#5

PROCESSING NEUROELECTRIC DATA

by

COMMUNICATIONS BIOPHYSICS GROUP OF
RESEARCH LABORATORY OF ELECTRONICS AND
WILLIAM M. SIEBERT

TECHNICAL REPORT 351

JULY 7, 1959

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS
CAMBRIDGE, MASSACHUSETTS

The Research Laboratory of Electronics is an interdepartmental laboratory of the Department of Electrical Engineering and the Department of Physics.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the U. S. Army (Signal Corps), the U. S. Navy (Office of Naval Research), and the U. S. Air Force (Office, of Scientific Research, Air Research and Development Command), under Signal Corps Contract DA36-039-sc-78108, Department of the Army Task 3-99-06-108 and Project 3-99-00-100.

by

Communications Biophysics Group of
Research Laboratory of Electronics and
William M. Siebert

PROCESSING NEUROELECTRIC DATA

The Technology Press of
The Massachusetts Institute of Technology

TECHNOLOGY PRESS RESEARCH MONOGRAPHS

NONLINEAR PROBLEMS IN RANDOM THEORY
By Norbert Wiener

CIRCUIT THEORY OF LINEAR NOISY NETWORKS
By Hermann A. Haus and Richard B. Adler

MATHEMATICAL PROGRAMMING AND ELECTRICAL NETWORKS
By Jack B. Dennis

PROCESSING NEUROELECTRIC DATA
By Communications Biophysics Group of
Research Laboratory of Electronics and
William M. Siebert

FOREWORD

There has long been a need in science and engineering for systematic publication of research studies larger in scope than a journal article but less ambitious than a finished book. Much valuable work of this kind is now published only in a semiprivate way, perhaps as a laboratory report, and so may not find its proper place in the literature of the field. The present contribution is the fourth of the Technology Press Research Monographs, which we hope will make selected timely and important research studies readily accessible to libraries and to the independent worker.

J. A. Stratton

PREFACE

This Technical Report (which is also an integral part of the new series of Technology Press Research Monographs) was written for several purposes. Primarily, it attempts to bring together in one place the views that this group has developed over the past half dozen years on processing the electrical data that we record from the nervous system. The methods that one employs in processing data are, of course, intimately related to the substantive problems that one is interested in, to the models that one has formulated, and to the capabilities of the instrumentation that is at one's command. Prevailing publication policies of most specialized journals make the discussion of these topics difficult, if not impossible. Since there is little overlap between those who read neurophysiological journals and those who are interested in communications systems, we felt that it might be useful to provide material of common interest to the various branches of the Communication Sciences.

Another motivating influence was the need to clarify our own views by making an over-all assessment of the techniques that we use in everyday research. The presentation that is given here will permit us in future research papers to refer to this monograph instead of having to include lengthy and perhaps not always appropriate discussions of data-processing techniques.

This is hardly the place to present a case history of how an experimentally oriented group evolves, how it works, and the factors that determine its choice of problems and its methods. We have greatly benefited from belonging to the intellectual community that the Research Laboratory of Electronics constitutes within "the Institute." The interest that Professor Norbert Wiener has shown in certain of our experiments has proven stimulating, and our association with Dr. Mary A. B. Brazier's group at the Massachusetts General Hospital has been beneficial. The assistance that the Digital Computers Group of the Lincoln Laboratory, under W. N. Papian, and in particular, Wesley Clark, Jr., have given us has been invaluable.

The present report bears the traces (and indeed the scars) of multiple authorship. A group such as ours cannot claim to be a "team" nor is it something akin to a committee. The data contained in this monograph were collected by present or former members or associates of the group. The chapters were primarily written by those whose names follow the chapter titles. Among those who are not specifically acknowledged as co-authors have

been some of our most faithful critics. They read and reread the successive versions of the different chapters and appendixes with almost as much ego-involvement as the authors. Among them should be singled out George L. Gerstein, who has performed above and beyond the call of his duty as a postdoctoral fellowship holder. Murray Eden, Belmont Farley, George Forsen, and Jan Kuiper helped us by their insistence on clarity. Professor William M. Siebert made an extraordinarily generous contribution to this monograph by writing what is to us an ideal introduction to the difficult topic of random processes. Appendix A reflects his sensitivity to the problems that concern us.

The technical services of the Research Laboratory of Electronics made their usual important contribution to the appearance and readability of this report. The Publications Office, headed by Mrs. Miriam C. Smythe, has had long experience in dealing with our prose in connection with the Quarterly Progress Reports of the Laboratory. It was in these reports that much of the material that is presented here was first discussed, usually in rather succinct form. Mr. Phokion Karas (from RLE's Photographic Service) and the personnel of the Drafting Room, under Mr. Charles P. Navedonsky, were of great assistance in the preparation of figures and graphs. Mrs. Aurice Albert and Mrs. Norma Getty assisted in many ways in the preparation of this monograph; Mrs. Albert, in particular, assembled Appendix D. It is a pleasure to acknowledge here the skill and the responsible manner in which Frank Nardo has collaborated in much of the processing of the data that are reported here. Miss Constance D. Boyd and Mrs. Ann Martin of the Technology Press exhibited unusual patience in the various phases of the manufacture of this monograph.

Finally, this series of acknowledgments would be extraordinarily incomplete were we not to express here our gratitude to one of our authors, who, with his other responsibilities, carried that of being the co-ordinating editor for the monograph: C. Daniel Geisler. Without his persistent, gentle prodding, without his industry and his common sense, we would have never come within striking distance of our deadlines; we might never have finished the job.

Walter A. Rosenblith

June 18, 1959

CONTENTS

Chapter 1

THE QUANTIFICATION OF NEUROELECTRIC ACTIVITY

W. A. Rosenblith

Throughout the history of science experimenters from different fields have dealt with the problem of the quantification of their data in a variety of ways. Technological necessities and the prevailing theoretical structure of a given field determine to a high degree the techniques of measurement that are developed and the choice of variables that are quantified. Experimenters concerned with problems of "organized complexity"[1] often made little effort to report their observations in quantitative or even systematic form. They were too aware of the limited range of experimental facts that they could ascertain with a sufficient degree of invariance and of the narrow realm in which they could actually verify predictions from mathematical models.

These difficulties and an overly narrow interpretation of Lord Kelvin's doctrine* may be largely responsible for the fact that neurophysiologists, for instance, have often been hesitant to go beyond reporting raw data in a somewhat phenomenological manner. Such an attitude renders communication with fellow scientists hazardous. If verbal statements alone are made to carry the informational burden of large bodies of data, friendly model-makers from the physical sciences are tempted to construct theories of "how the brain works" on the basis of a few isolated and easily mathematized facts.

But it was just not caprice or lack of farsightedness among the data-rich and theory-poor scientists that produced this mismatch between their vast labors and the relatively small amount of theoretically integrable knowledge that became available. They were handicapped by a lack of adequate data-processing facilities and by the fact that the mathematical models of classical physics (and certainly those of quantum physics) had little to offer to the student of the nervous system or of human behavior. Hence, many

* "I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of Science, whatever the matter may be." Contrast this view of Lord Kelvin's with Gödel's contention[2] according to which "it is purely an historical accident that it [mathematics] developed along quantitative lines."

among them were interested by cybernetics, which emerged as the philosophical expression of the communications technology of the postwar period. It was under cybernetics' influence that many problems relating to the behavior of complex living systems were reconsidered. Only too often these reconsiderations turned out to be not only suggestive but also frustrating. At that stage a search for general principles of the behavior of the nervous system could not help but be somewhat superficial. The neuroanatomical, the neurophysiological, and the behavioral data extant were not in a form that made theorizing at a fairly general level meaningful.

1.1 Problems of Measurement and Analysis in Electrophysiology

For more than two centuries - thanks to various species of electric fish - men have been aware of the existence of "animal electricity."[3] More than a century ago Helmholtz measured the conduction velocity of nerve, and throughout the second half of the nineteenth century an appreciable amount of knowledge concerning brain potentials accumulated. A recent review article on the "Rise of Neurophysiology in the 19th Century"[4] summarized the situation at the end of that century as follows: "It was known that the brain had "spontaneous" electric activity, that potential shifts could be elicited in the appropriate cortical areas by sensory stimulation, that these potentials could be recorded from the skull and that anesthesia abolished them." However, electrophysiology entered its period of rapid growth only after the technology of the vacuum tube gave us amplifiers and oscilloscopes. These two instruments permitted electrophysiologists to increase the sensitivity of their observations and to display even rapid fluctuations in voltages as a function of time.

The characteristic deflections or patterns in voltage-versus-time-displays constitute the electrophysiologist's basic data. But how are these characteristics of a waveform to be assessed? As long as scientists deal with DC potentials or sinusoids, an instrument that yields one or two characteristic numbers is perfectly satisfactory, but when they attempt to assess arbitrary waveforms containing sharp "transients" and "noise," several questions arise. Is the voltmeter (even the vacuum-tube voltmeter) the appropriate instrument of measurement? Is it necessary to display the complete waveform by photographing it from the face of an oscilloscope? Can we find selective transformations upon the data that yield meaningful descriptions while reducing the total amount of information displayed?

A further discussion of appropriate methods for the quantification of electrophysiological data leads us to consider issues that the physical sciences have faced - sometimes quite explicitly and sometimes less so - throughout their history. Before we make measurements reflecting the behavior of complex systems, it may be wise to ask ourselves two sets of questions. Why do we make a particular measurement? What conclusions (regarding the phenomena under investigation) shall we be able to draw on the basis of the measurement?

The first set of questions inquires into the purposes of the experimenting electrophysiologist: Is he interested in relating the electrical events that he records from an isolated nerve fiber to the physico-chemical processes that occur in the transmission of a nerve impulse? Is he using the electrical events in order to trace certain pathways in the nervous system? Is he trying to study the responses of certain neural structures to carefully controlled sensory stimuli? Is he investigating the behavior of neural structures in relation to a mathematical model that he has formulated? Is he studying the way in which certain chemical substances affect synaptic transmission? Is he trying to relate changes in an organism's electrical activity to conditioning or learning? Or is he concerned with the presence or absence of certain patterns in this activity, with a view towards clinical diagnosis? Neurophysiology includes all of these experiments. The experimenter's purpose determines the choice of his variables, the display technique for his data, and affects the very definition of what constitutes an experiment: Which parameters are to be held constant, how replicable must a phenomenon be, ... ? Neurophysiology - which has, compared to the physical sciences, little theoretical structure of its own - is thus characterized by an aggregate of techniques for the study of the nervous system or of its component parts. As a science it stands in close relation to fields such as neuroanatomy, sensory physiology, biochemistry, psychology, biophysics, and medicine, and the significance of neurophysiological findings is often assessed in terms of their relevance to the neighboring fields.

The second set of questions deals with the inferences that can be drawn from electrophysiological "pointer readings." It is here that our lack of understanding of the organizational principles and of the mechanisms of the nervous system is felt most seriously. The organizational structure of this nonhomogeneous medium that consists of large numbers of highly specific elements has so far defied useful description in terms of the over-all physical properties of the medium. Much effort has gone into analyzing the fine structure of its various components in terms of current biophysical and biochemical knowledge, but up to the present these efforts have not yielded an approach that is capable of dealing with the unique properties that characterize the nervous system of higher animals. Here is a system that is composed of many interacting units (all of which are by no means alike), that is organized both flexibly and hierarchically, that consists of subsystems (enjoying various degrees of autonomy) that are capable of fulfilling specific and/or nonspecific functions. Here is a system that reacts more reliably and predictably to informationally rich stimuli than to "simple" ones. Here is a system that is capable of learning and of giving reasonably reliable performance throughout an extended period of time, with all the safety factors and maintenance and repair requirements that such performance demands.

If we want to understand the "systems neurophysiology" that underlies the behavior of information-processing organisms, what is the type of electrical activity that we should study? What type of strategy should we adopt in dealing with the signals that we record from the nervous system - signals whose code is known so incompletely? Should we attempt to isolate a single neuron and and study its behavior in great detail, hoping that we will pick the "right" (representative) one out of a not-too-well defined population? Should we at the other extreme, work only with the muffled polyneural roar that is able to make itself "heard" through man's thick skull? Should we limit ourselves to studying recordings of "spontaneous" activity of a neuron (or of neuronal populations), that is, the activity that we can still observe when we have turned off all the stimulus generators that are under our control? Or should we study stimulus-response relations, that is, those response events whose occurrence is by some criterion (usually a temporal one) linked to the delivery of a definable stimulus? Can we assume that these latter stimulus-evoked events will always simply add to the "spontaneous background activity," or must we study their interaction in different physiological states of the organism?

Are the biggest voltages, especially when recorded at the outside of the skull, the most important ones to study? If we compare this situation with the facts of speech communication, we find that it is the vowels (yea, their first formants) that carry most of the energy among the speech sounds, although - in English at least - it is the consonants (whose clamor for attention is much less loud) that carry most of the linguistic information. There are perhaps other lessons to be drawn from the study of speech communication. When a Fourier analysis of speech signals is carried out, the vowels (whose duration is of the order of 1/10 second) seem to be represented much more meaningfully by Fourier components than the consonants. The latter can be viewed as "transients" or "transitionals," whose spectral composition depends much more upon the vowels that precede or follow them. The problem of where the vowels end and the consonants start (technically known as the segmentation problem) presents a challenge all of its own, comparable perhaps to that of defining the duration of an evoked response. An "ah" will exhibit rather different spectral components when pronounced by a man, a woman, or a child; it will even exhibit appreciable differences when pronounced repeatedly, and in different context, by the same individual. And yet there is something invariant about it that makes it recognizable as an "ah." This "ah"-ness is not anything that is easily characterizable by absolute numbers, but rather by distinctive features or parametrically defined patterns, by certain relations among the components of a sound, especially in relation to other sounds that might have been emitted. Lest this analogy be carried too far, let us not pretend that we are waiting for somebody to break "the" code of the nervous system. Let us realize that we are trying to discover the units of analysis, the distinctive features of neural signals, that will help us order the innumerable data of the nervous system.

What are the techniques of analysis that are readily available to electrophysiologists when they record data to deal with the range of experimental problems that we have mentioned above? Let us briefly mention some sample techniques that have been used. The mathematics of circuit analysis (at least in its simpler forms) assumes that the circuits and their components are linear, lumped, finite, passive, and bilateral.[5] It would, of course, be absurd to pretend that the nervous system has these properties, though it may be possible to find, by applying circuit theory, in what manner the behavior of a sensory system, for instance, deviates from this model.

If we restrict ourselves to dealing with whatever waveforms may have been recorded, we must ask whether the specific techniques such as Fourier analysis or correlation analysis are actually appropriate to the particular experimental question. Such techniques imply that the time series analyzed satisfy certain conditions.

Obviously, the assumptions implicit in these analytical techniques are a price that we have to pay for their use. Physical scientists also pay this price. They, however, know so much more about the processes that underlie the phenomena they study than we know about the mechanisms that underlie neuroelectric phenomena. Thus, in physical science there is a better chance of adjusting and correcting models than there is in neurophysiology. And yet the student of the nervous system has little choice until more appropriate techniques of analysis have been developed. He must utilize those that are available in order to find out where they cease to fit. It may, nevertheless, be wise to take the precaution of assembling a sufficient body of apparently consistent data before getting involved in ambitious computations.

Is there a moral that imposes itself on the basis of the preceding tedious and yet incomplete enumerations of problems that one faces in this type of research? We believe that there is, and we believe that it can be stated in a single word: pluralism. Only a pluralistic strategy guarantees, at this stage of our knowledge of the nervous system, that we shall not blind ourselves to useful approaches because we have oversold ourselves on one of them. The very multiplicity of purposes precludes our prescribing experimental design or methods of data processing and analysis too rigidly on intrinsic grounds. We must, rather, be prepared to make our choice on the basis of extrinsic values or influences: Given the biases of interest that we - as a group - have, given the physical and intellectual surroundings in which we work, we have developed certain methods of data processing and certain types of mathematical models. We believe that these techniques are capable of coming to grips with the statistical character of neural activity which is one of the essential features of the nervous system. We have, furthermore, a preference for packaging our results in a form that is reasonably quantitative; that is, we try to express as many of our findings as we can in some mathematical representation without always trying to fit our data to analytical functions. Since we are dealing with a multivariate system, we are not surprised that the patterns and relationships that we find are often statistical. Finally, it is fair to say that, while we feel more secure when we have the guiding influence of a mathematico-physiological model in our experiments, we are not so narrow-minded as to ignore the usefulness and even the beauty of a good classification scheme that relates to variables whose importance to the organism is undeniable.

1.2. A Statistical View of Neuroelectric Phenomena

No matter which aspect of the electrical activity of the nervous system we study, we always face the task of defining "typical events" among those we observe experimentally. This task confronts the experimenter, whether his concern is with evoked responses or with the EEG (electroencephalograph). He has to establish certain criteria of judgment. These criteria will be different when he records with the aid of gross electrodes than when he studies the activity of a single cell with the aid of a microelectrode. The electrophysiologist has the further problem of deciding whether two observations are "identical." Here the identity-defining operation may range from identity in one aspect of the event only (such as occurrence or nonoccurrence of a spike potential) to identity in all measurable aspects)average spike latency, distribution of spike latencies, and so on).

In order to decide whether an event is typical or whether two events differ, we really have to know something about the distribution of possible events. This distribution might be obtained by observing responses to a large number of identical stimuli or by repeatedly sampling an EEG (electroencephalographic) trace. Actually, experimenters rarely have such information available to them, and yet, if they are well trained, they choose representative records as illustrations for their papers. It is, nevertheless, necessary to realize that few, if any, systematic studies have been made to assess an experimenter's information-handling capacity as applied to his ability to view oscilloscopic traces or examine film records. In other words, we do not really know how safe the current procedures are.

We have tried to present and review elsewhere[6, 7, 8] some of the available evidence on the statistical character of input-output relations in either single units or for responses from populations of neuronal elements. Here we shall try to summarize the essential arguments only.

We faced this problem first when we tried to find criteria for deciding what constitutes a typical evoked response (a response that is evoked by the presentation of a discrete stimulus, most often a sensory one). There exists, to our knowledge, no generally accepted operational definition of what is meant by an evoked response although the concept has been exceedingly useful in electrophysiological and neuroanatomical studies of the nervous system.

Let us briefly see how evoked responses are recorded. The experimenter usually knows when a stimulus is being presented. He then most often establishes the presence or absence of an evoked response by either of two methods or by the two methods conjointly: (1) In recording with gross electrodes, he detects visually the presence of a characteristic waveform or deflection. (2) In recording with microelectrodes, he detects aurally (and/or visually) a change in the acoustic signals that represent the electrical events "seen" by the microelectrode after these events have been appropriately amplified and transduced.

As should be clear from this description, the experimenter's ability to detect such changes in visual and/or aural displays depends upon how stable these changes are in relation to the patterns of "background activity."* These changes will be most easily detected when they have short latencies (that is, when they occur right after the presentation of the stimuli). The more these changes exceed the experimenter's just-noticeable-difference for the visual or aural displays involved, the more reliable their detection will be.

For responses that are recorded with gross electrodes, there is variability both with respect to amplitude and with respect to time. The evoked responses of the classical afferent pathways exhibit relatively short latencies and little variability in latency. It is this relative stability of the temporal aspects of these responses that makes the use of averaging by computing devices (such as the ERD and the ARC-1) possible and useful. It goes without saying that latencies determined from the average evoked response permit us to say little about the latencies of the individual responses. So far no adequate techniques have been developed to deal with electrical events that have longer and more variable latencies (such as the so-called "blocking of the alpha rhythm").

*We have already mentioned the problems of the typicality of a response and of the identity of two responses. These problems include in some sense decisions of how typical the background activity is in which these responses are imbedded. Amassian and his co-workers emphasized only recently[9] how the presence of spontaneous cell discharges complicates the analysis of the effect of stimulus variables.

For responses that are recorded from single units with the aid of microelectrodes, the variability problem is rather different: Here we are dealing with a set of discrete events that are quite comparable in waveshape and amplitude but that occur at latencies that are governed by both stimulus parameters and the existing sequences of "spontaneous" firings of the cell. The changes in the patterns of "spontaneous" firing that do occur may result in either increases ("excitation") or decreases ("inhibition") in average firing frequency; thus variability may now affect (a) changes in number of firings (how many spikes does a given stimulus elicit or inhibit), (b) "first"-spike latency (latency of the spike whose occurrence is most directly linked to the delivery of the stimulus), (c) interspike intervals, and so on.

An overview of the problem of adequate detection and description of evoked responses leads thus to procedures in which computers are instructed to "look" for changes in patterns of ongoing activity that are somehow linked to the delivery of stimuli. "Looking" for changes in averages, such as means, or for changes in distributions within several time intervals becomes thus a method of search in which the properly instructed computer supplements human capacities.

From all that precedes, it should be clear that we must find ways of dealing with the undeniable fact that repeated presentations of the same stimulus do not yield "identical" neuroelectric responses in many physiological preparations. Instead of abdicating before this fact by declaring that neuroelectric activity is thus not tryly quantifiable, one can take advantage of this difficulty.

The variabilities that one observes seem to have their own regularities, which are in turn related to both stimulus and organismic variables. By constructing a model that had relevant statements to make with respect to both mean and variance of population responses, Frishkopf[10] was able to give a much deeper interpretation of neural events at the periphery of the auditory system than had been possible previously.

If we look for an interpretation of this statistical behavior, we must first of all consider the complexity of the system or subsystem under study, the multiplicity of possible interactions,* and the lack of adequate description of state in which a cell or a neuronal population finds itself at the time when a stimulus is presented.

A recent article of Bullock[12] gives a thoughtful discussion of the present status of the neuron doctrine and suggests several

*Sholl,[11] who has discussed the quantification of neuronal connectivity, states, for instance, that "impulses arriving along a single primary visual fiber will be dispersed among the 5000 neurons distributed around its terminal branches."

major revisions. Many of the ideas expressed by Bullock force a reconsideration of what is meant by the state of a neuron and emphasize the necessity for looking beyond the occurrence of the spike potential as the sole indicator of neuronal function.

Although there will undoubtedly become available more adequate descriptions of the state of single neurons or of neuronal populations, there is serious doubt whether we shall, in the foreseeable future, be able to dispense with statistical descriptions of neuroelectric phenomena. Given this prognosis, we shall endeavor to develop and use the most appropriate methods available in order to elucidate the statistical aspects of neuroelectric activity.

REFERENCES

1. W. Weaver, "Science and Complexity," Am. Scientist, 36, 536-544 (1948).
2. R. Oppenheimer, "Analogy in Science," Am. Psychologist, 11, 127-135 (1956).
3. A. v. Muralt, Neue Ergebnisse der Nervenphysiologie, Springer-Verlag, Berlin, Chapter I, 1958.
4. M. A. B. Brazier, "Rise of Neurophysiology in the 19th Century," J. Neurophysiol., 20, 212-226 (1957).
5. E. A. Guillemin, The Mathematics of Circuit Analysis, John Wiley & Sons, Inc., New York, 1949.
6. W. A. Rosenblith, "Some Electrical Responses from the Auditory Nervous System," Proceedings of the Symposium on Information Networks, Polytechnical Institute of Brooklyn, pp. 223-247, April 12-14, 1954.
7. L. S. Frishkopf and W. A. Rosenblith, "Fluctuations in Neural Thresholds," Symposium on Information Theory in Biology, H. P. Yockey, R. L. Platzman, and H. Quastler, Eds., Pergamon Press, New York, pp. 153-168, 1958.
8. W. A. Rosenblith, "Some Quantifiable Aspects of the Electrical Activity of the Nervous System (with Emphasis upon Responses to Sensory Stimuli)," Revs. Mod. Phys. 31, 532-545 (1959).
9. V. E. Amassian, L. Berlin, J. Macy, Jr., and H. J. Waller, "II. Simultaneous Recording of the Activities of Several Individual Cortical Neurons," Transactions of the New York Academy of Sciences, Ser. II, 21, 395-405 (1959).
10. L. S. Frishkopf, A Probability Approach to Certain Neuroelectric Phenomena, Technical Report 307, Research Laboratory of Electronics, M. I. T., March 1, 1956.

11. D. A. Sholl, The Organization of the Cerebral Cortex, John Wiley & Sons, Inc., New York, 1956.

12. T. H. Bullock, "Neuron Doctrine and Electrophysiology," Science, 129, 997-1002 (1959).

GENERAL BIBLIOGRAPHY

The following list of books and articles is intended to supplement the references given at the end of the various chapters. This list should provide enough variety to afford the interested reader an introduction to current research on the nervous system. A few general references have also been included.

1. E. D. Adrian, F. Bremer, and H. H. Jasper, A CIOMS Symposium: Brain Mechanisms and Consciousness, Charles C. Thomas, Springfield, 1954.

2. G. H. Bishop, "Natural History of the Nerve Impulse," Physiol. Revs., 36, 376-399 (1956).

3. M. A. B. Brazier, Ed., The Central Nervous System and Behavior, The Josiah Macy, Jr. Foundation, New York, 1959.

4. M. A. B. Brazier, The Electrical Activity of the Nervous System, MacMillan, New York, 1953.

5. F. Bremer, "Cerebral and Cerebellar Potentials," Physiol. Revs., 38, 357-388 (1958).

6. B. Delisle Burns, The Mammalian Cerebral Cortex, Williams & Wilkins Co., Baltimore, 1958.

7. P. E. K. Donaldson, Electronic Apparatus for Biological Research, Academic Press Inc., New York, 1958.

8. J. C. Eccles, The Physiology of Nerve Cells, The Johns Hopkins Press, Baltimore, 1957.

9. J. F. Fulton, Ed., A Textbook of Physiology, 17th ed., W. B. Saunders Co., Philadelphia, 1955.

10. R. Granit, Receptors and Sensory Perception, Yale University Press, New Haven, 1956.

11. H. F. Harlow and C. N. Woolsey, Eds., Biological and Biochemical Bases of Behavior, The University of Wisconsin Press, Madison, 1958.

12. B. Katz, "Nature of the Nerve Impulse" in Biophysical Science, a Study Program, Revs. Mod. Phys., 31, 466-474 (1959), and John Wiley & Sons, New York, 1959.

13. B. Katz, "Mechanisms of Synaptic Transmission" in Biophysical Science, a Study Program, Revs. Mod. Phys., 31, 524-531 (1959), and John Wiley & Sons, New York, 1959.

14. H. W. Magoun, The Waking Brain, Charles C. Thomas, Springfield, 1958.

15. D. Nachmansohn and H. H. Merrit, Eds., Nerve Impulse, Vols. 1-5, Transactions of the Conferences Sponsored by the Josiah Macy, Jr., Foundation, Corlies, Macy & Company, Inc., New York, 1951-1956.

16. J. von Neumann, The Computer and the Brain, Yale University Press, New Haven, 1958.

17. J. W. Papez, Comparative Neurology, Thomas Y. Crowell Company, New York, 1929.

18. C. S. Sherrington, The Integrative Action of the Nervous System, Cambridge University Press, 1947.

19. D. A. Sholl, The Organization of the Cerebral Cortex, Methuen, London, 1956.

20. S. S. Stevens, Measurement and Man," Science, 127, 383-389 (1958).

21. S. S. Stevens, Ed., Handbook of Experimental Psychology, John Wiley & Sons, New York, 1951.

22. A. E. Walker, C. E. Henry, and J. K. Merlis, Eds., A Symposium: Recent Developments in Electroencephalographic Techniques; Third International Congress of Electroencephalography and Clinical Neurophysiology, Electroencephalog. and Clin. Neurophysiol., Suppl. 4, 1953.

23. G. E. W. Wolstenholme and C. M. O'Connor, Ciba Foundation Symposium on the Neurological Basis of Behavior, Little, Brown and Company, Boston, 1958.

24. J. Z. Young, The Life of Mammals, Oxford University Press, Oxford, 1957.

Chapter 2

EVOKED RESPONSES

M. H. Goldstein, Jr., and W. T. Peake

with

C. D. Geisler, N. Y-S. Kiang, T. T. Sandel, and J. S. Barlow

2.1. Introduction

The importance of studies of evoked responses in electrophysiological research has been discussed in Chapter 1. Responses evoked by controlled stimuli can be recorded by microelectrodes from one or a few nerve cells, or by gross electrodes from many cells. It was pointed out in Chapter 1 that in either case the responses have a variable character. This chapter presents certain approaches to quantitative descriptions of evoked responses that are recorded by gross electrodes.

Examples of evoked responses are shown in Figs. 2.1 and 2.2. The responses in Fig. 2.1 were recorded from the auditory cortex of an anesthetized cat. Although the stimuli were identical, the individual responses show appreciable fluctuation. Fig. 2.2 illustrates activity recorded at two different times from the scalp of a sleeping human subject. The five traces at the top of each column show the activity recorded following the presentation of five consecutive click stimuli. In the set of individual responses on the right, a characteristic deflection is observable at approximately the same time in each trace. In the set of traces in the left column, taken 6 minutes earlier, characteristic deflections are difficult, if not impossible, to detect.

The variable nature of evoked responses which is illustrated in these figures makes it difficult to obtain representative measurements from any one response; sometimes it is even impossible to determine by visual examination of the recordings whether the presentation of stimuli has in any way changed the pattern of the recorded potentials. Some of the methods which have been used to cope with this problem change physiological state so much that it is difficult to determine whether the results obtained under these somewhat special conditions apply to more nearly normal situations. These special methods include the use of deep anesthesia to reduce

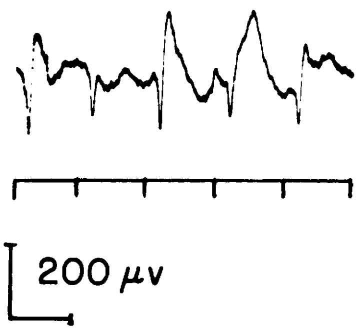

Fig. 2.1. Cortical responses to periodic auditory clicks recorded from an anesthetized cat (Dial anesthesia). The upper trace shows the cortical responses; the lower trace indicates the times at which stimuli were presented. The clicks were presented monaurally at a rate of 10 per second. The intensity level was 35 db above VDL (Visual Detection Level, the intensity level at which a response is just visually detectable in the recorded activity). The recording was from a monopolar electrode in the middle ectosylvian gyrus of the contralateral auditory cortex (indifferent electrode on exposed neck muscle).

In the experiments reported in this chapter, acoustic stimuli were produced by an earphone (PDR-10), and, unless otherwise stated, presented monaurally. Clicks were produced by the application of electrical pulses, 0.1 msec in duration, to the earphone. In all figures, unless otherwise noted, upward deflection indicates negativity of the recording electrode with respect to the reference electrode.

"ongoing" activity[1,2*], and the use of locally applied strychnine to "potentiate" responses.[3]

Another approach to obtaining stable measurements of responses involves the use of special techniques in the processing of recorded activity. These techniques include: (1) selecting particular portions of the response activity for display; (2) selective filtering so as to enhance the characteristic deflections expected in the evoked responses[4]; (3) using electrode geometry in order to enhance certain features of the responses[5]; (4) superposing of response traces with the sweep synchronized to the stimulus delivery[6]; and (5) averaging of individual responses to identical stimuli.[7]

*These references will be found listed at the end of this chapter. No attempt has been made to be complete or to list in each case the earliest reference in the literature. We have rather attempted to give accessible, illustrative examples.

Averaging of evoked responses has been facilitated in recent years by the use of electronic devices.[8] Two advantages of averaging are that (1) characteristic deflections may be detected in the waveform of the average of responses where such deflections are not visually detectable in any single response trace and (2) the average yields more stable measurements than those obtained from any individual response. Both aspects of averaging are illustrated in Fig. 2.2. The bottom trace in each column is the average of 100 responses. In the column on the left, the averaging process brings out a characteristic deflection that is not visually detectable in the individual responses. In the column on the right, although characteristic deflections are visible in the individual traces, there is considerable fluctuation that is reduced by averaging. An illustration of the stability of the average response evoked by optical stimuli is given in Fig. 2.3. Seventy identical stimuli were

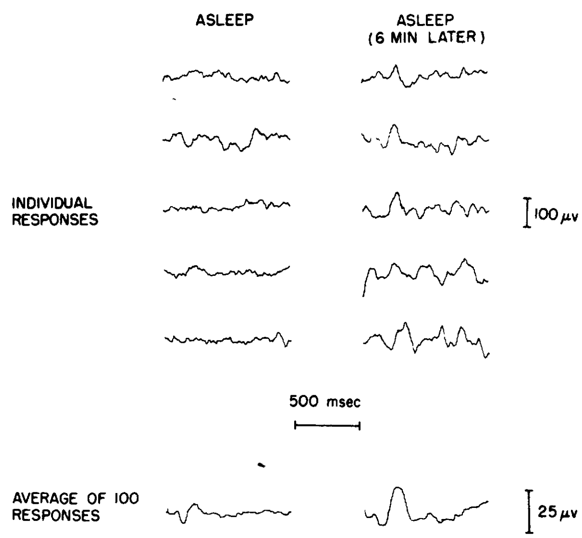

Fig. 2.2. Responses to periodic clicks recorded from a sleeping human subject. The five traces in the upper left show consecutive responses taken approximately 40 minutes after the subject was instructed to sleep. The five traces in the upper right are consecutive responses taken 6 minutes later. Each of the lower traces is the average of 100 consecutive responses (computed by ARC-1), and includes the five shown above it. Clicks were presented binaurally at a rate of 0.75 per second. Upward deflection represents positivity of an electrode at the subject's vertex with respect to an occipital electrode (Subject H-432).

