is not necessary to control the position and the size of the area accurately because the original selective attention model has the ability to segment and recognize patterns by itself, provided the number of patterns present is small. The only requirement is that the search area covers at least one pattern. It does not matter if it covers a couple of patterns simultaneously. Once a pattern has been recognized and segmented, a monitor in the network sends a command to switch attention to recognize another pattern. The search controller again seeks a place in which a larger number of line extracting cells u_{C1} are activated, and shifts the search area to the new place. If all of the responses from the cells of layer U_{C1} are small enough because of fatigue, however, the model stops working, and assumes that all characters in the input string have already been processed. In order to prevent the model from recognizing the same character twice, the fatigue of the cells after attention switching is forced to continue longer than in the case of the original model. A preliminary experiment was performed using computer simulation to check the ability of the model. In this experiment, the model was taught only a small number of characters, instead of the whole set of 26 alphabetical characters. The network was trained with unsupervised learning, and the five training patterns shown in Fig. 17(a) were repeatedly presented to the network during the training phase. These training patterns were presented only in this shape, and anything like a deformed version of them was not presented during the training. In this experiment, the same pattern '*e*', shown in Fig. 17(a), is used not only as the training pattern for '*e*' but also as the training pattern for 'ℓ'. It should be noted that both '*e*' and 'ℓ' have almost the same shape when written in script style, and the only difference between them resides in their heights. After finishing the training, the same recognition cell in layer U_{CL} comes to be activated by both '*e*' and 'ℓ', because our selective attention model can recognize the shape of patterns robustly, with little effect from deformation. The two characters can easily be discriminated, however, by comparing the heights of the segmented pat-

terns, which appear in layer W_{C0}. Hence we can say that, in this experiment, our model has been taught to recognize, not five, but six characters. Figure 17(b) shows how the response of layer W_{C0}, in which the result of segmentation appears, changed with time when a handwritten character string *'late'*, shown at the top of the figure, was presented to the input layer U_{C0}. Time t after the first presentation of the character string is indicated in the figure. The mark $^{\vee}$ indicates the position to which the center of the search area was moved. It can be seen from this figure that character *'t'* was recognized first and segmented, then followed by *'l'*, *'a'*, and *'e'*. Attention was switched just after $t = 3$, 8, and 12. The model stopped working after $t = 21$, when all the characters in the input string had been completely recognized and segmented. Although the characters in the input string are different in shape from the training characters, recognition and segmentation of the characters was successfully performed. Figure 17(c) shows some examples of input character strings which have been successfully recognized and segmented. It can be seen from the figure that the input strings are processed correctly, even if the spacing between the characters changes. A string which contains two of the same characters with somewhat different shapes can also be processed successfully.

Chinese character recognition

Machine recognition of Chinese characters is also difficult, but for another reason. There are an enormous number of different characters, and each of these characters has a very complicated shape. However, it is fortunate that most of the Chinese characters can be decomposed into a couple of fundamental parts, such as left-hand radical (i.e. the side), right-hand radical (i.e. the body), upper radical (i.e. the crown part), and so on. In other words, a large number of different characters can be composed by a simple combination of a small number of radicals. If each of these radicals is recognized as a single pattern, the number of categories, to which patterns have to be classified, can be greatly reduced. A preliminary experiment has shown that the selective attention model can also be successfully used for this purpose (Fukushima *et al.*, 1991). In some Chinese characters, the same pattern may be a radical placed at several different positions. In other words, the information on the shape of a radical, only, is not always enough to decide whether it is in the left-hand, right-hand, or upper radical position. In order to get the information on the position of the radical in a character, the model contains a circuit to detect the center of gravity of the pattern segmented in layer W_{C0}.

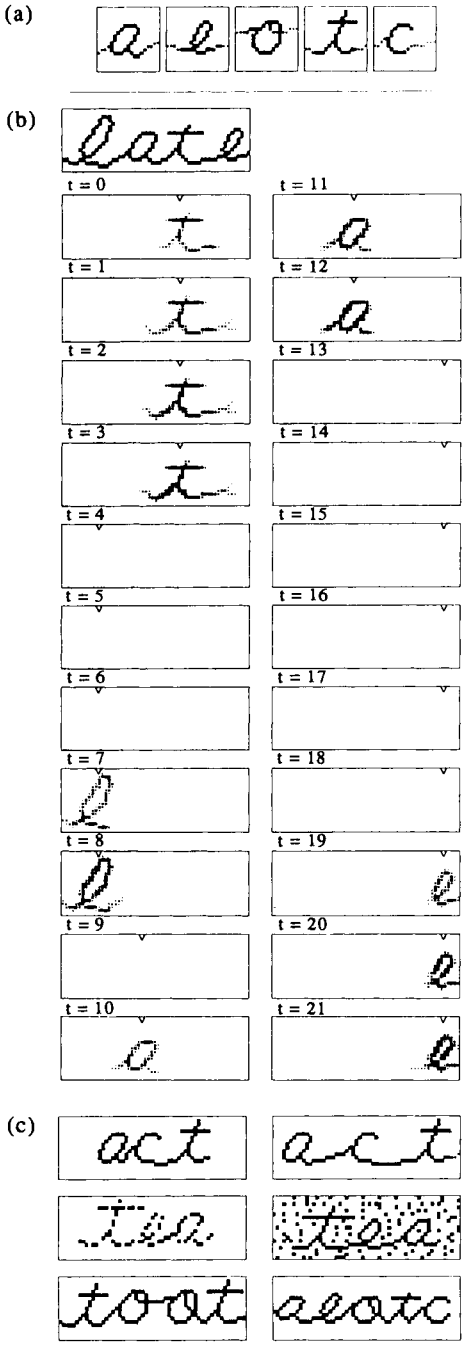


Fig. 17. Performance of the model for connected character recognition (Wake and Fukushima, 1992). (a) Five training patterns used for learning. The same pattern '*e*' is used as the training pattern for both '*e*' and '*ℓ*'. (b) Time course of the response of layer W_{C0}, in which the result of segmentation appears. A character string presented to the input layer is shown at the top. (c) Some examples of character strings which have been successfully recognized and segmented.

Conclusion

In order to design systems which have a high ability to process information like human beings, it is indispensable to construct networks with the same architecture as the real biological brain. This chapter has offered several examples of the systems which have been designed with the ideas acquired from modeling neural networks of the visual systems of the brain. For further development of the systems, however, we still have to continue our effort to construct neural network models more faithful to the biological brain.

Bibliography

Abbott, L. F. 1991a. Realistic synaptic inputs for model neural networks. *Network: Computation in Neural Systems* **2**, 245–258.

Abbott L. F. 1991b. Firing-rate models for neural populations, In: *From Biology to High Energy Physics* , Benhar, O., Bosio, C., Del Giudice, P. and Tabet, E. (Eds), pp. 179–196. Pisa, ETS Editrice.

Abbott, L. F. 1992. Simple diagrammatic rules for solving dendritic cable problems. *Physica* **A185**, 343–356.

Abbott, L. F., E. Farhi and S. Gutmann. 1991. The path integral for dendritic trees. *Biol. Cybern.* **66**, 49–60.

Abbott, L. F. and Kepler, T. B. 1990. Model neurons: from Hodgkin–Huxley to Hopfield. In: *Statistical Mechanics of Neural Networks,* Garrido, L. (Ed.), pp. 5–18. Berlin, Springer-Verlag.

Abbott, L. F. and LeMasson, G. 1992. Analysis of neuron models with dynamically regulated conductances. *Neural Comp.* (*submitted*).

Abbott, L. F., Marder, E. and Hooper, S. 1991. Oscillating networks: control of burst duration by electrically coupled neurons. *Neural. Comp.* **3**, 487–497.

Abbott, L. F. and Van Vreeswijk, C. 1992. Asynchronous states in networks of pulse-coupled oscillators. *Phys. Rev.* (*submitted*).

Abeles, M. 1982a. *Local cortical circuits.* Springer Verlag, Berlin.

Abeles, M. 1982b. Quantification, smoothing, and confidence limits for single-units' histograms. *J. Neurosci. Meth.* **5**, 317–325.

Abeles, M. 1983. The quantification and graphic display of correlations among three spike trains. *IEEE Trans. BME* **30**, 235–239.

Abeles, M. 1991. *Corticonics: Neural Circuits of the Cerebral Cortex.* Cambridge University Press, New York.

Abeles, M., Prut, Y., Bergman, H., Vaadia, E. and Aertsen, A. 1993. Integration, synchronicity and periodicity. In: *Spatio-temporal aspects of Brain Function,* Aertsen, A. and von Seelen,W. (Eds) Elsvier, Amsterdam (*in press*).

Abeles, M., de Ribaupierre, F. and de Ribaupierre, Y. 1983. Detection of single unit responses which are loosely time-locked to a stimulus. *IEEE Trans. SMC* **13**, 683–691.

Adams, P. R. and Gavin, M. 1986. Voltage-dependent currents of vertebrate neurons and their role in membrane excitability. *Adv. Neurol.* **44**, 137–170.

Aertsen, A. M. H. J. and Gerstein, G. L. 1985. Evaluation of neuronal connectivity: sensitivity of cross-correlation. *Brain Res.* **340**, 341–354.

Aertsen, A. M. H. J., Gerstein, G. L., Habib, M. K. and Palm, G. 1989. Dynamics of neuronal firing correlation: modulation of "effective connectivity". *J Neurophysiol.* **61**, 900–917.

Ahissar, E. and Vaadia, E. 1990. Single cell cortical oscillators in a somatosensory cortex of awake monkey. *Proc. Natl. Acad. Sci.* **87**, 8935–8939.

Ahissar, E., Vaadia, E., Galun, M., Bergman, H., Arieli, A. and Abeles, M. 1992. Dependence of cortical plasticity on correlated activity of single neurons and on behavioral context. *Science* **257**, 1412–1415.

Amaral, D. G., Ishizuka, N., Claiborne, B. 1990. Neurons, numbers and hippocampal network. In: *Understanding the brain through the hippocampus* (*Progress in Brain Research*), Vol. 83, Storm-Mathisen, J., Zimmer, J. and Ottersen, O. P. (Eds), pp. 1–11. Amsterdam, Elsevier.

Amari, S. 1974. A method of statistical neurodynamics. *Kybernetik* **14**, 201–215.

Amari, S., Maginu, K. 1988. Statistical neurodynamics of associative memory. *Neural Networks* **1**, 63–73.

Amit, D. J. 1989. *Modeling Brain Function.* Cambridge University Press, NY.

Amit, D. J., Gutfreund, H., Sompolinsky, H. 1985. Spin-glass models of neural networks. *Phys. Rev. A* **32**,1007.

Amit, D. J., Gutfreund, H., Sompolinsky, H. 1985. Storing infinite number of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.* **55**, 1530.

Amit, D. J., Gutfreund, H., Sompolinsky, H. 1987. Statistical mechanics of neural networks near saturation. *Ann. Phys.* **173**, 30.

Amit, D. J. and Tsodyks, M. V. 1991a. Quantitative study of attractor neural network retrieving at low spike rates I: Substrate–spikes, rates and neuronal gain. *Network* **2**, 259.

Amit, D. J. and Tsodyks, M. V. 1991b. Low-rate retrieval in symmetric networks. *Network* **2**, 275.

Amitai, Y., Friedman, A., Connors, B. W. and Gutnick, M. J. 1991. Dendritic electrogenesis in neocortical neurons in vitro. *Soc. Neurosci. Abstr.* **17**, 311.

Amoh, R. K. 1984. Estimation of parameters in mixtures of Inverse Gaussian distributions. *Commun. Statist.-Theor. Meth.* **13**, 1031–1043.

Anderssen, R. S., DeHoog, F. R. and Weiss, R. 1973. On the numerical solution of Brownian motion processes. *J. Appl. Prob.* **10**, 409–418.

Anninos, P. A., Beek, B., Csermely, T. J., Harth, E. and Pertile, G. 1970. Dynamics of neural structures. *J. Theor. Biol.* **26**, 121–148.

Arbib, M. A. 1981. Visuomotor coordination: from neural nets to schema theory. *Cognition and Brain Theory* **4**, 23–39.

Arbib, M. A., and House, D. H. 1987. Depth and detours: An essay on visually-guided behavior. In: *Vision, Brain, and Cooperative Computation* , M. A. Arbib and A. R. Hanson (Eds), pp. 129–163. Cambridge, MA: A Bradford Book/MIT Press.

Aronson, D. G., Ermentrout, G. B., and Kopell, N. 1990. Amplitude response of coupled oscillators. *Physica* **41D**, 403–449.

Ashwood, T. J., Collingridge, G. L., Herron, C. E. and Wheal, H. V. 1987. Voltage-clamp analysis of somatic gamma-aminobutyric acid responses in adult rat hippocampal neurones in vitro. *J. Physiol. (Lond.)* **384**, 27–37.

Babloyantz, A., Salazar, J. M. and Nicolis, C. 1985. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Phys Lett.* **111A**, 152–156.

Babloyantz, A. and Destexhe, A. 1986. Low-dimensional chaos in an instance of epilepsy. *Proc. Natl. Acad. Sci USA* **83**, 3513–3517.

Barna, G. 1992. Statistical approach for computing the largest Lyapunov exponent. In: *Proc. 2nd Int. Conf. Fuzzy Logic and Neural Networks* , pp. 921–924. Iizuka, Japan.

Barna, G. and Érdi, P. 1986. Pattern formation in neural systems II. Noise-induced selective mechanisms for the ontogenetic formation of ocular dominance columns. In: *Cybernetics and Systems '86.* Trappl, R. (Ed.), pp. 343–350. D. Reidel.

Barna, G. and Érdi, P. 1988. 'Normal' and 'abnormal' dynamic behaviour during synaptic transmission. In: *Computer simulation in brain science* , Cotteril, R. J. (Ed.), pp. 293–301. Cambridge, Univ. Press.

Barndorff-Nielsen, O., Blaesild, P. and Halgren, C. 1978. First hitting time models for the generalised Inverse Gaussian distribution. *Stoch. Proc. Appl.* **7**, 49–54.

Basar, E. 1980. *EEG–Brain Dynamics.* Amsterdam, Elsevier

Basar, E. 1990. *Chaos in Brain Function.* Berlin, Springer-Verlag.

Bear, M. F., Kleinschmidt, A., Gu, Q. A. and Singer, W. 1990. Disruption of experience-dependent synaptic modifications in striate cortex by infusion of an INMDA receptor antagonist. *J. Neurosci.* **10**, 902–925.

