A COMPUTER-AIDED INVESTIGATION
OF MOTION DETECTION UNITS
IN THE FLY

Thesis by
John Cedric Dill

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1970
(Submitted November 26, 1969)
ACKNOWLEDGEMENTS

I would first like to express thanks to Professor Gilbert D. McCann, my advisor, for the continuing guidance and support he has given me during the course of this investigation. I am also grateful to Dr. K.-I. Naka for many valuable discussions and for his comments on parts of this manuscript.

G. Verkaik and D. Hodgetts provided valuable technical assistance during the experiments, as did D. Aranovich and B. Ellert.

Finally, I wish to express my appreciation to the National Institutes of Health for their generous financial support throughout the course of this work. This support was from Trainee Grant Number GM 01335.
ABSTRACT

This work deals with two related areas: processing of visual information in the central nervous system, and the application of computer systems to research in neurophysiology.

Certain classes of interneurons in the brain and optic lobes of the blowfly Calliphora phaenicia were previously shown to be sensitive to the direction of motion of visual stimuli. These units were identified by visual field, preferred direction of motion, and anatomical location from which recorded. The present work is addressed to the questions: (1) is there interaction between pairs of these units, and (2) if such relationships can be found, what is their nature. To answer these questions, it is essential to record from two or more units simultaneously, and to use more than a single recording electrode if recording points are to be chosen independently. Accordingly, such techniques were developed and are described.

One must also have practical, convenient means for analyzing the large volumes of data so obtained. It is shown that use of an appropriately designed computer system is a profitable approach to this problem. Both hardware and software requirements for a suitable system are discussed and an approach to computer-aided data analysis developed. A description is given of members of a collection of application programs developed for analysis of neurophysiological data and operated in the environment of and with
support from an appropriate computer system. In particular, techniques developed for classification of multiple units recorded on the same electrode are illustrated as are methods for convenient graphical manipulation of data via a computer-driven display.

By means of multiple electrode techniques and the computer-aided data acquisition and analysis system, the path followed by one of the motion detection units was traced from one optic lobe through the brain and into the opposite lobe. It is further shown that this unit and its mirror image in the opposite lobe have a mutually inhibitory relationship. This relationship is investigated. The existence of interaction between other pairs of units is also shown. For pairs of units responding to motion in the same direction, the relationship is of an excitatory nature; for those responding to motion in opposed directions, it is inhibitory.

Experience gained from use of the computer system is discussed and a critical review of the current system is given. The most useful features of the system were found to be the fast response, the ability to go from one analysis technique to another rapidly and conveniently, and the interactive nature of the display system. The shortcomings of the system were problems in real-time use and the programming barrier—the fact that building new analysis techniques requires a high degree of programming knowledge and skill. Directions for further development are suggested. It is concluded that computer systems of the kind discussed will play an increasingly important role in studies of the central nervous system.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>The Computer in Neurophysiological Research</td>
<td>7</td>
</tr>
<tr>
<td>III</td>
<td>The Visual System of the Fly</td>
<td>37</td>
</tr>
<tr>
<td>IV</td>
<td>Experimental Methods and Apparatus</td>
<td>63</td>
</tr>
<tr>
<td>V</td>
<td>Results</td>
<td>109</td>
</tr>
<tr>
<td>VI</td>
<td>Discussion and Conclusions</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>LIST OF REFERENCES</td>
<td>211</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The work described in this thesis deals with two major areas: information processing in the nervous system and the application of computer systems to research in neurophysiology. The goals of this work were threefold: (1) to contribute to our understanding of the mechanisms involved in the processing of visual information in the central nervous system, (2) to develop the computer as an aid in this endeavor, and (3) to explore the problems involved in using a computer system in neurophysiological research.

The nervous system of the insect has been found to be a convenient vehicle for investigating the mechanisms involved in neural processing of visual information [8, 36, 57]. In particular, certain members of the order Diptera have been subjected to intensive studies on a number of levels: histological [10, 80, 85], electrophysiological [6, 8, 35, 88] and behavioral [22, 28, 59].

One of the more important findings was the discovery of motion detection units in the optic lobes and brain of various dipterans [6]. These units were sensitive not only to the velocity of a moving pattern, but also to the direction of motion. Several classes of such units have been found [7, 8, 58]. These classes have been identified by the following characteristics:
visual field, direction of motion for maximal response and anatomical location from which recorded. Because of observed correlations with behavior, it has been suggested that these directionally sensitive units might be part of the neural mechanism underlying optokinetic responses [7].

Of the various classes of motion detection units found, some differed only in location recorded from, others only in the sense of direction of preferred motion (i.e. direction of maximum sensitivity differs by $180^\circ$), still others had none of the identifying characteristics in common. These two sets of observations, existence of identifiable units and varying commonality of identifying characteristics, suggested the following questions, to which this work is addressed: (1) is there any interaction between various pairs of these units, (2) if such relationships can be found, what is their nature.

To answer these questions, it is essential to be able to record signals from two or more units simultaneously. It is also necessary to use more than a single recording electrode if the recording points are to be chosen independently. Although there are innumerable questions that can and need to be investigated using a single recording electrode, there has been an increasing awareness [63, 66, 87] of the need for studies involving two or more recording points, and it has been suggested that major progress toward a deeper under-
standing of neural mechanisms is limited unless means in addition to single electrode studies are sought, developed, and used. The use of multiple electrodes wherein relationships among interacting units are investigated is one such approach that appears promising. Ferkel et al. [66, p. 420] for example say that

it is ... in the comparison of two or more simultaneously observed spike trains that we believe lies the greatest potential usefulness for the statistical analysis of the data provided by precise measurement of times of spike events.... These experimental techniques [of recording from several neurons simultaneously] have been motivated by the promise that intercomparison of spike trains will reveal details of synaptic connections and other sources of interaction among the observed neurons.

The use of multiple electrodes in our investigations into the possibility of interaction between motion detection units enabled us to demonstrate that in fact interesting relationships do exist, for example, in certain mirror image pairs of motion detectors. The use of more than one electrode had the important additional benefit that physiological pathway tracing was possible; that is, the course of certain fibers could be traced by means of recording simultaneously from the same unit at two different locations.

Multiple electrode recordings from single units have not been common, not only because of difficulties in obtaining signals from more than a single electrode, but also in
analyzing the data so obtained. In particular, the volume of data increases greatly and there are problems in dealing with such large quantities of data, and in identifying and understanding relationships between simultaneously recorded spike trains. For instance, what tools are available for analyzing such data? Traditional techniques appear to be of limited use. Examination of film strip recordings, besides being terribly tedious, cannot produce any quantitative measure of the relationship between two units. Measurements can be made by hand [32], but the accuracy and reliability of such measurements is not high, and of course the time required is sufficiently long so as to make them impracticable for statistically significant numbers of spikes.

The same workers that mention the desirability of looking at many points simultaneously also describe the difficulty in so doing, and the need for new tools and techniques. Special purpose devices, digital and analog, have been developed and used for various purposes (averaging, signal generation, event recording, etc.), but their limited capabilities restrict their use to a narrow range of applications. A common comment [20, 73, 87] is that the computer appears to be worth considering as a potential aid in such endeavors. Taken with the current growing use of computers in biology, these recommendations suggest
that the use of a computer may be a profitable approach, allowing a better, more thorough exploration of the problem.

Computers however, do not come from the manufacturer as ready-made data acquisition and analysis systems. Rather it is the choice of a suitable hardware configuration, and even more important, the design of an appropriate software system, that transform the machine into a useful aid. Does research of the kind proposed above impose particular demands on a computer system? What kinds of hardware configuration are needed for the computer to be a useful tool in neurophysiology? What are desirable software characteristics? Appropriate answers to these questions will help in answering our biological questions. Thus, part of our work has been an attempt to develop the computer as an aid to the neurophysiological researcher by means of devising a collection of application programs in the environment of, and with support from, an appropriate computer system.

Presented in the following chapter, then, is a discussion of the requirements for a suitable computer system, and of the possibilities and problems in realizing such a system. Chapter III deals with the visual system of the fly and discusses the results of previous investigations on which the present work is based. In Chapter IV we describe the experimental apparatus and techniques, including the computer system. Results of the neurophysiological investigation are then presented in Chapter V, along with pertinent
information on computer system performance. Finally, in Chapter VI a discussion is given of the neurophysiological results and of the experience gained from use of the computer system, along with a critical review of the current system and recommendations for future modifications.
CHAPTER II

THE COMPUTER IN NEUROPHYSIOLOGICAL RESEARCH

Introduction

In the previous chapter, we outlined briefly the neurophysiological questions to be examined and introduced some problems and difficulties. Having suggested the computer as a possible aid in overcoming these difficulties, in this chapter we investigate the potentials and limits for computer applications in electrophysiological research. We shall be primarily concerned with those investigations for which the basic observations consist of spike potentials recorded from a relatively small number of neurons, or possibly various forms of slow wave potentials. We shall also attempt to define properties of computer systems useful in this and other areas of neurophysiology.

The first question to which we shall address ourselves then is can the computer be of significant help? Given encouragement by an affirmative finding, we are then led to ask what it is we want the computer to do--what are the areas in which it can be of most use. Finally, we will want to know how to get the computer to perform the desired tasks--what are reasonable design criteria for a computer system for neurophysiological research.
The fact that the use of computers in biology is neither new nor limited to a narrow range of applications might be thought of either as making the first question unnecessary, or the answer obvious. Opinions on the extent of their usefulness differ, however. According to Ledley, for example, [49, p. 2] "... the electronic computer makes feasible the solution to [biological] problems that could not otherwise be approached." (Italics his.) Other viewpoints vary from acceptance of limited usefulness to the feeling that the time and effort spent on the computer system may detract from (or be distracting to) the biological work, or may not be worthwhile. In one recent study [3] it was concluded that the effort involved in using very precise painstakingly careful statistical techniques via a (special purpose) digital computer was not worth the small increase in information over that obtainable with analog equipment or with more subjective methods. In view of these varying opinions, it seems worthwhile to review briefly the capabilities and limitations of computer systems and to show positive evidence for the usefulness of computer systems in biological research.

Can computers be useful? Certainly computers can process large amounts of data and can perform a great variety of operations on data in much less time than a human investigator, with resulting improvements in accuracy and reliability of the calculations. Advantages accruing generally fall into one of two
categories: the investigator is relieved of routine tedious chores that are completed faster and with increased accuracy and reliability; and perhaps more important, certain tasks are feasible that were previously very difficult or impossible. A simple example of the former would be routine statistical processing of large data volumes with calculation and plotting of associated distributions. Automated plotting alone can provide immense savings in time and effort. For an example of the latter, in modern quantum-chemical calculations, an essential role is played by the computer, without which only the most elementary models can be studied [56].

Other examples may be taken from medicine where, for example, analysis of electrocardiograms [72] can be performed by computer. Another system, recently developed, does automatic analysis of chromosomes [50]. The latter system is more automatic and advanced than many in that the whole process is computer-controlled. The input to the system is a film roll of photomicrographs. The system controls the film advance, asking for the next frame when processing of the previous one is complete. Scanning of the film, digitizing, and input to the computer memory takes less than half a second, with resolution equal to that of an optical microscope at 1000 power. Once data are in memory, the picture is processed; first each object is separated and identified; if it is recognized as a chromosome, its characteristics are saved.
for further processing. After all chromosomes are separated out, homologous chromosomes are paired according to arm length, area, etc., and grouped into classes. Finally, statistical calculations are made after a number of frames have been processed.

The computer has its limitations, some of which are obvious. Although computers can handle large amounts of data, their capacity is not unlimited; they can perform calculations rapidly, but processing still takes time and it is not difficult to think of computations that would require an impossibly long time (the classic example is the time required to compute the consequences of all possible moves in a chess game--an astronomical number of years). Even for the most powerful of today's machines, the simulation of chemical reactions requires many hours. The full services of an IBM 7090 are required for the chromosome recognition task described above. As a specific example, consider computer processing of an electroencephalogram (EEG). An EEG record consists of 8-16 channels, each having frequency components of interest from 1 to 50 Hz [76]. In practice, sampling rates for each will be 250-500 samples/sec., a total of 2-5 million samples for a 10-minute period. Unless some attention is given to the problem of sorting out the useful information, the volume of data will quickly overwhelm even the largest systems. Similarly, for analysis of electrocardiograms, sampling rates up to 1000/sec are necessary, resulting in an overall rate of 12,000 samples/sec for the twelve leads [72].
The processing time for such data in a small machine is lengthy indeed, over twelve hours for a complete record [72].

More difficult limitations to overcome are those associated with recognizing the "reasonableness" of given input data, whether the latter are raw experimental data, or results of a preliminary calculation to be used as input to a succeeding stage. While it could be obvious to a human observer that a preliminary curve-fitting result was nonsensical, unless special time and memory consuming precautions are taken (which might or might not be possible), computer processing of succeeding stages in the computation would continue unabated. Recent developments in conversational or interactive systems, computer controlled display devices, and their combination into "interactive computer graphics systems" have been important in helping to overcome some of these problems.

The term "computer graphics" is usually taken to mean the set of display devices and computer techniques by means of which data in the form of line drawings or graphs is accepted as input from, or presented as output to, the user. The display device most commonly used in interactive systems is, by reason of its speed, the computer-controlled cathode ray tube (CRT), first used as a computer output terminal in the 1950's [42]. By interactive system, we mean one in which there is give-and-take between user and system; the user can choose a particular
analysis technique for his data, examine the result, modify the technique, and try again, all in one continuous session on the computer.

As a simple example of the use of a graphics terminal, consider the problem described earlier of analyzing electrocardiograms. Instead of having the computer decide which inputs are artifacts, or otherwise unacceptable for processing, the system could first display raw data from the electrocardiogram leads. The user could then indicate which areas were to be used for analysis either by pointing at them with a light pen, or by shifting the data back and forth on the display screen. The latter method was also found to be particularly effective in analysis of pressure and flow data in studies of the lung [17].

Thus, although the computer cannot replace the human investigator's imagination, intuition and judgment, interactive or symbiotic systems can be developed in which man and machine cooperate, with each performing the tasks for which best suited.

To develop additional insight into the potential for computer use, let us review briefly what the experiences have been in a particular area. Because the physicists were among the first to make extensive use of computers in their investigation [81], they have probably progressed further than researchers in most other areas in making use of computer systems, and it will be instructive to look at what their problems were, how the use of computers
helped them to overcome these problems, and what other benefits were obtained in the process of doing so.

One of the basic problems mentioned in several recent works [14, 24, 30, 91] was that of reducing the time gap between the production of raw data and obtaining useful, physically meaningful, results. If results, even preliminary ones, could be made available rapidly enough, not only would mistakes resulting in the loss of hours of expensive reactor or accelerator time be avoided, but the results could be used to guide the path of the experiment while it was in progress [24], with consequent increases in productivity. Similarly, in the study of difficult to prepare isotopes with relatively short half-lives [91], to check the validity of the experiment, it was essential that results of analysis of raw data be made available while the sample was still viable.

The physicists were also faced with the problem of cataloging growing amounts of data, and increasingly often, extensive analysis of experimental data was required to bring out the physics of the experiment [30], which resulted in the tedious time consuming process of punching data on cards for submission to a remote computer center. Noise problems had also become more troublesome [30] as increasingly subtle effects were studied. Development of on-line systems capable of accepting data directly from the experiments has helped to overcome most of these problems. In fact, by enabling the physicist to repeat an experiment many times over a
relatively short time span, the computer has become ". . . the principal tool of the experimenter in allaying noise problems." [30, p. 16]

More interesting, however, is the fact that without the on-line computer facilities, many experiments would not have been attempted; in fact, some were not even thought of until after the systems had been implemented [5, 62]. The problems mentioned most often in implementing and using these systems were expense of a big machine and the "programming barrier" (the fact that a considerable effort is required in learning to use and program the computer effectively--more so for the smaller systems). Also referred to is the problem discussed above of the tradeoff between having decisions made by the program, making for a more automatic system, and an interactive system which can make use of the experimenter's intuition and judgment [14, 5].

We should note here that the physicists have had a number of advantages in developing and making use of computer systems. First, the experiments were already well instrumented, with much data directly available in digital form. Thus, the tasks of obtaining data and transforming it into computer-acceptable forms were already well in hand, with the remainder, if somewhat complicated, not difficult in principle. For example, much of the equipment used to recognize and analyze interesting events in bubble chamber photographs was readily adapted to the computer [4]. Moreover, selec-
tion criteria for reduction of raw digitized data were in many cases well defined and amenable to computer processing [4, 5, 30]. Further, highly structured theories were available by means of which the data could be organized, enabling precise correspondences to be made between theory and experiment.

Certainly then the physicist has found computers to be of great help, not only in improving the quality and reducing the tedium of pre-computer approaches, but in many cases producing results that would otherwise have been difficult or impossible to obtain.

Let us now consider the potential for computer applications in neurophysiology, using the foregoing general remarks on capabilities and limitations and particular examples drawn from physics as a framework. With regard to the problem of getting data into a computer system, there are some aspects similar to the case for physics. Many experiments are now well instrumented, with only minor difficulties preventing others from becoming so. Not only is it feasible to get data into the computer, it is also possible in some instances to have inputs to the system being studied controlled by the computer, as in the case for physics [74]. Once raw data are available for machine processing, there are numerous benefits to be obtained in the areas of data reduction and analysis in helping the investigator deal with his problems of variability of responses and high data rates and volumes. Consider for instance the obvious advantages over the procedure reported in [32], wherein spike
firing times were obtained by measuring distances along a filmstrip with a ruler. Many of the techniques normally used in analyzing such neurophysiological data (histograms of inter-spike intervals, latency times, etc.) are readily adapted to computer processing with consequent increases in speed, accuracy and reliability, and decreases in time and effort spent by the investigator.

On-line use of the computer can also make possible some techniques not often otherwise available, such as watching the formation of a histogram with time to detect nonstationarities in data. Three dimensional displays can help the experimenter to visualize system responses. The potential speed and flexibility allow him to consider his data from several points of view instead of just one: firing rate versus time may suggest a latency histogram; should this prove unrewarding, a post-stimulus-time histogram could be examined; if the latter is interesting, it can be compared with responses of other units; averages of such histograms can be formed and examined.

**Nature of the Problem**

Having shown the potential of the computer, we turn now to the question of what it is we want the computer to do; what does the neurophysiologist want to know, what are the procedures he uses, what are the problems confronting him,
and what are the areas in which a properly implemented computer system can help him.

In neurophysiology, as in other areas, one attempts to build an understanding of a system as a basis on which to explain certain observations. In other words [83, p. 2], "... theories must be formed which account for behavioral evidence in terms of mediating mechanisms." In forming such theories, the available data and evidence are examined in the light of current knowledge of the system to be studied. On this usually very incomplete basis, one attempts to build a complete picture, filling in gaps with best guesses, and then seeks a test of the validity of such guesses or working hypotheses. Usually such tests provide information in addition to that sought, which is then used to modify and refine the existing theory structure. In making the tests, stimuli or inputs whose general nature and form are determined on the basis of a preconceived model are applied to the system to be studied. The investigator must then determine, again initially on the basis of his preconceived model, which of the measurements he can make are appropriate. He must also devise suitable, practicable methods for analyzing the raw data produced by the experiments in order to bring out the underlying physiological results with which to further modify and develop the model. Note that the choice of measurements depends not only on the
model, but will also be a function of which level the investigator is working on (considering details of physical and chemical reactions in spike propagation or synaptic transmission at the subcellular level, studying firing rates with respect to sensory input at the single unit level, measuring activity in large groups of cells via EEG or ERG, or looking at the behavior of the whole animal), of which experimental techniques are available, and equally important, of the availability of suitable, feasible methods for analyzing the data so produced. The process continues until either all new strategies or experiments suggested by the model produce only minor modifications, or the results are repeatedly inconsistent with the structure proposed for the model. In the first case, the investigation is considered complete; in the second, information and insight gained from the "unsuccessful" experiments are used to rebuild the model or to replace it with a more adequate one.

Within this framework, the neurophysiologist is faced with some particularly troublesome problems. One of the most difficult is that of making observations directly relevant to theoretical constructs. For example, having discovered the existence of a relationship between neural units, one may propose a network model to account for the nature of the observed interaction, but because of technical constraints be unable to identify experimentally other components of the network. At least part
of this problem is the extreme difficulty of recording from known neurons, or of being able to record at all from members of a particular population, due at least partly to the small size and high packing density of neural elements. Horridge et al. [35] in summarizing some of the methodological problems, mention the difficulty of doing more than describing the activity in a few large units having big spike potentials while much of the activity occurs in smaller neurons that cannot be recorded from. Appropriate analysis techniques for improving signal-to-noise ratios, such as the averaging of a number of repetitions or moving linear (or weighted) averages used by the physicist, could be helpful here. Also necessary are powerful methods for identifying and separating multiple spikes recorded on the same electrode.

Other difficulties discussed by Horridge [35] are in knowing what appropriate stimuli are, the problems of recognizing patterns in sequences of impulses and of determining their significance while lacking knowledge of the use (if any) of those patterns to succeeding stages in the neuronal chain, and difficulty in understanding interactions among many and partially unidentified units. Analysis techniques such as latency or post-stimulus-time histograms and cross-correlograms will be invaluable aids in showing the existence of patterns in and relationships among spike trains. In determining appropriate stimuli and the effects of varying stimulus parameters, the stimulus space that must be
searched may be large and the time and effort involved in a comprehensive search prohibitive, unless powerful, flexible analysis techniques can be brought to bear during the course of an experiment.

In addition, there is the pervasive problem of variability inherent in most biological work. Apparently identical systems vary considerably in their responses to the same stimuli; even variations in the same system with repetitive stimuli are the rule rather than the exception. In this respect, the biologist faces much more severe problems than does the physicist. The difficulties involved in obtaining a "good" preparation seem to make this task much more of an individual enterprise than is the case in physics. Electrode construction techniques vary widely; a given method may work well for one investigator, another may find success only with a quite different technique. It is perhaps surprising that there are not more differences of opinion, considering the variations in electrode positioning, environmental conditions, and so on.

Consider, for example, the problems in studying responses to a particular stimulus. Whether recording from single units via microelectrodes or from populations via gross electrodes (EEG, ERG), one is confronted with the problem of defining a typical response (or equivalently in determining if two responses are the same or different) in the face of variations in
both the evoked response itself and in background or spontaneous activity. One must then be able to characterize and work with distributions of events. For gross electrodes, both the amplitude and time characteristics may vary. In some cases [73], the response latency may be relatively short and thus fairly constant, in which case response averaging can be useful and is quite straightforward. For spike trains, the stimulus may result in an increase or decrease in firing rate, or more generally, some change in firing pattern. Not only will there be variability in latency of response, but in latency to the first spike (identification of which can be difficult if there is background, spontaneous firing), rise or decay time of firing rate changes, net change in firing rate, duration of change for transient responses, and so on.

Thus, for a given set of stimuli, averages of responses to many repetitions may be required to obtain a reliable outcome. Once an investigator believes promising conditions have been established, he tends to repeat an observation many times, and the length of such a sequence may be limited only by preparation lifetime. By performing averaging as the data are collected, the volume of data can be reduced. It may, however, be desirable to keep the raw as well as the averaged data since significant information may be lost in the averaging process, or it may prove necessary to use a number of different averaging techniques.
in order to extract certain response characteristics. These considerations, together with the initial effort required to get a good preparation, make the temptation to record huge volumes of data almost overwhelming. Opinions of many workers [35, 66, 87] on the necessity of obtaining simultaneous recordings from increasing numbers of points also indicate a trend toward larger data volumes. The analysis methods available to the investigator and the environment in which they operate must therefore be capable of managing large quantities of experimental data in ways that permit him not only to summarize the results of a number of repetitions, but also to be able to compare several different averages, and to look at details of a single response; not only to be able to see the forest, but to make comparisons between forests, or to examine the structure of a single tree.

