# Adaptive Techniques for Signal Enhancement in the Human Electroencephalogram

by

Richard Winski B.Sc.

A thesis submitted to the University of Keele for the Degree of Doctor of Philosophy

November 1985

Device Applications Group,
Department of Physics,
University of Keele,
Staffordshire,
England.

TO MY FAMILY

## ABSTRACT.

This thesis describes an investigation of adaptive noise cancelling applied to human brain evoked potentials (EPs), with particular emphasis on visually evoked responses. The chief morphological features and signal properties of EPs are described. Consideration is given to the amplitude and spectral properties of the underlying spontaneous electroencephalogram and the importance of noise reduction techniques in EP studies is emphasised. A number of methods of enhancing EP waveforms are reviewed in the light of the known limitations of coherent signal averaging. These are shown to be generally inadequate for enhancing individual EP responses.

The theory of adaptive filters is reviewed with particular reference to adaptive transversal filters using the Widrow-Hoff algorithm. The theory of adaptive noise cancelling using correlated reference sources is presented, and new work is described which relates canceller performance to the magnitude-squared coherence function of the input signals. A novel filter structure, the gated adaptive filter, is presented and shown to yield improved cancellation without signal distortion when applied to repetitive transient signals in stationary noise under the condition of fast adaption. The signal processing software available is shown to be inadequate, and a comprehensive Fortran program developed for use on a PDP-11 computer is described.

The properties of human visual evoked potentials and the EEG are investigated in two normal adults using a montage of 7 occipital electrodes. Signal enhancement of EPs is shown to be possible by adaptive noise cancelling, and improvements in signal to noise in the range 2-10 dB are predicted. A discussion of filter strategies is

presented, and a detailed investigation of adaptive noise cancelling performed using a range of typical EP data. Assessment of the results confirms the proposal that substantial improvement in single EP response recognition is achieved by this technique.

**ACKNOWLEDGMENTS.**

I would like to record my gratitude to a number of people who have contributed significantly to the completion of this research, and in particular to the following:-

Professor W. Fuller, Head of the Depatartment of Physics, for the provision of research facilities and for valuable teaching opportunities within the department.

Professor D.M. MacKay who initiated the project while Head of the Department of Communication and Neuroscience, for much valuable supervision up to and beyond his retirement, and for freely offered hospitality.

Mr. N.M. Allinson who jointly supervised the project until the final stages, for providing much of the initial impetus and direction, and for establishing the research facilities.

Dr. R.E. Challis who supervised the latter stages of the research, for many valuable and detailed discussions, and for careful reading of several drafts of this thesis.

Professor E.F. Evans and the staff of the Department of Communication and Neuroscience for friendly advice and discussions, and in particular to Dr. D.A Jeffreys and Dr. M.J. Musselwhite for help with evoked potential experiments and analysis, and for acting as experimental subjects.

Mrs. E. Durrant, Mrs. H. Moors and Mrs. C. Owen for many helpful secretarial services throughout this project, and to Miss K. Pickerill and Mrs D. Evans for assistance in the preparation of this thesis.

Mr. S. Cartledge for professional drafting of several of the diagrams.

The technical staff of the Department for their skills in circuit construction, photography and mechanical construction.

My colleagues, in particular Mr. A H. Mukalaf and Mr. J.D. White for helping to create a pleasant working environment and for many hours of interesting if not always relevant discussions, and Mr D. Ball and Mr. E. Adem for assistance in the preparation of the bibliography.

A number of friends whose support and encouragement has been much valued, and in particular to Miss. C. Case and Mr. R. de Souza.

My family for their continual interest, encouragement and patience, to whom this thesis is dedicated.

## CONTENTS

## ABBREVIATIONS.

AEP - average evoked potential

AF - adaptive filter, adaptive filtering

ANC - adaptive noise canceller, adaptive noise cancelling

APWF - a posteriori Wiener filtering

ATF - adaptive transversal filter

CAI - Computer Automation Incorporated

CCA - cross-correlation averaging

DEC - Digital Equipment Corporation

DFT - discrete Fourier transform

DISPAC - digital signal processing package

EEG - electroencephalogram

EP - evoked potential

FFT - fast Fourier transform

GAF - gated adaptive filter

LCA - latency corrected averaging

LMS - least mean square

MGAF - multiple-reference gated adaptive filter

MMSE - minimum mean square error

MSC - magnitude-squared coherence

MSE - mean square error

PEP - pattern-evoked potential

SNPDR - signal-to-noise power density ratio

SNR - signal-to-noise ratio

TSAF - time-sequenced adaptive filter

TVF - time-varying filter

UPA - unwrapped phase averaging

VEP - visual evoked potential

# CHAPTER ONE.

# INTRODUCTION.

## 1.1 Introduction and historical background.

The field of evoked potential research has been intensively studied since 1947 when Dawson first demonstrated a technique that enabled the minute electrophysiological signals from peripheral nerve to be investigated [1]. Since then and particularly following the advances in instrumentation afforded by the development of semiconductor technology, many hundreds of papers have been written which describe attempts to characterise the nature of the signals that can be recorded from the peripheral and central nervous system in response to sensory stimulation. Present understanding of these evoked potentials (EPs) has made significant contributions to our understanding of brain sensory processes and the technique is now an important complement to traditional EEG analysis in fundamental research in neurophysiology, psychology and pharmacology, as well as offering unique diagnostic capability in clinical care. Typical examples of EPs are shown in fig. 1.1.

A fundamental difficulty in the investigation of these signals is the presence of unwanted signals in the raw data record. These have been generally referred to as 'noise' in traditional communications engineering, as opposed to the 'signal' representing the desired component [37]. Though this terminology is open to ambiguity, it commonly appears in the literature and will be retained for that reason. 'Noise' thus includes any deterministic or stochastic interference component, and will specifically include the background (unevoked) EEG

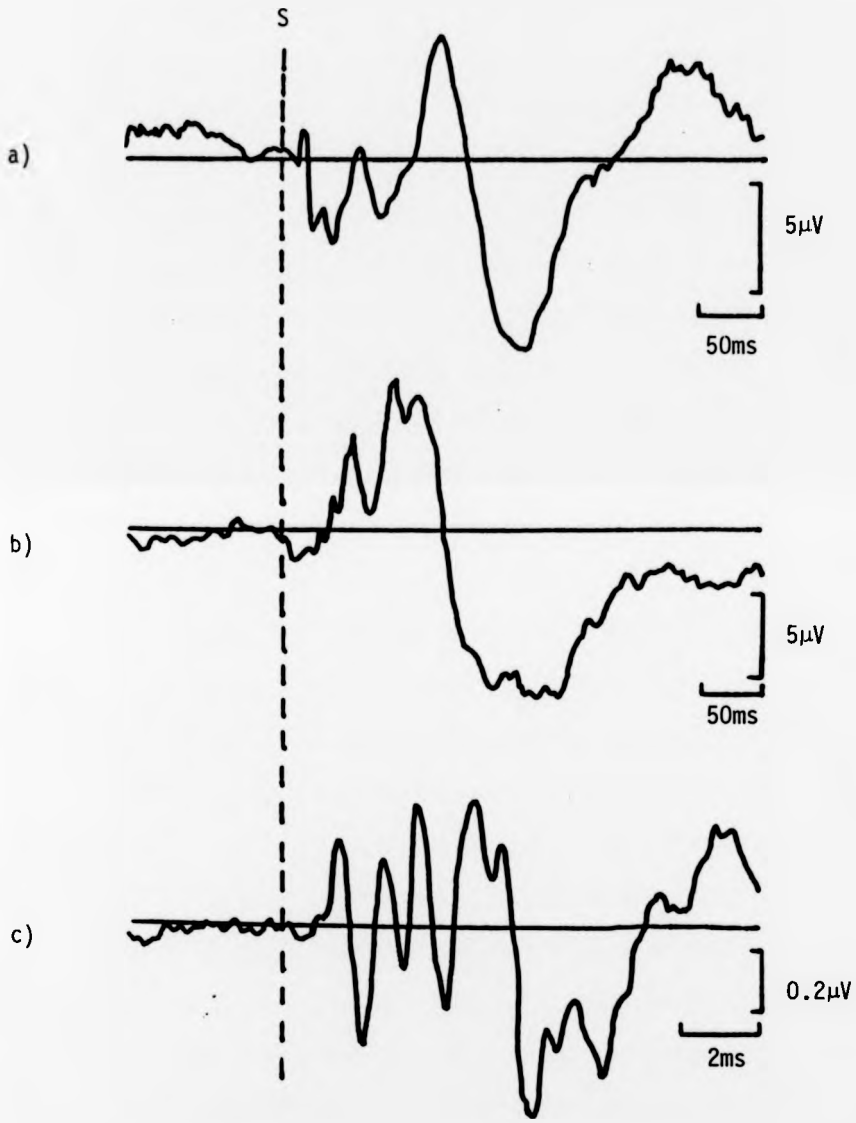



FIGURE 1.1 Typical examples of evoked potentials from the human brain (adapted from fig.5.5 in de Weerd & Kap [24]).
a) somatosensory evoked potential
b) visual flash evoked potential
c) brainstem auditory evoked potential
S indicates moment of stimulation. Note differences in calibration bars.

itself, random instrumentation noise and deterministic components such as mains interference.

While the amplitude of external interference can generally be reduced to acceptable levels by appropriate experimental techniques, such as use of differential amplifiers and correct grounding and screening procedures, it is not possible to avoid recording the spontaneous activity of the brain when scalp electrodes are used. The EEG remains the most serious interference component in recording cortical EPs due to its relatively high amplitude and generally similar spectral characteristics. Techniques are therefore required to eliminate or at least attenuate this activity.

## 1.2 Coherent signal averaging.

One of the earliest of such techniques, still in routine use today, is coherent signal averaging [38]. This technique gained wide popularity as it is straightforward to analyse, and implementation in dedicated instrumentation or general purpose computing equipment is not difficult. The basis of the method is that the amplitude of repetitive signals buried in uncorrelated random noise can be enhanced with respect to the mean noise level by the square root of the number of averages taken [39]. This performance can be achieved provided that certain assumptions concerning the signals are satisfied, principally that the signal is homogeneous (i.e. the signal mean at all sample points is invariant), and the noise is stationary and uncorrelated with the signal or with itself for lag times greater than the sampling epoch.

These assumptions are only approximately met in practice, and several examples are cited in the literature of significant variations

occurring in EP or EEG properties [26],[42]. In addition there are inherent limitations to the technique which render it unsuitable in some instances. The major limitation is that due to the poor signal to noise ratio a large number of responses are required (typically about 50 for cortical EPs and several hundred for brain-stem EPs) to obtain reliable and reasonably noise-free averages. This can result in long data recording sessions which a) limit the amount of information that can be gained in an investigation, b) increase the possibility of external artifact or nonstationary EEG activity giving erroneous results, c) increase the possibility of inaccurate results due to EP inhomogeneity and d) render the method impractical or difficult such as when testing babies or infants. Fundamental studies which require the individual responses to be made available are also unsuited to this method.

For these reasons interest has recently focussed on the development of methods that are more appropriate in these circumstances. Three different approaches have been taken. The first is concerned with developing modified averaging procedures that are relatively insensitive to changes in waveform latency or other signal inhomogeneity. These include temporal alignment of responses prior to averaging [41], weighted averaging [42] and selective averaging of homogenous subsets of data [43]. The second approach attempts to improve the efficiency of averaging to obtain a reduction in the number of averages required or to improve the statistical reliability of the results. These have included 'optimal' linear filtering methods such as Wiener filtering [45], and its time-varying counterpart [83], although de Weerd [113] recommends that these be employed to obtain better quality averages rather than to obtain a reduction in the number of averages employed. Finally others have abandoned averaging altogether and have sought ways of measuring individual response characteristics, generally peak amplitudes and

latencies, usually by statistical methods such as cross-correlation with a suitable template [46]. These latter methods do not furnish an estimate of the actual waveform. While these have extended the range of possible situations that can be investigated, they are often specific solutions to particular problems, rather that general purpose alternatives to signal averaging. In addition they can be rather demanding computationally which restricts their wider application in routine practice.

There is still considerable interest in developing other methods of processing EP signals which do not suffer from these disadvantages. The methods described are generally single dimensional, that is they take no account of the relationships between spatially different EEG signals. This inspired research into the possibility of exploiting redundant information in the multi-channel EEG to obtain signal enhancement. Since most routine EEG and EP studies of significance employ a number of electrodes, the possibility of obtaining improvements in signal estimation by this means are attractive and would not necessarily involve significant departures from existing practice. Improved signal estimation might lead to reduced recording durations, with consequent advantages for both research and clinical use of EPs.

### 1.3 Signal enhancement using correlated reference sources.

Previous studies have shown that significant correlations are present in the EEG over extensive scalp areas, and the work of Walter et al is typical of these [31]. Evoked potentials tend to be distributed about the cortical area devoted to each sensory process. This suggests that EEG activity in areas distant from the sensory region may be sufficiently well-correlated with the selected EP recording site to

permit some degree of cancellation to be gained. This obviously requires the EP to be considered as an independent signal added to the spontaneous brain activity which, though not strictly true in all cases [44], is nevertheless accepted to be a reasonable assumption for stimulus intensities that are much higher than threshold [45].

In order to assess the potential use of this technique, it was important to determine the nature and extent of correlations over the scalp which could be used to yield a reduction in spontaneous activity at other locations. Typical scalp distributions of common EPs were also required, to enable reference sites to be chosen that did not contain significant EP components. Visual evoked potentials were chosen to assess the viability of this approach as these are widely used in vision research as well as clinical care, and the experimental facilities were immediately available at Keele.

Previous experience and reports in the literature testify of the changing character of EEG signals, such as the waxing and waning of rhythmic components. Most workers in the EP field would agree that brain activity appears to take place in loosely coupled networks or neurons, which exhibit continually changing patterns of inter-relationship [7]. It was therefore decided at an early stage to develop a general method of cancelling background EEG activity using a correlated reference source, which makes no assumptions about the nature of the correlations with regard to frequency or time properties.

## 1.4 Adaptive techniques in signal processing.

Fortunately adaptive techniques exist which are potentially able to cancel noise activity by means of a correlated reference source. A

large part of this thesis is devoted to their investigation in order to determine their general suitability to EEG processing. Adaptive filters permit the amplitude and phase relations between correlated signals to be continuously modelled, and for this reason offer potential improvements over simpler techniques that assume fixed signal relationships. They are inherently self-designing and are well-suited to applications when a fixed filter is difficult or impossible to specify, as well as offering the capability of dealing with changing signal statistics. Pioneering work in adaptive filters took place over two decades ago [87], and it has since attracted widespread research interest. As a result the field is advancing rapidly, and such filters enjoy wide application, notably in communications and control engineering. Applications to biomedical signals have also been demonstrated, and include cancelling of mains interference [92], separation of maternal and fetal ECG activity [108] and filtering of electrogastrographic signals [109]. A similar approach has been described using non-adaptive techniques for the removal of ECG artifact in EMG signals [48].

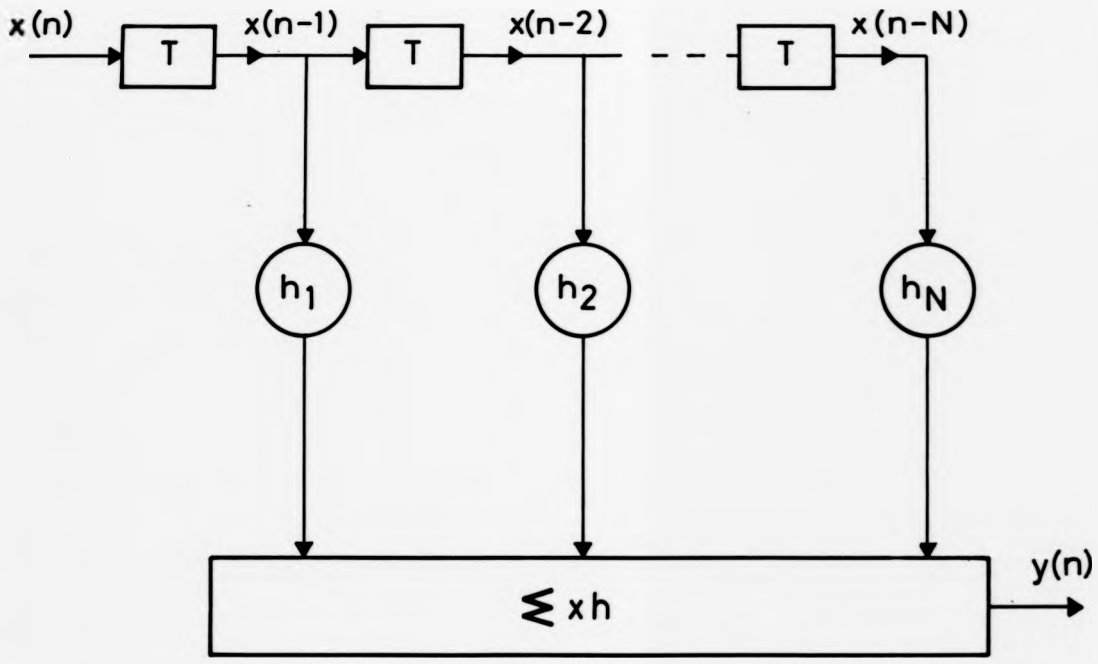
Adaptive filters have generally been implemented digitally, though analogue designs also exist. Digital filters have a number of attractive properties, such as simple design procedures, freedom from component and environmental influence upon performance, and the ability to accurately specify the parameters of high performance filters with precision. They can be implemented either in special-purpose hardware or in general-purpose computers. The latter possibility greatly facilitates experimental studies during the development of many different filter strategies whose operation can be analysed in considerable detail.

Fig. 1.2 shows a particularly useful structure which uses a tapped delay line filter whose coefficients are adjustable by a control algorithm to enable some criterion of filter performance to be met. Commonly this criterion is to minimise the mean square error of the filter when compared with some desired signal. Adaptive transversal filters of this form have been used in echo cancellation, system modelling and noise cancelling [50]. It is this latter mode of use that is relevant to this thesis as it permits a reference signal to be optimally filtered to permit cancellation of correlated activity in another channel.

This is illustrated in fig. 1.3 which shows two inputs to the adaptive noise canceller, a primary input containing a signal s(t) embedded in noise $n_p(t)$, and a reference channel containing correlated noise $n_r(t)$. If the degree of correlation between the two noise signals is high, effective cancellation of $n_p(t)$ can be obtained by appropriately filtering $n_r(t)$ so that it very nearly matches $n_p(t)$. Significant reduction of the noise in the primary channel can be obtained even if the signal and noise have similar spectral characteristics, and it is this property which makes this approach attractive as a means to signal enhancement of EP records. The main thrust of this thesis will be devoted to the investigation of this possibility.