Bekkers, J. M. and Stevens, C. F. 1989. NMDA and non-NMDA receptors are co-localized at individual excitatory synapses in cultured rat hippocampus. *Nature* **341**, 230–233.

Benardo, L. S., Masukawa, L. M. and Prince, D. A. 1982. Electrophysiology of isolated hippocampal pyramidal dendrites. *J. Neurosci.* **2**, 1614–1622.

Bernander, O., Douglas, R. J., Martin, K. A. C. and Koch, C. 1991. Synaptic background activity determines spatio-temporal integration in single pyramidal cells. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11569–11573.

Beurle, R. L. 1956. Properties of a mass of cells capable of regenerating pulses. *Phil. Trans. Roy. Soc.* **240A**, 55–94.

Bower, J. M., ed. 1992. Modeling the Nervous System. *TINS* **15**, 411–464.

Bradley, W. G., Daroff, R. B., Fenichel, G. M. and Marsden, C. D. 1991. *Neurology in Clinical Practice* , Vol. **2**. Boston, Butterworth-Heinemann.

Brady, R. J. and Swann, J. W. 1986. Ketamine selectively suppresses synchronized afterdischarges in immature hippocampus. *Neurosci. Lett.* **69**, 143–149.

Brady, R. J. and Swann, J. W. 1988. Suppression of ictal-like activity by kynurenic acid does not correlate with its efficacy as an NMDA receptor antagonist. *Epilepsy Res.* **2**, 232–238.

Braitenberg, V. 1984. *Vehicles: Experiments in Synthetic Psychology.* Cambridge, MA: Bradford Books/The MIT Press.

Braitenberg, V., Schutz, A. 1991. *Anatomy of the Cortex: Statistics and Geometry.* Berlin, Springer-Verlag.

Brannan, J. R. 1982. Spatially localized interactive neural populations–II, Stability and dynamics of excitatory sets. *Bull. Math. Biol.* **44**, 43–56.

Brannan, J. R. and Boyce, W. E. 1981. Spatially localized interactive neural populations–I, A mathematical model. *Bull. Math. Biol.* **43**, 427–446.

Bressler, S. L. 1984. Spatial organization of EEGs from olfactory bulb and cortex. *Electroencephalography Clin. Neurophysiol.* **57**, 270–276.

Bressler, S. L. 1987. Functional relation of olfactory bulb and cortex: I. Spatial variation of bulbo-cortical interdependence. II. Model for driving of cortex by bulb. *Brain Res.* **409**, 285–301.

Brooks, R. A. 1986. A robust layered control system for a mobile robot. *IEEE J. Robotics Automation* **RA-2**, 14–23.

Buchholtz, F., Golowasch, J., Epstein, I. R. and Marder, E. 1992. Mathematical model of an identified stomatogastric ganglion neuron. *J. Neurophysiol.* **67**, 332–340.

Buhmann, J., Divko, R. and Schulten, K. 1989. Associative memory with high information content. *Phys. Rev.* **A39** 2689.

Buonocore, A., Nobile, A. G. and Ricciardi, L. M. 1987. A new integral equation for the evaluation of first-passage-time probability densities. *Adv. Appl. Prob.* **19**, 784–800.

Buonocore, A. Nobile, A. G., Ricciardi, L. M. and Sato, S. 1989. On the evaluation of first-passage-time probability densities via non-singular integral equations. *Adv. Appl. Prob.* **21**, 24–36.

Buonocore, A., Di Crescenzo, A., Iardino, F., Nakamura, A., Ricciardi, L. M. and Rinaldi S. 1992. A vectorized simulations of procedure for computations of first crossing time densities of normal processes with oscillatory covariances. *Tecnical Report No. 1/93.* Progetto finalizzato Sistemi Informatici e Calcolo Parallelo.

Bures, J., Buresova, O. and Koroleva, V. I. 1983. Spreading depression, epilepsy and memory. In: *Neurophysiological Mechanisms of Epilepsy* , Okujava, V. M. (Ed.), pp. 120–130. Metsnierba, Tiblisi.

Butz, E. G. and Cowan, J. D. 1974. Transient potentials in dendritic systems of arbitrary geometry. *Biophys. J.* **14**, 661–689.

Buzsáki, G. 1986. Hippocampal sharp waves: their origin and significance. *Brain Res.* **398**, 242–252.

Buzsáki, G., Horváth, Z., Urioste, R., Hetke, J. and Wise, K. 1992. High-frequency network oscillation in the hippocampus. *Science* **256**, 1025–1027.

Buzsáki, G., Leung, L.-W. S. and Vanderwolf, C. H. 1983. Cellular bases of hippocampal EEG in the behaving rat. *Brain Res.* **287**, 139–171.

Calabrese, R. L. and De Schutter, E. 1992. Motor-pattern-generating networks in invertebrates: modeling our way toward understanding. *TINS* **15**, 439–445.

Cannon, S. C., and Robinson, D. A. 1985. An improved neural-network model for the neural integrator of the oculomotor system: More realistic neuron behavior. *Biological Cybernetics* **53**, 93–108.

Cao, B. J. and Abbott, L. F. 1992. A new computational method for cable theory problems. *Biophys. J.* (*in press*).

Capocelli, R. M. and Ricciardi, L. M. 1971. Diffusion approximation and first passage time for a model neuron. *Kybernetik* **8**, 214–223.

Capocelli, R. M. and Ricciardi, L. M. 1972. On the inverse of the first passage time probability problem. *J. Appl. Prob.* **9**, 270–287.

Capocelli, R. M. and Ricciardi, L. M. 1973. A continuous Markovian model for neuronal activity. *J. Theor. Biol.* **40**, 369–387.

Cerbone, G., Ricciardi, L. M. and Sacerdote, L. 1981. Mean, variance and skewness of the first passage time for the Ornstein-Uhlenbeck process. *Cybern. Syst.* **12**, 395–429.

Chevalier, G., Vacher, S., Deniau, J. M., and Desban, M. 1985. Disinhibition as a basic process in the expression of striatal functions. I. The striato-nigral influence on the tecto-spinal/tecto-diencephalic neurons. *Brain Research* **334**, 215–226.

Chhikara, R. S. and Folks, J. L. 1989. *The Inverse Gaussian Distribution, Theory, Methodology and Applications.* Marcel Dekker, New York.

Clarke, E. and Dewhurst, K. 1972. *An illustrated History of Brain function.* Oxford, Sanford.

Cobas, A. and Arbib, M. A. 1991. Prey-catching and predator-avoidance in frog and toad: defining the schemas. In: *Visual Structure and Integrated Functions* , Arbib, M. A. and Ewert, J.-P. (Eds), pp. 153–166. New York, Springer-Verlag.

Colonnier, M. 1966. The structural design of the neocortex. In: *Brain and Conscious Experience.* Eccles, J. C. (Ed.), pp. 1–23. Berlin-Heidelberg-New York, Springer Verlag.

Connors, B. W., Gutnick, J. M. and Prince, D. A. 1982. Electrophysiological properties of neocortical neurons *in vitro. J. Neurophysiol.* **48**, 1302–1320.

Connors, B. W., Malenka, R. C. and Silva, L. R. 1988. Two inhibitory post-synaptic potentials and $GABA_A$ and $GABA_B$ receptor-mediated responses in neocortex of rat and cat. *J. Physiol.* **406**, 443–468.

Constantine-Paton, M., Cline, H. T., Debski, E. 1990. Patterned activity, synaptic convergence, and the NMDA receptor in developing visual pathways. *Ann. Rev. Neurosci.* **13**, 129–154.

Cooper, K. E., Gates, P. Y. and Eisenberg, R. S. 1988. Surmounting barriers in ionic channels. *Quart. Rev. Biophys.* **21**, 331–364.

Cope, T. C., Fetz, E. E. and Matsumura, M. 1987. Cross-correlation assesment of synaptic strength of single Ia fibre connections with triceps surae motoneurons in cats. *J. Physiol. (Lond.)* **390**, 161–188.

Cowan, J. D. 1991. Stochastic neurodynamics. In: *Neural information processing* , Vol. 3, Lippmann, M. P., Moody, J. E. and Toretzky, D. S. (Eds), pp. 62–69. Morgan Kaufmann Publ.

Cowan, J. D. and Sharp, D. H. 1988. Neural Nets. *Quart. Rev. Biophys.* **21**, 365–427.

Cox, D. R. 1970. *Renewal theory.* Methuen, London.

Cox, D. R. and Isham, V. 1980. *Point processes.* Chapman and Hall, London.

Crutchfield, J. P., Farmer, J. D., Packard, N. H. and Shaw, R. S. 1986. Chaos. *Sci. Am.* **255**, 38–49.

Cugliandolo, L. 1992. Correlated attractors from uncorrelated stimuli, *Talk given at the meeting "Mathematical Methods motivated by Neural Networks and Learning Processes"* Les Houches Centre de Physique.

Davies, S. N. and Collingridge, G. L. 1989. Role of excitatory amino acid receptors in synaptic transmission in area CA1 of rat hippocampus. *Proc. Roy. Soc. Lond.* **236**, 373–384.

Deiters, D. F. K. 1865. Posthumous citation in Max Schultze's Textbook, *Untersuchungen über Gehirn und Rückenmark des Menschen und der Säugetiere* , Braunschweig.

Del Castillo, J. and Katz, B. 1954. Quantal components of the end-plate potential. *J. Physiol. (Lond.)* **124**, 560–573.

Derrida, B., Gardner, E. and Zippelius, A. 1987. An exactly soluble asymmetric neural network model. *Europhys. Lett.* **4**, 167.

DeRoberties, E. D. P. and Bennett, H. S. 1957. Some features of the submicroscopic morphology of synapses in frog and earthworm. *J. Biophys. Biochem. Cytol.* **1**, 47–58.

Desai, M. A. and Conn, P. J. 1991. Excitatory effects of ACPD receptor activation in the hippocampus are mediated by direct effects on pyramidal cells and blockade of synaptic inhibition. *J. Neurophysiol.* **66**, 40–52.

Domann, R., Dorn, T. and Witte, O. W. 1991. Afterpotentials following penicillin-induced paroxysmal depolarizations in rat hippocampal CA1 pyramidal cells in vitro. *Pflüg. Arch.* **417**, 469–478.

Dominey, P. F., and Arbib, M. A. 1992. A cortico-subcortical model for generation of spatially accurate sequential saccades. *Cerebral Cortex* , **2**, 153–175.

Douglas, R. J., Martin, K. A. C. and Whitteridge, D. 1991. An intracellular analysis of the visual responses of neurones in cat visual cortex. *J. Physiol. (Lond.)* **440**, 659–696.

Durbin, J. 1971. Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov–Smirnov test. *J. Appl. Prob.* **8**, 431–453.

Dvorak, I. and Holden, A. V. 1991. *Mathematical approaches to brain functioning diagnostics.* Manchester, Univ. Press.

Eccles, J. C. 1957. *The Physiology of the nerve Cell.* Johns Hopkins Univ. Press, Baltimore.

Eccles, J. C. 1984. The cerebral neocortex: a theory of its operation. In: *Cerebral Cortex,* Vol. 2, Jones, E. G. and Peters, A. (Eds), pp. 1–36. Plenum Press, New York.

Eccles, J. C., Ito, M., Szentágothai, J. 1967. *The Cerebellum as a Neuronal Machine.* Berlin, Springer Verlag.

Eckhorn, R., Bauer, R., Jorden, W., Brosch, M., Kruse, W., Munk, M. and Reitboeck, H. J. 1988. Coherent oscillations: A mechanism for feature linking in the visual cortex? *Biol. Cybern.* **60**, 121–130.

Eeckman, F. H., Freeman, W. J. 1991. Asymmetric sigmoid nonlinearity in the rat olfactory system. *Brain Research* **557**, 13–21.

Eisenberg, J., Freeman, W. J., Burke, B. 1989. Hardware architecture of a neural network model simulating pattern recognition by the olfactory bulb. *Neural Networks* **2**, 315–325.

Eisenfeld, J. and DeLisi, C. 1985. On conditions for qualitative instability of regulatory circuits with applications to immunological control loops. In: *Mathematics and computers in biomedical applications* , Eisenfeld, J. and DeLisi, C. (Eds), pp. 39–53. Elsevier, New York.

Ekeberg, & ., Oumlaut;., Wallén, P., Lansner, A., Tråven, H., Brodin, L., Grillner, S. 1991. A computer based model for realistic simulations of neural networks. I. The single neuron and synaptic interaction. *Biological Cybernetics* **65**, 81–90.

Enright, J. T. and Winfree, A. T. 1987. Detecting a phase singularity in a coupled stochastic system. In : *Lectures on Mathemathics in Life Sciences.* Vol. 19, Carpenter, G. A. (Ed.), pp. 121–150. American Math. Soc., Providence, Rhode Island.

Érdi, P. 1991. Self-organization in the nervous system: network structure and stability. In: Dvorak, I. and Holden, A. V. , pp. 31–43.

Érdi, P. and Barna, G. 1984. Self-organizing mechanism for the formation of ordered neural mappings. *Biol. Cybernetics* **51**, 93–101.

Érdi, P. and Barna, G. 1985. Self-organization of neural networks: noise-induced transition. *Phys. Lett.* **107A**, 287–290.

Érdi, P. and Barna, G. 1987. Self-organization in the nervous system: some illustrations. In: *Lecture Notes in Biomathematics, Vol 71, Mathematical topics in population biology, morphogenesis and neuroscience* , Teramoto, E. and Yamaguti, M. (Eds), pp. 301–317. Springer.

Érdi, P. and Barna, G. 1991. 'Neural' model for the formation of the ocularity domains. In: *Artificial neural networks* , Vol 1, Kohonen *et al.* (Eds), pp. 513–518. North-Holland.

Érdi, P., Gröbler, T. and Kaski, 1992. Dynamic phenomena in the olfactory bulb I. Bifurcation sequences, coexistence of periodicity and chaos, synaptic modification induced transition. In: *Artificial neural networks* **II**, Aleksander, I. and Taylor, J. (Eds), pp. 873–876. Elsevier.

Érdi, P. and Szentágothai, J. 1985. Neural connectivities: between determinism and randomness. In: *Dynamics of Macrosystems* , Aubin, J.-P., Saari, D. and Sigmund, K. (Eds), pp. 21–29. Springer.

Érdi, P. and Tóth, J. 1990. What is and what is not stated by the May-Wigner theorem? *J. theor. Biol.* **145**, 137–140.