To this list of problems, we add those associated with attempts to model the neural systems being studied. Although conventional mathematical approaches have been used successfully in extremely limited areas, the Hodgkin-Huxley [34] model of the neural membrane for example, the analysis of networks—even very simple ones—is beyond the capabilities of conventional mathematical tools. Harmon [33] illustrates the problem of deducing network function from component function in the following way. An expert in circuit analysis, given a complex schematic diagram with no explanation of the purpose of the circuit, will
often be completely unable to describe the circuit's function, no matter how detailed his knowledge of individual components. Thus neural networks even of highly simplified components are too complex for contemporary methods of mathematical treatment. On the other hand, in some instances, it is possible to devise models which can be studied by means of simulation. That is, a program can be constructed which, when run on a computer, simulates operation of the model [21, 65]. Other means such as electronic and ionic models are also possible [33, 41].

Classification of Analysis Procedures

The preceding discussions have emphasized the necessity of the neurophysiologist's making data analysis an integral part of his investigations in helping to overcome some of the problems facing him. Not only should data analysis contribute to the process of bringing the physiology out of the raw data, it can and should make important contributions at the experimental level. Data analysis should aid the investigator in his search for proper experimental conditions, help him to determine quickly their appropriateness or to recognize changes in the system, help him avoid collection of huge amounts of useless data and instead reduce meaningful data to useful, manageable quantities. If the quality of an experiment can be determined rapidly enough, the investigator may be able to exploit a subject more efficiently and more
thoroughly by modifying experimental conditions, for example, in order to improve the isolation of certain phenomena, to study the subject under new sets of parameters, or to continue the present course after having established its adequacy.

In its contribution to the process of forming and modifying new concepts, data analysis should allow the investigator to view his data in a variety of ways, to shape the data to his hypotheses. Its fundamental purpose has been well expressed by Tukey and Wilk [86, p. 7]:

The basic intent of data analysis is simply stated: to seek through a body of data for interesting relationships and information and to exhibit the results in such a way as to make them recognizable to the data analyzer and recordable for posterity. Its creative task is to be productively descriptive, with as much attention as possible to previous knowledge, and thus to contribute to the mysterious process called insight.

To help us in determining desirable characteristics for a computer system capable of aiding the neurophysiological investigator, we summarize the preceding discussions in the following two schemes for classifying various methods employed during analysis of experimental data. For the first of these, the techniques are differentiated primarily on the basis of time. That is, we consider the differences between analysis techniques used while data is being collected, those used during an experiment, but not during actual data acquisition, and those used after completion of an experiment.
1.1 Pre-analysis
Techniques in this class serve for immediate evaluation or control of an on-going experiment. Analysis and data recording or collection occur simultaneously. A simple example would be direct observation of signals via loudspeaker or oscilloscope. Traditionally, pre-analysis has been based on such direct observation of raw data, or on reduction methods using hard-wired apparatus such as counters or average response computers. Although these devices have proven to be of value, indeed most are still a necessity, any basis for decision making on the part of the experimenter is restricted to observations of raw data or quantities derived from them in limited ways.

1.2 Sample-analysis
Methods in this category are used for evaluation of an experimental run immediately upon its completion. These "between-run" analysis techniques may aid the investigator in determining succeeding steps in his experiment. Sample-analysis is used when the desired method is based on data from the entire run or from several runs, or when there is not enough time to perform required calculations during the run, whether due to complexity of calculation or high data rates.
1.3 Post-analysis

In this stage highly time-consuming techniques are applied, series of experiments are compared and correlated, and various alternative approaches to interpretation of results are explored. In addition, new techniques for pre- and sample-analysis are developed and tested.

It is also useful to consider a classification of analysis techniques with respect to the role played during analysis and evolution of a theory. Accordingly, we present the following second classification scheme.

2.1 Editing

Functions performed by methods in this category are improvement in signal-to-noise ratio or isolation of signal from background noise, separation of simultaneously observed events, removal of unwanted portions of records, and forming various combinations of selected sections of records in preparation for more complex techniques.

2.2 Existence Tests

These disclose existence or absence of well-defined properties of the data.

2.3 Data Massaging

Characterization of this stage is difficult, possibly because it is a more creative or constructive stage, one from which
new concepts are more likely to emerge. Indeed, there may often be no specific reason for the analyst's choice of a particular method other than a vague expectation of an "interesting" outcome. Tukey and Wilk's comment [86, p. 7] that "... exposure, the effective laying open of the data to display the unanticipated, is to us a major portion of data analysis." applies to these methods. In general, these techniques are used in the search for unknown or suspected effects by helping the analyst explore a number of strategies, to vary parameters of given ones, and to expose the data to a variety of treatments. These might include procedures for counting, averaging, generation of distributions, curve fitting, and most important, for graphical presentation.

2.4 Modeling

Overall descriptions are developed here. Results of analysis are interpreted into the current structure, the latter is examined in the light of the former and guidance is sought for further analysis and experimentation.

The Role of the Computer

So far, we have indicated the usefulness of computer aid to the neurophysiologist, outlined some of the problems facing him in his endeavors, emphasized the value of the computer in data analysis, and developed two ways of characterizing methods of data analysis.
We turn now to the question of the role of the computer in data acquisition, reduction, and analysis. It should be stressed that we view data analysis as an activity or on-going process, rather than an object or entity. We repeat that it is an iterative process, beginning with very simple techniques and proceeding slowly, with many loops and dead ends through more and more complex formalized procedures, in attempting to clarify or confirm hypotheses and to discover new phenomena or relations. Briefly, data analysis is the name we give to the activity involved in employing a collection of procedures in searching through a set of data for interesting relationships relative to some theory.

Now, what are the implications of the classifications of the preceding section for computer aided data analysis? Consider the overall analysis process as consisting of the following stages:

1. collect the data or make the observations
2. decide on an analysis technique (choose or construct a technique)
3. perform the technique
4. examine the result
5. decide on next analysis technique or modify theory accordingly.

It would appear that the major contribution of a computer system would be step 3. Although this is indeed an important contribution, others of equal or greater importance are possible.
Consider first the post-analysis stages, assuming step 1 is complete and raw data are in the machine. The two most obvious benefits of computer aid are that the machine can perform some of the tasks done by the experimenter better, more reliably, and usually much faster, and that the machine makes feasible certain techniques not otherwise possible. In other words, in many cases, the machine performs step 3.

What about contribution to step 2? The decision to use a particular method is certainly based in part on the time and effort required for, or the practicability of, performing the necessary calculation. By increasing the size of the space of possible methods then, significant benefits can be obtained. How should a system be designed to realize these? Step 2 says choose or construct a technique. Although the bulk memory of a computer system can contain a large number of analysis programs, the analyst will always want something not available, or want some variation of an existing program, both of which will require much effort on the part of an experienced programmer.

One way of resolving this difficulty, as indicated in step 2 itself, would be to allow the analyst to construct techniques of his own design from basic components supplied by the system. However, there are fundamental problems [84] not only in designing a system that permits the non-programmer to construct his own language, but also in providing building blocks or components that
are natural and convenient for the analyst to use. A step towards this goal would be a compromise solution in which the analyst is presented with a reasonable choice of very flexible "preprogrammed" techniques and is allowed some form of composition of these. The system to be described in Chapter IV represents such an approach.

How can the computer help with the fourth step? In step 4 the result of some procedure is looked at, thought about, interpreted with respect to other results or with respect to current hypotheses. Almost by definition, this is entirely a task for the investigator. However, it is much more difficult for him to "examine the result" and to understand what is going on by looking at a table of numbers than by looking at the same data in the form of a graph. It is even easier if the graph can be rescaled, perhaps several times, quickly and conveniently, or if the data can be viewed in both smoothed and unsmoothed states, or compared easily with a number of other results or with the output of a model. With fast, flexible display and data manipulation capabilities, the computer can greatly enhance the analyst's abilities to "examine the results" as he attempts to gain insight into the system being studied. We see then that the computer's role is by no means limited to just performing calculations, that it can contribute to the whole process of data analysis.
For the pre- and sample-analysis stages of data analysis, much of the foregoing still applies, but with greater emphasis on speed and convenience. The primary objective for both stages is to give the investigator quick-look facilities for rapid visual summaries. For pre-analysis this means looking at the data as it is being acquired and stored for further processing during sample- and post-analysis, to provide the investigator with means of monitoring the progress of his experiment. Simple techniques such as a display of firing rate versus time from the beginning of the run can be helpful. For a repetitive stimulus, displaying both the current and accumulated average response could aid in monitoring changes in the preparation. For sample-analysis, existence tests can be used to help direct the course of the experiment, but more flexibility is needed than for pre-analysis. Speed and convenience are the main concerns for the latter: speed to keep up with the data; convenience because the investigator must be free to concentrate on the experiment and should not be distracted by having to perform complicated actions on the computer.

Three basic roles can now be identified, corresponding approximately to the pre-, sample-, and post-analysis stages. The first two apply during the course of an experiment, the later, after or between experiments. The pre-analysis function of the computer system is analogous to that of the oscilloscope and
loudspeaker, to help monitor the preparation. The sample-analysis function is to help evaluate the experiment's progress, to provide preliminary analysis while the preparation is still viable, to help the experimenter determine which direction to take in continuing his experiment. Post-analysis techniques should aid the investigator in looking at his data in a variety of ways, to help him shape his data to his hypotheses and to help him abstract the physiological meaning out of the raw data.

**Design Criteria**

Having outlined what it is we want the machine to do, we turn now to the question of how to design such a system. Until recently, most uses of computers in biology were limited to submitting punched cards to a distant computer center for processing by some pre-packaged program. Although such procedures are still widely used, and indeed still useful [65], they provide no facilities for making symbiotic use of the investigator's abilities, and are completely unsuited for any kind of pre- or sample-analysis techniques [55]. However, with the recent introduction of small (relatively) inexpensive machines, the so-called laboratory instrument computer has become commonplace, and there is now a moderately large class of investigators using them. The laboratory computers have done much in the way of reducing the time and effort of getting data from an experiment into a machine, as well as providing the investigator with a limited amount of computing
power over which he has more immediate control. In our terms, they have helped at the pre- and sample-analysis levels.

These machines are not without their problems though, some of which arise simply from their small size, which of necessity limits their abilities, especially with regard to post-analysis techniques and manipulating large data volumes. Other than for rather simple straightforward applications, programming generally requires much more effort for the smaller machines.

What we wish to suggest, then, is that the computer be integrated into all phases of data analysis. Special emphasis should be placed on furnishing tools which permit a much larger percentage of massaging and isolation to take place during pre- and sample-analysis. Methods must be developed which deal more efficiently with larger bodies of information, allow the introduction of more complex models, and add the capability of exploring a number of alternate routes quickly and easily. Further, the computer system must preserve the identity of the investigator during all stages; it must be responsive to his needs, complementing his creativity and imagination with fast, precise verification.

To serve as a useful tool in neurophysiological research, a computer system should be designed against the background developed in the preceding sections. On this basis, principle criteria are:
1. Ease of communication between user and system. In making the computer respond to his needs, the investigator should not have to perform inconvenient or unfamiliar actions. It is much more natural for him to interact with the system via push buttons, switches, potentiometers, etc. than by having to type lengthy statements. In addition, the time required to "tell" the system what to do is greatly reduced in this way. While convenience can be dramatically effective at the post-analysis level [86], it is a necessity for sample-analysis and crucial for pre-analysis. Since expressing results in graphical form is essential, some form of computer-controlled display is needed. To obtain the desired speed and responsiveness, this should be some kind of CRT display unit. It is important also to be able to interact with the display. Because of the central importance of graphical interaction, the design of the system should be focused on presenting the system's capabilities to the user through the display terminal and its associated buttons, switches, etc., and on accepting user input in the same manner. Consideration should also be given to some inexpensive means of obtaining permanent records of oscilloscope displays.

2. Interaction at all levels of analysis. In order to fill the needs of pre- and sample-analysis, the system must furnish the investigator with results, both complete and partial, as fast as possible both for a more complete understanding of the immediate
experimental situation and as a basis for evaluating which course he should pursue in continuing the experiment. Rapid response by the system is also invaluable during post-analysis, either to converge quickly on an appropriate description of observations, or to recognize and abandon unrewarding approaches long before all data have been considered. At all levels, selection of an overall strategy will alternate with adjustment of individual characteristics of the chosen procedure. The researcher must be able to accomplish these variations in kind and degree with ease.

3. Step-wise analysis. The outcome of complex formal techniques may be difficult to verify intuitively. Since a computer system makes the use of such techniques possible, it should also permit analysis in steps small enough that intuitive appraisal of the validity of each is possible.

4. Addition of new techniques. Normally, the user starts out with a set of standard techniques. As his experience with the system grows, he will want to make additions or modifications. With present techniques unfortunately, this usually requires the aid of a programmer. However, the system can and should provide services that minimize the problems involved in adding a new analysis algorithm. In addition, it should support insofar as practicable the composition of larger programs from smaller ones used as building blocks.
5. Storage of data. Data storage facilities of the system should provide and maintain a single common pool of data in which all data are kept, be they directly acquired experimental data or results obtained through analysis. Since such permanent storage gives rise to bookkeeping chores, the latter should be automated as much as possible.

Based on the foregoing general design criteria, there has been developed here an experimental computer system to aid a group of neurophysiological investigators in data acquisition and analysis. The system itself is experimental in that the design goal was not only to provide a tool for the neurophysiologist, but also to investigate the problem of what a desirable computer system ought to be like to effectively aid the experimenter. Thus, rather than presenting the user with a finished product, the effort has been to design a very plastic system, to study its use, to modify the system appropriately as the need arose, and to encourage users themselves to suggest additions or modifications.

We shall postpone a description of system hardware and software to Chapter IV since it will follow more naturally at that point.
CHAPTER III

THE VISUAL SYSTEM OF THE FLY

Introduction

As discussed in chapter II, one of the fundamental objectives of neurophysiology is to explain behavioral observations in terms of underlying neuronal mechanisms. One of the first questions to be considered is the choice of a preparation suitable for the purposes expressed in previous chapters--study of the function of central nervous system neurons with the aid of the appropriate data processing system. In choosing a suitable preparation, the following general requirements should be met:

(1) There should be some clearly defined input such as light or sound.
(2) It should have been established that the animal utilizes the input to modify its behavioral patterns.
(3) The animal should have a CNS which is of reasonable complexity--simple enough that it can be studied, complex enough to be interesting.
(4) Information on the nervous system structure and on the behavior of the animal should be readily available.
(5) The animal should be readily available.
The visual systems in certain insects of the order Diptera meet appropriate criteria in all of these areas. The blowfly *Calliphora phaenicia* and housefly *Musca domestica* were chosen as the subjects for a series of concentrated studies in our laboratories [6, 8, 58, 59, 60]. As for CNS complexity, one of the major reasons cited by several workers [12, 38, 45, 71] for choosing preparations from the arthropods is the relatively small number of central neurons, especially considering the complexity of integration of which the system is capable. Although the only actual count is for the crayfish ([89], cited by Bullock and Horridge [12]), which has on the order of $10^5$ neurons, Hoyle [36] estimates that the number for most insects does not differ greatly from this figure.

Generally a high proportion of the neuronal population is sensory, but in spite of this feature there seems to be an emphasis on economy of neurons. Bullock and Horridge [12] also point out that where speed of response is important, the transmission of information between widely separated points is accomplished by a single axon (or a few) rather than a tract of many. As an example of this economy of operation, Kennedy [44] cites the claw bearing arm of the crab which uses only 24 cells to support its complex activity. A mammal on the other hand would probably use thousands of cells for the same purpose, although some simple functions such as the monosynaptic reflex might well be performed
with relatively few motor neurons. Thus, the more parsimonious systems offer the advantage of greater simplicity to the investigator of groups of cells.

In addition, arthropods generally tend to have clearly separated ganglia joined by connectives which contain fibers only. This facilitates the study of units in identifiable ganglia. The situation is similar of course in the vertebrate CNS where various nuclei are connected by tracts. Further, the neurons in arthropod ganglia have their cell bodies distributed around the outside, making them more accessible to the investigator.

The basic physiology and anatomy of *Musca* and *Calliphora* are quite well known. Histological studies at two different levels, light and electron microscopy, have provided detailed information on the ultrastructure of the receptor layers (the first stages in the visual system) of the insect compound eye [23, 29, 85, 92]. Detailed studies on the more proximal side of the central nervous system have also been carried out and information on CNS structure is readily available [10, 48, 68, 75, 80].

Both *Musca* and *Calliphora* have well defined behavioral responses to visual stimuli, the turning and landing responses being among the most widely studied. Further, there is no behavioral evidence that any permanent learning takes place in the fly as a result of sensory experience [16]. To the best of our knowledge, the properties of the nervous system that we are trying to
characterize do not change as a result of the experimental procedures we impose upon it. It therefore seems reasonable to make the assumption that the system does not change as we study it. Finally, electrophysiological activity, some of which appears to be related to behavioral responses, can be observed in various places in the nervous system.

In the following sections of this chapter, we present a brief description of the basic structure of the visual system of these insects and review some of the behavioral and electrophysiological work relevant to our study of a particular class of units involved in the detection of motion.

Structure

The visual pathway of the insect is composed of three major components: the receptor (compound eye), the optic lobe and the brain (figure 3.1).

The receptor layer is composed of the optical system and receptors. Compared with the vertebrate retina, the compound eye is characterized by its unitary structure, i.e. the eye is composed of functionally independent units called ommatidia. The dipteran eye is of the type called apposition, in which the ommatidia are optically separate. The primary eyes thus consist of a mosaic of distinct sampling stations. The optical system of each ommatidium (there are approximately 3200 in each of the two
Figure 3.1 Horizontal cross-section through head of Calliphora phaenicia showing brain and optic lobes. Also shown are locations of the major classes of units recorded. The insert is a section through the medulla and lobula-lobula plate regions parallel to the fiber plane connecting them.
primary eyes) is shared by several receptor cells. The latter are called retinula cells and have two major regions, retinula cell soma and rhabdom or rhabdomere. It is the general consensus that the photochemical pigment is in the rhabdomere microvilli, and that the first photochemical reaction takes place in the rhabdom. This reaction is supposed to cause depolarization of the retinula cell soma which is conducted by passive propagation down through the retinula cell axon to the first synaptic layer, the lamina. Some of the retinular cell axons synapse at this point with second order neurons. The others, the so-called long visual fibers [85], continue on to the medulla.

The lamina is composed of elements called cartridges or neuro-ommatidia (there are the same number of cartridges as ommatidia) and can be considered as a relay station, accepting inputs from the ommatidia and sending outputs to the next station, the medulla. The long visual fibers, however, remain outside the cartridge [85]. Because of the geometry of the optics, each retinular cell has a slightly different field of view. It has been shown, however, that each cartridge accepts inputs from those retinular cells (in different ommatidia) having the same visual field [10]. In summary, there is a one-to-one relationship between the ommatidia of the retina and cartridges of the lamina. There also appears to be nearly a one-to-one relationship [58] between fiber bundles in medulla and laminar cartridges. Details of the medulla-lobula connections are not currently available.
In discussing the finer anatomy of the insect brain, Bullock and Horridge [12] mention that although much of the older anatomical work is useful for general naming of easily recognized topological areas, it is of limited usefulness for present day requirements for detailed information about particular units and pathways. In general, while a considerable body of knowledge is available on tracts, relatively little is known about individual paths [12]. In addition, some of the paths that have been described have not been directly observed, rather they have been inferred from tract observations. As for major tracts between opposite optic lobes, Bullock and Horridge report on the work of Power [68] and Satija [75] who describe two tracts. More recently Larsen [48] has described three tracts between opposite lobes; one between the anterior portion of the two medullas, and two between lobulas. He also mentions the possibility of a fourth between opposite lobular plates. Of particular interest here is the one which he calls the posterior tract of giant fibers, consisting of a small number of very large fibers. After leaving the lobular plate, the tract separates, one branch moving ventrally and posteriorly to the subesophageal mass and then out through the brain stem and into the thoracic nerve center. According to Larsen, the other branch ends in the central brain.

Strausfeld [79] has identified a similar system of giant fibers in Calliphora. Based on our electrophysiological findings,
he has attempted to trace this tract and has tentatively concluded that it runs from one lobular plate through the brain to the opposite lobula, as well as sending a branch down to the subesophageal mass (figure 3.2). We shall refer to this finding again as it relates closely to our work on pathway tracing via physiological rather than histological means.

**Physiology**

A variety of physiological approaches have been used in studying the visual system. These have ranged from optomotor response work through hook electrodes on the ventral cord, single and multiple unit recording from the optic neuropiles and brain to studies of single retinular cell responses. We consider first the optomotor work.

It has long been known that, when placed in the center of a rotating drum having alternating light and dark vertical stripes, the animal generally tends to follow the drum rotation. This fact has been used in making a number of measurements. One of the most common measures is that of yaw torque; the animal is suspended and his response (as evidenced by torque) to various drum patterns is measured. Other procedures include those of Reichardt [69] who used the beetle *Chlorophanus* in a series of investigations. Instead of measuring torque, the beetle was suspended and held a Y-maze globe in his feet. As
Figure 3.2  Diagrammatic representation of tentative location of class IIa-in unit [79].

(a) Horizontal cross-section through brain and optic lobes. Both members of a pair of mirror-image IIa-in units are shown. Arrows indicate direction of conduction.

(b) Vertical cross-section showing one IIa-in unit.
the beetle "walks" the globe he periodically must choose a right or left branching path. By counting the numbers of times the animal turns with and against the drum directions, a response measure is obtained. By using different combinations of spatial wavelengths and slits through which the moving drum was viewed, Reichardt [69] was able to develop a model for motion detection in the beetle in which he demonstrated that at least two ommatidia must take part in motion detection. Although the model accurately predicts certain phenomena found in Chlorophanus, such as the stroboscopic effect, it is difficult to identify model components with known anatomical elements in the nervous system. The experimental results of Reichardt's work showed that the interommatidial angle is quite uniform over the eye and that only adjacent or next-to-adjacent ommatidia interact to produce a response to motion.

For other species, such as the housefly Musca, blowfly Calliphora and fruitfly Drosophila, the situation is not so clear cut. McCann and MacGinitie [59] showed that for Musca, due to the relatively wide half sensitivity angle of the ommatidia, the magnitude of the negative reaction (torque opposite to drum rotation) was from 3 to 6 orders of magnitude less than the positive reaction. They went on to show how previously reported positive responses to patterns of wavelengths small enough to produce negative reactions were due to inaccuracies in the patterns
used. It has also been shown that motion detection is primarily due to adjacent ommatidia for *Musca*, with little or no interaction between ommatidia more than 2 or 3 spacings apart [22], and more importantly that responses to large field patterns were simple saturable summations of unit responses [22, 59]. The same workers found, as did Goetz [27] for *Drosophila*, that the effective sampling station derived from optomotor responses corresponded to the directly measured interommatidial angle. Goetz also showed that for the best image from a sampling station with a given light transmission there is an optimum ratio of interommatidial spacing \((\Delta \theta)\) to sensitivity angle \((\alpha)\). The existence of this optimum depends on the fact that the total flux through an element, which is proportional to \((\alpha \cdot \Delta \theta)^2\), must be greater than some threshold to excite neural activity [27]. For *Drosophila* the measured value of \((\alpha / \Delta \theta)\) was very close to optimal.