In all figures in this chapter, unless otherwise noted, stimuli were presented at the times corresponding to the beginning of each trace (that is, the left-hand edge).

presented consecutively. The first ten and final ten responses of the series were averaged separately. These averages and the average of the seventy responses are shown together with ten individual sample responses. Note that the averages for even ten responses show less variability than the individual traces.

Averaging, of course, is not the only way of processing evoked activity to obtain a stable and quantifiable description of responses. For example, when the responses have a clearly identifiable waveform, as in Fig. 2.1, one can measure particular characteristics of each response. Examples of such characteristics are the amplitudes of peaks in the response waveform and the time intervals from each stimulus onset to these peaks. When a number of responses is processed in this way, averages, histograms, and other statistics can be computed and displayed.

* Thus far we have taken an operational view of certain methods of processing electrical activity that has been recorded by gross electrodes. This type of data reduction can be viewed more mathematically by attempting to find a mathematical model which closely "fits" the experimental data. Models which seem to be the best candidates for describing certain aspects of the electrical activity recorded from the nervous system are the mathematical models of random processes. An introductory discussion of random processes is given in Appendix A. The methods of mathematical statistics which are used to estimate the parameters of the assumed models and to check the "fit" of the models are discussed in Appendix B.

If stimuli are presented periodically, the appropriate model may be a periodically time-varying random process. The averaging technique would then be a method of estimating the mean of this process. For a time-varying random process the mean is generally a function of time, and in the periodic case the mean is a periodic function of time.

The usefulness of a mathematical model is measured by its ability to predict results of processing the data in ways different from those used to estimate the parameters of the model. If a simple model "fits" the data well, it provides an efficient way of characterizing the data and may suggest new experiments.

In certain of the cases presented below, the computations have been carried out to test specific mathematical models which postulate a certain kind of behavior for "neural units." In other cases, the mathematical theory of random processes has been used as a descriptive model without reference to any particular neurophysiological mechanism or structure.

An important facet of the description of electrophysiological activity in terms of random-process models - the choice of appropriate models - lies outside the realm of mathematics. Within the theory lies the means for estimating parameters of models, and testing the probability that the models could have generated the recorded activity. But the choice of models is in the hands of the research worker.

---

*Passages in fine print in chapters 2 and 3 are written from the viewpoint of random process models. A knowledge of the material contained in Appendixes A and B is assumed.

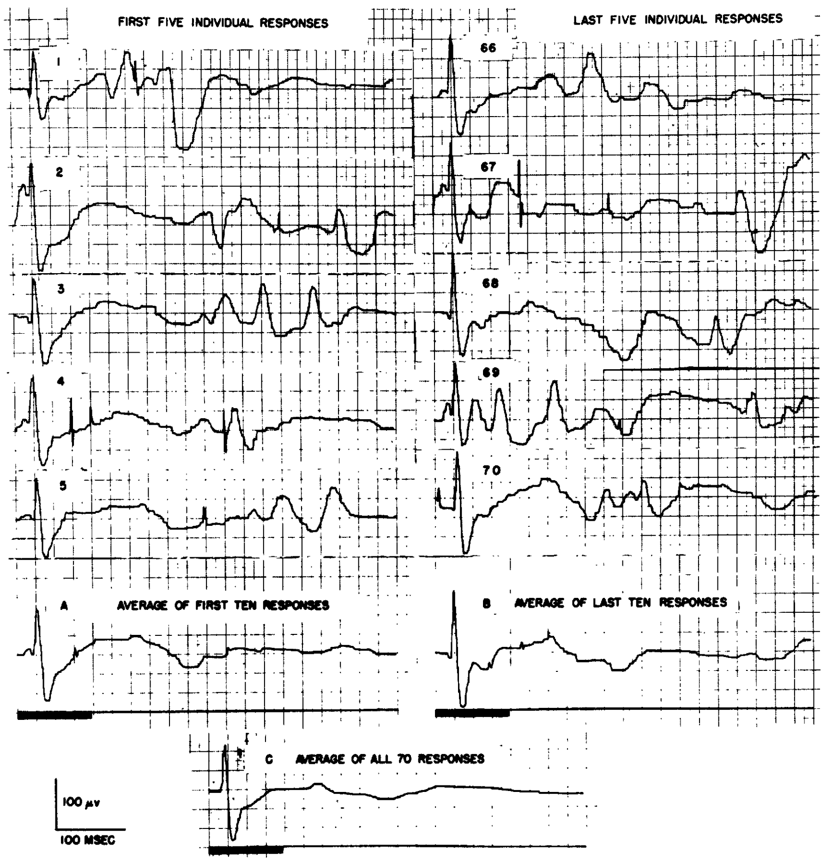

Fig. 2.3. Cortical responses to periodic light flashes recorded from an anesthetized cat (Dial anesthesia). 70 monocular light pulses about 100 msec long and 7 log units above VDL were presented at a rate of 0.2 per second. The lower traces in A, B, and C show the duration of the light flash. The beam supplied by the photic stimulator was focused at the cornea of the cat where it subtended a visual angle of 29°. Averages computed by ARC-1. Courtesy of E. Berger and O. Gutierrez.

2.2. Processing of Evoked Responses

In this section, several examples are discussed in which neuroelectric evoked response data have been processed in various ways. Emphasis is placed on the techniques used in treating the data rather than upon interpretation of results. (Readers are referred to the References for formal reports of the research mentioned.)

All of the operations to be described could be carried out, in principle, by manual computation. However, electronic computers, because of their great speed and reliability, have been used for most of this work. Brief descriptions of some of the computers that have been used are included here; more details are given in Appendix C.

Measurement of Response Characteristics

One method of dealing quantitatively with well-defined neural responses such as those of Figs. 2. 1 and 2. 4 is to measure the amplitude and latency of certain deflections. An electronic computer called ALMIDO (Amplitude and Latency Measuring Instrument with Digital Output) was constructed to make measurements of this kind, and its function is described here. (A description of the device, including specifications, is given in Appendix C.)

A particular time segment of each incoming response is selected by ALMIDO. The delay between the stimulus onset and the start of this "window" and also the length of time for which the window is open can be adjusted by the experimenter. An amplitude measurement is made from the largest positive peak to the most negative peak in the selected time segment. This measurement is made digitally as a number from 0 to 99, and it is printed on paper tape. The results of a series of such measurements can be presented as an amplitude histogram on a set of counters. Figure 2. 5 shows amplitude histograms which were computed by ALMIDO.

If the amplitudes of the responses are assumed to be samples of a random process, the normalized histogram is an unbiased estimate of a step approximation to the probability density function of the random process (see Appendix B). It is shown in Appendix B that the estimate will probably be close to a steplike approximation of the density function if the number of samples is large enough. For physiological data it is desirable to use the smallest number of samples consistent with the desired accuracy of the estimate, since it is difficult to maintain preparations in a constant state for the time required to obtain a large number of samples. The number of samples necessary to obtain a given accuracy can be determined experimentally by repeating the computation of the histogram several times with the same number of samples and observing the variability in the estimate. If the variability in the histogram is small, it is very likely that the estimate is accurate.

Each time that the amplitude measurement is made, ALMIDO also measures the time interval from the stimulus onset to the largest peak in the selected time segment. Scatter diagrams of peak latency versus peak-to-peak amplitude for auditory nerve responses to clicks and short bursts of noise are shown in Fig. 2. 6.

A test of association can be made on the data in these diagrams. A corner test [10] shows that the hypothesis of independence cannot be rejected at the 5 per cent level for the click responses in Fig. 2. 6a. The same hypothesis can be rejected at the 1 per cent level in Fig. 2. 6b. Since short noise bursts are similar to clicks having random amplitudes, the negative correlation of latency and amplitude might be explained in terms of the demonstrated relationship [11] between the latency and amplitude of responses to clicks. (The latency decreases and the amplitude increases with increasing intensity.)

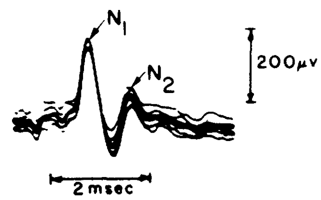

Fig. 2.4. Superposition of ten auditory nerve responses to periodic clicks recorded from an anesthetized cat (Nembutal anesthesia). Click rate was 1 per second; intensity was 35 above VDL. The recording electrode was near the round window of the cochlea, and the reference electrode was attached to the headholder, (Cat 376). After Goldstein and Kiang [9].

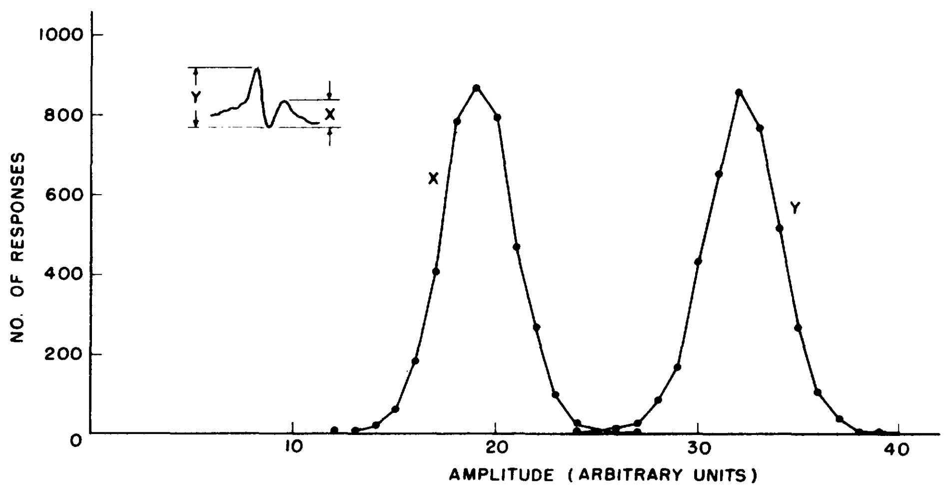

Fig. 2.5. Histograms of peak-to-peak amplitudes of auditory nerve responses. An anesthetized cat (Dial anesthesia) was stimulated monaurally by periodic clicks at a rate of 10 per second. Stimulus intensity was 30 db above VDL. The recording electrode was near the round window of the cochlea, and the reference electrode was attached to the headholder. The two amplitudes X and Y are defined in the inset diagram (Cat 494).

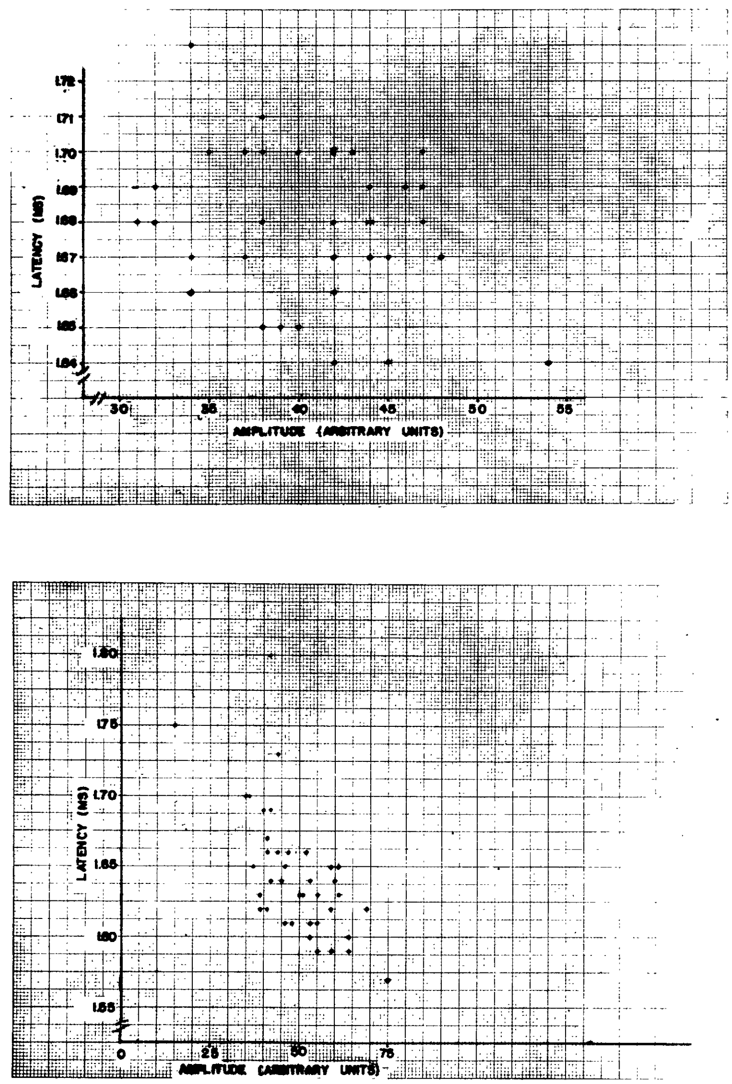

Fig. 2.6. Scatter diagrams of peak latency of N_1 versus peak-to-peak amplitude (see inset of Fig. 2.5). (a) Responses to 0.1-ms square pulses applied to earphone at a rate of 1 per second. Stimulus intensity is 15 db above VDL. (b) Responses to 0.1-msec noise bursts applied to earphone at the same rate. Intensity is 20 db above VDL. (Cat 491.) Experimental conditions are the same as described in Fig. 2.5. Latency is measured from the time of the application of the electrical pulse to the earphone to the time of the peak of N_1.

Figure 2.7 is another example of computations performed on the amplitude of auditory nerve response. The average of the peak-to-peak amplitude of the auditory nerve responses to clicks (dimension Y in Fig. 2.5) and the standard deviation of this amplitude are plotted as a function of the intensity of the clicks. These data have been interpreted in terms of a probability model for the behavior of "neural units." The theoretical curve of the standard deviation, with confidence limits (based on this model), is presented with the standard deviations of the experimental data. For a more detailed description of this work see Ref. 12.

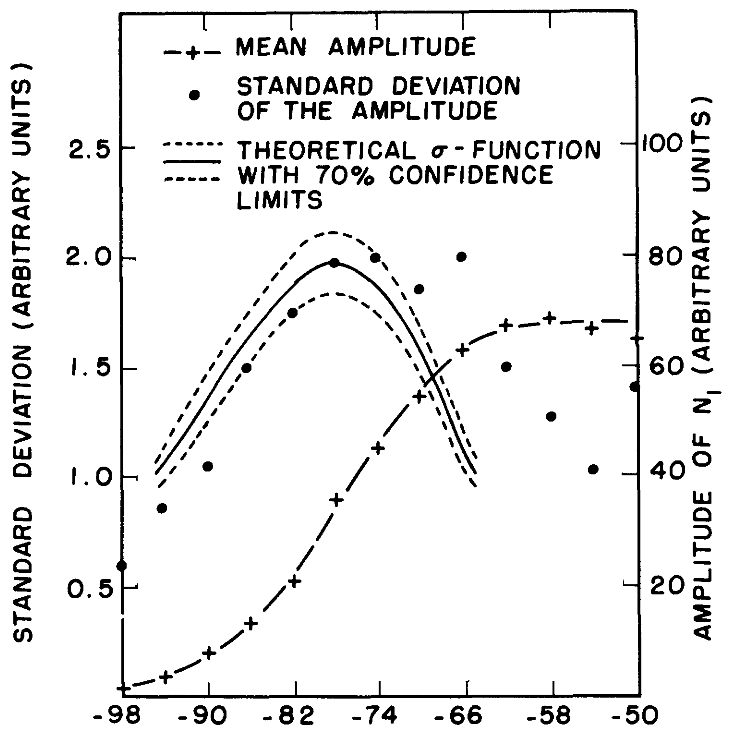

Fig. 2.7. The mean and standard deviation of the peak-to-peak amplitude of the auditory nerve response, as a function of click intensity, recorded from an anesthetized cat (Urethane anesthesia). The dimension Y in the insert of Fig. 2.5 shows the amplitude that was measured. Each data point was computed from the responses to 100 periodic acoustic clicks presented at 1 per second. The recording electrode was near the round window, and the reference electrode was placed on exposed neck muscle (Cat 349). After Frishkopf and Rosenblith.[12] (Reprinted by courtesy of Pergamon Press.)

Averaging of Evoked Responses

In the following sections, results of several experiments are presented in which evoked responses have been averaged. The examples were chosen to illustrate some of the kinds of problems to which this technique is applicable. The operation of the two computers that were used to calculate response averages is described first.

In order to compute an average response, the recorded activity that is to be processed and a signal indicating the onsets of the stimuli must be available in their correct temporal relationship. If the activity is recorded onto magnetic tape for subsequent processing, the stimulus signal must be recorded in such a way that the temporal relationship of these two signals will be preserved when the recording is played back. Figure 2.8 is a sketch of the two signals. The stimulus presentations are marked $t_{0,n+1}$, $t_{0,n+2} \cdots t_{0,n+k}$. The average of the responses is computed using these two signals in the following way: (1) The amplitude of the signal to be averaged is measured at a number of instants following each stimulus onset (such as t_0, $t_1, \cdots, t_j$ for the responses illustrated in Fig. 2.8). (2) Samples taken at the same delay after the stimulus onset are added, and their sum is stored. For instance, the samples taken at $t_{0,n+1}$, $t_{0,n+2}, \cdots, t_{0,n+k}, \cdots$ are added, the samples taken at $t_{1,n+1}$, $t_{1,n+2}, \cdots, t_{1,n+k}, \cdots$ are added, and so forth. (3) After the desired number of responses N has been added in this way, the average of the responses is given by the computed sum times an appropriate scale factor. With use of the notation of Appendix B,

$$M_N(t_j) = \frac{1}{N}\sum_{k=1}^{N} x(t_{j,n+k}), \qquad j = 0, 1, \cdots, P. \qquad (2.1)$$

The averaging process is illustrated in Fig. 2.9. In Fig. 2.9a, all of the individual "responses" are identical, and therefore the waveform of the average is the same as that of any single response. In Fig. 2.9b, random noise has been added, and the individual responses no longer are identical. The amplitude of the noise is

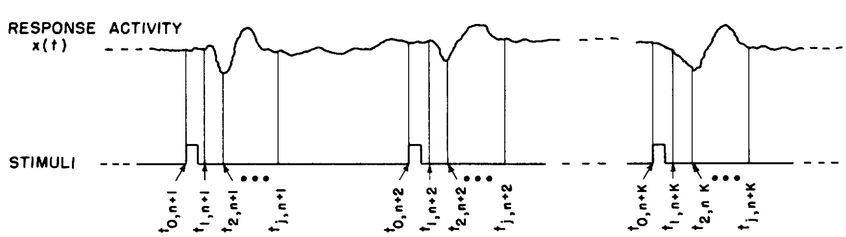

Fig. 2.8. Schematic representation of response activity x(t) and stimuli to aid in the explanation of averaging.

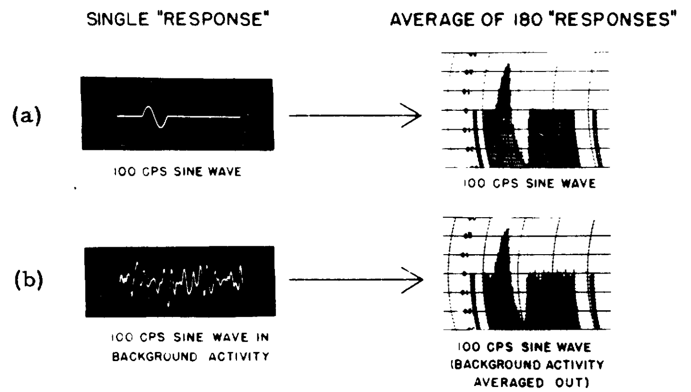

Fig. 2.9. Illustration of the averaging process. (a) The left-hand trace shows a single artificial "response." The right-hand trace shows the average of 180 such responses. (b) The left-hand trace shows a single artificial "response," the sum of a sinusoidal pulse and random noise. The right trace shows the average of 180 such responses (computed by ERD). After Barlow.[8]

large relative to the amplitude of the sinusoidal pulse, so that the pulse cannot be detected visually in the individual responses. In the average, however, the sinusoidal pulse is clearly evident. Since the sinusoidal pulse occurs in each response in exactly the same way, whereas the noise varies randomly about a mean of zero, the sinusoidal pulse will tend to be emphasized relative to the noise as more and more of these responses are added.

An appropriate random process model for responses to stimuli periodically presented is a periodically time-varying random process. It can be shown (see Appendix B) that for such a process, the average of N responses (for a particular value of t_j) is a random variable with a mean equal to the mean of the random process, and a variance that decreases as the number of responses is increased. Hence, if the number of responses is large enough to make the variance small, the average will, with high probability, be very nearly the process mean. For the example of Fig. 2.9, in which the process is the sum of a fixed component (sinusoidal pulse) and a random component with zero average (noise), the sample average looks like the sinusoidal pulse for large N. Hence in this case, the average is an estimate of the pulse waveform. However, neural responses are not necessarily describable by this simple additive model (see p. 29).

Engineering considerations determine the particular form of the electronic computer that is used for the calculation of averages of responses. Considerations important to the design include: (1) the desirability of computing averages during experiments ("on-line" computation), (2) the amount of computer-caused drift and noise that is allowable, (3) the desired resolution in time and amplitude, and (4) the resources available for construction and maintenance of the computer.

Two quite different average-response computers have been designed for use in this laboratory. One is an analog device called the ERD (Evoked Response Detector).[8] For processing by ERD, the data must be recorded onto magnetic tape, because the average $M_N(t_j)$ is computed for only one value of t_j each time the data are fed through the machine. The resulting sum is displayed by a recording meter. To obtain the waveform of the average, the process must be repeated for each value of t_j. A more detailed description of the ERD is given in Appendix C.

The other device is a digital computer called the ARC (Average Response Computer).[13] In this machine, the response activity is sampled at a number of instants after each stimulus onset. The sums $\sum_{k=1}^{N} x(t_j, n+k)$ for all j's are computed and displayed on a cathode-ray tube after each stimulus has been presented. This allows the experimenter to observe how the set of sums builds up as the number of responses N increases. (The set of sums is N times the average of the responses.) When the desired number of responses has been added, the result can be (1) displayed on the oscilloscope, (2) plotted by a pen recorder, or (3) punched onto paper tape. A more detailed description of ARC is given in Appendix C.

Responses to Acoustic Stimuli in Humans. - Electrical activity of the brain can be recorded from electrodes located on the scalp of an awake person, but it is usually difficult to see evoked responses to sensory stimuli in the records of this activity. This difficulty arises because of the presence of activity in the recorded potential, which is not directly related to the stimulus. By averaging responses, the activity which is time locked to the stimulus, tends to be emphasized, so that characteristic deflections can be detected in the average.

Figure 2.10a shows both an ink record and a computed average of the potential between two scalp electrodes obtained while acoustic clicks were being presented to the subject periodically. Figure 2.10b shows the average when there is no acoustic stimulus. The dependence of the early component of the average of the evoked responses on click intensity is illustrated in Fig. 2.11.

A method for detecting evoked responses in awake human subjects is of particular interest, since this makes it possible to obtain psychophysical and physiological responses from the same subject under the same conditions. In the experiment shown in Fig. 2.11, the subject's psychophysical threshold for clicks was approximately -85db. This corresponds quite closely to the lowest stimulus intensity at which the characteristic deflection can just be detected.[14]

Responses to Photic Stimuli in Humans. - In activity recorded from the cortex of anesthetized animals, many workers have observed a series of apparently rhythmic waves following the primary evoked response to sensory stimuli. A similar pattern can sometimes be observed in the activity recorded from the scalp

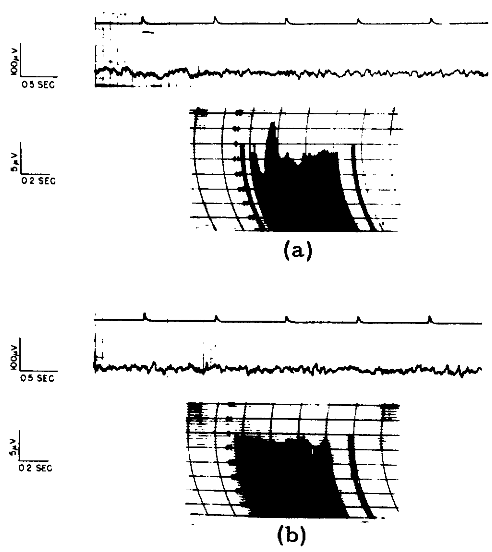

Fig. 2.10. Responses to periodic clicks recorded from the scalp of a human subject. The upper records of (a) and (b) show ink traces obtained from the scalp with and without clicks, respectively. The pulse channel, recorded simultaneously, indicates in (a) the times of click presentations, and in (b) serves as a comparable time reference. The lower records in each case are the average (computed by ERD) of 250 consecutive responses, a few of which are shown in the ink traces directly above. The clicks, about 70 db above psychophysical threshold, were presented at a rate of 1 per second. Upward deflection signifies positivity of an electrode at the vertex with respect to an occipital electrode. (Subject H-436.)

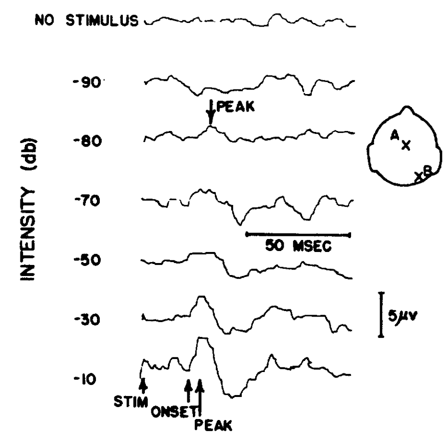

Fig. 2.11. Average responses to periodic clicks obtained from the scalp of a human subject. Each trace represents the average (computed by ARC-1) of 400 responses to identical click stimuli, presented at a rate of 1.5 per second. Upward deflection indicates that electrode A is positive with respect to electrode B. (Subject H-480, awake, eyes open.) After Geisler, Frishkopf, and Rosenblith.[14]

of humans during stimulation by flashes of light. This rhythmic activity appears to be about the same frequency (approximately 10/sec) as the "alpha activity" that is often observed in the EEG of resting subjects. (See Chapter 3 for a discussion of alpha rhythm.) It has been demonstrated that alpha activity tends to disappear for a short time following stimulation by a light flash. Hence rhythmic activity occurring after a flash might be the reappearance of activity that is not time-locked to the stimulus.

Averaging responses tends to emphasize activity that is time-locked to the stimuli; however, even if the rhythmic activity were not time-locked to the stimuli, the rhythmic component in the average of the response will not, in general, be exactly zero. Hence it is important to be careful in deciding from the average of responses whether or not the activity is time-locked to the stimuli. In particular, one can observe the way in which the average changes as the number of responses is increased. If the rhythmic activity were not time-locked to the stimuli, it should decrease in amplitude as the number of responses in the average is increased. In Fig. 2.12, averages are shown for 45 and 180 responses. Since the amplitude of the rhythmic activity is about the same in both cases, it can be stated that the rhythmic activity was time-locked to the stimuli. The prominence of this activity varies considerably among subjects.

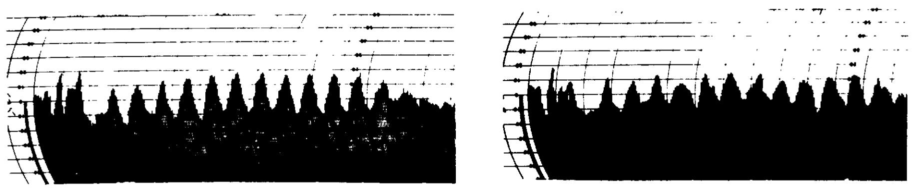

Fig. 2.12. Average responses obtained from the scalp of a human subject stimulated by randomly delivered stroboscopic flashes. The figure on the left shows the average (computed by ERD) of about 180 responses, while that on the right shows the average of about 45 responses. Potentials were obtained from electrodes placed on the left parietal and occipital regions of the scalp. The time markers are separated by 100 msec.

Changes in Evoked Response Waveform with Anesthesia. - Responses to acoustic stimuli recorded from the auditory cortex of cats can be detected visually after almost every stimulus presentation. However, the variability of the response is so great that it is difficult to quantify changes in the waveform of evolved responses. Averaging of responses provides more stable measures and permits studies of response parameters as a function of both stimulus and "state" variables.

Figure 2.13 shows several examples of single responses to clicks recorded during three different conditions: (a) before anesthetization, (b) five minutes after injection of Nembutal into the femoral vein, and (c) one hour later. Averages of 300 responses taken during each of the above conditions are shown in Figure 2.14. It is clear that the shape of the average of responses changes with the state of anesthesia and that quantitative statements can be made about differences in the waveforms.

Responses from the Auditory Nervous System to Repetitive Acoustic Stimuli. - Although the electrical response of the auditory nerve to impulsive acoustic stimuli is well defined at low stimulus repetition rates (see Fig. 2.15), the response amplitude decreases considerably at high rates, and measurement of the response from oscilloscope records becomes difficult. This difficulty arises because the neural component of the recorded potential is approximately the same amplitude as both the background activity and the cochlear microphonic potential. Figure 2.15 shows that, whereas at low stimulus repetition rates and moderate intensities the microphonic potential does not interfere appreciably with the neural component, at high rates it is difficult to see the neural component. Although the random background activity can be reduced relative to the neural component by averaging responses, the microphonic potentials will be emphasized by averaging in the same way as the neural potentials since they are also time-locked to the stimuli. The microphonic potential is linearly

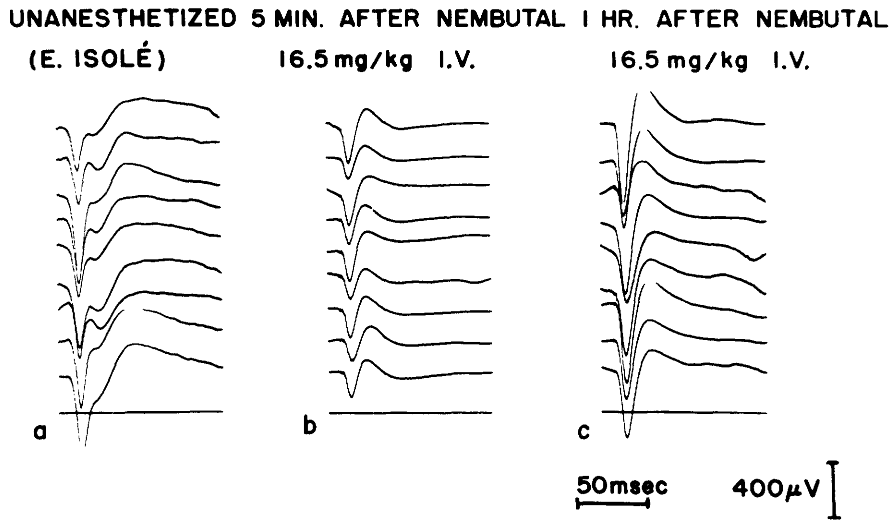

Fig. 2.13. Effects of intravenous Nembutal on the waveform of cortical responses obtained from a cat. Responses to periodic clicks presented at a rate of 1 per second are illustrated. The responses were recorded (a) before injection of anesthesia (encephale isolé, the cat was immobilized by a high spinal section); (b) 5 minutes after injection of anesthesia; and (c) 1 hour after anesthesia injection. The potentials were recorded from the middle ectosylvian gyrus of the contralateral cortex using a concentric bipolar electrode (center electrode 2 mm below surface of the cortex, sleeve on surface). Intensity was 35 db above VDL. (Cat 459.)