Ermentrout, G. B. 1986. The behavior of rings of coupled oscillators. *J. Math Biol.* **23**, 55–74

Ermentrout, G. B. 1990. Oscillator death in populations of "all to all" coupled nonlinear oscillators. *Physica* **41D**, 219–231.

Ermentrout, G. B. 1991. An adaptive model for synchrony in the firefly Pteroptyx malaccae. *J. Math. Biol.* **29**, 571–585.

Ermentrout, G. B. 1992. Periodic solutions in discrete and continuum arrays of coupled nonlinear oscillators. *SIAM J. Appl. Math.* (*in press*).

Ermentrout, G. B. and Cowan, J. D. 1979. Temporal oscillations in neuronal nets. *J. Math. Biol.* **7**, 265–280.

Ermentrout, G. B. and Kopell, N. 1986. Parabolic bursting in an excitable system coupled with a slow oscillation. *SIAM J. Appl. Math.* **46**, 233–253.

Ermentrout, G. B. and Kopell, N. 1990. Oscillator death in systems of coupled neural oscillators. *SIAM J. Appl. Math.* **50**, 125–146.

Ermentrout, G. B. and Kopell, N. 1991. Multiple pulse interactions and averaging in systems of coupled neural oscillators. *J. Math. Biol.* **29**, 33–58.

Ewert, J.-P. 1987. Neuroethology of releasing mechanisms: prey-catching in toads. *Behavioral Brain Sci.* **10**, 337–405.

Fatt, P. and Katz, B. 1952. Spontaneous subthreshold activity of motor nerve endings. *J. Physiol. (Lond.)* **117**, 109–128.

Favella, L., Reineri, M. T., Ricciardi, L. M. and Sacerdote, L. 1982. First passage time problems and some related computational problems. *Cybernetics Systems* **13**, 95–128.

Feller, W. 1951a. Diffusion processes in genetics. In: *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability.* pp. 227–246.

Feller, W. 1951b. Two singular diffusion processes. *Ann. Math.* **54**, 173–182.

Feller, W. 1952. Parabolic differential equations and semigroup transformations. *Ann. Math.* **55**, 468–518.

Feller, W. 1954. Diffusion processes in one dimension. *Trans. Amer. Math. Soc.* **77**, 1–31.

Feller, W. 1966. *An Introduction to Probability Theory and its Applications.* Vol. 2. Wiley, New York.

Fetz, E., Toyama, K. and Smith, W. 1991. Synaptic interactions between cortical neurons. In: *Cerebral Cortex* , Vol. 9. Peters, A. (Ed.), pp. 1–47. Plenum Publ. Corp.

Fienberg, S. E. 1974. Stochastic models for single neuron firing trains, A survey. *Biometrics* **30**, 399–427.

Fisher, B. 1973. A neuron field theory: Mathematical approach to the problem of large number of interacting nerve cells. *Bull. Math. Biol.* **35**, 345–357.

FitzHugh, R. 1961. Impulses and physiological states in models of nerve membrane. *Biophys. J.* **1**, 445–466.

Forsythe, I. D. and Westbrook, G. L. 1988. Slow excitatory postsynaptic currents mediated by *N* -methyl-D aspartate receptors on cultured mouse central neurones. *J. Physiol. (Lond.)* **396**, 515–533.

Frankel, P. and Kiemel, T. Relative phase behavior of two slowly coupled oscillators. *SIAM J . App. Math. (submitted*).

Freeman, W. J. 1960. Correlation of electrical activity of prepyriform cortex and behavior in cat. *J. Neurophysiol.* **23**, 111–131.

Freeman, W. J. 1975. *Mass Action in the Nervous System.* New York, Academic Press.

Freeman, W. J. 1979. Nonlinear gain mediating cortical stimulus–response relations. *Biological Cybernetics* **33**, 237–247.

Freeman, W. J. 1983. Dynamics of image formation by nerve cell assemblies. In: *Synergetics of the Brain* , Basar, E., Flohr. H., Haken, H., Mandell, A. J. (Eds), pp 102–121. Berlin, Springer-Verlag.

Freeman, W. J. 1987. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biological Cybernetics* **56**, 139–150.

Freeman, W. J. 1987. Techniques used in the search for the physiological basis of the EEG. In: *Handbook of Electroencephalography and Clinical Neurophysiology,* Vol. 3A, Part 2, Ch. 18, Gevins. A., Remond, A. (Eds), pp. 583–664. Amsterdam, Elsevier.

Freeman, W. J. 1988. Analysis of strange attractors in EEGs with kinesthetics and computer graphics in 4-D. Lecture at *2nd Intern. Conf. on Dynamics of Sensory and Cognitive Processing in the Brain* , Basar. E., Stowell H (Eds), pp. 512–520. Berlin, Springer-Verlag.

Freeman, W. J. 1991. The physiology of perception. *Scientific Amer.* **264**, 78–85.

Freeman, W. J., Baird, B. 1987. Relation of olfactory EEG to behavior: Spatial analysis. *Behavioral Neuroscience* **101**, 393–408.

Freeman, W. J., Davis, G. W. 1990. Olfactory EEG changes under serial discrimination of odorants by rabbits. In: *Chemosensory Information Processing* , NATO ASI Series, Vol. H39, Schild, D. (Ed.), pp. 375–391. Berlin, Springer-Verlag.

Freeman, W. J., Grajski, K. A. 1987. Relation of olfactory EEG to behavior: Factor analysis. *Behavioral Neurosci.* **101**, 766–777.

Freeman, W. J., Schneider, W. 1982. Changes in spatial patterns of rabbit olfactory EEG with conditioning to odors. *Psychophysiol.* **19**, 44–56.

Freeman, W. J., Skarda, C. A. 1985. Nonlinear dynamics, perception, and the EEG; the neo-Sherringtonian view. *Brain Res. Rev.* **10**, 147–175.

Freeman, W. J., Van Dijk, B. 1987. Spatial patterns of visual cortical fast EEG during conditioned reflex in a rhesus monkey. *Brain Res.* **422**, 267–276.

Freeman, W. J., Viana di Prisco, G. 1986. Relation of olfactory EEG to behavior: Time series analysis. *Behavioral Neurosci.* **100**, 753–763.

Freeman, W. J., Yao, Y. and Burke, B. 1988. Central pattern generating and recognizing in olfactory bulb: a correlation learning rule. *Neural Networks* **1**, 277–278.

Friedland, R. P., Budinger, T. F., Jagust, W. J., Koss, E., Derenzo, S., Huesman, R. H., Yano,Y. 1985. Positron tomography and the differential diagnosis and pathophysiology of Alzheimer's disease. In: *Senile Dementia of the Alzheimer Type* , Traber, J. and Gispen, W. H. (Eds), pp. 124–133. Springer.

Friedlander, M. J., Sayer, R. J. and Redman, S. J. 1990. Evaluation of long-term potentiation of small compound and unitary EPSPs at the hippocampal CA3-CA1 synapse. *J. Neurosci.* **10**, 814–825.

Frigessi, A. and Den Hollander, F. 1989. A stochastic model for the membrane potential of a stimulated neuron. *J. Math. Biol.*

Frolov, A. A. and Medvedev, A. V. 1986. Substantiation of the 'point approximation' for describing the total electrical activity of the brain with the use of a simulation model. *Biophysics* **31**, 332–335.

Fuchs, A. F., Kaneko, C. R. S., and Scudder, C. A. 1985. Brainstem control of saccadic eye movements. *Ann. Rev. Neurosci.* **8**, 307–337.

Fukushima, K. 1970. An electronic model of the retina. *Proc. IEEE* **58**, 1950–1951.

Fukushima, K. 1970. A feature extractor for curvilinear patterns:A design suggested by the mammalian visual system. *Kybernetik* **7**, 153–160.

Fukushima, K. 1971. A feature extractor for a pattern recognizer—A design suggested by the visual system (*in Japanese*). *NHK Tech. J.* **5**, 351–367.

Fukushima, K. 1975. Cognitron: A self-organizing multilayered neural network. *Biological Cybernetics* **20**, 121–136.

Fukushima, K. 1976. *Physiology and Bionics of the Visual Systems* (*in Japanese*). Tokyo, Inst. Electronics Commun. Engrs.

Fukushima, K. 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics* **36**, 193–202.

Fukushima, K. 1981. Cognitron: A self-organizing multilayered neural network model. *NHK Tech. Monograph* , **30**. Tokyo, NHK Tech. Res. Labs.

Fukushima, K. 1986. A neural network model for selective attention in visual pattern recognition. *Biological Cybernetics* **55**, 5–15.

Fukushima, K. 1987. A neural network model for selective attention in visual pattern recognition and associative recall. *Appl. Optics* **26**, 4985–4992.

Fukushima, K. 1988. A neural network for visual pattern recognition. *Computer* **21**, 65–75.

Fukushima, K. 1988. Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks* **1**, 119–130.

Fukushima, K., Imagawa, T. 1992. Recognition and segmentation of connected characters with selective attention. *Neural Networks* **5** (*in press*).

Fukushima, K., Imagawa, T., Ashida, E. 1991. Character recognition with selective attention. *IJCNN-91-Seattle (International Conference on Neural Networks)* , Vol. I, pp. 593–598. Seattle, WA, U.S.A.

Fukushima, K., Miyake, S. 1982. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition* **15**, 455–469.

Fukushima, K., Wake, N. 1991. Handwritten alphanumeric character recognition by the neocognitron. *IEEE Trans. Neural Networks* **2**, 355–365.

Fukushima, K., Wake, N. 1992. Improved neocognitron with bend-detecting cells. *IJCNN '92-Baltimore* , Vol.IV, pp. 190–195. Baltimore, MD, U. S. A. .

Fukushima, K., Wake, N. 1992. An improved learning algorithm for the neocognitron. In: *Artificial Neural Networks* Vol. 2, Aleksander, I., Taylor, J. (Eds), pp. 497–505. Amsterdam, North-Holland.

Fuster, J. M. 1985. The prefrontal cortex and temporal integration. In: *Cerebral Cortex*, Vol. 4, Peters, A. and Jones, E. G. (Eds), pp. 151–177. Plenum Press, New York.

Gallez, D. and Babloyantz, A. 1991. Predictability of human EEG: a dynamic approach. *Biol. Cybernetics* **64**, 381–391.

Gardner, E. 1987. Maximum storage capacity in neural networks. *Europhys. Lett.* , **4**, 481.

Gardner, M. R. and Ashby, W. R. 1970. Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature* **228**, 784.

Geisler, C. D. and Goldberg, J. M. 1966. A stochastic model of the repetitive activity of neurons. *Biophys. J.* **6**, 53–69.

Gerstein, G. L. and Mandelbrot, B. 1964. Random walk models for the spike activity of a single neuron. *Biophys. J.* **4**, 41–68.

Gerstein, G. L., Perkel, D. H. and Dayhoff, J. 1985. Cooperative firing activity in simultaneously recorded populations of neurons: detection and measurement. *J. Neurosci.* **5**, 881–889.

Gierer, A. and Meinhardt, H. 1972. A theory of biological pattern formation. *Kybernetik* **12**, 30–39.

Giorno, V., Lánský, P., Nobile, A. G. and Ricciardi, L. M. 1988. Diffusion approximation and first-passage-time problem for a model neuron. III. A birth-and-death process approach. *Biol. Cybern.* **58**, 387–404.

Giorno, V., Nobile, A. G., Ricciardi, L. M. and Sato, S. 1989. On the evaluation of the first-passage-time densities via non-singular integral equations. *Adv. Appl. Prob.* **21**, 20–36.

Giorno, V., Nobile, A. G. and Ricciardi, L. M. 1990. On the asymptotic behaviour of first-passage-time densities for one-dimensional diffusion processes and varying boundaries. *Adv. Appl. Prob.* **22**, 883–914.

Giorno, V., Nobile, A. G. and Ricciardi, L. M. 1992. Instantaneous return processes and neuronal firings. In: *Cybernetics and Systems Research* , Vol. 1, Trappl, R. (Ed.), pp. 829–836. World Scientific, Singapore.

Glass, L. and Mackey, M. 1988. *From Clocks to Chaos.* Princeton, New Jersey, Princeton University Press.

Gluss, B. 1967. A model for neuron firing with exponential decay of potential resulting in diffusion equations for probability density. *Bull. Math. Biophys.* **29**, 233–243.

Goel, N. S. and Richter-Dyn, N. 1974. *Stochastic Models in Biology* , Academic Press, New York.

Goldman, P. S. and Nauta, W. J. H. 1977. Columnar distribution of cortico-cortical fibers in the frontal association limbic and motor cortex of the developing Rhesus monkey. *Brain Res.* **122**, 393–413.

Golgi, C. 1873. Sulla struttura della sostanza grigia del cervello. *Gazetta Medica Lombarda* **33**, 244–246.

Golomb, D., Gueron, S., Harris-Warrick, R. and Guckenheimer, J. 1992. Reduction of LP cell model (*submitted*).

Goodhill, G. 1992. Correlations, competition, and optimality: modelling the development of topography and ocular dominance. *CSRP* **226**.

Gorelova, N. A. and Bures, J. 1983. Spiral waves of spreading depression in the isolated chicken retina. *J. Neurobiol.* **14**, 353–363.

Grassberger, P. and Procaccia, I. 1983. Characterisation of strange attractors. *Phys. Rev. Lett.* **50**, 346–349.

Gray, C. M., Freeman, W. J., Skinner, J. E. 1986. Chemical dependencies of learning in the rabbit olfactory bulb: acquisition of the transient spatial-pattern change depends on norepinephrine. *Behavioral Neurosci.* **100**, 585–596.

Gray, C. M. and Singer, W. 1989. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **86**,1698–1702.

Griffith, J. S. 1963. A field theory of neural nets:I. Derivation of fields equation. *Bull. Math. Biophys.* **25**, 111–120.

Griniasty, M., Tsodyks, M. V., Amit, D. J. 1993. Conversion of temporal correlations between stimuli to spatial correlations between attractors. *Neural Computation* **5**, 1–17.

Gustafsson, B. and Wigstrom, H. 1988. Physiological mechanisms underlying long-term potentiation . *Trends Neurosci.* **11**, 156–162.

Hablitz, J. J. 1984. Picrotoxin-induced epileptiform activity in the hippocampus: role of endogenous versus synaptic factors. *J. Neurophysiol.* **51**, 1011–1027.