In other experiments on *Drosophila*, Goetz [28] considered both yaw torque and thrust measures of the optokinetic response. He showed the two systems were independent and that the torque motion detection system must simultaneously excite and inhibit opposite sides of the motor output system while the thrust system must simultaneously excite or inhibit both sides. This implies at least four contralateral and four ipsilateral connections, each capable of exciting or inhibiting one side of the motor system [28].
Goetz further showed this could be reduced to a minimum model with two contralateral and ipsilateral channels if common use were made of two orthogonal motion detection systems suitably oriented with respect to the animal's long axis. Similar results were obtained for Musca.

In summary, each retinula cell in an ommatidium looks in a different direction [47]. However, each neuro-ommatidium (cartridge) has inputs from retinula cells whose visual axes are parallel [10]. This agrees with the optomotor response experiments of McCann and MacGinitie [59], Goetz [27], Fermi and Reichardt [22], and others who show in essence that the effective sampling station for detection of motion is the ommatidial unit. Further, the information output by the cartridge is equivalent to that represented by the total average flux entering an ommatidium [60].

A number of investigators have also examined the neural response to motion, both at the sub-system level and in more detail at the single cell level. McCann and co-workers [60] measured the electrical response of the locust ventral cord fibers to moving stripes. The response was found to be essentially similar to the optomotor responses for this species and for Musca.

Single unit activity in response to visual stimuli has long been well established for both vertebrates and invertebrates. In
the vertebrate retina, ganglion cells have been found which are sensitive to movement of a spot of light in certain directions. The first demonstration of single units capable of coding optomotor responses in the visual system of Diptera was that of Bishop and Keehn [6, 7]. These investigators found units in the optic lobes of Musca, Calliphora, and Eucalliphora that were directionally sensitive and whose behavior as a function of spatial wavelength of a striped drum, drum velocity, and light intensity was, with one or two minor exceptions, the same as the optomotor response. The directionally sensitive units' response increased from a background rate of 5-15 spikes/sec to a maximum of 60-80 spikes/sec for motion in the preferred direction and decreased to 0-5 spikes/sec for motion in the null (opposite) direction (figure 3.3). This is in contrast to the vertebrate motion detection units in which no background firing is observed. In the vertebrate, motion in the preferred direction simply causes a higher discharge rate than motion in the null direction [2]. The fact that pairs of oppositely directed motion detection units were occasionally recorded simultaneously on the same electrode provides some clues as to the structural, and possibly also functional, organization of the motion detection scheme. Collett and Blest [15] also report finding oppositely directed units close together in the optic lobe of the moth.
Figure 3.3 Typical response of motion detection unit to striped drum (vertical stripes) rotating as shown. Abscissa is time in seconds. Ordinate is firing rate in spikes/sec (averaged over 0.4 sec bins).
The Form and Motion Detection Units

The motion detection units mentioned above, along with other units to be described, were investigated in more detail by Bishop, Keehn and McCann [8] and McCann and Dill [58]. In addition, a preliminary hypothesis to explain the motion detection mechanism was set forth.

The units examined by these researchers were classified according to the following criteria:

1. Anatomical location of unit.
2. Field of view.
3. Response to different classes of visual stimuli (see below).

As shown in Table I, classes I, II, and III refer to units found in the medulla, lobula-lobular plate and brain respectively (see also figure 3.1). For lobe and brain units, the letter "a" or "b" refers to units with contralateral or ipsilateral fields of view respectively. The suffix indicates the direction of preferred motion for these units. The stimuli were patterns projected onto the interior surface of a sphere and can be characterized by the following (see next chapter for details):

- **Intensity patterns**: circular spots of light, 20-60° in diameter.
- **Form patterns**: alternating light and dark stripes within a circular boundary.
<table>
<thead>
<tr>
<th>Recording Location</th>
<th>Class</th>
<th>Field Location</th>
<th>Field Size</th>
<th>Response Type</th>
<th>Details of Sub-classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medulla</td>
<td>I</td>
<td>ipsilateral</td>
<td>15-60°</td>
<td>dynamic form detection</td>
<td>Ia1 vertical edges</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ia2 horizontal edges</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ic non-selective</td>
</tr>
<tr>
<td>Lobula,</td>
<td>II</td>
<td>ipsilateral (b)</td>
<td>full eye</td>
<td>selective motion</td>
<td>IIa-in vertical edges moving</td>
</tr>
<tr>
<td>Lobular plate</td>
<td></td>
<td>contralateral (a)</td>
<td>or ca. 60°</td>
<td>detection</td>
<td>IIb-in moving inward from</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>periphery to center</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IIa-out vertical edges moving</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IIb-out outward</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IIa-up horizontal edges</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IIb-up moving up or down</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IIa-down</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IIb-down</td>
</tr>
<tr>
<td>Brain</td>
<td>III</td>
<td>ipsilateral (b)</td>
<td>same as</td>
<td>same as</td>
<td>IIIc nonlinear summation of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>contralateral (a)</td>
<td>class II</td>
<td>class II</td>
<td>IIIa, IIIb [8]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>binocular (c)</td>
<td>both eyes</td>
<td>same as</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>class II</td>
<td></td>
</tr>
</tbody>
</table>

Table I. Summary of classification scheme for units in optic lobe and brain.
motion patterns: as for form, but stripes moving within boundary in direction perpendicular to stripes.

The basic responses to these three classes of visual stimuli are shown in figure 3.4 for the IaI and IIa-in units. These responses are typical of all class I and II units in Table I. The responses shown are the average responses to a particular pattern suddenly presented at time zero. A summary of the class IIa-in response to motion as a function of pattern position is given in the contour map of figure 3.5. The coordinate system represents the orientation of the rows and columns of the array of ommatidial facets, with approximately 20 degrees between lines. The contour map shows the response of the IIa-in unit (firing rate) to a small pattern swept along a great circle trajectory [8]. Also shown are peak sensitivity points and directions for other motion detection units. For small (22°) patterns, the firing rate was shown to be approximately proportional to the cosine of the angle between the direction of pattern motion and the preferred direction of motion. For angles of ±90°, the rate approximated that for no motion. By using various combinations of patterns with edges moving in the same or opposite directions, these investigators also showed that the class II units performed a quasi-linear summation of the various motions present in their fields of view.
Class Ia1

Class IIa-in

Figure 3.4 Summary of basic responses of classes Ia1 and IIa-in to intensity, form, and motion patterns. Curves are averages of 20 repetitions of patterns presented at time zero. Abscissa is time in seconds from start of stimulus. Ordinate is firing rate in spikes/sec, averaged over binwidths of 0.1 sec for 4 sec and over 0.01 sec for 0.2 sec.
Figure 3.5 Contour map showing response of class IIa-in unit as function of pattern position. Also shown are peak response points and directions for other motion detection units. Coordinate system represents orientation of rows and columns of ommatidial facet array. Spacing is approximately 20 degrees.
In a further series of experiments [58] the same form pattern was turned on and off twice in succession, with the pattern displaced slightly for the second presentation. By measuring the response as a function of displacement, it was shown that for the class I units, the response was essentially independent of displacement, that the unit was responding more as a "novelty" detector. It was also shown that the response increased with number of edges and was orientation sensitive. It will also be seen in figure 3.4 that the increased response to form over intensity is transient only, lasting about two seconds. Thus, the unit was more a form detector than motion detector and because of its transient form response and novelty response was called a "dynamic" form detector.

The tentative hypothesis put forward [58] to explain the mechanism of motion detection was that the so-called large field units (class I and II units with 60° to full-eye fields) could be considered as receiving inputs from relatively large number of units having smaller fields of a much more local nature. The latter were called "elemental" units and two were proposed. The first of these was supposed to have a stronger response to form than to motion and to be sensitive to angular orientation. The Ia units were thought of as integrating the outputs of a number of such elemental form detectors. The second class of elemental units was supposed to integrate the output of a small
number, two to four, of elemental edge detectors and was appropriately called an elemental motion detector. Finally, the class II units were supposed to be operating on the outputs of a number of elemental motion detectors.

Although the elemental units were proposed partially as conceptual aids, some evidence, both anatomical and electrophysiological, can be adduced. Recent work in our laboratories [1] has disclosed the existence of units in the lobula and medulla having very small fields (on the order of a few degrees wide). Their responses to various patterns suggest them as possible candidates for the elemental units.

That the proposed theoretical structure is anatomically reasonable follows from the work of Braitenberg [10] and Strausfeld [80]. Braitenberg's work shows that there is a nearly one-to-one relation between the ommatidial facets and fiber bundles in the medulla. The insert in figure 3.1 shows a regular array of fibers running between medulla and lobula. These bundles could provide for transmission of local-field information from the area where form detection units were found to that where motion detection units were recorded from. It is possible then that this structure could account for the elemental form and motion detection processes. Strausfeld [80] has described cells traveling around the circumference of the medulla that could be the class I units summing the elemental
form detectors. He has also described units in the lobula [79] which could correspond to the class II units. We shall consider these units again in Chapter VI and discuss their relation to the results of the present work.

The foregoing description has considered primarily one-way processing (centripetal) in the visual system. That is, we have considered the processing chain from facets through retinula cells, cartridges, medulla and finally to the wide field selective motion detection cells, of which there were several classes. Although units with contralateral as well as ipsilateral fields have been studied, in essence only one "side" of the visual system has been examined. Strictly speaking, this is not quite true since the implied symmetry assumption that both sides were the same has been verified [8]. However, there has been no work on questions as to the relationship between similarly or oppositely directed motion detectors. For example, since both IIa-in and IIb-in units respond in the same manner to horizontal inward motion (motion from periphery inward in a horizontal plane) and have the same fields of view, it would be important to know if their activity were related in some way. Are they parallel units, each carrying essentially the same information? Or are they in series, with the IIb unit making synaptic contact with the IIa unit? Or, since their responses are so much the same, are they the same unit, recorded from two different
locations? The relative ease with which these units are found tends to indicate either a large number of units in parallel, or a very large widely branching unit. Since IIa units have been found in both lobes, one can also ask if there is any relationship between the IIa units in opposite lobes. If any interaction is to be found, one might expect it to be of an inhibitory nature, since each unit responds maximally to inward motion seen through opposite eyes (that is, the IIa unit in the left lobe would respond to right-to-left motion in the field of the right eye, the right lobe unit to left-to-right motion in the left eye). In their work on binocular units in the hawk moth, Collett and Blest [15] reported that on the one occasion when they recorded from oppositely directed units simultaneously, the units appeared either to have inhibitory connections between them, or to be fed by a common interneuron. Some progress toward answering such questions could be made with single electrode studies via latency or post-stimulus time histograms [35]. However, such evidence would be at best very indirect, and considering the wide variability in response, of extremely limited accuracy.

An approach using multiple recording electrodes would seem the best way of obtaining information pertinent to the class of questions mentioned above. It is the opinion of some workers in fact, that the simultaneous recording of several cells will in general yield information of much more value than is possible
from sequences of single unit recordings [35, 63, 66, 87]. In spite of the apparent desire for this kind of data, the literature on multiple unit and especially multiple electrode work is not overabundant [63]. Most multiple electrode recordings have been done using the mass (EEG) response. Moreover, much of the work that does exist on multiple "single" units has come about from recordings of several units on a single electrode [15, 25, 32, 70, 77], usually by chance, and has been primarily concerned with separating the original spikes into separate classes [6, 25, 77]. As a consequence, such instances of multiple unit recordings have usually been presented in the form of the original raw data and little has been said about any interaction beyond a brief description of the recording. Recently, however, with improving techniques, the use of two and more electrodes to investigate synaptic connections and multi-neuron interactions has become more frequent [9, 61, 64, 78, 82, 93]. Wyman, for example, has been quite successful in his investigations dealing with interactions among several motor output units in the dipterans and has been able to develop an interesting hypothesis explaining the apparent diversity of patterns observed in the flight motor system [93, 94, 95].

Other notable exceptions to the lack of multi-electrode work include Perkel et al. [66] and Strumwasser [78]. In experiments on the cochlear nucleus of the cat, it was shown that the interaction
between two units being stimulated could be predicted from the unstimulated interaction and the separate post-stimulus-time histograms of each unit [66]. Strumwasser was able to record from both symmetrically and asymmetrically placed units in Aplysia [78]. He showed that both members of a symmetric pair had similar firing patterns and in fact that they were fed from common sources. Also demonstrated was inhibitory interaction between neighboring units.

Functional Anatomy

It is clearly important to be able to relate the fibers observed in histological sections to those which produce the spikes observed on an oscilloscope during an experiment, that is, to relate structure to function. With two or more electrodes it is sometimes possible to obtain such direct anatomical information in addition to that relating to the effects of one neuron on another. Hughes [37] calls this technique "electrophysiological anatomy." That is, if one is fortunate enough to have a preparation in which one can record from the same unit at two points in space simultaneously, then electrophysiological pathway tracing is possible. Several benefits arise from this technique. First of course is the obvious anatomical result that a particular unit has (at least) an identifiable spatial spread, or that it runs between some set of identifiable points. This
possibility is especially important in extracellular recording since with this technique it is not possible to stain the recorded unit for later histological identification as is the case for some intracellular methods. Further, estimates of conduction velocity can readily be made. Finally, there is the somewhat intangible but nevertheless useful by-product of being able to give the anatomist some confidence about what to look for and where to look for it. He can be given hope that if he loses a recalcitrant fiber in one section, he will be able to find it again in a later one.

The use of this technique clearly depends heavily on the ease with which one can obtain simultaneous recordings of the same unit and since in general such recordings are difficult, reports of its use are not common. Although it is much easier in preparations with large readily identifiable cells [78], there are occasional examples of success with this method in the locust and optic lobes of the fly [35, 51].
CHAPTER IV

EXPERIMENTAL METHODS AND APPARATUS

Introduction

This chapter describes the experimental methods and apparatus, beginning with a description of preparation of the animal. Following sections describe the visual stimuli used, major components of the experimental environment, signal processing equipment, and computer hardware and software systems.

Specimen Preparation

The animals used were adult red-eyed (wild type) specimens of both sexes of the blowfly Calliphora phaenicia. The animals were laboratory bred and raised and were selected randomly from a cage containing a large number of both sexes at between 5 and 15 days post-emergence. Using a hot wire, the specimen was mounted on a ball-jointed stand and secured with dental wax in such a way that respiratory movements were not restricted. No anesthesia was used and except as described below, no surgery was done.

Access to the optic lobes and brain was obtained by tilting the head approximately $30^\circ$ forward with respect to the
thorax and removing an appropriate portion of the exoskeleton from the back of the head. The amount removed depended on the particular preparation and varied from a small triangular section to expose one lobe to approximately 70% of the posterior head exoskeleton to expose both lobes and brain. Ringer solution was put in the opening so made, and a platinum indifferent electrode positioned so as to make good contact with the Ringer solution.

Electrodes

Recording was extracellular, using for the most part, stainless steel microelectrodes with a tip diameter of approximately 1 micron. Occasionally, electrolyte filled glass micropipettes were used. The stainless steel electrodes were prepared according to a modification [6] of Green's [31] technique. Briefly, a stainless steel pin was acid-etched to the desired diameter and coated with Insul-X. Insulation at the tip was removed by placing the electrode in Ringer solution and passing current through it. The electrodes had a d-c impedance of 1-10 Megohms. Initial placement of the electrode was made by viewing the tip through a dissection microscope (approximately 50 power) as the electrode was advanced into the preparation by means of a 3 axis micromanipulator (Narishige Model MM-3). Final positioning was done by
monitoring the response on an oscilloscope and loudspeaker as the electrode was slowly advanced through the tissue. When a satisfactory signal-to-noise ratio had been obtained (figure 4.1), further tests to identify the unit were made as described in the results chapter.

Most of the experiments to be described made use of two electrodes. Preliminary studies used a single micro-manipulator with both electrodes on the same "bed." Each of the pair had individual fine control for x and y position and micrometer advance. This arrangement proved unsatisfactory because as the position of one electrode was being adjusted, mechanical vibration caused motion in the other, usually resulting in the frustrating experience of losing the cell on one electrode every time the other was adjusted. As a consequence it was almost impossible to find and keep two good units, and the setup was modified to include two completely separate manipulators. By mounting them independently on a half-inch steel plate, the mechanical interaction was reduced below that caused by the tissue itself. After preparation as described above, the mounted animal and micromanipulator assembly were placed on a stand in the center of a six foot diameter sphere, as illustrated in figure 4.2. The animal's position was then aligned to make the center of his visual field correspond to the co-ordinate system in figure 4.3.
Figure 4.1 Oscilloscope record of spike potentials for typical class IIa-in unit.
Determination of electrode position by dead reckoning is not satisfactory since as the electrode is advanced, it tends to drag tissue along with it [6]. Only approximate location was possible on viewing the preparation through a dissection microscope. That is, it was possible to determine whether the electrode was in the lobe or brain by inspection, and with experience, approximate position within those areas. However, as it was desired to know the position more accurately than this, the post-experiment localization technique described by Bishop and Keehn [6] was used. In this technique, iron from the electrode tip is deposited by means of passing current through the electrode (electrode positive, 5-15 \( \mu \text{a} \) for 10-20 sec). Perfusion of the specimen with 3% potassium ferrocyanide in Ringer solution results in a blue spot (5-15 \( \mu \) diameter) being formed which can be identified in histological sections. The obvious disadvantage of this method is that the unit being recorded is destroyed.

**Visual Stimuli**

Three sources of visual stimuli were used in these experiments. First, the sphere itself (figure 4.2) contained indirect light sources which provided constant uniform background illumination. Second, directly beneath the specimen stand was a gimbal-mounted projector by means of which various patterns were
Figure 4.2  Experimental environment showing sphere, central platform and micromanipulator system. The vertically striped pattern is from the gimbal-mounted projector, directly beneath the specimen platform.
projected onto the interior surface of the sphere. Using the
gimbal system, such patterns could be positioned anywhere in
the animal's visual field. The projector contained a servo-
motor system for imparting arbitrary periodic motion to the
patterns. Finally, two other projectors were used in certain
cases to provide for patterns with constant motion. For these,
a continuous film strip was moved at constant angular velocity
through the projector. All light sources had the color spectrum
of incandescent light.

All patterns used in this work had circular boundaries
and were, as described above, projected onto the interior
surface of the sphere. The three basic classes of patterns
were intensity, form and motion as defined in [58] and were
superimposed on a uniform background intensity $I_b$. For
completeness, the definitions are repeated here. In addition
to pattern position, the following parameter sets define the
three classes (see also figure 4.3):

1. Intensity patterns were uniform intensity circular
   spots of light. Defining parameters were diameter
   $(d)$ and intensity $(I_1)$. Because of the pattern boundary,
some form information was present, but was shown
[58] to have negligible effect for $d \geq 30^\circ$. 

Figure 4.3  Coordinate system for defining pattern position. Parameters for defining visual stimuli are also shown. Intensity patterns had uniform intensity $I_1$. Form patterns had stripes (wavelength $\lambda$) with intensities $I_0$ and $I_1$. Motion patterns had constant velocity $V$ in direction $\alpha$. 
2. Form patterns consisted of uniform parallel light and dark stripes within a circular area. Intensities of light and dark stripes were $I_1$ and $I_0$ respectively. Defining parameters were $I_0$, $I_1$, spatial wavelength ($\lambda$) and orientation angle ($\alpha$).

3. Motion patterns were similar to form patterns, but with the stripes moving within the circular boundary. Motion was constant velocity at right angles to the stripes. Defining parameters were $I_0$, $I_1$, $\lambda$, $\alpha$ and velocity ($V$).

We can summarize the defining parameters in the following manner, writing $I$, $F$ and $M$ for the parameter sets corresponding to intensity, form, and motion, and letting position be represented by $P$:

$$I = <P, d, I_1>,$$
$$F = <I, I_0, \lambda, \alpha>,$$
$$M = <F, V>.$$  

For some of the experiments it was necessary to use a sequence of alternating patterns. For example, it was often desirable to alternate motion in one direction with motion in the opposite direction. In such cases, each pattern was turned on suddenly, and after a given time, turned off. Rise and decay times of light on and light off were less than 3 msec. It has been shown [58] that if a sufficient adaptation period was
allowed between pattern presentations, the response to a given pattern was independent of the response to the previous pattern. Thus by presenting a repeating cycle of a sequence of patterns it was possible to study responses averaged over a number of cycles.

**Signal Processing**

This section contains a description of pertinent details of the signal processing equipment, the general organization of which is shown in figure 4.4. Excluded from the discussion is the computer system itself, which will be described in the following section. We shall further restrict the comments here to describing the characteristics and capabilities of the equipment and postpone to a later section a discussion of how various components fit into our data analysis methods.

Details of the amplifier chain are shown in figure 4.5. Because recording was extracellular, the signal strength at the electrode tip was of the order of 50 μvolts. Normally, an overall gain of 100 to 1000 would suffice for such signals. However, in order to obtain the signal strength needed for input to the A/D converter, additional amplification was used, resulting in a nominal overall gain of $10^4$. The preamplifier stage was a high input impedance ($10^9$ ohms) neutralized capacity amplifier (Bioelectric Co., NF 1). It was used in
Figure 4.4  Schematic diagram of signal processing equipment.
<table>
<thead>
<tr>
<th>Function</th>
<th>From electrode</th>
<th>Gain</th>
<th>Frequency Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal pickup</td>
<td>Preamplifier</td>
<td>2</td>
<td>0.2 Hz - 10 KHz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(nominal)</td>
<td></td>
</tr>
<tr>
<td>Primary amplification</td>
<td>Wide band amplifier</td>
<td>100</td>
<td>80 Hz - 10 KHz</td>
</tr>
<tr>
<td>noise reduction</td>
<td>60 Hz notch filter</td>
<td>1</td>
<td>350 Hz - 4.5 KHz</td>
</tr>
<tr>
<td></td>
<td>Band pass amplifier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>secondary amplification</td>
<td>Variable gain amplifier</td>
<td>40</td>
<td>20 Hz - 15 KHz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(nominal)</td>
<td></td>
</tr>
</tbody>
</table>

To oscilloscopes, etc.

Figure 4.5  Schematic diagram of amplifier chain.
its a-c coupled mode with a frequency response extending from 0.2 Hz to 10 KHz. The remaining components of the amplifier chain, with the exception of the wideband amplifier (Tektronix Inc., FM 122) were constructed in our laboratories. Completely separate, identical chains were used in all experiments involving two electrodes.

Three different oscilloscopes were used in this work (figure 4.6). First, a dual beam oscilloscope (Tektronix, model 565) was used to display individual spike waveforms (figure 4.1) and as a preliminary measure of synchronzation (see Chapter V). A second oscilloscope (Tektronix Inc., RM 561A, with 4-trace input amplifier, model 3A74) was used for monitoring electrode responses and stimulus signals. Finally, a third unit (same as previous one, except with a storage CRT, model RM 564) was occasionally used to provide a preliminary estimate of average firing rate by connecting the electrode signal to the oscilloscope through a simple R-C integrating circuit.

The other devices in the Analog and Special Purpose section of figure 4.4 are sufficiently specialized that their description is best postponed until the particular experiments in which they are used are discussed (Chapter V).