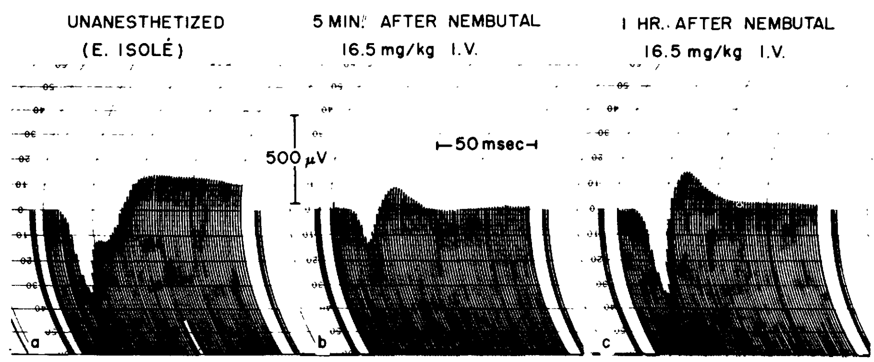

Fig. 2.14. Effects of intravenous Nembutal on the waveform of cortical responses obtained from a cat. Each record represents the average (computed by ERD) of 300 responses to successive clicks. Data taken during experiment described in Fig. 2.13.

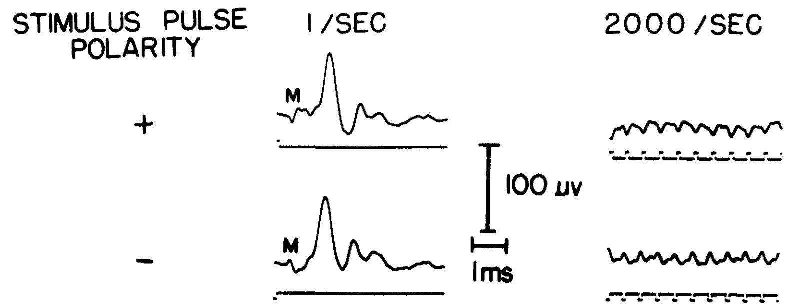

Fig. 2.15. Auditory nerve responses to acoustic clicks obtained from an anesthetized cat (Dial anesthesia). Responses were recorded between a wire electrode near the round window and a reference electrode attached to the headholder. Note that the microphonic component (labeled M in the left-hand responses) reverses polarity with reversal of stimulus polarity. Intensity was 35 db above VDL. (Cat 498.)

related* to the acoustic stimulus for low intensity stimuli, whereas the neural potential is a nonlinear function of the stimulus. Hence, if the stimuli are bursts of random noise, the microphonic potential will vary randomly about a zero average value, and in the average of responses the microphonic potentials will be deemphasized because they will tend to be negative just as often as they are positive. On the other hand, since the neural potentials always have the same polarity (as seen in Fig. 2.15), they will tend to be emphasized in the average of responses. Figure 2.16 shows that at low stimulus repetition rates the ratio of the microphonic component to the neural component is maintained in the average of responses to pulses, whereas this ratio is reduced considerably in the average of responses to bursts of noise. Figure 2.17 shows both superimposed single responses and averages of responses for several repetition rates. With use of this technique, neural responses to repeated bursts of noise have been detected at rates up to 3000/second.

Cortical responses likewise become increasingly difficult to detect as the stimulus repetition rate is increased. In the unanesthetized preparation, stimulus-locked activity tends to be obscured at a lower rate than in the anesthetized animal because of greater background activity, making it difficult to compare the two situations. Results from a study using averaging of responses to overcome the detection problem are shown in Fig. 2.18. These

* That is, the microphonic potential can be thought of as the response of a linear system to the stimulus. When the stimulus polarity is reversed, the polarity of the microphonic potential also reverses.

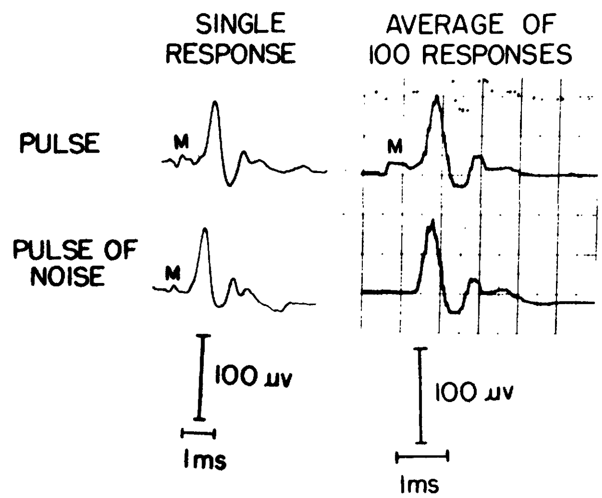

Fig. 2.16. Single and average auditory nerve responses. 0.1-msec rectangular pulses and 0.1-msec pulses of noise were applied to the earphone at a repetition rate of 1 per second. Same preparation as that of Fig. 2.15.

data indicate time-locked activity is present at higher stimulation rates in the unanesthetized preparation than in the anesthetized.[15]

"Off" Responses from the Auditory Cortex. - Another response which is difficult to detect because of the large background activity present in unanesthetized preparations is the cortical response that is evoked by abruptly turning off a burst of acoustic noise. This response appears to be present only in unanesthetized preparations. Figure 2.19 shows averages of these responses recorded before and after the administration of Nembutal. Figure 2.20 is an example of the changes in the averages of these responses as a function of the intensity of the noise bursts. The stability of the average response makes it possible to measure quantitatively the way in which characteristics of the average of off-responses changes with intensity.

Measurement of Other Statistics of Evoked Responses

In the preceding section, several examples showing the utility of averaging have been described. Although averaging is useful for the detection of stimulus-locked activity, and though it yields stable measures of response characteristics, it does not provide a complete statistical description of the responses. In this section some results are presented from computations that were designed to give a more complete statistical description.

The computation was carried out in the following way. A number of responses to identical stimuli were sampled at particular times t_j after stimulus onset. These samples were quantized so that

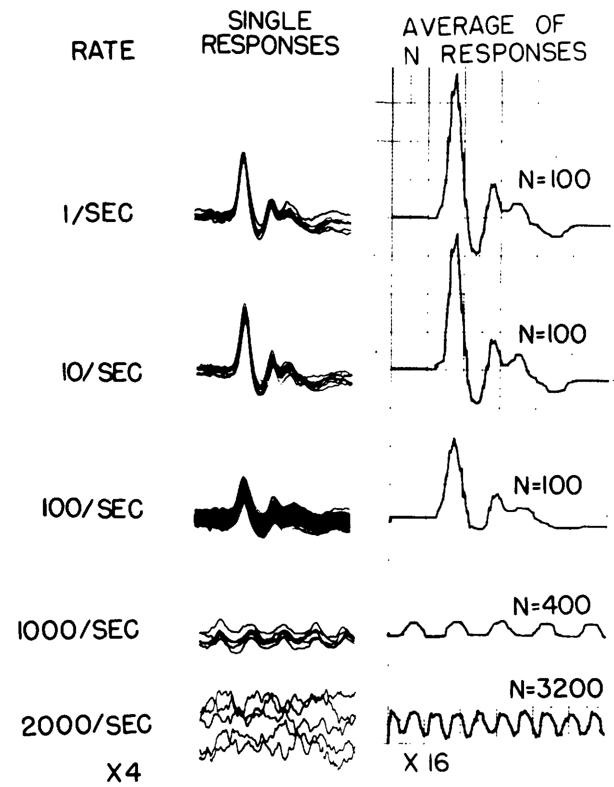

Fig. 2.17. Single and average auditory nerve responses to bursts of noise 0.1 ms in duration. N is the number of responses averaged. (The gain is increased by 4 in the lower left-hand figure and by 16 in the lower right-hand figure.) Same preparation as that of Fig. 2.15.

there was a finite number of amplitude levels which the sample could assume. The number of responses occurring at each level was computed. The resulting set of numbers was used to plot a histogram such as that shown in Fig. 2.21.

The family of histograms obtained for many values of t_j are estimates of approximations to the first-order probability density functions at these times. The sample average response, which has been discussed in the previous sections, is an estimate of the first moments of these distributions.

In order to describe the way in which these distributions change as a function of time t_j, it is convenient to use measures of certain aspects of the distributions such as central tendency, dispersion, and skewness. In this case, a general-purpose digital computer (the Lincoln Laboratory TX-0) was programmed to calculate certain statistics and to display them with the his-

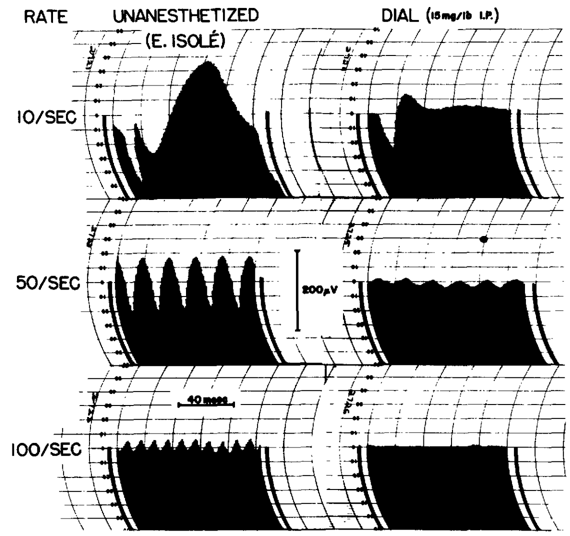

Fig. 2.18. Average of cortical responses to repeated clicks. The responses in the left column were recorded while the cat was unanesthetized (encephale isolé). The responses in the right column were recorded after Dial was injected into the peritoneal cavity. Number of responses averaged: 10/sec, 600; 50/sec, 3000; 100/sec, 6000. Averages were computed by ERD. The size of responses to clicks presented 1 per second was approximately equal before and after anesthetization. Clicks were presented monaurally at an intensity of 25 db above VDL. Recording from a concentric bipolar electrode in contralateral middle ectosylvian gyrus. (Cat 446.) After Goldstein, Kiang, and Brown.[15]

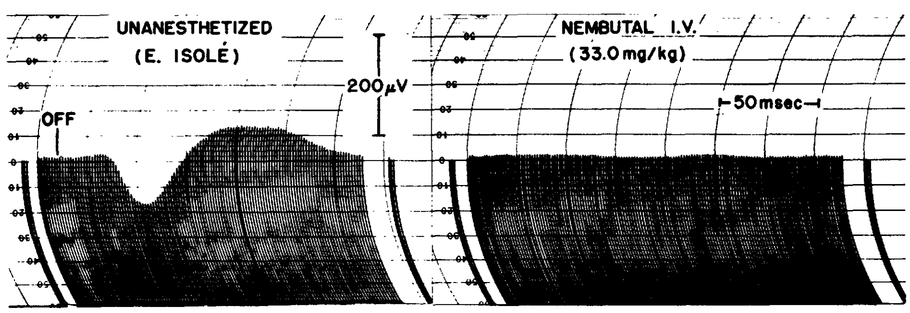

Fig. 2.19. Effects of intravenous Nembutal on "off" responses to bursts of noise obtained from a cat. Each trace is the average (computed by ERD) of 300 responses. Bursts 375 msec in duration, having a 20 msec rise and fall, and at an intensity 60 db above VDL, were presented at a repetition rate of 1 per second. Recordings were made with a concentric bipolar electrode in area II of the contralateral auditory cortex. (Cat 461.)

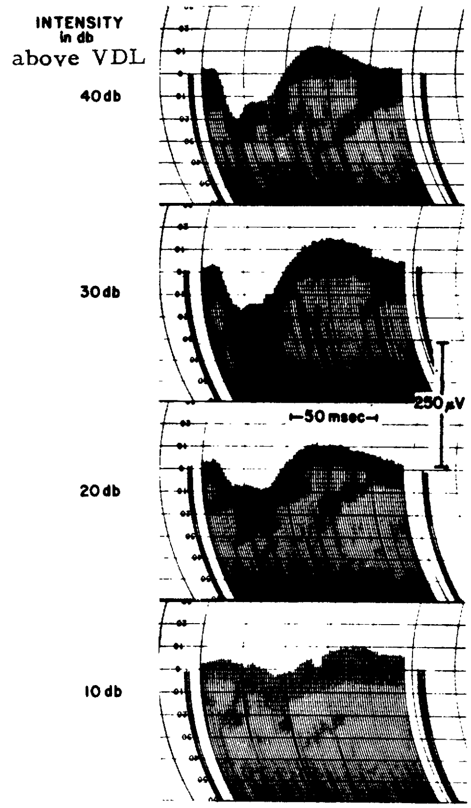

Fig. 2.20. Effects of intensity of stimulation on "off" responses to bursts of noise recorded from area II of the auditory cortex of an unanesthetized cat (Encephale isolé). Bursts 50 msec in duration, having an abrupt rise and fall, were presented at a repetition rate of 1 per second. (Cat 467.) Each trace is the average of 300 responses (computed by ERD).

tograms. The numbers in Fig. 2.21 represent (reading from top to bottom) the first, second, and third quartiles* (Q_1, Q_2, Q_3), the interquartile range (Q_3 - Q_1), the number [(Q_3 - Q_2) - (Q_2 - Q_1)], and the number of responses. The second quartile, which is also called the median, is a measure of the central tendency of the distribution; the interquartile range is a measure of the dispersion; and [(Q_3 - Q_2) - (Q_2 - Q_1)] is a rough measure of skewness. Fig. 2.22 is a plot of the interquartile range and the median as a function of time after the stimulus t_j. It is clear from this picture that the dispersion is not the same for all values of t_j. A mathematical model consisting of an invariant response plus independent random noise (such as that illustrated in Fig. 2.9) has a dispersion which is constant as a function of t_j. Hence, these results indicate that this simple model does not adequately describe these data.

* The quartiles are defined as the three values of amplitude that separate the histogram into four sections, with one quarter of the samples in each section.

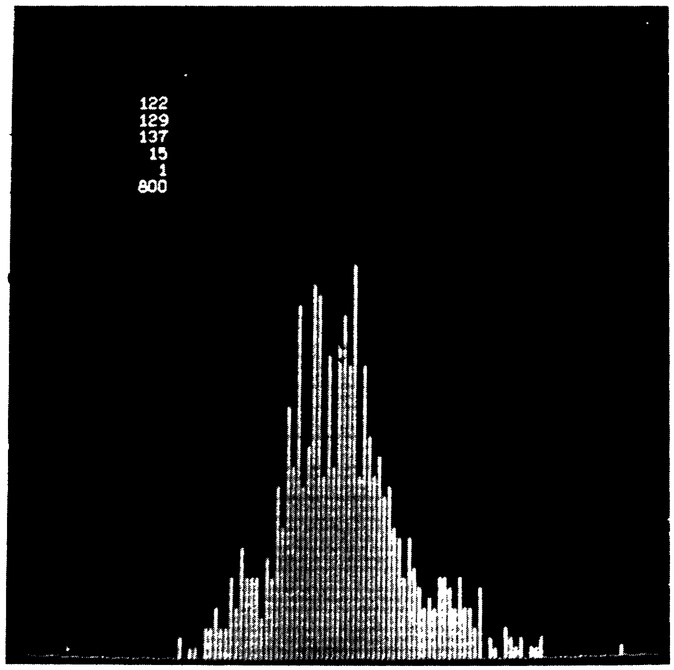

Fig. 2.21. Histogram of the amplitudes of 800 cortical responses taken 1.125 msec after the presentation of identical auditory stimuli. The stimuli were 2900-cps tone bursts, having a duration of 150 milliseconds and rise and decay times of 5 msec. Stimulus intensity was 40 db above VDL. The preparation was an unanesthetized cat, immobilized by a high spinal section. (Cat 496.)

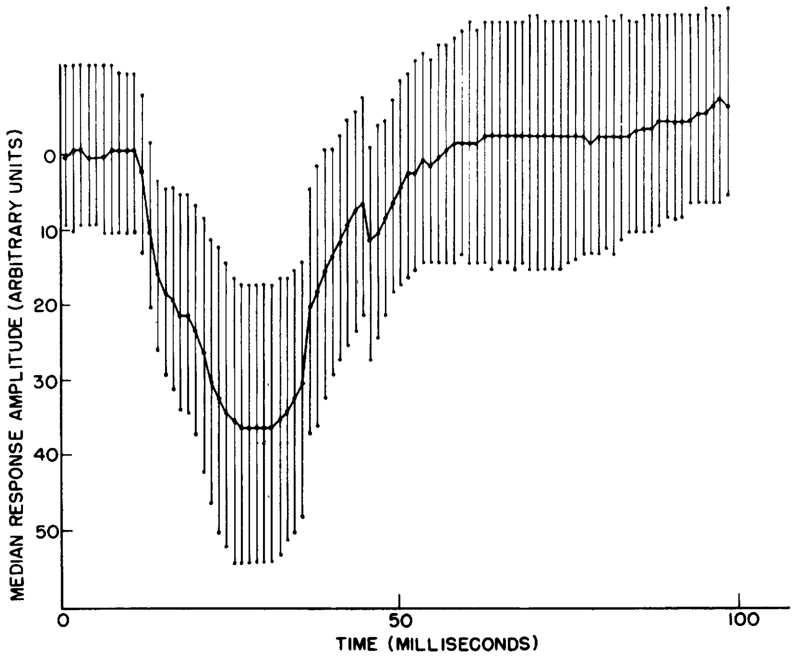

Fig. 2.22. The median and interquartile range of 800 cortical responses obtained from an unanesthetized cat. Points were calculated every 1.175 msec. The parameters of the first point in time were calculated from Fig. 2.21. Same preparation as that of Fig. 2.21.

REFERENCES

1. C. N. Woolsey, "Patterns of Sensory Representation in the Cerebral Cortex," Federation Proc., 6, 437-441 (1947).

2. W. H. Marshall, C. N. Woolsey, and E. P. Bard, "Cortical Representation of Tactile Sensibility as Indicated by Cortical Potentials," Science, 85, 388-390 (1937).

3. A. R. Tunturi, "Physiological Determination of the Boundary of the Acoustic Area in the Cerebral Cortex of the Dog," Am. J. Physiol., 160, 395-401 (1950).

4. R. Granit and A. Wirth, "A Scotopic 'Blue Shift' Obtained by Electrical Measurements of Flicker Resonance," J. Physiol. (London), 122, 386-398 (1953).

5. E. D. Adrian and D. W. Bronk, "The Discharge of Impulses in Motor Nerve Fibers," J. Physiol. (London), 67, 119-151 (1929).

6. G. D. Dawson, "Cerebral Responses to Electrical Stimulation of Peripheral Nerve in Man," J. Neurol., Neurosurg., Psych., 10, 134-140 (1947).

7. G. D. Dawson, "A Summation Technique for the Detection of Small Evoked Potentials," Electroencephalog. and Clin. Neurophysiol., 6, 65-84 (1954).

8. J. S. Barlow, "An Electronic Method for Detecting Evoked Responses of the Brain and for Reproducing Their Average Waveforms," Electroencephalog. and Clin. Neurophysiol., 9, 340-343 (1957).

9. M. H. Goldstein, Jr., and N. Y-S. Kiang, "Synchrony of Neural Activity in Electric Responses Evoked by Transient Acoustic Stimuli," J. Acoust. Soc. Am., 30, 107-114 (1958).

10. A. McF. Mood, Introduction to the Theory of Statistics, McGraw-Hill Book Company, New York, 1950.

11. W. A. Rosenblith, "Some Electrical Responses from the Auditory Nervous System," Proceedings of the Symposium on Information Networks, Polytechnical Institute of Brooklyn, pp. 223-247, April 12-14, 1954.

12. L. S. Frishkopf and W. A. Rosenblith, "Fluctuations in Neural Thresholds," Symposium on Information Theory in Biology, H. P. Yockey, R. L. Platzman, and H. Quastler, Eds., Pergamon Press, Inc., New York, pp. 153-168, 1958.

13. W. A. Clark, Jr., "Average Response Computer (ARC-1)," Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., pp. 114-117 (April 15, 1958).

14. C. D. Geisler, L. S. Frishkopf, and W. A. Rosenblith, "Extracranial Responses to Acoustic Clicks in Man," Science, 128, 1210-1211 (1958).

15. M. H. Goldstein, Jr., N. Y-S. Kiang, and R. M. Brown, "Responses of the Auditory Cortex to Acoustic Stimuli," J. Acoust. Soc. Am., 31, 356-364 (1959).

Chapter 3

TWO TECHNIQUES FOR THE PROCESSING OF EEG DATA

C. E. Molnar
T. F. Weiss
C. D. Geisler

3.1 Introduction

The previous chapter has discussed techniques for quantitatively describing the electrical activity of parts of the nervous system following the presentation of a known stimulus. In many cases, particularly when the preparation is not heavily anesthetized, fluctuating potentials may be observed when the animal is not being stimulated deliberately. These "nonevoked" potentials may be observed with microelectrode techniques in single nerve cells or groups of cells, with gross electrodes located in the central nervous system, or even with electrodes placed on the scalp of a human subject or experimental animal. The potentials observed in the last case are often referred to as the electroencephalogram (EEG).

Description of these nonevoked potentials poses problems that are not encountered in the study of evoked responses. All of the techniques described in Chapter 2, for the analysis of evoked responses, depend upon knowledge of the time at which a stimulus has been presented. The results of these analyses are generally in a form that shows the dependence of some quantity, calculated from the potentials, upon the time elapsed since the particular stimulus under study was presented. Even if the nonevoked potentials were caused solely by the continuous flow of sensory information originating in the environment of the subject, the techniques described earlier could not be applied readily; there is no exact way to determine the instant at which such "natural" stimuli occur or to ensure that the responses to various stimuli have enough in common to justify averaging them. Furthermore, fluctuating potentials are still observable when an experimental animal or human subject is isolated from visual and auditory stimuli in a soundproof and totally darkened room. Under these conditions, olfactory, tactile, or taste stimuli may still occur and produce effects within the nervous system. Moreover, the flow of information from proprioceptors and other internal sources may also produce effects like those caused by external stimuli. Isolating the brain of a cat from many of these inputs by severing the spinal cord still does not eliminate fluctuating brain potentials,[1] suggesting that perhaps these potentials represent an intrinsic property of the central nervous system and are not merely responses produced by external causes. In any case, the problem of describing the fluctuations remains essentially the same whether or not one regards them as being produced by causes external to the central nervous system.

Visual examination of the potentials recorded from the scalp of human subjects and from the cerebral cortex of cats suggests that certain features of these potentials depend upon whether the subject is awake or asleep; whether he is paying close attention to his surroundings or ignoring them; and upon other factors which appear to be related to the "physiological state" of the animal rather than to the presence or absence of specific stimuli. Figure 3.1 illustrates the differences between samples of EEG recordings taken from the

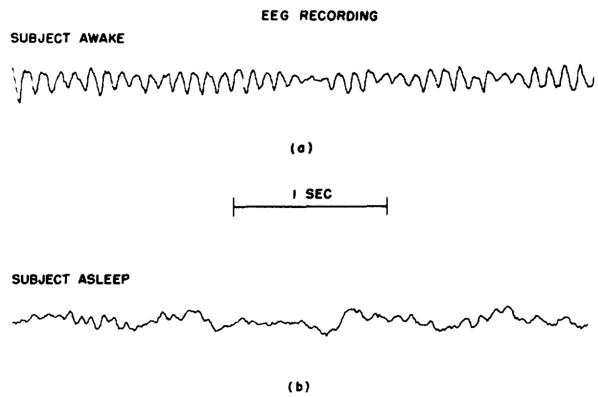

Fig. 3.1. EEG ink traces from the same subject; (a) awake and (b) asleep. Subject was in dark, anechoic chamber. Electrodes were located at vertex and left occipital area. The ink trace represents 3 1/2 seconds of data.

same subject while awake and asleep. Figure 3.2, on the other hand, shows the differences in samples of EEG recordings from seven different subjects, all taken under similar conditions A description of the recordings appears to require an examination of the potential over a period of several seconds or even minutes. It is therefore desirable that a data-processing technique be available which can extract information contained in several minutes of a recording and present it in a form which is more compact and easier to interpret than the original record.

It is not immediately apparent which properties of the record should be emphasized and retained by such a data-processing technique, and which should be minimized or eliminated. If the mechanisms producing the potentials were more completely understood,

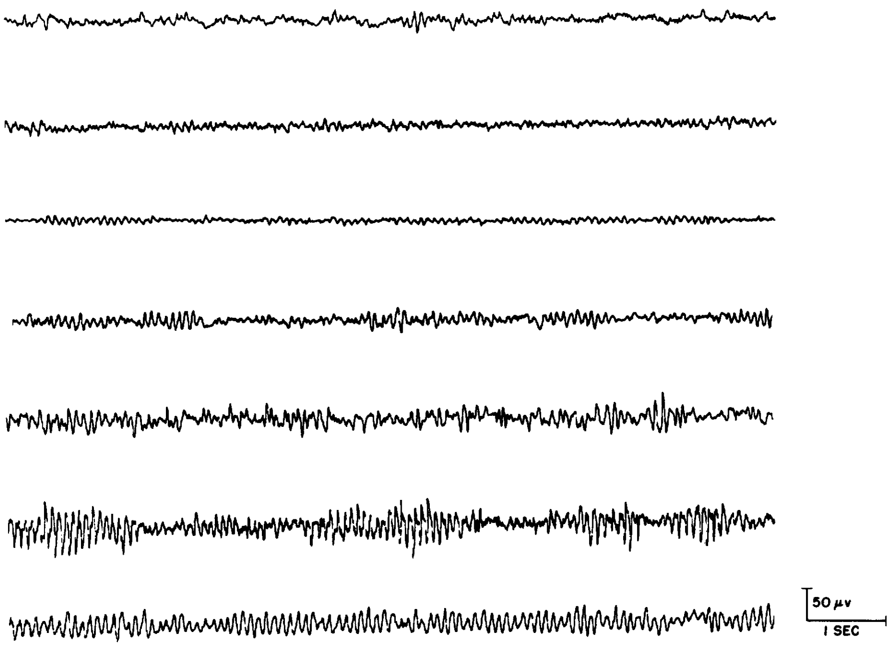

Fig. 3. 2. EEG ink traces of seven different subjects. All subjects were seated in a dark, anechoic chamber. Electrodes were located in the left parietal-occipital area. The ink traces represent 10 seconds of data in each case.

one could determine which properties of the potentials are capable of yielding the most information about the functioning of the mechanisms. However, there is at present no way to decide which type of analysis is most appropriate, other than an empirical comparison of different techniques.

An appropriate technique for the study of a particular experimental question should produce results that are repeatable, that is, the variability in the results obtained when an experiment is repeated as exactly as possible should be much smaller than the change in the results produced by a change in the experimental conditions of interest. Ideally, the data-processing technique should yield results that depend only upon those parameters of the experiment that are being investigated in a particular case, and not at all upon incidental parameters. Whether or not this ideal is approached in any particular case depends upon the experimental questions being asked as well as upon the data-analysis techniques.

Two distinct approaches to the analysis of EEG records will be discussed in detail. The first method represents an effort to perform electronically a type of analysis that is similar to that carried out by an electroencephalographer when he visually examines the EEG for certain rhythmic characteristics. The second method, correlation analysis, was suggested originally by the similarity of some EEG records to certain types of random signals encountered in communication engineering. The success of correlation analysis in the study of these random signals indicated that it might also be usefully applied to the EEG.[2] The results of this type of analysis present information that can be related to the results obtainable by a frequency analysis of the signal. The discussion of correlation analysis will depend to some extent upon the mathematical properties of random processes. An introduction to this latter subject may be found in Appendix A and in the references listed there.

3.2 An Examination of Some Temporal Patterns in the EEG

One of the most striking properties in the EEG records of many subjects is the occurrence of rhythmic "bursts," a typical example of which is shown in Fig. 3.3.[3]

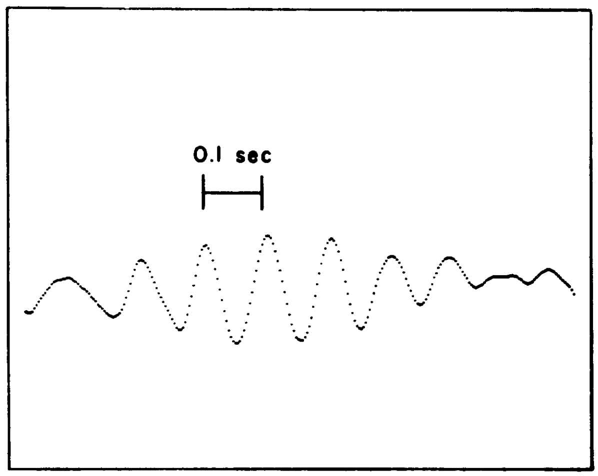

Fig. 3.3. An example of a rhythmic burst in an EEG record, as photographed from the TX-0 cathode-ray tube display. The data represented have been sampled at 300 samples/sec and a 10-point moving average has been used to smooth the data.

A classification of the amount of this activity present is sometimes obtained by a visual examination of a paper record of EEG,[4] but cannot be regarded as a fully objective and repeatable measure of the rhythmic burst activity. A refinement of the visual examination technique is the measurement of the "alpha index,"[5] the percentage of the record which contains alpha rhythm. This method provides a more quantitative measure of the alpha rhythm, and depends upon the manual measurement of the EEG wave as recorded on a paper record and is impractical for the examination of large amounts of data. Certain computer applications to the measurement of EEG data have also been made.[6]

Many limitations of the visual and manual techniques for investigating the rhythmic bursts in the EEG have been overcome by an application of modern computing techniques to the problem. This particular approach is an extension and mechanization of the general ideas involved in the methods just mentioned. The analysis was performed by the computer system shown in Fig. 3.4. Data recorded

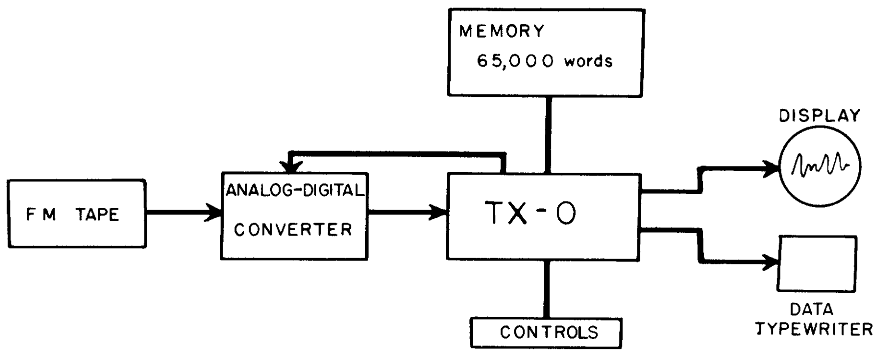

Fig. 3.4. Schematic diagram of the digital computer system used for the measurement of rhythmic burst activity.

on magnetic tape were periodically sampled by an analog-to-digital converter and transferred to the Lincoln Laboratory TX-O computer.[7] In effect, the computer stored a three-minute sample of an EEG record and then classified portions of the record according to three well-defined criteria. The detailed operation of the computer program which executed this analysis is described elsewhere[3] and will not be given here. The criteria used to define rhythmic burst activity and the way in which they are applied are of some interest, however, and will be discussed briefly.

First, a section of record must have at least a certain minimum peak-to-peak amplitude in order to be considered further. This minimum amplitude is defined in terms of the median peak-to-peak amplitude of the particular EEG record being analyzed. A fraction expressing the relation of the minimum amplitude to the median amplitude is called the amplitude parameter or AP.

The second criterion requires that the interval between two successive "zero crossings," lie within a specified range. The limits within which an interval must lie to be acceptable are defined as the median zero-crossing interval plus and minus a specified fraction (interval parameter) of this median interval. The interval parameter is denoted by IP.

The third criterion requires that at least five consecutive zero-crossing intervals meet both the amplitude and interval criteria. Those portions of the record that satisfy all three criteria are then defined as rhythmic burst activity.

The computer then tabulates the number of rhythmic bursts, and the total activity, which is defined as the percentage of the record that consists of rhythmic burst activity. Information about the distribution of burst lengths, peak-to-peak amplitudes, and zero-crossing intervals is also available.

This process was first applied to the study of individual differences in the EEG records taken from different human subjects under comparable conditions. Figure 3.5 indicates the range of the results obtained from the analysis of records from four different subjects. These results are shown in two graphs; (a) total (rhythmic) activity and (b) number of bursts, plotted as a function of the amplitude parameter.

Before the differences in the graphs for different subjects may be interpreted as significant, one must consider the variability of the results obtained from any one subject. Figure 3.6 shows the results of processing EEG records taken from the same subject on four different days. The thin lines in the figure show the results for each day, and the heavy line shows the average of the results for the four days. The shaded areas enclosing the curves represent a rough estimate of the limits within which 99% of graphs of the data points for this subject would be expected to fall.

Any pair of curves taken from Fig. 3.5 is found to have at least a few points lying outside of the shaded areas of Fig. 3.6. This indicates that the experiments (in other words, the same subject under similar conditions) that generated the data of Fig. 3.5 were different, with high probability, from those experiments that yielded the data of Fig. 3.6.