Hablitz, J. J. and Thalmann, R. H. 1987. Conductance changes underlying a late synaptic hyperpolarization in hippocampal CA3 neurons. *J. Neurophysiol.* **58**, 160–179.

Hansel, D., Sompolinsky, H. 1990. Learning from examples in a single-layer neural network. *Europhysics Letters* **11**, 687–692.

Hanson, F. B. and Tuckwell, H. C. 1983. Diffusion approximations for neuronal activity including synaptic reversal potentials. *J. Theor. Neurobiol.* **2**, 127–153.

Hanson, F. E. 1978. Comparative studies of firefly pacemakers. *Fed. Proc.* **37**, 2158–2164.

Harrison, L. G. 1987. What is the status of reaction-difusion theory thirty-four years after Turing? *J. theor. Biol.* **125**, 369–384.

Harth, E., Csermely, T. J., Beek, B. and Linsday, R. O. 1970. Brain functions and neural dynamics. *J. theor. Biol.* **26**, 121–148.

Hayashi, H., Ishizuku, S., Ohta, M. and Hirakawa, K. 1982. Chaotic behaviour in *Onchidium* giant neuron under sinusoidal stimulation. *Phys. Lett.* **88A**, 435–438.

Hayashi, H. and Ishizuka, S. 1990. Chaotic activity in hippocampus neural network and intracranial self-stimulation. In: *Proceedings of the International Fuzzy Logic and Neural Networks* , Vol. 2, pp. 583–586. Kyushu Institute of Technology, Iizuca, Fukuoka, Japan.

Hayashi, H. and Ishizuka, S. 1992. Chaotic nature of bursting discharges in the *Onchidium* pacemaker neuron. *J. theor. Biol.* **156**, 169–191.

Hebb, C. O. 1949. *The Organization of Behaviour.* Wiley, New York.

Herrick, C. J. 1924. *Neurological Foundations of Animal Behavior.* New York, Hafner Publishing.

Hestrin, S., Nicoll, R. A., Perkel, D. J. and Sah, P. 1990. Analysis of excitatory synaptic action in pyramidal cells using whole-cell recording from rat hippocampal slices. *J. Physiol. (Lond.)* **422**, 203–225.

Hestrin, S., Sah, P. and Nicoll, R. A. 1990. Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron* **5**, 247–253.

Hikosaka, O., and Wurtz, R. 1983a. Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *J. Neurophysiol.* **49**, 1268–1284.

Hikosaka, O., and Wurtz, R. 1983b. Visual and oculomotor functions of monkey substantia nigra pars reticulata. IV. Relation of substantia nigra to superior colliculus. *J. Neurophysiol.* **49**, 1285–1301.

Hille B. 1992. *Ionic Channels of Excitable Membranes.* Sunderland MA, Sinauer.

Hines, M. 1984. Efficient computation of branched nerve equations. *J. Bio-Med. Comp.* **15**, 69–76.

Hirai,Y. 1980. A new hypothesis for synaptic modification: an interactive process between postsynaptic competition and presynaptic regulation. *Biol. Cybern.* **36**, 41–50.

Hirsch, M. 1991. Network dynamics: principles and problems. In: *Neurodynamics* , Pasemann, F. and Doebner, H. D. (Eds), pp. 3–29. World Scientific.

Hodgkin, A. L. and Huxley, A. F. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.)* **117**, 500–544.

Holden, A. V. 1975. A note on convolution and stable distributions in the nervous system. *Biol. Cybern.* **20**, 171–173.

Holden, A. V. 1976. *Models of stochastic activity of neurones. Lecture Notes in Biomathematics* , Vol. **12**. Springer, Berlin.

Holden, A. V. 1978. Information transfer in a chain of model neurones. In: *Progress in Cybernetics and Systems Research* , Vol. III, Trappl, R., Klir, J. and Ricciardi L. M. (Eds), pp. 558–569. Hemisphere, Washington.

Holden, A. V. 1985. Why the nervous system not as chaotic as it should be? In: *Dynamic phenomena in neurochemistry and neurophysics: theoretical aspects* , Érdi, P. (Ed.), pp. 6–11. KFKI, Budapest.

Holden, A. V. 1991. Equilibria, periodicity, bursting and chaos in neural activity. In: *Neurodynamics* , Pasemann, F. and Doebner, H. D. (Eds), pp. 96–128. World Scientific.

Holden, A. V., Winlow, W. and Haydon, P. G. 1982. The induction of periodic and chaotic activity in molluscan neurone. *Biol. Cybernetics* **43**, 169–173.

Holmes, W. R. 1986. A continuous cable method for determining the transient potential in passive trees of known geometry. *Biol. Cybern.* **55**, 115–124.

Hopfield, J. J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554–2558.

Hopfield, J. J. 1984. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3088–3092.

Hopfield, J. J. and Tank, D. W. 1986. Computing with neural circuits: a model. *Science* **233**, 625–633.

Horsthemke, W. and Lefever, R. 1984. *Noise-induced transition. Theory and application in physics, chemistry and biology.* Springer, Berlin-Heidelberg-Tokyo.

Horwitz, B. 1981. An analytic method for investigation transient potentials in branched neurons with branching dendritic trees. *Biophys. J.* **36,** 155–192.

Hubel, D. H. 1982. Exploration of the primary visual cortex, 1955–78. *Nature* **299**, 515–524.

Hubel, D. H. and Wiesel, T. N. 1959. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol. (Lond.)* **148**, 574–591.

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

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

Hubel, D. H. and Wiesel, T. N. 1970. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J. Physiol.* **206**, 419.

Hubel, D. H. and Wiesel, T. N. 1972. Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *J. Comp. Neurol.* **146**, 421–450.

Hubel, D. H. and Wiesel, T. N. 1977. Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. Ser. B.* **198**, 1–59.

Imagawa, T., Fukushima, K. 1991. Character recognition in cursive handwriting with the mechanism of selective attention (*in Japanese*). *IEICE Transactions* **J74-D-II** 1768–1775.

Ingber, L. 1982. Statistical mechanics of neocortical interactions. I. Basic formulation. *Physica* **5D**, 83–107.

Ingle, D., and Hoff, K. V. 1990. Visually elicited evasive behavior in frogs: Giving memory research an ethological context. *BioScience* **40**, 284–291.

Ito, M. 1984. *The Cerebellum and the Neural control.* New York, Raven Press.

Jack, J. J. B, D. Noble and R. W. Tsien. 1975. *Electrical Current Flow in Excitable Cells.* Oxford, Oxford Univ. Press.

Jack, J. J. B. and Redman, S. J.,. 1971. The propagation of transient potentials in some linear cable structures. *J. Physiol.* **215**, 283–320.

Jackson, J. H. 1931. On the scientific and empirical investigation of epileptics. In: *Selected Writings of John Hughlings Jackson* , Vol. 1, Taylor, J., (Ed.), pp. 177. London, Hodder and Stoughton.

Jahr, C. E. and Stevens, C. F. 1990. A quantitative description of NMDA receptor channel kinetic behavior. *J. Neurosci.* **10**, 1830–1837.

Jahr, C. E. and Stevens, C. F. 1990. Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *J. Neurosci.* **10**, 3178–3182.

Jefferys, J. G. R. 1990. Basic mechanisms of focal epilepsies. *Exp. Physiol.* **75**, 127–162.

Jefferys, J. G. R. and Traub, R. D. 1992. Synchronization of CA3 pyramidal neurons by NMDA mediated excitatory synaptic potentials in hippocampal slices incubated in low-Mg^{2+} solutions. *J. Physiol. (Lond.)* **452**, 32P.

Johannesma, P. J. M. 1968. Diffusion models for the stochastic activity of neurones. In: *Neural Networks* , Caianiello E. R. (Ed.), pp. 116–144. Springer, Berlin.

Johnston, D. and Brown, T. H. 1981. Giant synaptic potential hypothesis for epileptiform activity. *Science* **211**, 294–297.

Jones, E. G. 1986. Connectivity of the primate sensory-motor cortex. In: *Cerebral Cortex* , Vol. 5, Jones, E. G. and Peters, A. (Eds), pp. 113–173. New York, Plenum Press.

Jürgens, R., Becker, W., and Kornhuber, H. H. 1981. Natural and drug-induced variations of velocity and duration of human saccadic eye movements: Evidence for a control of the neural pulse generator by local feedback. *Biological Cybernetics* **39**, 87–96.

Kallianpur, G. 1983. On the diffusion approximation to a discontinuous model for a single neuron, In: *Contributions to Statistics* , Sen P. K. (Ed.), pp. 247–258. North-Holland, Amsterdam.

Kandel E. R., Schwartz, J. H. and Jessell, T. M. 1991. *Principles of Neural Science*. New York, Elsevier.

Karlin, S. and Taylor, H. M. 1981. *A Second Course in Stochastic Processes.* Academic Press, New York.

Katchalsky, A. K., Rowland, V. and Blumenthal, R. 1974. Dynamic patterns of brain cell assemblies. *NRP Bull.* **12(1)**.

Kay, A. R. and Wong, R. K. S. 1987. Calcium current activation kinetics in pyramidal neurones of the CAl region of the mature guinea pig hippocampus. *J. Physiol. (Lond.)* **392**, 603–616.

Kepler, T. B., Abbott, L. F. and Marder, E. 1992. Reduction of conductance-based neuron models. *Biol. Cybern.* **66**, 381–387.

Khibnik, A., Borisyuk, R. M., and Roose, D. 1991. Numerical bifurcation analysis of a model of coupled neural oscillators. In: *Bifurcation and Symmetry*, Allgower, G., Boehmer, K., and Golubitsky, M. (Eds). Basel, Birkhauser.

Kievit, J., Kuypers, H. G. J. M. 1977. Organization of the thalamo-cortical connections to the frontal lobe in the Rhesus Monkey. *Exp. Brain Res.* **29**, 299–322.

King, R. B. 1983. Chemical applications of topology and group theory 14. Topological aspects of chaotic chemical reactions. *Theoret. Chim. Acta (Berl.)* **63**, 323–338.

King, R., Barchas, J. D. and Huberman, B. A. 1984. Chaotic behavior in dopamine neurodynamics. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1244–1247.

Kliemann, W. 1987. A stochastic dynamical model for the characteization of the geometrical structure of dendritic processes. *Bull. Math. Biol.* **49**, 135–152.

Koch, C. and Poggio, T. 1985. A simple algorithm for solving the cable equation in dendritic trees of arbitrary geometry. *J. Neurosci. Meth.* **12**, 303–315.

Koch, C. and Segev, I. (Eds). 1991. *Methods in Neuronal Modeling*. Cambridge MA, MIT Press.

Kopell, N. and Ermentrout, G. B. 1986. Symmetry and phaselocking in chains of weakly coupled oscillators. *Comm. Pure Appl. Math.* **39**, 623–660.

Kopell, N. and Ermentrout, G. B. 1990. Phase transitions and other phenomena in chains of coupled oscillators. *SIAM J. Appl. Math.* **50**, 1014–1052.

Kopell, N., Ermentrout, G. B., and Williams, T. 1991. Periodically forced chains of coupled oscillators. *SIAM J. Appl. Math.* **51**, 1397–1417

Kopell, N., Zhang, W., and Ermentrout, G. B. 1990. Multiple coupling in chains of oscillators. *SIAM J. Math. Anal.* **21**, 935–953.

Kriebel, M. E., Vautrin, J. and Holsapple, J. 1990. Transmitter release: prepacking and random mechanism or dynamic and deterministic process. *Brain Res. Rev.* **15**, 167–178.

Krinskii, V. I. and Kokoz, Y. M. 1973. Analysis of equations of excitable membranes—1. Reduction of the Hodgkin–Huxley equations to a second-order system. *Biofizika* **18**, 506–511.

Kruger, J. and Aiple, F. 1988. Multimicroelectrode investigation of monkey striate cortex: Spike train correlations in the infragranular layers. *J. Neurophysiol.* **60**, 798–828.

Kryukov, V. I. 1984. Short-term memory as a metastable state. I. Master equation approach. In: *Cybernetics and Systems Research*, Vol. 2, Trappl, R. (Ed.), pp. 261–266. North-Holland, Amsterdam.

Kuhn, R. 1990. Statistical mechanics of networks of analog neurons. In: *Statistical Mechanics of Neural Networks* , Garrido, L. (Ed.). Springer, Berlin.

Kuramoto, Y. 1984. *Chemical Oscillations, Waves, and Turbulence.* Berlin, Springer.

Lábos, E. 1977. Theoretical considerations of local neuron circuits and their triadic synaptic arrangements (TSA) in subcortical sensory nuclei. *J. Neurosci. Res.* **3**, 1–10.

Lábos, E. 1981. The formal neuronal network theory: applications. In: *Recent developments of Neurobiology in Hungary* , Lissák, K. (Ed.), pp. 9–33. Budapest, Akadémiai Kiadó 9.

Lacaille, J. –C. 1991. Postsynaptic potentials mediated by excitatory and inhibitory amino acids in interneurons of *stratum pyramidale* of the CA1 region of rat hippocampal slices *in vitro. J. Neurophysiol.* **66**, 1441–1454.

Lancaster, B. and Adams, P. R. 1986. Calcium-dependent current generating the afterhyperpolarization of hippocampal neurons. *J. Neurophysiol.* **55**, 1268–1282.

Lánský, P. 1983. Inference for the diffusion models of neuronal activity. *Math. Biosci.* **67**, 247–260.

Lánský, P. 1984. On approximations of Stein's neuronal model. *J. Theor. Biol.* **107**, 631–647.

Lánský, P., Giorno, V., Nobile, A. G. and Ricciardi, L. M. 1988. A diffusion neuronal model and its parameters. In: *Biomathematics and Related Computational Problems* , Ricciardi, L. M. (Ed.), pp. 27–37. Kluwer, Dordrecht.

Lánský, P. and Lánská, V. 1987. Diffusion approximation of the neuronal model with synaptic reversal potentials. *Biol. Cybern.* **56**,19–26.

Lánský, P. and Smith, C. E. 1989. The effect of a random initial value in neural first-passage-time models. *Math. Biosci.* **93**,191–215.

Lánský, P., Smith, C. E. and Ricciardi, L. M. 1990. One-dimensional stochastic diffusion models of neuronal activity and related first passage time problems. In: *Trends in Biological Cybernetics* , Menon, J. (Ed.), pp. 153–162. Research Trends, Trivandrum in Kerala State of India.