For the majority of the studies described herein, signals were sent directly to the computer during the experiment and stored on disk memory in digital form. In other words, the
Figure 4.6  Analog display oscilloscopes. Shown (from top to bottom) are storage, monitor, and dual beam oscilloscopes. The signals being displayed are from a dual-electrode experiment. Synchronization of spikes is shown on dual beam CRT, in which both traces are being triggered from upper signal. Signals are being played back from FM tape recorder, part of which can be seen in the bottom of the photograph.
computer system was "on-line," and the data were available immediately for (digital) pre-analysis (see Chapter II), or sample-analysis. It was occasionally desirable, however, to record the data in analog form as well. A seven channel FM tape recorder (Ampex, model FR1300) was available for this purpose. Signals could be tape-recorded and sent to the computer simultaneously, or either storage media could be used separately.

As will be discussed in Chapter VI, transmitting data directly to the computer was found to be much more convenient than the two step procedure of recording on tape and transmitting data at a later time. However, because of occasional unavailability of the computer, whether due to machine breakdown, competition from other users, or the need to have a back-up record (computers have been known to lose data), the latter procedure was sometimes used. In addition, it was sometimes necessary to send only partial information directly in order to obtain a preliminary result rapidly, postponing complete analysis until after the experiment. In such cases, a complete record was put on tape while simultaneously abstracting certain information (for example, spike firing times) to be sent directly to the computer.

The A/D conversion and digital preprocessing stage (LORI) was in essence a special-purpose computer designed and constructed in our laboratories. Because of the importance of this device as
the primary link between experiment and central computer, we consider briefly some of the design criteria before presenting a "numerical" description of its capabilities. The basic design goal was to produce an A/D conversion system that would have, in addition to the standard facilities, special capabilities for handling neurophysiological signals and could do some preprocessing of such signals in order to reduce the amount of time spent by the central computer on data reduction chores.

For more detailed criteria, it is necessary to consider the kinds of data to be transmitted to the computer and the data rates and volumes involved. For stimulus signals, two classes of signals are generated. First, is the standard continuously varying signal which can be treated by continuously sampling the signal at a suitable rate. Second, for repetitive presentation of a stimulus, often all that is required are the times of stimulus presentation. Certainly the previous method could be used, using the square wave generated by a photocell, for example. However, if it were necessary to know the time of presentation to within a millisecond, the sampling rate would have to be 1000/sec. Thus, for a stimulus presented once per second, 1000 data samples are necessary for each useful number, which is extremely wasteful of both the time and memory of the central computer. This problem was solved by putting digital clocks in LORI such that a suitable pulse could cause the clocks to be read and the resulting time-of-event (TOE) to be sent to the computer.
For response signals, there are again two basic types: continuously varying responses and spike potentials. The former are treated simply by continuous sampling. For the latter, there are two cases: only the time of the spike is of interest or both the time and waveform are of interest. If only the time is of interest, the digital clocks can be used as for the stimulus signals by means of threshold detection circuits. If the waveform is needed, we are faced with the same problem as in the case of the stimulus in that the signal must be sampled continuously for long periods of time relative to the time periods of interest.

To solve this problem, special circuitry was built into LORI so that at the time of threshold crossing, not only was the TOE sent to the computer, but also the ADC was started and a preset number of waveform samples were sent to the computer. Addition of a recirculating memory to LORI provided for the inclusion of a number of samples prior to threshold crossing in this group.

Let us now examine briefly what typical data rates and volumes are. Assume that two electrodes are used, both extracellularly, that one is recording from a single unit and the other from two units simultaneously. Assume further that two visual patterns are used, that the position of one must be monitored continuously at 50 samples/sec and that for both the time of presentation is required. For the stimuli then, approximately 50 samples/sec are necessary. For the responses,
assume a maximum firing rate of 200 spikes/sec over short intervals (on the order of a second or less) and an average of 20 spikes/sec over long intervals (several minutes to an hour).

For the electrode recording from two units, waveform information will be necessary in order to be able to distinguish the units. Assume that 20 samples per spike are used. These figures lead to short and long time rates of approximately 8000 and 900 samples/sec respectively, and to just over $0.5 \times 10^6$ samples for a 10 minute period. If each electrode is recording from a single unit and waveform samples are not needed, the rates and volume are reduced by approximately a factor of 10.

Note that the foregoing rates are both averages. For waveform sampling, if the spikes are approximately one millisecond in duration, the conversion rate must be between 5K and 20K samples/sec for the duration of the spike.

In detail, LORI's capabilities are as follows. The central components were a high speed (500 KHz) ADC, a 10-channel multiplexer, and a number of digital clocks. Each of up to 10 inputs to LORI was processed in one of the following three ways as illustrated in figure 4.7.

1. Continuously sampled. Signals can be continuously sampled at user-specified intervals ranging from 10 µsec to 1 sec. Different signals can be sampled at different rates.
Figure 4.7 LORI modes of operation. Upper line represents input to LORI from amplifier chain. Also shown are two possible thresholds. Next line represents output of A/D converter in continuously sampled mode. Following two lines show crosses for each event (spike) at the time of threshold crossing (Time Of Event or TOE) for Threshold 1 and 2 (top line) respectively. Bottom line represents a combination of A/D and TOE modes. See text for details.
2. Time-of-event. As described above, times of threshold crossings were read from digital clocks which were run at preset rates, usually at 50 KHz.

3. Combination mode. At the time of threshold crossing both the time and a preset number of samples are sent to the computer. From 1-56 samples per waveform can be used. The sampling rate for the waveform is independent of the clock rate for the TOE's. This third mode was given the unfortunate acronym TOAD, for TOE and A/D.

In addition to the three major modes described above, external threshold circuits could be used in conjunction with the digital clocks to produce TOE's. This was the method normally used when it was clear that only a single unit was being recorded. In this case, a pulse generated in the oscilloscope threshold circuit caused the digital clock to be read and the resulting TOE to be sent to the computer.

The digital clocks used held 27 bits (equivalent to between 8 and 9 decimal digits). Thus, with the normal clock rate of 50 KHz, spike times were accurate to 20 µsec over a period of up to 45 minutes without recycling the clock. The necessity for this order of accuracy arose from the need to measure times between spikes on different electrodes to within 50 - 100 µsec over intervals of up to 10 - 20 minutes.
Computer System

In this section we describe the computer hardware and software, beginning with the machine configuration. Figure 4.15 shows in block diagram form an overview of the entire system, from experiment through central computer (IBM 360/44 [40]) to display terminal. Details of the computer configuration are shown in figure 4.8. It should be evident that a simpler configuration would suffice for the applications considered here. The reasons for choosing the more powerful configuration were twofold. First, the system was to be used for other purposes in addition to neurophysiological research. Second, as was discussed previously, the system itself was to be considered an experiment and the larger configuration made the tasks of developing and modifying the system much easier.

It will be observed that the bulk memory is large, and consists entirely of disks, with no tapes. The large volumes of data being generated by several experimenters necessitated extensive bulk memory capabilities, while the reason for choosing disks derives from the design criteria developed in Chapter II, one of which stated that all data were to be kept in a common pool. Therefore, in order to keep data access times acceptably short, some form of random access bulk memory was necessary and disks were used exclusively. However, to minimize the consequences of inevitable occasional errors, disk contents were periodically copied onto tapes via a tape unit that could be shared with another machine.
Figure 4.8  Schematic diagram of computer hardware configuration.
The display unit was an IBM 2250 graphic terminal, located close to the laboratories. It has a 21-inch CRT on which information for the experimenter and results are presented in the form of text, graphs, etc. The researcher communicates with the system via the display terminal using its associated keyboard (a small panel with 32 "function keys" or push buttons). Additional means of communication are provided by a special panel of 11 potentiometers and 16 toggle switches, designed and constructed in our laboratory. In all instances, the number of steps necessary to make a request of the system is small, the actions themselves are not unlike those performed on more standard pieces of laboratory equipment.

We turn now to a description of the software system. As illustrated in figure 4.9, there are several software levels. The reasons for this modular approach, as opposed to a single system covering all levels, were disparity in user groups and ease of construction. Broadly speaking, there were two classes of users: the neurophysiological researchers who needed a working tool, and investigators in information science who needed access to the machine at a more basic level. Needless duplication was thus avoided by constructing several levels of systems rather than a number of complete and different systems.

Excluding the system of application programs, details of the other systems appear elsewhere [13, 19, 46, 54] and, with
Figure 4.9  Block diagram representation of software systems.
the exception of Phase 2, we present here only a cursory
description. Because the application system is heavily
dependent on its capabilities, Phase 2 is treated in somewhat
more detail. The lowest level is the basic operating system
[13] which provides an environment for programming research
and development of higher level systems. Furnished are support
of input-output devices, interrupt handler, etc. One of the
higher level processors is the Advanced Programming Support
(APS) system [46]. Rather than an executive system, it was
designed expressly to be a set of facilities for constructing
on-line systems to manipulate large collections of data by means
of a graphic console. Built on top of the basic system, using
the facilities of APS, are the two systems of direct interest to
the neurophysiological investigator. These are PLEXUS, a
neural network modeling system [19], and Phase 2, the experimental
data acquisition and analysis system [54]. In the sense of
figure 4.9, both systems are on the same level and can communicate
with each other as will be described below.

Using PLEXUS, the user may construct the topology of
and specify parameters for a neural network, simulate operation
of the network, and store and retrieve networks from a library.
Analysis of network output is performed using Phase 2. The
system operates in a conversational mode and interacts directly
with the user via the graphic display terminal. One of the
important features of PLEXUS is that it will accept experimental data. That is, a network can be "driven" from data obtained from real (as opposed to simulated) experiments, thereby creating a direct link between model and experiment.

The realization of the experimental system concept discussed in chapter II is the Phase 2 system, overall criteria for which were developed there. Final design specifications were, in brief:

1. To provide the user with an interactive, responsive system capable of expressing results in graphical form.
2. To keep communications between user and system as simple as possible to minimize the amount of training required.
3. To provide a large number of basic analysis techniques in order that the user might have a wide choice in building a tentative analysis strategy.
4. To provide means for handling large bodies of data conveniently.
5. To permit operation in two modes: real-time for examination of data being collected (corresponding to our pre-analysis), and non-real-time for more complete analysis (we have divided this into sample- and post-analysis).
It is convenient to think of Phase 2 as being composed of an underlying system and a collection of analysis programs (called "processes"), the system providing an environment for, and control of, analysis processes. The term "process" is used to distinguish these entities from complete analysis algorithms. That is, the processes are building blocks with which the user constructs a particular analysis strategy. This, of course, does not exclude the possibility of a single process being complete in itself.

Because of the decision to keep communication between user and system as simple as possible, there are essentially two different user groups. After Lockemann and Knutsen [52], we give the name "programmers" to those actually designing and writing analysis processes, and "users" to those making use of the facilities so provided. It is possible, of course, as in our own case, for an individual to be active in both categories (while this has decided advantages, it also provides one with some particularly frustrating experiences). Another result of the same decision is that the nature of the interaction between the user and his data is almost entirely a function of the application process. Thus, while the system furnishes tools for doing so, the actual task of providing the user intimate contact with his data and the ability to manipulate it conveniently falls to the designer of the application processes.
As a basis for discussing in a later chapter some of the problems faced by both the user and process designer, we present here a highly simplified description of the programmer's view of Phase 2, based on our experience with the system and on the detailed information in references 52, 53, and 54. The two basic components of the system, in addition to the graphic facilities of APS, are a multiprogramming monitor and a group of data management routines. Once the user has constructed a particular strategy from a group of processes (see example below), the monitor supervises execution of the group, passing control from one process to another as dictated by the status and requirements of the different processes. A process may deliberately return control while waiting for some user action, and will lose control when no data are available and occasionally during normal data access. Thus, there is no system-provided mechanism whereby the programmer can specify that one process is to be a subroutine of another. It is possible, although not simple, for the programmer to provide the user with limited subroutine-like or hierarchical structures. To do this requires essentially bypassing the system.

Data are stored and retrieved in files called "datasets" by which is meant a group of data items referenced under a single symbolic name assigned by the user when the data are generated. Dataset management routines provide convenient symbolic access to the data and relieve the programmer of most bookkeeping chores. They also supply facilities for associating attributes,
such as sampling rate for continuously sampled data, with a given dataset and for creating some internal structure. A dataset may, for example, be composed of a number of "subsets." Dataset names are always entered by the user when he first asks for a given process (see example below), and cannot be assigned during execution of a process, nor can they be created by a process itself.

We shall describe the user's view and general capabilities of Phase 2 by means of examples. The elements the user deals with are analysis processes, datasets, and parameters. The latter two are thought of as operands, the former as operators. Our first example is a contrived one and will illustrate the use of these elements and of one of the basic display processes.

Suppose the user has two continuously sampled datasets (STIM1 and STIM2, say) in the machine and wishes to examine them and to form their sum. He then constructs a suitable program out of a sequence of processes. Each process is added to the sequence by depressing its assigned function key (as each new process is first put into the system, it is assigned a unique function key). As each key is depressed, header messages, each containing a brief descriptive title and a list of required input and output datasets and parameters, are displayed as shown in figure 4.10 (a) for our example. The first process, SUMATD, performs the sum

\[ \text{SUM} = A \cdot \text{STIM1} + B \cdot \text{STIM2} + C, \]
Figure 4.10 Steps in forming sum of two continuously sampled signals.

(a) Headers for processes SUMATD and SCAN.
(b) Headers with dataset names and parameter values filled in by user.
(c) Resulting display. Upper curve is sum of lower two. DT is the sampling interval. For clarity, curves 1 and 3 have been shifted up. Other details in text.
where A, B, and C are constants (parameters), STIM1 and STIM2 are input datasets, and SUM is the result or output dataset. Since all three are to be examined, all three names are typed as inputs to the display process SCAN. Note that the output of one process can be input to another, giving the user facilities for composition of building blocks (figure 4.10(b)). Having checked the names and parameters, the user presses another function key to begin execution of the processes, with the result as shown in figure 4.10(c).

The process shown here, SCAN, is a flexible general purpose display process, capable of displaying up to four inputs or datasets simultaneously. The inputs may be of the continuously sampled variety, TOE, or TOAD. The latter two can be displayed either as average firing rate or as a series of dots, each representing a single spike. The averaging is over a binwidth which the user may vary. The central part of the display may be thought of as an infinitely variable window through which one looks at a strip chart recording. The window size (that is, the scales) may be varied for each curve separately or for all curves simultaneously. Similarly the curves may be easily and rapidly shifted back and forth or moved up or down. Shown in the lower right part of the display are the input names. The associated numbers (N, SS) refer to subset SS of input number N. For identification purposes, the input number is
displayed on the appropriate curve. In the upper part of the
display are the attributes of the dataset whose name is boxed,
along with other information such as the number of spikes for
TOE or TOAD data, the number of points used in a moving linear
average (AVG), and the binwidth for the current display. The
latter two may be altered simply by typing in new numbers.

Light-penning members of the list of words at the
bottom of the display causes appropriate actions to be taken.
By light-penning ONOFF, curves can be turned on or off. The
mode of TOE datasets may be changed from average firing rate
to dot mode or vice-versa (MODE). If the process for computing
interspike interval histograms was asked for at the same
time as the SCAN process, such histograms can be produced
(HIST). On light-penning a curve, one of the following two
actions takes place: (1) the number identifying the curve is
moved to the light-penned point; (2) horizontal and vertical
lines are drawn through the light-penned point and coordinates
of the point are displayed. In figure 4.10 (c) for example, the
coordinates are (62, 0). Light-penning VALUE an even or odd
number of times determines which of the above two actions is
taken. The words TOE and SUBS are associated with creating a
new output dataset from selected portions of input datasets—that
is, with facilities for editing input data.
The next example is presented in the context of an experiment in which we suppose the object is to investigate possible interaction between two neural units. We assume further that it is the interaction between two specific classes of units that is to be considered.

The first stage will, of course, consist of obtaining two units and making preliminary tests to see if they are the classes in question. The preliminary tests will probably be made using relatively crude stimuli and might be evaluated simply by monitoring oscilloscope and loudspeaker outputs. It may be desirable, however, to obtain a somewhat better measure, or may be necessary to view firing rate over a longer interval than is convenient with an oscilloscope. The latter being the case, the investigator could make use of the real-time version of the process SCAN. The major difference is that the firing rate display is updated continually at a rate specified by the user (usually between 1 and 20 times per second). The rate may be varied while the process is running. The result is similar to moving a film past a viewing window, but with window size controlled by the user. This process is clearly in the pre-analysis category.

The next stage of the experiment is to determine whether or not there is any interaction between the units during background or spontaneous firing. This is accomplished in two steps. The
processes forming the first step are shown in figure 4.11 (a). Here, two copies of the acquisition process LCOLLECT have been specified, one for each electrode, along with the display process SCAN. After execution is begun, each copy of LCOLLECT presents the display shown in figure 4.11 (b). The user specifies which LORI channels he wants, data types, etc. After filling these in, the system is ready to start data collection. After a short initial period during which the first data are entering the system, SCAN will display the first portions of the records, and the user may change scales, shift the records, etc., while data collection continues. By watching the spike count displayed by SCAN, and evaluating the amount of variation in firing rate, the user can determine when sufficient data have been acquired for the succeeding step. Note that SCAN is being operated in essence as a pre-analysis process. Data collection is terminated by pushing a button on LORI, after which the firing rate records may be examined in detail (figures 4.11 (c), (d)). Here SCAN is used as a sample-analysis process.

If the records are acceptable, the second step is the computation and display of a cross-correlogram (figure 4.12). The correlogram display provides manipulation facilities similar to those of SCAN. In addition, the correlogram may be recomputed with a different resolution (the binwidth of the
Figure 4.11  Steps in data acquisition and sample-analysis.
(a) Headers for data collection from two electrodes and simultaneous display.
(b) Channel specification display. User sets LORI channel number, data type, starting time, and sampling rate.
(c) Firing rate versus time (spontaneous firing) displayed via SCAN while data are being collected.
(d) Short section of (c) in single spike or "dot" mode. Each spike is shown as a single cross. With narrow viewing "window," the precision is automatically increased.
Figure 4.12  Cross-correlogram showing interaction between the two units of figure 4.11. Abscissa is lag in msec; ordinate is number of coincidences.
correlation histogram) simply by typing in a new value. The ease and rapidity (computation and redisplay for trains of a one or two thousand spikes requires only a few seconds) with which new resolutions may be obtained has proven to be of immense value. As will be seen in the results chapter, for different experiments the resolution required varied over three orders of magnitude.

Note that the correlation process is being used as an existence test in sample-analysis mode. If no peaks or valleys were observed, the electrodes would be repositioned and the procedure restarted.

The next stage might be to investigate the nature of the observed interaction, for example, by presenting a repetitive sequence of alternating visual patterns. The setup would be similar to that shown in figure 4.11 (a), but with a third copy of LCOLLECT for stimulus data. At the end of the run, the result would appear as in figure 4.13 (a) which shows the entire record. The user may now look at selected parts of the record (figure 4.13 (b)), check the validity of the responses, examine details, compare the two responses, etc. To determine which stimulus variations should be presented next, the user might wish to examine the average response. Two processes which compute post-stimulus-time (PST) histograms [26] are available. One of these, since it makes but a single pass
Figure 4.13 Responses to repetitive stimulus presentation.
(a) Responses to entire run. Dots show times when stimulus was turned on. Curve 1 has been shifted up for clarity.
(b) Close-up of single response.
(c) Average of all presentations.
through the data, is quite fast and thus suitable for quick-look or sample-analysis operations, and would be used at this point if the number of spikes were large. After evaluation of the PST histogram, new stimulus conditions, perhaps dependent on the evaluation, can be set up and the procedure repeated. The other PST process, while it takes considerably longer (an average of \( N/2 \) complete passes through the spike train must be made, where \( N = \) number of stimulus repetitions), is more accurate and produces an output that can be manipulated more flexibly via SCAN. This latter process would probably be used during post-analysis where it would be applied to all experimental runs in which repetitive stimuli were used. The resulting averages could then be examined and compared via SCAN (figure 4.13 (c)). Should errors have occurred during some of the stimulus presentations, the editing facilities of SCAN provide for easy removal of the offending sections. This procedure, too, would likely be used during post-analysis since unless several errors occurred, the average of all repetitions would be accurate enough for a quick qualitative evaluation.

Our next example is a method for separation of multiple units recorded on a single electrode. Particularly when recording extracellularly, a single electrode may pick up responses from several units. This situation occurred not infrequently in our
work and similar reports in the literature are not uncommon [25, 70, 77]. It is thus desirable to have procedures available for separating the spikes so recorded. While we make no claims for uniqueness in the separation algorithms themselves--the present ones are very simple--we believe the process to be described offers a number of advantages to the user by reason of its interactive, iterative nature.

For this process, the assumption is made that a "single" unit can be defined on the basis of spike waveform, that is, that spikes from different units have different waveforms. Elaborate spike separation techniques have been developed using methods adapted from communications theory [43], but application of such techniques is cumbersome and expensive. Since in the present case, as in many, there were but a few dominant sources of spikes of interest, our procedure uses simple methods and emphasizes interaction between experimenter and machine. This procedure assumes that it is sufficient to characterize the spike waveshape by a few simple properties such as position of maximum and minimum, or difference (called spike "height") between these two. If based on a small number of such characteristics, spike separation can be done quickly. Obviously, for any single one of these properties, separation can be done without computer aid. Our method was designed to take advantage of the experimenter's intuition and
judgment by providing him with a number of different properties among which he could choose at any time during the analysis. Although only three characteristics are presently available, adding more is trivial should the need arise. Since the method allows one to explore various alternatives, it is primarily employed during post-analysis.

The spike separation procedure consists of five major phases which may be performed repeatedly in any order.

1. Selection of characteristic on which to base classification (figure 4.14 (a)). The experimenter starts by choosing one of these characteristics or, if he wishes more information before choosing, he may EXAMINE the recorded signals by asking for the display in figure 4.14 (b), which is similar to the familiar film strip recording. Here spikes are represented by vertical lines (bars) between maximum and minimum points of the complete spike waveform. By light-penning the appropriate message, the strip may be moved or (de) magnified. Complete waveforms may also be displayed (figure 4.14 (d)). The experimenter may alternate between this mode and the bar form (SPIKE, TIMEHIST).

If there is a known relationship between stimulus

---

1 Completely capitalized words indicate an action taken by the computer when the corresponding word on the computer display is light-penned.
Figure 4.14 Spike separation procedure.

(a) The line indicates that the user has chosen to separate on the basis of height.

(b) Display of spikes in bar form before classification. If a stimulus curve were present, it would be displayed on the top line.

(c) Spike height histogram. Abscissa is height in arbitrary units. Ordinate is number of spikes in given height bin (of width 14 units). Spikes in the right peak (height 844-1200) are assigned the name IIAIN. Had MAX-MIN been chosen, histograms of positions of maximum and minimum would appear in upper and lower boxes.

(d) Post-classification examination of waveshapes.
and response(s) the experimenter may employ the bar form to determine the characteristic best suited for distinguishing between spikes. Typically, however, a combination of both display modes is used.

2. Identification of spike sources. After having chosen the characteristic, the user must specify ranges of allowable values of the characteristics for each source. To this end, a histogram is produced for each characteristic (figure 4.14 (c)). If the characteristics were chosen suitably, the histogram should exhibit distinct peaks, each corresponding to a unique source. To assign all spikes falling within a given peak to one of the data sets named on the display, the user light-pens the appropriate name and uses potentiometers to move two vertical lines across the histogram(s), placing the lines so that they enclose the peak. Limit values corresponding to line position are displayed opposite the name.