Another experiment was conducted to determine how the amount of rhythmic burst activity depends on time elapsed from the beginning of a single long EEG recording. EEG records were taken from subjects who were seated in a soundproof room and asked to close their eyes but to remain awake. The lights were then turned out, and a continuous recording of 12 minutes duration was made. At the end of this long recording, the lights were turned on, and the subject was allowed to talk and move freely for about 3 minutes. Following this intermission, the previous experimental conditions were restored and an additional 3-minute recording was taken.

This experiment was repeated six times for each of four subjects. The EEG records thus obtained were divided into 3-minute samples and processed, yielding the results shown in Fig. 3.7.

It is obvious from the data that the amount of rhythmic burst activity for these subjects decreases as a function of time. Note the recovery in the amount of burst activity after the intermission.

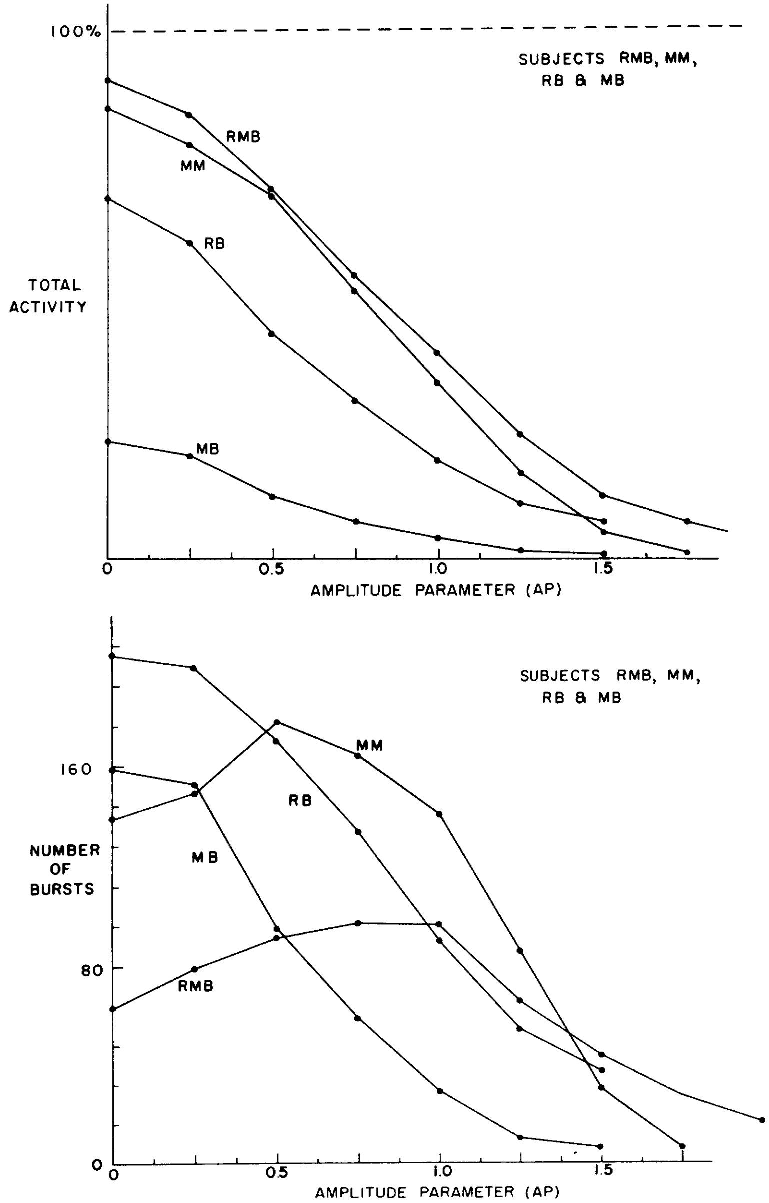

Fig. 3.5. (a) Total rhythmic burst activity and (b) number of bursts, as a function of the amplitude parameter for four different subjects. Each subject was seated in dark, anechoic chamber with electrodes located in the parietal-occipital area. The curves were computed from 3 minutes of data.

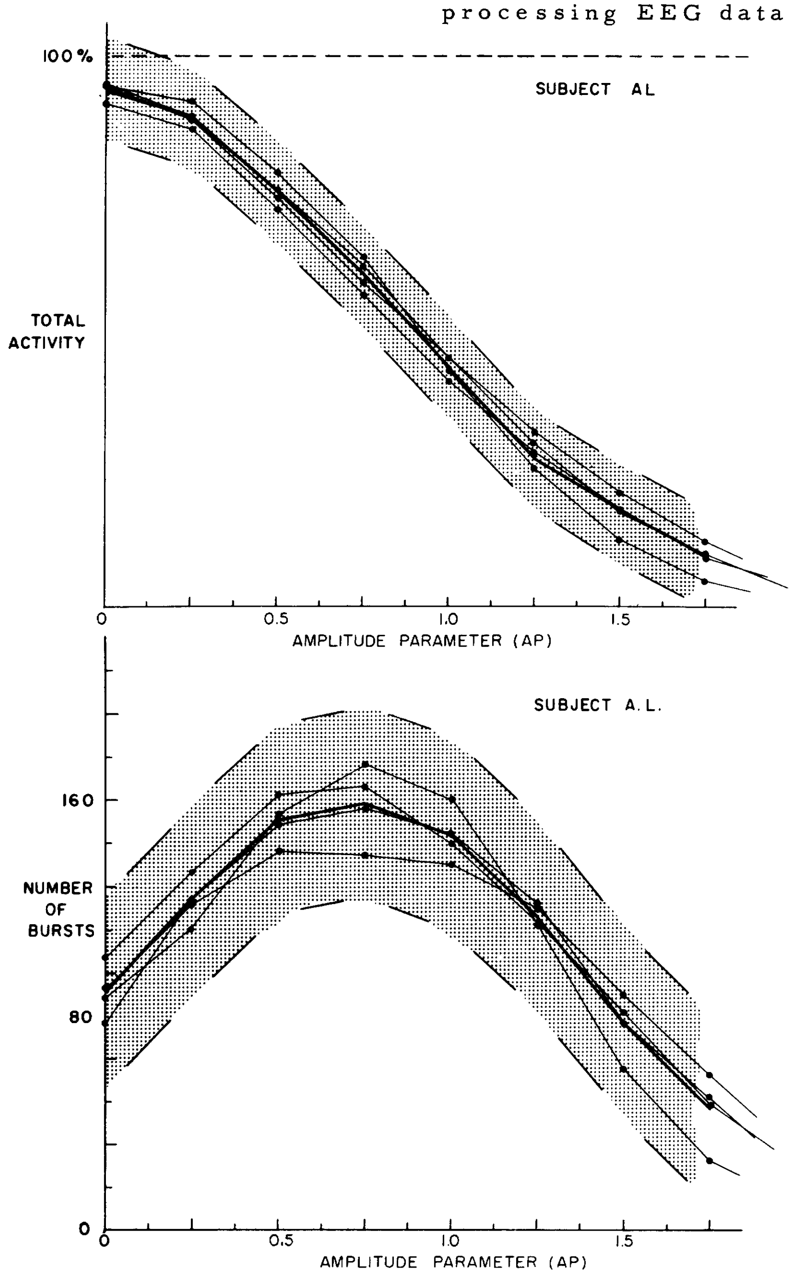

Fig. 3.6. (a) Total rhythmic burst activity and (b) number of bursts plotted as a function of the amplitude parameter for one subject on four different days. Data were taken while the subject sat in a dark, anechoic chamber with electrodes located in the parietal-occipital area. The curves were computed from 3 minutes of data.

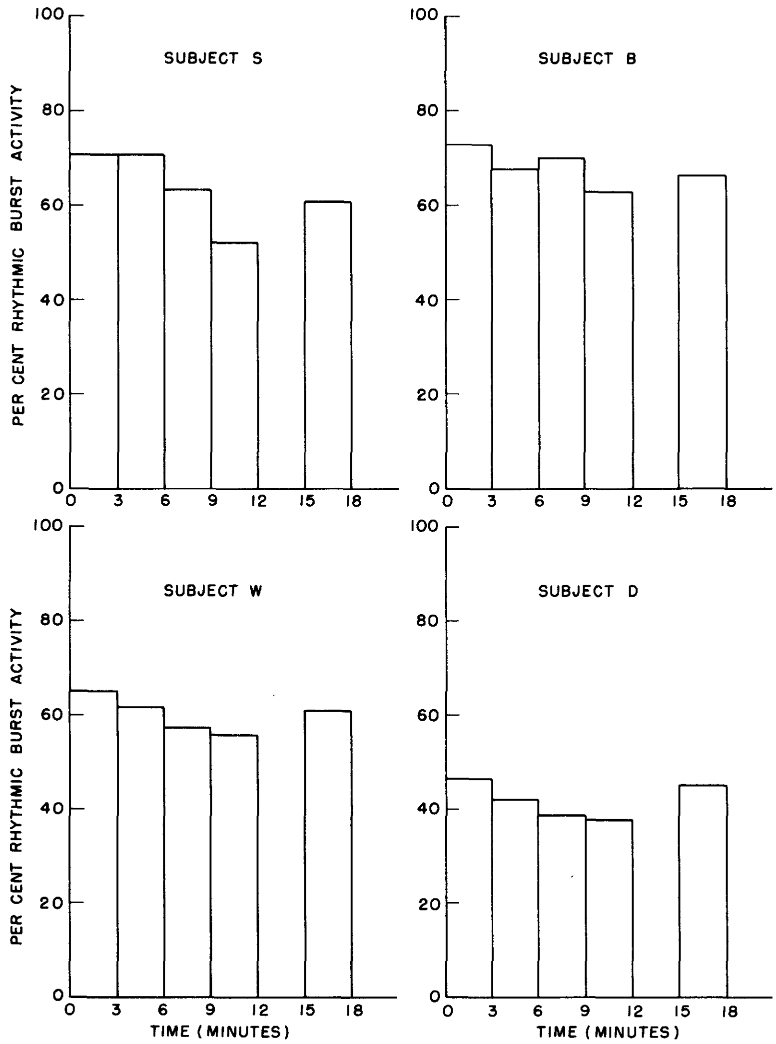

Fig. 3.7. Rhythmic burst activity as a function of time during the experiment for four different subjects. The results shown are the average of results for six different days for each subject. Data were taken while the subject sat in a dark, anechoic chamber with electrodes located in the left parietal-occipital area.

The data-processing technique just described has been an attempt to define a set of variables that describe certain features of EEG data. The approach has been empirical and the particular processing that was finally used evolved from a trial-and-error procedure.

The experience gained in the development and application of this technique suggests several thoughts both about the types of problems that may be amenable to this sort of processing and about the role of general-purpose digital computers in developing techniques for studying such problems.

The technique presented here differs from most of the other techniques discussed in this monograph in one important respect —attention was focused upon particular events in a waveform. That is, instead of doing some sort of averaging (which is the essence of the techniques presented in Chapter 2 and of the autocorrelation technique to be discussed later in the chapter), the TX-O computer was programmed to identify particular events on the basis of a set of criteria whose parameters were determined by a preliminary analysis of the data. The experiments discussed in detail in this section, have by no means exhausted the possibilities inherent in the general concept of "marking" various portions of a waveform and selectively processing these portions. We have, for example, considered the use of this sort of approach in the study of the very challenging and theoretically important problem of the interrelationship of evoked and nonevoked activity.

Studies such as the one just described need not be made with a general-purpose digital computer; the versatility and flexibility of such a computer constitute, however, a considerable advantage. This is particularly true when the particular method of data processing best suited to the problem at hand is not apparent from the start, and empirical comparisons of different techniques must be used in deciding which one is "best." In this respect, flexible input and output devices make it easier to monitor the operation of particular techniques at every step of their operation, thus guiding one toward a set of criteria that most closely approximates the results desired. Once a particular scheme is found to be appropriate for the study of some problem, devices other than those making use of general-purpose computers may turn out to be more economical, more rapid, or more efficient. During the early stages of an investigation, however, a general-purpose computer is a most helpful tool.

3.3 The Use of Correlation Techniques in the Study of EEG

The rhythmic activity often seen in the EEG had suggested earlier the use of still another data-reduction technique, correlation.[8] This technique can be viewed most simply as the performance of a specific operation on data that may be too complex or too long to interpret visually. For instance, consider Fig. 3.8. The left-hand curve in each case shows 3 1/2 seconds of EEG data. Now consider the problem of comparing two such EEG records of 100 seconds duration that might occupy 10 feet of paper record. By use of an analog correlator (described in Appendix C) the curves on the right of Fig. 3.8, autocorrelograms, were obtained.

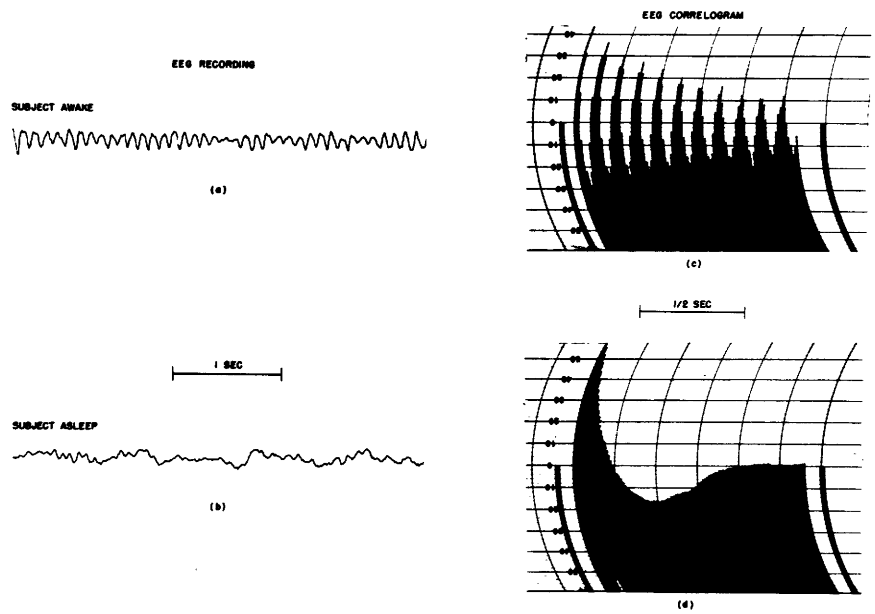

Fig. 3.8. EEG ink traces and autocorrelograms for subject, (a) awake and (b) asleep. The ink traces shown represent 3 1/2 seconds of data. The autocorrelograms were computed from a sample length (τ) of 100 seconds of data. The delay increment ($\Delta\tau$) is 10 milliseconds, and the maximum delay of the autocorrelogram is 1.0 second. The data were recorded from a subject in a dark, anechoic chamber. Electrodes were located in the parietal-occipital area. (Subject H-432.)

The rhythmic activity in the upper EEG record is emphasized by its correlogram, while the lower correlogram reflects the lack of rhythmic activity in the lower EEG record. In this case, the differences reflect the difference in EEG recordings obtained from the same subject when he is asleep and when he is awake. Figure 3.9 is another example of the use of the autocorrelogram as an aid in the interpretation of experimental data.[9] This figure shows autocorrelograms of EEG records taken from the left and right sides of the head of a normal subject, while Fig. 3.10 shows results of the same computational procedure[10] applied to records taken in the same way from the left and right sides of the head of a subject with a brain tumor. Another method of processing such data is the computation of crosscorrelograms; the computational procedure is similar to that used to obtain autocorrelograms, except that now two different signals are involved. The crosscorre-

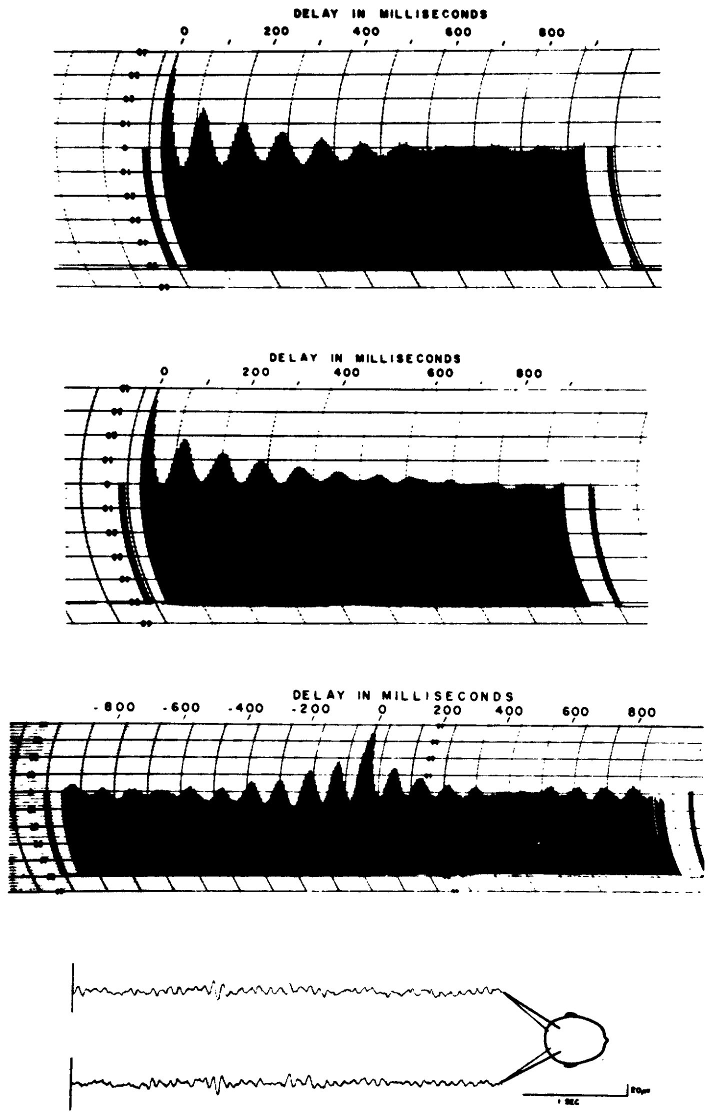

Fig. 3.9. Correlograms computed from the EEG of a normal human subject: (top) autocorrelogram of EEG from the left parietal-occipital area, (second from top) autocorrelogram of EEG from the right parietal-occipital area, (third from top) crosscorrelogram of data from left and right parietal-occipital areas and (bottom) inked trace of EEG showing electrode placement. The delay increment ($\Delta\tau$) is 5 milliseconds and 1.0 minute sample length (T) of data was used for the computation. After Barlow and Brown.[10]

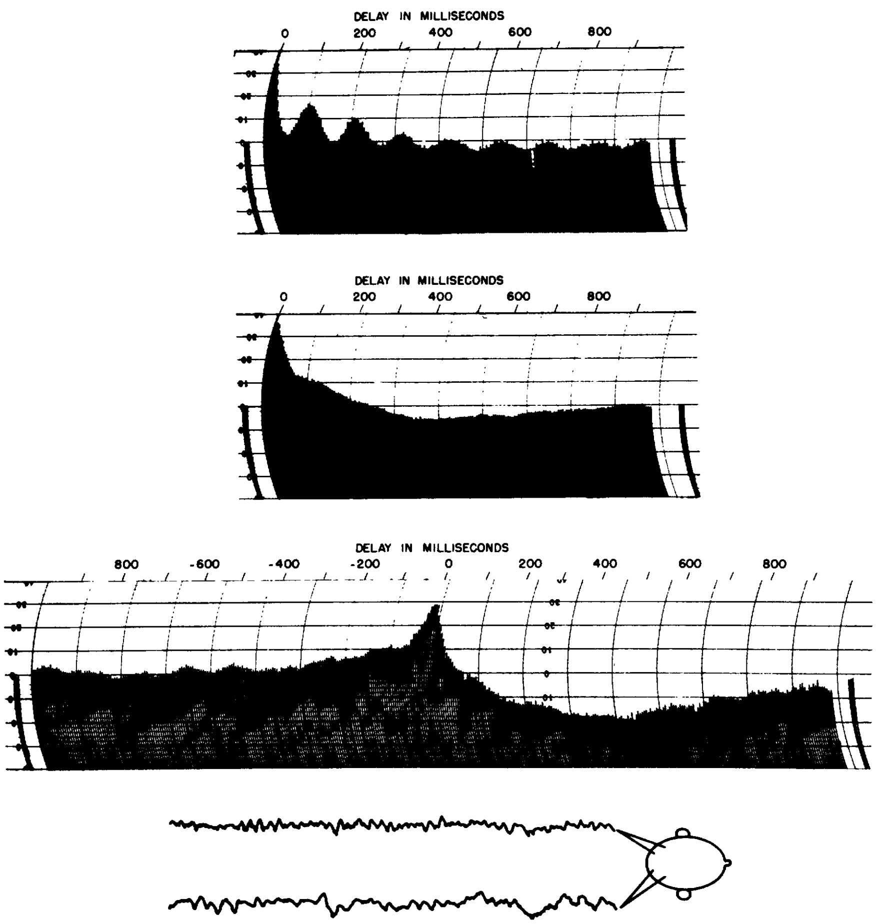

Fig. 3.10. Correlograms computed from the EEG of a subject with a tumor of the right cerebral hemisphere (P. C.): (top) autocorrelogram of EEG from the left parietal-occipital area, (second from top) autocorrelogram of EEG from the right parietal-occipital area, (third from top) crosscorrelogram of data from left and right parietal-occipital areas, and (bottom) inked trace of EEG showing electrode placement. After Barlow and Freeman.[10]

lation of records from the left and right sides of the head yields the crosscorrelograms shown in the bottom parts of Figs. 3.9 and 3.10 respectively.

Thus far the interpretation of the results of the computation of the correlograms has involved no assumptions about the nature of the data. The computation has been viewed merely as a method for abstracting some characteristics of the EEG into a convenient display that emphasizes these characteristics. However, a series of questions remain: What kinds of experimental questions can this approach answer, how does it answer them, and what are the implicit assumptions and limitations of this technique of data analysis? But before we can deal with these questions, we must define what we mean by an experiment and by a class of experiments. We shall consider an experiment to consist of obtaining and processing data by a particular technique for the purpose of testing some hypothesis. The experiments discussed so far in this chapter consist in recording EEG data and in either measuring burst activity or in computing correlograms from these data.

In order to draw valid conclusions about a particular hypothesis, one must repeat these experiments. One can argue from theoretical considerations that no physical experiment so defined can ever be repeated exactly. The best that one can do, under these circumstances, is to control certain of the known variables, thereby forming a class of experiments with respect to these controlled variables. Changing any of these variables and conducting a new set of experiments define a new class.

The concept of a class of experiments is not useful unless one can make certain assumptions about the similarities of the experiments within a given class. In particular, one requires that any class of experiments be represented by a reasonable number of experiments. For example, we compute the autocorrelogram of a 3-minute EEG record taken at 9 A.M. on Monday morning from subject A. We then repeat this on 30 different Monday mornings, keeping as many of the conditions as uniform as possible. The requirement of generalization within a class requires that the 30 experiments we have done will predict something about the outcome of the 31st experiment, at least to the extent of giving us a probability statement about the outcome. Suppose now that the hypothesis being tested was, subject A has alpha activity. Suppose further that an objective criterion was set up to decide this question on the basis of the autocorrelogram of his EEG record. Then if in 29 of the 30 experiments subject A was judged to have alpha activity, we could predict (by standard statistical inference tests) that with some probability p, he would exhibit this in the next experiment.

Thus the empirical approach to correlation discussed so far relies completely on the manyfold repetition of a particular experiment with each change of a parameter of the experiment. In other words, this approach does not generalize from one class of experiments to another. If we were interested in the outcome of the experiment on subject A on a Monday morning for which 15 minutes of data instead of 3 were correlated, it would be risky to use the results of the first 30 experiments to predict the new results. A new set of experiments with 15 minutes of data would have to be run. This procedure of rerunning a set of experiments each time an experimentally uninteresting parameter (one which may have nothing to do with the hypothesis under test) is changed is tedious, and it is clear that an alternative to this approach is desirable.

The use of mathematical models often enables one to extend the results of one class of experiments to describe those of a new class. The more general the model, the more classes of experiments it will be able to describe on the basis of the results of one class. A successful model, as an added bonus, may suggest new experiments.

At the present time, the models for physiological data are still quite simple and not very comprehensive in terms of the number of classes of experiments that are encompassed by them, but they are of some use when such simple parameters as sample length of data are changed. When more is known about neurophysiology, models may be constructed of a sufficiently general character so as to include such complicated experimental changes as the change of "state" of a subject or even the use of different subjects.

The theory of random processes appears to be one mathematical model that is suitable for describing certain physiological data. If one is willing to accept this mathematical model as a description of the data, then the correlation function may yield a considerable amount of information about the parameters of the model. Appendix A deals at length with the random-process model and with some of the assumptions implicit in it. Under certain conditions, the <u>correlation function</u> is defined there as

$$\phi(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, y(t+\tau)\, dt. \qquad (3.1)$$

In this expression, τ is the delay parameter, and $x(t)$ and $y(t)$ are the two time functions to be correlated. When $x(t)$ and $y(t)$ are two different functions of time, then Eq. 3.1 defines the <u>cross</u>-correlation function. When $x(t)$ and $y(t)$ are the same time functions ($x(t) = y(t)$) then the resultant function is the <u>auto</u>correlation function. Note that the correlation function is defined for an infinite sample length of data, and, therefore, it cannot be computed from physical data. One can, however, compute the <u>correlogram</u>, defined as

$$\phi_T(\tau) = \frac{1}{T} \int_{0}^{T} x(t)\, y(t+\tau)\, dt. \qquad (3.2)$$

Note that this function depends upon the sample length T . The relationship between the correlation function and the correlogram is discussed in Appendix B, where it is shown that the correlogram can be considered to be a statistical estimate of the correlation function. This estimate becomes better and better as the sample length is increased.

Appendix A gives some of the background necessary for a thorough understanding of the autocorrelation function. Some of the more elementary properties of this function, however, will be illustrated here by means of a few examples.

Figure 3.11 shows that the autocorrelation function of a sinusoid of arbitrary phase is a cosinusoid. In general, periodic signals have autocorrelation functions which have the same period. Figure 3.12 shows the autocorrelation function of a wide-band noise passed through a low-pass filter. In this case the rate of decay (time constant) of the autocorrelation function is inversely proportional to the bandwidth of the filtered noise. Figure 3.13 shows the autocorrelation function of noise passed through a narrow-band filter. In this case, the frequency of the oscillations in the autocorrelation function is approximately equal to the central frequency of the narrow-band filter, and the decay (time constant) is again inversely proportional to the bandwidth of the filter.

Figure 3.14 shows the autocorrelation function of a mixture of a sine wave and random noise, and illustrates that the autocorrelation function of a sum of independent signals is the sum of their autocorrelation functions.

The crosscorrelation function, which results from the correlation of two different signals, emphasizes the frequency components that are common to two signals. Figure 3.15 shows the crosscorrelation function of two sinusoids of the same frequency but different phase. The crosscorrelation function has the same frequency as the sinusoids and shows the phase difference between the two sinusoids. It can further be shown that sinusoids of different frequency have zero crosscorrelation. Thus it is suggested that correlation functions deal basically with the frequency content of a signal. Briefly stated, one can say that the autocorrelation function results from putting all the frequency components of a time series into cosine phase and assigning to each component a value equal to its mean-square value.

Figures 3.10 through 3.15 have illustrated the correlation functions of some known signals. These functions have been computed mathematically and do not come from any physical data. Now consider Fig. 3.16, which shows the autocorrelogram of the filtered output of a noise source. A comparison between this display and Fig. 3.13 reveals the differences between a machine-computed autocorrelogram and the autocorrelation function predicted by a random-process model of the noise. The differences can be

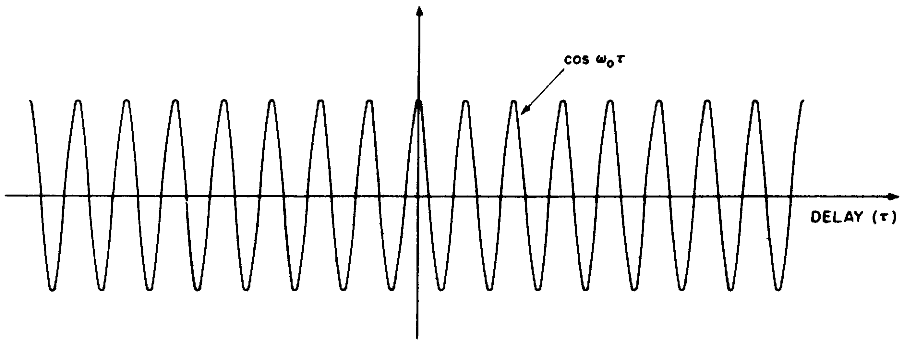

Fig. 3.11. Autocorrelation function of a sinusoid of arbitrary phase.

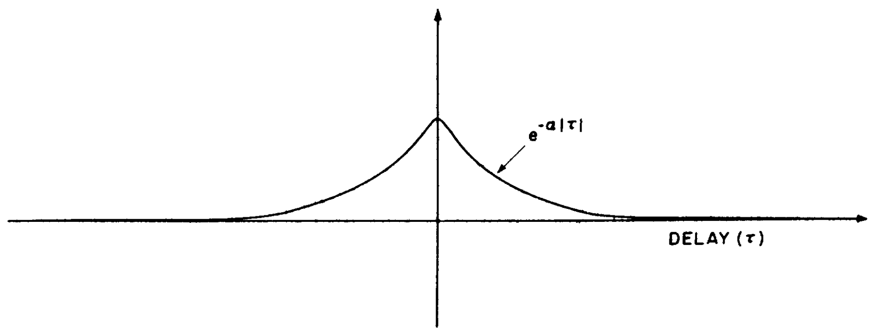

Fig. 3.12. Autocorrelation function of noise with a low-pass frequency spectrum.

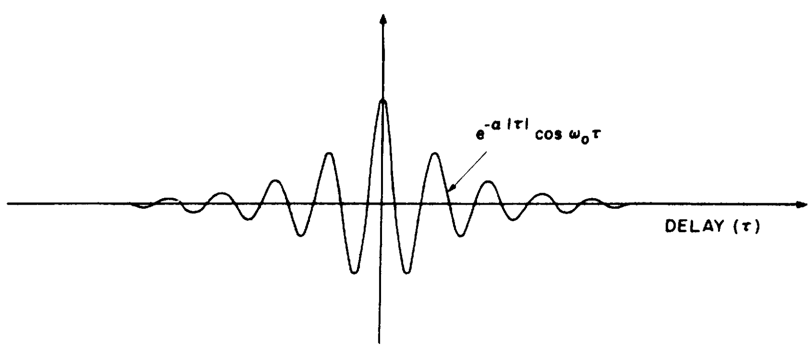

Fig. 3.13. Autocorrelation function of noise with narrow-band (quadratic) frequency spectrum.

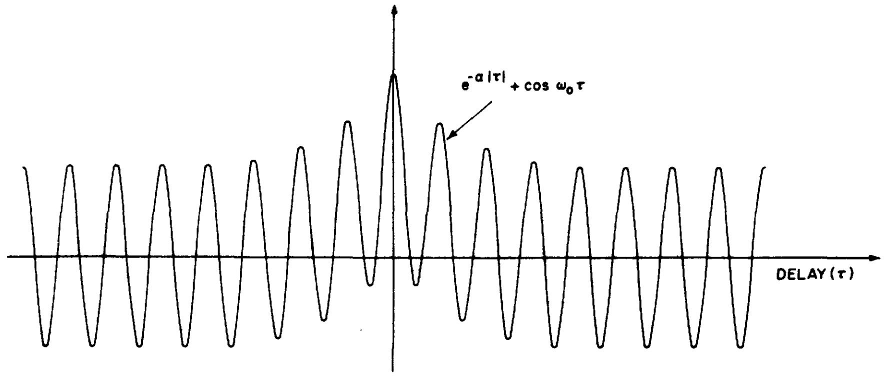

Fig. 3.14. Autocorrelation function of a sum of a sinusoid of arbitrary phase plus noise with a low-pass frequency spectrum.

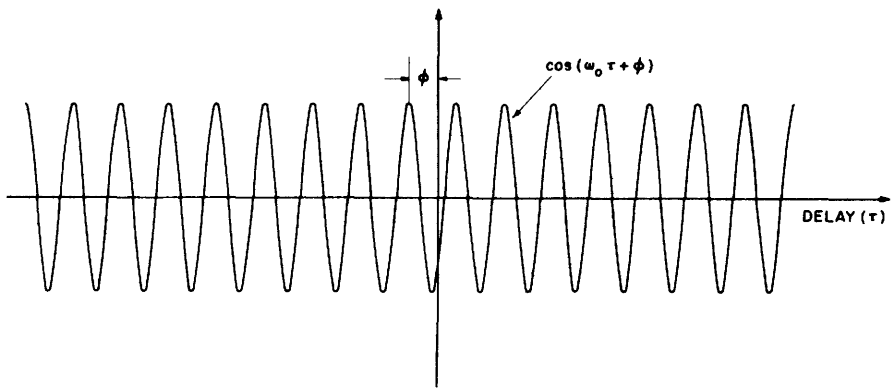

Fig. 3.15. Cross-correlation function of two sinusoids of identical frequencies, with phase difference ϕ.