Lashley, K. S. 1929. *Brain Mechanisms and Intelligence.* Chicago, University of Chicago Press.

Lashley, K. S. 1950. In the search of the engram. In: *Symposium of the Society of Experimental Biology.* Vol. 4. New York, Cambridge University Press.

Lee, P. A. 1979. Some stochastic problems in neurophysiology. *S.E. Asian Bull. Math.* **11**, 205–244.

Lee, W.-L. and Hablitz, J. J. 1989. Involvement of non-NMDA receptors in picrotoxin-induced epileptiform activity in the hippocampus. *Neurosci. Lett.* **107**, 129–134.

Lee, W.-L. and Hablitz, J. J. 1990. Effect of APV and ketamine on epileptiform activity in the CA1 and CA3 regions of the hippocampus. *Epil. Res.* **6**, 87–94.

LeMasson, G., Marder, E. and Abbott, L. F. 1992. Activity-dependent regulation of conductances in model neurons. *Science (in press*).

Lettvin, J. Y., Maturana, H., McCulloch, W. S., and Pitts, W. H. 1959. What the frog's eye tells the frog brain. *Proc. IRE* **47**, 1940–1951.

Levitan, I. B. and Kaczmarek, L. K. 1991. *The Neuron: Cell and Molecular Biology.* Oxford, Oxford University Press.

Li, Z. and Hopfield, J. J. 1989. Modeling the olfactory bulb and its neural oscillatory processings. *Biol. Cybernetics* **61**, 379–392.

Liaw, J.-S., and Arbib, M. A. 1991. A neural network model for response to looming objects by frog and toad. In: *Visual Structures and Integrated Functions* , Arbib, M. A. and Ewert, J.-P. (Eds), pp. 167–180. New York, Heidelberg, Springer-Verlag.

Linsker, R. 1990. Perceptual neural organization: some approaches based on network models and information theory. *Ann. Rev. Neurosci.* **13**, 257–281.

Llinás, R. 1988. The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* **242**, 1654–1664.

Llinás, R., Ribary, U., Capell, J., Yamamoto, T., Winar, R. 1991. An anatomical localization revealed by MEG recordings of the human somatosensory system. *Electroencephalography Clin. Neurophysiol.* **78**, 185–196.

Lorente de Nó, R. 1938. The cerebral cortex: Architecture, intracortical connections and motor projections. In: *Physiology of the Nervous System* , Fulton, J. F. (Ed.), pp. 291–321. London-New York-Toronto, Oxford University Press.

Luhman, H. J. and Prince, D. A. 1990. Control of NMDA receptor-mediated activity by GABAergic mechanisms in mature and developing rat neocortex. *Dev. Brain Res.* **54**, 287–290.

MacGregor, R. J. 1987. *Neural and Brain Modeling.* Academic Press, San Diego.

Makara, G. B., Palkovits, M., Szentágothai, J. 1980. The endocrine hypothalamus and the hormonal response to stress. In: *Selye's Guide to Stress Research* , Vol. 1, Selye, H. (Ed.), pp. 280–337. New York, Van Nostrand Reinhold Co.

Martin, K. A. C. 1984. Neuronal circuits in cat striate cortex. In *Cerebral Cortex* , Vol. 2, Jones, E. G. and Peters, A. (Eds), pp. 241–284. New York, Plenum Press.

Mascagni, M. V. 1989. Numerical methods for neuronal modeling. In: *Methods in Neuronal Modeling,* Koch, C. and Segev, I. (Eds), pp. 439–484. Cambridge MA, MIT Press.

Matsumoto, G., Aihara, K., Ichikawa, M. and Tasaki, A. 1984. Periodic and nonperiodic responses of membrane potentials in squid giant axons during sinusoidal current stimulation. *J. Theoret. Neurobiol.* **43**, 1–4.

Matsumoto, H. and Ajmone Marsan, C. 1964. Cortical cellular phenomena in experimental epilepsy: ictal manifestations. *Exper. Neurol.* **9**, 305–326.

Matsuyama, Y. 1976. A note on stochastic modeling of shunting inhibition. *Biol. Cybern.* **24**, 139–145.

Matsuyama, Y., Shirai, K. and Akizuki, K. 1974. On some properties of stochastic information processes in neurons and neuron populations. *Kybernetik* **15**, 127–145.

May, R. M. 1962. Will a large complex system be stable? *Nature* **238**, 413–414.

McBain, C. and Dingledine, R. 1992. Dual-component miniature synaptic currents in rat hippocampal CA3 pyramidal neurons. *J. Neurophysiol.* **68**, 16–27.

McCauley, J. L. 1988. An introduction to nonlinear dynamics and chaos theory. *Physica Scripta* **T20**, 5–57.

McCormick, D. A. 1990. Membrane properties and neurotransmitter actions. In: *The Synaptic Organization of the Brain* , Shepherd, G. M. (Ed.), pp. 32–66. Oxford, Oxford Univ. Press.

McCormick, D. A., Connors, B. W., Lighthall, J. W. and Prince, D. A. 1985. Comparative electrophysiology of pyramids and sparsely spiny stellate neurons of the neocortex. *J. Neurophys.* **54**, 782–806.

McCulloch, W. S. and Pitts, W. A. 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Mat. Biophys.* **5** , 115–133.

McEliece, R. J., Posner, E. C., Rodemich, E. R. and Venkatesh, S. S. 1987. The capacity of the hopfield associative memory. *IEEE Transactions Inform. Theory* **33**,461–482.

Mead, C. A., Mahowald, M. A. 1988. A silicon model of early visual processing. *Neural Networks* **1**, 91–97.

Meakin, P. 1986. A new model for biological pattern formation. *J. theor. Biol.* **118**, 101–113.

Meinhardt, H. 1982. *Models of biological pattern formation.* Academic Press, London.

Michelson, H. B. and Wong, R. K. S. 1991. Excitatory synaptic responses mediated by $GABA_A$ receptors in the hippocampus. *Science* **253**, 1420–1423.

Miles, R. 1990a. Synaptic excitation of inhibitory cells by single CA3 hippocampal pyramidal cells of the guinea-pig *in vitro*. *J. Physiol. (Lond.)* **428**, 61–77.

Miles, R. 1990b. Variation in strength of inhibitory synapses in the CA3 region of guinea-pig hippocampus *in vitro*. *J. Physiol. (Lond.)* **431**, 659–676.

Miles, R., Traub, R. D. and Wong, R. K. S. 1988. Spread of synchronous firing in longitudinal slices from the CA3 region of the hippocampus. *J. Neurophysiol.* **60**, 1481–1496.

Miles, R. and Wong, R. K. S. 1983. Single neurones can initiate synchronized population discharge in the hippocampus. *Nature* **306**, 371–373.

Miles, R. and Wong, R. K. S. 1984. Unitary inhibitory synaptic potentials in the guinea-pig hippocampus*in vitro*. *J. Physiol. (Lond.)* **356**, 97–113.

Miles, R. and Wong, R. K. S. 1986. Excitatory synaptic interactions between CA3 neurones in the guinea-pig hippocampus. *J. Physiol. (Lond.)* **373**, 397–418.

Miles, R. and Wong, R. K. S. 1987a. Inhibitory control of local excitatory circuits in the guinea-pig hippocampus. *J. Physiol. (Lond.)* **388**, 611–629.

Miles, R. and Wong, R. K. S. 1987b. Latent synaptic pathways revealed after tetanic stimulation in the hippocampus. *Nature* **329**, 724–726.

Miles, R., Wong, R. K. S. and Traub, R. D. 1984. Synchronized afterdischarges in the hippocampus: contribution of local synaptic interaction. *Neuroscience* **12**, 1179–1189.

Mishkin, M. 1982. A memory system in the monkey. *Phil. Trans. R. Soc. Lond.* **B298**, 85–95.

Miyashita, Y. and Chang, H. S. 1988. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* **331**, 68.

Miyashita, Y. 1988. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* **335**, 817.

Mody, I., Lambert, J. D. C. and Heinemann, U. 1987. Low extracellular magnesium induces epileptiform activity and spreading depression in rat hippocampal slices. *J. Neurophysiol.* **57**, 869–888.

Mountcastle, V. B. 1957. Modalities and topographic properties of single neurons of cat's sensory cortex. *J. Neurophysiol.* **20**, 408–434.

Mountcastle, V. B. 1978. An organizing principle for cerebral function: The unit module and the distributed system. In: *The Mindful brain* , Edelman, G. R. and Mountcastle V. B. (Eds), pp. 7–50. MIT-Press, Cambridge.

Muller, D., Joly, M. and Lynch, G. 1988. Contributions of quisqualate and NMDA receptors to the induction and expression of LTP. *Science* **242**, 1694–1697.

Muller, R. U., Kubie, J. L. and Saypoff, R. 1991. The hippocampus as a cognitive graph. *Hippocampus* **1**, 243–246.

Murray, J. D. 1989. *Mathematical Biology* , Springer-Verlag, Berlin.

Murthy, V. N. and Fetz, E. E. 1992. Coherent 25–35 Hz oscillations in the sensorimotor cortex of the awake behaving monkey. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5670–5674.

Nelson, J. I., Salin, P. A., Munk, M. H. J., Arzi, M. and Bullier, J. 1992. Spatial and temporal coherence in cortico-cortical connections: A cross-correlation study in areas 17 and 18 in the cat. *Visual Neurosci.* **9**, 21–37.

Nicholls, J. G., Martin, A. R. and Wallace, B. G. 1992. *From Neuron to Brain.* Sunderland MA, Sinauer.

Nicolis, J. S. and Tsuda, I. 1985. Chaotic dynamics of information processing- The "magic number seven plus–minus two" revisited. *Bull. Math. Biol.* **47**, 343–365.

Nilsson, H. G. 1977. Estimation of parameters in a diffusion neuron model. *Comput. Biomed. Res.* **10**,191–197.

Nobile, A. G., Ricciardi, L. M. and Sacerdote, L. 1985a. Exponential trends of Ornstein-Uhlenbeck first-passage-time densities. *J. Appl. Prob.* **22**, 360–369.

Nobile, A. G., Ricciardi, L. M. and Sacerdote, L. 1985b. Exponential trends of first-passage-time densities for a class of diffusion processes with steady-state distribution. *J. Appl. Prob.* **22**, 611–618.

Numann, R. E., Wadman, W. J. and Wong, R. K. S. 1987 Outward currents of single hippocampal cells obtained from the adult guinea-pig. *J. Physiol. (Lond.)* **393**, 331–353.

Optican, L. M., and Robinson, D. A. 1980. Cerebellar-dependent adaptive control of primate saccadic system. *J. Neurophysiol.* **44**, 1058–1076.

Ott, E., Grebogi, C. and Yorke, J. A. 1990. Controlling chaos. *Phys. Rev. Lett.* **64**, 1196–1199.

Pantev, C., Makeig, S., Hoke, M., Galambos, R., Hampson, S., Gallen, C. 1991. Human auditory evoked gamma band magnetic fields. *Proc. Natl Acad. Sci. U.S.A.* **88**, 8996–9000.

Pardoux, E. and Talay, D. 1986. Discretization and simulation of stochastic differential equations. *Acta Appl. Math.* **3**, 23–47.

Paré, D., de Curtis, M. and Llinás, R. 1992. Role of the hippocampal-entorhinal loop in temporal lobe epilepsy: extra- and intracellular study in the isolated guinea pig brain *in vitro*. *J. Neurosci.* **12**, 1867– 1881.

Park, C. and Schuurmann, F. J. 1976. Evaluations of barrier-crossing probabilities of Wiener paths. *J. Appl. Prob.* **13**, 267–275.

Park, C. and Schuurmann, F. J. 1980. Evaluations of absorption probabilities for the Wiener process on large intervals. *J. Appl. Prob.* **17**, 363–372.

Pellionisz, A. J. 1984. Coordination: a vector-matrix description of transformations of overcomplete CNS coordinates and a tensorial solution using the Moove–Peurose generalized inverse. *J. Theor. Biol.* **110**, 353–375.

Pellionisz, A. J. 1991. Discovery of neural geometry by neurobiology and its utilization in neurocomputer theory and development. In: *Artificial Neural Networks* , Vol. 1, Kohonen, T. *et al.* (Eds), pp. 485–493. North-Holland, Amsterdam.

Pellionisz, A. and Llinás, R. 1982. Space–time representation in the brain. The cerebellum as a predictive space–time metric tensor. *Neuroscience* **7**, 2949–2970.

Peretto, P. and Niez, J.-J. 1986. Stochastic dynamics of neural networks. *IEEE Trans. Syst. Man. Cyber.* **16**, 73–83.

Perkel, D. H., Gerstein, G. L., Moore G. P. 1967a. Neuronal spike trains and stochastic point processes. I. The single spike train. *Biophys. J.* **7**, 391–418.

Perkel, D. H., Gerstein, G. L., Moore, G. P. 1967b. Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys. J.* **7**, 419–440.

Perkel, D. H., Gerstein, G. L., Smith, M. S. and Taton, W. G. 1975. Nerve impulse patterns: A quantitative technique for three neurons. *Brain Res.* **100**, 271–296.

Pernier, J. 1972. Adjustement automatique des densites de probabilité d'intervalles entre potentiels d'action selon la loi de Wiener. *Biometrics* **28**, 737–745.

Peters, A., Sethares, C. 1991. Layer IV.A of rhesus monkey primary visual cortex. *Cerebral Cortex* **1**, 445–462.

Petrov, V., Peng, B. and Showalter, K. 1992. A map-based algorithm for controlling chaos. *J. Chem. Phys.* **10**, 7506–7513.

Rakic, P. 1975. Local circuit neurons. *Neurosci. Res Prog. Bull.* **13**, 289–446.

Rall, W. 1959. Branching dendritic trees and motoneuron membrane resistivity. *Exp. Neurol.* **2**, 503–532.

Rall, W. 1964. Theoretical significance of dendritic trees for neuronal input-output relations. In: *Neural Theory and Modeling,* Reiss, R. (Ed.), pp. 73–97. Stanford CA, Stanford Univ. Press.

Rall, W. 1977. Core conductor theory and cable properties of neurons. In: *Handbook of Physiology,* Vol. 1, Kandel, E. R. (Ed.), pp. 39–98. Bethesda MD, Amer. Physiol. Soc.

Ramón y Caial, S. 1894. Les nouvelles idées sur la structure du Système nerveux chez l'homme et les vértebrés. Paris, C. Reinwald et Cie.