3. Evaluation of the selected property. If the histograms have no distinct peaks, the user may RETURN to stage 1 and try another characteristic, or may EXAMINE the spike train again.
4. Classification. Light-penning CLASSIFY results in spikes being sorted on the chosen characteristics and their time of event only being saved in the proper dataset. On completion of classification, the user may EXAMINE the results. Appropriateness of classification may be judged on the basis of:

(1) similarity of shape of spikes in any one dataset.
(2) comparison of any pair of datasets.
(3) comparison of any dataset with stimulus.

Spikes not so far classified may also be displayed.

5. Post-classification examination. Three alternatives are possible:

(1) Results unacceptable. Process repeated by RETURNing to stage 1 and RESTARTing (all previous results erased).

(2) Results acceptable but some spikes remain unclassified. RETURN to stage 1 and ITERATE (previous results are saved and the procedure is repeated for unclassified spikes).

(3) Results acceptable, STOP. Classified spikes may now be used for further analysis. Spikes still not identified are available for treatment by more complex methods of spike classification.

When histograms of characteristics show distinct peaks, the above method is usually sufficient. If such clear features are
not exhibited, it may be necessary to add operational criteria to evaluate any given classification. In one such case [18], it was discovered that, if the limit values for heights were suitably chosen, certain classes showed pronounced cyclic activity as evidenced by peaks in an autocorrelogram. Classification and autocorrelation were thus used together, iteratively, to finally separate spikes into distinct classes. Other such methods will be discussed where appropriate in the following chapter.

As mentioned in Chapter I, one of the aims of this work was to develop a system of application processes that would be of use to the neurophysiological investigator. Thus, as the need arose, we designed and implemented various additional processes during the course of this work. These will be described at suitable points in the following chapter.

We conclude with a diagrammatic summary of the data acquisition and analysis system as shown in figure 4.15. Both hardware and software systems are illustrated.
Figure 4.15 Summary block diagram of hardware and software systems. Double arrows indicate user interaction. Single arrows indicate data paths and control lines. All hardware except central computer is located in or near laboratory.
CHAPTER V

RESULTS

Introduction

This chapter presents the results of two series of experiments in which the data processing system described in previous chapters was fully utilized in studying the processing of visual information in the insect central nervous system. Both series of experiments were concerned with the motion detection units described in Chapter III. The first section presents results on tracing the path or fiber tract of these units from one lobe to the other by means of simultaneous microelectrode recordings. The second series of experiments was an investigation as to whether or not there is interaction between different motion detection units.

As mentioned in Chapter I, one of the objectives of this research was the development of computer-based methods for analyzing electrophysiological data. The computer programs and analysis methods described in the previous chapter are quite general and are widely applicable. As more information was obtained on the neuronal units being studied, this information was taken advantage of in using these programs in specific ways. It was also used as a guide in modifying existing programs and
in developing new programs and methods to help solve particular problems and meet new demands arising from later work. Parts of this chapter therefore consist of descriptions of various methods which were developed during the course of this work. One of these was a set of procedures for classifying multiple units recorded from a single electrode, one technique for which was described in the preceding chapter. Logically we should first present a description of the spike separation procedures, since those available at a given time were, when needed, a standard part of all analysis procedures. However, one of the techniques depends on the pathway tracing results. In order to be able to present a more unified treatment of the spike separation procedures therefore, we first describe the pathway tracing experiments, then the spike separation procedures. Following the latter, we shall present results from the study of interaction between units.

Pathway Tracing

This study first started as a search for interaction between various pairs of the class II and III motion detection units. Initially we expected that corresponding pairs of class II and III units or class IIa and IIb units in opposite lobes (for example, a class IIIa-in and class IIa-in, or IIb-in and IIa-in) might be correlated since all of their identifying characteristics—except for recording location—were the same. We therefore expected that the curves of firing rate versus time for simultaneous
recordings from such pairs would be similar. Thus if the units were interacting, the interaction would show up as a rather broad peak in the cross-correlogram (on the order of a few milliseconds in width at a lag of at least a few milliseconds) as for example in figure 4.12 or as discussed by Perkel [67]. Contrary to expectations, however, many cases were found in which from 80 to 100 per cent of all spikes from both electrodes had the same time history except for one spike train being shifted in time with respect to the other by a few tenths of a millisecond. Figure 5.1 illustrates this "synchronization" of spike trains. In this example the two electrodes were recording from a class IIa-in unit in the left lobe and a IIb-in unit in the right lobe. The only stimulus was a constant low level of background illumination. Although only a single curve is visible in figure 5.1(a) there are actually two, superimposed. Figure 5.1(b) shows a short section of these two curves with the mode of display changed to dot mode so that each spike appears as a cross. (On the CRT, each spike appears as a point or a dot, hence the label "dot mode" for this form of display. Because of an idiosyncracy of our plotting routines, points appear as crosses on the figures here.) The synchronization of spike trains is apparent here, but the interval displayed is too long with respect to the time delay for the latter to be apparent. A still shorter interval is shown in figure 5.1(c) with the times of two spikes displayed. Observe that the upper spike is delayed by 0.64 msec. Similar measurements on other spike pairs yield the
Figure 5.1  Synchronization between spike trains. Units were IIa-in in left lobe, IIb-in in right.
(a) Average firing rate versus time. Binwidth = 0.5 sec. There are actually two curves superimposed.
(b) Short section of (a) in dot mode. Upper line is IIa-in.
(c) Short section of (b) showing time delay. With shorter time window, precision of time axis is automatically increased.
same delay, plus or minus about 50 $\mu$sec.

Thus, contrary to expectations, the correlogram peaks found at this stage were very sharp (in the fractional millisecond range) and occurred at lags of a millisecond or less. Figure 5.2 illustrates the use of the correlogram to measure the overall average correlation or delay between spike trains. Shown here is a sequence of correlograms for different resolutions. The data used were from the same units as for figure 5.1 but the stimulus conditions were changed. In addition to background illumination, there was a motion pattern in the field of the right eye. Except for actual ordinate values, correlograms for the unstimulated case were identical to those in figure 5.2.

The abscissa of the correlogram is lag time in milliseconds, the ordinate is number of coincidences for a given lag time (actually for a lag "bin" of width RESOLUTION). The fact that the correlogram has non-zero values for large lag times for the two larger resolutions simply reflects the non-zero firing rate of the two units. The flatness of the correlograms for the largest resolution, except for the peak, shows there is no other interaction and essentially reflects the random nature of the distribution of interspike intervals. Had the two spike trains been exactly periodic, say with intervals of $T$ msec between spikes, then the correlogram would have consisted of a series of peaks $T$ msec apart.
Figure 5.2  Sequence of cross-correlograms showing appearance of synchronization between spike trains. Correlation of IIb-in in right lobe with IIa-in left lobe. Resolutions for (a), (b), and (c) are 2.51 msec, 0.51 msec, and 0.05 msec, respectively.
The tight synchronization of the two spike trains suggested that the electrodes were recording spike potentials from the same neuronal element at two different locations some distance apart. If this hypothesis were true, the correlogram peak would simply indicate spike travel time along the unit. The synchronization could also be observed directly on the oscilloscope as shown in figure 5.3. Signals from both electrodes were displayed on the dual beam oscilloscope, with both beams triggered from the lower trace. The records are the same as those used for figure 5.1.

It will be seen that the spikes on the upper trace (class IIa-in unit, left lobe) are delayed by approximately 0.6 msec from those on the lower trace (class IIb-in, right lobe). Notice that with this technique the absence of even a single spike on the upper trace corresponding to one of the "trigger" spikes will show up as a relatively straight line. Triggering the traces in this manner demonstrates that for every spike on the lower trace there exists a corresponding synchronized spike on the upper trace. Triggering both traces on the other beam shows the relationship to be one-to-one.

The first experimental evidence of this synchronization was obtained with both electrodes quite close together in the same lobe. Therefore, since two completely separate amplifier chains were used, the possibility existed that both electrodes were recording from the same point in space and that the observed delay between spike trains was an artifact due to differences between
Figure 5.3 Oscilloscope record showing synchronization between two simultaneously recorded spike trains. Upper trace is class IIa-in unit in left lobe. Lower trace is class IIb-in in right lobe. In the right-hand part of the figure, both beams of a dual beam oscilloscope were triggered on the lower trace.
amplifier chains. However, reversing the inputs to the pre-amplifiers caused no change in delay. Spurious electronic effects were thus ruled out.

It should be pointed out that the synchronization observed via the computer dot mode display was not always perfect. Occasionally one or the other spike train would have extra spikes. This problem was particularly apparent when recording from class III units as the signal-to-noise ratio in the brain recording was poor compared to that for the lobula-lobular plate area. Also, the number of smaller spike potentials from unidentified units was greater in the brain and it was not uncommon to record from a IIIb-in and IIIa-in unit simultaneously on the same electrode, with both units having approximately the same spike waveshape. Figure 5.4 presents an example of a class IIIa and class IIa unit recorded simultaneously. In this example, instead of a one-to-one relationship, there is a one-to-many relationship between lobe and brain units. That is, for every IIa spike there is a corresponding brain spike, but there are some brain spikes for which there is no IIa (figure 5.4(b), line 3). We became convinced, however, that the unsynchronized spikes recorded from the brain electrode were from a unit or units other than the IIIa-in unit. For example, by removing the synchronized spikes (via a special program written for that purpose—see next section), the remaining unsynchronized spikes were shown not to respond as did the IIIa unit to motion patterns. Also, correlograms showed the remaining spikes not
Figure 5.4  Simultaneous recording of IIIa-in and IIa-in units. Stimulus was inward motion pattern. Motion started at 61 sec, stopped at 123 sec.

(a) Overall response of both units. Average firing rate versus time.

(b) Short section of (a) in dot mode.

(c) Short section of (b) showing time delay.

(d) Cross-correlogram of IIIa spikes with IIa spikes showing IIIa leading by 0.4 msec.
only to be unsynchronized but unrelated to the spikes on the other electrode. That is, the correlogram was essentially flat. Only in a few very noisy records was other than a one-to-one or one-to-many relationship seen.

As further verification of the synchronism, in a few cases film strip recordings were made. Examination of several hundred spikes failed to turn up a single unsynchronized spike pair. Altogether a total of 145 simultaneous recordings were made from pairs of motion detection units (from classes II and III) having the same or similar fields of view and responding to inward motion. In approximately 12 per cent of the cases the spike trains were not synchronized. We shall present the results of these cases in a later section. Table II presents a numerical summary of the cases in which synchrony was observed. In every case, the sign of the delay time corresponded to that which would be expected from spikes travelling in the directions indicated by the solid lines in figure 3.1. For example, when recording from a IIb-in in the left lobe and a IIa-in in the right lobe, the IIb spikes always occurred before the IIa, never after.

In a number of cases, the electrode positions were marked and the location of their tips determined by examination of histological sections as described in Chapter IV. Figure 5.5 presents a summary of these results, showing the area where the units were found and a plot of distance between electrodes versus lag time. In some histological sections, the lobes had been torn
Table II. Summary of cases in which synchronization between spike trains was observed. All units were selective motion detection units whose preferred direction was inward. Subscripts L, R for the case of one electrode in each lobe refer to left and right lobe.
Figure 5.5 Results of timing study of synchronized spike trains.

(a) Horizontal cross-section (diagrammatic) of optic lobes and brain. Shaded area shows approximate locus of points recorded from.

(b) Plot of distance between electrodes versus lag time for spikes recorded at two different points.
away from the brain by the sectioning procedure. In these cases, the positions with respect to the brain or lobula were noted and distance estimates made via appropriate measurement of a more suitable section. The straight line fit was by eye. Its slope is approximately unity.

In several experiments, a number of different positions were tried with the same electrode pair. That is, having found a synchronized pair, one electrode was moved to a new position and a second lag measurement made. This procedure was repeated, moving only one (either one) electrode each time. In every case, the lag time increased or decreased as the distance between electrodes was increased or decreased. These results are included in the entries in table II.

It was not uncommon in these experiments to record from pairs of motion detectors on the same electrode. In such cases the pair consisted of a IIa-in and IIb-in (or IIIa-in and IIIb-in) having fields of view in opposite eyes. This happened more often for class III than class II units. Moreover, for the class III cases, most of the time the spikes from the "a" and "b" units were approximately the same height while for the class II units the spikes were more likely to be of different heights. Also, for the class II units, it was sometimes possible to selectively alter the heights of the two classes by moving the electrode slightly in or out. That is, as the electrode was moved, the height of one of the units would increase while the other decreased.
It was observed in a number of cases that after marking the position of the "leading" electrode--by passing current through it--not only did the signal on that electrode disappear but so also did the signal on the other. In two cases out of nine this did not occur.

Two cases of synchronism were observed for outward motion detection units. One of these was a IIIa-out - IIa-out pair with the brain unit leading by 0.25 msec. The other was a pair of IIa-out's in the same lobe, with the more central unit leading by 0.05 msec.

Two explanations for the synchronization were possible:

1. The records were from different units, either in parallel as for example, from two parallel fibers in the same tract--or in series--with one unit making synaptic contact on the second.

2. As mentioned above, the records were from two different locations on the same unit.

Because of the one-to-one relationship between spike trains, it does not seem likely that the records were from two units in series. If such were the case, it should have been possible to alter the synchrony or to destroy the one-to-one relationship by presenting a visual stimulus which would cause the two units to be differentially excited. A variety of stimuli were tried but as illustrated in figure 5.6, no change was observed. Thus, for two units in series, a one-to-one synapse would be required. While
Stimulus: constant velocity inward motion.

Brain spikes lead by 0.4 msec.

Stimulus: no motion.

Brain spikes lead by 0.4 msec.

Figure 5.6 Correlograms between IIIa-in and IIA-in units for two stimulus conditions. Note similarity.
this type of synapse has been observed in both vertebrates and invertebrates, it does not appear to be as common as the integrative type of synapse [11]. In addition, the short observed delay times (0.1 msec to 1 msec) would appear to require an electrical rather than chemical type of synapse since the synaptic delay for the latter (0.5 msec to 10 msec [11]) is typically longer than for the former. The delay for the sensory-giant fiber synapse in the locust, for example, is 2-3 msec [12, p. 1137]. Of course one could propose a one-to-one electrical synapse, but there seems little point as there is no functional difference between this and an uninterrupted fiber, at least from the "black box" point of view.

The fact that delay times increased in approximate proportion to the distance between electrodes with no sudden jumps also argues against the existence of a synapse.

The case against parallel fibers is not quite so strong. However, for the reasons given above, both would have to be carrying identical information. In view of the general economy of neurons this does not seem likely. In addition, the parallel fiber hypothesis implies the existence of very nearly identical input synaptic structures which again does not seem likely to be the case, considering the invariance in the synchrony under widely varying conditions of visual stimuli.

Thus we concluded that we were in fact recording from the same neural element at different points and that class IIa-in motion
detection neurons send their axons to the opposite lobe via the brain. Class IIIb-in and IIIa-in units were therefore identified as the projections in the brain (in the sides ipsilateral and contralateral respectively) of IIb-in units. Similarly class IIa-in units were identified as continuations of these same projections into the contralateral lobe. Strictly speaking we should give the IIb-in, IIIb-in, IIIa-in and IIa-in elements the same new name. For uniformity however, we shall keep the present naming conventions.

An estimate of the velocity of propagation is given by the straight line in figure 5.5, approximately 1 m/sec. This figure is lower than that for the giant fiber systems in the cockroach and locust where the velocity ranges from 3 m/sec in the locust to 7 m/sec in the cockroach ([12], p. 1137). Leutscher-Hazelhoff and Kuiper [51] on the other hand, report a figure of 0.5 to 1 m/sec for spikes in the optic lobes of the blowfly Calliphora, with which the present figure is in good agreement. Because our measurements were straight line distances while the actual path followed by the cell was probably not along a straight line, our velocity estimates may be on the low side.

The marking procedures showed, as indicated in figure 5.5, that the transit time from the brain or ipsilateral lobe to contralateral lobe always increased as the "downstream" electrode was moved toward the medulla. Histological sections indicate two possible paths shown as dashed and solid lines in figure 3.1. If the
path marked by the dashed line had been the one recorded from, at some point transit times would have decreased as the electrode was moved toward the medulla. Since this was not the case, the path indicated by the solid line was the more likely course.

Spike Classification Procedures

As it was not uncommon to record from more than one unit on the same electrode, it was important to have suitable procedures for separation or classification of spikes from different units, especially since one of our objectives was to study the relationship between different units. Classification of multiple units on a single electrode thus increases the number of possible simultaneous relations that can be examined. A number of procedures, both digital and analog, were therefore developed for use during pre-, sample- and post-analysis, and were made an integral part of the experimental data collection and analysis system. This section describes these methods and presents some examples of their use.

After initial placement of the electrode, the signal being recorded was examined on the oscilloscope. If it was clear that only a single unit was present, the oscilloscope threshold circuit was used as described in Chapter IV to obtain times of spike firings (TOE's). If two units were observed whose heights were sufficiently different (by a minimum of about 25%) and if the signal to noise ratio were high enough, classification by analog means was possible via the window method using two threshold circuits and a logical "nand" circuit.
In other cases, it was necessary to use computer aided procedures, most of which required some waveshape information. The TOAD mode (see Chapter IV) of data transmission was therefore used—both TOE and from 12 to 24 waveform samples were sent to the computer. A number of techniques were developed and their most effective use was usually through some combination. Following a brief description of the techniques, examples illustrative of their application will be presented. The techniques are conveniently grouped into the categories waveshape-related and operational. The former are described first.

**Simple-Waveform Criteria.** These techniques were described in Chapter IV. Briefly, after examination of the waveforms and choice of suitable criteria, histograms of numerical values of the chosen characteristics are computed and displayed. Spikes in separate peaks of the histogram are then assigned to different classes. Available criteria included spike height, position of maximum, position of minimum, and width (time between maximum and minimum points). Any combination of these could also be used. This procedure was quite rapid, requiring but a few minutes (usually less than five) for one or two thousand spikes (if less than 25 waveform samples per spike were used).

**Full Waveform Classification.** Since details of this procedure appear elsewhere [43] we present an outline only. After examination of spike waveforms, the user makes an initial classification by hand of from five to ten spikes for each class. Hand
classification is done by light-penning individual spikes shown on the computer display unit. The program averages spikes within each class to form a template. The remainder of the spikes are machine classified by comparing them against the templates with a least-squares technique. While powerful, this method is somewhat cumbersome and very time consuming. Its use is therefore usually restricted to post-analysis.

The following are considered operational techniques in that they depend on some characteristic of the whole spike train. These techniques are usually used in conjunction with one of the waveshape methods.

**Classification by Stimulus.** If the units being recorded have different fields of view or respond to motion in different directions, selectively varying the visual stimulus will cause the relative populations of the units to change. For example, for selective motion detection units with the same field and opposite directions, a pattern with alternating motion can change the population completely (see figure 5.9).

**Classification by Synchronization.** This procedure depends on the results of the preceding section in which it was shown that the same unit could be recorded from at two different points. Therefore if one electrode is recording from two units, say L and R, and a second electrode is also recording from L, the spikes on the first
electrode can be separated by "subtracting" the synchronized spikes in the following way. Call the spikes on the first electrode \( L = \{L_1, L_2, \ldots, L_n\} \), and those on the second electrode \( R = \{R_1, R_2, \ldots, R_m\} \). If a correlogram is computed and the time lag measured, then the spikes on the first electrode can be expressed as a sum of unsynchronized and synchronized spikes:

\[
L = LU + LS
\]

\[
= \{L_1, L_2, \ldots, L_k\}
+ \{R_1 + \Delta \pm \delta, R_2 + \Delta \pm \delta, \ldots, R_m + \Delta \pm \delta\}
\]

where \( \Delta = \) time lag between synchronized spikes.

\( \delta = \) resolution, usually set to approximately one-half the base width of the correlogram peak.

It is then possible to design a program which accepts \( L \) and \( R \) as inputs and produces as output a dataset containing only those \( L \) spikes for which there is no synchronized \( R \) spike. It is convenient to think of this as a subtraction process, writing

\[
LU = L - (R + \Delta \pm \delta)
\]

\( LU \) now has the unsynchronized portion of the \( L \) spikes and \( LS \),

\[
LS = L - (LU \pm \delta)
\]

has the synchronized spikes. Figure 5.4(b) shows an example of the use of the subtraction program. The unsynchronized brain spikes have been subtracted out and are displayed on line 3.

Other operational methods are also possible. For example, autocorrelograms have been used together with histograms of simple waveform criteria [18].
Although the foregoing techniques can be and have been used separately for classification of units, they are most useful when various combinations are used in conjunction with one another. It is here that the power of the computer system becomes apparent in allowing rapid convenient combination and re-combination of techniques. The following examples illustrate various groupings of techniques.

Example 1. Because the classifier based on simple waveform criteria allows for output of unclassified spikes, the two waveform techniques can be used together. Suppose for instance that attempting to classify on height alone produces a histogram with three peaks, one clear and distinct, and two that overlap. Suppose further that after classifying the spikes in the distinct peak, histograms based on other available criteria also fail to separate the overlapped peaks. These spikes can be output into a new dataset which can be input to the full waveform classifier where there may be a better chance for separation. By doing as much as possible with simple techniques, the amount of time required of both experimenter and machine can be significantly reduced.

Example 2. This example is taken from an experiment in which one electrode, in the left lobe, was recording from both a IIa-in and a IIb-in unit. While they could not be separated by analog means, their heights were sufficiently different that a
height histogram produced two distinct peaks. The other electrode was in the right half of the brain. Initially the brain unit's response appeared to be that of a IIIc unit; that is, it responded to inward motion in either eye. However, as some synchronization was observed, a correlogram was computed and displayed as shown in figure 5.7(a). The appearance of two distinct sharp peaks suggested the brain "unit" was not a IIIc, but a combination of IIIa-in and IIIb-in units. The similarity of the waveforms on the brain electrode made them unseparable by waveform methods. However, after classification of the lobe spikes into IIa-in and IIb-in spikes (via height histogram), separate correlograms for each with respect to the brain were computed and displayed (figure 5.7(b) and (c)). It will be seen that in each case the correlogram shows but a single peak. Shown in figure 5.8(a) is a display of firing rate versus time for a sequence of pattern presentations. Note that there are no IIb spikes for the first pattern. Synchronization of brain and IIa spikes is evident in figure 5.8(b). For the third pattern (right-hand side of figure 5.8(a)), both IIa and IIb units responded. Synchronization of both can be seen in figure 5.8(c). It was therefore concluded that the brain spikes were from two units, a IIIa-in synchronized with the IIb-in and a IIIb-in synchronized with the IIa-in. The brain units could then be classified via the subtraction process.

Varying stimulus conditions could also be used in this experiment. Notice that in figure 5.7(a) the left peak is smaller.
Figure 5.7  Cross-correlograms between spike trains recorded in brain and left lobe. Brain electrode was recording from IIIa-in and IIIb-in units. Lobe electrode was recording from IIa-in and IIb-in units. The latter were separable by height.

(a) For all spikes on both electrodes.

(b) For IIb-in spikes only on lobe electrode.

(c) For IIa-in spikes only on lobe electrode.
Figure 5.8  Same units as figure 5.7, after classification of lobe spikes into classes IIa-in and IIb-in. Curve 1 is from brain electrode, 2 and 3 from IIb-in and IIa-in units on lobe electrode.