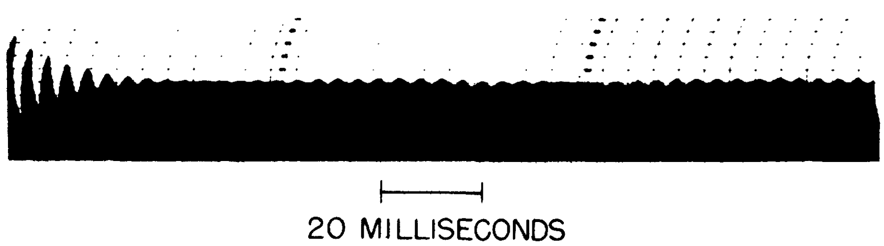

Fig. 3.16. Autocorrelogram of narrow-band noise, that is, wide-band noise filtered by a narrow-band quadratic filter (central frequency 237 cps, bandwidth 13.2 cps). Sample length (T) of data is 7.5 seconds. Delay increment ($\Delta \tau$) is 1/4 millisecond, and the maximum value of delay is 185 milliseconds.

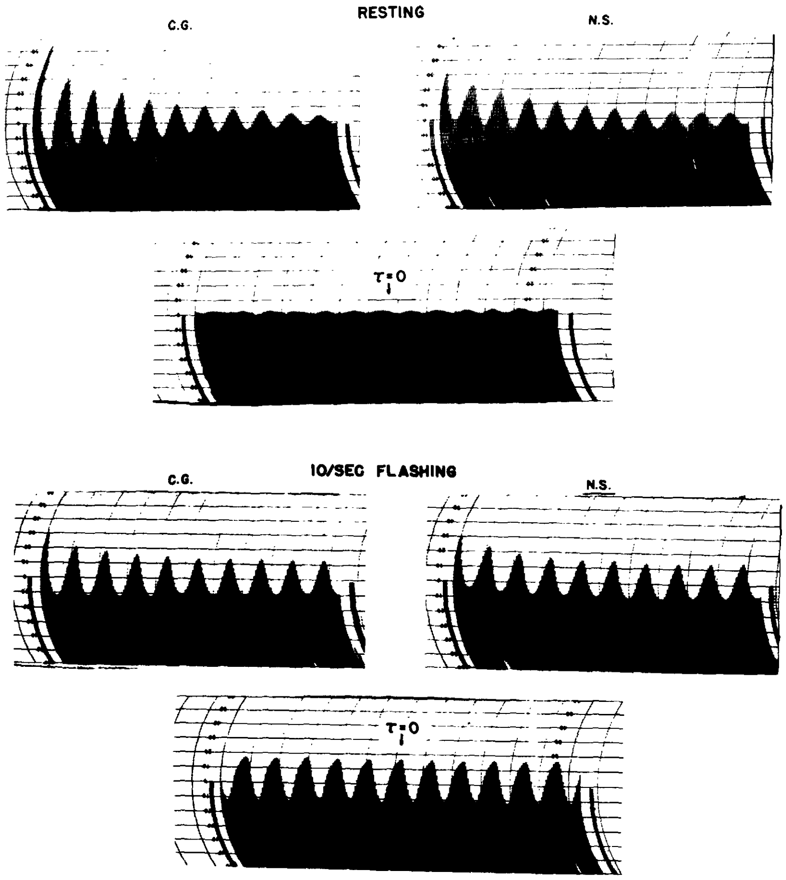

Fig. 3.17. Autocorrelograms and crosscorrelograms computed from simultaneous recordings from two normal subjects at rest and during 10/sec photic stimulation. The electrodes are located in the right parietal-occipital area. The correlograms are computed from one-minute of data. After Barlow and Freeman.[9]

ascribed to the finite resolution of the equipment and the errors of estimation (dealt with in Appendix B).

The investigation of known signals (signals for which mathematical models have been known to fit well in the past) has yielded some information about the properties of correlation functions

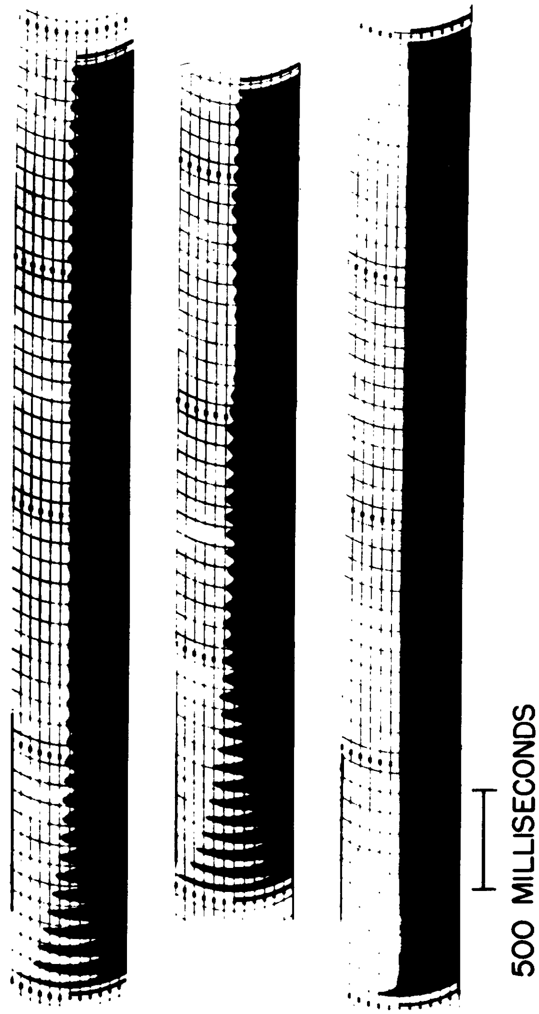

Fig. 3.18. Autocorrelograms of the EEG of three resting subjects. Electrodes are located in the left parietal-occipital area. The sample length of data is 3 minutes. The delay increment is 6.25 milliseconds, and the maximum value of delay is 4.6 seconds. After Barlow and Freeman.[9]

and the correlograms that estimate them. Knowledge of these signals allows one to make some statements about the results of correlating EEG data. For instance, consider Fig. 3.17. The top part of the figure shows the autocorrelograms of two subjects whose resting EEG's (from parietal-occipital area) show a prominent alpha rhythm with periodicities that are approximately equal. Cross-correlating these two records yields a crosscorrelogram that is essentially zero. If repeatable, this result implies that the two EEG's are incoherent - they have no common spectral components. In the bottom part of the figure are shown the autocorrelograms for the same two subjects when they are being flashed by a light at a rate of 10 flashes per second. Note first of all, that the autocorrelograms now exhibit a marked periodic component at 10 cps. Cross-correlating these two yields a periodic component at 10 cps, thus indicating that there are common spectral components in the two EEG signals. These results indicate that the rhythmic stimulation is "driving" the alpha rhythm of both subjects at the stimulus frequency.

One can go even a little further, by applying the random-process model to the EEG. For instance, consider the behavior of the autocorrelograms of EEG for large values of delay. Figure 3.18 shows three autocorrelograms computed out to delays of 4.6 seconds. All three of these records were computed from EEG data that had been recorded from the parieto-occipital area of the scalps of normal subjects who had their eyes closed and were seated in a dark anechoic chamber. The top two autocorrelograms are typical of the autocorrelograms of many subjects. Note the large cyclic component that these two autocorrelograms exhibit at large delays. The presence of such a cyclic component in the autocorrelation function of a signal might lead one to expect a narrow peak in the spectrum of the signal. The following discussion (in fine print) is an effort to show that a particular random-process model, narrow-band noise, predicts that estimates of the correlation function computed from finite lengths of data may contain some long-delay cyclic activity. * One can conclude from the model that the mere existence of long-delay cyclic activity in the autocorrelogram of EEG does not necessarily imply the existence of a narrow peak in the spectrum of the signal.

Now let us see exactly how the random-process model might help us to investigate the question of whether the cyclic component of the correlograms of some EEG's necessarily indicates that there is a narrow peak in the spectrum of the signal. As a first approach, let us construct a model for the EEG time series that will resemble it in some sense. In particular, let us choose a model in which correlograms are very similar to those of the EEG and study this model to see what the correlograms tell us about the correlation function. Consider the correlograms of narrow-band, Gaussian noise, as shown in Fig. 3.19. This signal was constructed by passing wide-band

*See T. F. Weiss[11] for full development.

noise through an RLC filter, and it can be seen that the correlograms at first glance look rather similar to those of subjects with a prominent alpha rhythm in their EEG (see Fig. 3.18).

To investigate the behavior of the finite-time-sample autocorrelation function (autocorrelogram) of narrow-band noise, one can start from the results of Appendix B. There it was shown that $\phi_T(\tau)$ (defined in Eq. 3.2) is itself a random variable, dependent upon the random variable x_t. It was shown that $\phi_T(\tau)$ is an unbiased and consistent estimate of $\phi(\tau)$ (the autocorrelation function of x_t). $\phi(\tau)$ for the narrow-band noise process can be shown to equal

$$\phi(\tau) = \exp(-a|\tau|) \cos \omega_0 \tau \tag{3.3}$$

where $2a$ is the bandwidth and ω_0 the central frequency of the noise. By starting with Eq. B.14 in Appendix B and by assuming a zero-mean Gaussian distribution for x_t, the variance of $\phi_T(\tau)$ around its mean $\phi(\tau)$ can be shown to approximate

$$\sigma_T^2(\tau) \approx \frac{1}{2aT} \tag{3.4}$$

for values of T and τ that are large compared with $1/a$. This result suggests that at large values of delay the variance of the random variable $\phi_T(\tau)$ or the error caused by the finite sampling length remains constant even though the mean of $\phi_T(\tau)$ is decreasing exponentially as τ increases. This result is shown plotted in Fig. 3.20 for a particular value of a and for a normalized time scale. In this figure the mean of $\phi_T(\tau)$ is plotted along with a schematic representation of the $3\sigma_T(\tau)$ confidence limits of $\phi_T(\tau)$. Note that the relative errors of finite sampling (ratio of $\sigma_T(\tau)$ to $E[\phi_T(\tau)]$) get larger for large τ.

In addition to giving us some estimate of the errors due to finite sampling, the random process model can also predict the form of these errors for the narrow-band Gaussian noise signal. This can be done by computing the crosscorrelation of successive samples of the correlogram for values of delay that are large compared to $1/a$. Thus consider the function

$$\mu = E[\phi_T(\tau)\, \phi_T(\tau + \tau')]. \tag{3.5}$$

Under the assumption that $\tau \gg 1/a$ such that

$$E[x_t x_{t+\tau}] \approx [E(x_t)]^2, \tag{3.6}$$

the result of this computation for the noise signal gives

$$\mu \approx \frac{1}{2aT} \exp(-a\tau') \quad \cos \omega_0 \tau' \qquad \text{for } \tau' \geq 0. \tag{3.7}$$

Equation 3.7 indicates that the errors of estimation (errors due to finiteness of record length) have the same spectrum as the noise process.

In summary, it has been shown that a physical signal, namely noise passed through an RLC filter, which has correlograms that appear to be similar upon visual inspection to many EEG correlograms, has a very good representation in terms of the random-process model. On the basis of this stationary Gaussian model of the noise, predictions can be made about the behavior of the estimate of the correlation function as a function of sample length and delay. In particular, it has been shown that the long-delay cyclic

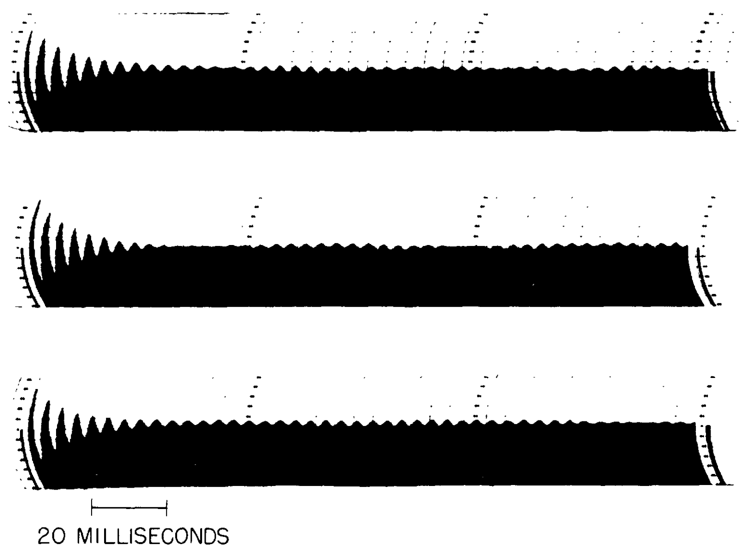

Fig. 3.19. Three autocorrelograms of narrow-band noise: wide-band noise filtered by a narrow-band quadratic filter (central frequency 237 cps, bandwidth 13.2 cps). Sample length (T) of data is 7.5 seconds. Delay increment ($\Delta\tau$) is 1/4 millisecond and the maximum value of delay is 185 milliseconds.

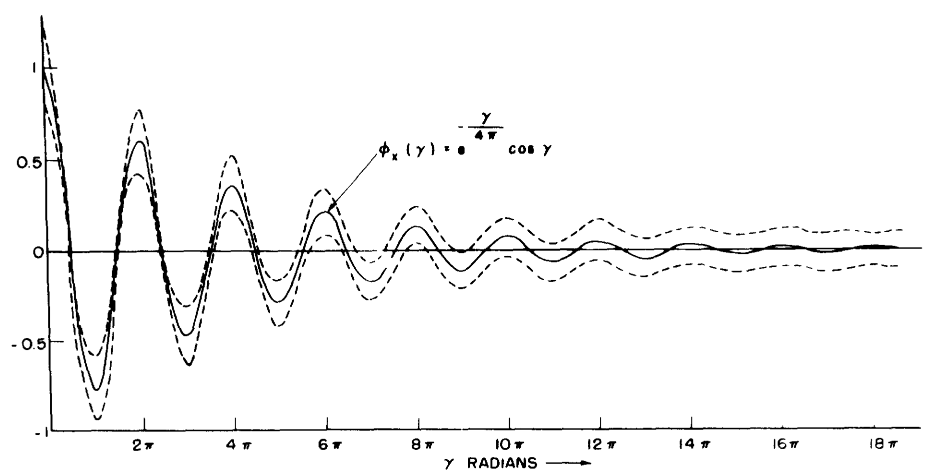

Fig. 3.20. Mean of autocorrelogram plotted on a normalized scale with a schematic representation of the $3\sigma_T(\tau)$ confidence limit (ratio of central frequency to bandwidth is 4π).

activity of these correlograms is a result of the finite sampling time.

The implication of these results for the EEG problem are not quite so clear-cut. One can say, however, that the mere existence of long-delay cyclic activity in an autocorrelogram of EEG does not necessarily imply the existence of a sharp peak in the spectrum of the signal.

Some further work has been done to test this representation of some EEG records by a narrow-band, Gaussian noise model. The amplitude histograms of these EEG records have been studied, and at this time it appears that the results are not inconsistent with the Gaussian hypothesis. The problem of representing EEG records as coming from statistically time-invariant or stationary processes is, however, hazardous. For the particular EEG record shown in Fig. 3.21, the narrow-band, Gaussian noise model appears to be a very good one with respect to prediction of the behavior of the correlogram as a function of sample length. Note the decrease of the long-delay cyclic activity as a function of sample length.

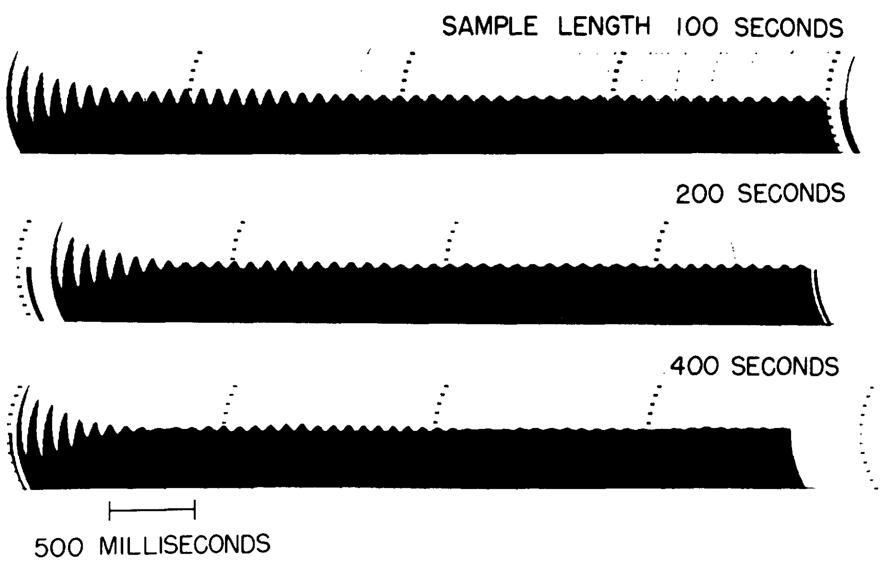

Fig. 3.21. Autocorrelogram of the EEG of a resting subject as a function of sample length. Electrodes located on left parietal-occipital area. Subject seated in a dark, anechoic chamber. Delay increment is 6.25 milliseconds and maximum value of delay is 4.6 seconds.

This example of the use of the random-process model to answer an experimental question has been given in some detail to show the utility of such an approach. If the explicit assumptions of this model are justified, then the utility of this approach can be realized. As a result, a large number of questions about the data can be answered on the basis of a smaller number of experiments. For instance, for the noise signal, the effect of finite time samples on the autocorrelograms is known, and many experiments plotting these effects as a function of all the sample lengths of interest need not be done. In addition, many other functions of this noise signal that may be of interest can be computed mathematically. Tolerance limits of the expected range of variations can also be placed on these functions without doing a large number of experiments.

REFERENCES

1. F. Bremer, "Effets de la déafférentation complète d'une région de l'écorce cérébrale sur son activité électrique spontanée," Compt. rend. soc. biol., 127, 255-358 (1938).

2. Y. W. Lee, T. P. Cheatham, Jr., and J. B. Wiesner, "Application of Correlation Analysis to the Detection of Periodic Signals in Noise," Proc. I. R. E. 38, 1165-1171 (1950).

3. B. G. Farley, L. S. Frishkopf, W. A. Clark, Jr., and J. T. Gilmore, Computer Techniques for the Study of Patterns in the Electroencephalogram, Technical Report 337, Research Laboratory of Electronics, M. I. T., Nov. 6, 1957 (Lincoln Laboratory Technical Report 165).

4. D. B. Lindsley, "Psychological Phenomena and the Electroencephalogram," Electroencephalog. and Clin. Neurophysiol., 4, 443-456 (1952).

5. P. A. Davis, "Technique and Evaluation of the Electroencephalogram," J. Neurophysiol., 4, 92-114 (1941).

6. H. Childers and N. R. Burch, "A Special Purpose Computer for the Period Analysis of the Encephalogram," presented at 11th Annual Conference on Electrical Techniques in Medicine and Biology, Minneapolis, Nov. 19-21, 1958.

7. J. L. Mitchell and K. H. Olsen, "TX-0, A Transistor Computer with a 256 x 256 Memory," Proceedings of the Eastern Joint Computer Conference, December 10-12, 1956, American Institute of Electrical Engineers, New York, pp. 93-101, 1957.

8. M. A. B. Brazier and J. U. Casby, "Crosscorrelation and Autocorrelation Studies of Electroencephalographic Potentials," Electroencephalog. and Clin. Neurophysiol., 4, 201-211 (1952).

9. J. S. Barlow and R. M. Brown, An Analog Correlator System for Brain Potentials, Technical Report 300, Research Laboratory of Electronics, M.I.T., July 14, 1955.

10. J. S. Barlow and M. Z. Freeman, "Comparison of Electroencephalographic Activity Recorded from Different Locations on the Scalp by Means of Autocorrelation and Crosscorrelation Analysis," presented at Eastern Association of Electroencephalographers, St. Margurite, P.Q., Canada, February, 1959.

11. T. F. Weiss, Some Properties of the Finite Time Sample Autocorrelation of the Electroencephalogram, S.M. Thesis, Department of Electrical Engineering, M.I.T., 1959.

Chapter 4

CONCLUDING REMARKS

The preceding chapters and the appendixes that support them constitute a loosely strung chain of discussions and illustrations that relate to certain methods for the data processing of neuroelectric activity.* In Chapter 1 we tried to depict some of the reasons for our attempts to study the nervous system as a communication system. The authors have, however, little doubt that this motivation does not shine through every page of the present monograph.

The illustrative examples of data processing that are given in the preceding pages were not thought up in order to illustrate the capabilities of computers. They derive from studies of specific aspects of the electrical activity of the nervous system (or of parts thereof), in which we tried to deepen our grasp of the observable phenomena by quantifying them in the broad sense of the term. The phenomena that we have dealt with range from electrical potentials recorded from the round window of the inner ear of an anesthetized cat to the electroencephalogram and evoked responses to sensory stimuli that are recorded from the scalp of awake humans.

In most of these problems a more-or-less explicit set of hypotheses had been formulated, and a class of experiments had been designed in order to test the hypotheses. The experimenter then tried to find a method of data processing that would serve these purposes most adequately. An experimenter must not deceive himself regarding the range of data-processing techniques that are available to him or, more importantly, the range of such techniques that he will readily think of. Once one has special-purpose computers in the laboratory or has experience with certain programs on a general-purpose computer, one will tend to think of

*As emphasized in the preface, the present monograph neither attempts to discuss the processing of electrophysiological data in general, nor does it claim a monopoly of insight for the methods described here; they are simply the methods that we have used in recent years.

the available techniques first and preponderantly. The sequence of formulation of hypotheses, design of experiments to test these hypotheses, and search for adequate methods of data processing is, of course, a textbook ideal. The working scientist in the laboratory operates rather less formally and more heuristically. He thinks in terms of what he has learned to do and tries to stretch his limited skills to cover a considerable range of problems. Nevertheless, it is important to emphasize the necessity of looking for problems of intrinsic scientific interest instead of letting the choice be guided by the capacities of available computing devices.

Research on the nervous system demands the mastery of a variety of techniques. Here we have concentrated on techniques for the processing of data from sizable populations of neural units. We are all aware that the quantification of such data remains rudimentary. We must underline that the results that can be obtained by even the most sophisticated methods of data processing are only as good as the scientist's other experimental techniques (such as control of stimuli and of physiological preparations) and the degree of relevant knowledge that he possesses - in the domains of neuroanatomy or of behavior, for instance. In particular, the quality of his results will depend upon how penetrating a question he has really asked. No amount of data processing is a substitute for the selection of significant problems that are formulated sharply enough to justify the use of powerful techniques in order to obtain nontrivial answers. To use electronic data processing for "fishing expeditions" or population studies with many uncontrolled variables will allow one to pile up numbers at a fast rate but will not help necessarily in interpreting the findings.

The computer's speed - important though it is - is not the only asset that the computer has to offer to students of neuroelectric phenomena. Computer programs require that problems be formulated with a definiteness that transcends that commonly found in verbal statements. Such definiteness leads often to a more searching examination of the planned experimentation.

In the absence of theoretical schemes for the operation of the nervous system (and in the presence of a multitude of possible neural mechanisms), there is a premium upon being able to tackle an experimental problem from different angles; a general-purpose computer has the flexibility that permits the quantification of a variety of measures. (Chapters 2 and 3, which dealt with the evaluation of a half-dozen statistical parameters, have by no means exhausted this particular approach.) On the other hand, a special-purpose computer can often perform a set of given calculations more "efficiently," and may thus offer the experimenter an advantage if he can commit himself to one particular measure as being of paramount importance in a large variety of experiments.

Recent evolution in data-processing techniques has made it possible to build computers into experimental setups for "on-line use."

(See, for example, Chapter 2 where the use of ARC-1 as an "on-line" averager is described.) When as complex a system as the nervous system is being investigated, it is to the experimenter's advantage to monitor at least some of his results during the course of the experiment. This use of "on-line" computation can be carried farther by using computers not merely to monitor data processing but to control experimental parameters (such as stimulus patterns) in relation to the organism's response behavior.

As we look to the future a question is often raised: "What are the important obstacles that we might run into when we use computers in the study of the nervous system?" Clearly, there is a gap between the data-processing techniques that are now technically feasible and the techniques that we might want to have at our disposal. The gap involves our desire to do multichannel recording and to have multichannel displays available; the gap exists in the size of currently available memories, in the input and output equipment, and, let us not forget, in the cost of computer facilities. But this technological gap is small compared to the more serious gap between the possible experimental problems that we can adequately handle and the relevant and realistic mathematical models that are available. A discussion of such models deserves an entire monograph, and all that we can do here is to refer to some of the initial steps[1-3] that have been taken in this direction.

We need also to bring about a rapprochement between studies of biological nervous systems and studies in which the behavior of aggregates of neural units is simulated on computers.[4-7] Such studies may provide us with highly useful catalogues of possible mathematical models in areas in which, up to now, not enough collaborative work has been done by students of "dry" and "wet" brains. To convince young researchers from the physical and life sciences of the desirability of acquiring the gamut of necessary skills is one of the purposes of this monograph.

REFERENCES

1. L. S. Frishkopf, A Probability Approach to Certain Neuroelectric Phenomena, Technical Report 307, Research Laboratory of Electronics, M. I. T., March 1, 1956.

2. J. Macy, Jr., A Probability Model for Cortical Responses to Successive Auditory Clicks, Ph. D. Thesis, Department of Mathematics, M. I. T., 1954.

3. M. H. Goldstein, Jr. and N. Y-S. Kiang, "Synchrony of Neural Activity in Electric Responses Evoked by Transient Acoustic Stimuli," J. Acoust. Soc. Am., 30, 107-114 (1958).

4. S. Frankel, "On the Design of Automata and the Interpretation of Cerebral Behavior," Psychometrika, 20, 149-162 (1955).

5. B. G. Farley and W. A. Clark, "Simulation of Self-Organizing Systems by Digital Computer," Trans. IRE, PGIT-4, 76-84 (1954).

6. W. A. Clark and B. G. Farley, "Generalization of Pattern Recognition in a Self-Organizing System," Proceedings of the Western Joint Computer Conference, Institute of Radio Engineers, New York, p. 86-91, 1955.

7. N. Rochester, J. H. Holland, L. H. Haibt, and W. L. Duda, "Tests on a Cell Assembly Theory of the Action of the Brain Using a Large Digital Computer," IRE Transactions on Information Theory, 1956 Symposium on Information Theory, Cambridge (Sept. 10-12, 1956).

Appendix A

THE DESCRIPTION OF RANDOM PROCESSES

William M. Siebert

Neurophysiological phenomena, like many other natural phenomena, often appear to be haphazard, disorganized, unpredictable, random, or "noisy." But these are relative words. They imply only a degree of uncertainty or irregularity. We are not able to predict the weather with precision, but still it will probably be warmer in the summer than it is in the winter. Buried in the "noise" accompanying many of these random phenomena there are often attributes that on the average demonstrate considerable regularity and that can be usefully, or even causally, related to external conditions or stimuli, as the average temperature can be related to the earth's orientation with respect to the sun. Thus it is possible sometimes to discover a degree of order in apparent chaos.

But if the resulting descriptions of the average order are to be quantitative, we must begin with a quantitative description of the observed data. Observations of random phenomena, such as an EEG record, can have many mathematical descriptions. Perhaps the most familiar is to consider the data in terms of a specific function of time, that is, for each time instant we assign a value to the function. But there is another means of representation - the random process - which is often more useful, particularly if we are interested in the regularities or averages that are often the only apparently meaningful attributes of the observed data. As contrasted with a time function, a random process is defined not by its values at various instants but by certain probability distributions and averages such as means, moments, spectra, correlation functions, and so forth. The principal purpose of this appendix is to suggest briefly what a random process is and how it can be described.

Our treatment of random processes, although using mathematical notation where necessary, will be largely discursive, aimed at qualitative comprehension rather than manipulative skill. It is an introduction rather than an outline. Moreover we shall have little or nothing to say about the application of the theory, so far as neurophysiological data are concerned, this topic belongs to the main body of this monograph. But before proceeding it is necessary to make in this regard one point that is often overlooked: A random process is a mathematical model, not a physical reality. To say that some set of data can be usefully considered for some purpose as if it came from a certain random process is not by any means the same as saying that the physical process involved is that random process. Indeed the physical process need not actually be "random" at all in any ordinary sense. The problem may merely be so complex in formulation or analysis (for example, the flip of a coin, behavior of gas molecules, height of ocean waves, and so on) as to render the calculation of presumably deterministic effects unfeasible. The theory of random processes bears the same relationship to the "real world" as does any other mathematical theory, such as geometry. Whether an observed shape can be usefully considered as a triangle and, if so, what values to assign to the angles are problems in surveying, not geometry. Similarly whether an observed fluctuating quantity can appropriately be represented as a sample function of a random process, and, if appropriate, what numerical values to assign to the parameters of the random process are not really mathematical questions. And we shall not discuss them here. But it is only fair to point out that both of these questions (which belong properly to the broad domain of statistics) are often very hard to answer, particularly in neurophysiological applications. The final test, of course, is "in the eating"; if a random-process model leads to useful results, that is proof enough of its utility. Questions of "truth" lie outside mathematics.

1. Random Processes and Random Experiments

A time function which has interesting, "noisy," characteristics can be constructed from the results of a sequence of random experiments. For example, suppose that we spin the spinner shown in Fig. A.1a. The result is one of the six integers 1, 2, 3, 4, 5, or 6 (we agree to ignore any spin that lies "on the line"). Suppose

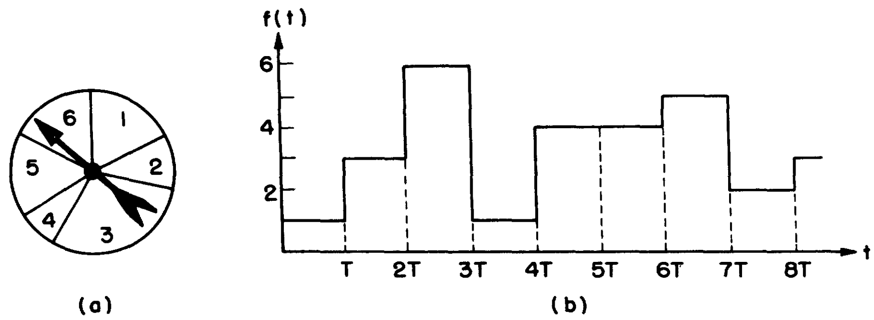

Fig. A.1

further that we construct a pulse, starting at $t = 0$, of duration T, and with an amplitude equal to the number on the spinner (see Fig. A.1b). If, now, we perform a second similar experiment at $t = T$, we can construct, starting at that time, a second pulse of amplitude equal to the result of the second experiment, and so on. The result might be to generate the particular time function shown in Fig. A.1b. But suppose that we were to repeat the entire process. We should not expect to get exactly the same time function, and instead might obtain the time function shown in Fig. A.2. If we were

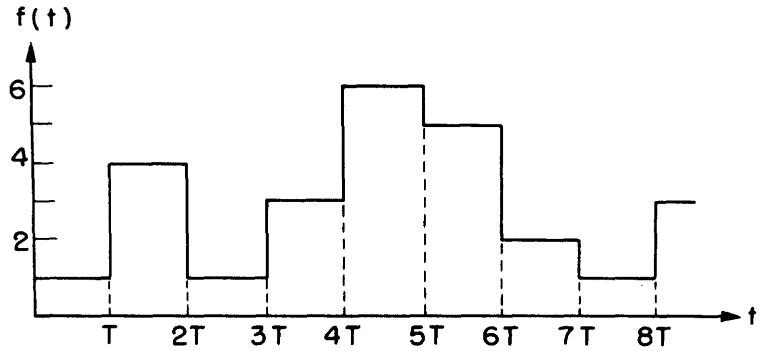

Fig. A.2

to repeat the entire process an indefinitely large number of times, we should obtain a very large number of time functions, some of which might be very much alike, of course. The entire (infinite) collection of waveforms obtained in this way is called an ensemble, and each member of the ensemble is called a sample function. Provisionally, we shall accept the complete ensemble as describing a particular example of a random process.