## 1.5 The structure of this thesis.

In the next chapter the nature of EP and EEG signals is reviewed. Sections are included which discuss their origin, physiological attributes, applications and signal properties. This serves to illustrate the uses to which EPs can be put, and demonstrates




Transversal filter.

FIGURE 1.2 Diagram of a transversal filter based on a tapped delay line structure. The output y(n) is formed by summing weighted contributions of successively delayed input samples x(n), x(n-1)...x(n-N). T represents a time delay of one sample period, and $h_1$, $h_2$...$h_N$ are the filter coefficients.

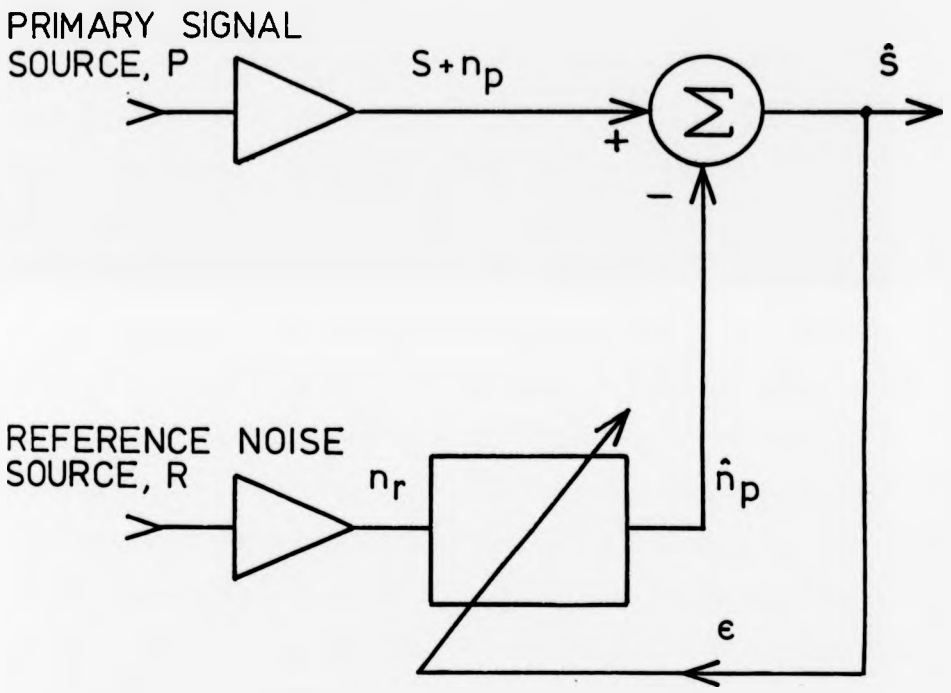



FIGURE 1.3 Block diagram of an adaptive noise canceller. The reference noise source $n_r$ is filtered by the adaptive filter to form a best least squares estimate of the noise $n_p$ in the primary signal source, yielding a best least squares estimate of the signal S at the canceller output.

that in the majority of experiments signal extraction procedures must be employed to recover them from the raw record.

Chapter 3 reviews a number of such signal processing methods. Coherent signal averaging is discussed first, particularly with regard to the main assumptions implicit in its use. Subsequent material deals with alternative or additional procedures that can be employed when these assumptions are not fully met. This discussion reveals that few methods of signal processing have been advanced to deal with multi-channel EEG derivations in a manner which exploits their inter-relationships. It clearly illustrates the necessity for this extended study.

Adaptive filters are introduced in chapter 4, and the main theoretical results derived or otherwise presented. Consideration is restricted to adaptive transversal filters using tapped delay line architectures controlled by the Widrow-Hoff Least Mean Square adaptive algorithm, though brief mention is made of other strategies. The theory of adaptive noise cancelling is introduced, and the main results presented that indicate how effective the method can be in different situations. Analysis of the main noise cancelling properties is extended in Appendix C where expressions are derived to predict noise cancellation performance using the coherence function. Chapter 4 concludes with a brief overview of previous applications of adaptive filters to biomedical signals, though adaptive noise cancellation of EP signals is considered to be novel.

This theory is extended in chapter 5 to incorporate modifications to the basic adaptive filter which form the proposed gated adaptive filter (GAF), which is based on the time-sequenced filter

[102]. The GAF was specifically developed to enable EP signals to be filtered using fast adaption which would otherwise result in distortion of the signal when using the basic adaptive filter or time-sequenced adaptive filter. Computer studies are described which compare each of these filters and show that the GAF is the best filter to apply when filtering repetitive transient signals using fast adaption.

Chapter 6 precedes the main experimental investigations by describing the experimental methodologies and equipment employed. Details are presented of EP experiments, including the stimulation and recording arrangements, and the off-line data processing capability. The latter had to be developed for this research programme, and discussion of the hardware and software strategies is presented. In particular, a general purpose signal processing software package developed by the author is described. This is written in Fortran IV and is suitable for a DEC PDP-11 computer having the appropriate harware configuration. Additional details concerning this package are contained in Appendix A.

Chapter 7 is concerned with a preliminary investigation of EP and EEG characteristics to determine whether the method of adaptive noise cancelling is feasible and what improvement in signal quality is predicted. The study considers the scalp distributions of visual evoked potentials and the coherence properties of the EEG over transverse occipital sites. Spectral functions are computed and primarily include auto-spectral power densities, cross-power spectral densities and the magnitude squared coherence for all possible electrode combinations in two subjects. From these results it is possible to specify the conditions necessary for effective use of an adaptive noise canceller, and several possibilities are forwarded for further investigation.

These possibilities are examined in detail in chapter 8 by the application of suitable adaptive filters. Practical considerations such as filter choice, parameter value selection and baseline removal methods are first dealt with and a comparison of the GAF and basic adaptive filter is undertaken using data selected from typical EP experiments. A detailed analysis of the adaptive noise cancelling procedure is given using several data records from each experimental subject under a variety of stimulus conditions. The conclusions of this study are presented in the final chapter, which considers the interpretation of the results and gives recommendations regarding the use of this technique. Methodological difficulties are also discussed and the chapter concludes with suggestions for future research.

Four appendices are included. The first contains additional material concerning the software package described in chapter 6. Appendix B summarises the computational procedures involved in performing the spectral analyses. Appendix C presents the theory which shows how the magnitude squared coherence function can be used to predict signal improvement through use of an adaptive noise canceller. Appendix D presents the results of a preliminary investigation into the use of a multi-reference adaptive filter for EP signal enhancement.

# CHAPTER TWO.

## THE ORIGIN, NATURE AND APPLICATIONS OF EVOKED POTENTIALS.

### 2.0 Introduction.

It is neither desirable, nor possible, to attempt a comprehensive review of the field of evoked potentials (EPs) in a work of this nature. For this the reader is referred to a number of monographs, review articles and other published material [2],[3],[4],[5],[6],[7]. A brief introduction and review will however be included to provide a sufficient background for the reader unfamiliar with this subject.

The material in this chapter includes a description of the nature of evoked potentials (primarily visual evoked potentials), and will discuss how they arise, what fundamental properties they possess and what value they are to the visual neurophysiologist and practising clinician. Following this, the remainder of the chapter will focus on the properties of EPs from a signal analysis perspective, from which it will become apparent that signal extraction procedures must be employed for their successful isolation. This discussion will also be concerned with the main features of the EEG which presents a fundamental obstacle to eliciting the noise-free EP. It is important to characterise the nature of both EP and background EEG signals as these must be adequately understood if signal extraction methods are to be correctly applied. Previous work in this area will be reviewed in the following chapter.

## 2.1 Definitions and nomenclature.

The electrical activity recorded by electrodes located on the scalp or cerebral surface is known as the electroencephalogram (EEG) and since its discovery in humans in 1929 [8] it has found widespread application as a research tool in many scientific disciplines, such as neurophysiology, psychology and pharmacy, as well as being a firmly established technique in clinical care. The EEG primarily reflects the general electrochemical activity of the brain and as such provides a useful monitor of cerebral state.

The method of evoked potentials is a more specific probe of brain functioning and is used to investigate the activity arising from externally applied stimulation. The electrical record following the stimulus is composed of endogenous activity reflecting the brain state of the individual at the time of stimulation, known as the spontaneous or background EEG, together with activity correlated with the stimulating event, variously known as the evoked potential (EP) or evoked response (ER). Evoked potentials are commonly defined as electrical potentials which are time-locked to the sensory stimulation which produces them. This definition, although not wholly adequate, has largely arisen from the use of signal averaging techniques, which permit small amplitude signals to be recovered from independent background activity on the basis of coherence with the applied stimulus. The difficulties of this definition are that it includes activity that may not be related to the brain's response such as stimulus-related artifacts, and does not include stimulus-related activity which is not coherent with the stimulus or is not of a deterministic form.

### 2.1.1 Cortical EPs to sensory stimulation.

The term EP is in fact used to describe a wide range of potentials ranging from the compound action potential of peripheral nerve lasting about 6 ms to highly complex cortical waveforms involving undetermined psychological factors and having a duration of several hundred ms. EPs may be recorded from the peripheral and central nervous systems. In the latter case it is possible to obtain spinal, brainstem and cortical EPs, according to the assumed structures from which they originate. In this thesis, attention will be focussed upon cortical visually evoked potentials.

Of the various kinds of EP, sensory EPs are the most familiar and extensively studied. They are elicited by stimulating any of the visual, auditory, somatosensory or olfactory organs (or their corresponding afferents), alone or in combination. Gustatory stimulation is difficult to arrange experimentally, though EPs have been obtained for these. The most common are visual evoked potentials (VEP), auditory evoked potentials (AEP) and somatosensory evoked potentials (SEP). Brain stem evoked potentials (BSEP) are also in this category and enable responses to be obtained, usually to auditory stimulation from the brain stem, a structure difficult to investigate by any other means. In this thesis examples will be shown mainly using VEPs, though some of the techniques described are also likely to be relevant to the other EPs. Typical examples of sensory EPs were shown in fig. 1.1.

The term event-related potential (ERP) is retained to describe electrical activity concomitant with an identified event (stimulation or action), whether or not it can be regarded as evoked by the stimulus. Examples of ERPs are motor potentials which accompany or antecede all

voluntary muscular contractions, and slow-wave potentials. The latter is a general category for endogenous potentials that typically arise when the stimulus undergoes evaluation by the subject for significance or has some cognitive significance. These have quite different spatial and temporal properties compared with sensory EPs, and will not be discussed further in this thesis. The reader is referred to the review by Hillyard and Picton [14] for further information regarding these.

### 2.1.2 'Transient' and 'Steady-state' EPs.

A further distinction may be drawn between so-called 'transient' and 'steady-state' EPs. The former arise when low stimulation rates are employed such that the response is largely attributable to each individual stimulus contains a minimal contribution from previous stimuli. They permit investigation of the temporal structure of responses and have generally required averaging techniques to recover them from the background activity.

Steady state EPs on the other hand are obtained when a train of periodic stimulation is employed at such a rate that the individual brain responses are confounded and the initial transient to the application of the train has decayed, leaving a steady state continuous response. Though temporal information is now largely lost it is nevertheless possible to obtain useful information from the amplitude and phase of the frequency components. Fourier analysers locked to the stimulus frequency or harmonics provide a convenient means of analysing only those components that are frequency locked to the stimulus, and is the means by which steady-state signal extraction is performed[2].

voluntary muscular contractions, and slow-wave potentials. The latter is a general category for endogenous potentials that typically arise when the stimulus undergoes evaluation by the subject for significance or has some cognitive significance. These have quite different spatial and temporal properties compared with sensory EPs, and will not be discussed further in this thesis. The reader is referred to the review by Hillyard and Picton [14] for further information regarding these.

### 2.1.2 'Transient' and 'Steady-state' EPs.

A further distinction may be drawn between so-called 'transient' and 'steady-state' EPs. The former arise when low stimulation rates are employed such that the response is largely attributable to each individual stimulus contains a minimal contribution from previous stimuli. They permit investigation of the temporal structure of responses and have generally required averaging techniques to recover them from the background activity.

Steady state EPs on the other hand are obtained when a train of periodic stimulation is employed at such a rate that the individual brain responses are confounded and the initial transient to the application of the train has decayed, leaving a steady state continuous response. Though temporal information is now largely lost it is nevertheless possible to obtain useful information from the amplitude and phase of the frequency components. Fourier analysers locked to the stimulus frequency or harmonics provide a convenient means of analysing only those components that are frequency locked to the stimulus, and is the means by which steady-state signal extraction is performed[2].

For a linear system either would yield a complete descriptor; but for a highly nonlinear system such as the brain both approaches can be required and can yield complementary information. Each has its own advantages and disadvantages, but transient EPs are generally preferred because they allow the temporal sequence of each response to be studied and related to known cortical structures and mechanisms on the basis of their anatomical distributions and temporal characteristics. For an introduction to the methods and applications of steady-state EPs the reader is referred to Regan's monograph [2].

### 2.1.3 Common conventions for labelling EP features.

Before proceeding to describe the origins of cortical VEPs, it will be appropriate at this stage to briefly describe the nomenclatures in common use to identify major EP features. These invariably concentrate on the major peaks and troughs in the response, and there are three ways commonly employed to do so

a) simple sequential ordering of components or peaks, e.g. C1, C2, C3
b) sequential ordering with polarity, e.g. N1, P2 represents the first major (negative) peak and the second (positive) peak
c) polarity with latency, e.g. P300 represents a positive peak occurring approximately 300 ms following the stimulus

Those which are purely sequential become inconvenient when intermediate features need to be included later, while those based on polarity or latency do not account for differences in response in different individuals or to different stimulus conditions. There is no generally adopted labelling scheme which is free from pitfalls, and all three are used by different groups. In this thesis the first will

generally be adopted, following the convention of Jeffreys [13] and others, though occasional use is made of the alternative nomenclatures.

## 2.2 Physiological and anatomical basis for the Visual Evoked Potential

### 2.2.1 The visual pathway.

Though the material in this section is contained in many elementary texts that describe visual neurophysiology such as [40], the main features of the visual pathway, shown in fig. 2.1, are now briefly reviewed. Fig. 2.1 also shows the major anatomical and physiological subdivisions of the human cortex. Receptor cells in the retina, both cones and rods, receive incoming light energy and convert this into graded electrical potentials. These are transmitted via other cells to the retinal ganglion cells which perform elementary image processing upon the signals and transmit image information in the form of spike trains. The axons of these cells converge to form the bundle of fibres known as the optic fibre. At the optic chiasm, the optic fibres from each eye meet and the nasal half-fields of each eye cross so that left-half field information from each eye proceeds to the right hemisphere and the right-half field maps to the left hemisphere.

From the optic chiasm, the fibres of the optic tract converge upon the lateral geniculate nucleus (LGN), which is primarily a processing station from which the signals arrive at the visual cortex. Some contrast enhancement is thought to take place at the LGN, and the nucleus may have a role in mechanisms of eye movement and binocular interactions, but otherwise little transformation of the visual image

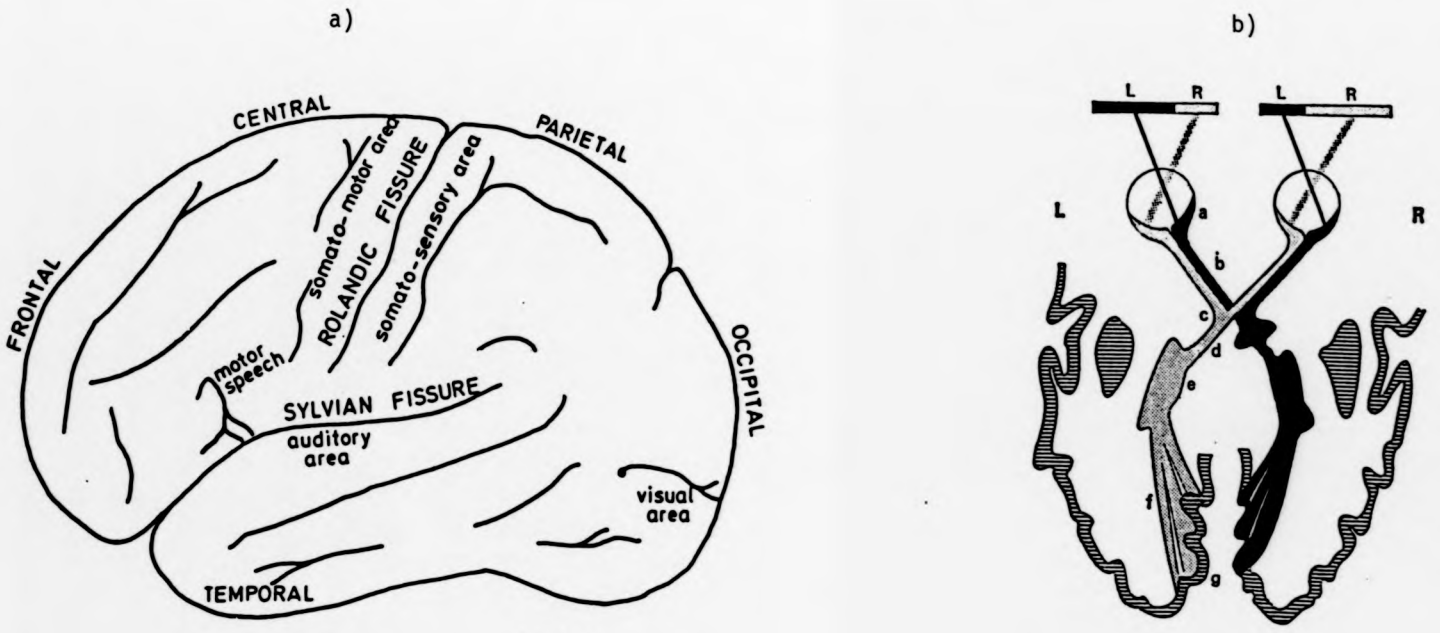



FIGURE 2.1 a) Lateral view of the human cortex showing the major anatomical and physiological features. Note the locations of the primary cortical regions responsible for sensori-motor function (adapted from fig.1.1 in Cooper, Osselton and Shaw [11]).

b) Diagram of the visual pathways of the brain
a. retina, b. optic nerve, c. optic chiasm, d. optic tract, e. lateral geniculus, f. optic radiation, g. primary visual area of the cortex. L and R signify Left and Right.

(Adapted from fig.21 in Luria, A.R., The Working Brain, Penguin Press, 1973).

takes place. A web of fibres streams out from the LGN to area 17 of the visual cortex in each hemisphere, which is the first cortical region involved in visual processing. Neighbouring regions are also involved, probably at a higher level of abstraction. Thus the secondary visual cortex appears to be the first region where binocular fusion takes place. Before this point, information from each eye is thought to be processed separately.