(a) Firing rate (spikes/sec) versus time (sec). Response to sequence of patterns in field of right eye: inward motion; outward motion; stationary.

(b) Short section of response to inward motion (in dot mode). No IIb spikes are present.

(c) Short section of response to stationary pattern. Some brain spikes synchronized with IIa spikes, others with IIb spikes.
The difference in size was a consequence of the stimulus being in the right eye field. Thus, the IIb unit (and hence the IIa unit) were strongly excited relative to the IIIa (and the IIb) unit. For inward motion in the left eye field, the relative sizes of the correlogram peaks would be reversed.

Similar conditions occurred in a number of experiments. On several occasions, we recorded from a single IIa or IIb unit in the lobe and from a unit in the brain that could have been either a IIIc or a IIIa and IIIb together. Synchronization of lobe spikes with some of the brain spikes indicated the latter were from IIIa and IIIb units.

**Example 3.** Our final example shows varying stimuli giving confirmation to classification by histogram. Here, an electrode in the right lobe was recording from both a IIa-in and a IIb-in unit. The stimulus consisted of a sequence of pattern presentations. The pattern in the left eye field was always inward motion. Figure 5.9(a) and (b) show the appearance of the spikes before and after classification by height for the two stimulus conditions. Observe that the stimulus "inward left eye, outward right eye" has in effect already "classified" the IIb-in spikes by inhibiting the IIa-in spikes, thereby confirming the classification by height histogram. Figure 5.9(c) shows firing rate versus time for both unclassified and classified spikes.
Figure 5.9 Example of varying stimuli confirming classification by height histogram. Two units—IIa-in and IIb-in—were recorded on a single electrode in the right lobe. Stimulus was sequence of pattern presentations. Pattern in left eye was inward motion, pattern in right eye alternated between inward and outward motion.

(a) Bar form of representation of spikes before classification for the two stimulus conditions shown. By inhibiting one of the units, the stimulus during the period 7.7 to 8.1 sec has in effect "classified" the IIb unit.

(b) Bar form of spikes after classification by height histogram.

(c) Average firing rate versus time for both unclassified and classified spikes. Upper two curves have been shifted up for clarity.
Inward motion: both eyes

Inward motion: left eye
Outward motion: right eye

---

Before Classification

After Classification

---
With the exception of the full waveform classifier, all of the above processes are sufficiently rapid and convenient to be used during sample analysis if the number of spikes is not too large—on the order of a thousand spikes. When multiple spikes were observed or suspected, the usual procedure was to make a preliminary run with a suitable set of stimuli. The spikes were examined and appropriate methods used to attempt to separate them. If the attempts were successful—in that the spikes were classifiable or shown to be from a single unit—the experiment was continued unmodified. Otherwise, electrode positions were changed and the procedure restarted. Classification of spikes from succeeding experimental runs was sometimes postponed to conclusion of the experiment and performed during post-analysis, especially if it were necessary to use either of the waveform classifiers on large numbers of spikes—5–10,000 or more. Often, however, the preliminary examination showed that classification by synchronization was possible, in which case it could be performed in sample- or pre-analysis mode. The flexibility and power of these methods provided invaluable aids in our efforts to examine multiple units simultaneously.

**Mutual Inhibition between Pairs of Motion Detector Units**

This section presents results on the interaction observed between pairs of "opposite" IIa-in selective motion detection units. The term "opposite" is used to mean that the two IIa-in's recorded
from were in opposite optic lobes and had their fields of view in opposite eyes. Figure 5.10 illustrates in diagrammatic form the locations and fields of a pair of opposite IIa-in units. These units have been shown to be excited by inward and inhibited by outward horizontal motion and to have a spontaneous no-motion discharge rate between 5 and 15 spikes/sec [6, 8].

The mirror symmetry in both location and response of opposing IIa-in's and the ubiquity of mutual or reciprocal inhibition in sensory system [33] suggested that such a relationship might exist between pairs of these units. It was therefore decided to attempt to record simultaneously from a pair of opposite IIa-in's to test this hypothesis. Appropriate data were taken and analyzed as described below. Analysis results indicated that IIa-in units recorded in opposite lobes indeed appeared to have a mutually inhibitory effect on one another.

The inhibitory effect can be seen both in displays of average firing rate versus time and in dot mode displays (figure 5.11). Shown here are the records from a typical pair of opposite IIa-in's. For this experiment, no visual patterns were used and the background illumination was zero. Approximately 2000 spikes were recorded on each electrode and each unit had an average firing rate of approximately 11 spikes/sec. Note that as the rate of one unit increases, that of the other decreases, and vice-versa. As shown in the figure, the SCAN program can display data in average and dot modes simultaneously. The dot mode portion of the display is
Figure 5.10 Diagrammatic representation of locations and fields of view of a pair of opposite IIa-in units. Both are excited (E) by inward motion and inhibited (I) by outward motion. The unit going from left to right has its field in the left eye and is being recorded from the right electrode (in the right lobe). The other unit is just the opposite.
Figure 5.11  Spontaneous discharge from pair of IIa-in units recorded in opposite lobes. No visual stimulus was present. Lower two curves (1 and 2) are average firing rate versus time. For clarity, curve 1 has been shifted up. Upper two lines (3 and 4) are dot mode display of first 5 seconds of lower curves. Note that as the rate of one unit increases, that of the other decreases.
a short section from the left-most part of the average firing rate curves.

The cross-correlogram shown in figure 5.12 was computed from the same data as that used for figure 5.11. Note that the correlogram shows the interaction between units for no visual stimuli. In the figure, the expected number of coincidences for uncorrelated records is marked by the letter "E". This number is given by (see reference [66])

\[ E[N(\tau, \delta)] = \rho_A \rho_B T \delta \]

where \( N(\tau, \delta) \) = number of coincidences,
\( \tau \) = lag,
\( \delta \) = resolution or binwidth of correlogram,
\( T \) = record length,
\( \rho_A \) = average firing rate of unit A,
and \( \rho_B \) = average firing rate of unit B.

It is important to note that the features of the correlogram (valleys in this case) are wide with respect to the correlogram resolution. Because of neuron refractoriness, successive bins tend to be negatively correlated, even for independent cells. That is, if a given bin has a relatively large number of coincidences compared to the expected value, the next bin will be more likely to have a smaller than expected number. Thus, significant features in the correlogram should have a width of several bins, and should be relatively independent of minor changes in resolution [67]. The
CORRELATION OF 11/212 R 1 WITH 11/212 L 1

MAXIMUM SHIFT = 240 msec  RESOLUTION = 2.51 msec

Figure 5.12  Cross-correlogram between IIa-in units recorded in the left (11/212 L 1) and right (11/212 R 1) lobes. Data from same records as used for figure 5.11. Abscissa is lag time in msec, ordinate is number of coincidences in bin of width RESOLUTION for given lag. The letter "E" indicates the expected number of coincidences for uncorrelated spike trains. The valley on the right indicates that the unit in the left lobe has an inhibitory influence on the unit in the right lobe, and vice versa. No visual patterns were used. Background illumination was zero. Each spike train contained approximately 2000 spikes.
features of correlograms in the present work were verified by recomputation for several different resolutions. The valley on the right (at a lag of approximately 15 msec) indicates that the unit in the left lobe has an inhibitory influence on the unit in the right lobe, and vice versa. More precisely, the right valley indicates a decreased probability of a spike from the unit in the right lobe given the occurrence of a spike in the other unit. The fact that the correlogram is almost symmetrical about zero lag indicates that each unit has nearly the same effect on the other. The relative flatness of the correlogram for large lags (both positive and negative) indicates a lack of interaction at these lags. Note that the expected and observed numbers of coincidences agree for large lags.

One other test of the correlogram is to randomize the intervals of one spike train and recompute the correlogram. If the features are not "artificial"--as for example in the case of two independent pacemakers--the recomputed correlogram should be essentially flat [67]. For this test, one method of randomizing the intervals of one spike train with respect to the other is simply to shift the spike train in time. Figure 5.13 shows an example of performing this operation. The lower correlogram is for un-shifted data; in the upper, one spike train has been shifted in time by five seconds. While the inhibitory valleys are clear in the lower, the upper correlogram is essentially flat. Repeating the procedure for different shifts produced no changes. To make the test convenient, facilities for shifting spike trains in time were
Figure 5.13  Test of correlogram features. Cross-correlograms of spike trains recorded from pair of opposite IIa-in units. Lower part of figure shows normal correlogram with inhibitory valleys. For upper part, spikes of record "L" were randomized with respect to spikes of record "R" by shifting the L spikes in time. Note the lack of significant features in the upper (test) correlogram.
added to SCAN. With this facility, the whole procedure from initial computation and display of the correlogram through shifting of one spike train to recomputation and redisplay of the test correlogram required but a few minutes to perform. The time required was of course dependent on spike train size. For a few thousand spikes one or two minutes were usually sufficient.

The question also arose as to whether or not the features of the correlogram were natural or usual for units mutually inhibiting each other. Although models of mutually inhibitory units have been studied (see for example Harmon [33]), and mutual or reciprocal inhibition seems to be characterized more by its ubiquity in nervous systems than by scarcity, no one seems to have taken the trouble to examine correlograms. It was decided that it would be of some interest to simulate a pair of mutually inhibiting units via the PLEXUS system and see if there was any resemblance between correlograms of simulated units and those observed experimentally. Figure 5.14 shows the network used for this simulation and the resultant correlograms between cells C01 and C02 for inhibiting connections in place and removed (the latter as a control). The squares represent sources of spikes. Both sources were independent and produced spikes at intervals drawn from normal distributions. Cell bodies are represented by triangles. Axons originate from the apex of the triangle. The model was also tested with single inhibitory connections and several variations in parameters. The model proved quite robust.
Figure 5.14  Simulation of mutually inhibitory cells via PLEXUS.
(a) PLEXUS network. S01, S02 are independent spike sources with intervals drawn from normal distribution (mean 40 msec, standard deviation 10 msec). Triangles represent cell bodies. Parameters adapted from those used in [67]. Membrane resting potential -70 mv, reset to -100 mv following a spike. Potential recovers exponentially, \( T = 144 \) msec. Resting threshold -40 mv, reset to -20 mv after a spike. Recovers exponentially, \( T = 289 \) msec. Refractory period 2 msec. EPSP's (from S01, S02) normally distributed: mean 20 mv, standard deviation 8 mv. Inhibitory fibers: conduction time normally distributed, mean 10 msec, standard deviation 1 msec; IPSP's normally distributed with mean -15 mv, standard deviation 5 mv. Approximately 1000 spikes produced by each cell.
(b) Cross-correlograms for inhibitory fibers present (lower) and removed (upper).
(a) MUTUAL INHIBITION NET

(b) CORRELATION OF (xxx)C02 WITH (xxx)C01
MAXIMUM SHIFT = 250 msec
RESOLUTION = 2.51 ms

INHIBITORY FIBERS REMOVED
in the sense of producing essentially the same result for most variations. The resemblance between the correlogram in figure 5.14 and those for experimental cases is clear.

The origin of the central peak is not completely certain. Harmon [33], using electronic models, noted some degree of near simultaneity during the period when neither unit was "dominant", that is, during the period when one unit was changing from a higher rate than the other to a lower one. There may also be some dependence on the refractory periods of the units.

Contrary to the case for synchronization, the inhibition effect was not detectable by simple analog means—that is, by viewing the oscilloscope traces or listening to the loudspeakers. The effect could only be seen by looking at the results of computer processing. In the majority of cases the effect could be seen in both SCAN display modes (figure 5.11). However, the most powerful indicator was the appearance of inhibitory valleys in the cross-correlogram.

The mutual inhibition effect as observed via inhibitory valleys in the cross-correlogram was found in over 50 preparations. The depth of the valleys and the general appearance of the correlogram (sometimes called the "signature" of the relationship in the correlogram [66]) varied over a considerable range. Figure 5.15 shows some representative examples of this variation. Notice that the effect is not always symmetric (figure 5.15 (d)). Some cases of this mutual inhibition were also found when recording from a IIa-in
Figure 5.15  Correlograms for various cases of simultaneously recorded opposite IIa-in's.

(a)  No visual stimulus.
(b)  Same as (a) but at finer resolution to show details near zero lag.
(c)  Both units stimulated by inward motion patterns. For both patterns diameter = 22°, λ = 10°, α = 0°, V = 15°/sec, I = 30cd/m², I₀ = I₀ = 10cd/m². Positions were (10, L25) and 10, R25).
(d)  Two units (IIa-in and IIb-in) recorded on same electrode in right lobe. Sharp central dip due to spike separation process. No visual stimulus.
and IIb-in in the same lobe (i.e. inward motion detection units with contralateral and ipsilateral fields, respectively). This was to be expected since our previous results showed the existence of at least one member of this class of unit which could be recorded in both lobes. An example of this is presented in figure 5.15 (d). Here, both units were recorded on the same electrode. The very sharp dip in the center was caused by the spike classification process. That is, if both spikes occurred nearly simultaneously, the resulting "spike" as recorded by the electrode would be rejected by the separation process as having too large an amplitude.

In an attempt to categorize the approximate strength of the inhibitory features of the correlogram, the ratio $G(\delta)$ was used:

$$G(\delta) = N(\tau_{V}, \delta)/E[N(\tau, \delta)]$$

where $\tau_{V}$ = lag time of valley,

and $N(\tau_{V}, \delta)$ = number of coincidences at lag $\tau_{V}$.

This ratio varies between 0 and 1 with deeper valleys giving rise to smaller values. Correlograms were categorized as having strong, medium, or weak features for the ratios 0 to 0.3, 0.3 to 0.6, and 0.6 to 0.8 respectively. Approximately 40% of the cases were in the medium category, with the remainder evenly divided between strong and weak categories. In 3 cases the ratio was greater than 0.8.

In the following discussions, when it is necessary to identify a particular member of a pair of IIa-in units, the labels
"left" and "right" will be used according to whether the unit responds to inward motion in the field of the left or right eye.

The inhibiting effect of one unit upon the other was also shown by presenting different combinations of motion patterns in the fields of one or both units. Three kinds of experiments were performed on a total of 25 preparations. Given in figures 5.16 and 5.17 are typical results of an experiment in which a single pattern was used. The stimulus was a vertically striped motion pattern (motion perpendicular to stripes). The stripes were oscillated back and forth according to a triangular wave. The pattern was positioned so that it was in the field of only one of the IIa-in units. Figure 5.16 shows a portion of the record so obtained. The upper curve (3) is pattern position as a function of time. The lower two (2 and 1) show firing rates for the left and right units. The crosses (line 4) are timing signals used in the computation of PST histograms. As the right unit responds strongly to inward motion, the rate of the left unit decreases. When the right unit is inhibited by outward motion, the left unit's rate is increased. The average change in rate is shown in figure 5.17 (curves 1 and 2) by means of PST histograms.

A control experiment was also done to verify that the inhibition was of neural origin and not from both units responding directly to the patterns. For the control case, the right eye was blocked so that the right IIa-in could not respond to the pattern. The left eye was not blocked. Stimulus conditions were unchanged.
Figure 5.16 Experiment showing effect of inhibition between opposite IIa-in units by stimulating one unit with a motion pattern. Units were IIa-in's in opposite lobes. Stimulus was motion pattern in field of right unit only. Vertically striped pattern (diameter 20°, λ = 7°) was oscillated horizontally according to a triangular wave (curve 3). Curves 1 and 2 are firing rate versus time for right and left units. Line 4 is a timing signal.
Figure 5.17 Post-stimulus-time histograms of data for figure 5.16 and for control case. For the latter, the right eye was covered and the same stimulus presented. Curves 1 and 2 are averages for conditions of figure 5.16. Curve 2 shows the left unit was inhibited during strong firing of the right unit (0 - 2.75 sec) and returned approximately to its resting level when the left unit was shut off by the pattern (2.85 - 5.5 sec). Curves 3 and 4 are for the control case. Note that there was no change in average rate for either unit as the direction of the pattern was reversed. Data were averaged in bins of width 0.0275 sec after which a moving linear average of 7 points was performed. The effective binwidth was thus 0.1925 sec.
Curves for this control case are shown in figure 5.17 (3 and 4) along with those for the non-control case. For the former, there was no net change in firing rate in either unit as the direction of pattern motion was reversed. This has two implications. First, the control was effective in that the right unit was in fact not affected by pattern motion. Second, the absence of a rate change for the left unit (the right eye was covered) showed that for the non-control case, it was not affected directly by the stimulus, but indirectly through the response of the right unit.

In the two other types of experiments involving inhibition of one IIa-in by visual stimulation of an opposite one, both units were stimulated independently. Both types of experiments involved presenting a sequence of stimuli to both eyes. Stimulus sequences are illustrated schematically in figure 5.18. Two motion patterns were used, each only in the field of the corresponding IIa unit. Both patterns were presented suddenly; that is, turned on, then off. For both types of experiments, one pattern ($P_R$, in the field of the right eye in figure 5.18) had constant velocity.

For the first type of experiment, the other pattern ($P_L$ in figure 5.18) was oscillated back and forth once during each presentation. In figure 5.19 (a) are shown PST histograms of typical responses to such a pattern sequence. Curve 2 shows the response of the left unit to the oscillating pattern. As the initial direction of motion for $P_L$ was outward, the response of the left unit was inhibited during the first half of the pattern presentation--except for a
Figure 5.18 Stimulus sequences for two types of experiments to test inhibition. Pattern $P_L$ ($P_R$) is in the field of the left (right) eye. Both patterns turned on and off together; $P_R$ has constant velocity inward motion. Stripes moved within fixed pattern boundaries.

(a) Relative positions of $P_R$ and $P_L$.

(b) Sequence for oscillating $P_L$. During each "on" period, $P_L$ was oscillated back and forth once according to a triangular wave. The on-off cycle was repeated a number of times, usually 16.

(c) Sequence for alternating $P_L$. Motion of $P_L$ alternated between inward and outward motion every other "on" period. The inward-outward cycle was repeated a number of times, usually 16.
Figure 5.19  PST histograms of responses of opposite IIa-in's for pattern sequences in figure 5.18.
(a) Oscillating $P_L$. Curves 1 and 2 are responses of right and left units, respectively. For $P_L$, position = (10, L30), diam. = 22°, $\lambda = 7.5°$, $\alpha = 0°$, $I_L = 60$ cd/m$^2$, $I_b = I_0 = 10$ cd/m$^2$, amplitude = 16°, period = 3.2 sec. For $P_R$, position = (110, R45), diam. = 20°, $\lambda = 10°$, $\alpha = 0°$, $V = 15°$/sec, $I_L = 50$ cd/m$^2$.
(b) Alternating $P_L$. Curves 1 and 2 (3 and 4) are responses of left (right) unit for inward and outward motion of $P_L$, respectively. Pattern parameters similar to those above except $P_L$ is alternating instead of oscillating.
brief transient response to pattern "on". When the direction of motion was reversed, the left unit responded strongly to the inward motion. The right unit's response was strong initially but decreased sharply when the left unit's rate increased due to pattern motion reversal.

Stimuli for the second type of experiment were similar to that described for the first except that instead of the direction of one pattern reversing during its on period, the direction was changed when the pattern was off (figure 5.18). In other words, both patterns were turned on and off periodically and both had constant velocity while they were on. One pattern had inward motion for each presentation, the other alternated between inward and outward motion. The upper two curves of figure 5.19 (b) show the response of the left unit to the alternating pattern, the lower two are the responses of the right unit to the constant pattern. Of the latter two, the one with the highest response is for outward motion of the alternating pattern—that is, for the cycle when the left unit was inhibited by the pattern.

In all cases, the parameter used to measure the amount of inhibition was average firing rate. This was done by using the SCAN display of PST histograms. A flat portion of the response (of duration 1 to 2 seconds) was chosen and average number of spikes per pattern presentation per second computed. For the experiments in which oscillating patterns were used, two intervals were chosen immediately adjacent to and on either side of the point of pattern
motion reversal. When slow overall trends in firing rate (in addition to the sharp decrease in rate due to inhibition) were observed, several pairs of intervals of varying duration were used. The computed average rates were plotted against interval length and extrapolated to zero interval duration.

A general increase in the amount of inhibition with increasing rate of the inhibiting unit was observed. There was, however, considerable variation. It was then observed that the amount of inhibition seemed to depend not only on the firing rate of the inhibiting unit, but also on the inhibited unit's rate. Thus the plot in figure 5.20 shows the results of a number of inhibition measurements as decrease in rate of inhibited unit (normalized) versus the difference between the inhibiting unit's rate and the inhibited unit's rate prior to being inhibited. That is, the ordinate is percent decrease in rate for the inhibited unit. Let $\rho_{D}^{in}$ and $\rho_{D}^{out}$ be the rates of the unit responding directly to the pattern with oscillating or alternating inward and outward motion. Let $\rho_{I}^{in}$ and $\rho_{I}^{out}$ be the corresponding rates of the inhibited unit. Then the abscissa is $\rho_{D}^{in} - \rho_{I}^{in}$. While there is considerable variation, in general, the amount of inhibition increases with increasing values of $\rho_{D}^{in} - \rho_{I}^{in}$. The plotted points were from a number of experiments using different combinations of patterns as described above. Although inhibition was occasionally present for $\rho_{D}^{in} < \rho_{I}^{in}$, it was more often the case that none could be observed. These cases were not included in figure 5.20.
Figure 5.20 Amount of inhibition versus difference between rates of inhibiting and inhibited units. Ordinate is per cent decrease in rate of inhibited unit. Let $\rho_D^{in}$ and $\rho_D^{out}$ be rates of unit responding directly to the pattern with alternating (or oscillating) inward and outward motion. Let $\rho_I^{in}$ and $\rho_I^{out}$ be the corresponding rates of the inhibited unit. Then the abscissa is $\rho_D^{in} - \rho_I^{in}$. Different symbols correspond to different preparations.
Initial Oscillation. In previous work [58] an initial oscillation in firing rate upon sudden presentation of a pattern was observed. This oscillation was also present immediately following pattern motion reversal. In addition, opposite-phase oscillations were often detected in the opposite IIa-in unit.

Convergence. The results of the pathway tracing experiments as well as the results of previous work [8] indicated that the effect of the inhibiting IIa unit on the inhibited should not depend on the particular motion pattern or patterns used to stimulate the inhibiting unit. That is, two different patterns stimulating the inhibiting unit should produce the same effect if the firing rate was the same for both patterns. This hypothesis was tested with the set of stimuli in figure 5.21. The patterns were $P_R$ in the right field and $P_{LA}$ and $P_{LB}$ in the left field. The experiment consisted of two steps. The first was conducted with the stimuli as shown, except that the directions of $P_{LA}$ and $P_{LB}$ were always inward. Under these conditions, the intensities, velocities, and position of $P_{LA}$ and $P_{LB}$ were adjusted so that the response of the left unit was approximately the same for both. Adjustments were made on the basis of oscilloscope and loudspeaker output, and by monitoring the SCAN display (pre-analysis). A test run of a few cycles was then made and PST histogram averages compared (sample-analysis) to verify that the responses of the left unit were essentially the same for both left patterns. The second stage was presentation of the
Figure 5.21 Stimulus sequence for convergence experiments. Upper part of figure shows relative positions of patterns: $P_R$ in field of right unit; $P_{LA}$ and $P_{LB}$ in field of left. Lower part of figure shows timing sequence for pattern presentations. All patterns had constant velocity horizontal motion during "on" period. $P_R$ always had inward motion. $P_{LA}$ and $P_{LB}$ were presented alternately; their on- and off-times coincided with those for $P_R$. $P_{LA}$ and $P_{LB}$ alternated between inward and outward motion.
pattern sequence as in figure 5.21. PST histograms for the left (inhibiting) unit were computed for the four stimulus variations: $P_{LA}$ inward and outward; $P_{LB}$ inward and outward. The two inward and outward cases for the inhibiting unit were again compared to verify that they did not differ. The corresponding four histograms for the inhibited unit were then computed and the amount of inhibition for $P_{LA}$ and $P_{LB}$ measured as described above. The experiment was performed on four preparations with a total of nine different pairs of patterns. The ratio of difference in amount of inhibition to firing rate of the inhibited unit varied between 0.5% and 10%. The average was 5%. In all cases, the difference appeared negligible on visual examination of the PST histogram.