There are many ways in which the procedure we have described can be generalized. We need not, of course, limit ourselves to spinning a spinner. We could toss a die, or toss ten coins and count the number of heads, or pick a name at random from the phone book, or sample the output of a noise generator such as a diode. Also, we need not make the height of the pulse equal to the number on the spinner but could make it equal the square or logarithm of this number, or equal to the number of letters in the last name of the person selected from the phone book, or in general any real number associated with the particular result of the random experiment which occurs. Such a number is called a random variable. If the random variable can have only a set of discrete values (for example, the random variable corresponding to the spin of the spinner in Fig. A.1a), then it is said to be of discrete type, whereas if it could have any value in some interval (for example, a random variable equal to the voltage output of a noisy diode), then it is a random variable of continuous type. Furthermore, it is clear that the random experiment used to determine the pulse height between 0 and T need not be the same as the random experiment which determines the pulse height between T and 2T, and so forth. For example, a whole set of spinners with different divisions of the circle could be used in sequence, subject to the condition that they shall be used in the same sequence to determine each sample function of the ensemble, or, alternatively, the experiment used to determine the height of the second pulse could be dependent on the result of the first experiment. For example, suppose that a number of different spinners were labeled with the integers 1, 2, 3, 4, 5, 6. Then, if the height of the first pulse for a particular sample function were a 3, the spinner labeled 3 would be spun to determine the height of the second pulse of that same sample function, and so on. Finally, if the random variables representing the results of the various experiments are of continuous type, if T is made very short, and if successive experiments are made heavily dependent on the results of preceding experiments, we can by an appropriate limiting process obtain sample functions that are continuous functions of time.

2. Probability Distributions

The description of a random process in the preceding section is open to several criticisms. In the first place a complete listing of every possible sample function of the ensemble is clearly a ridiculous procedure physically. This difficulty can be circumvented either by specifying the process in terms of a certain set of spinners and instructions for their use, or alternately by abstractly defining the properties of an idealized set of spinners. For example, we might specify each spinner by defining the angular sectors associated with each value of the corresponding random variable, together with the requirement that the result of each spin is equally likely to be any angular position. This latter definition, which we shall explore further in this section, clearly depends on an (intuitive) notion of probability.

But, in the light of our introductory discussion of the distinction between a mathematical theory and the "real world," our provisional definition of a random process is in still more serious trouble. We have defined a mathematical concept in terms of a physical experiment. This difficulty, too, is partially removed by replacing the physical experiment with an abstract experiment involving idealized spinners. But the success of this artifice is by no means complete because the description of an idealized spinner requires a definition of probability, a notoriously difficult problem in mathematical philosophy. To show how this final trouble can be resolved, we shall first examine our intuitive notion of probability in terms of what is called the relative frequency approach.

Let us return to the random process as described in the preceding sections in terms of a sequence of experiments. The value of each sample in the first interval $0 < t < T$ is determined by spinning a certain spinner separately for each sample function. Suppose that we examine N different sample functions in the interval $0<t<T$

and suppose that the value "1" occurs n_1 times, the value "2" occurs n_2 times, and so forth. The ratio n_1/N is called the relative frequency of occurrence of the event "1," n_2/N is the relative frequency of "2," and so forth. As $N \to \infty$, it seems reasonable that n_1/N should tend to some limiting value, $p(x = 1)$, which could be called the "probability of occurrence of the event '1.'" In a similar way we could define $p(x = 2)$, $p(x = 3)$, $\cdots$, $p(x = 6)$. This approach has the weakness, from the point of view of the mathematician, that the existence of a limit for the relative frequency can never be proved, in the sense that the limit depends on the unknowable future. But let us pass on for the moment; it is a familiar fact that many random experiments do exhibit what is called statistical regularity; that is, the relative frequencies of events appear to be approaching limits as the number of trials is increased.

To generalize, suppose that a discrete-type random variable x, associated with a certain experiment, can have any one of the different values, ξ_1, ξ_2, $\cdots$, ξ_M. If the particular value ξ_k occurs n_k times out of N trials of the experiment, then, assuming that the limit exists,

$$p(x = \xi_k) = [\text{probability that } x = \xi_k] = \lim_{N\to\infty} n_k/N.$$

The set of values of $p(x = \xi_k)$ is called the probability distribution for the random variable x. There are a number of consequences of this "definition":

1. Probability of a certain event (that is, one that is bound to occur) = 1.
2. Probability of a null (or impossible) event = 0.
3. $0 \leq p(x = \xi_k) \leq 1$.
4. Since the different values of a random variable are mutually exclusive (that is, x cannot simultaneously have two different values) $p(x = \xi_i \text{ or } \xi_j \text{ or } \xi_k \cdots) = p(x = \xi_i) + p(x = \xi_j) + p(x = \xi_k) + \cdots$.
5. If the random variable x can take on any one of no more than M different values, then

$$\sum_{k=1}^{M} p(x = \xi_k) = 1 .$$

6. The probability distribution function for a discrete random variable is defined by

$$P(x \leq X) = \sum_{\substack{\text{all } k \\ \text{for which } \xi_k \leq X}} p(x = \xi_k)$$

7. $P(x \leq +\infty) = 1, \quad P(x \leq -\infty) = 0.$

8. If $b \geq a$,

$$P(x \leq b) - P(x \leq a) = P(a < x \leq b) \geq 0 .$$

In words, $P(x \leq X)$ is a nondecreasing function of X. As an example, let a random variable x be defined as the number of "heads" that occur if a biased coin is tossed M times. (A single random experiment then consists of M tossings of the coin.) Suppose that the probability of "heads" occurring on a single toss is p. The various possible values which x can have are

$$\xi_0 = 0, \; \xi_1 = 1, \; \xi_2 = 2, \cdots, \xi_M = M .$$

It can readily be shown that

$$p(x = k) = \frac{M!}{k!(M-k)!} p^k(1-p)^{M-k}, \quad 0 \leq k \leq M$$

which is plotted in Fig. A. 3 for a particular choice of M and p and is known as the Binomial Distribution. The probability distribution function for this random variable is illustrated in Fig. A. 4.

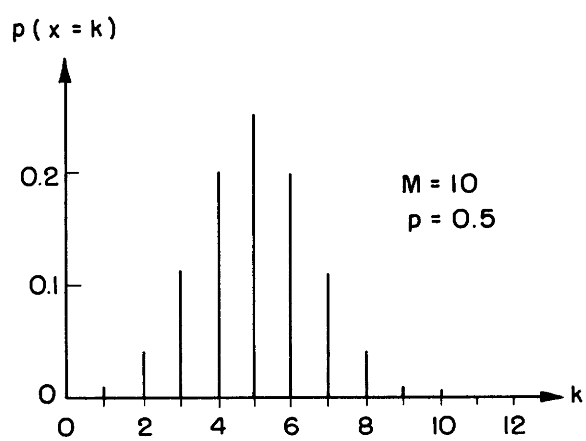

Fig. A. 3 Probability distribution.

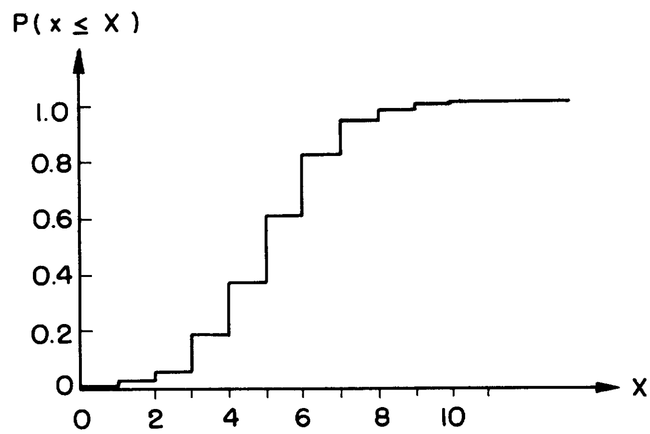

Fig. A.4 Probability distribution function.

In particular it is easy to see that

$$P(x \leq M) = \sum_{k=0}^{M} p(x = k)$$

$$= p^{M} + Mp^{(M-1)}(1 - p) + \frac{M(M - 1)}{2!} p^{(M-2)}(1 - p)^{2}$$

$$+ \cdots + (1 - p)^{M}$$

$$= \left[p + (1 - p)\right]^{M} = 1$$

(by the Binomial Theorem), as it should.

The relative-frequency approach to probability is familiar and intuitive, but as we have suggested it suffers from the difficulties that the existence of the limit can never be proved and, once again, we are guilty of confusing theory and experiment. A way out of this dilemma (suggested by Kolmogorov and now almost universally accepted) is merely to define the probability as a positive real number associated with an event. Axioms (essentially numbers 1 and 4 of our list of "consequences") for the manipulation of probabilities are then specified so that probabilities "behave" like relative frequencies. From the standpoint of the axiomatic theory, then, the specific value of the probability attached to a specific event is of no importance; indeed, the determination of the appropriate value for the probability of some (real) event specifically lies outside the domain of the theory of probability. There is no harm in thinking loosely about probabilities as relative frequencies, provided that we are careful to distinguish between two quite different quantities:

1. The (abstract) probability that some event "has" (for instance, the probability that a tossed coin will come up heads).

2. The estimate of this probability that we might derive from some statistical experiment (for instance, the fraction of heads in 1000 tossings). Whether such an estimate is a "good" estimate, or indeed whether the notion of probability is appropriate to such an experiment at all, are subjective matters for which the theory of probability can at best provide guides. With this understanding we shall assume that our metaphysical foundations are reasonably solid and shall continue for reasons of intuitive simplicity to discuss probability and random-process notions in relative-frequency language.

The concepts of the preceding paragraphs can be extended to continuous-type random variables. Suppose, for example, that a random experiment consists of sampling the output of a diode noise source and that the associated random variable x is the amplitude of the sample. Clearly, for every number X, the event $x \leq X$ occurs some number of times n_X in N samples. Thus we can "define" the probability distribution function $P(x \leq X)$ for a continuous-type random variable as the limit of a relative frequency. As before, $P(x \leq X)$ will be a nondecreasing function of X ranging from 0 for $X = -\infty$ to 1 for $X = +\infty$. If $P(x \leq X)$ is differentiable, we may define the probability density function $p(x)$ as

$$p(X) = \frac{dP(x \leq X)}{dX} \geq 0$$

or

$$p(X)\, dX = P(X < x \leq X + dX).$$

The probability density function for continuous random variables is analogous to the probability distribution for discrete random variables and has similar properties, for example;

$$\int_{-\infty}^{\infty} p(x)\, dx = 1$$

$$\int_{-\infty}^{X} p(x)\, dx = P(x \leq X)$$

$$\int_{a}^{b} p(x)\, dx = P(a < x \leq b) \quad \text{for } b \geq a.$$

By far the most important distribution for continuous random variables is the Gaussian or Normal Distribution

$$p(x) = \frac{1}{\sqrt{2\pi}\ \sigma} \exp[-(x - m)^2 / 2\sigma^2] .$$

The density and distribution functions for this random variable are plotted in Fig. A. 5 and Fig. A. 6, respectively.

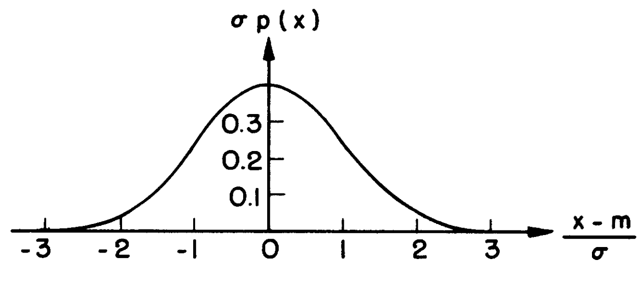

Fig. A. 5 Normal probability density function.

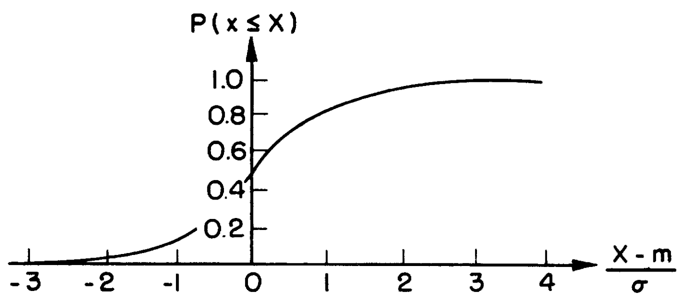

Fig. A. 6 Normal probability distribution function.

It should now be clear that our original random process, which was constructed from independent spins of a spinner for each interval $(n - 1)T < t < nT$, can be neatly specified by the simple procedure of giving the probability distribution $p(x_n = \xi_k)$ that determines the relative frequency of the various possible values of the discrete random variable x_n, which in turn is equal to the height of a sample function in the nth time interval. Of course, if the same spinner is used for each time interval, $p(x_n = \xi_k)$ will not depend on n , and only a single probability distribution $p(x = \xi_k)$ is required.

But, as we have already pointed out, it is not necessary that successive random experiments be independent. For example, the particular spinner to be used in determining the value of a sample function in the second time interval could depend on the value of the same sample function in the first interval. In this event, although it will still be possible to "define" $p(x_n = \xi_k)$ from relative-frequency measurements on the members of the ensemble during the nth interval, the set of $p(x_n = \xi_k)$ for each n and k will not completely describe the random process because these distributions say nothing about the interdependence of successive values of a single sample function. To describe this interdependence, we need to introduce the notions of joint and conditional probability distributions.

Consider a pair of time intervals, and let x_1 with particular values ξ_k be the random variable corresponding to the first interval, and x_2 with the same particular values ξ_k be the random variable corresponding to the second interval. Suppose now that we examine N sample functions of the ensemble and we observe that the particular pair of values $x_1 = \xi_3$ and $x_2 = \xi_1$ occur together in the same sample function $n_{12}^{(31)}$ times. Then we shall define the joint probability of $x_1 = \xi_3$ and $x_2 = \xi_1$ as

$$p(x_1 = \xi_3, x_2 = \xi_1) = \lim_{N\to\infty} \frac{n_{12}^{(31)}}{N}$$

or if there are M different possible values of ξ_k, then the set of M^2 numbers $p(x_1 = \xi_i, x_2 = \xi_j)$ is called the joint probability distribution for x_1 and x_2. From the definition it is clear that

1. $0 \le p(x_1 = \xi_i, x_2 = \xi_j) \le 1.$
2. $\sum_{i=1}^{M} \sum_{j=1}^{M} p(x_1 = \xi_i, x_2 = \xi_j) = 1.$

Furthermore, if $p(x_1 = \xi_i)$ is the probability of the ith value of the first random variable considered alone and "defined" as the limit of a relative frequency as before, then

$$p(x_1 = \xi_i) = \sum_{j=1}^{M} p(x_1 = \xi_i, x_2 = \xi_j) \ge p(x_1 = \xi_i, x_2 = \xi_k) \quad \text{for all } k$$

$$p(x_2 = \xi_j) = \sum_{i=1}^{M} p(x_1 = \xi_i, x_2 = \xi_j) \ge p(x_1 = \xi_k, x_2 = \xi_j) \quad \text{for all } k$$

so that from the joint distribution we can obtain the first-order distributions $p(x_1 = \xi_i)$ and $p(x_2 = \xi_j)$. Finally, it should be obvious that by considering more experiments simultaneously we can "define" in an exactly identical manner joint distributions $p(x_1 = \xi_i, x_2 = \xi_j, x_3 = \xi_k, \cdots x_s = \xi_\ell)$ of any order.

The dependence between a pair of random variables can be described in another way. Let us consider the question, "What is the probability that $x_2 = \xi_1$, if we know that the value $x_1 = \xi_3$ has occurred?" This is the conditional probability of $x_2 = \xi_1$ given $x_1 = \xi_3$ written $p(x_2 = \xi_1/x_1 = \xi_3)$. To see how to define $p(x_2 = \xi_1/x_1 = \xi_2)$, let $n_{12}^{(31)}$ equal the number of simultaneous occurrences of $x_1 = \xi_3$ and $x_2 = \xi_1$ in N sample functions from the ensemble. Similarly, let $n_1^{(3)}$ represent the total number of occurrences of the value $x_1 = \xi_3$ in the same N sample functions independent of what value of x_2 occurs. Then

$$\frac{n_{12}^{(31)}}{n_1^{(3)}} = \frac{n_{12}^{(31)}/N}{n_1^{(3)}/N}$$

is the fraction of those times when $x_1 = \xi_3$ that $x_2 = \xi_1$ also and in the limit should be the desired conditional probability. But the ratio $n_{12}^{(31)}/N = p(x_1 = \xi_3, x_2 = \xi_1)$ in the limit and the ratio $n_1^{(3)}/N = p(x_1 = \xi_3)$ in the limit, so that

$$p(x_2 = \xi_1/x_1 = \xi_3) = \frac{p(x_1 = \xi_3, x_2 = \xi_1)}{p(x_1 = \xi_3)}$$

or in general

$$p(x_1 = \xi_i, x_2 = \xi_j) = p(x_1 = \xi_i)p(x_2 = \xi_j/x_1 = \xi_i) = p(x_2 = \xi_j)p(x_1 = \xi_i/x_2 = \xi_j)$$

which is known as the product rule.

In some cases it might be true that

$$p(x_2 = \xi_j/x_1 = \xi_i) = p(x_2 = \xi_j)$$

for all i and j. In this event it is clear that knowledge of x_1 tells us nothing about what to expect of x_2. Then x_1 and x_2 are said to be statistically independent random variables, and the product rule becomes

$$p(x_1 = \xi_i, x_2 = \xi_j) = p(x_1 = \xi_i)\, p(x_2 = \xi_j) .$$

If the experiments which determine the value during each time interval of each sample function of the ensemble are dependent, then it will be necessary in general to specify the complete joint distribution $p(x_1 = \xi_i, x_2 = \xi_j, x_3 = \xi_k \cdots)$ in order to describe the random process. But if the random variables corresponding to different intervals are statistically independent (as we implicitly assumed initially), then the joint distribution factors into a product of first-order distribution, that is, a specification of $p(x_n = \xi_i)$ is sufficient.

The concepts of joint and conditional probabilities can be extended to continuous random variables in a more or less obvious way. Thus, for example,

$$p(X_1, X_2)\, dX_1\, dX_2 = P(X_1 < x_1 \leq X_1 + dX_1,\ X_2 < x_2 \leq X_2 + dX_2)$$

$$P(x_1 \leq X_1,\ x_2 \leq X_2) = \int_{-\infty}^{X_1} \int_{-\infty}^{X_2} p(x_1, x_2)\, dx_1\, dx_2$$

$$p(x_1) = \int_{-\infty}^{\infty} p(x_1, x_2)\, dx_2$$

$$p(x_1, x_2) = p(x_1/x_2)\, p(x_2) = p(x_2/x_1)\, p(x_1).$$

By analogy with the discrete-parameter discrete-valued random process, if the random variable x_t (corresponding to the amplitude of the sample functions at time t) is of continuous type and if each sample function of the ensemble is a continuous function of time, then a complete description of the random process requires a specification of the joint probability density function

$$p(x_{t_1}, x_{t_2}, \cdots, x_{t_n})$$

for every choice of the times $t_1, t_2, \cdots, t_n$ and for every finite n.

3. Averages, Correlation Functions, and Spectra

One of the principal ways in which random-process theory could be useful in neurophysiology would be to provide a language for the description of observed data in terms of the language used to describe a related random process. But, the discussion of the preceding section, necessarily brief, has perhaps been detailed enough to suggest that the complete specification of a general random process in terms of joint probability density functions is a very complex and difficult matter. In fact, for a completely arbitrary random process, it is practically impossible. Fortunately it is often possible to compromise with a complete description in such a way that the description becomes tractable while retaining much of its significance. Specifically we nearly always limit our description of a random process to the specification of a few simple averages (means, correlation functions, spectra, and so forth). It must be emphasized that these averages do not in general completely describe the process. But, by analogy with the first few terms of the power series expansion of a function, the limited description can often be very good for some purposes, and there are cases of great importance (as we shall see in the next section) in which a small number of averages completely specify a random process.

Let us consider a discrete random variable x with M possible values $\xi_1, \xi_2, \cdots, \xi_M$ corresponding perhaps to the possible values of the sample functions of some random process at a particular instant of time. If we examine N sample functions, let us further suppose that the particular value ξ_1 occurs n_1 times, ξ_2 occurs n_2 times, and so on. Then the average height of the sample functions at this instant is clearly

$$\frac{\xi_1 n_1 + \xi_2 n_2 + \cdots + \xi_M n_M}{N}.$$

Passing now to the "limit" as $N \to \infty$, we shall define the statistical average, ensemble average, or expectation of the random variable x as*

$$E[x] = \sum_{k=1}^{M} \xi_k \, p(x = \xi_k).$$

For a continuous random variable the corresponding definition is

$$E[x] = \int_{-\infty}^{\infty} x p(x) \, dx.$$

If, now, we wish the statistical average of some function of x, such as x^2, $\log x$, or, in general, $f(x)$, the appropriate operation is clearly

$$E[f(x)] = \int_{-\infty}^{\infty} f(x) \, p(x) \, dx .$$

(Hereafter, we shall usually write formulas for the continuous case only.) A most important class of functions is the set x^n, $n=1, 2, 3, \cdots$, $E[x^n]$ is called the nth moment of the random variable x. The name, of course, comes from analogy with mechanics, since, if $p(x)$ plotted as in Fig. A.5 is imagined cut out of sheet metal, then $E[x]$

*$E[x]$ should be read "expectation of the random variable x." The bracket notation is used to distinguish $E[x]$ from a function of x, such as $f(x)$, where x is the independent variable.

is the location of the center of gravity, $E[x^2]$ is the moment of inertia about the origin

$$E[x^2] = \int_{-\infty}^{\infty} x^2 \, p(x) \, dx$$

and so forth. Sometimes instead of the moments $E[x^n]$, the central moments $E[(x - E(x))^n]$ are more convenient. Two moments are particularly important and are given special names and symbols:

1. $E[x]$ is called the mean of x, $E[x] = m$.
2. $E\left[\left(x - E[x]\right)^2\right]$ is called the variance of x,

   $E\left[\left(x - E[x]\right)^2\right] = \sigma^2$.

It is easy to show that

$$\sigma^2 = E[x^2] - (E[x])^2 = E[x^2] - m^2$$

which is analogous to the parallel-axis theorem of mechanics. As we shall see shortly, σ^2 is (in electrical engineering terminology) usually equal to the average "a-c" noise power in a sample function of the ensemble and m is the "d-c" value of the sample function, while $E[x^2]$ is the total average power in the sample function.

As an illustration of the computation of the moments of a probability distribution, consider the Binomial Distribution defined earlier:

$$p(x = k) = \frac{M!}{k!\,(M-k)!} \, p^k (1-p)^{M-k}, \quad 0 \le k \le M$$

$$E[x] = \sum_{k=0}^{M} k p(x=k) = \sum_{k=0}^{M} \frac{M!\,k}{k!\,(M-k)!} \, p^k (1-p)^{M-k}$$

$$= Mp \sum_{\ell=0}^{M-1} \frac{(M-1)!}{\ell!\,(M-1-\ell)!} \, p^{\ell} (1-p)^{M-1-\ell}$$

$$= Mp$$

$$E[x^2] = \sum_{k=0}^{M} k^2 p(x=k) = \sum_{k=0}^{M} \frac{M'\,k^2}{k!\,(M-k)!} \, p^k (1-p)^{M-k}$$

$$= \sum_{k=0}^{M} \frac{M!\,[k(k-1) + k]}{k!\,(M-k)!} \, p^k (1-p)^{M-k}$$

$$= M(M-1)\,p^2 + Mp$$

$$\sigma^2 = E[x^2] - (E[x])^2 = Mp(1-p) .$$

The moments are important, of course, because they partially describe the shape of the probability distribution. Thus for a single-peaked or unimodal probability density function (such as that in Fig. A.5), m is approximately equal to the location of the peak and

$$\sqrt{\sigma^2} = \sigma = \underline{\text{standard deviation}}$$

is an approximate measure of the width of the peak. Indeed, it can be shown that if all the moments are finite (and they need not be, even for probability distributions related to well-defined physical processes), then the moments completely determine the corresponding probability distribution.

Analogously, if we consider a pair of random variables, x_1 and x_2, representing a random process at two different instants of time, then the statistical average of any function of x_1 and x_2, such as x_1x_2, $x_1 \exp(x_2)$, or in general $f(x_1, x_2)$, is

$$E[f(x_1, x_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\, p(x_1, x_2)\, dx_1\, dx_2 .$$

Again, powers of x_1 and x_2 are important, and $E\left[(x_1 - m_1)^{\ell} (x_2 - m_2)^n\right]$ is called the ℓ-nth joint central moment of x_1 and x_2. In particular, the case $\ell = n = 1$ is of special importance, and $E[(x_1 - m_1)(x_2 - m_2)]$ is called the covariance. A related quantity, if t_1 and t_2 are any two instants of time, is

$$E[x_{t_1} x_{t_2}] = R(t_1, t_2)$$

which is called the autocorrelation function for the random process.

The covariance is a measure of the interdependence of x_1 and x_2. But like σ^2 and m, the covariance does not provide a complete specification of $p(x_1, x_2)$ and thus does not describe completely the interdependence of x_1 and x_2. For example, if $E[(x_1 - m_1)(x_2 - m_2)] = 0$, then x_1 and x_2 are said to be uncorrelated. But this does not imply that x_1 and x_2 are unrelated. Indeed, suppose that $p(x_1)$ is symmetrical about the origin and that $x_2 = (x_1)^2$. Then x_2 is exactly determined by x_1, yet it can be shown that x_1 and x_2 are uncorrelated; that is, $E[(x_1 - m_1)(x_2 - m_2)] = 0$. On the other hand, if x_1 and x_2 are statistically independent, that is, if $p(x_1, x_2) = p(x_1)p(x_2)$, then they are also uncorrelated, since

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - m_1)(x_2 - m_2)\, p(x_1, x_2)\, dx_1\, dx_2$$

$$= \int_{-\infty}^{\infty} (x_1 - m_1)\, p(x_1)\, dx_1 \int_{-\infty}^{\infty} (x_2 - m_2)\, p(x_2)\, dx_2 = 0 .$$

The statistical averages that we have been discussing are averages at a given time over a large number of sample functions. But it is clear that we could define another sort of average, the time average of a particular sample function,* as

$$<f\big(x(t)\big)> \; = \; \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f\big(x(t)\big)\, dt \; .$$

An important question, then, is under what circumstances will the time averages and the ensemble averages be equal; that is,

$$E[f(x_{t_1})] \; = \; < f\big(x(t)\big) > \; .$$

Since the time average cannot by construction be a function of time, it is obvious that the two averages can be equal only if the ensemble average also does not depend on time. In general it can be shown that the ensemble average will be equal to the time average of almost every sample function, provided that the random process is both stationary and satisfies the ergodic condition. A random process is stationary if the joint probability density

$$p(x_{t_1}, x_{t_2}, x_{t_3}, \cdots, x_{t_m})$$

depends only upon the time differences $t_2 - t_1$, $t_3 - t_2$, and so forth, and not upon the actual time instants; that is, stationarity implies that the ensemble will look just the same if the time origin is changed.

$$p(x_{t_1}, x_{t_2}, \cdots, x_{t_m}) = p(x_{t_1+\tau}, x_{t_2+\tau}, \cdots, x_{t_m+\tau}).$$

The ergodic condition essentially requires that almost every sample function shall be "typical" of the entire group.

Next to the concept of probability itself, no concepts in random-process theory have been the subject of quite so much confusing discussion and misinterpretation as stationarity and ergodicity. Much of this confusion disappears if it is acknowledged that both stationarity and ergodicity are attributes which relate to the abstract mathematical random-process model and only to this model. To ask

*It is often illuminating to think of a random process as a function of two variables, one being the time variable which runs along each sample function, and the other being an index variable which indicates which sample function is being considered. This idea is implicit in our notation; thus x_t is to be thought of as (loosely) a function of the index variable, over the ensemble, at a particular time t, whereas $x(t)$ is to be thought of as a function of time, along a particular sample function.

whether some "real-world" process "is" stationary or ergodic is to ask a virtually meaningless question. This becomes evident when one recognizes that our total experience with a "real-world" process typically corresponds to part of one sample function, whereas stationarity and ergodicity have significance only in relation to the whole ensemble. Putting the matter operationally, given almost any arbitrary "real-world" record of finite length, it is possible to construct random processes that are either stationary or nonstationary, ergodic or nonergodic, as desired, and that will contain the given record as a reasonably likely, typical part of some sample function. Whether the random process corresponding to some observed data should be chosen so as to be stationary or ergodic depends on what is known about the specific situation. Thus a random process representing hourly temperature readings should clearly be nonstationary if the diurnal or seasonal changes are the phenomena of principal interest.

For stationary random processes, we may write the autocorrelation function as

$$E[x_{t_1} x_{t_2}] = R(t_1, t_2) = R(t_2 - t_1) = R(\tau)$$

since the average can depend only on time differences. We may also define a time-average autocorrelation function as

$$\phi(\tau) = \langle x(t)\, x(t+\tau) \rangle = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, x(t+\tau)\, dt.$$

If the process is ergodic, then

$$\phi(\tau) = R(\tau).$$

The autocorrelation function plays an important role in applications, and it is worth considering some of its more important properties. We assume the processes to be both stationary and ergodic:

1. $R(0) = E[x_t^2] = \sigma^2 + m^2$

   $= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x^2(t)\, dt =$ average power in a sample function.

2. $R(0) \geq R(\tau) = R(-\tau)$.

3. For many physical processes, samples taken a long distance apart are uncorrelated, with the result that

   $$\lim_{\tau\to\infty} R(\tau) = m^2.$$

4. The Fourier transform, $S(\omega)$, of $R(\tau)$ is always real and positive. $S(\omega)$ is defined by

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) \exp(-j\omega\tau)\, d\tau$$

and is called the power density spectrum. The justification for this name will become more apparent shortly.

5. From the Fourier inversion formula,

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \exp(j\omega\tau)\, d\omega .$$

In particular,

$$R(0) = \text{total average power} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)\, d\omega$$

which is part of the justification for calling $S(\omega)$ a power density spectrum.

6. Let $H(\omega)$ be the complex frequency response of a linear system (such as, for example, an electrical network). Then, if $S_i(\omega)$ is the power density spectrum of the input (stimulus),

$$S_o(\omega) = |H(\omega)|^2 S_i(\omega)$$

is the power density spectrum of the output (response). In particular, the total average output power is

$$R_o(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_i(\omega)\, |H(\omega)|^2\, d\omega$$

which provides the final justification for calling $S(\omega)$ a power density spectrum. Thus suppose for example that $H(\omega)$ represents a narrow bandpass filter,

$$H(\omega) = \begin{cases} 1 & \text{in band of width } \Delta f \text{ centered on } \omega_0 \\ 0 & \text{elsewhere.} \end{cases}$$

Then

$$R_0(0) = \text{power of input process in band } \Delta f \text{ at } \omega_0$$

$$= \frac{1}{2\pi} \int_{\omega_0 - \pi\Delta f}^{\omega_0 + \pi\Delta f} S_i(\omega)\, d\omega \approx S_i(\omega)\, \Delta f$$

for small Δf.

Thus, operationally, $S_i(\omega)$ has precisely the significance we desire for a power spectral density.

Ensemble averages are important because they are the quantities in terms of which we usually describe (more or less completely) a random process. Time averages of a single sample function are important because they are closely related to quantities we can actually measure. However, the relation between the (infinite) time averages that we have been discussing and the corresponding ensemble averages depends on stationarity and ergodicity. What, then, are we to do in a situation that clearly must be considered to be nonstationary (as, for example, the hourly temperature record mentioned previously)? No general answer can be given, of course, but often it is possible to justify the significance of some special sort of time average. Thus consider a periodically time-varying random process whose joint probability density is not in general independent of shifts in time, but which is independent of shifts that are multiples of a period T. Thus

$$p(x_{t_1}, x_{t_2}, \cdots, x_{t_n}) = p(x_{t_1+kT}, x_{t_2+kT}, \cdots, x_{t_n+kT})$$

for all integer values of k. For such a nonstationary process, we can define a special sort of time average; for example,

$$\lim_{N\to\infty} \frac{1}{2N} \sum_{k=-N}^{N} x(t + kT)$$

which is a function of t but which should be equal generally to the ensemble average

$$E[x_t]$$

for each t .