### 2.2.2 The cortical representation of visual information.

As scalp-recorded VEPs are believed to arise mainly in the visual cortex, little further will be said about the early visual pathway, apart from those aspects which have some bearing on the nature of the VEP. One such aspect concerns the projection of the retinal field upon visual cortex. In daylight levels of illumination, the rod cells, being much more sensitive to light, are completely saturated and play no part in vision. Although they are more numerous than the cone cells, (approximately 130 million compared with 7 million cones in each eye), they are evenly distributed over the visual field. The cone cells on the other hand are concentrated in the fovea, the central 1-2° of the visual field, and their density decreases inversely with eccentricity, being approximately 150,000/$mm^2$ at the fovea. It is the foveal region which is most sensitive to spatial discrimination, and this has consequences for the cortical representation of visual information.

The afferent fibres do not project haphazardly onto visual cortex, but are arranged retinotopically, that is in some correspondence to retinal location. The area of cortex is not proportional to the area of the retina it subserves, but more nearly to the number of receptors, whose density varies with retinal location. Foveal regions project to

disproportionately large areas of cortex which may be more than ten times larger than the area devoted to a corresponding solid angle in the periphery. This is partly attributable to the greater density of receptors and smaller receptive fields of ganglion cells serving the fovea, and is known as the 'magnification factor'. The same principle is demonstrated in the mapping of body surfaces to the somatosensory regions of cortex, where the cortical representations of the lips and fingers is far greater in area than that of the trunk, for example.

As a result of this it is the central 6-12° field of vision which is predominantly responsible for VEP generation. This is also due to the cortical topography of the primary visual area which is located at the occipital pole so that macular regions project to the exposed surface of the occipital lobe, but peripheral regions are located in the medial surfaces in the interhemispheric gap. A diagram showing the projection of retina to cortex may be seen in fig. 2.2 illustrating the magnification factor, retinotopic mapping and the transverse and lateral reversal of the visual field.

### 2.2.3 The effect of cortical structure upon VEP form.

The cortical representation of visual image information has a number of consequences for VEP studies which use patterned stimulation. Firstly approximate fixation of a well-defined pattern is necessary to ensure that the chosen retinal region is stimulated, as changes in fixation can lead to the cortical EP generator moving around the curved surface of the cortex and producing variable scalp potentials. Secondly the choice of area stimulated will have a bearing on the resulting VEP form, and interpretation of results must be made with attendant regard to the underlying anatomy and physiology. For example use of lower

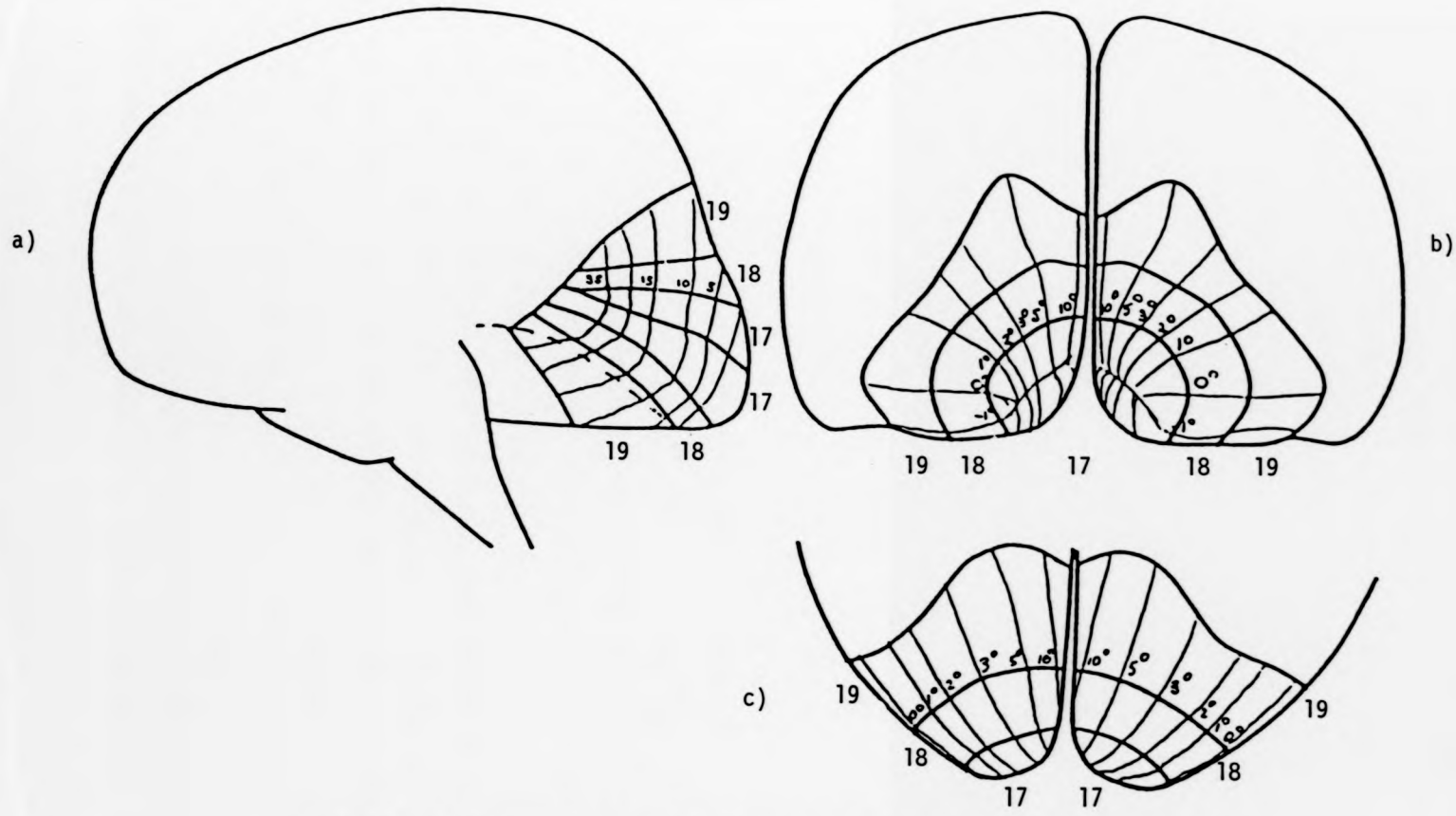



FIGURE 2.2 Diagram showing the retinotopic mapping of the visual field onto visual cortex, viewed a) laterally, b) posteriorly and c) vertically. Visual areas 17, 18 and 19 are labelled. Lines of constant eccentricity are drawn in each case, and show clearly that foveal regions project onto the most accessible cortical regions for scalp electrodes. (Adapted from Drasdo fig.1, in Barber [3]).

half-field stimulation is often preferable to upper-half field stimulation as there is less infolding of the cortex, resulting in larger responses over the scalp surfaces at the occiput. A third consequence is that responses from different individuals are likely to show a good deal of variability as the cortical topography is not exactly the same in each individual. These points will be illustrated in a later section describing gross properties of VEPs.

Data on the processes that form the VEP signals are sparse. While it is generally accepted that the VEP reflects organised activity of neuronal populations in the visual cortex, it is not clear how they arise nor what processes are involved, though some clues to these are available. This is partly due to our profound ignorance of the visual cortex, unlike the retina whose structure and functional properties are now largely understood.

## 2.3 Electrogenesis of the EEG and Evoked Potential.

This topic is subject to considerable confusion and uncertainty as reflected in the literature, so this discussion will be necessarily brief. Nonetheless it is acknowledged that research into the fundamental relationships between gross scalp records and underlying neural processes is crucial to the future of EP studies particularly in sensory physiology research, and therefore warrants further study [ 6].

### 2.3.1 Genesis of the EEG.

From a consideration of several factors, such as scalp distributions and correlations with cortical slow waves, it is generally accepted that the EEG and EP are predominantly of neural origin [6],

though it is possible for non-neurogenic components to be present, such as the EMG, EOG and ECG. The former two can generally be avoided in VEP experiments by careful precautions to avoid eye movements and to minimise muscular tension. The ECG is not generally present in VEP recordings. The use of high unnaturally intensity stimuli must also be avoided as these can generate stimulus-related myogenic activity [9].

The basic mechanism underlying the EEG is the movement of ions across the semi-permeable membranes of fibres or neurons in the cerebral cortex, subcortex or deeper nuclei. This ionic current gives rise to current flow in the tissues overlying the cortex - the dura, cerebrospinal fluid, skull, scalp, and hence to potential differences that may be measured with voltage amplifiers by means of electrodes attached to the scalp.

Initially it was thought that the EEG was the envelope of summated action potentials originating in cohesive neuronal assemblies. This hypothesis arose out of the success gained in relating neoronal firing patterns to motor activity. This gave way to the present hypothesis that the EEG is a complex summation of graded post-synaptic potentials (PSPs). This is a far more likely hypothesis as the relatively slow time courses of PSPs matches the EEG, and animal studies have demonstrated that an EEG is still obtained when spike activity is abolished by drugs. At present it appears that the pyramidal cells are responsible for the EEG, and summation of their PSPs is facilitated by the columnar organization of these cells extending from lower to upper cortical surfaces [10]. EPs are also thought to primarily reflect graded PSPs, though some EP components may be due to summed action potentials which are sufficiently well synchronised, such as axon volleys [2].

In general there is poor understanding of the locations and mechanisms of EEG generators. Alpha activity is thought to originate in visual association areas rather than primary visual cortex, but little is known of other sites of EEG origin. The locations of EP generating sites are also generally uncertain, with the exception of a small number of components. Several of the components of the BSEP to auditory stimulation have now been related to the passage of afferent signals through various sub-cortical nuclei [5].

### 2.3.2 Modelling of EEG transmission processes.

Following the successful application of field theory to modelling the ECG using dipole models and volume conduction, various workers have attempted the same for EPs. This theory is fundamentally unable to provide unique models for generative processes solely on the basis of observed scalp fields, and it is possible to infer equivalent dipole sources only in an infinite homogeneous conducting medium, which is not strictly the case for the brain. Nonetheless studies based on these assumptions have been used to predict amplitudes of cortical signals for sources of different sizes and at different depths in the cortex [47], and to hypothesize likely sites for some EP components [46]. This has had good success in some cases, particularly BSEPs which apparently do fit these assumptions quite well.

There is still considerable disagreement about this approach as brain tissue is not a good conductor, and volume conduction is only likely to be effective over short distances. It is also considered unlikely that EP generators can be modelled by large dipole surfaces under the scalp. Other theories to explain transmission of the EEG are based on conduction of potentials through specific pathways. The latter

is able to account for observations that closely spaced cortical electrodes can sometimes display very different activity whereas widely spaced electrodes can have similar activity. This is not easily explained if volume conduction were the main transmission mechanism. But this theory cannot account for the observation that BSEPs are transmitted over fairly large distances, apparently by volume conduction, when no other pathway exists.

### 2.3.3 The relation between the cortical and scalp EEG.

The scalp record does not contain as much information as the cortical record, and there is no clear picture demonstrated of a simple relationship existing between the two. There are marked similarities observed, with the scalp record generally related to underlying activity especially when it is synchronous over a cortical area subtending a large angle [11], but possibly also containing contributions from other cortical sites. Activity arising from sources of small angular subtense is heavily attenuated and may not even appear at all, as in the case of sharply localised spike activity [10]. Different workers report quite different degrees of attenuation through the scalp ranging from 50% [11] to orders of magnitude smaller [12], though this depends on the type of signal and its hypothesized source.

It is accepted by a number of workers that the material overlying the cortex acts as a spatial lowpass filter. Theoretical studies have shown this and are apparently supported by observations, though there is not universal agreement on this. As these studies were based on dipole models, it is not clear how much reliance can be placed on them. Some spatial filtering is likely to occur as this would account for the loss of detail in the scalp record as compared with the

cortical record, and the apparent enhancement of features in the scalp record when spatial deconvolution procedures have been employed [48]. Until this is resolved it is not clear how well the scalp record reflects underlying cerebral processes, though in a recent review Childers concludes that at present there is no general predictive model, nor one immediately foreseen, which successfully relates neural events to the EEG [7]. Nevertheless the ease with which the EEG can be recorded from the intact human scalp, and the variety of uses to which it has been put ensure a continuing and important role for EEG studies.

## 2.4 The gross nature of Visual Evoked Potentials.

In this section the nature of transient visually evoked potentials will be described. The description in this section closely follows a review by Jeffreys [4,Chapter 6], and will concentrate on the most important temporal and spatial properties of the VEP.

### 2.4.1 VEPs to flash and pattern stimuli.

Much of the early work was performed using unpatterned luminance-change stimuli produced by gas discharge tubes. The resulting flash evoked potentials have a complicated structure which has only relatively recently been comprehended as described in [3]. An example of their form is shown in fig. 2.3. In spite of a large body of experimental work, luminance EPs have not proved very useful in diagnostic procedures because of their poor correlation with perceptual thresholds. Furthermore the relative strengths of the channels giving rise to the primary, secondary and late responses varies substantially among subjects and with stimulus parameters. The variation in amplitude and latency that is found to exist makes the flash EP less suitable for

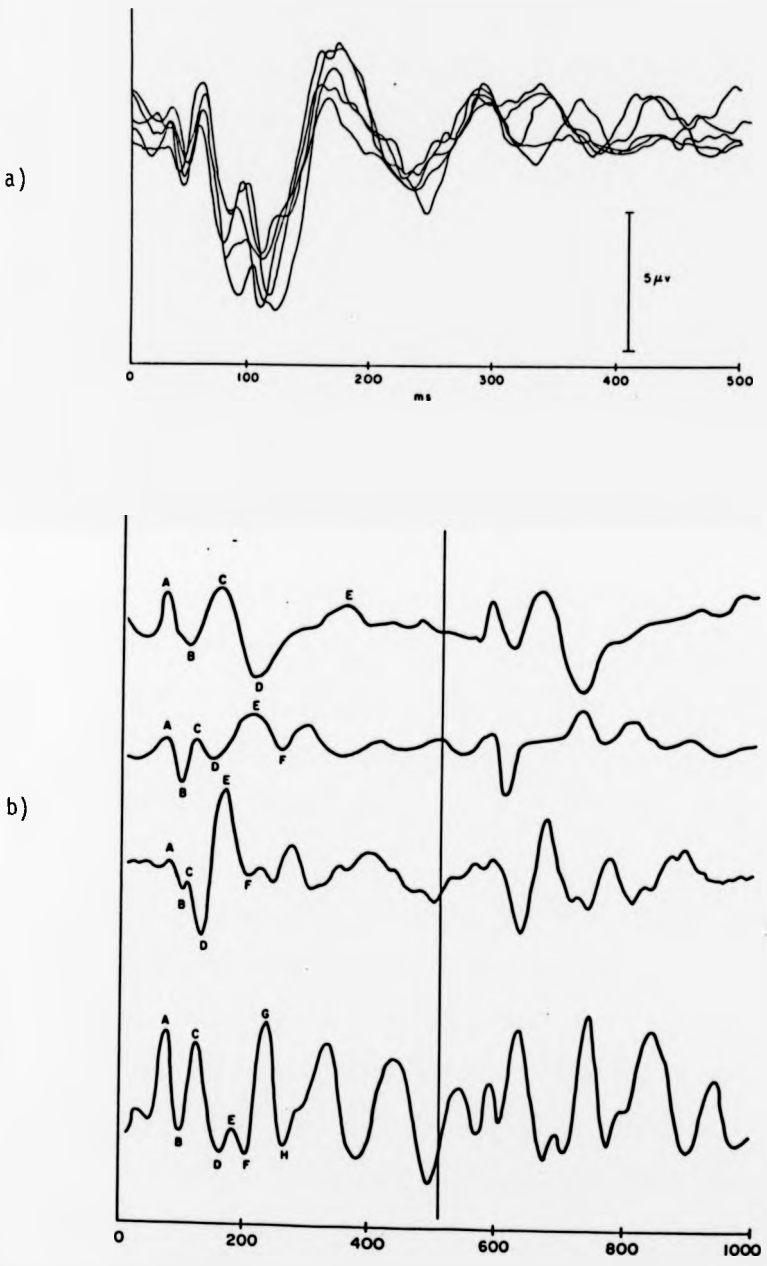



FIGURE 2.3 Illustration of the variability of human flash-evoked potentials in a) one subject and b) in four different subjects. In b) two successive VEPs are shown in each record. (Adapted from figs.2 and 3, Kinney JAS, J.Opt.Soc.Am., 67, p.1456 (1977) ).

differential diagnosis than for example the pattern EP. This is illustrated in fig. 2.3. Pattern evoked potentials have a number of advantages in both clinical and research environments which has made them by far the most popular choice of EP. Little further attention will be paid to flash EPs for this reason though it is noted that in certain situations they retain advantages over the latter, particularly when used with patients who cannot be relied upon to fixate, such as young children [20].

Patterned stimuli generally elicit much larger amplitude responses and provide a much richer range of stimulus possibilities than unpatterned stimuli. This has a firm physiological basis, as the (now classical) studies performed by Hubel and Wiesel in the early 1960's showed that (in cat and monkey at least) the majority of visual cortical neurones are sensitive to spatially structured stimuli and relatively insensitive to overall luminance changes. Patterned stimuli are much more useful to the visual scientist as they permit stimulus attributes such as contrast, stereoscopic depth, motion, colour etc. to be investigated.

The visual evoked response to patterned stimuli is a composite response having contributions from different cortical regions and/or processes. These often overlap in time and space and so must be investigated with careful attention to the use of multiple recording sites and location of retinal stimulation. Differences in form over the scalp are rather significant and have important consequences for the work in this thesis. Three different contributions to the VEP can be identified: the pattern-related onset and offset components, and non-pattern-related components.

### 2.4.2 Pattern-related components of the VEP.

Jeffreys and Axford [13] showed that pattern-related components of anatomically separate origin can be identified by studying the influence of retinal location of stimulation upon their scalp distributions. Three components related to pattern onset were identified, which they termed C1, C2 and C3 (C1 and C2 are shown in fig. 2.4). The components labelled as C3 are now thought to be composites of several components and will henceforth be referred to as 'C3'. These occur in all subjects tested and have similar peak latency times which are relatively independent of electrode position and retinal location. Their form and polarity do vary with these experimental conditions, which allows them to be identified. Though these show variability among individuals in their detailed form, they display many characteristic properties in general. One such property is the linear additivity of responses obtained to stimulation of different retinal regions. Thus the response to whole-field stimulation closely resembles the linear sum of the responses obtained to stimulation of each quadrant alone.