Divergence. The variation in amount of inhibition suggested that the effect on the inhibited unit might depend on the particular motion patterns in its field. This was tested in the following way. Consider the two very schematic models in figure 5.22. The left and right units represent a pair of opposite IIa-in's, with firing rates $\rho_L$ and $\rho_R$, respectively. Each has two input branches. Each branch is assumed to be separately excited by patterns P1 and P2. That is, for model A, it is assumed that if only P1 is present, the rates at the point "a" are $\rho_{R1}$ and 0 for the upper and lower branches. For pattern P2 only, the rates are 0 and $\rho_{R2}$ for the upper and lower branches. For model A, each branch is separately "inhibited" by the left unit such that the rates at point "b" are $\rho_{R1} - k_1\rho_L$ for the
Figure 5.22  Schematic representation of models for testing divergence hypothesis. Left and right units represent a pair of opposite IIa-in units.
upper branch and $\rho_{R2} - k_2\rho_L$ for the lower branch. (The $k_i$ can be thought of as "coefficients of inhibition".) That is:

$$\rho_R = \rho_{R1} - k_1\rho_L(P1) \quad \text{for P1 only},$$

and

$$\rho_R = \rho_{R2} - k_2\rho_L(P2) \quad \text{for P2 only},$$

where the notation $\rho_L(Pi)$ is used to allow for dependence of $\rho_L$ on pattern. For model A, the general equation for $\rho_R$ is

$$\rho_R = [\rho_{R1} - k_1\rho_L(P1)] + [\rho_{R2} - k_2\rho_L(P2)]$$

where the brackets indicate that each term cannot be less than zero. For model B, the corresponding equation is

$$\rho_R = \rho_{R1} + \rho_{R2} - k\rho_L.$$ 

The problem is to devise an experiment that will distinguish between models. Suppose $P1$ only is presented under the conditions of two different firing rates for the left unit: $\rho_L > 0$; and $\rho_L = 0$. Then

$$\rho_R(P1) = \rho_{R1} \quad \text{for } \rho_L = 0,$$

$$\rho_R(P1) = \rho_{R1} - k_1\rho_L(P1) \quad \text{for } \rho_L > 0.$$ 

Writing $\Delta\rho_R(P1)$ for the difference in $\rho_R(P1)$ under these two conditions,

$$\Delta\rho_R(P1) = k_1\rho_L(P1),$$
or

\[ k_1 = \frac{\Delta \rho_R(P1)}{\rho_L(P1)} \]  \hspace{1cm} (1)

Similarly, for pattern P2 only,

\[ k_2 = \frac{\Delta \rho_R(P2)}{\rho_L(P2)} \]  \hspace{1cm} (2)

Then for model A,

\[ \frac{\Delta \rho_R(P1)}{\rho_L(P1)} - \frac{\Delta \rho_R(P2)}{\rho_L(P2)} = k_1 - k_2 \]  \hspace{1cm} (3)

Note that if \( \rho_L(P1) = \rho_L(P2) \), then the ratio of the \( k_i \)'s is simply the ratio of the \( \Delta \rho_R \)'s:

\[ \frac{k_1}{k_2} = \frac{\Delta \rho_R(P1)}{\Delta \rho_R(P2)} \]  \hspace{1cm} (4)

For model B, the right hand side of equation (3) is zero, and the ratio in equation (4) is unity.

Thus, an experiment in which the inhibited unit (the right unit in the model) is stimulated with two different patterns should differentiate between model A for unequal \( k_i \)'s and either model B or model A for equal \( k_i \)'s. For each pattern presented to the inhibited unit, the inhibiting unit (left unit in the model) is presented both an inward motion pattern (for \( \rho_L > 0 \)) and an outward motion pattern (for \( \rho_L = 0 \)). The rates can be measured and the quantities on the right hand sides of equations (1) and (2) computed. If they
are significantly different, model A is a better representation.

A typical arrangement of patterns is shown in figure 5.23, with P1 and P2 in the field of the right unit and P_L in the field of the left unit. Special timing circuitry was used to produce the pattern sequence shown. P1 and P2 always had constant inward motion and, as described earlier, P_L was either oscillated back and forth once per presentation or had alternating inward and outward motion.

Figure 5.24(a) and (b) show typical responses in the form of PST histograms for oscillating and alternating P_L, respectively. The responses of the inhibiting unit for alternating P_L (figure 5.24(b)) are not shown, but were similar to the upper curves in figure 5.19(b). The amount of inhibition was measured as described earlier. Rather than using the SCAN program, however, a special program was written to compute the average number of spikes per pattern presentation for a time interval whose endpoints with respect to stimulus-on time were specified by the user. The program also computed the root-mean-squared deviation.

A total of 17 experiments were performed as described above on 8 preparations. The experiments were approximately evenly divided between the inhibited unit being the left or right Ila-in. The combinations of patterns used to stimulate the inhibited unit (P1 and P2) are illustrated schematically below.
Figure 5.23 Stimulus sequences for divergence experiments. P1 and P2 have constant velocity inward motion. $P_L$ is either oscillated back and forth once per presentation or has alternating inward and outward motion during alternate presentations.

(a) Relative pattern positions.
(b) Timing sequence for oscillating $P_L$. Two cycles shown.
(c) Timing sequence for alternating $P_L$. One cycle shown.
Figure 5.24 Typical PST histograms from divergence experiments for stimuli as shown in figure 5.23. Both (a) and (b) are firing rate (spikes/sec) versus time (sec).

(a) For oscillating $P_L$. Average of 16 cycles.

(b) For alternating $P_L$. Average of 9 cycles. Responses of left unit not shown.
Cases (1) and (2) were performed with the oscillating pattern only. Case (3) was performed with both oscillating and alternating patterns in the field of the inhibiting unit. Approximately half the experiments were for case (3). The remainder were approximately evenly divided between (1) and (2). In comparing the amounts of inhibition, the percentage decrease in rate was used to attempt to normalize effects that may have been dependent on absolute rate. The differences in inhibition were tested using the two-sided $t$-test at the 5% level of significance. Table III shows typical results for each case illustrated in the above diagram.

For case (1), no significant difference in the amount of inhibition was found. For cases (2) and (3), there was a significant difference with pattern 2 giving on the average approximately 20% less inhibition for (2) and 50% less for (3). Thus, for the patterns positioned in approximately equal response regions of the inhibited unit's field, no difference was detected. For patterns positioned so that one was in a weaker region, there was more
Table III  Typical data from divergence experiments. Case number refers to positions of patterns P1 and P2 (see diagram on previous page). Firing rate symbols are: $\rho_L$ = rate of inhibiting unit for inward ($\rightarrow$) and outward ($\leftarrow$) motion; $\rho_R$ = rate of inhibited unit for inward and outward motion in field of inhibiting unit. All rates are in spikes/sec. Ratio of $k_i$'s calculated using equation (4)---since from the table, $\rho_L(P1) \div \rho_L(P2)$. 

<table>
<thead>
<tr>
<th>Case</th>
<th>P1 or P2</th>
<th>$P_L$ in $\rightarrow$ out $\leftarrow$</th>
<th>$\rho_L$</th>
<th>$\rho_R$</th>
<th>Percent decrease in $\rho_R$ ($% \Delta \rho_R$)</th>
<th>$k_1 = \frac{% \Delta \rho_R(P1)}{% \Delta \rho_R(P2)}$</th>
<th>$k_2 = % \Delta \rho_R(P2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P1</td>
<td>$\rightarrow$ 141 0 20 42</td>
<td>53</td>
<td>10</td>
<td></td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>$\leftarrow$ 140 0 20 38</td>
<td>48</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>P1</td>
<td>$\rightarrow$ 113 0 24 44</td>
<td>45</td>
<td>14</td>
<td></td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>$\leftarrow$ 113 0 56 70</td>
<td>20</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>P1</td>
<td>$\rightarrow$ 116 0 29 48</td>
<td>39</td>
<td>15</td>
<td></td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>$\leftarrow$ 118 0 53 66</td>
<td>20</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
inhibition for the patterns in the weaker region.

**Unsynchronized Inward Motion Detectors**

In the pathway tracing study, a large number of cases were found in which it was possible to record simultaneously from a single horizontal motion detection unit at two widely separated points in space. The fact that it was relatively easy to obtain such records led us to ask if in fact it was the same large cell that was repeatedly recorded or if there were two or more units in parallel. With present methods, a positive answer to the first question is extremely difficult to obtain. Electrophysiologically, one could only say that out of a large number of records, none were found to be unsynchronized. Extracellular marking procedures (see Chapter IV) can supply some additional evidence, but again, it would be difficult to make definite claims. On the other hand, even a few cases in which no synchrony was observed would enable one to say with some confidence that more than one unit existed.

The following presents results obtained which bear on the above question. The experiments considered here are restricted to those in which two inward motion detection units having fields of view in the same eye were recorded simultaneously. Explicitly excluded for example are combinations such as a IIa-in and IIa-out recorded in the same lobe and IIa-in's recorded in opposite lobes.

As mentioned in the section on pathway tracing, a total of 145 simultaneous recordings were made from the classes of units
considered here. In no instance were the spike trains unsynchronized when both electrodes were recording from class IIa-in units in the same lobe. In approximately 12% of the total number of recordings, no synchronization was detected. Table IV below shows the recording locations for these cases along with number of preparations for each. Of these, cross-correlograms were obtained for

TABLE IV

<table>
<thead>
<tr>
<th>Location</th>
<th>Type of Unit</th>
<th>Number of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>same lobe</td>
<td>IIb - IIb</td>
<td>2</td>
</tr>
<tr>
<td>brain - lobe</td>
<td>IIIb - IIb</td>
<td>3</td>
</tr>
<tr>
<td>opposite lobes</td>
<td>IIb - IIa</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>IIa - IIb</td>
<td>9</td>
</tr>
</tbody>
</table>

II cases, only one of which indicated little or no interaction. In the majority of IIa - IIb cases, the IIb's response to motion patterns was less than that of the IIa, although with one or two exceptions, the fields of view appeared to be approximately the same. In the exceptions, the IIb's field was smaller. Figure 5.25 shows the response of a typical IIa-IIb pair in dot mode for constant velocity inward motion. Typical correlograms for this combination are shown in figure 5.26 for no pattern and for a constant velocity (inward) motion pattern, each for two resolutions. For the coarser resolutions (figure 5.26 (b) and (d)), the peak appears to occur at zero lag. The only differences between motion and no-motion are that for
Figure 5.25  Firing patterns of simultaneously recorded IIa-in in left lobe and IIb-in in right lobe. Stimulus: background illumination plus constant velocity inward motion pattern in right eye field.
Figure 5.26  Cross-correlograms of IIb-in unit with IIa-in unit (same as for figure 5.25). For (a) and (b), background illumination only; approx. 4000 and 1000 spikes for IIa and IIb, respectively. Rates were 17 and 3.5 spikes/sec. For (c) and (d), constant velocity motion pattern (same as figure 5.25); 6500 and 2900 spikes, 51 and 23 spikes/sec for IIa and IIb, respectively. Expected numbers of coincidences marked by arrow "E".
the former the expected number of coincidences is higher (due simply to the increased number of spikes) and the small valley at a lag of -6 msec. The latter may be due to refractoriness of one or both units having a greater effect at higher firing rates. The central location of the peaks is an indication that the interaction is more likely due to a common source of input than to one unit directly affecting the other. Some asymmetry can be seen however at the finer resolution. Here (figure 5.26 (a), (c)) the peak appears centered at a lag of 0.5 msec. The cause of the sharp dip near zero lag is not clear. That it was not because of an artifact due to the particular choice of resolution was verified by recomputation for a number of different resolutions. It occurred in about half of the cases. Although the correlograms in figure 5.26 are typical of most of the records obtained, in 3 instances the peaks occurred at lags of approximately 4 msec, indicating increased probability of a IIb spike given the occurrence of a IIa spike. That is, it appeared that either the IIa had an excitatory effect on the IIb, or that if both were receiving input from a common source then the IIb had a longer latency. Correlograms for pairs of IIb's in the same lobe were not unlike those for a IIa-IIb pair in opposite lobes, as illustrated in figure 5.27. Also shown are short sections of the dot mode representation of the spike trains. Although a correspondence between the spike trains can be seen in the dot mode representation, there is no absolute synchrony as was observed in the pathway
Figure 5.27  Firing patterns and correlograms (of central IIb with respect to peripheral) for pairs of IIb-in units. (a) and (b) are from different preparations.

(a) No illumination. Central and peripheral electrodes recorded approx. 2500 and 2300 spikes.
Correlogram resolution = 1.01 msec.

(b) Constant inward motion in field of right eye. Central and peripheral electrodes recorded approx. 6700 and 5500 spikes.
Correlogram resolution = 0.15 msec.
Resolution = 1.01 msec

Resolution = 0.15 msec

Central electrode

Peripheral electrode

1 sec time window

Central electrode

Peripheral electrode

2 sec time window
tracing results. Figure 5.27 (b) shows that the sharp dip can occur for \( \text{IIb} \) pairs as well as \( \text{IIa-IIb} \) pairs. Here again, the closeness of the peak to zero lag tends to indicate interaction via common input or inputs. In figure 5.27 (a) on the other hand, there is a very strong excitatory peak indicating a much higher probability of spike on the more central electrode 4 msec after one on the more peripheral electrode.

To summarize the results of this section, consider the schematic cross-section shown below. In a large number of cases,

it was possible to record from the same inward motion detector \( \alpha \) at two spatially separated points (a and b, say) simultaneously. The question then was, is there another such unit, \( \beta \)? If so, recordings made from pairs of points such as a-a' or a-b' should be unsynchronized. From the results, no pair a-a' was found. That is, every time a pair of inward motion detectors was recorded in the contralateral lobe, they were synchronized. However, unsynchronized pairs a-b' and b-b' were found. This suggests the above
diagram be modified as shown below. Although the results showed the units represented by $\alpha$ and $\beta$ were not synchronized, there did appear to be interaction between them. Correlograms showed the interaction to be of an excitatory or common input nature.

**Outward Motion Detectors**

The class IIa-out units respond maximally to outward motion. Their field of view is similar to that of the IIa-in, but is centered somewhat higher \[8\]. Although the outward motion detection units were considerably more difficult to record from than the inward, on several occasions a IIa-in and IIa-out were recorded simultaneously from two electrodes placed in opposite lobes. As for the IIa-in, the IIa-out units usually exhibited a spontaneous firing rate of between approximately 1 and 10 spikes/sec in the dark. Occasionally a low level of background illumination was required to obtain firing.

Cross-correlograms were computed and examined either for conditions of no illumination or for constant background
illumination when necessary. No difference in correlograms for the
two cases was observed, nor between those computed for both cases
for the same preparation. In every case, the correlogram showed
a large peak at a lag of approximately 4 msec (figure 5.28), indi­
cating an increased probability of a IIa-out spike given the occurrence
of a IIa-in spike. In the 10 preparations in which the IIa-out - IIa-in
combinations were recorded, the location of the peak varied from
3 msec to 6 msec and occurred most often between 3.5 msec and
4.5 msec. The small dip following the peak may be due to the
refractory period of the IIa-out. That is, once the unit has fired,
the probability that it will immediately fire again is less than the
expected value for a random point in time.

Due to the difficulty of obtaining and keeping the outward
motion detection units for any length of time, it was possible to
perform direct tests on only three preparations. These tests were
similar to those described earlier for pairs of opposite IIa-in units.
Because the correlogram indicated that the IIa-in was affecting the
IIa-out rather than vice-versa, the oscillating (or alternating)
pattern was placed in the field of the IIa-in. The pattern in the field
of the IIa-out had constant outward motion. In 2 of the 3 experiments,
PST histograms showed the IIa-out's rate to be greater for inward
motion in the field of the IIa-in than for outward by approximately
20 percent. In the third case, the difference was 3 percent.
Figure 5.28 Cross-correlogram between IIa-in and IIa-out units recorded in opposite lobes. No visual stimuli were present. Lag time of peak is approximately 3 msec indicating increased probability of IIa-out spike given occurrence of IIa-in spike.
Other Interactions

In two preparations, different inward-outward combinations from those presented earlier were found. In one case, a IIb-in - IIb-out pair were recorded in the left and right lobes, respectively. No interaction was detected--the correlogram was essentially flat. In the second case, a IIa-in - IIa-out pair were recorded in the same lobe. Thus, both units had their fields in the same eye, one unit responding to inward, the other to outward motion. The correlogram for this pair had a weak inhibitory valley extending approximately 10 msec on either side of zero lag.

The correlogram between a IIa-down and IIa-in recorded in the same lobe was flat, indicating little or no interaction. On one occasion a class IIa-up was recorded simultaneously with a ventral cord unit that responded primarily to upward motion. As their spontaneous firing rates were almost zero, a motion pattern with constant upward velocity was used. The correlogram showed a peak near zero lag.

Pairs of IIb-in's were recorded in opposite lobes in two preparations. In one, no correlation was observed; in the other, a weak inhibitory valley similar to that for opposite IIa-in's was noted.

In eight preparations a combination of class I and either class II or III units were recorded. In every instance but one, the cross-correlogram was flat. In the exception, a very weak peak at lag zero was observed.
CHAPTER VI

DISCUSSION AND CONCLUSIONS

Introduction

In the course of this work we have been concerned with two areas: (1) processing of visual information in the central nervous system of the fly, and (2) development and application of the computer as an aid in such research. Accordingly, following a brief review of the neurophysiological findings a discussion of their significance is given along with possible behavioral correlates and suggestions for areas of future research. Performance of the data acquisition and analysis system is then reviewed and various problems that arose during its development and use are discussed. Suggestions for improvements and directions for further development are also given.

The Motion Detection Units

The following summarizes very briefly the major findings presented in the foregoing chapter. First, previous work [8] had demonstrated the existence of inward motion detector units with similar fields of view in optic lobes ipsilateral (class IIb units) and contralateral (class IIa units) to the field. By means of multiple electrode techniques we have shown these two units to be ipsilateral and contralateral ends of the same neural element. The path followed by this element was traced from the distal end of
one lobe to the central end, through the brain, and into and through the contralateral lobe. Both members of the symmetrical pair were traced, i.e. recordings were made from units conducting spikes from left to right and right to left. Some similar evidence was also obtained for the outward motion detection units. This seems to be the first case in which the pathway of a single neuronal element in the central nervous system has been verified using functional characteristics as criteria.

When both electrodes were recording contralaterally (IIa-IIa), no unsynchronized pairs were found. However, some unsynchronized pairs were found when both electrodes were ipsilateral (IIb-IIb) or when one was ipsilateral and one contralateral (IIb-IIa). For these two pairs, peaks in the cross-correlogram demonstrated the existence of interaction between members of the pairs. The peaks in the correlograms were of two types: (1) near zero lag indicating both units had a common source of input, (2) at a lag of about 4 msec indicating a more causal type of relationship. In other words, the units appeared to be in parallel for (1), in series for (2). The peak location indicated the IIa (or peripheral IIb) had an excitatory effect on the (more central) IIb.

Mutual inhibition between pairs of opposing IIa-in's was shown via direct examination of firing patterns, via characteristic signature in the correlogram, and by various direct experiments in which a visual stimulus caused one unit either to fire strongly or to be inhibited from firing. For the two cases in which IIb's were
recorded in opposite lobes, in one instance they were mutually inhibiting, in the other no interaction was seen. The above relationships along with those between inward and outward motion detection units are shown schematically in Table V.

What are the implications of the pathway tracing experiments? Clearly the primary result is that the contralateral inward motion detector is not an \((n+1)^{\text{th}}\) order unit with the ipsilateral inward motion detector being an \(n^{\text{th}}\) order unit. That is, information is transmitted by an axon of the class II neuron from one optic lobe to the opposite one without any intervening synaptic junction. What purpose is served by this mechanism, and how can we use this fact to further our understanding of the fly's visual system? One of the functions served is to inhibit spike generation in the opposite inward motion detecting neuron. We shall return to this point below. One way to apply the pathway tracing results is as an aid to detailed investigations of structure. With extracellular marking of the electrode tip location as a guide, histological examinations have in fact been begun [79]. The unit tentatively identified as the inward motion detector has been tracked by such means into the brain where a branch has been found going down toward the subesophageal mass. In the brain the fiber appears to be part of a tract of either eight or twelve similar fibers. If one assumes the former figure to be correct and that the eight are composed of four conducting each way, then each can be assigned to one of the four directions for selective
### Table V

Summary of relationships observed between pairs of horizontal motion detectors. Left column shows schematic diagram of units. Arrows show direction of motion to which units respond.

<table>
<thead>
<tr>
<th>Class of Units</th>
<th>Number of Cases</th>
<th>Type of Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="IIb-in-IIIb-in-IIIa-in-IIa-in" alt="Diagram" /></td>
<td>128</td>
<td>Same neural element. Delay is approximately 0.1 msec for distance of 0.1 mm.</td>
</tr>
<tr>
<td><img src="IIb-in-IIa-in-IIa-in" alt="Diagram" /></td>
<td>50</td>
<td>Mutually inhibitory. Valleys at lags of ±10 msec.</td>
</tr>
<tr>
<td><img src="IIb-in-IIa-in" alt="Diagram" /></td>
<td>17</td>
<td>Excitatory. Peaks near zero lag or at lag of 4 msec.</td>
</tr>
<tr>
<td><img src="IIa-out-IIa-in" alt="Diagram" /></td>
<td>10</td>
<td>Excitatory. Peak at 4 msec.</td>
</tr>
<tr>
<td><img src="IIa-in-IIa-out" alt="Diagram" /></td>
<td>1</td>
<td>Inhibitory. Central valley extending ±10 msec.</td>
</tr>
</tbody>
</table>
motion detectors: upward, downward, inward and outward. So far, however, the histological identification remains tentative. In addition, in the present work it has been possible to show the identity of ipsilateral and contralateral motion detectors for the inward direction only, with two instances of synchronization for outward motion detectors.

Should intracellular recording techniques and marking procedures prove feasible in this preparation, one other benefit may accrue from the pathway tracing results. It is often quite difficult for the histologist to determine whether a terminal structure is pre- or post-synaptic on the basis of anatomical information from classical Golgi staining techniques. In the present case however, since the direction of spike propagation follows from our results, determination of which end of the element is pre- and which post-synaptic is straightforward.