Ramón y Caial, S. 1900. Studien über die Hirnrinde des Menschen. Leipzig, J A. Barth.

Ramón y Caial, S. 1911. Histologie du Système Nerveux de l'Homme et des Vertébrés. **I–II**. Paris, Maloine.

Ramon-Moliner, E. and Nauta, W. J. H. 1966. The isodendritic core of the brain stem. *J. Comp. Neurology* **126**, 311–335.

Rapoport, A. 1948. Cycle distribution in random nets. *Bull. Math. Biophys.* **10**, 145–157.

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

Rapp, M., Yarom, Y. and Segev, I. 1991. The impact of parallel background activity on the cable properties of cerebellar Purkinje cells. *Neural Comp.* **4**, 518–533.

Rapp, P. E., Zimmerman, I. D., Albano, A. M., deGuzman, G. C. and Greenbaum, N. N. 1985. Dynamics of spontaneous neural activity in the simian motor cortex: The dimension of chaotic neurons. *Phys. Lett.* **110A**, 335–338.

Rashevsky, N. 1938. *Mathematical Biophysics.* Chicago University Press, Chicago.

Rensing, L., van der Heiden, U., Mackey, M. C. 1987. *Temporal Disorder in Human Oscillatory Systems.* Springer, Berlin.

Réthelyi, M. 1976. Central core in the spinal grey matter. *Acta Morph. Acad. Sci. Hung.* **24**, 63–70.

Rexed, S. 1954. A cytoarchitectonic atlas of the spinal cord in the cat. *J. Comp. Neurol.* **100**, 297–379.

Ricciardi, L. M. 1976. Diffusion approximation for a multi-input model neuron. *Biol. Cybern.* **24**, 237–240.

Ricciardi, L. M. 1977. *Diffusion Processes and Related Topics in Biology. Lecture Notes in Biomathematics* , Vol. 14. Springer, Berlin.

Ricciardi, L. M. 1982. Diffusion approximation and computational problems for single neurons activity. In: *Competition and Cooperation in Neural Nets. Lecture Notes in Biomathematics* , Vol. 45, Amari, S. and Arbib, M. A. (Eds), pp. 143–154. Springer, New York.

Ricciardi, L. M. and Sacerdote, L. 1979. The Ornstein–Uhlenbeck process as a model for neuronal activity. *Biol. Cybern.* **35**, 1–9.

Ricciardi, L. M., Sacerdote, L. and Sato, S. 1983. Diffusion approximation and first passage time problem for a model neuron, II. Outline of a computation method. *Math. Biosci.* **64**, 29–44.

Ricciardi, L. M., Sacerdote, L. and Sato, S. 1984. On an integral equation for first-passage-time probability densities. *J. Appl. Prob.* **21**, 302–314.

Ricciardi, L. M. and Sato, S. 1983. A note on the evaluation of first-passage-time probability densities. *J. Appl. Prob.* **20**, 197–201.

Ricciardi, L. M. and Sato, S. 1986. On the evaluation of first passage time densities for Gaussian processes. *Signal Processing* **11**, 339–357.

Ricciardi, L. M., Di Crescenzo, A., Iardino, F. and Rinaldi S. 1991. Vectorized simulations of normal process and the first passage time probability problem. In: *Proceedings of the International Conference on Noise in Physical Systems and 1/f Fluctuations* , Musha, T., Sato, S. and Yamamoto, M. (Eds), pp. 383–388. Ohmsha, Ltd, Tokyo.

Ricciardi, L. M., Di Crescenzo, A., Giorno, V. and Nobile, A. G. 1992. On the instantaneous return process for neuronal diffusion models (*submitted*).

Rinzel, J. 1985. Excitation dynamics: insights from simplified membrane models. *Fed. Proc.* **44**, 2944–2946.

Rinzel, J. M. and Ermentrout, G. B. 1989. The geometry of excitability. In: *Methods of Neuronal Modeling* , Koch, C. and Segev, I. (Eds), pp. 135–170. Boston, MIT.

Rinzel, J. and Frankel, P. 1992. Activity patterns of a slow synapse network predicted by explicitly averaging spike dynamics. *Neural Computation* **4**, 534–535.

Rinzel, J. and Rall, W. 1974a. Voltage transients in neuronal dendritic trees. *Biophys. J.* **14**, 759–790.

Rinzel, J. and Rall, W. 1974b. Branch input resistance and steady attenuation for input to one branch of a dendritic neuron model. *Biophys. J.* **13**, 648–688.

Rinzel, J., Sherman, A. and Stokes, C. L. 1992. Channels, coupling, and synchronized rhythmic bursting activity. In: *Analysis and Modeling of Neural Systems* , Eeckman, F. (Ed.), Kluwer Academic Publishers.

Robinson, D. A. 1964. The mechanics of human saccadic eye movement. *Journal of Physiology* **174**, 245–264.

Robinson, D. A. 1970. Oculomotor unit behavior in the monkey. *Journal of Neurophysiology* **33**, 393–404.

Robinson, D. A. 1975. Oculomotor control signals. In:*Basic Mechanisms of Ocular Motility and their Clinical Implications* , Lennerstrand, G. and Bach-y-Rita, P. (Eds), pp. 337–374. Oxford: Pergamon.

Rodieck, R. W., Kiang, N. Y.-S. and Gerstein, G. L. 1962. Some quantitative methods for the study of spontaneous activity of single neurons. *Biophys. J.* **2**, 351–368.

Rose, R. M. and Hindmarsh J. L. 1982. A model of the nerve impulse using two first-order differential equations. *Nature* **296**, 162–174.

Rosenblatt, F. 1962. *Principles of Neurodynamics.* Washington, DC, Spartan Books.

Roy, B. K. and Smith, D. R. 1969. Analysis of the exponential decay model of the neuron showing frequency threshold effects. *Bull. Math. Biophys.* **31**, 341–357.

Rumelhart, D. E., Hinton, G. E., Williams, R. J. 1986. Learning internal representations by error propagation. In: *Parallel Distributed Processing* , Vol. 1, Rumelhart, D. E., McClelland, J. L. and the PDP Research Group (Eds), pp. 318–362. Cambridge, MA., A Bradford Book, MIT Press.

Rümelin, W. 1982. Numerical treatment of stochastic differential equations. *SIAM J. Numer. Anal.* **3**, 604–613.

Sah, P., Hestrin, S. and Nicoll, R. A. 1990. Properties of excitatory post-synaptic currents recorded *in vitro* from rat hippocampal interneurones. *J. Physiol. (Lond.)* **430**, 605–616.

Sakai, K. and Miyashita, Y. 1991. Neural organization for the long-term memory of paired associates. *Nature* **354**, 152.

Sakmann, B. and Neher, E. (Eds). 1983. *Single-channel recording* , Plenum Press, New York.

Sammon, J. W. 1969. A nonlinear mapping for data structure analysis. *IEEE Trans. Computers* **18**, 401–409.

Sampath, G. and Srinivasan, S. K. 1977. *Stochastic Models for Spike Trains of Single Neurons. Lecture Notes in Biomathematics* , Vol. 16. Springer, New York.

Sato, S. 1978. On the moments of the firing interval of the diffusion approximated model neuron. *Math. Biosci.* **39**, 53–70.

Sbitnev, V. J. 1975. Transport of spikes in statistical neuron ensembles. Conception of phase transition. *Akad. Nauk CCCP* **176**, 1–24 (in Russian).

Scarfman, H. E., Kunkel D. D. and Schwartzkroin, P. A. 1990. Synaptic connections of dentate granule cells and hilar neurons: Result of paired intracellular recordings and intracellular horseradish peroxidase injections. *Neuroscience* **37**, 693–707.

Scheibel, M. E., Scheibel, A. B. 1958. Structural substrates for integrative patterns in the brain stem reticular core. In: *Reticular Formation of the Brain* , Jaser, H. H., Proctor, L. D., Knighton, R. S., Noshay, W. C., Costello, R. T. (Eds), pp. 31–55. Boston, Mass., Little Brown and Co.

Schierwagen, A. 1987. *Neuronal Form-Function Relationship, Fractals and Diffusive Growth Models.* KMU-Leipzig, NTZ.

Schmidt, R. F. (Ed.). 1985. *Fundamentals of Neurophysiology* , 3rd Edn. Springer, N. Y.

Schmitt, F. O, Dev, P. and Smith, B. H. 1976. Electronic Processing of information by brain cells. *Science* **193**, 114–120.

Schneiderman, J. H. 1986. Low concentrations of penicillin reveal rhythmic, synchronous synaptic potentials in hippocampal slice. *Brain Res.* **398**, 231–241.

Schneiderman, J. H. and MacDonald, J. F. 1987. Effects of reduced magnesium on hippocampal synchrony. *Brain Res.* **410**, 174–178.

Schwartzkroin, P. A. and Haglund, M. M. 1986. Spontaneous rhythmic synchronous activity in epileptic human and normal monkey temporal lobe. *Epilepsia* **27**, 523–533.

Schwartzkroin, P. A. and Prince, D. A. 1978. Cellular and field potential properties of epileptogenic hippocampal slices. *Brain Res.* **147**, 117–130.

Scudder, C. A. 1988. A new local feedback model of the saccadic burst generator. *J. Neurophysiol.* **59**, 1455–1475.

Segev, I., Fleshmann, J. W. and Burke, R. E. 1989. Compartmental models of complex neurons. In: *Methods in Neuronal Modeling*, Koch, C. and Segev, I. (Eds), pp. 63–96. Cambridge MA, MIT Press.

Segundo, J. P. 1986. What can neurons do to serve as integrating devices. *J. Theor. Neurobiol.* **5**, 1–59.

Shepherd, G. M. 1983. *Neurobiology.* New York, Oxford University Press

Shepherd, G. 1990. *The Synaptic Organization of the Brain.* New York, Oxford University Press.

Sherman, A. and Rinzel, J. 1992. Rhythmogenic effects of weak electrotonic coupling in neuronal models. *Proc. Natl. Acad. Sci U.S.A.* **89**, 2471–2474.

Shiino, M. and Fukai, T. 1990. Replica-symmetric theory of the non-linear analogue neural networks. *J. Phys. A: Math. Gen.* **23**, L1009.

Siebert, W. M. 1969. On the stochastic neural models of the diffusion type. *Quart. Prog. Rep. M.I.T. Res. Lab. Electronics* **94**, 281–287.

Singer, W. 1990. The formation of cooperative cell assemblies in the visual cortex. *J. Exp Biol.* **153**, 177–197.

Skarda, C. A. and Freeman, W. J. 1987. How brains make chaos in order to make sense of the world. *Behav. Brain. Sci.* **10**, 161–195.

Škvaŕil, J., Radil-Weiss, T., Bohdanecký, Z. and Syka, J. 1971. Spontaneous discharge patterns of mesencephalic neurones, interval histogram and mean interval relationship. *Kybernetik* **9**, 11–15.

Slack, J. M. W. 1987. Morphogenetic gradients—past and present. *Trends Biochem.* **12**, 200–204.

Smith, C. E. and Goldberg, J. M. 1986. A stochastic afterhyperpolarization model of repetitive activity in vestibular afferents. *Biol. Cybern.* **54**, 41–51.

Sompolinsky, H. 1987. The theory of neural networks: The Hebb rule and beyond. In: *Heidelberg Colloquium on Glassy Dynamics*, van Hemmen, L. and Morgenstern, I. (Eds), Heidelberg, Springer-Verlag.

Soong, A. C. K. and Stuart, C. I. J. M. 1989. Evidence of chaotic dynamics underlying the human alpha-rhythm electro-encephalogram. *Biol. Cybernetics* **62**, 55–62.

Sparks, D. L., and Jay, M. F. 1985. The functional organization of the primate superior colliculus: a motor perspective. In: *Progress in Brain Research* , Vol. 64, H.-J. Freund, U. Büttner, B. Cohen, and J. Noth (Eds), Amsterdam: Elsevier.

Sparks, D. L., and Mays, L. E. 1980. Movement fields of saccade-related burst neurons in the monkey superior colliculus. *Brain Res.* **190**, 39–50.

Squire, L. R. 1986. Mechanisms of memory. *Science* **232**, 1612–1619.

Squire, L. R, Shimamura, A. P, Amaral, D. G. 1989. Memory and the hippocampus. In: *Neural Models of Plasticity* , Byrne, J. H. and Berry, W. O. (Eds), pp. 208–239. Academic Press, San Diego.

Standing, L. 1973. Learning 10,000 pictures, *Quart. J. Exp. Psychol.* **25**, 207.

Stanley, H. E. and Ostrowsky, N. (Eds). 1986. *On Growth and Form.* Nijhoff Publ, Boston.

Storm, J. 1988. Temporal integration by a slowly inactivating K^+ current in hippocampal neurons. *Nature* **336**, 379–381.

Strogatz, S. and Mirollo, R. 1990. Synchronization of pulse coupled oscillators. *SIAM J. Appl. Math.* **50**, 1645–1662.

Stryker, M. P. and Harris, W. A. 1986. Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *J. Neuroscience* **6**, 2117–2133.

Sugiyama, H., Moore, G. P. and Perkel, D. H. 1970. Solutions for a stochastic model of neuronal spike production. *Math. Biosci.* **8**, 323–341.

Swindale, N. W. 1990. Is the cerebral cortex modular? *Trends Neurosci.* **13**, 487–492.

Székely, G., Czéh, G. 1971. Muscle activities recorded simultaneously from normal and supernumerary forelimbs in ambystoma. *Acta Physiol. Acad. Sci. Hung.* **40**, 287–301.

Székely, G., Szentágothai, J. 1962. Experiments with "model nervous systems". *Acta Biol. Acad. Sci. Hung.* **12**, 253–269.

Szentágothai, J. 1962. On the synaptology of the cerebral cortex. In: *Structure and Function of the Nervous System* , Sarkissov, S. A. (Ed.), pp. 6–14. Moscow, Medgiz.

Szentágothai, J. 1969. Architecture of the cerebral cortex. In: *Basic mechanisms of the Epilepsies* , Jasper, H. H., Ward, A. A., Pope, A. (Eds), pp. 13–28. Boston, Little, Brown and Co.

Szentágothai, J. 1975. The module concept in cerebral cortex architecture. *Brain. Res.* **95**, 475–496.

Szentágothai, J. 1978. Specificity versus (quasi-) randomness in cortical connectivity. In: *Architectonics of the Cerebral Cortex.* , Brazier, M. A. B., Petsche, H. (Eds), pp. 77–79. New York, Raven Press.