Fig. 2.4 shows the temporal characteristics of C1 and C2. Their peak latencies are approximately 75 and 110 ms after stimulus onset. Note the temporal overlap of these components which renders consideration of EP peaks alone inadequate for proper investigation of EPs. It is all too possible for amplitude and latency differences to be observed in peaks which are due to variations in the relative contributions of these components to the EP, and thus confuse interpretation of the EP [49]. A number of studies have been made of the properties of these components, and the main conclusions are now presented.

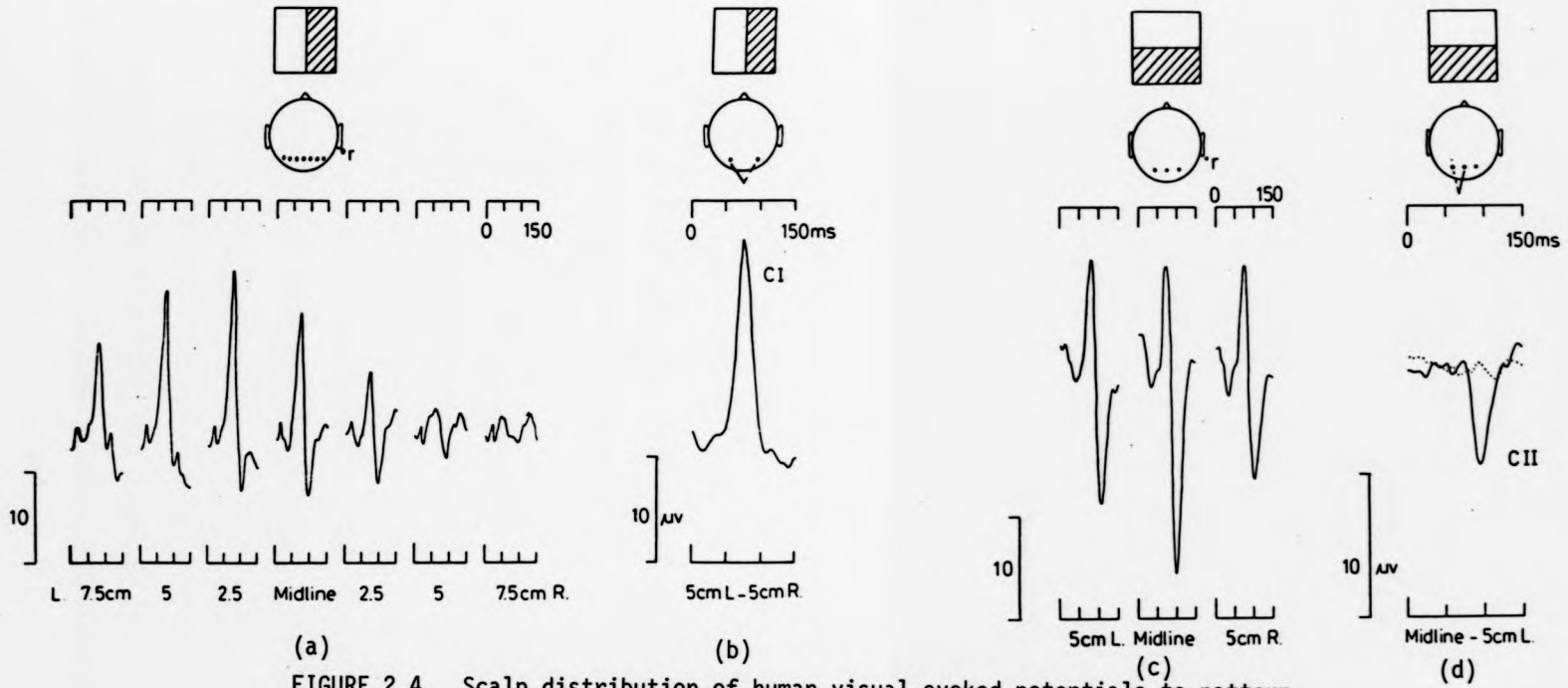



FIGURE 2.4 Scalp distribution of human visual evoked potentials to pattern onset using (a) right half field, and (c) lower half field stimulation. Temporal properties of VEP components C1 and C2 are shown in (b) and (d) with the electrode arrangement used to obtain enhanced derivation of each of these components. Note the temporal overlap of C1 and C2. Electrodes mounted on an arc 4 cm above inion. (Adapted from figs 3 and 4 in Barber, Chapter 15 [3]).

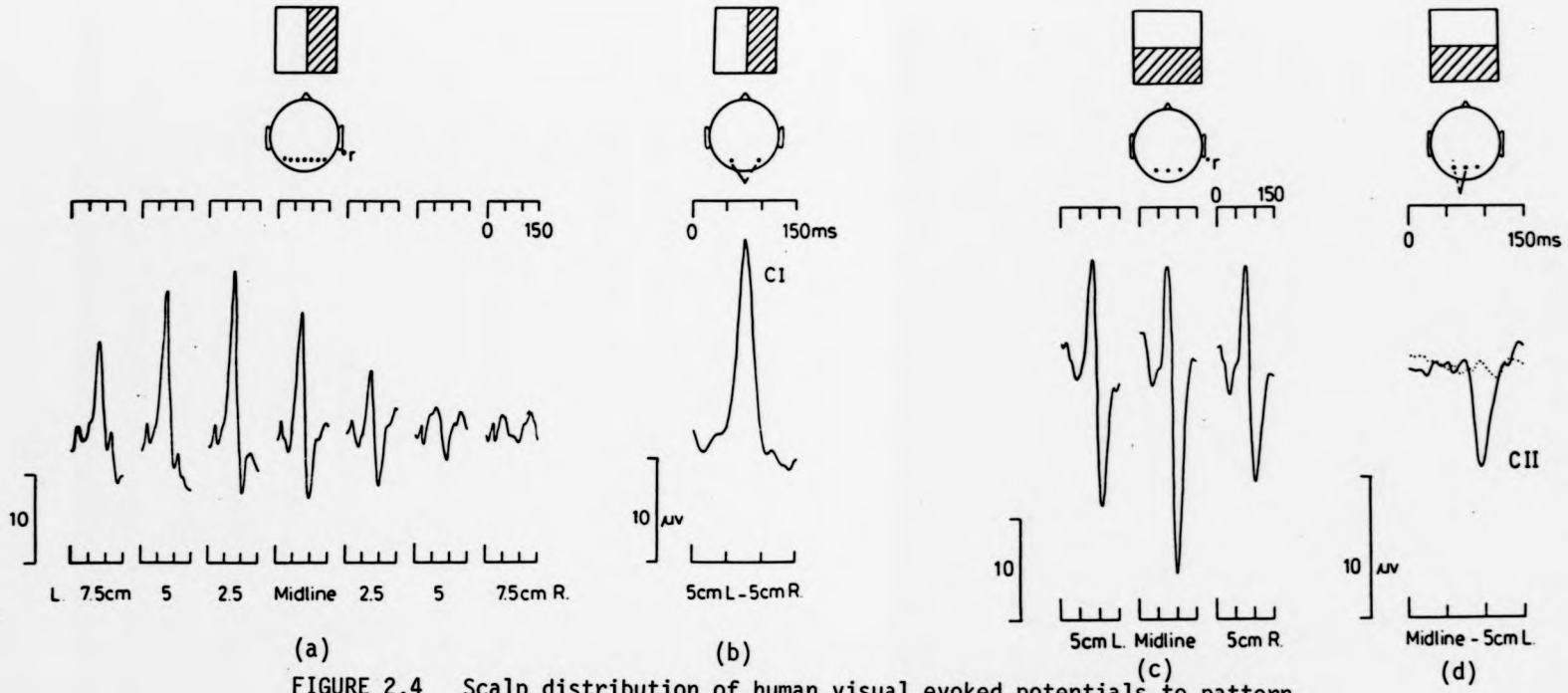



FIGURE 2.4 Scalp distribution of human visual evoked potentials to pattern onset using (a) right half field, and (c) lower half field stimulation. Temporal properties of VEP components C1 and C2 are shown in (b) and (d) with the electrode arrangement used to obtain enhanced derivation of each of these components. Note the temporal overlap of C1 and C2. Electrodes mounted on an arc 4 cm above inion. (Adapted from figs 3 and 4 in Barber, Chapter 15 [3]).

Firstly, they have quite different amplitude distributions over the scalp according to the region of retina stimulated. This allows experiments to be conducted that selectively favour recording of one particular component, and thus permit its properties to be deduced. To illustrate this fig. 2.4a shows a set of EPs obtained from a transverse row of 7 electrodes to a patterned stimulus presented to the lower visual half-field. The C1 component reverses polarity across the scalp, whereas C2 remains relatively unchanged. Leads 2 and 6 in this case can be used as a bipolar pair to enhance C1. A similar scheme to isolate C2 is shown in fig. 2.4b, though the C2 component has a similar distribution to the 'C3' components and so is more difficult to isolate.

Experiments designed to investigate the nature of these components have concluded that C1 largely reflects contrast-specific mechanisms of vision. These are sensitive to patterned stimuli (both to onset and offset) but not overall luminance changes, and they exhibit a systematic increase in response amplitude with increasing contrast over a relatively large range of contrast. They do not require the presence of well-defined contours.

C2 (and 'C3') on the other hand apparently reflect contour-specific mechanisms which respond best to well-defined contours in the stimulus pattern but are less sensitive to actual contrast. They are orientation and dimension specific and are most sensitive to discontinuous contours. Unlike the former contrast-specific response, they appear only at stimulus onset and are highly adaptive. Several similarities in these properties suggests a possible relationship with the 'contour-sensitive' neurones found in cat and monkey visual cortex. It is possible to obtain clues to their likely anatomical sources from a consideration of their scalp distributions to retinal stimulation and

stimulus properties, but at present there is no general agreement on the sites of origin.

### 2.4.3 Pattern-related offset components.

The most consistent pattern-related offset components appear to be two peaks occurring at about 110 and 150 ms following stimulus offset, often preceded by a component corresponding to C1. These main offset peaks have quite different properties from C2 and 'C3', though their latencies correspond. For example they vary less with retinal location stimulated and exhibit less adaptation. Brief stimulation, which minimises adaptation effects in the onset response and hence permits faster stimulation rates, has the disadvantage that the offset components are confounded with the onset components. As these are relatively small compared with the onset components in most cases this is generally not a difficulty, provided one is only interested in the onset components.

### 2.4.4 Non-pattern-related components.

In addition to the pattern-related components there are also non-pattern-related components which are sensitive to overall luminance changes. These vary from subject to subject, but in most cases are relatively small in the latency range 0-150 ms where the most important pattern-related components occur, though they may produce peaks at approximately 100 and 150 ms. These correspond to components of similar character but larger amplitude in the flash EP, and so are likely to be luminance-related components. An important finding is that in contrast to the pattern-related onset components, these are relatively unaffected by retinal location, nor do they show an additive amplitude relationship

of the type exhibited by C1 and C2. This suggests that they have different cortical sources and that these sources are essentially independent of each other.

The greatest non-pattern-related contributions occur in the later response with peaks at about 200 and 300 ms. These have a more anterior distribution with a maximum near the vertex, and have been shown to be non-stimulus specific. Unlike the components described so far which vary systematically with stimulus parameters, these later endogenous components are affected by the cognitive state of the subject and appear to reflect stimulus evaluation [14]. Examples of these are the P300 wave which appears to be correlated with the improbability of the stimulus event, and the Contingent Negative Variation or Expectancy wave which appears when a warning stimulus is very likely to be followed by a second stimulus demanding a response. Further details of these are available in the review by Hillyard and Picton [14].

## 2.5 Applications of Evoked Potentials.

In spite of the severe experimental difficulties that are encountered in recording EPs there are a number of situations for which this method is well-suited and even uniquely so. Users of EPs tend to be mainly interested in sensory physiology and perceptual mechanisms or are clinical users, though there is increasing interest in the later cognitive brain responses among psychologists and others.

### 2.5.1 The use of EPs in brain sensory research.

Research into sensory processes has made most significant advances through traditional psychophysical methods and more recently through neurophysiological studies at the level of individual neurones. The advent of microelectrode recording techniques has probably been most responsible for our present understanding of mechanistic brain science leading for example to the pioneering discoveries of Hubel and Wiesel in the early 1960's. Although this has paved the way for extraordinarily fruitful though painstaking research, it is limited in that it can only provide information regarding the behaviour of a single cell (or at most a small number of cells). This is necessarily inadequate if we are to understand the organized behaviour of a large number of functionally related units.

Psychophysics has also greatly contributed to our present understanding of sensory perception, largely through detailed investigation of the limits of perception, and has provided good evidence for some higher level features that can be expected to be found in the visual system. For instance a body of psychophysical evidence has been accumulated to support the notion that the sensory pathways of the brain break down complex sensory stimuli into a small number of abstract features that are processed virtually independently in different 'channels'. These may be for perception of form, motion, colour etc., and neurophysiological evidence exists for cells that could embody these processes. In the main, the psychophysical findings have relied on soundly based, though often restricted, studies performed mainly on humans using threshold stimuli and requiring statistical treatment of a large number of subjective reports of sensory experience.

Evoked potentials on the other hand are thought to reflect the organized activity of functional sub-units in the brain, though some critics are sceptical of the potential value of EP recording, owing to uncertainty about what they represent and the poor resolution they offer. They are however the only electrophysiological means available to investigate brain function in man, as intra-cerebral recordings may not be made except in rare cases when they are advised for medical purposes.

As well as providing independent electrophysiological evidence to support psychophysical measures, EPs offer possibilities that may not be attained through psychophysical methods alone. In situations where there is a close correlation between EPs and sensory perception, EPs offer a single means of objectively investigating sensory processes at super-threshold and sub-threshold perception levels as well as threshold levels. As the former corresponds more closely with normal experience, it may be more meaningful than investigating perception at threshold when the processes involved are close to the limit of operation.

Some of the areas where VEPs correlate well with perception are contrast threshold and adaption, real and apparent motion, spectral sensitivity of the eye and stereoscopic depth perception [15]. EP studies of these phenomena provide further information on brain processes as well as enabling the nature of the VEP to be elicited. Some of this research has had ramifications for clinical and ophthalmic practice; for example VEP methods have been developed to perform routine refractions. These can offer advantages over traditional methods in young or difficult patients.

Research is also proceeding into other possible uses of EPs. Carmon at al [16] have used EPs to obtain an objective measure of pain sensation. To minimize suffering of the experimental subject it is obviously necessary to use the minimum number of (painful) stimuli. This in turn demands advanced signal extraction techniques to obtain reliable results from single responses. Other research areas that have a similar requirement are studies into the possibility of feedback control based on evaluation of the stimulus [17] and studies on dynamic properties of single EP responses [18].

When poor correlations exist, EPs offer a means of studying sensory processes which are necessary for perception but which are not accessible to psychophysical means because they occur at peripheral level or do not directly intrude into conscious experience. An example of the former situation is the processing required to perform 'housekeeping' duties such as accomodation in the visual system. The latter situation is found in patients who are cortically blind through removal of some or all of the visual cortex, but who may nonetheless have some EP response to visual stimuli, showing that some visual processing apparatus remains, though insufficient to support normal vision. This discussion demonstrates that EP techniques provide experimental possibilities that may not be realised through psychophysics or single-cell studies alone, and that fruitful integration of these methods is possible in present-day brain research.

### 2.5.2 Applications of EPs in clinical medicine.

The use of EPs in clinical practice suffered from an initial loss of credibility due to unfounded optimism initially placed in the technique when it was still in its infancy. They have since proved to

be valuable in a number of situations including diagnosis, treatment monitoring and direct involvement in surgical procedures, though it is fair to say that at present they only occupy a minor role in clinical practice compared with the more widely-used EEG.

EP techniques have proved useful in the diagnosis of disorders of the visual, auditory and somatosensory pathways. Lesions in these are often shown up as abnormalities of the EP, and in general there is good correlation between the severity of the sensory loss and EP abnormality. Information on the site of the loss can also be obtained from the nature of the abnormality in some cases. For example, large and significant delays are present in the VEP in conditions such as optic neuritis and in progressively demyelinating diseases such as multiple sclerosis. These delays can be reliably detected at an early stage even before clinical signs are present, and the use of EPs in the diagnosis of multiple sclerosis is one of the most successful applications of the EP technique [3].

A large number of applications take advantage of the objectivity of EP measures. Thus diagnoses can be made in cases when the patient is unable or unwilling to cooperate. This can be the case with babies and young children, coma patients or adults suffering from senile dementia or mental disorders. Harding [19] has demonstrated the value of EPs in assessing the integrity of the eye and visual pathway in accident victims who have suffered severe optic trauma. This can aid in treatment planning. Much interest has also focussed on the use of EPs in assessing visual and auditory function in babies and young children from the first weeks of life onward. This aids detection of abnormal neural development at an early stage when the system is still plastic and can be reversed. It also aids early diagnosis of diseases that

affect these pathways and permits detection of defects that can be corrected through the use of eyeglasses or hearing aids in time to prevent retarded development. Some examples of this work are described by Harden et al [20]

Other applications have been the use of EPs in surgery to locate more effectively the site to be removed, and to measure brain function under the influence of drugs or pathway dysfunction. Applications in psychiatry, in particular to the discrimination of psychogenic from pathological disorders, is one area where EPs provide a unique objective tool. Although it is fair to say that the applications of EPs in the clinical field are not as developed nor as widely adopted as for instance those of the EEG, there has been growing interest in the technique which is likely to continue in the foreseeable future.

## 2.6 Gross properties of Evoked Potential and EEG signals.

This section is concerned with a description of the EP and EEG in terms of their gross signal properties. A summary of these is important to enable suitable signal processing methods to be used and applied correctly. Indeed the very need for such methods is indicated by a consideration of this material. Following a general consideration of interference sources, this section will be concerned with the most important properties of the EP and then with the EEG, which is the most serious interference component as it is unavoidably present in the scalp record and has many similarities to EP activity.

The very low amplitudes of EP signals requires elaborate procedures to record them with minimum instrumentation noise and external artifacts. This involves thorough scalp and electrode

preparation, use of low noise differential amplifiers, careful screening and earthing procedures and care in siting the experimental equipment to avoid a.c. mains interference. With sufficient attention to these it is possible to reduce extraneous interference to acceptable levels.

In addition to these, interference from biological sources must also be guarded against. The four most common sources of biological interference in EP recordings are the ECG, EOG, EMG and unevoked EEG activity itself. ECG signals are generally not recorded at scalp sites unless a non-cephalic reference is used, which is generally avoided for this reason. EOG activity can be quite serious when the eyes are rotated due to a standing potential of several mV that exists between the cornea and retina of each eye. This can give rise to large artifacts, as rotation of the eyes through 10° can generate a d.c. potential shift of about 120 μV. This is not a major problem when visual fixation is employed, as it is in most VEP experiments, provided that fixation is maintained. Several techniques exist to remove these artifacts should this be necessary, e.g. [21],[22]. Examples of some of these artifacts are shown in fig. 2.5.