4. The Gaussian Process

The most important random process is certainly the Gaussian random process. Formally, a Gaussian process is defined by the requirement that the joint probability density function of every order must have a certain form; for example, the second-order joint probability density function for the random variables x_{t_1} and x_{t_2} must have the form

$$p(x_{t_1}, x_{t_2}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ \frac{1}{2}\left[\frac{\left(x_{t_1}-m_1\right)^2}{\sigma_1^2(1-\rho^2)} - \frac{2\left(x_{t_1}-m_1\right)\left(x_{t_2}-m_2\right)}{\sigma_1\sigma_2(1-\rho^2)} + \frac{\left(x_{t_2}-m_2\right)^2}{\sigma_2^2(1-\rho^2)} \right] \right\}$$

for all t_1 and t_2, where

$$m_1 = E\left[x_{t_1}\right],\ m_2 = E\left[x_{t_2}\right],\ \sigma_1^2 = E\left[x_{t_1}^2\right] - m_1^2,$$

$$\sigma_2^2 = E\left[x_{t_2}^2\right] - m_2^2,\quad \rho = \frac{R(t_1, t_2) - m_1 m_2}{\sigma_1 \sigma_2}.$$

From the standpoint of applications, the importance of the Gaussian process depends on two characteristics:

1. A wide variety of physically observed random waveforms can be usefully represented as sample functions selected from Gaussian processes.
2. The Gaussian process has a number of mathematical features that make analysis relatively simple in many situations which otherwise would present great difficulties. The analyst would certainly be justified if he added to his familiar prayer, "Oh, Lord, please keep the world linear," the words, "and Gaussian."

The first of these characteristics is easy to understand in the light of one of the most important theorems of probability, the Central Limit Theorem. With some qualifications, the central limit theorem states that any random process, each of whose sample functions is constructed from the sum of a large number of sample functions selected independently from some other random process, will tend to become a Gaussian random process as the number of sample functions added tends to infinity. This result is essentially independent of the characteristics of the original random process. Since so many physical random effects can be considered as the superposition of innumerable random elementary causes (for example, the hiss of a radio receiver, the sound of rain on the roof, the vibrations induced by a rocket motor), it should hardly be surprising that the Gaussian process model can often be used quite successfully to represent these effects.

Contrary to the general rule of the preceding section, a Gaussian process is completely specified by one simple average: its correlation function or (if the process is stationary) its spectrum. Furthermore, an ensemble whose sample functions are sums of sample functions from Gaussian random processes is, as might be suspected from the central limit theorem, the ensemble of a Gaussian random process. Thus in particular, if the input to a linear filter (that is, a device whose output at any time is a sum of past inputs) is a sample function from a Gaussian random process, then the output random process will also be a Gaussian random process that is completely described by its spectrum determined from the spectrum of the input, as in the preceding section.

Finally, since a Gaussian process is completely described by its correlation function, it must be possible to express any joint moment of any order in terms of the correlation function. For example, if x_{t_1}, x_{t_2}, x_{t_3}, and x_{t_4} are random variables corresponding to the amplitudes of the sample functions of a Gaussian random process at any four times, and if

$$E\left[x_{t_1}\right] = E\left[x_{t_2}\right] = E\left[x_{t_3}\right] = E\left[x_{t_4}\right] = 0$$

then

$$E\left[x_{t_1} x_{t_2} x_{t_3} x_{t_4}\right] = R(t_1, t_2)\, R(t_3, t_4) + R(t_1, t_3)\, R(t_2, t_4) + R(t_1, t_4)\, R(t_2, t_3) \,.$$

This property of a Gaussian process can be exploited to determine the correlation function of the output of a nonlinear device if the input is Gaussian with a known correlation function. For example, let $x(t)$ be a sample function of a stationary Gaussian process with zero average value and correlation function $R_x(\tau)$, and let $y(t)$ be a sample function of the output of a square-law device when $x(t)$ is the input; that is,

$$y(t) = ax^2(t) \,.$$

Then the correlation function of the output (non-Gaussian) random process is

$$\begin{aligned} R_y(\tau) &= E\left[y(t)\, y(t+\tau)\right] = a^2 E\left[x^2(t)\, x^2(t+\tau)\right] \\ &= a^2 R_x^2(0) + 2a^2 R_x^2(\tau) \end{aligned}$$

where use has been made of the preceding general formula.

It is much less easy to justify the Gaussian process as a "realistic" model for neurophysiological phenomena than it is for many other physical processes. But even so, the Gaussian process is pre-eminently important in neurophysiology. It is certainly the simplest continuous process to describe. And, in part because of this simplicity, it is an easy process to manipulate. If we desire to know the result of some complex data-processing operation, such as the squaring operation above, calculations that might otherwise be very difficult can often be easily carried out for the Gaussian case. Pragmatically, such calculations may often lead to useful predictions or suggestions for new experiments, procedures, analyses, and so forth, and thus are justified even though the Gaussian-process assumption may be hard to rationalize ab initio. Indeed, this procedure works so often that a demonstration that a Gaussian-process assumption is not justified in a particular case is a conclusion of some importance.

5. Conclusion and Bibliography

The purpose of this chapter has been to provide an introduction to the notion of a random process as a model for the description of neurophysiological data and for the analysis of various forms of data processing. It is clear that in the space available no more than an introduction to the methods, concepts, and nomenclature could be achieved. In particular the important question of statistical estimates of the random-process parameters has been reserved for Appendix B. For further study, a recent book is recommended:

W. B. Davenport and W. L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Company, New York, 1958.

Briefer treatments are given in:

H. M. James, N. B. Nichols, and R. S. Phillips, The Theory of Servomechanisms, Chap. 6, McGraw-Hill Book Company, New York, 1947,

or W. R. Bennett, "Methods of Solving Noise Problems," Proc. IRE, 44 609-637 (1956).

The first of these contains an excellent general bibliography. A more complete bibliography is

F. L. Stumpers, "A Bibliography of Information Theory (Communication Theory - Cybernetics)," IRE Trans. on Information Theory, PGIT - 2, (Nov. 1953), and IT - 1(2), (Sept. 1955).

For more detailed study, the following three books and papers are classics in the field:

W. Feller, Probability Theory and its Applications, John Wiley and Sons, New York, 1950.

H. Cramer, Mathematical Methods of Statistics, Princeton University Press, Princeton, 1946.

S. O. Rice, "Mathematical Analysis of Random Noise," Bell System Tech. J., 23, 282-332 (1944); 24, 46-156 (1945). Also reprinted in N. Wax, Selected Papers on Noise and Stochastic Processes, Dover Publications, New York, 1954.

Appendix B

MATHEMATICAL STATISTICS

M. H. Goldstein, Jr.
T. F. Weiss

1. Introduction

Statistics is the middleman between the real world of experiments and the abstract world of mathematics. Like most middlemen, statistics works at more than one level. There is the branch of statistics that deals with wheat prices in Cambodia and the number of blue-eyed baby boys born in the Bahamas. This branch of statistics is purely descriptive and has been of great importance to life insurance companies, economists, and so forth. The other level at which statistics operates is the level of mathematical statistics, which is a branch of probability theory. The models of mathematical statistics, like those of probability theory, are abstract.

The raison d'être of mathematical statistics is to make inferences concerning experimental data. Here we meet a paradox. Although we would _like_ to make inferences in terms of the experimental data, we are _obliged_ to make our inferences in terms of probability models. Inferential statements concerning experimental data must take the form: "_If_ we assume the data to have been generated by such and such a class of probabilistic models, _then_ a good value for some parameter of the models would be . . . ; or _if_ we assume that the data were generated by random process A or B, _then_ the probability that model A generated the data is · · · ."

There is an analogous situation in the measurements made in surveying. The angles that the surveyor reads from his transit are experimental data, and he uses a handbook to make inferences regarding the topography of the terrain. In this analogy, the handbook plays the role of mathematical statistics. The tables in the handbook are computed according to the rules of geometry, an abstract mathematical model. Strictly speaking, the surveyor's inferences should be couched in terms of the model ("if I assume this city block is represented by a polygon, then its area is · · ·"). The extent to which the surveyor's inferences apply to the experimental situation depends on the "fit" of the assumed models. Luckily for the surveyor, it is usually easier to choose an appropriate geometrical model for a section of land than it is to choose appropriate random-process models for neurophysiological data.

2. Estimation of Statistical Parameters

Independent Sampling. To explore further the relationship of mathematical statistics and experimentally obtained data, let us consider an example. Suppose that the solid-line waveform in Fig. B.1 represents experimental data (they happen to be the

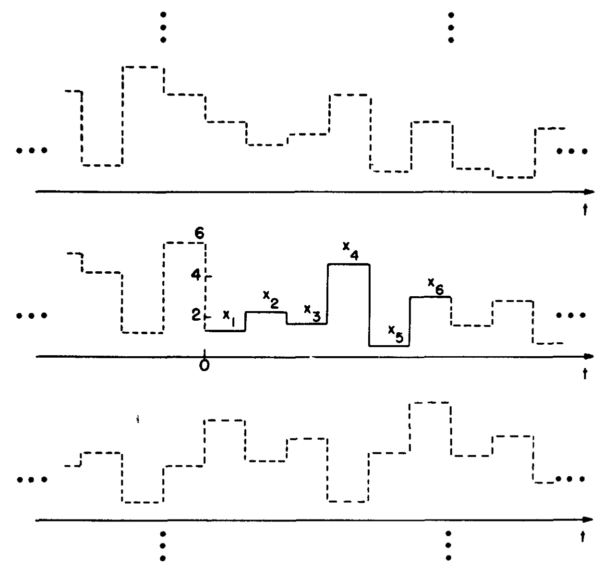

Fig. B.1. Relationship of experimentally obtained data and ensemble of assumed random process. The dots indicate infinite extension of the ensemble and of the member functions.

amount won by a $2 "place" bet on the winning horses in the first six races at Jamaica, April 3, 1959) and that we wish to make inferences about the mean and the distribution of the experimental process which generated these data. What we actually are obliged to do is to assume that the data were generated by some random process and to estimate from the data the mean and the distribution of the assumed process.

If we assume that the data were generated by a random process, the data obtained constitute a finite section of one of the infinite number of infinitely long time functions in the ensemble that describes the random process. This situation is indicated in the figure. The six values of the solid-line waveform $x_1, x_2, \cdots, x_6$ represent six samples from one member of the ensemble generated by the random process. What we desire are functions of these samples which enable us to obtain good estimates of particular parameters of the random process. Such functions are called statistics or estimators. (Hereafter, instead of referring to the six samples in the example, we shall refer more generally to N samples.)

First let us assume that the random process is generated by independent repetitions of a random experiment. Each of the N samples has the same statistical properties; in particular,

$$E[x_k] = m \qquad \sigma^2[x_k] = \sigma^2, \qquad k = 1, 2, \cdots, N. \qquad (B.1)$$

We should like a statistic, a function of the N samples, that will be a good estimate of the mean of the process. An obvious first choice would be the function

$$M_N = \frac{1}{N} \sum_{k=1}^{N} x_k \qquad (B.2)$$

which is sometimes called the sample mean.

Since M_N is a constant times a sum of random variables, it is itself a random variable. In fact, since statistics themselves are in general functions of random variables, statistics are random variables. The "goodness" of a statistic as an estimator depends upon the way in which it approximates the estimated parameter. Since the estimator is a random variable, a "good" estimator could be crudely described as one whose probability density is narrow and well centered on the parameter being estimated. (See Fig. B.2.) The center of gravity of the probability density of the estimator is given by its expectation. In the case under consideration

$$E[M_N] = \frac{1}{N} \sum_{k=1}^{N} E[x_k] = m. \qquad (B.3)$$

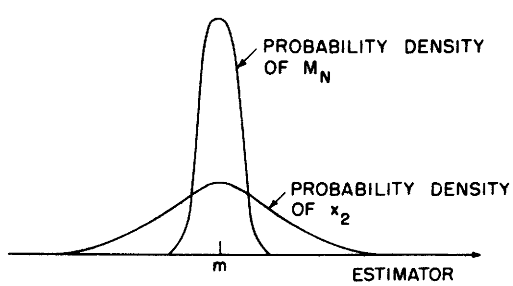

Fig. B.2. Probability densities of two estimators of the mean. M_N is the sample mean, and x_2 is the second of a set of sample observations. Although both estimators are correct on the average, M_N, with the narrower distribution, is the better estimator.

A statistic whose expectation equals the quantity being estimated is said to be an unbiased estimator; M_N is therefore an unbiased estimator of the mean.

M_N is not the only unbiased estimator of the mean. For example, the single sample x_2 is also an unbiased estimator of the mean. However, as illustrated in Fig. B. 2, an unbiased estimator may be a poor estimator if for a given sample size N its probability density function is much broader than those of other estimators. Both of the estimators illustrated will be correct on the average, but it is highly probable that the estimator M_N will be closer to m than is the estimator x_2. An important criterion of "goodness" of an estimator is the spread of its probability density function for a given sample size. A measure of this spread is the variance of the estimator. The variance of M_N is

$$\sigma^2 [M_N] = E[(M_N - E[M_N])^2]$$

$$= E[M_N^2] - E^2[M_N].$$

From Eqs. B. 2 and B. 3,

$$\sigma^2[M_N] = E\left[\left(\frac{1}{N}\sum_{k=1}^{N} x_k\right)^2\right] - m^2$$

$$\left.\begin{aligned}\sigma^2 [M_N] &= \frac{1}{N^2}\sum_{k=1}^{N}\sum_{n=1}^{N} E[x_k x_n] - m^2\\ &= \frac{1}{N^2}\sum_{k=1}^{N}\sum_{n=1}^{N}\left(E[x_k x_n] - m^2\right)\\ &= \frac{\sigma^2}{N}\ .\end{aligned}\right\} \qquad (B.4)$$

The variance of x_2 is σ^2, and, as is evident in Fig. B. 2, M_N is a better estimator of m than is x_2.

Of course, the mean m is an incomplete description of a random process, and we may want to estimate other statistical parameters that yield a more complete description. One set of statistics is the normalized histogram, which estimates the step approximation of the probability density function, as illustrated in Fig. B. 3. The range of the variable x is divided into a number of segments by the points x_α, $x_\beta \cdots$. The samples $x_1, x_2, \cdots, x_N$ are grouped according to the interval in which they fall, and the area under any rectangle in the normalized histogram is set equal to 1/N times the number of samples falling in that interval.

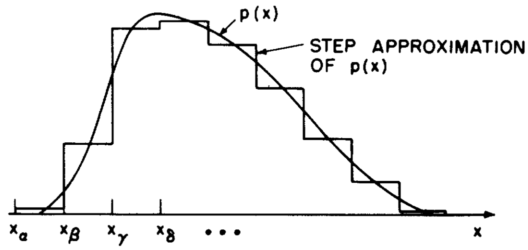

Fig. B.3 Step approximation of probability density function. Area under any interval of step approximation equals area under that interval of p(x).

Consider the interval x_γ to x_δ of the histogram $H_N(x)$. Out of N independent samples the number of samples falling into this interval will be described by the binomial distribution, in which the probability of any sample falling into the interval is

$$p = \int_{x_\gamma}^{x_\delta} p(x)\, dx.$$

Since the mean of the binomial distribution is pN, the expectation of the area of the histogram within the interval x_γ to x_δ is

$$E\left[\left(H_N(x)\right)\left(x_\delta - x_\gamma\right) \right] = \frac{pN}{N} = p, \quad x_\gamma < x \leq x_\delta.$$

Therefore,

$$E[H_N(x)] = \frac{p}{x_\delta - x_\gamma} = \frac{\int_{x_\gamma}^{x_\delta} p(x)\, dx}{x_\delta - x_\gamma}, \quad x_\gamma < x \leq x_\delta. \tag{B.5}$$

This expectation is the value of the step approximation to p(x) between x_γ and x_δ illustrated in Fig. B.3.

The variance of the histogram between x_γ and x_δ, for independent samples, is easily derived from the variance of the binomial distribution $\sigma^2 = Np(1 - p)$:

$$\sigma^2 [H_N(x)] = \frac{p(1 - p)}{N(x_\delta - x_\gamma)^2}, \quad x_\gamma < x \leq x_\delta \tag{B.6}$$

where

$$p = \int_{x_\gamma}^{x_\delta} p(x)\, dx.$$

For intervals in which $1 - p \approx 1$,

$$\sigma^2 [H_N(x)] \approx \frac{p}{N(x_\delta - x_\gamma)^2} = \frac{E[H_N(x)]}{N(x_\delta - x_\gamma)}, \quad x_\gamma < x \leq x_\delta. \tag{B.7}$$

Dependent Sampling. The preceding discussion assumed that the samples $x_1, x_2, \cdots, x_N$ were generated by N independent repetitions of a probabilistic experiment. This assumption is not always justified in models of electrophysiological activity.

Before we consider probabilistic models specifically related to electrophysiological data, let us generalize the preceding results to include sampled random processes for which the samples are not independent. Thus the random-process model for the data of Fig. B.1 would be one in which the values of the sample functions still represent repetitions of random experiments but where the probabilistic parameters in each repetition depend on the previous values of the sample functions. Such a random process was described in Appendix A.

The data of Fig. B.1 could be generated by a random process having statistical dependence between samples. If this model is assumed, the data at hand can be used to estimate parameters of the model. It is important to note that the assumption of a different probabilistic model does not in any way change the experimental data but changes the characteristics of the ensemble in which we assume the data are imbedded. When the model of a process is changed, the assumed statistical characteristics of a given estimator may also change. As an example, we shall now consider the mean and variance of the sample mean computed from dependent samples. The expression for the sample mean remains

$$M_N = \frac{1}{N} \sum_{k=1}^{N} x_k \tag{B.2}$$

If we assume a process in which the mean and variance of each sample are the same

$$E[x_k] = m \qquad \sigma^2[x_k] = \sigma^2, \quad k = 1, 2, \cdots, N \tag{B.1}$$

then it follows that the expectation of M_N,

$$E[M_N] = \frac{1}{N} \sum_{k=1}^{N} E[x_k] = m. \qquad \text{(B.3)}$$

Thus, as in the case of independent samples, M_N is an unbiased estimator of m .

From Eq. B.4, the variance has been shown to be

$$\sigma^2[M_N] = \frac{1}{N^2} \sum_{k=1}^{N} \sum_{n=1}^{N} \left(E[x_k x_n] - m^2 \right). \qquad \text{(B.8)}$$

If the assumed random process is stationary in such a way that $E[x_k x_{k+j}]$ is the same for all k and $E[x_k x_{k+j}] = E[x_k x_{k-j}]$, then B.8 simplifies to

$$\sigma^2[M_N] = \frac{\sigma^2}{N} + \frac{2}{N} \sum_{j=1}^{N-1} \left(1 - \frac{j}{N}\right)\left(E[x_k x_{k+j}] - m^2\right). \qquad \text{(B.9)}$$

For uncorrelated samples (that is, $E[x_k x_{k+j}] = m^2$), Eq. B.9 reduces to

$$\sigma^2[M_N] = \frac{\sigma^2}{N}, \qquad \text{(B.10)}$$

the result for independent samples (see Ref. 3).

An example of estimating a mean from dependent samples is found in estimating the mean of a periodically time-varying process. Such processes, which were discussed in Appendix A (p. 84), are of special interest when we consider responses evoked by periodically presented sensory stimuli. (See Chapter 2.) When it can be assumed that the various physiological factors influencing a neuroelectric potential define a state of equilibrium, then responses to sensory stimuli can be studied by studying changes in the statistical characteristics of the potential. Thus, if the sensory stimulus is presented periodically, an appropriate model for the neuroelectric potential is one with periodically varying statistical characteristics, that is, a periodically time-varying random process.

Consider the upper waveform in Fig. B.4 to be responses to sensory stimuli presented at times illustrated in the lower waveform.

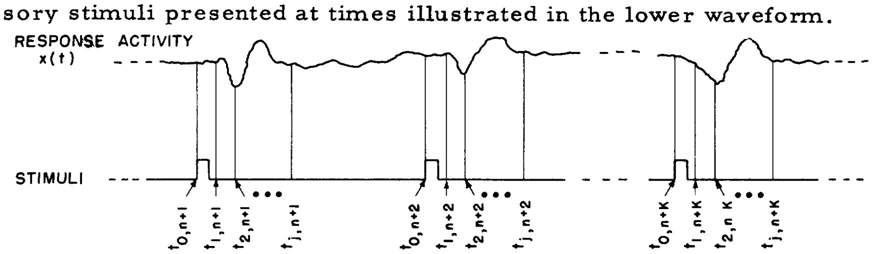

Fig. B.4 "Responses" to periodic stimuli. Upper trace shows simulated neuroelectric potential, lower trace shows stimulus timing markers.

The responses to the n + 1th, n + 2th, and n + kth stimuli in a periodic train are drawn. The mean of a periodically time-varying random-process model of responses, such as these, is generally periodic, and, to estimate the mean at one point, for example t_0, in its period, we might periodically sample the data and average these samples:

$$M_N(t_0) = \frac{1}{N} \sum_{k=1}^{N} x(t_0, n + k).$$

To estimate the mean at other points, the data must be resampled with other delays relative to the instants of stimulus onsets:

$$M_N(t_j) = \frac{1}{N} \sum_{k=1}^{N} x(t_j, n + k).$$

It is by just this process that the "average responses" computed by the devices described in Appendix C estimate the mean of a random-process model of evoked responses.

Continuous Time Averages. We have discussed the problem of estimating the statistical parameters of a process from discrete samples of data. For continuous, ergodic, random-process models we can estimate these parameters by a continuous time average of the data. In such a case, the parameters of the model are related to averages over infinite lengths of time, and we are forced to make our estimates on the basis of finite lengths of data.

Suppose that we have an experimentally obtained time function $x(t)$, $0 \leq t \leq T$. Assume that $x(t)$ is part of one sample function of the ergodic process x_t and that the mean of this process is m and its correlation function $\phi(\tau)$. A reasonable estimator of m might be

$$M_T = \frac{1}{T} \int_0^T x(t)\, dt. \qquad \text{(B.11)}$$

We may consider the integral in Eq. B.11 as the limit of the sample mean obtained by sampling $x(t)$ with more and more closely spaced samples in the interval $0 \leq t \leq T$. In fact, the resulting expressions for the parameters of M_T are limits of the expressions obtained for the sample mean in the case of dependent sampling.* The expectation of M_T is

$$E[M_T] = E\left[\frac{1}{T} \int_0^T x(t)\, dt\right]. \qquad \text{(B.12)}$$

*Reference 3, p. 80-81.

Interchanging the order of averaging and integration gives

$$E[M_T] = \frac{1}{T} \int_0^T E[x_t]\, dt = m. \tag{B.13}$$

Therefore, M_T is an unbiased estimator of m. The variance of M_T is

$$\sigma^2[M_T] = \frac{1}{T^2} \int_0^T \int_0^T \left\{ E[x_{t_1} x_{t_2}] - E^2[x_t] \right\} dt_1\, dt_2$$

$$= \frac{2}{T} \int_0^T \left(1 - \frac{\tau_0}{T}\right) \left\{ E[x_t x_{t-\tau_0}] - E^2[x_t] \right\} d\tau_0 \tag{B.14}$$

where $\tau_0 = t_1 - t_2$.

Note that the variance of M_T approaches zero as $T \to \infty$, and therefore, $M_T \to m$.

We can generalize the results of the estimation of the mean of a random process to the problem of estimating the correlation function. Consider the function $x(t)$ of Eq. B.11 to be defined as

$$x(t) = z(t)\, y(t + \tau).$$

Then M_T becomes a function of τ and can be designated as $\phi_T(\tau)$. Thus

$$\phi_T(\tau) = \frac{1}{T} \int_0^T z(t)\, y(t + \tau)\, dt. \tag{B.15}$$

This function, $\phi_T(\tau)$, can be used to estimate the correlation function and as $T \to \infty$ approaches the correlation function, defined in Appendix A as

$$\phi(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} z(t)\, y(t + \tau)\, dt. \tag{B.16}$$

Note that when $z(t) = y(t)$, the equation defines the autocorrelation function, whereas for $y(t) \neq z(t)$, Eq. B.16 defines the cross-correlation function. We should once more like to know the properties of $\phi_T(\tau)$ as an estimate of the correlation function. Since $\phi_T(\tau)$ is itself a random variable, we should like to know its probability distribution as a function of the two parameters T and τ . Unfortunately, this problem has not been solved, and we must resort to computing the expectation and variance of this function to get some measure of its properties. Thus

$$E[\phi_T(\tau)] = E\left[\frac{1}{T}\int_0^T z(t)\ y(t+\tau)\ dt\right]. \quad (B.17)$$

Interchanging the order of averaging and integration, we obtain

$$E[\phi_T(\tau)] = \frac{1}{T}\int_0^T E[z_t y_{t+\tau}]\ dt \quad (B.18)$$

$$E[\phi_T(\tau)] = E[z_t y_{t+\tau}] = \phi(\tau). \quad (B.19)$$

The computation of the variance of $\phi_T(\tau)$ is equivalent to the results of Eq. B.14 if x_t is replaced by $z_t y_{t+\tau}$. Thus, $\phi_T(\tau)$ is an unbiased and consistent estimate of the correlation function, since the mean of $\phi_T(\tau)$ is $\phi(\tau)$, and the variance of $\phi_T(\tau)$ decreases to zero as T becomes infinite.

3. Concluding Remarks on Mathematical Estimation and Electrophysiology

We have discussed the role of mathematical statistics in the estimation of parameters of probabilistic models on the basis of experimental data. The importance of the assumption that the data were generated by some random process has been stressed. It has also been pointed out that the choice of a specific random-process model does not lie in the realm of mathematics.

In the description of electrophysiological data in terms of random-process models, the research worker encounters a basic difficulty. The applicability of these models depends on the extent to which experimental data exhibit statistical regularity. On the one hand, it is difficult to maintain physiological preparations in a relatively constant "state" for prolonged experimental periods (see Chapter 3, p. 41); on the other hand, it seems tenuous to assume that statistical regularity applies to data recorded from a preparation undergoing significant changes of state. Thus one is often forced to work with short samples and to accept correspondingly high variances in the estimates of the parameters of the models.

One of the primary uses of probabilistic models is to obtain efficient descriptions of experimental data. When we replace the mass of original data by a small number of descriptive statistics, we perform a reduction of the data. This reduction may take the form of an estimate of the parameters of a simple random process. An advantage of this type of data reduction is that it allows prediction of the results of performing other operations on the original mass of data. The extent to which such predictions are correct measures the "fit" of the simple model.

Of course, the usefulness of any particular scheme of data reduction depends upon the relevance of the reduced data to the phenomena of concern. It may be that no simple random-process model will fit the data at hand. In such a case we may hope that simple statistics of some (not simple) random process would be closely related to the phenomena of interest. Even if this hope is not realized, we may view the data as generated by some random process, and the data reduction as an estimate of some statistical parameters of the process, but very little is gained by doing so! However, since research workers have just started to apply statistical methods in electrophysiology, it might prove worth while to consider the simple models first.

REFERENCES

1. A. McF. Mood, Introduction to the Theory of Statistics, McGraw-Hill Book Company, New York, 1950.
2. H. Cramér, Mathematical Methods of Statistics, Princeton University Press, Princeton, 1946.
3. W. B. Davenport and W. L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Company, New York, 1958.
4. W. B. Davenport, Jr., R. A. Johnson, and D. Middleton, "Statistical Errors in Measurements of Random Time Functions," J. Appl. Phys., 4, 377-388 (1952).

Appendix C

DESCRIPTION OF COMPUTERS

R. M. Brown
J. S. Barlow
D. F. O'Brien
W. T. Peake
C. D. Geisler

1. Description of the Correlator[1] Used by the Communications Biophysics Group

The mathematical definition of a correlogram, as given in Chapter 3, is

$$\phi_T(\tau) = \frac{1}{T}\int_0^T x(t)\, y(t+\tau)\, dt.$$

Hence, the computation of correlograms involves the operations of time delay, multiplication, and integration. The particular device presently used by this laboratory for computing correlograms operates in the following manner. First, the two signals* to be correlated, $x(t)$ and $y(t)$, are read into the correlator and shifted in time with respect to one another by τ_0 seconds. Their product is computed and then integrated over T seconds. The result is proportional to the value of the correlogram at one particular value of τ, $\tau = \tau_0$. The correlogram is thus evaluated for a single value of τ, and the process must be repeated for each value of the correlogram which is desired. The usual procedure is to perform the computation for values of τ separated by a constant increment $\Delta\tau$ (that is, $\tau = \tau_0, \tau_0 + \Delta\tau, \tau_0 + 2\Delta\tau, \cdots, \tau_0 + j\Delta\tau, \cdots, \tau_0 + P\Delta\tau$). These results, when plotted, yield a time-sampled correlogram. Because the same signals must be reprocessed for each different value of time delay τ desired, the two signals are permanently recorded and repeatedly read into the computer. The complete correlator system, which performs the various operations already mentioned, is shown schematically in Fig. C.1 and is now described.

*If a crosscorrelogram is desired, two different signals are read into the correlator. If an autocorrelogram is desired, then the same signal is read into both inputs. In either case, the operation of the computer is exactly the same, only the signals operated upon are different.

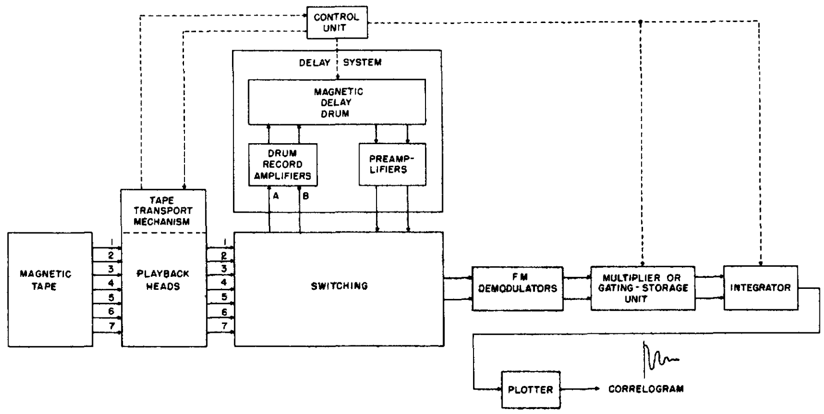

Fig. C.1. Block diagram of correlator system used by the Communication Biophysics Group. Solid lines indicate signal channels, broken lines indicate control functions.

The signals to be correlated are recorded simultaneously onto magnetic tape. Because the capabilities of conventional tape-recording machines are not sufficient to handle the low frequencies often encountered, a frequency-modulation system is used, in which a carrier wave, frequency modulated by the neuroelectric potentials, is the signal actually recorded. The particular tape recorder employed is an Ampex 7-channel FM recorder, and allows the recording of signal frequencies that extend down to zero.

The time delay is accomplished by means of a rotating magnetic drum (Fig. C.2) patterned after one designed by Goff.[2,3] A signal is recorded onto the surface of the drum by means of one head,

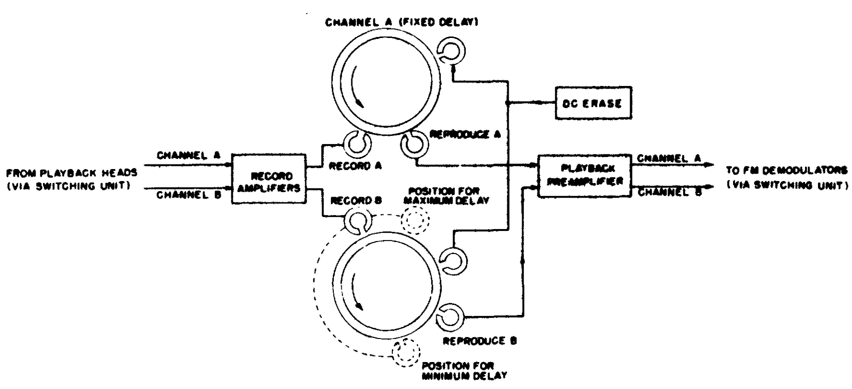

Fig. C.2. Block diagram of magnetic drum delay system. After Goff.[3]

and is recovered by another head, which has some known angular displacement from the recording head. Hence, the signal is delayed by the time required for a point on the surface of the drum to rotate through the angular displacement between the recording head and playback head. The delay may be varied by changing the angular displacement between the two heads. The magnetic drum itself has two channels. The recording and playback heads of one channel are fixed, and the over-all delay in that channel is approximately 15 milliseconds. In the other channel, the playback head is fixed, but the recording head can be moved around a large portion of the circumference of the drum to produce delays of from 15 to 200 milliseconds. Thus the time delay between the outputs of the two channels is variable from 0 to 185 milliseconds. In the variable-delay channel, a stepping solenoid under the control of the computer drives the record head around the circumference of the drum, allowing the computer to change τ in increments ranging from 1/20 to 5 milliseconds.