Szentágothai, J. 1981. Visual cortex circuitry: revisited. *Freiburger Universitätsblütter* **74**, 47–49.

Szentágothai, J. 1983. The modular architectonic principle of neural centers. *Rev. Physiol. Biochem. Pharmacol.* **98**, 11–61.

Szentágothai, J. 1983. Functional anatomy of the visual centers as cues for pattern recognition concepts. In: *Pontif. Acad. Sci. Scripta Varia* , **54**, Chagas, Ch., Gattass, R., Gross, Ch. (Eds), pp. 39–52.

Szentágothai, J. 1987. The brain–mind relationship. In: *The Brain–Mind Problem* , Gulyas, B. (Ed.), pp. 61–80. Leuven, University Press.

Szentágothai, J. 1990. "Specificity versus (quasi-) randomness" revisited. *Acta Morphol. Hung.* **38**, 159–167.

Szentágothai, J. 1993. (*In press*).

Szentágothai, J. and Arbib, M. 1974. Conceptual models of neural organization. *NRP Bull.* **12(3)**.

Szentágothai, J., Érdi, P. 1989. Self-organization in the nervous system. *J. Social. Biol. Struct.* **12**, 367–384.

Szentágothai, J., Réthelyi, M. 1973. Cyto and neuropil architecture of the spinal cord. In: *New Developments in Electromyography and Clinical Neurophysiology* , **3**, Desmedt, J. E. (Ed.), pp. 20–37. Basel, Karger. .

Tancredi, V., Avoli, M. and Hwa, G. G. C. 1988. Low-magnesium epilepsy in rat hippocampal slices: inhibitory postsynaptic potentials in the CA1 subfield. *Neurosci. Lett.* **89**, 293–298.

Thalmann, R. H. 1984. Reversal properties of an EGTA-resistant late hyperpolarization that follows synaptic stimulation of hippocampal neurons. *Neurosci. Lett.* **46**, 103–108.

Thompson, R. F. 1986. The neurobiology of learning and memory. *Science* **233**, 941–947.

Thompson, R. F. and Berger, T. W., Berry, S. D., Hoeler, F. K. 1980. The search for the engram, II. In: *Neural Mechanisms in Behavior* , McFadden D. (Ed.), pp. 172–222. New York, Springer.

Thompson, W. J., Bicker, G., Changeaux, J. P., Ebendal, T. L., Heisenberg, M., Henderson, C. E., Huttner, W., Kandel, E. R., Mallet, J. B., Stent, G. S., Thoenen, H., Yaniv, M. 1987. Activity-dependent regulation of gene expression. In: *The Neural and Molecular Bases of Learning* , Changeux, J.-P. and Konishi, M. (Eds), pp. 13–30. Wiley, New York-London.

Thompson, J. M. T., Stewart, H. B. 1986. *Nonlinear Dynamics and Chaos.* New York, Wiley.

Thomson, A. M., Girdlestone, D. and West, D. C. 1988. Voltage-dependent currents prolong single-axon postsynaptic potentials in layer III pyramidal neurons in rat neocortical slices. *J. Neurophysiol.* **60**, 1896–1907.

Thomson, A. M., Walker, V. E. and Flynn, D. M. 1989. Glycine enhances NMDA-receptor mediated synaptic potentials in neocortical slices. *Nature* **338**, 422–424.

Thorendike E. L. 1911. *Animal Inteligence.* pp. 244–250. Hafner, New-York (*reprinted 1965*).

Traub, R. D. and Miles, R. 1991. *Neuronal Networks of the Hippocampus.* New York, Cambridge University Press.

Traub, R. D., Miles, R. and Buzsüki, G. 1992. Computer simulation of carbachol-driven rhythmic population oscillations in the CA3 region of the *in vitro* rat hippocampus. *J. Physiol. (Lond.)* **451**, 653–672.

Traub, R. D., Miles, R. and Jefferys, J. G. R. 1993. Synaptic and intrinsic conductances shape picrotoxin-induced synchronized afterdischarges in the guinea-pig hippocampal slice. *J. Physiol. (Lond.)* , (*in press*).

Traub, R. D., Miles, R., Muller, R. U. and Gulyás, A. I. 1992. Functional organization of the hippocampal CA3 region: implications for epilepsy, brain waves and spatial behaviour. *Network* **3**, 465–488

Traub, R. D., Miles, R. and Wong, R. K. S. 1987. Models of synchronized hippocampal bursts in the presence of inhibition. 1. Single population events. *J. Neurophysiol.* **58**, 739–751.

Traub, R. D., Miles, R., Wong, R. K. S. 1988: "Large scale simulations of the hippocampus". *IEEE Eng Med. Biol Magazine* , 31–38.

Traub, R. D., Miles, R. and Wong, R. K. S. 1989. Model of the origin of rhythmic population oscillations in the hippocampal slice. *Science* **243**, 1319–1325.

Traub, R. D., Miles, R., Wong, R. K. S., Schulman, L. S. and Schneiderman, J. H. 1987. Models of synchronized hippocampal bursts in the presence of inhibition. 2. Ongoing spontaneous population events. *J. Neurophysiol.* **58**, 752–764.

Traub, R. D. and Wong, R. K. S. 1982. Cellular mechanism of neuronal synchronization in epilepsy. *Science* **216**, 745–747.

Traub, R. D., Wong, R. K. S., Miles, R. and Michelson, H. 1991. A model of a CA3 hippocampal pyramidal neuron incorporating voltage-clamp data on intrinsic conductances. *J. Neurophysiol.* **66**, 635–650.

Ts'o, D. Y., Frostig, R. D., Lieke, E. E. and Grinvald, A. 1990. Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science* **249**, 417–420.

Tsodyks, M. V. and Feigel'man, M. V. 1988. The enhanced storage capacity in neural networks with low activity level. *Europhys. Lett.* **46**, 101.

Tsonis, P. A. 1987. The nature of positional information. *Trends Biochem.* **12**, 249.

Tsuda, I. 1984. A hermeneutic process of the brain. *Prog. Theor. Phys. Suppl.* **79**, 241–259.

Tsuda, I. 1991. Chaotic itinerancy as a dynamical basis of hermeneutics in brain and mind. *World Futures* **32**, 167–184.

Tsuda, I. 1992. Dynamic link of memory—chaotic memory map in monequilibrium meural networks. *Neural Networks* **5**, 313–326.

Tsuda, I. and Shimizu, H. 1985. Self-organization of the dynamics channel. In: *Complex systems—Operational approaches* , Haken, H. (Ed.), pp. 240–251. Springer, Berlin.

Tuckwell, H. C. 1988. *Introduction to Theoretical Neurobiology* , Vol. 2. Nonlinear and Stochastic Theories. Cambridge Univ. Press, Cambridge.

Tuckwell, H. C. 1989. *Stochastic Processes in the Neurosciences.* SIAM, Philadelphia.

Tuckwell, H. C. and Cope, D. K. 1980. Accuracy of neuronal interspike times calculated from a diffusion approximation. *J. Theor. Biol.* **83**, 377–387.

Tuckwell, H. C. and Richter, W. 1978. Neuronal interspike time distributions and the estimation of neurophysiological and neuroanatomical parameters. *J. Theor. Biol.* **71**, 167–183.

Turing, A. M. 1952. The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc. Lond B* **237**, 37–72.

Tyson, J. J. 1975. Classification of instabilities in chemical reaction systems. *J. Chem. Phys.* **62**, 1010–1015.

Vaadia, E. and Aertsen, A. 1992. Coding and computation in the cortex: Single-neuron activity and cooperative phenomena. In: *Information Processing in the cortex,* Aertsen, A. and Braitenberg, V. (Eds), pp. 81–121. Springer, Berlin.

van Gisbergen, J. A. M., Robinson, D. A., and Gielen, S. 1981. A quantitative analysis of generation of saccadic eye movements by burst neurons. *J. Neurophysiol.* **45**, 417–442.

Ventriglia, F. 1974. Kinetic approach to neural systems. I. *Bull. Math. Biol.* **36**, 534–544.

Ventriglia, F. 1978. Propagation of excitation in a model of neural system. *Biol. Cybern.* **30**, 75–79.

Ventriglia, F. 1980. Numerical investigations for one-dimensional neural systems. *Biol. Cybern.* **36**, 125–130.

Ventriglia, F. 1983. Kinetic theory of neural systems: Study of the two-dimensional model. *Biol. Cybern.* **46**, 93–99.

Ventriglia, F. 1988a. Learning in a cortical-like neural system. In: *Biomathematics and Related Computational Probems* , Ricciardi, L. M. (Ed.), pp. 99–108. Dordrecht, Kluwer Academic Publ.

Ventriglia, F. 1988b. Computational simulation of cortical-like neural systems *Bull. Math. Biol.* **50**, 143–185.

Ventriglia, F. 1990. Activity in cortical-like neural systems: short-range effects and attention phenomena. *Bull. Math. Biol.* **52**, 397–429.

Ventriglia, F. 1990. Kinetic modeling in neural systems. In: *Trends in Biological Cybernetics* **1**, Menon, J. (Ed.), pp. 103–114. CRSI, Sreekanteswaram, Trivandrum, India.

Ventriglia, F. 1990. Towards a kinetic theory of some global brain activities. *Acta Neurol.* **52**, 1–17.

Ventriglia, F. 1991. Attention and memory according to a kinetic theory of neural systems. In: *Mathematical Approaches to Brain Functioning Diagnostics (Proceedings in Nonlinear Sciences)* , Dvorak, I. and Holden, A. V. (Eds), pp. 93–106. Manchester University Press, Manchester.

Ventriglia, F. and Érdi, P. 1987. Statistical approach to the dynamics of cerebral cortex: learning aspects. In: *Cybernetics and Systems: the way ahead,* Vol. 1, Rose, J. (Ed.), pp. 443–447. Thales Publications Ltd, Lytham St. Annes, Enland.

Ventriglia, F. and Raddi, A. 1993. Synaptic activity in a kinetic theory of neural systems *(submitted)*).

Vertes, R. P. 1986. Brainstem modulation of the hippocampus: anatomy, physiology and significance. In: *The Hippocampus* , Vol. 4, Isaacson, R. L. and Pribram, K. H. (Eds), pp. 41–75. Plenum Press, New York.

Viana Di Prisco, G., Freeman, W. J. 1985. Odor-related bulbar EEG spatial pattern analysis during appetitive conditioning in rabbits. *Behav. Neurosci.* **99**, 962–978.

Vinogradova, O. S. 1975. Functional organization of the limbic system in the process of registration of information: facts and hypotheses. In: *The Hippocampus* , Vol. 2, Isaacson, R. L. and Pribram, K. H. (Eds), pp. 3–69. Plenum Press, New York-London.

Vinogradova, O. S. and Brazhnik, E. S. 1986. Organization and modulation of theta-rhythm by the rhythmically bursting septal neurons. In: *Learning and Memory* , Matthies, H. (Ed.), pp. 141–142. Pergamon Press, Oxford-New York.

Wake, N., Fukushima, K. 1992. Improved learning method for the neocognitron. (*in Japanese*). *IEICE Transactions* , D-II (Inst. Electronics, Information, Commun. Engineers, Japan) **J75-D-II**.

Walter, W. G. 1953. *The Living Brain.* London: Duckworth.

Weisbuch and Fogelman-Souliè. 1985. Scaling laws for the attractors of Hopfield networks. *J. Physique Lett.* **2**, 337.

West, B. 1986. *An essay on the importance of being nonlinea* r. Springer, Berlin.

White, E. L. *Cortical Circuits: Synaptic Organization of the Cerebral Cortex.* Birkhauser, Boston.

Wiesel, T. N. 1982. Postnatal development of the visual cortex and the influence of environment. *Nature* **299**, 583–591.

Wigner, E. P. 1959. Statistical properties of real symmetric matrices with many dimensions. *Proceedings of the Fourth Canadian Mathematics Congress* , MacPhail, M. S. (Ed.), pp. 174–184. Univ. Toronto Press, Toronto.

Williams, S. H. and Johnston, D. 1991. Kinetic properties of two anatomically distinct excitatory synapses in hippocampal CA3 pyramidal neuron. *J. Neurophysiol.* **66**. 1010–1020.

Williams, T., Sigvardt, K., Kopell, N., Ermentrout, G. B., and Remler, M. 1990. Forcing of coupled nonlinear oscillators: Studies of intersegmental coordination of the lamprey locomotor central pattern generator. *J. Neurophys.* **64**, 862–871.

Williamson, S. J., Kaufmann, L. 1989. Advances in neuromagnetic instrumentation and studies of spontaneous brain activity. *Brain Topography* **2**, 9–18.

Wilson, H. R. and Cowan, J. D. 1972. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical J.* **12**, 1–24.

Wilson, H. R. and Cowan, J. D. 1973. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80.

Winfree, A. T. 1967. Biological rhythms and the behavior of populations of coupled oscillators. *J. Theoret. Biol.* **16**, 15–42.

Winfree, A. T. 1980. *The Geometry of Biological Time.* New York, Springer-Verlag.

Wise, M. E. 1981. Spike interval distributions for neurones and random walks with drift to a fluctuating threshold. In: *Statistical Distributions in Scientific Work* , Vol. 6, Taillie, C., Patil G. P. and Baldessari, B. (Eds), Reidel, Dordrecht-Holland.

Wolpert, L. 1969. Positional information and pattern formation. *J. theor. Biol.* **25**, 1–47.

Wong, R. K. S., Miles, R. and Traub, R. D. 1984. Local circuit interactions in synchronization of cortical neurones. *J. Exp. Biol.* **112**, 169–178.

Wong, R. K. S. and Prince, D. A. 1981. Afterpotential generation in hippocampal pyramidal cells. *J. Neurophysiol.* **45**, 86–97.

Wong, R. K. S., Prince, D. A. and Basbaum, A. I. 1979. Intradendritic recordings from hippocampal neurons. *Proc. Natl Acad. Sci. U.S.A.* **76**, 986–990.

Woodbury J. W. 1960. Potentials in volume conductor. In: *Medical physiology and biophysics* , 18th Edn, Ruch, T. C. and Fulton, J. F. (Eds.), pp. 83–91. Saunders, Philadelphia.

Woody, C. D. 1982. *Neurons, Learning and Higher Functions.* Springer, New York-Heidelberg-Berlin.

Yamada, W. M., Koch, C. and Adams, P. R. 1989. Multiple channels and calcium dynamics. In: *Methods of Neuronal Modeling* , Koch, C. and Segev, I. (Eds), pp. 97–134. CambridgeMA, MIT Press.