Activity generated by cranial or facial muscles can arise in the EP record, though again these can be minimised by suitable precautions. This includes adoption of a relaxed posture by the subject and avoidance of unnecessary body, head or facial movements. These can not always be avoided completely, particularly if inexperienced subjects or patients are involved, but are not usually a major problem when experienced subjects are used. Most muscle activity recorded at the scalp occurs in the main signal bandwidth, from about 14 Hz extending to beyond 100 Hz, and generally peaking in the range 30-60 Hz [23]. The EEG exhibits increased noise activity when these muscles are contracted, but there

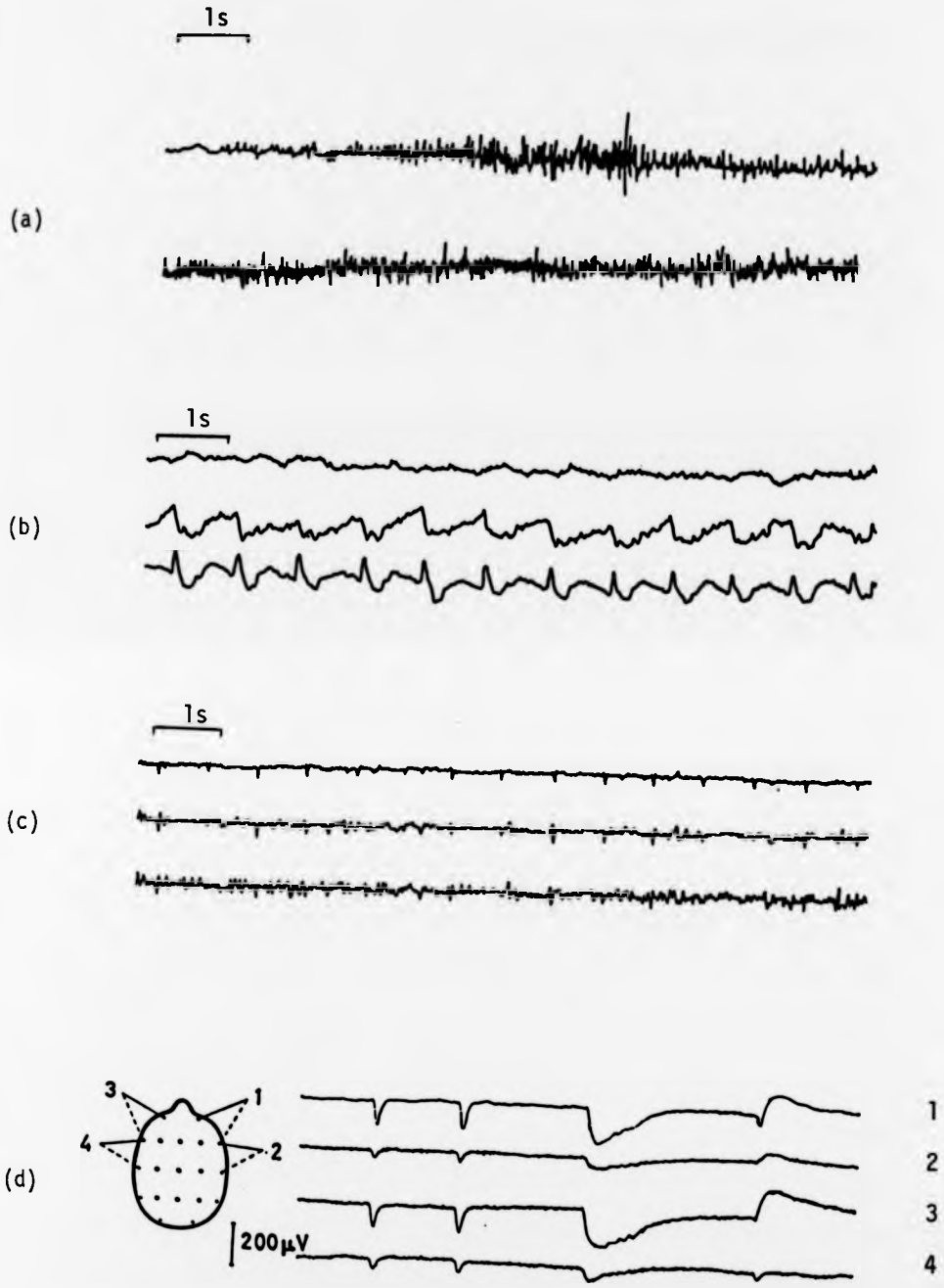



FIGURE 2.5 Examples of common artifacts in EEG records a) muscle activity, b) pulse artifact caused by electrode overlying an artery, c) ECG artifact, d) eyeblink and eyelid closure. (Adapted from figs. 5.9, 5.11 and 5.12 in Cooper, Osselton & Shaw [10]).

seems to be little correlation between this activity and the EMG recorded for the muscles involved, unless the scalp electrode lies directly over the muscle, indicating that nonlinear transmission processes are involved. The ongoing activity of the brain presents a fundamental obstacle to eliciting noise-free EPs, and for this reason will be considered in detail in a separate section later.

### 2.6.1 EP signal characteristics.

As previously discussed, many different types of EP can be recorded which have quite different signal properties. It is therefore intended to restrict this discussion to cortical EPs, with VEPs being the main examples in view of their importance to this thesis. In discussing these signals it is important to remember that the precise form of the response is greatly affected by the stimulation procedure, and that it varies in different individuals. The description must therefore be rather general, though sufficiently detailed to allow the merits of different techniques to be assessed. The discussion of EP signals centres on their amplitude range, temporal and spectral properties, and on the extent of the variability encountered. The significance of these will become evident when they are considered in relation to the EEG.

The main components of cortical EPs rarely exceed 30 μV and are often considerably less. The amplitudes are commonly measured from one peak to an adjacent trough, though some workers use a pre-response baseline to establish a reference level. The latter avoids confounding the effects of different components in the measurement. The EP is generally polyphasic in form, typically lasts 100-300 ms, and generally has several peaks of varying duration in the range 10-100 ms. Owing to

the transient nature of the response (and the variation in characteristic form) it is generally inappropriate to use traditional spectrum analysis methods which assume time-invariance of the signal, though some estimate of spectral occupancy can be gained from these. It is better to examine how the frequency structure of the signal varies with signal duration, and this can be achieved by passing the signal through a series of bandpass filters of different centre frequency, or by employing joint time-frequency spectral methods [24].

Fig. 2.6 shows the result of applying such a method to a typical VEP and AEP, and clearly shows that a time-invariant spectral analysis can yield misleading results. From these examples it is clear that a fairly large bandwidth (approx. 200 Hz) is required for an accurate representation of the signal. Nevertheless the majority of clinical and research work employs much lower bandwidths, typically about 50 Hz. This choice of bandwidth, though sub-optimal, is generally sufficient to characterise the main components of the EP and affords some attenuation of higher frequency interference, such as muscle-related activity.

The variability in form of the responses for different subjects and for different stimuli has already been mentioned. The latter clearly reflects physiological differences in perceiving different stimuli, and thus permits these processes to be studied. This has already been considered in a previous section of this chapter. Differences in the form of responses to the same stimuli in different individuals also exist, and hinder direct comparison of responses from a pool of subjects. Statistical criteria can be defined by which abnormal responses can be identified, but signals from different subjects cannot be directly pooled for processing as though they belonged to one population. Apart from multivariate statistical procedures, techniques

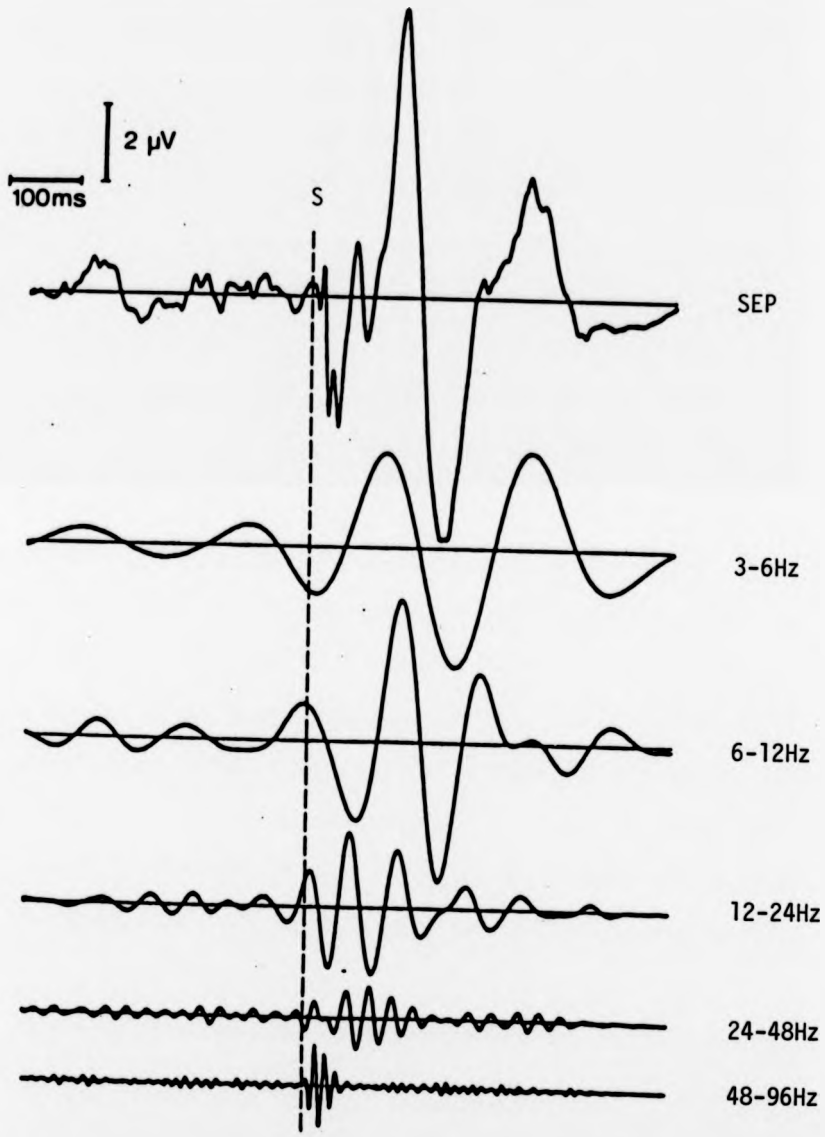



FIGURE 2.6 Spectral decomposition of a somatosensory EP obtained by passing the EP through a bank of bandpass filters with the indicated bandwidths. S denotes stimulus onset time. (Adapted from fig.5.1 in de Weerd and Kap [24]).

of signal processing are generally confined to a set of responses from a single individual to one stimulus.

Even within a set of such responses it is possible that trial-to-trial variations exist in the responses. It is important to consider this possibility for two reasons. Firstly some signal processing methods, such as averaging, implicitly assume that the underlying responses are identical and their efficiency or accuracy are impaired if this is not the case. This possibility is considered in more detail in the next chapter. The second reason is that variability can be an intrinsic source of information which may help our understanding of sensory processes, just as the variation in form arising in different individuals or with different stimuli can yield useful information. For this reason there is a growing interest in methods of analysis that are sensitive to individual response detail, though these are still at an early stage of development.

It is not clear from inspection of the raw records themselves whether variability arises in the responses themselves or is due to the background EEG. In general the average EP is quite consistent and does not show marked variations in form. A substantial body of evidence does however indicate that the responses vary on a trial to trial basis.

Many studies e.g. [25] have shown that the first response to a train of stimuli is significantly different from the remaining ones. This is illustrated in fig. 2.7 using VEP data from our own lab which were averaged according to trial number. Several investigations have shown that habituation of EPs does occur, with an exponentially decaying amplitude response and other characteristics of habituation [18]. Apart from differences in the first response this generally manifests itself

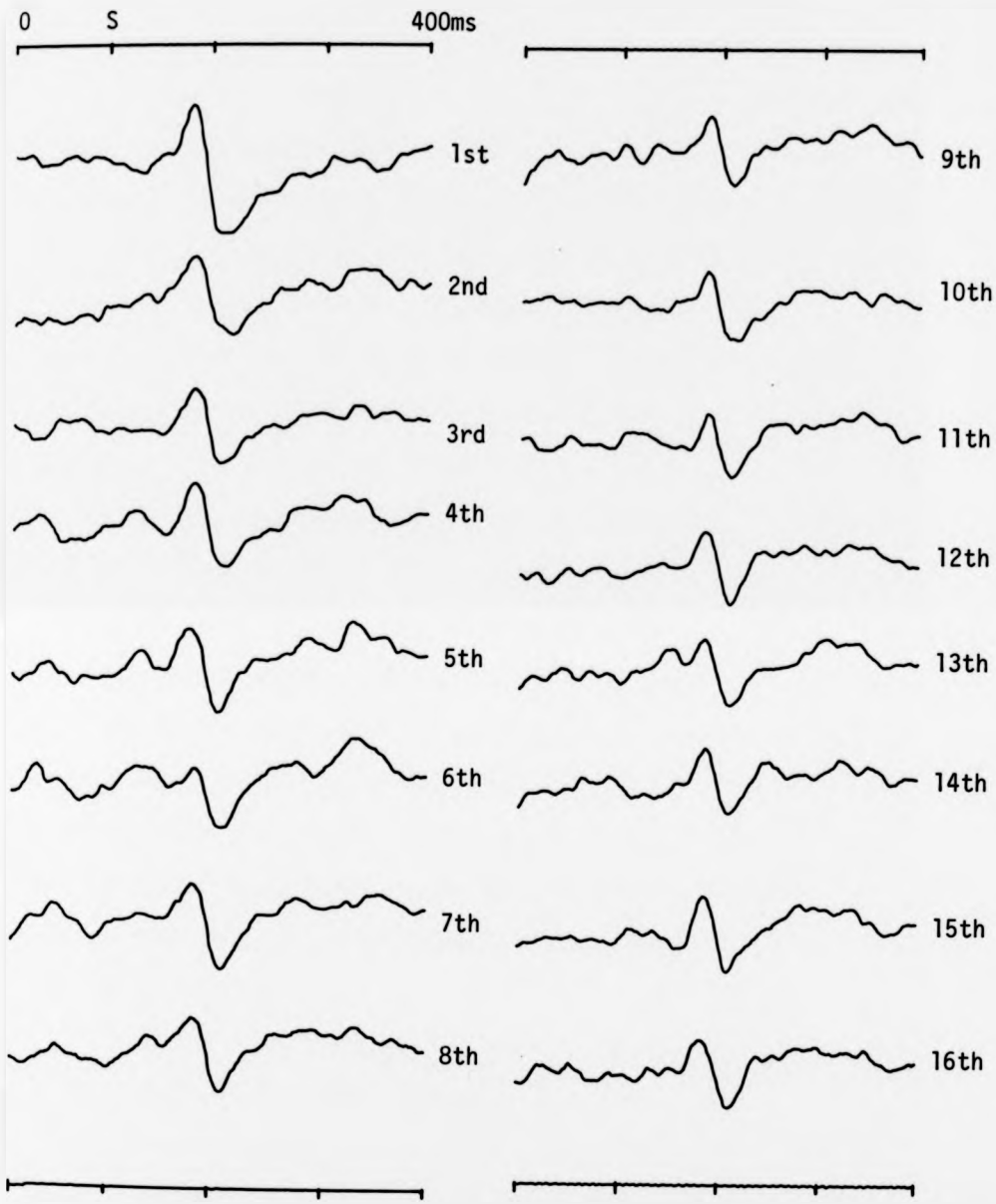



FIGURE 2.7 Average VEPs obtained as a function of stimulus order. The first waveform is the average of the first response in each of 16 separate runs, each run comprising 16 stimulus presentations. There is little evidence of habituation except that the first response is larger than succeeding ones. EPs obtained by presenting the stimulus for 500ms, at approximately 1Hz repetition rate.

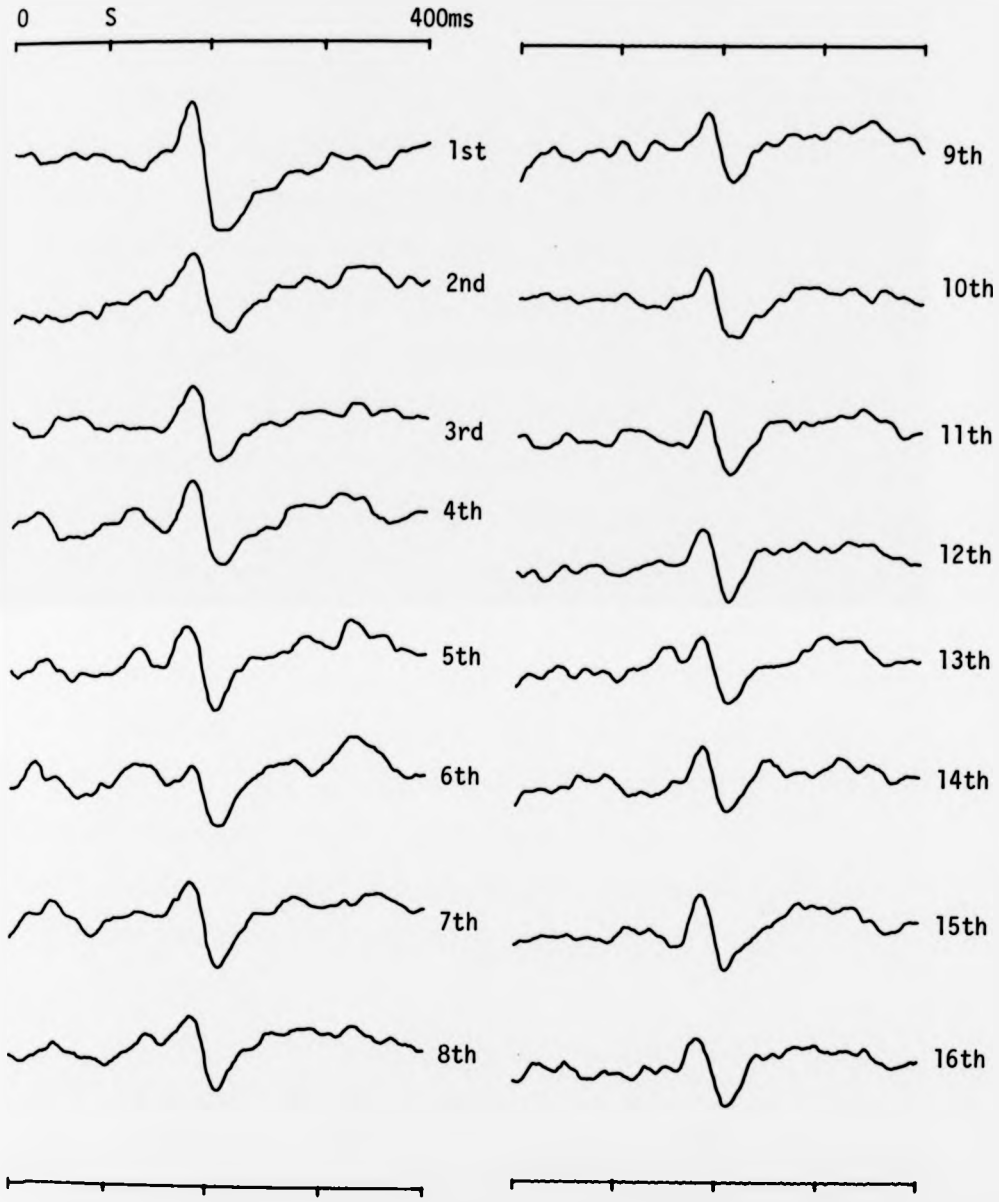



FIGURE 2.7 Average VEPs obtained as a function of stimulus order. The first waveform is the average of the first response in each of 16 separate runs, each run comprising 16 stimulus presentations. There is little evidence of habituation except that the first response is larger than succeeding ones. EPs obtained by presenting the stimulus for 500ms, at approximately 1Hz repetition rate.

as a slowly decaying amplitude over the course of the experiment, and usually does not seriously affect the form of the averaged response.