The two signals from the drum, after demodulation, are fed into a quarter-squaring multiplier. This multiplier takes the sum and difference of both signals, squares both the sum and the difference, and subtracts the two squares. This process produces a result which is proportional to the product, since

$$(x + y)^2 - (x - y)^2 = x^2 + 2xy + y^2 - x^2 + 2xy - y^2 = 4xy.$$

The integrator is a Miller integrator. The design of both the multiplier and integrator is taken from those used on a correlator for speech waveforms that was built at the Imperial College of Science and Technology in London.[4] The output, the value of correlogram for a given delay τ, is plotted on a recording milliammeter (Esterline-Angus).

For each value of τ at which the correlogram is to be computed, the signals to be correlated must be processed by the machine (that is, read in, delayed, multiplied, and integrated). The repeated reintroduction of the data, which this process necessitates, has been performed by reading in a section of tape, rewinding it back to the starting point, reading it in again, and so on. Another, and faster, way is to record the signals onto an endless tape loop. This tape loop is then continually cycled; no rewinding is necessary, and higher tape speeds are attainable. In both of these procedures, the beginning and the end of the tape section on which the data are recorded are marked with pieces of silvered tape. As the silvered tape passes under a carefully placed light source, light is reflected onto a photocell that in turn sends pulses to the control unit, signalling the beginning or end of a section. The control unit then either begins the computation or prepares the correlator for the next read-in of the data.

The tape-recording speeds most frequently used have been from 30 to 0.3 inch per second. The correlator, on the other hand, reads the tape at a speed of 30 or 60 inches per second. Therefore, it is possible to shorten data-processing time by recording the signals at a slower speed and by reading the tapes at 30 or 60 inches per second. Speed-up factors of from 1 to 200 are available by this method. The effective value of time shift obtainable by means of the drum is also multiplied by the speed-up factor. The range of possible delay τ is thus increased up to approximately 35 seconds. Recording at lower speeds, however, decreases the upper frequency limit of the recorded signals. If the recording is made at 30 inches per second, signals in the spectral band 0 to 5000 cps are recorded. At slower speeds, the higher cutoff frequency is 5000 cps divided by the speed-up factor. At 0.3 inch per second, therefore, the upper cutoff frequency is only 50 cps.

2. Description of ERD (Evoked Response Detector)[5]

The correlator system just described, with a few modifications, can be used to compute the average of responses evoked by specific sensory stimuli. Consider the experiment idealized in Fig. C.3.

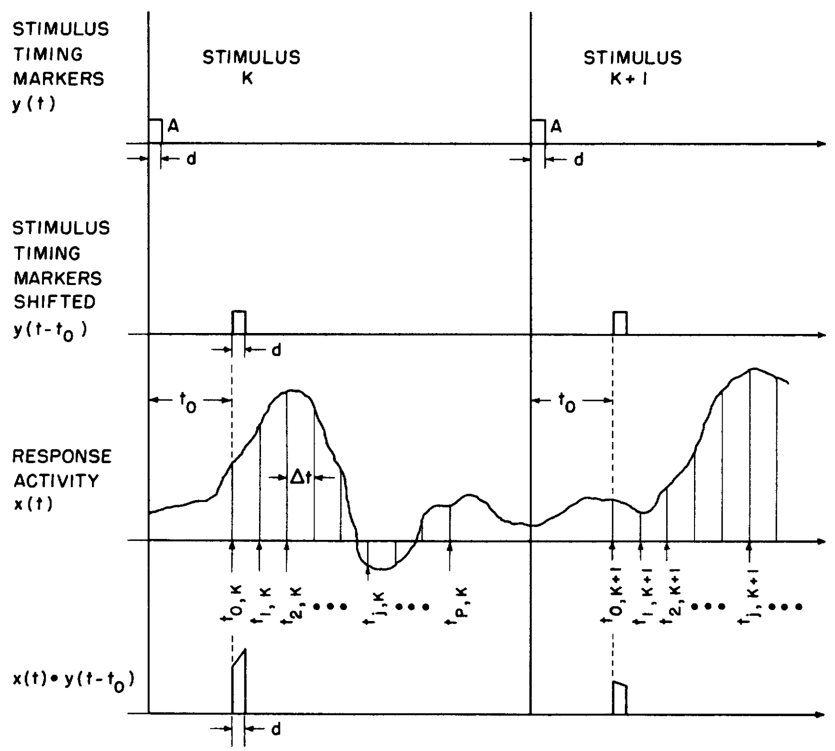

Fig. C.3. Waveforms encountered in the crosscorrelation of "response activity" and stimulus timing markers.

The third trace from the top shows simulated "response activity" following the presentation of two stimuli, k and $k + 1$, that are part of a long train of such stimuli. The onsets of these particular stimuli are marked in the upper trace by small rectangular pulses of constant amplitude.

The crosscorrelation of the two "signals," the stimulus markers and response activity, could be accomplished by shifting the two signals with respect to one another, multiplying them, and then integrating their product over some time T. To illustrate this process, let us calculate the value of the crosscorrelogram for one particular value of time shift, $\tau = -t_0$. First of all, the stimulus markers are shifted by t_0, as shown in the second trace t from the top in Fig. C. 3. Then, $x(t)$ and $y(t - t_0)$ are multiplied, yielding the waveform shown in the bottom trace A of Fig. C. 3. The response activity is multiplied by the height of the pulse A, as long as the timing pulse lasts, and is multiplied by zero for the stretch between pulses. Because the stimulus markers are rectangular pulses of constant width and amplitude, the multiplication of the two waveforms is equivalent to multiplying the response activity by A and then gating a small fixed time interval d of the activity occurring t_0 seconds after each stimulus. If the time interval d is small, compared to variations in the response activity, the height of pulse k in the lower waveform is approximately constant, and proportional to the value of the response activity at time $t_{0,k}$. Likewise, the height (and hence the area) of pulse $k + 1$ is proportional to $x(t_{0,k+1})$, and so forth. Finally, the crosscorrelogram of $x(t)$ and $y(t)$ at $-t_0$ is computed to within a constant by integrating $x(t) \cdot y(t - t_0)$ over some sample time T. This integral is just the area under all the pulses and is therefore proportional (approximately) to the sum of response activity occurring at times t_0 following the presentation of the stimuli. It is clear, then, that

$$\phi_T(-t_0) = \frac{1}{T} \int_0^T x(t)\, y(t - t_0)\, dt \approx \frac{Ad}{T} \sum_{k=1}^{N} x(t_{0,k}).$$

If this process is repeated for $\tau = -(t_0 + \Delta t), -(t_0 + 2\Delta\tau), \cdots, -(t_0 + j\Delta\tau), \cdots, -(t_0 + P\Delta\tau)$, the whole correlogram is computed. According to Eq. 2.1, however, the value of this correlogram is proportional to the average of responses. It follows, therefore, that the average of responses may be computed by crosscorrelation techniques.

The actual operation of ERD, though similar to correlating, involves only part of the correlator. To compute one point of the average, the stimulus marker signal and the response activity are time-shifted by means of the magnetic drum of the correlator.

However, after being shifted, instead of multiplying the response activity, the stimulus markers are used as signals for a gating and storing unit (see Fig. C.1). This unit samples the response activity at the incidence of a stimulus marker, and the corresponding voltage is stored until the incidence of the next stimulus marker, at which time a new sample is taken, its value stored, and so on. The waveform of the stored samples is integrated by means of the integrator in the correlator and recorded by the recording milliammeter. The computer is then reset for the next value of time shift τ, and the process is repeated, and so on. This computation, while mathematically similar to the crosscorrelation of response activity and stimulus markers, avoids the operation of multiplication by substituting the more easily accomplished operations of sampling and storing.

Since the present ERD incorporates much of the correlator system, the capabilities of the two systems are similar. The available time delays, delay increments, speed-up factors, and so on, mentioned in the description of the correlator apply also to computation done by ERD.

Recently, an electronic system, using commercially available components, has been constructed to replace the magnetic drum.[6] Other modifications have also been introduced so that ERD may now be operated independently of the correlator. A magnetic-tape loop is used to read in the data. The design of the sampling and storage circuits has also been altered to provide improved operation over a wider range of operating parameters. In addition, ERD can now compute a measure of the variance at particular times t_j after the onset of stimuli.

3. Description of ARC-1 (Average Response Computer) [7]

The Evoked Response Detector just described requires that the data must be reintroduced into the computer for each value of time t_j at which the average of responses is to be computed. In addition to requiring long computation times, this feature means that the ERD cannot be used during experiments, since under "on-line" conditions the data may be read in only once. The desirability of computing averages of responses during experiments led to the design of ARC.

ARC is a high-speed transistorized special-purpose digital computer using a magnetic-core memory. The computer consists of 256 magnetic-core registers, each having a capacity of 18 binary digits, and of transistorized logic that controls this array. ARC has two principal modes of operation. In one mode it computes an average of responses and in the other, amplitude histograms.

The operation of the computer in the averaging mode is illustrated in Figs. C.4 and C.5. In this mode of operation, the onset of the stimulus triggers the computer. After the preset initial delay t_0 (see Fig. C.4), the computer triggers an analog-to-digital

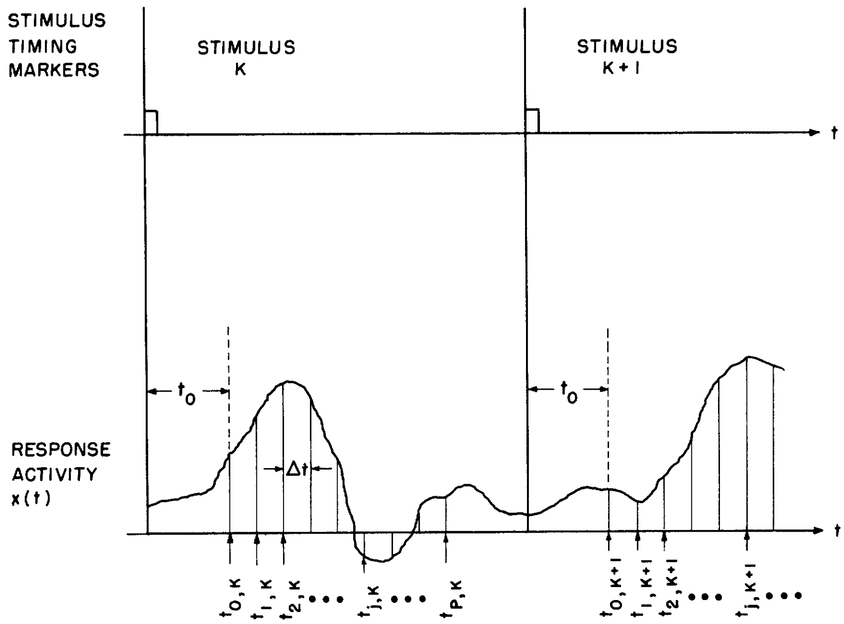

Fig. C.4. "Responses" to successive stimuli. The upper trace shows stimulus markers, the lower trace shows a simulated neuroelectric potential.

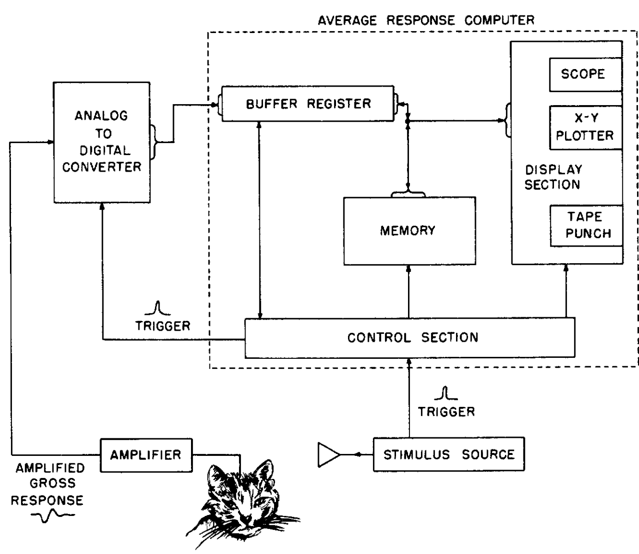

Fig. C.5. Block diagram of Average Response Computer System.

converter (Epsco Datrac). (Since the logic is controlled by an 800 kc clock, this process will be carried out with a maximum error of 1.25 microseconds in its timing.) When triggered, the converter samples the response activity [point $x(t_{0,k})$ of Fig. C.4] and sends to the computer a number that is the value of $x(t_{0,k})$ accurate to one part in 256. The number is coded as a seven-bit binary number and a sign bit. The computer stores the number in a memory register that has been assigned to the first sample following each stimulus. After a time interval Δt elapses, the converter is again triggered by the computer, and a number proportional to the amplitude of the response at that instant, $x(t_{1,k})$, is sent to the computer and stored in the next memory register. When this has been accomplished for all desired sample points, the computer waits until the next stimulus is presented. Then, after the same initial delay t_0, the process repeats. The number representing the voltage $x(t_{0,k+1})$ is added to the number representing the voltage $x(t_{0,k})$, and the sum is stored in the appropriate memory register; the number representing the voltage $x(t_{1,k+1})$ is added to the number representing the voltage $x(t_{1,k})$, and the sum is stored in the next register, and so on.

Thus, at any moment, each data-handling memory register stores the current sum of the responses sampled at some fixed point in time relative to the stimulus onset. These current sums, multiplied by a scale factor, are displayed on a CRO after each stimulus, allowing the experimenter to observe the actual build-up of the sum. The process is repeated until a predetermined number of responses (as many as 262,143) have been summed.

The scale factors of the display are integral powers of 1/2 [up to $(1/2)^{10}$]. Hence if 2^n (where n is an integer between 0 and 10) responses are summed, the final display may be shown as the <u>average</u> of the responses.

Prior to the computing of an average of responses, the number of points P to be sampled in each response (up to 254) is preset, as is the interval between sample points Δt (which may be chosen between 80 microseconds and 2.62 seconds in 10-microsecond increments). This assigns a particular memory register to some fixed point in time after the application of each stimulus. Thus, for example, the same memory register is assigned to points $x(t_{2,1})$, $x(t_{2,2})$, $\cdots$, $x(t_{2,k})$, $\cdots$, $x(t_{2,N})$. The initial delay t_0 (variable from 0 to 2.62 seconds in 10-microsecond increments) and the number of responses to be averaged N are also preset.

The second mode of operation is one in which an amplitude histogram of the responses is compiled. In this mode, each of the data-handling memory registers is assigned the function of counting the number of times that the activity at some specified instant t_j after each stimulus falls into a given amplitude range. As before, the analog-to-digital converter periodically samples the response and sends a number to the computer. However, in this mode of operation, the response is sampled at only one point in time t_j after the application of each stimulus. (This delay may be preset from 0 to 2.62 seconds in 10-microsecond increments.) When the number representing the voltage at that point $x(t_{j,k})$ arrives in the computer, ARC adds 1 to the particular memory register assigned to that numerical value. Upon presentation of the next stimulus, the system samples the response after the same preset delay, and adds 1 to the memory register which has an assigned number equal to the numerical value of the response amplitude $x(t_{j,k+1})$. The computer stops after some preset number of responses are processed this way.

Besides these principal modes of operation, the computer has a test mode designed to serve as check on proper operation of the memory-logic circuitry, the 800-kc clock, the memory and the display devices.

Three methods are available to display the end result of either the average or histogram computations. (1) An oscilloscope (Tektronix) presents the average of responses as a graph of amplitude versus time. The same oscilloscope also displays the completed histogram as a graph of number of responses versus amplitude. (2) The oscilloscope display may be permanently recorded by an X-Y plotter (Moseley Autograph) (see Fig. 2.3 for an example of an average of responses recorded by the plotter). (3) The results of these computations may be punched on paper tape by a motorized tape punch (Commercial Controls). This paper-tape record is useful if the data are to be further processed by general-purpose computers that accept this type of input.

4. Description of ALMIDO (Amplitude and Latency Measuring Instrument with Digital Output).

One particular instrument for measuring characteristic well-defined neural responses, as discussed in Chapter 2, page 17, is ALMIDO. This machine was designed and built by R. L. Koehler,[8] and is an improvement and combination of two earlier devices.[9,10] ALMIDO is designed to measure latencies of peaks and peak-to-peak amplitudes in neural responses. It presents the results of the measurements in digital form. Some results obtained using this device have been presented in Chapter 2.

The general operation of ALMIDO is indicated by the block diagram in Fig. C.6. The measurement is started by a pulse synchronized with the stimulus. A preset interval after this pulse, the gate opens and passes the amplified neural signal for a certain interval. The stimulus pulse also gates a 100-kc oscillator (in counter control) into the counter. The counter registers the number of cycles of the oscillator output occurring between the stimulus pulse and the output pulse from the peak detector. The peak detector delivers an output pulse every time the gated signal has a peak that is larger than any previous peak. The number in the counter at

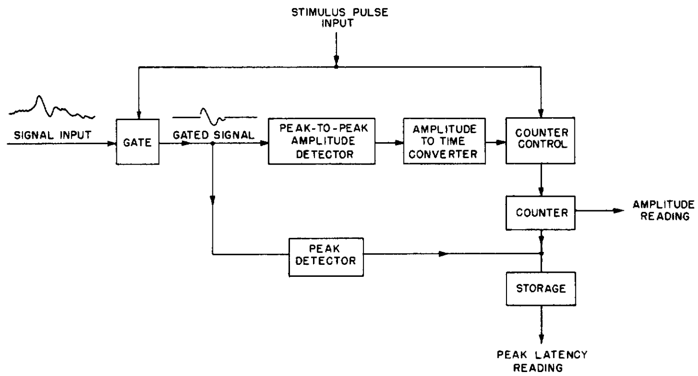

Fig. C.6. Block diagram of ALMIDO

the time of the peak is then transferred to the storage registers without interrupting the count. Hence, the number in the storage register after the gate closes indicates the latency of the largest peak in the gated interval. The counter stops and resets at the end of each gated interval. An optional inverter stage is included in the amplifier so that the peak detector can operate on either positive or negative peaks of the input signal.

The amplitudes of the largest positive and largest negative voltages during the gated interval are stored in the peak-to-peak amplitude detector. When the gated interval is over, this voltage is compared to a linearly rising voltage in an amplitude-to-time converter. A voltage ramp starts from a negative value; when it reaches the level of the negative peak, an amplitude comparator puts out a signal which gates the oscillator into the counter. When the ramp reaches the value of the positive peak, a comparator stops the counter. Hence, a number is held in the counter which is proportional to the peak-to-peak amplitude.

Specifications of ALMIDO

Amplitude measurement.- The amplitude measurement is indicated digitally as a number from 0 to 99. A gain adjustment on the amplifier provides for four different scales 0.5, 1.0, 2.0, and 5 volts full scale.

Latency measurement. - Two ranges are available for latency measurements. For responses of short latency (for example, responses to acoustic clicks recorded from near the round-window of cats) the reading can be made from 0.00 to 9.99 milliseconds in steps of 0.01 millisecond. For responses of longer latency (for example, responses to clicks recorded from cat cortex), the reading can be made from 00.0 to 99.9 milliseconds in steps of 0.1 millisecond. The delay of the onset of the gate and the gate interval are independently adjustable from 0.5 to 30 milliseconds. The gated signal is available for observation on an oscilloscope so that the gate can be adjusted visually to the desired delay and duration.

Output. - The latency and amplitude measurements are indicated by neon bulbs. This indication is held until another stimulus pulse is sensed by the machine. A display timer is included which prevents a new measurement from being made for an interval which is adjustable from 0.02 to 7 seconds. With this feature the machine can be used to sample responses periodically for high stimulus repetition rates.

Output Recorders

Histogram recorder. - This device has a set of twenty-three electromechanical counters (Veeder Root). Twenty of these are connected in such a way that they add 1 to their count every time an amplitude measurement falls into a particular level. These twenty levels are adjacent and can be placed anywhere in the range of one hundred levels into which the machine quantizes amplitude measurements. The other three counters indicate (1) total number of responses, (2) responses larger than the highest of the twenty levels, and (3) responses smaller than the lowest of these levels. The histogram recorder will record at rates up to 20 per second.

Printer recorder. - An automatic printer (Hewlett-Packard) can be connected to ALMIDO. This device prints the amplitude and latency measurements on paper tape. It can be used at rates up to 5 per second. Although the automatic printer is slower than the histogram recorder and does not reduce the data into histograms, it is sometimes desirable to have the data in this form in order to test for correlations between latency and amplitude, or between latencies or amplitudes from two different signals (recorded from different locations in the nervous system).

REFERENCES

1. J. S. Barlow and R. M. Brown, An Analog Correlator System for Brain Potentials, Technical Report 300, Research Laboratory of Electronics, M.I.T., July 14, 1955.
2. K. W. Goff, "The Development of a Variable Time Delay," Proc. IRE, 41, 1578-1584 (1953).
3. K. W. Goff, "An Analog Electronic Correlator for Acoustic Measurements," J. Acoust. Soc. Am., 27, 223-236 (1955).
4. J. N. Holmes and J. M. C. Dukes, "A Speech-Waveform Correlator with Magnetic-Tape Delay and Electronic Multiplication," Proc. Inst. Elec. Engrs. (London), 101, Part III, 225-237 (1954).
5. J. S. Barlow, "An Electronic Method for Detecting Evoked Responses of the Brain and for Reproducing their Average Waveforms," Electroencephalog. and Clin. Neurophysiol., 9, 340-343 (1957).

6. J. S. Barlow, "A Small Analog Electronic Averager for Evoked Potentials," Second International Conference on Medical Electronics, Paris, June 24-27, 1959.

7. W. A. Clark, Jr., "Average Response Computer (ARC-1)," Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., 114-117, 15 April 1958.

8. R. L. Koehler, An Amplitude and Latency Measuring Instrument with Digital Output, S.M. Thesis, Department of Electrical Engineering, M.I.T., 1956.

9. K. Putter, A Time-Gated Amplitude Quantizer for Neural Signals, Technical Report 275, Research Laboratory of Electronics, M.I.T., Jan. 13, 1954.

10. A. K. Hooks, Automatic Measurement of Neural Pulse Latency, S.M. Thesis, Department of Electrical Engineering, M.I.T., 1954.

Appendix D

A SELECTION OF PUBLICATIONS BY MEMBERS AND/OR ASSOCIATES OF THE COMMUNICATIONS BIOPHYSICS GROUP

The following bibliography has been appended primarily for illustrative purposes. It represents a selected sample of the written output of the Communications Biophysics Group and of some of the scientists and engineers who are and/or have been associated with the group. Most of the full-length papers referred to deal with problems of the nervous system: They are thus written from a viewpoint that is complementary to that of the present monograph, in which certain aspects of methodology and techniques have been emphasized. It is perhaps worth while to stress again the fact that the record of the activities of the Communications Biophysics Group must be viewed in the context of the neurophysiology and the data-processing technology of the corresponding period. To view it in any other way would seriously distort one's perspective on the nature and the scope of what we were attempting to do.

A fuller account of the activities of the Communications Biophysics Group is to be found in the Quarterly Progress Reports of the Research Laboratory of Electronics, beginning with the issue of January, 1952.

REFERENCES

1. W. J. McGill, A Statistical Description of Neural Responses to Clicks Recorded at the Round Window of the Cat, Ph. D. Thesis, Department of Psychology, Harvard University, 1953.
2. J. S. Barlow and M. A. B. Brazier, "A Note on a Correlator for Electroencephalographic Work" Electroencephalog. and Clin. Neurophysiol., 6, 321-325 (1954).
3. K. Putter, A Time-Gated Amplitude Quantizer for Neural Signals, Technical Report 275, Research Laboratory of Electronics, M. I. T., January 13, 1954.
4. K. S. Harris, Serial Characteristics of Neural Responses to Acoustic Clicks Recorded at the Auditory Cortex, Ph. D. Thesis, Department of Psychology, Harvard University, 1954.
5. A. K. Hooks, Automatic Measurement of Neural Pulse Latency, S. M. Thesis, Department of Electrical Engineering, M. I. T., 1954.

6. J. Macy, Jr., A Probability Model for Cortical Responses to Successive Auditory Clicks, Ph.D. Thesis, Department of Mathematics, M.I.T., 1954.

7. K. H. Pribram, B. S. Rosner, and W. A. Rosenblith, "Electrical Responses to Acoustic Clicks in Monkey: Extent of Neo-cortex Activated," J. Neurophysiol., 17, 336-344 (1954).

8. W. A. Rosenblith, "Electrical Responses from the Auditory Nervous System," Ann. Otol., Rhinol., & Laryngol., 63, 839-860 (1954).

9. W. A. Rosenblith, "Some Electrical Responses from the Auditory Nervous System," Proceedings of the Symposium on Information Networks, Polytechnical Institute of Brooklyn, pp. 223-247, April 12-14, 1954.

10. J. S. Barlow and R. M. Brown, An Analog Correlator System for Brain Potentials, Technical Report 300, Research Laboratory of Electronics, M.I.T., July 14, 1955.

11. G. S. Fahringer, An Evoked Response Detector for Brain Potentials, S.B. Thesis, Department of Electrical Engineering, M.I.T., 1955.

12. D. H. Raab and N. Y-S. Kiang, "Variability of Electrical Activity of the Cat's Auditory Cortex," Am. J. Physiol., 183, 652 (A) (1955).

13. A. Albert, Analysis of Variance in Reaction Time Experiments, S.B. Thesis, Department of Mathematics, M.I.T., 1956.

14. M. A. B. Brazier, "The Control of Input from Sensory Receptors by Output from the Brain," Proceedings of the First International Congress on Cybernetics, Namur, Gauthier-Villars, Paris, pp. 857-861, 1956.

15. M. A. B. Brazier and J. S. Barlow, "Some Applications of Correlation Analysis to Clinical Problems in Electroencephalography," Electroencephalog. and Clin. Neurophysiol., 8, 325-333 (1956).

16. L. S. Frishkopf, A Probability Approach to Certain Neuroelectric Phenomena, Technical Report 307, Research Laboratory of Electronics, M.I.T., March 1, 1956.

17. R. L. Koehler, Amplitude and Latency Measuring Instrument with Digital Output: Application to Neurophysiological Events, S.M. Thesis, Department of Electrical Engineering, M.I.T., 1956.

18. W. A. Rosenblith, "A Probabilistic Approach to the Study of the Electrical Activity of the Auditory Nervous System," 20th International Physiological Congress: Abstracts of Communications, pp. 774-775, 1956.

19. J. S. Barlow, "An Electronic Method for Detecting Evoked Responses of the Brain and for Reproducing Their Average Waveforms," Electroencephalog. and Clin. Neurophysiol., 9, 340-343 (1957).

20. J. S. Barlow, M. A. B. Brazier, R. Kjellberg, and W. H. Sweet, "Electrographic Recording and Correlation Analysis from Deep Structures within the Human Brain," Conference on Electrical Studies on the Unanesthetized Brain, Georgetown University, June 10-13, 1957 (in press).

21. M. A. B. Brazier, "Depth Recordings with Correlation Analysis from the Basal Ganglia in the Unanaesthetized Human Brain," Symposium on Depth Recording from the Unanaesthetized Human Brain, Georgetown University, June 10-13, 1957 (in press).

22. M. A. B. Brazier, "EEG Studies of Flicker in Normal Man," ERDL Symposium on Flicker, Fort Belvoir, Virginia, and Tulane University, pp. 199-221, April, 1957.

23. M. A. B. Brazier, "A Study of the Late Response to Flash in the Cortex of the Cat," Acta Physiol. Pharmacol. Neerl., 6, 692-714 (1957).

24. B. G. Farley, L. S. Frishkopf, W. A. Clark, Jr., and J. T. Gilmore, Jr., Computer Techniques for the Study of Patterns in the Electroencephalogram, Technical Report 337, Research Laboratory of Electronics, M. I. T., November 6, 1957 (Lincoln Laboratory Technical Report 165).

25. G. E. Forsen, An FM Transmitter for Recording the Electrical Activity of the Nervous System, S. M. Thesis, Department of Electrical Engineering, M. I. T., 1957.

26. M. H. Goldstein, Jr., Neurophysiological Representation of Complex Auditory Stimuli, Technical Report 323, Research Laboratory of Electronics, M. I. T., February 19, 1957.

27. J. R. Hughes and W. A. Rosenblith, "Electrophysiological Evidence for Auditory Sensitization," J. Acoust. Soc. Am. 29, 275-280 (1957).

28. J. F. Kaiser and R. K. Angell, New Techniques and Equipment for Correlation Computation, Technical Memorandum 7668-TM-2, Servomechanisms Laboratory, M. I. T., December, 1957.

29. W. B. Lehman, Design of Stimulus Programing Device, S. M. Thesis, Department of Electrical Engineering, M. I. T., 1957.

30. W. A. Rosenblith, "Relations between Auditory Psychophysics and Auditory Electrophysiology," Trans. N. Y. Acad. Sci., 19, 650-657 (1957).

31. T. T. Sandel and N. Y-S. Kiang, "'Off' Responses in the Auditory Nervous System of Anesthetized Cats," J. Acoust. Soc. Am., 29, 184(A) (1957).

32. T. T. Sandel and N. Y-S. Kiang, "Off-Responses to Noise Bursts at the Auditory Cortex of the Cat," J. Acoust. Soc. Am. 29, 1249(A) (1957).

33. M. A. B. Brazier, "Studies of Evoked Responses by Flash in Man and Cat," Reticular Formation of the Brain, H. H. Jasper, L. D. Proctor, R. S. Knighton, W. C. Noshay, and R. T. Costello, Eds., Little, Brown and Company, Boston, pp. 151-176, 1958.

34. M. A. B. Brazier, "The Development of Concepts Relating to the Electrical Activity of the Brain," J. Nervous and Mental Disease 126, 303-321 (1958).

35. L. S. Frishkopf and W. A. Rosenblith, "Fluctuations in Neural Thresholds," Symposium on Information Theory in Biology, H. P. Yockey, R. L. Platzman, and H. Quastler, Eds., Pergamon Press, Inc., New York, pp. 153-168, 1958.

36. C. D. Geisler, L. S. Frishkopf, and W. A. Rosenblith, "Extracranial Responses to Acoustic Clicks in Man," Science, 128, 1210-1211 (1958).

37. M. H. Goldstein, Jr. and N. Y-S. Kiang, "Synchrony of Neural Activity in Electric Responses Evoked by Transient Acoustic Stimuli," J. Acoust. Soc. Am., 30, 107-114 (1958).

38. W. Thompson, Jr., Human Reaction Time to a Masked Acoustic Stimulus, S. B. Thesis, Department of Electrical Engineering, M. I. T., 1958.

39. J. S. Barlow, M. A. B. Brazier, and W. A. Rosenblith, "The Application of Autocorrelation Analysis to Electroencephalography," Proceedings of the First National Biophysics Conference, Columbus, Ohio, March 4-6, 1957, H. Quastler and H. J. Morowitz, Eds., Yale University Press, New Haven, 622-626, 1959.

40. C. D. Geisler, C. E. Molnar, W. T. Peake, T. F. Weiss, and T. T. Sandel, "Loudness Judgments on Clicks," J. Acoust. Soc. Am., 31, 128(A) (1959).

41. M. H. Goldstein, Jr., N. Y-S. Kiang, and R. M. Brown, "Responses of the Auditory Cortex to Acoustic Stimuli," J. Acoust. Soc. Am., 31, 356-364 (1959).

42. J. L. Hall, II, A Psychoacoustic Study of the Mechanism of Binaural Fusion, S. M. Thesis, Department of Electrical Engineering, M. I. T., 1959.

43. W. T. Peake, M. H. Goldstein, Jr., and N. Y-S. Kiang, "'Steady-state' Auditory Nerve Potentials for Different Stimulus Repetition Rates," J. Acoust. Soc. Am., 31, 124(A) (1959).

44. T. F. Weiss, Some Properties of the Finite Time Sample Autocorrelation of the Electroencephalogram, S. M. Thesis, Department of Electrical Engineering, M. I. T., 1959.

45. W. A. Rosenblith, "Information and Control in Organ Systems," IRE Trans. on Med. Elec., ME-6, 23(A) (1959).

46. W. A. Rosenblith, "Sensory Performance of Organisms," Revs. Mod. Phys., 31, 485-491 (1959).

47. W. A. Rosenblith, "Some Quantifiable Aspects of the Electrical Activity of the Nervous System (with Emphasis upon Responses to Sensory Stimuli)," Revs. Mod. Phys. 31, 532-545 (1959).

48. N. Y-S. Kiang and M. H. Goldstein, Jr., "Tonotopic Organization of the Cat Auditory Cortex for Some Complex Stimuli," J. Acoust. Soc. Am., 31, 786-790 (1959).

49. J. S. Barlow, "A Small Analog Electronic Averager for Evoked Potentials," Second International Conference on Medical Electronics, Paris, June 24-27, 1959.

INDEX