Yang, G. L. and Chen, T. C. 1978. On statistical methods in neuronal spike-train analysis. *Math. Biosci.* **38**, 1–34.

Yao, Y., Freeman, W. J. 1990. Model of biological pattern recognition with spatially chaotic dynamics. *Neural Networks* **3**, 153–170.

Yasuda, M., Fukushima, K., Nagata, S., Yamaguchi, Y. 1971. An electronic model of the visual system (*in Japanese*). *NHK Technical Journal* **23**, 368–377.

Subject Index

absolute novelty 240
absolute refractory period (ARP) 229, 247, 260, 277
absorbing boundary 136
absorption coefficient 223
action potential 44, 68, 69, 77, 112, 113
    fast 112
    slow 112
Ad fibers 6
afterdischarges 117
afterhyperpolarization 117
alpha rhythm 197
AM pattern 193, 204
AMPA 113, 117, 120
    conductances 120
    current 67
amplitude
    modulation (AM) 193
    of the EPSP 49
anisotropic coupling 106
anterior
    nucleus 205
    thalamic nuclei 8
    ventral temporal (AVT)
cortex 256, 257
antiphase 83
approach 24
archicortex 12
Arnold tongues 87
array of electrodes 190, 196
associative
    learning 198, 210
    reinforcement 241
asymmetric
    dilution 282
    disruption 272
asymptotic
    phase shift 80
    stability 104
asymptotically stable periodic solution 89
attention process 239
attractors 139, 143, 201, 255, 256, 259, 261, 274
    wings 206
    of Lorenz 206
auditory cortex 197
auto associative memory 261
auto-correlation 46, 197
autonomous systems 175
autopoietic 6
averaged evoked potential 189
avoidance 24, 27
axo-axonic interneuron 13
back propagation 213, 286
basin of attraction 201, 262, 275
basket cells 13
bifurcation theory 258
biological model 24
Boltzmann factor 62
brainstem 2, 6, 26, 197
bulb cortex 198
burst of action potentials 115
CA1 111, 218
CA3 111, 218
    network model 122
cable
    equation 59, 60, 74
    length constant 60
carrier wave 197, 212
catecholaminergic neurons 6
caudate nucleus 40
cell
    assembly 198, 206
    membrane 58
cerebellar cortex 8, 9
cerebellum 6, 32
cerebral cortex 10
chain of oscillators 102, 105
channels 60, 62, 112
chaos 173, 212
    deterministic 173, 202
    controlling 174

chaotic
activity 180
behavior 87
information processing 177
itinerancy 197
chemical synapse 87, 90, 237
climbing fibers 6
closed states 61
cognitive
computations 251
psychology 258
cognitron 286
collective behavior 259
commisural path 13
common inputs 48
compartment 74
compartmental equations 74
competitive learning procedure 288
complex cells 286, 290
spikes 113
computational experiments 231, 237, 244
computer simulation 72, 285, 289, 301, 304
conductance 90
noninactivating 63
inactivating 64
connected characters
segmentation 275
recognition 275
conservation of probabilistic mass 228, 246
continuum neural network 108
continuous-time stochastic process 140
cooperative
collective action 193
effects 54
correlation 257, 265
coefficient 50
of non-stationary processes 53
cortical column 193
mappings 234
module 11
cortico-cortical connectivity 11
covariance function 50
cross-
correlation 46
correlogram 46, 48
intensity function 46
renewal density 46
crystalline structure 12
current to rate transduction function 277
cursive handwriting words 302
cytoarchitectonic 10
delay memory saccade 37
delayed-rectifier potassium conductance 64
dendrites 112
apical 112
basilar 112
dendritic
arborization 163
Ca spike 121
currents 187
dentate gyrus 12
depolarization 58, 64
depolarizing hump 120
deterministic autonomous differential equations 80
developmental mechanisms 164
diffusion process 133, 135
diffusive coupling 88, 104
dipole 41
direct current (DC) potentials 115
dispersion relation 104
distorted patterns 290
divergence 220
dopaminergic neurons 6
dorsomedial medullary reticular formation 32
double
barrel structure 3, 6
-period limit cycle 178
drift 104
dripping faucet 169
dynamic
remapping 14, 37
disease 174
electroencephalogram (EEG) 115, 187, 202, 223, 244
records 205
spectrum 202
electric field 41
electrode 58
electronic network 285
electroolfactogram 187

electrotonic length 77
engrams 15, 214
ensemble
    frequency 104
    time average 189, 190
entorhinal cortex 12, 238
epigenetic stabilization 241
epilepsy 111, 115, 235
epileptic 113, 173
    seizures 13
equivalent cylinder 77
event related activity 51
evoked potential 202
excitatory burst neurons 40
extracellularly recoded field 43
extreme random asymmetric dilution 272
failures of transmission 220
false alarm 258
fascia dentata 12, 238, 239
fast conducted pain 4
fast sodium conductance 66
fatigue 301, 304
feature-extracting cells 290
feed-back
    inhibition 13
    negative 13
    positive 13
feed-forward inhibition 13
field potential recording 114
firing rates 78
first passage time (FTP) 129, 277
Floquet multipliers 80, 84, 94
focus 83
focused attention 240, 300
fornix 13
fractal dimension 173, 213
fractals 164
Fredholm alternative 91
frequency gradient 106
frontal eyefields 38
FTP
    problem 136, 146
    random variables 133
GABA 112, 119
$GABA_A$ 68, 114, 115, 117, 122
$GABA_B$ 68, 114, 118
gain of the synapse 49
ganglion cells 30
gap
    junctional coupling 93
    junction 88
global
    attention 240, 244
    attractor 26
    brain activity modulation 237
    chaotic attractor 206
glutamate 67, 113
Golgi 3, 6
    type I neurons 9
    type II neurons 9, 10
grandmother cell 213
granule cells 9, 13, 141, 142
gravity method 50
growth phase 118
guinea pig hippocampal slice 117
Hebb rule 49, 67, 123, 198, 241, 242, 254
Hebbian learning 114, 226
hierarchical multilayered network 290
high activity neural systems 229
higher order correlations 50
hippocampal system 12
hippocampus 111, 238
    proper 239
Hodgkin-Huxley
    equations simplification 81
    model 69, 81, 97
        reduction 123, 124
Hopf bifurcation 82, 180, 201
Hopfield model 260
horse-radish-peroxydase 11
hypercomplex cells 286
hyperpolarization 58, 64
    -activated conductances 65
hypothalamus 6
impulse conduction velocity 219
impulses 221, 222
in vitro epilepsy 111
inferior olive 6
informational circuit 238
integrate and fire neuron 45, 276
integrated synaptic currents 275
interspike interval 133, 162
intracellular

calcium 67
recording 114
"secondary" metabolic processes 112
intrinsic bursts 113, 119
ippocampal slice 114
joint peri stimulus time histogram (JPSTH) 54
kinetic
equation 224, 226, 230, 249
model 220
theory of neural fields 218
lamprey spinal cord central pattern generator 107
lateral connections 177
leading edge 30, 65
leakage currents 63
learning 67, 204, 228, 238, 244, 258
algorithm 297
supervised 286, 291 297
unsupervised 286, 291, 297
length constant 44
lesion experiments 24
limbic system 186
limit cycle 180
attractor 201
lipidarchitectonics 10
local
circuit neurons 7
mean field potential 204
repellor 26
localized
engrams 241
modules 259
long distance, deterministic effect 220, 227
long lasting activities 237
long-lived elevated spike rate 259
long-term
memory traces 238, 239
potentiation (LTP) 67, 123
longitudinal
current 59
resistance 59
looming stimulus 27, 29
lower brainstem 6
Lyapunov exponents 173
Mach-band 285
macrocolumn 234
macroscopic cross-section 223
magnetic sensors 196
Markov processes 135
maximum hyperpolarization level (MHL) 219, 248
mean field equation 278
medial forebrain bundle 7
mediators 1
medulla oblongata 5, 6
membrane
conductance 61, 62
current 59
depolarization 276
potential 41, 42, 58, 61, 81, 133, 143
resistance 44, 60
synaptic currents 112
time constant 60
memory 12, 13, 247, 281
mesencephalon 62
metabotropic 113
micro-electrode 41, 43
midbrain 5, 36
miniature end-plate potentials (MEPPs) 168
mitral cells 177, 178
modular architectonic principle 2, 11, 165
modulatory peptides 112
module 232, 255
moments of the firing time 153
monkey 256
attending 50
monoaminergic neurons 6
morphogenetic gradients 164
mossy fibers 6
motoneurons 3
motor
heading map 31
pattern generator 26
schema 24
system 186
moving-object 24
multi-
schema 39
layered network 286

multiple
microelectrode 234
simultaneously recorded
EEGs 196
myeloarchitectonics 10
negative feedback 81
neocognitron 290
Nernst equation 62
neural network 54, 78, 90, 96, 175, 177
model 117
simulations 116
oscillators 87, 93
neurodynamic models 166, 171
neuron
electrotonically compact 68
network 260
omnipause 33
Onchidium pacemaker 173
neuronal
burst activity 147
cables 58
models 139
modules 2
threshold 137
neurons and impulses
interaction 221
neuropeptide
modulators 7
transmitters 7
neurotransmitter 65
concentration 246
activity 245
NMDA 113, 117
conductances 120
receptors 67, 123
node 83
noise 266, 296, 298
non-stationary stochastic processes 51
novel perceptions 238, 240
nullclines 72
numerical
simulations 281
evaluations 148
ocular
dominance 166
motoneurons 32
off-centre receptive fields 284
olfactory
bulb 8, 177, 188, 205
coding 197
cortex 193, 197 198, 205
system 186, 187
on-center receptive fields 284
open states 61
optical recorders 196
orbitally asymptotically stable solution 80, 89, 106
orbitally stable oscillator 83
origin-destination matrix 228, 241
Ornstein-Uhlenbeck
neuronal model 142, 148
process 156
oscillations 81, 82
oscillators 80, 104
out-of-phase solution 89
Papez's circuit 238
paramedian pontine reticular formation 32
pattern recognition 291
patterns 284
percept 188, 210
perceptron 286
perceptual schemas 23
perforant path 12
peri event time histograms (PETH) 51, 53
period-doubling bifurcation 180
periodic boundary 156
persistent activity 254, 256
phase
diagram 180
gradient 202
patterns 201
plane 72
response curve (PRC) 83
shift 83
space 173
-difference models 99
-differences 102
-locked solution 192, 105, 107
phaselocking 84
phaseolus lectin 11
photoreceptors 284
physiological sharp waves 113
point processes 50
pons 5, 6

population
activity 188
oscillation 122
position feedback 81
positron emission tomography (PET) 234
posterior parietal cortex 39
predator 27
presynaptic burst 121
pretectum 24
prey and barrier 26
primary visual cortex 197
priming 258
probability density function (pdf) 133, 221
Purkinje cells 6, 9
putamen 6
pyramidal
cells 13, 111, 173
neuron 60
model 119
quasi-periodic attractor 201
quasi-segmental pattern 5
rabbit 186
Rana computatrix 40
randomly, symmetrically, diluted network 237
random walk 140
raw EEG records 210
re-entry 220
reaction-diffusion model 164
realistic model 250, 273
receptive field 30
recognition 24
processes 250
of Chinese characters 305
of connected characters 302
recruiting 120
recruitment 237, 244
rectangular lattice structure 9
recurrent connections 112, 117
reflection condition 136
refractoriness 227
regularity condition 136
regulatory circuit 238
renewal
density function 45
process 162
replica
method 280
symmetry 281
response latency 186
resting potential 64
retina 24, 31, 284
VLSI chip of 285
retinal ganglion cells 284
retinotopic
code 39
maps 24, 28, 31
retrieval 279
retrieve patterns 271
reverberation 255, 256
reversal potential 62, 65, 89, 97, 159, 161
diffusion models 160
reverse reproduction 289
ring of oscillators 104
robustness 271
rotating waves 234
saccade 32, 33, 35, 36, 187
Schaffer collaterals 13
search controller 304
second messenger 67, 87
segmental arrangement 2
segmentation of connected characters 302
seizure 173
selective
weakening rule 242
attention 297, 304
model 298, 301, 305
self-
organization 288
repair 289
sensitivity analysis 175
septum 12
medial 240
serial dependence 162
serotoninergic neurons 6
shared inputs 48
shifted patterns 290
short-range, random effects 220, 221, 227
simple cells 286
simulation experiments 167
single unit recording 45, 255
sleep cycle 173
slow IPSP 245, 250
slow synapses 95
sniff 187
snow-flakes 50

spatial
coherence 196
ensemble average 190
map 123
spikes 45
doublets 113
spinal
cord 6, 7, 10
gray matter 3
squid giant axons 173
squirrel monkey 173
stable
equilibrium point 175
limit cycle 180
oscillation 234
state transition 201
statistical theories of neural fields 217
steady-state phase 118
stochastic differential equation (SDE) 133
storage capacity 235
store patterns 236
strange attractor 180
striatum 6
strong attraction 94, 95
subicular complex 239
substantia nigra pas reticulata 39
supercritical Hopf bifurcation 101
superior colliculus 34, 36
swimming lamprey 108
symmetric
connection 236
synapses 8
symmetry 232
synapse 66, 90, 187, 220, 237, 245
synaptic
conductance 65
coupling 92, 93, 104
current 59,65
glomeruli 8
matrix 261, 274, 276
potentials 44
triads 8
synchronized afterdischarges 111
synchronous
solution 88
state 89
T-periodic operator 91
tectum 24, 28, 31
temporal correlations 259
thalamus 40
theta rhythm 197
threshold 219
time
delay 81
-homogeneous 137
series 45
-varying threshold 142
topographic map 187, 188
toy models 226
training pattern set 296
transition
front 42
p.d.f. 130
translational
invariance 81
symmetry 291
transmitter 66
peptides 112
travelling wave solution 104
trigger zones 187
two
coupled oscillators 102
-dimensional neural model 231
-layered neural network 284
ultra-stability 234
unanesthetized
cats 45
monkeys 45
uncorrelated patterns 268
upper brainstem 6
visual
image matching 256
pattern recognition 297
voltage-
gated channels 63
sensitive dyes 196
wave packet 188, 197, 210
weak coupling 91, 95
Wigner - May theorem 175
Wiener
neuronal model 148
process 139, 141
working memory 38
Zeitgeber 87