Variations in form have also been shown to occur, and these have greater affect upon the average, as they 'smear' the fine detail of the response. The most easily illustrated example of this is variation of the latency of waveforms with regard to the stimulus onset. Studies have shown that the latency is affected by stimulus intensity and duration, such that near-threshold stimuli exhibit a significantly delayed latency compared with high intensity stimuli. It has been demonstrated by a number of workers that variability in the latency of individual responses to identical photic stimuli occurs in the range 5-20 ms. Some workers have also explored the possibility that individual peaks vary in their latency and good evidence has been shown for this, in for example the work of Aunon and McGillem [26], though this has not yet been confirmed for other stimuli types or sensory modalities. It does indicate however that caution is required before assuming that each response is identical, though in general the variation in individual responses has not been sufficiently serious to affect widespread use of procedures that assume response invariance, such as averaging.

### 2.6.2 EEG signal characteristics.

The EEG is a continuous signal that is more readily analysed by traditional signal analysis methods. As with EPs the form of the EEG varies considerably with scalp recording position and from individual to individual, and is generally quite stable in frequency composition within a recording session provided that the subject remains in a stable state of wakefulness. The EEG takes on markedly different character

when the subject becomes drowsy or falls asleep, but these changes are not of concern here as VEP experiments are generally carried out using alert subjects.

The characteristics of EEG signals are described in elementary texts such as [27],[28], but the main features of relevance to this study will be be presented here. The EEG appears to be a randomly varying signal, of approximate amplitude in the range 20-100 μV, with main signal power in a bandwidth of about 50 Hz, though frequency components beyond this can be recorded. A typical amplitude spectrum of a normal adult EEG is shown in fig. 2.8. Within the signal there are often strong rhythmic components, and early EEG investigations were primarily concerned with characterising these and discovering their relation to cerebral function and clinical abnormality. While clinical electroencephalography has advanced a great deal since then, routine analysis of records still include a description of EEG activity in terms of the main alpha (8-13 Hz), beta (>13 Hz), theta (4-8 Hz) and delta (<4 Hz) bands. These will now be briefly described as they constitute the main activity of the normal waking EEG. Examples of these rhythms can be seen in fig. 2.9.

The most common rhythmic component in the waking EEG is the alpha rhythm, and this has received a great deal of research interest. A distinction must be drawn between general activity in the alpha band and the alpha rhythm. The latter is defined as "rhythmic activity in the frequency range 8-13 Hz, being most prominent in the posterior regions, present most markedly when the eyes are closed and attenuated during attention, especially visual" [29].

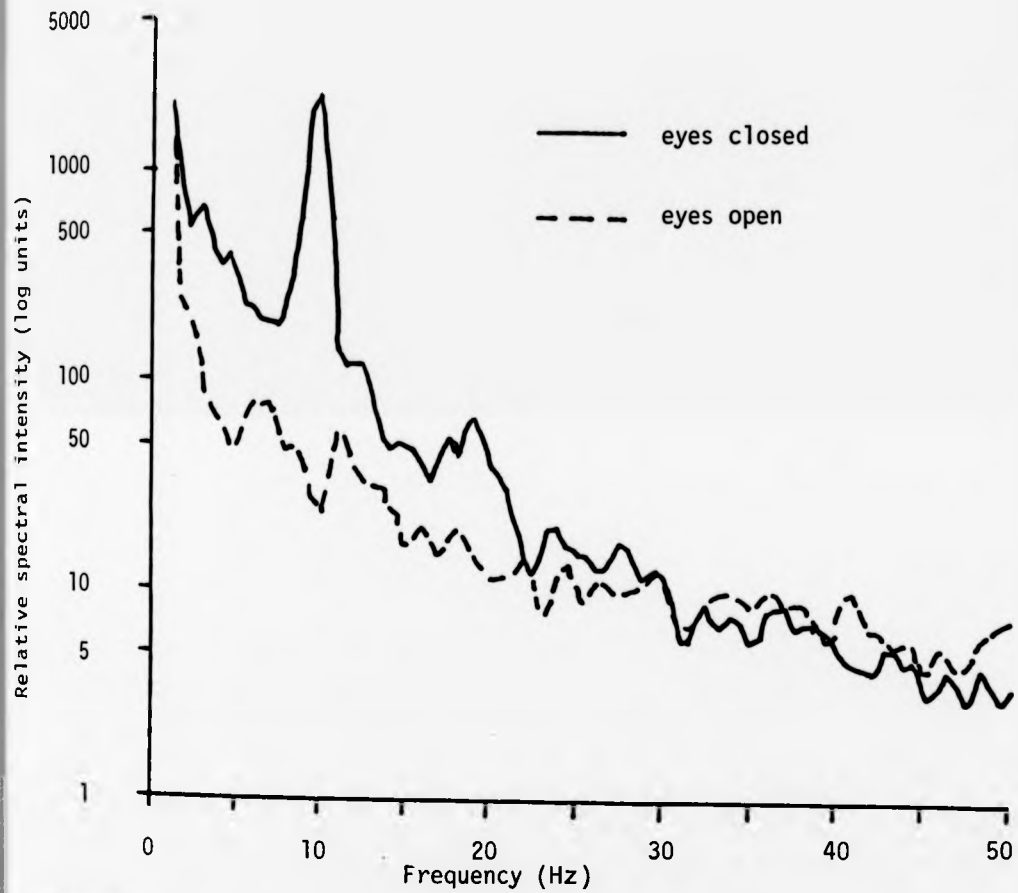



FIGURE 2.8 Power spectral intensity for normal adult waking EEG with open and closed eyes. Recordings made at occipital scalp referentially to earlobe. Note prominent alpha (8-13Hz) and beta (>24Hz) activity. (Adapted from fig.5.1 in Kooi [27]).

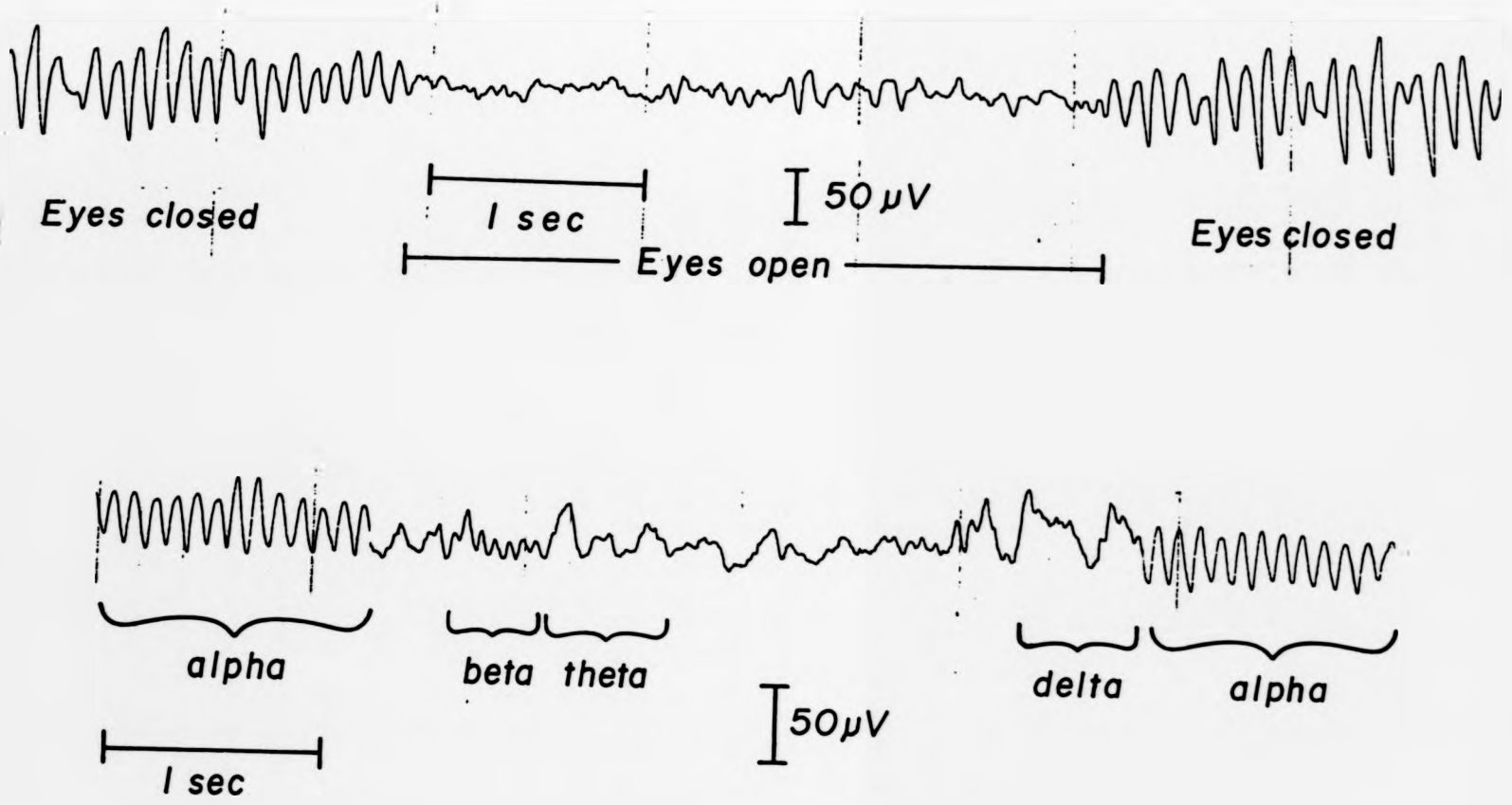



FIGURE 2.9 Examples of different EEG rhythmic activity. Note attenuation of alpha activity in a) when eyes are opened, b) shows an EEG record which, unusually, exhibits each of the alpha, beta, theta and delta rhythms. (Adapted from figs.1.2 and 1.4 in Tyner, Knott and Mayer [28]).

About 85% of healthy adults have an alpha rhythm in the frequency range 9.5-10.5 Hz, though some rare individuals have a mixture of the normal alpha rhythm and a different frequency rhythm, often half its frequency. The centre alpha frequency is quite stable in individuals, even from day to day, and varies less than 0.5 Hz, except at the onset of drowsiness or sleep. In about 66% of the population the amplitude is in the range 20-60 μV and it rarely exceeds 100 μV peak to peak.

The amount of alpha activity present in the waking EEG varies from person to person, and usually waxes and wanes in bursts or spindles lasting 1-2 s. In most individuals it is present for 20-80% of the record, though about 10% of normal adults show very little regular alpha, and a small proportion have a steady unvarying rhythm. Occipital alpha tends to be slightly asymmetrically distributed, and usually occurs with a greater amplitude and higher incidence over the right hemisphere, though the reason for this is unknown. Although it is often thought of as being sinusoidal, this is incorrect and the alpha rhythm can depart significantly from a sinusoid, even to having a spiky appearance. Asymmetry of negative and positive half-cycles is also common.

In addition to the occipital rhythm, independent foci of alpha activity can arise in fronto-central and temporal regions in about 20% of normal adults. In both locations the frequency is often lower than the alpha rhythm, generally in the range 7-11 Hz with a mean frequency of 9 Hz. In the central region the wave may have an n- or u-shaped form and is known as the mu rhythm. As with the alpha rhythm it can be fairly constant or intermittent. Unlike the occipital alpha rhythm it is not affected by eye closure or visual stimulation, but is blocked by

real, intended or imagined motion, or tactile stimulation, particularly of the hands. It may be either bilateral or unilateral, and if the former, may appear independently in either hemisphere.

Beta activity is defined as activity occurring with a frequency greater than 13 Hz. Normally it is of fairly low amplitude and rarely exceeds 20 μV. it is commonly seen in central and frontal regions where it appears as an 18-24 Hz rhythm. One form of beta activity becomes evident when the alpha rhythm becomes desynchronized, and it appears either as rhythmic or randomly organized activity. In general beta activity is bilaterally synchronous, though it may have asymmetric features.

Theta activity (that is activity occurring in the 4-8 Hz band), is generally inconspicuous in the normal waking EEG. It is not uncommon to find it in central areas, though it is seldom the main pattern and tends to be of low amplitude (<15 μV) and intermixed with alpha and beta waves. It is rare for rhythmic activity to occur. Large amplitude theta activity (greater than 50% of the alpha amplitude) is evidence of clinical abnormality.

Rhythmic delta activity (that is frequency components less than 4 Hz) is also rarely found, though it can occur episodically over occipital regions. Brief mixed alpha-theta bursts or single theta or delta waves may occasionally occur independently in the left or right temporal region or posterior region in some adults. These are often seen mixed with alpha waves in young adults and can be quite large, in the range 75-150 μV, but are attenuated by opening of the eyes. They often occur in both hemispheres but are usually asymmetric, being larger over the right hemisphere. They have no known clinical significance.

### 2.6.3 EEG correlations over the scalp.

While there have been very many papers written on spectral analysis of the EEG, there have been relatively few comprehensive studies of the correlation properties of the EEG over the scalp in normal subjects, with respect to frequency. This is partly due to the large quantity of computations required and the difficulty in interpreting them. The results of some studies will be reviewed here to provide a general picture of what is known.

One of the first of such studies considered the correlated activity over posterior regions of the brain, using standard 10-20 electrode placements [30] during an auditory vigilance task [31]. This yielded 6 channels of data that were subjected to detailed spectral analysis, including autospectral and cross-spectral power and coherence estimates. These were presented in the form of a combined time-frequency plot to permit ready assessment of the nature and extent of inter-channel correlations. An example of their results is shown in fig. 2.10. The study was performed on 30 normal subjects and the main results are now summarised.

The highest intensities occurred in two frequency ranges, below 1 Hz and from about 9-11 Hz. Though the amplitudes varied over the course of the experiment, this general pattern was maintained. The coherences tended to be highest when the spectral intensities were high, and were most consistently high in the alpha band, except for channels cross-coupled with a transverse occipital channel. This occurred in spite of the fact that the occipital bipolar pair exhibited as much or more alpha activity compared with other channels, and was evidence for orthogonal transverse and longitudinal components of alpha activity. In

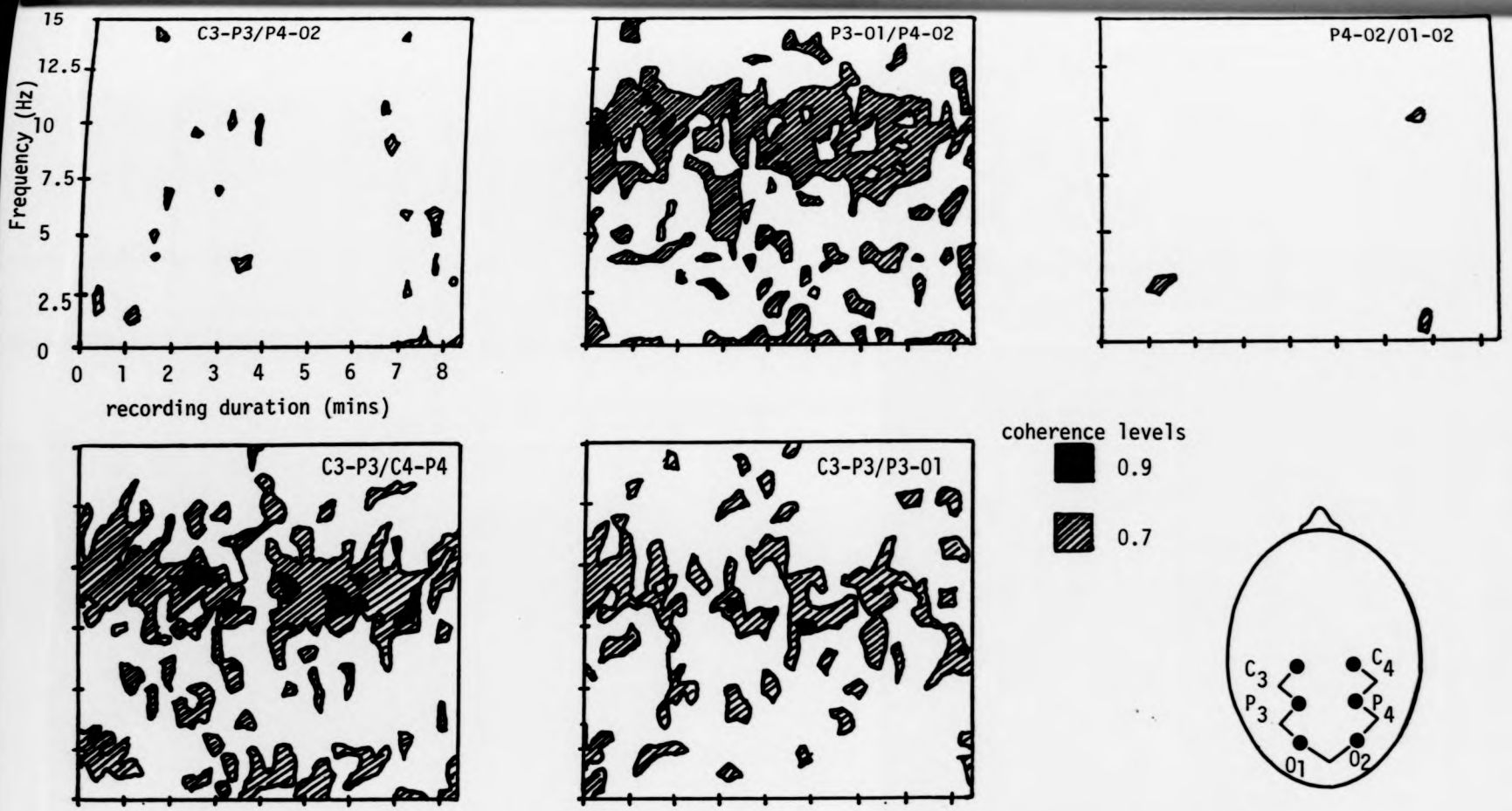



FIGURE 2.10 Selected plots from an early study of EEG coherence over posterior scalp regions. Regions of high coherence occur at frequencies near the alpha band and near d.c. Note that left-right coherences (C3-P3/C4-P4 and P3-01/P4-02) are higher than antero-posterior ones (C3-P3/P3-01). Coherences with 01-02 are low. (From fig.5, Walter et al [31]).

some subjects there was greater coherence between longitudinally derived activity, while in others coherence was greater between transverse derivations. Of the 30 subjects considered, only 3 showed no evidence of orthogonality of transverse and longitudinal alpha activity, showing that this activity may not be considered to be transmitted isotropically over the scalp.