Let us consider now the results from pairs of unsynchronized inward motion detectors. Earlier work [8] indicated there were a number of IIb-in's (different fields and peak response points were found). No information was given concerning the number of IIa-in's. Although the present findings do not prove conclusively that for each side of the visual system there is but a single inward motion detector sending a fiber to the contralateral lobe, they do
suggest that this may be the case. Further, none of the results contradict the claim, as would for example, finding an unsynchronized pair of IIa's in the same lobe. On the other hand, the results do show directly that there are at least two inward motion detectors on the ipsilateral side--two IIb-in units--one of which sends a fiber contralaterally. Recall that two cross-correlogram signatures were found for simultaneously recorded IIa and IIb units in opposite lobes. Because it is not possible in general to distinguish between a pair of units in parallel and a pair in series with the cross-correlogram, either signature can be interpreted either way. It is not likely that both signatures are due to units in series however, since in the one case, not only is the peak near zero lag, but it overlaps zero lag. More probable is that one correlogram is due to units in series, the other to units in parallel. A possible interpretation showing three "IIb" units is given in the diagram below. Recordings at a₁ and b₁ would give rise to a peak with a

![Diagram of brain lobes showing connections between IIa and IIb units.](image)
large lag (due to synaptic delay); recordings at $a_1$ and $b_2$ would have a peak near zero.

Additional evidence for the existence of separate ipsilateral horizontal inward motion detectors may also be adduced. In one of the two cases in which opposite IIb's were recorded, mutual inhibition was observed. This is reasonable since the IIb's were traced through to the opposite lobe--i.e. IIa's and IIb's can be the same. In the other case, no interaction was seen. This supports the case for more than one IIb. For the pair $a_3 - b_3$ in the above diagram, inhibition would be expected. For $a_2 - b_2$, no interaction would be seen. It is possible that $a_3 - b_1$ could occur, in which case the correlogram should show some interaction. However, since opposite IIb's were only found in two instances, it is not reasonable to do more than include them as supporting evidence for two or more inward motion detectors, one of which sends a fiber contralaterally.

With regard to the opposite IIa's, it seems clear that they do have a mutually inhibitory relationship, the exact nature of which appears quite complex. The amount of inhibition does increase with the rate of the inhibiting unit, and is independent of the particular pattern used to stimulate the inhibiting unit. This is as expected for inhibition by means of spikes conducted from one lobe to the other. The amount of inhibitory chemical substance released is a function of the incoming spikes, independent of how the spikes were generated. If Dale's principle is assumed to hold
here, then one or more interneurons may be involved. The
detailed nature of the inhibitory relationship and possible inter-
action with or dependence on other "parallel" IIb units remains a
question for future research. In addition, since the correlogram
between inward and outward motion detectors with fields of view
in the same eye showed an inhibitory relationship, the effect of
the outward unit may also have to be taken into account. This
of course raises the question of how to deal with multiple inter-
actions--those involving three or more units recorded simultaneously.
Although the question is being investigated [66], this problem is
largely unsolved and one that will require considerable effort, both
theoretical and experimental.

It is interesting to interpret the existence of both inward and
outward motion detectors in the light of the previously mentioned
optomotor work by Goetz [28] on Drosophila and Musca. Torque
and thrust were found to be described approximately by cosine and
sine of the angle between pattern motion and the long axis of the fly
(recall that the IIa-in unit's response also obeys a cosine law [8]).
Goetz concluded that torque and thrust motion detectors were
oriented in at least two linearly independent directions. In the
model he derived from these results, in order to modify torque
without changing thrust, it was necessary that the excitation to
one side of the motor output system be increased while that to the
other is decreased. Goetz interpreted this as arising from
a single detector exciting one side and at the same time inhibiting
the opposite side of the motor system.
Notice though that since inward and outward motion detectors exist, both effects need not be due to a single detector. Inward and outward detectors could operate in a push-pull fashion. Relative inward motion would increase the output of the inward detector and decrease the output of the outward detector.

It is also possible to interpret mutual inhibition between opposite inward motion detectors in a similar manner. That is, inward motion excites one side of the motor system via the branch going down to the subesophageal ganglia, and inhibits the opposite side via the inhibitory effect on the opposite inward motion detector. However, this does not seem as likely as the above interpretation for two reasons. First, outward motion detectors would seem superfluous. Second, our results indicate the inhibitory effect to be a diffuse one which seems an unnecessary complication just to inhibit the opposite motor output system.

It therefore seems more likely that the observed mutually inhibitory relationship may be associated with some more complex function such as depth perception. Future research may provide more information on this question.

The excitatory relationship observed between IIa-in and IIa-out units in opposite lobes indicates an increased probability of observing a IIa-out spike given the occurrence of a IIa-in spike, with a time lag of about 4 msec. One interpretation of this is that objects entering the field of view of one eye are of more immediate interest than those leaving the field of the other. Alternatively, or additionally, the relationship could serve to "prepare" the other side for an
object about to enter its field. Although the rather large lag time of the peak is indicative of one unit having an excitatory effect on the other, the possibility of a common source of input to the two units cannot be excluded. That is, the existence of a peak in the correlogram is a necessary but not sufficient condition to show that two units are in series.

More generally, it is interesting to note first that there is a remarkable lack of independence between the classes of horizontal motion detection units, and second that the observed relations are generally what might be expected. That is, four out of the six pairs possible with inward or outward motion detectors having fields of view in the same or opposite eyes have been recorded simultaneously (the missing combinations are pairs of IIa-out's having fields in the same or opposite eyes). In every instance, the correlogram has shown the existence of some relationship. For "opposed" detectors--inward in opposite lobes or inward-outward in the same lobe--the relationship is inhibitory in nature. For "non-opposed" detectors--inward in the same lobe or inward-outward in opposite lobes, the relationship is excitatory.

Results of the present work indicate three areas of interest for future investigations. First, if techniques for recording the intracellular response of the class IIa unit could be developed, several interesting studies are possible. If it is possible to stain these units intracellularly by dye injection, the pathway tracing results will
provide a link between histological and electrophysiological investigations. Further, intracellular techniques will allow recording EPSPs and IPSPs. This, along with direct intracellular stimulation of one IIa while recording from the opposite one, would provide a means for further more directly controlled studies of the inhibitory relationship.

The second area is the problem of obtaining more information on the nature of the relationship between neural elements from cross-correlograms. In the present work, this tool has been used primarily to demonstrate the existence of a relationship and to label the latter as inhibitory or excitatory. It is not possible, in general, to predict details of neuronal interconnections from a correlogram as several structures may have essentially the same correlogram signatures. Under certain limited conditions however, some inferences can be drawn from correlogram features [67]. The current PLEXUS - Phase 2 system could provide invaluable aid in pursuing this line of investigation, particularly in view of the fact that appropriate experimental data with which to test the results of modeling is readily available.

Finally, a related problem that was mentioned earlier involves development of methods for dealing with interaction between three or more units recorded simultaneously. The only work in this area of which we are aware is that of Perkel [66]. Even his methods appear to be limited to those cases in which a periodic stimulus is used. Here again, the PLEXUS - Phase 2 system could be of assistance.
Computer System Review

In addition to the neurophysiological research, an important objective of this work was investigation into and development of methods whereby profitable use could be made of a computer-aided data acquisition and analysis system, and more generally to explore the feasibility and usefulness of such a system by applying it to a specific set of biological questions. In this section we present a review of system performance along with a discussion of those features found most useful or valuable and of various difficulties encountered in using the system. We extend this discussion to include suggestions for modifications and additions to improve system performance, and directions for further development.

As we said before, as a result of certain characteristics of the Phase 2 system, two classes of individuals were associated with the system: programmers--those designing and implementing application processes; and users--those making use of the facilities so provided. Most individuals belonged exclusively to one group or the other. On the other hand, we have been by our own choice part of both groups. Thus we are in a position to consider the points of view of the programmer as well as the user in our review.

The ultimate decision about success or failure of the system rests with satisfaction of the user. Hence we consider his point of view first. The fundamental question is, did the system aid the investigator with the task at hand, was he able to accomplish tasks that he could or would not have otherwise undertaken.
Consider the following two examples taken from procedures used for detecting and measuring interaction between neural elements. During the experiment, SCAN and the cross-correlation process were used in pre- and sample-analysis mode to perform existence tests. By using SCAN while data were being collected, any or all of the data from the beginning of the run to within a few seconds of the current time could be displayed. Both average firing rate and dot mode displays could be examined for evidence of interaction. When a sufficient number of spikes had been obtained—determined by visual examination of the SCAN display—the cross-correlation process was called to compute and display a correlogram. Two factors are important here. First, it was not found possible to detect the existence of interaction via analog oscilloscope—computer processing of the data was essential for this. Second, the power of the system enabled producing results rapidly enough to help determine the course of the experiment. The time from the end of the run to completed correlogram was of the order of a minute or less for records of several thousand spikes. If no interaction were detected, electrodes were repositioned and a new run was begun immediately. If interaction were detected, its nature as shown by the correlogram signature was used to determine which combination of stimuli to use in order to investigate the interaction in more detail. The total time required from the end of the background or spontaneous firing run to the point where sufficient information had been obtained to make a decision about the next step of the experiment was thus but a minute or two.
Our second example illustrates application of the concept of "semi-automatic" analysis we introduced in Chapter II. Stimuli used in direct tests of interaction were repetitive presentations of various patterns. Resulting records were processed by computing PST histograms. Because equipment is rarely perfect, not all of the cycles in such a record were always useable. The problem was how to edit the data so that input to the PST process contained no faulty cycles. In hand calculations it is a simple matter just to "not include" such cycles. Of course hand processing tens of thousands of spikes is not feasible. On the other hand, blindly feeding all data into the PST process is not satisfactory either. But a program to recognize and delete artifacts automatically would, as we pointed out before, require an extremely long execution time if indeed it could be written at all. The solution developed was a combination of human and machine processing. Since recognition of "bad" cycles is easy for the user, appropriate editing facilities were added to SCAN. By means of a few simple operations, the user indicates those cycles of interest; the machine then forms them into a new record for analysis by the PST process--the other cycles are "not included".

These two examples clearly illustrate analysis methods that would not have been feasible without the system. In turn, these techniques formed an integral part of the methods used to obtain the results presented in Chapter V. For the present work then, the computer system played an essential role. From our observations of
other investigator's use of the system, they have also found it to be of considerable help in their research. Whether or not their results could have been obtained without the system is a moot point—in some cases, similar findings might have been made. However, without exception, it would have taken a great deal longer and the findings would have been based on fewer data. One investigator, for instance, reported that although he would still have been examining the same question, had he not had the system's facilities available it would not have been practicable to examine the response curve of interest at more than a small fraction of the number of points he has used.

In general, system features found most useful and powerful were, as might be expected, the very fast response, the ability to go rapidly and conveniently from one analysis technique to another, and the interactive aspects of the display system—particularly the latter. It is pertinent to consider these with respect to the stages of the overall analysis process as set forth in Chapter II: collect the data, decide on an analysis technique, perform it, examine the result, and decide on next technique. With regard to the latter two of these, the value of rapid visual feedback provided by the digital display system cannot be over-emphasized, especially when performing operations of the data massaging class in attempting to develop some insight into the nature of a particular response. In designing the various analysis processes, a great deal of effort was put into providing facilities for manipulating displayed curves in ways that would be convenient and
natural, and yet fast enough that users would feel that the system was
reacting to their actions in a continuous fashion. However, the capa-
bilities thus provided to us as a user (and to other users) have made
our efforts as a designer more than worthwhile.

The other features mentioned—speed of response and con-
venience in applying varying combinations of techniques—provided
for the necessary continuity in moving through the stages of analysis.
They also made it possible to bring different combinations of tech-
niques to bear during different stages of an experiment. During pre-
analysis, SCAN was used primarily as a flexible display tool. During
sample-analysis, SCAN and the cross-correlation processes were used
to detect the existence of interaction between neuronal elements. If
interaction were observed the faster of the two PST processes was
used in later sample-analysis periods. During post-analysis, spike
separation techniques were combined with the editing capabilities of
SCAN and the other PST process (see Chapter IV) to make more com-
plete, detailed measurements of the interaction.

One of the penalties of a successful system is that it is used—and this has certainly been the case for our system. Because it has
been used extensively, certain problems have become apparent with
increasing experience. In fact, one problem that has been growing
more severe has been simply the availability of machine time.

Because relatively little was known during the design stages
about how the data analyst would use such a system, a number of
assumptions had to be made. One assumption was that an investigator would move rapidly from collection of raw data through reduction of these to successively smaller amounts of data. These would be analyzed by increasingly refined, formal techniques. This expectation has been realized only to a limited extent by experience so far. Most analysis techniques are at a level not far removed from initial data reduction and consist, for the most part, of fairly straightforward processing of rather large data volumes with much emphasis on graphical manipulation and "eyeballing" [57] techniques. As a result, some of the shortcomings of Phase 2 that have become apparent with experience are problems with real-time use, deficiencies in secondary storage, and inconveniences in certain operating procedures.

Another assumption made in the design stages, that of the desirability and difficulty of constructing new and modifying current analysis processes, has been borne out by experience. The designers had hoped this problem might be at least partially alleviated by the modular approach of providing facilities for the user to construct a "program" from a sequence of more elementary analysis processes [52]. This might have been the case had the first expectation been realized. As it was not, one of the more serious drawbacks has been the programming barrier mentioned in Chapter II.

In the following, we examine these shortcomings in more detail in order to develop suggestions for improved system performance and
derive conclusions about directions for future systems.

**Secondary Storage.** The secondary storage medium used by the system consists entirely of disks, for the reasons given in Chapter IV. With this type of bulk memory it is extremely difficult to avoid occasional mishaps which result in portions of the data on disk being destroyed. The current solution to this problem has been to periodically copy the entire disk contents onto tape--usually once a week. This of course means that it is still possible to lose a week's data. The copying procedure itself is complex enough that it is usually done only by systems personnel.

A second minor disadvantage of the current disk storage scheme is that because of the large data volumes generated by a number of users, the disks must be treated somewhat as a temporary storage medium. Once data have been acquired, there is some pressure to finish whatever analysis is required as soon as possible so that old data may be deleted from the system to make room for new. Thus there is a definite need for a more archival form of secondary storage which goes beyond the simple back-up procedure. Although a single dataset can be saved and later restored in this way, the procedure is so unwieldy as to be unusable. As a result, much time is spent in attempting to ensure that one has obtained all desired information from a given dataset before it is deleted. Even if this is done, occasions will inevitably arise when one would like to test some new idea on data taken previously--and be unable to do so simply because the data no longer exist. A significant improvement
would be convenient means for the user to save and restore selected datasets or groups of datasets on tape. Not only would this provide desired archival storage, it would replace and improve the existing back-up procedure, since the user could back up just those data of interest, and at times of his choosing.

**System State.** With any computer system, and especially with an experimental one such as ours, there are occasional hardware and software malfunctions. Since some procedures take several minutes to execute, it is impossible to tell if things are proceeding normally or not without being in the same room as the central processor. Some effort should therefore be put into providing both visual and audible means whereby the user can monitor the current state of the system from his station, be it the experimental theater or a remote display terminal. This is particularly important in a time-sharing environment. Audible signals are especially desirable in the experimental theater because no effort is required to monitor them.

**Real-time Problems.** Certain areas of difficulty in using the system in real-time mode became apparent as experience with this mode of use grew. By real-time use is meant a combination of pre-analysis and sample-analysis—those procedures used during the course of an experiment. As we saw in Chapter IV, data were either sent directly to the computer during the experiment—in which case real-time use was possible—or were recorded on analog magnetic tape and sent to the computer some time after completion.
of the experiment and the computer was used for post-analysis only.

Difficulties in real-time use of the system arose from two sources. The first was simply that both LORI and the computer display unit were in a different room from the experiment itself. Although operating controls for LORI were in the experiment room, effective real-time use meant continually moving back and forth between rooms which proved to be a distinct disadvantage. A further inconvenience was that although LORI could be started and stopped from the experiment room, re-initializing the computer system for a new experimental run required performing various operations on the display unit, including typing in new dataset names. This last point will be discussed further below. In short, incomplete integration of the system into the experimental environment is not satisfactory.

Some efforts of an experimental nature were made in order to gain experience with a more completely integrated system. A closed circuit television system was set up to monitor the computer display. In addition, a primitive version of the SCAN program was written to be operated by means of the potentiometers and switches described in Chapter IV. Both monitor and potentiometers could be placed in the experiment room. It became apparent that while such a system could be made to work, rather extensive redesign both of Phase 2 and of the analysis processes would be required and the project was not continued. The brief experience with this scheme emphasized a second source of difficulties, namely, the
problems of operating additional equipment. It was, for example,
found quite difficult to manipulate the computer system while
performing other experimental procedures. Because of the nature
of many preparations, a great deal of the investigator's time during
an experiment must be spent on monitoring the preparation as
directly as possible. Although the computer can supplement the
oscilloscope and loudspeaker, it cannot completely replace them.
Operating additional equipment thus represents an extra drain on the
experimenter's energy and must be made as absolutely convenient as
possible. For example, it is essential that all of the controls for
operation of the computer during pre-analysis should be push buttons,
switches, potentiometers, etc. No typing of commands or names
should be required. Similar considerations also apply to sample-
analysis, although requirements are not as stringent. Even if such
requirements were met, it may still be that the efforts of two people
will be needed to perform the experiment and operate equipment if
the potential real-time capabilities of the computer are to be realized.

We mentioned above that typing in new dataset names for
each experimental run was found grossly inconvenient. This problem
appears in all stages of analysis. Each time a new process is called
in, dataset names must be typed in anew. This makes the set-up
part of Phase 2 quite general and quite tedious. However, it is our
experience, confirmed by other users, that during any given session
with Phase 2, only a relatively limited number of datasets is of
interest. A better method would be to require the system to store dataset names entered during one set-up phase and to re-display them at the start of the next set-up phase.

Although this modification improves the situation for post-analysis, it still requires action to be taken at the console, and is therefore not completely satisfactory for pre- or sample-analysis. More desirable would be a method enabling data collection facilities of the computer to be used as simply as a tape recorder—and from the experiment room. A simple solution is possible by making use of the subset facilities in Phase 2. With minor modifications to LORI and to the LCOLLECT data acquisition process, the set-up procedure need be performed but once during an experiment, instead of once for each experimental run—assuming LORI channel specifications are unchanged.

**Programming Barrier.** A serious drawback that the current system has in common with other present-day data collection and analysis systems (e.g. in experimental physics [62]) is the programming barrier referred to in Chapter II. The problem arises from the fact that the set of techniques used by the data analyst is not static, but changes as his understanding of his data changes. However, to modify a given analysis process or to construct a new one, changes or additions must be programmed in assembly language or FORTRAN. For the non-programmer user this is a major obstacle since he must either learn to program
in one of these languages or find someone to do his programming. Neither choice is satisfactory. For the first, the user must spend considerable time and effort to learn a language that operates at a level far removed from that of interest to him. Although the second choice is easier, it has two drawbacks. First, the user does not explore at first hand the potentials and limits of what the computer can do for him and thus can develop no feel for what is possible and what is not, what can be easily accomplished and what is a major programming effort. Second, since he must explain to a programmer what he wants done, there is every likelihood that the result may not be what he wants.

In the current system, even the programmer-user faces a programming barrier of sorts. Because of the desire to keep as much machine time as possible free for actual data collection and analysis, it was decided that all assemblies and compilations would be done on other machines. In order to modify a program, it was therefore necessary to go through several steps, including submitting a deck of cards to a batch processing system. As both programmer and user, the current system would have been considerably more useful to us had an assembler or compiler been an integral part of the system. Had this been the case, improved debugging aids would also have been possible. As an example of the need to modify techniques while using them, on occasion we resorted to hand patching in absolute machine language in order to see immediately the effect of a change.
In attempting to reduce the programming barrier, two questions arise. First, what facilities should the user be provided with so that he can vary his choice of techniques? Second, how should these facilities be presented? Two approaches to the first question are: (1) provide the user with a fairly complete set of processes, some of which may be combined in various ways to produce different analysis techniques, and (2) provide facilities so he may construct his own processes. The first approach is the one taken by Phase 2 and it has certainly provided investigators with a useful tool. Notice though that if the first approach is to be completely successful, the system designers must somehow discover all of the processes needed by an investigator or must discover which subset of them will provide a basis (in the mathematical sense) for the space of useful processes. Because, as has been said before, processes have been restricted primarily to graphical manipulation and straightforward statistical techniques, we have not been able to find a suitable subset. Since the analysis processes do operate at a level not far removed from the raw data, it is questionable whether such a subset exists. Future research may provide more insight into this problem. On the other hand, we are now beginning to develop a need for higher level methods which may be more amenable to the first approach. Even in this case, however, it is not clear just what would constitute an appropriate subset. For this reason and because of the dynamic nature of the set of
processes used by an investigator, we feel that the second approach of supplying facilities for the user to construct his own processes would be a more satisfactory one. The problem of just what facilities or what language or languages to provide the user is not a solved problem. Indeed, this general area is the subject of current research [84].

We wish to emphasize however that, whatever tools or languages are provided for construction of new techniques, they should be made an integral part of the analysis system, since building new analysis techniques is very much a part of the analysis procedure.

From the programmer's point of view, the facilities provided within Phase 2 for writing analysis processes have, with one or two exceptions, proven adequate and convenient. The exceptions arise from the restrictions imposed by the "block-diagram" nature [54] of the structure of a set of processes forming a "program", and from multiprogramming aspects of the system. In our experience, the two features most restricting were that it was not possible to form hierarchical structures of processes and that a dataset could not be both input to and output from the same process. The former problem was recognized by the designers of Phase 2 [52]. A possible means of improvement would be a method for processes to pass control explicitly rather than implicitly as is the case now. Addition of such features should also make it easier to remove the above restrictions on datasets.
Hardware. Reviewed here are those aspects of hardware performance and configuration relevant to considerations for future systems. With regard to central processor usage, two factors are important. First, in data analysis mode (as opposed to data acquisition or pre-analysis mode), the system is operated entirely on a demand basis—the system waits for and responds to user actions. Second, because most analysis methods consist of straightforward processing of large data volumes, for a given user action, far less time is spent in computation than in performing input-output operations. As a consequence, the central processor is idle (waiting for commands from the user) for considerable periods of time, and spends a significant fraction of its non-idle time doing input and output—and central processor time is available even while the latter is taking place. In our experience, these consequences are not atypical for the kinds of processing discussed.

Conclusions. What implications can we derive from the foregoing discussions for hardware and software considerations? First, accommodation of a Phase 2-like system within a more modest machine configuration than the present one should be feasible, although it is clear that more than a minimal configuration is required. In choosing a machine for this kind of work, the primary considerations are size and speed of bulk memory and central processor input-output capabilities. Other central processor characteristics, such as arithmetic capabilities, are less critical. Suitable means for data
acquisition (e.g. A/D converter, multiplexor), for display of results, and for interaction with the system are also essential. The latter is particularly important for real-time use of the system.

Second, what changes might be indicated for the present system? Two of the points discussed above bear on this question: (1) increasing use of the system has made availability of machine time a growing problem, and (2) due to the characteristics of current analysis processes, considerable central processor time is not used. A future project of some interest then would be a time-sharing version of Phase 2. Our experience would suggest that this be limited however, to two or perhaps three users at most, with the further restriction that only one be collecting data at a given time. Thus, one user's analysis and development of new analysis techniques could proceed concurrently with another's data acquisition.

Finally, despite the shortcomings discussed earlier, the Phase 2 data acquisition and analysis system has proven itself to be an important aid in neurophysiological investigations, and is fundamentally a satisfactory system.
LIST OF REFERENCES


39. IBM System/360 Component Description, "IBM 2250 Display Unit Model 1", IBM Systems Reference Library, File No. S360-03, Form A27-2701-0.