Dumermuth and Fluhler [32] performed a similar study using more electrodes over a larger scalp area, but used only one subject. Their results were similar and confirmed highest coherences in the alpha band, though several paired channels showed high coherences over much of the 0-20 Hz bandwidth analysed. The results varied with location and orientation of electrodes, though neighbouring medial and lateral regions generally showed highest correlations.

### 2.6.4 Stationarity and normality of the EEG.

Any purely random process can be characterised in terms of its statistical properties; indeed these form the only valid descriptors that may be employed. Because EEG signals largely appear to have random signal characteristics, much of the signal analysis of the EEG has been based on this approach. Two key notions regarding the statistical treatment of random signals are stationarity, and to a lesser extent, normality. Analysis is greatly simplified if one or both of these are true of the data.

A random process can be characterised by the ensemble of all possible sample functions that could be realized by that process. In dealing with time series data, the sample functions are an ordered sequence of data samples in time. The statistical relations governing

the data samples can be used to characterise the process. Stationarity simply requires that the average statistical properties computed over the ensemble of all possible realisations must not vary with respect to translations in time.

The condition of strict stationarity requires that all statistical moments computed over the ensemble, that is the mean, variance and higher moments, be invariant with respect to translations in time. If this is true of the mean and variance alone, the process is termed weakly stationary, or stationary in the wide sense. A further distinction can be drawn in that if the ensemble-averaged statistics computed for a stationary process are identical to the time-averaged statistics for each sample function, the process is said to be ergodic. Ergodicity is useful as it allows estimates of ensemble statistics to be computed from a single record. In practice, it is often only possible to obtain a single time record, for which it is appropriate to define the condition of self-stationarity. This is obtained if the time averaged statistics for each of a number of sub-records show no statistical variation, and is the appropriate condition for EEG signals. This will be referred to throughout this thesis simply as stationarity.

Another useful statistical description of a signal is its amplitude distribution function. A particularly useful situation exists when the probability distribution is gaussian or normal, which often occurs with real data. In this case the amplitude distribution can be characterised by the mean and variance alone, as higher moments are directly related to these. This ensures that strict stationarity is obtained when weak stationarity alone occurs. Many useful mathematical results are available for normal processes and for this reason they are often assumed to exist, even when this is not strictly true, as much

useful information can be obtained. The signal can be defined in terms of the mean and autocorrelation alone, or in the frequency domain, by the power spectrum. Confidence limits for both time and spectral estimates are readily computed, although departures from normality are not thought to seriously affect the latter [33]. In addition linear transformations on the data are sufficient to yield optimal processing in the least mean square sense.

Each of these properties has been investigated in the EEG in a number of studies, owing to their importance in the statistical treatment of these data. There is general agreement that the waking EEG can be considered stationary for moderate intervals of 12-25 s, though normality is not so well obtained. In one study of 30 subjects, McEwan and Anderson [34] showed that 50% of 8 s epochs can be considered both wide sense stationary and normal, but stationarity alone was demonstrated for 50% of 64 s epochs. Kawabata [35] confirmed this by reporting that the mean, variance and power in the range 0-17.5 Hz were stationary for 25 s epochs and for 90% of 50 s epochs. This study was based on two subjects, though another study of 104 subjects by Cohen [36] showed that the mean may be considered stationary only for 12 s epochs at the 1% significance level. Each of these studies employed different statistical tests, and the results were reported to be approximately the same for different scalp regions including occipital sites. Nonetheless it is fairly clear from an examination of the EEG record that short term nonstationarities do occur, so care needs to be taken before automatically invoking the assumption of stationarity.

## 2.7 Summary.

In this chapter evoked cortical potentials were introduced and described in terms of their physiological origin, relation to stimulus attributes and their signal characteristics. Visual evoked potentials to patterned stimuli are a particularly informative response, owing to the wide range of stimulus parameters that may be controlled, and much of the discussion was presented in terms of these signals. In addition material was presented to show the range of applications to which EPs have been put, both in neurophysiology research and in clinical medicine.

General characteristics of both cortical EP and unevoked EEG signals were summarised. These show clearly the difficulty in obtaining estimates of the low amplitude EP from the raw signal record, and indicate the need for signal extraction methods. This section included a discussion of the correlation properties of the EEG over the scalp which is of primary relevance to the projected aims of this thesis.

Before proceeding with the theoretical and experimental details concerning this approach, the next chapter reviews some commonly applied signal enhancement methods used with EPs. This includes a discussion of their main advantages and limitations and provides further justification for the investigation of novel signal enhancement methods.

## 2.7 Summary.

In this chapter evoked cortical potentials were introduced and described in terms of their physiological origin, relation to stimulus attributes and their signal characteristics. Visual evoked potentials to patterned stimuli are a particularly informative response, owing to the wide range of stimulus parameters that may be controlled, and much of the discussion was presented in terms of these signals. In addition material was presented to show the range of applications to which EPs have been put, both in neurophysiology research and in clinical medicine.

General characteristics of both cortical EP and unevoked EEG signals were summarised. These show clearly the difficulty in obtaining estimates of the low amplitude EP from the raw signal record, and indicate the need for signal extraction methods. This section included a discussion of the correlation properties of the EEG over the scalp which is of primary relevance to the projected aims of this thesis.

Before proceeding with the theoretical and experimental details concerning this approach, the next chapter reviews some commonly applied signal enhancement methods used with EPs. This includes a discussion of their main advantages and limitations and provides further justification for the investigation of novel signal enhancement methods.

# CHAPTER THREE.

# A REVIEW OF EVOKED POTENTIAL SIGNAL ENHANCEMENT METHODS

## 3.1 Introduction.

In the last chapter the basic nature of evoked potentials in terms of neurophysiology and signal characteristics was briefly introduced. In this chapter I intend to review various methods of estimating EPs from raw records. Firstly signal averaging will be described, with regard to the underlying assumptions, followed by a number of techniques and modifications which may be employed when these assumptions do not hold. A number of alternative methods will also be described which attempt to improve upon averaging and even make the individual responses available. The coverage of this material is broadly similar to the review chapters of Aunon, McGillem and Childers in the Critical Reviews in Bioengineering series [51] [52] which may be consulted for a more detailed description of the various techniques discussed.

## 3.2 Coherent signal averaging.

Signal averaging is the most widely used signal extraction technique, as it is simple to apply and is effective in many applications. As it has been described in detail elsewhere [39] only the main results will be presented here. The continuous-time signal $x_i(t)$ recorded following the i'th stimulus may be considered as the sum of a neuronal response $s_i(t)$, the background EEG activity $e_i(t)$, and an artifact signal $a_i(t)$, each drawn from an ensemble of sample functions representing a random process to include the effects of signal

variability. This can be written:

$$x_i(t) = s_i(t) + e_i(t) + a_i(t) \tag{3.1}$$

where the mean signal component is s(t), and both $e_i(t)$ and $a_i(t)$ are assumed to be zero mean processes.

The ensemble average v(t) is the mean of all responses in the ensemble at each sample time. This may be written as

$$v(t) = 1/N \sum_{i=1}^{N} x_i(t) \tag{3.2}$$

The rationale is that the underlying signal will be enhanced preferentially to the incoherent 'noise' activity such that given a large enough sample size N, reliable noise-free estimates are obtained of the average signal s(t). It is easy to show that in the limit the mean response will be the underlying signal s(t) and hence that averaging is an unbiased estimation procedure, viz:

$$\begin{aligned} E[\, v(t) \,] &= E[\, s_i(t) + a_i(t) + e_i(t) \,] \\ &= E[\, s_i(t) \,] \\ &= s(t) \end{aligned} \tag{3.3}$$

noting that the EEG activity is assumed to have zero mean and assuming that the artifact term is identically zero. Such an assumption may be justified for this analysis, as careful experimental procedure can largely eliminate these potential artifacts.

It is well known that when certain conditions to be described apply, the expected amplitude of the noise component is reduced by a factor of $\sqrt{N}$ in the average. This result may be stated as:

$$\rho_i(t) = \sqrt{N} \cdot \rho_o(t) \tag{3.4}$$

where

$$\rho_i(t) = s(t) / \sigma_x(t) \tag{3.5}$$

is the input amplitude signal to noise ratio and

$$\rho_o(t) = s(t) / \sigma_v(t) \tag{3.6}$$

is the output amplitude signal to noise ratio. $\sigma_x(t)$ and $\sigma_v(t)$ are the standard deviations of $x_i(t)$ and $v(t)$ respectively. The conditions required for these results to hold are as follows. The signals $s_i(t)$ must be

(a) synchronous with the stimulus

(b) homogeneous, though nonstationary in general

(c) of finite duration less than the minimum stimulation period and the noise samples $e_i(t)$ must be

(d) mutually independent for periods equal to or greater than the stimulation period

(e) independent of the signal

(f) stationary

It is thus implicitly assumed that the record $x_i(t)$ is the linear sum of the signal and noise records. No constraint is made upon the EEG to be a normal random process. The requirement that the signal be homogeneous though not necessarily stationary simply means that each sample function may be nonstationary, as it will be for a deterministic transient signal, but nonstationarity with respect to the fixed lag times across the ensemble of sample functions. Although these conditions are assumed to hold for a broad range of signals, they are often satisfied only approximately and sometimes may not even be applicable. A number of workers have therefore proposed modified or alternative procedures for use in such cases. These will now be described in the following sections, treating each condition separately.

### 3.2.1 Signal not synchronous with the stimulus.

#### 3.2.1.1 Latency correction of responses prior to averaging.

As has been demonstrated by Brazier and others [53] [26] the single EP can vary on a stimulus to stimulus basis which obviously degrades the performance of averaging, though these appear to be more serious for intracerebral recordings than for scalp recordings [42]. Variability in the time of occurrence of each EP is more serious since the resolution of the average is reduced and important features can be 'smeared out'. In an attempt to deal with this, Woody [41] developed an iterative procedure to identify the time of occurrence of each EP by crosscorrelation with a suitable template in order to align each corrected response in time prior to normal averaging. This process can then be repeated using the improved average as the new template until some convergence criterion is attained. His results on cat EP data showed a significant improvement in resolution, and many workers have since applied this technique.

Wastell [54] applied a mathematical correction to Woody's original technique and studied its performance on human pattern VEPs. His results suggested that iteration was proceeding beyond the point of meaningful improvement, due to an inadequate convergence criterion. It was recommended that the method be applied with caution by selecting a good initial template and employing a more sensitive criterion. Aunon and Sencaj [55] also studied the performance of Woody's filter using human flash VEPs, and found that although it sometimes yielded a larger response compared with signal averaging, it tended to lock onto and enhance the largest component present and attenuate others. More seriously it was unable to discriminate between signal and noise and

performed badly when applied to noise-only records, probably by locking onto the alpha waves. Steeger and Herrmann [61] also reported that application of Woody's method gave poorer results when applied to auditory EPs in man, and suggested that this was due to erroneous synchronization of EP-like transients in the EEG. and proposed an alternative method of detecting the signal which is optimally adapted to EP and EEG signal characteristics. They also recommended that a practical reliability test be performed which would estimate the number of erroneous synchronizations for a given SNR.

It has also been suggested that apart from variations in the latency of the overall response, significant variability in the latency of individual components of the response may be present. This has been supported by the evidence of a number of workers [59] [53]. It seems clear [26] that genuine latency variations are responsible rather than the statistical effects of the background EEG. may not be simply attributed to the background EEG. McGillem and Aunon [58] proposed a technique they call latency-corrected averaging (LCA) to effectively deal with this difficulty. Their method, briefly, is to first apply a filter to the raw data to remove components characteristic of the EEG but not EP. The single trial data are then processed using a triangular window to identify significant peaks in each, and a histogram constructed of their distribution. The LCA is formed by aligning and averaging segments in the single trials that correspond to these. If desired, a continuous representation of the disjoint LCA may be obtained using a polynomial interpolation.

The authors claim significant improvements over both averaging and Woody's method [55] using both artificial and human VEP data. As previously mentioned, Woody's filter tends to lock onto the largest peak

present, but if individual components are of variable latency then these will be attenuated. LCA does not suffer from this drawback as each significant peak is independently compensated for. The morphology of the average peak is preserved, but often peaks not evident otherwise are disclosed. Also the method works well in the case of noise-only trials in that only a small number of peaks are spuriously detected, unlike Woody's filter.

### 3.2.1.2 Deconvolution of the average.

Another approach, that of temporal deconvolution, was suggested independently by Sjontoft [57] and McGillem et al [58]. The basis of the method is that the (smeared) average can be considered as the sum of signals which have undergone random delays, and thus can be mathematically represented as the convolution of the underlying signal with a suitable smearing function $h_i(t)$:

$$x_i(t) = s(t) * h_i(t) \tag{3.7}$$

where $h_i(t)$ is defined as a dirac pulse with random delays $\tau_i$ having probability distribution (p.d.f.) $p(t)$:

$$h_i(t) = \delta(t - \tau_i) \tag{3.8}$$

Taking the Fourier Transform of $v(t)$, the expectation of $x_i(t)$, yields

$$V(f) = S(f) \cdot P(f) \tag{3.9}$$

where $V(f)$, $S(f)$ and $P(f)$ are the Fourier transforms of $v(t)$, $s(t)$ and $p(t)$ respectively. The underlying signal may be estimated by applying the inverse filter $P^{-1}(f)$ to the smeared average, but this method suffers from instabilities whenever noise is present in the records, which requires constraints to be applied in the solution.

Sjontoft considered two types of convolving function having Gaussian and rectangular p.d.f.s whose mathematical properties enabled him to derive a computationally simple method of extracting S(f) without the use of the Fourier Transform. To deal with the noise problem, he suggested fitting a polynomial to the average and using it instead of the average itself to estimate the higher derivatives, though rigorous justification for this was not presented. McGillem and Aunon applied the inverse filter directly using the Fourier Transform, and simply truncated computation of its coefficients beyond significant signal frequencies, thereby avoiding the instability at higher frequencies. While, again, this seems intuitively reasonable, no formal justification for this was provided.

Twyman and Lastimosa [60] applied the method of constrained least squares deconvolution to BSEPs. This method seeks an optimum solution based on some criterion, here taken to be a measure of the smoothness of the solution, to suppress high frequency components. It requires a priori knowledge of the convolving function and also the noise mean and variance. As these are not known in this application, the authors modelled the EEG as a stationary zero-mean process and estimated its variance statistically. Then from a consideration of the likely mechanisms of spatial and temporal distortions, they used a Gaussian convolving function to approximate these for simplicity. Enhanced resolution of previously indistinct components was reported and the authors claimed that this enabled more confident clinical assessments to be made.

### 3.2.1.3 Phase averaging.

Two other methods of correcting for variations in response latency have been proposed. The first is the phase-locked time average developed by Auerbach et al [62]. This method computes the Fourier coefficients of each response and instead of summing these algebraically (which yields the Fourier transform of the normal signal average), averages them vectorially, using an augmented phase angle defined by the authors. This yields an average which may be inverse transformed to obtain the phase-locked time average. The method was developed to reproduce the average signal of a set of responses of varying amplitude suffering random delays to onset, buried in band-limited noise. Within certain operational constraints, the method has been shown to work well using an artificially generated test pulse buried in pink noise. Though this would appear to be a useful and rewarding technique, no further work has been reported regarding its application to evoked potentials or to determine its performance analytically or by other simulations.

Rodriguez et al [63] proposed a similar method known as unwrapped phase averaging (UPA). This also obtains an estimate of a deterministic signal experiencing random delays and buried in noise of similar spectral characteristics by operating on the Fourier coefficients of the raw data. Empirical evidence was presented showing that the average signal estimate could be derived from averages of the amplitude and phase spectra. However their method requires that the phase spectrum of the signal be a reasonably smooth function and that the signal amplitude spectrum be related to the noise spectrum by a known scale factor. The method of UPA gave good results using artificial signals in noise with SNR in the range 0 - 10 dB. Application to VEP data yielded a more distinct signal than conventional

averaging, but appeared to distort the EP estimates. This could be due to incorrect estimation of the SNR, or incorrect assumptions regarding signal invariance or noise statistics. No comparisons were presented with adaptive crosscorrelation averaging or deconvolution, so it is difficult to form any estimate of the applicability of the method to EPs. However both of these methods suggest that consideration of the phase response may well be important in the estimation of EP signals.

In summary, then, there are a number of procedures that may be performed when latency variations occur in the ensemble of records. If these occur in the whole response, then adaptive alignment by crosscorrelation, deconvolution or phase averaging methods can be used. The former requires a good template relatively free from noise and is best when SNR is fairly good, otherwise it tends to lock onto noise components and may give unsatisfactory results. Deconvolution by a number of methods is appropriate when some knowledge of the delay distribution exists and has the added advantage of being able to compensate for temporal smearing if the transmission properties of the intervening tissues are known or can be modelled by a linear process. The main disadvantages of deconvolution are that it is prone to instability as high frequency noise in the raw records tends to be amplified, and the nature of the delay distribution is usually not precisely known and has to be estimated. Although enhanced resolution was claimed by each group, no thorough case studies or simulations were presented to demonstrate the reliability convincingly. Indeed the inclusion of a deconvolution result using a deliberately large dispersion value by Sjontoft demonstrated the possibility of spurious enhancement.

The phase averaging methods have not been sufficiently well developed to permit a fair assessment. Both are appropriate for fairly poor SNR data, but phase-locked time averaging appears to be more suited to EPs as UPA sets constraints which are less suited, such as signal invariance and known SNR which is frequency independent. The former only requires a priori knowledge of the approximate duration of the signal and extent of latency jitter, and yields the mean response if amplitude variability is present. Each of these methods can only compensate for delay experienced by the whole waveform, but LCA can be applied when individual shifts in latency occur in the different components. Although this has been given the fullest study of all of these methods, with the possible exception of Woody's filter, it has not featured in the reports of others' work and so independent confirmation of its potential benefits is unavailable.

### 3.2.2 Signal Inhomogeneity.

While averaging of a non-stationary signal is a simple extension of averaging of a stationary signal, the requirement of homogeneity is more stringent. Mere variability of the signal can be accommodated, but a fundamental requirement is that there be a well-defined signal to estimate, and by well-defined we can mean having a stationary mean value at each time point. It is impossible to form a single estimate of the signal otherwise as some distribution of waveforms would be implied. A good example where inhomogeneity occurs is in the case of habituation. It is clear that in such a case averaging may yield a mean response, but one which is inaccurate and misleading, though it may serve as an approximation. If the nature of the inhomogeneity is known, e.g. an exponentially decreasing amplitude of otherwise invariant responses, then this can be dealt with appropriately in the averaging process.

Bendat [39] allowed for this possibility in his analysis by permitting the signals in an ensemble to have time-wise correlations and computed the key results of averaging for this case. He also suggested that by comparing the averages of two ensembles with respect to the expected confidence limits (assuming homogeneity) it would be possible to identify non-homogeneous ensembles which could than be analysed on the basis of some predictive model of homogeneity.

Ruchkin [64] presented a technique to form separate averages corresponding to different categories or modes. Firstly an ensemble is examined to see if it is likely that there are several modes, and if so what latency ranges are involved. The ensemble may then be examined at these particular latencies to determine the modes present and at what point the boundaries are most likely to have occurred. Note that this technique requires the responses to be analysed in strictly ordered sequence, and thus is only useful for modes occurring in discrete sub-sequences. John [65] developed the Sort method as an extension of the previous method. The main difference is that it is not dependent on the contiguity of the various EP types. Briefly the method works by examining the amplitude histograms of the single trials for each sample point, identifying those ranges of amplitude that differentiate non-homogeneous components, then classifying and averaging each response depending on the amplitude sub-range to which it belongs at selected time-points. This method has been used successfully to identify EPs having different characteristics that would otherwise have been lumped together to provide a much less sensitive estimate. However it requires careful application, and the author recommends independent validation of the results.

Brazier [53] illustrated the distortions that can occur for each of latency and amplitude variability and showed that variations in amplitude are less serious, as the overall morphology of the signal is not seriously affected, which is the object of most EP studies. Cummins [66] has shown that signal averaging is not the best linear estimator when significant amplitude variability occurs in the signal, and shows that a better estimate is obtained if each response is weighted by the value of its crosscorrelation with the deterministic response. Several workers have proposed methods of selective or weighted averaging that are based on this. Pfurtscheller and Cooper [43] analysed auditory EPs using intracerebral electrodes, and by crosscorrelating the single trial responses with a suitable template, they showed that significant amplitude and latency variations are present in these. By selectively averaging only those responses that matched given amplitude criteria, an average was obtained that was 30% larger than the ordinary average, suggesting that information is lost by the latter method. More recently Gasser et al [42] used similar methods to obtain averages of surface-recorded EPs which are less affected by variability in responses. They employed a weighted averaging scheme, where each response is assigned a weight according to whether it can be classified as typical or atypical, on the basis of crosscorrelation with the probable response template. A study of 42 children using flash VEPs demonstrated convincingly that weighted averaging of inhomogeneous EP data yields more reliable estimates of these signals.

So to summarise, an ensemble may be first assumed homogeneous and then tested to see if this is valid. Whereupon various techniques may be applied to determine the number and nature of the different modes occurring and by identification of these in the raw data, form averages for each homogeneous set. Only Bendat considers the case where there is

a continuously varying inhomogeneity that may be modelled and analysed, as the other methods assume that a finite set of homogeneous modes exists.

### 3.2.3 Signal duration exceeding stimulation period.

Though EPs are generally considered as having finite duration, the desirability for short analysis epochs can cause a problem when long latency components persist beyond this, as distortions are introduced into the average. This is the case with the alpha-like "after-discharge" that can follow photic stimulation. If this is still occurring at the time of the next stimulus, the responses will be clearly confounded. This is a particular hazard when periodic stimulation is employed. Ruchkin [67] has analysed this situation and recommended the use of aperiodically delivered stimuli of some appropriate random characteristic. By estimating the duration of the signal from previous experiments it is usually possible to ascertain what stimulation rate is permissible to avoid overlapped responses. Use of this procedure with aperiodic stimulation is usually sufficient to avoid this problem.

### 3.2.4 Correlation between EEG records.

While short-term correlations do exist in the background EEG which itself may be approximated as narrow band noise, the assumption of uncorrelated EEG activity across sample records is universally made for simplification. It is not altogether unreasonable, particularly for longer records, though it is to be noted that short records are best if signal repeatability and noise stationarity are to be ensured. Records obtained using fast stimulation rates are more prone to suffer from

having correlated noises, and lower (aperiodic) stimulation should be employed to minimise this problem.

The effect of correlated EEG activity in different records is not to introduce bias into the average, but rather to render the noise reduction process less effective such that the $\sqrt{N}$ improvement is no longer obtained. Cummins [66] has indicated that signal averaging remains the best linear estimator even when the noise is correlated and time dependent, provided that variability of the neuronal responses is small. In such cases the variance of the estimate approaches a minimum value as averaging proceeds, rather than decreasing as 1/N. This suggests that there is a limit beyond which useful gains in noise-reduction may be obtained.

### 3.2.5 Correlation between the EP and the EEG.

This question is at the very heart of EP research, and is concerned with the appropriateness of the commonly used signal in additive noise model. There is evidence that the brain response (at least to auditory stimulation) is dependent on psychological factors such as arousal and anxiety. Changes in shape and amplitude have been obtained depending on task relevance and subject attentiveness [68]. Pattern VEP components are less sensitive to these factors, and the assumption of a brain response primarily affected by stimulus parameters is more nearly justified. Nonetheless changes in VEP form can be related to the nature of the background EEG, for example the flash EP varies in form depending upon the phase of the alpha rhythm when stimulation is applied [68]. Application of the stimulus can also affect the background EEG, exemplified by the attenuation of alpha activity following stimulus presentation [70]. It is clear then that

the assumption of independence of EP and EEG activity must be treated with caution, though attempts to model this are uncommon.

It is not clear to what extent the EP is correlated with the EEG owing to the difficulty in differentiating between these signals in the raw responses. Salomon and Barfod [71] hypothesized that the EEG record following the stimulus is the sum of an impulse (filtered to represent the EP transient signal), and white noise, representing underlying random EEG processes, each processed by a common time-varying filter to represent the changing nature of brain processes. By applying a whitening filter to the data which effectively forms the inverse of the hypothesized common filter they were able to demonstrate that the inverse filtered EPs showed less variability, which provides some experimental justification for considering EP and EEG processes jointly. Basar and colleagues [73] have supported a model of EP generation processes based on coupled oscillators that are brought into synchrony by the stimulus. They report experimental evidence in support of this model based on intracerebral studies in cat using photic stimulation, though it is not clear how well these results may be extrapolated to surface-recorded activity in man using more subtle stimuli. Sayers and Beagley [44] have shown that better detection of near-threshold auditory EPs is possible by considering the relative phases of EEG components before and after stimulation rather than the amplitude response, on the grounds that the stimulus acts to alter the phase relationships of the spontaneous EEG frequency components. This has been challenged by Jervis and colleagues [73] who argue that a purely additive signal in noise model will also result in phase re-ordering of frequency components of the post-stimulus EEG. Their analysis of EP data in fact suggested that additive processes were more likely to be the main

mechanism than phase re-ordering.

In addition to this study, evidence in support of an additive relationship between the EP and EEG comes primarily from the relative success with which well-defined EPs are reliably obtained using averaging procedures. EP components can often be reliably related to stimulus parameters and to neurophysiology. This is particularly true of the brainstem EPs though cortical EPs are acknowledged to be more susceptible to variability and influence of higher order mental processes. In spite of the evidence for superior detection of low-level auditory EPs reported by Sayers and Beagley [44], these authors admit that the hypothesis of additive EP generation processes is nonetheless a useful one in practice, and one that becomes more reasonable for stimulus intensities beyond threshold. Until this situation becomes clearer following further research, it is proposed to continue using the simple signal plus noise model implicit in averaging, while being aware of its possible limitations.

### 3.2.6 Nonstationarity of the EEG.

Stationarity of the EEG has already been discussed in some detail in chapter 2 where it was reported that short data segments in the range 12-25 s may be considered stationary. Stationarity is usually assumed in the analysis of averaging for simplification. If this condition is not strictly met the improvement predicted by expressions such as (3.6) may no longer be obtained. Little mention of this difficulty occurs in the literature concerning signal processing of EPs, and the matter behoves further investigation, although in practice many are content with the results gained from averaging, so it may not be considered too serious. The effects of suboptimal averaging are

unlikely to be important for routine work when a certain amount of residual noise is tolerable in the average, but may affect long runs when small responses are being recorded such as to threshold stimuli.

## 3.3 Optimal Filtering.

### 3.3.1 A Posteriori Wiener Filtering

In attempting to obtain better estimates of the signal than averaging alone can provide, attention has recently focussed on the possible application of optimal linear filters, mainly based on Wiener's theory [37]. Although not an alternative to averaging, they may be considered as a useful adjunct to it. The motivation for such a technique is that a unique filter can be formulated on the basis of known signal and noise characteristics to optimally improve the SNR according to some criterion. Thus the filter is not a fixed filter designed on the basis of typical data, rather it is matched to the characteristics of a given data set to provide the best obtainable filtering operation in some sense. The potential gains lie in a reduced number of averages required to attain an arbitrary level of accuracy in the average, or alternatively, a more accurate estimate from a given data set. This work will be briefly reviewed, particularly with regard to the main conclusions that have been reached, but for a fuller review the reader is referred to [74].

Wiener's theory of optimal filtering [37] states that for a signal in noise such that both are uncorrelated, stochastic stationary processes with zero mean, application of a filter H(f) defined as

$$H(f) = S(f) / ( S(f) + N(f) ) \qquad (3.10)$$

results in an output signal which is optimally filtered in the least mean square sense. S(f) and N(f) denote the signal and noise power spectra respectively. The signal estimate $\hat{s}(t)$ is :-

$$\hat{s}(t) = F^{-1}[\, X(f) \,.\, H(f) \,] \qquad (3.11)$$

where $F^{-1}[.]$ denotes the inverse Fourier transform operation. The resulting LMS error e(t) is

$$e(t) = E[\, (\hat{s}(t) - s(t))^2 \,] \qquad (3.12)$$

The effect of the filter is thus to weight the spectral components proportionally to the SNR at each frequency, (i.e. noise-only components are completely attenuated and signal-only components completely passed, assuming no spectral overlap).

Walter [75] first proposed the use of a form of Wiener filter to EPs where the signal and noise spectra were estimated from the raw data in an ensemble. His formulation proved to be incorrect for this application and was modified by Doyle [76]. I shall refer to this filter as an a posteriori Wiener filter (APWF). Since then a number of research workers have examined the use of APWF to both simulated and real data, with conflicting reports on the gains achieved, some describing it as extremely valuable and others dismissing it. De Weerd [74] has discussed the application of APWF to EP estimation and noted the following points:

a) Wiener's filter as originally proposed requires that the signal and noise be stationary, stochastic processes and yet in the estimation of the signal and noise power spectra by APWF, recourse is made to a deterministic model. This is inconsistent, though there are indications that this is not a very serious difficulty as long as the signal has a spectrum which is relatively time invariant. McGillem and Aunon [77] using essentially the time-domain equivalent of the APWF

concluded that the resulting filter would be virtually the same for deterministic signals as for random stationary signals.

b) Because of the considerable spectral overlap between EPs and EEGs the Wiener coefficients generally lie within the range 0 - 1. This has the double effect of passing some noise and yet attenuating the signal at those frequencies, resulting in a distorted estimate. In marked contrast to averaging which introduces little or no bias, this can considerably degrade the morphology of the EP. It is even possible to obtain filtered estimates from noisy data which are better than the average in the least mean square sense, but which are so heavily smoothed that large distortions are present. This is a particular hazard of all optimal filters and necessitates the exercise of caution and control tests.

c) Related to this is the serious difficulty that as the signal and noise properties are unknown a priori, these have to be estimated from the data and thus only an estimate of the Wiener filter can be obtained, which obviously involves a departure from optimality. This question was extensively investigated by Strackee and Cerri [78] and De Weerd and Martens [45] who concluded that the estimated transfer function suffers from a large variance and negative bias at low SNR. It is often found that negative values are obtained for the filter coefficients due to chance correlations between signal and noise which is clearly a departure from the original assumptions. Commonly these negative values are simply clipped to zero, though this is a somewhat arbitrary procedure. More rigourously, Woestenberg et al [80] performed a statistical test on each filter coefficient to retain only those which could be confidently attributed to the signal and reject all others by clipping them. This does ensure that only genuine signal components are

passed by the filter and thus represents an improvement, however the rejection of all coefficients below 0.8 (as happens for a confidence level of 95%) does mean that signal components are removed. De Weerd and Martens suggested that proper smoothing of the spectral estimates (with clipping of any residual negative coefficients) can significantly overcome this difficulty with much improved results. Nonetheless the resulting filter can still only be regarded as an approximation to the optimal filter.

d) From studies of the time-varying spectral composition of EPs of various modalities, De Weerd and Kap [24] have shown that an empirical inverse relationship approximately exists between frequency and duration of EP activity. This means that a short burst of high frequency activity will have its power smeared out over the whole analysis epoch, resulting in a significant under-estimation. This has the effect of attenuating these components while also passing noise at the same frequency which might have been removed. This prompted research into a time-varying filter, which will be described later.

In spite of these drawbacks, APWF in one of its modified forms can nonetheless yield improved estimates over averaging. It seems clear that the cases when this improvement did not obtain are particular examples of the difficulties described. Thus Carlton and Katz [79] failed to obtain improved estimates because they used EPs with a highly transient characteristic, and Wastell [81] demonstrated the problem of spurious signal components in the estimate that were actually due to noise. Better results are obtained when the signal is more periodic such as the EP employed by Hartwell and Erwin [82]. However the somewhat paradoxical situation is that as the SNR decreases, which is when an alternative to averaging is most desired, the performance of the

APWF becomes poorer as the filter coefficient estimates become increasingly less accurate. In particular it is less effective, or even quite inappropriate, when sharply transient signals are being used, when a fixed filter is no longer adequate.

### 3.3.2 Time-varying filtering.

Stemming from this analysis of APWF de Weerd [83] developed a more generalised method, known as time-varying filtering (TVF). It differs from APWF in that it may be used to estimate transient signals as effectively as periodic signals. The process consists of estimating the power of the data record in the combined time-frequency domain and then applying a filter whose transfer function varies in time over the duration of the sample record. This may be represented mathematically as follows:

$$G(t,f) = \frac{S(t,f)}{S(t,f) + 1/N\ N(t,f)} \qquad (3.13)$$

where G(t,f) is the time-varying filter coefficient, and S(t,f) N(t,f) are the signal and noise time-varying power spectra respectively.

Because time and frequency are inversely related quantities there is no physical measure that can simultaneously localize the power of a signal within an arbitrarily small region in time and simultaneously within an arbitrarily narrow band in frequency. The solution adopted by de Weerd and Kap was to employ a bank of filters whose bandwidth is proportional to their centre frequency so that accurate time localization of short duration high frequency bursts would be obtained, while longer duration low frequency activity would be accurately resolved in frequency. This scheme is compatible with the general characteristics of EPs. Instead of a continuum in the

time-frequency domain, a quantized scheme is employed with regions of constant time-frequency product as basic units. The power of the signal and noise can then be estimated in each of these regions over the entire observation interval. The ensemble average record in each frequency band is modulated by the time-varying filter coefficients and the final estimate formed from the sum of these. This may be described mathematically:

$$\hat{s}(t) = \sum_{q=1}^{Q} g_q(t)\, x_q(t) \tag{3.14}$$

where $x_q(t)$ is the ensemble average record in the q'th band, $g_q(t)$ is the time-varying filter coefficient in the q'th frequency band, and $\hat{s}(t)$ is the TVF estimate.

Simulations have clearly demonstrated that TVF is superior both to APWF and to averaging in most conditions, and yields significantly improved estimates in terms of reduced noise activity. The method has also been applied to a large number of EPs of all modalities which have undergone clinical evaluation, with the conclusion that when used correctly it does offer advantages over APWF and averaging in nearly all cases. Blind use of the technique is cautioned against, and the authors recommend that it be employed in conjunction with averaging, and only for data that contain a discernible EP. Its use is also recommended for obtaining improved estimates rather than a reduction in the number of sweeps as a small data set is more liable to give amplitude distortion.

time-frequency domain, a quantized scheme is employed with regions of constant time-frequency product as basic units. The power of the signal and noise can then be estimated in each of these regions over the entire observation interval. The ensemble average record in each frequency band is modulated by the time-varying filter coefficients and the final estimate formed from the sum of these. This may be described mathematically:

$$\hat{s}(t) = \sum_{q=1}^{Q} g_q(t)\, x_q(t) \tag{3.14}$$

where $x_q(t)$ is the ensemble average record in the q'th band, $g_q(t)$ is the time-varying filter coefficient in the q'th frequency band, and $\hat{s}(t)$ is the TVF estimate.

Simulations have clearly demonstrated that TVF is superior both to APWF and to averaging in most conditions, and yields significantly improved estimates in terms of reduced noise activity. The method has also been applied to a large number of EPs of all modalities which have undergone clinical evaluation, with the conclusion that when used correctly it does offer advantages over APWF and averaging in nearly all cases. Blind use of the technique is cautioned against, and the authors recommend that it be employed in conjunction with averaging, and only for data that contain a discernible EP. Its use is also recommended for obtaining improved estimates rather than a reduction in the number of sweeps as a small data set is more liable to give amplitude distortion.

### 3.3.3 Adaptive filtering.

Friedman and Carmon [110] applied optimal filtering to visual, auditory and somatosensory EPs using an adaptive technique to determine the filter coefficients for each data set. They were initially concerned with making accurate measurements on a small number of single trials to painful stimulation with a view to determining the relationship between the levels of pain and EP amplitudes, which for practical and ethical reasons requires a single trial approach.

The Adaptive Filter (AF) will be described in detail in the following chapter, but briefly it is a self-adapting linear filter designed to achieve some specified filter performance, in this case to minimise the mean square error between the filtered responses and a given EP template. These authors first applied Woody's method to obtain a good quality average. Using this as the reference template they applied the adaptive filter to the single trials to remove activity unrepresentative of the template.

The results seem impressive, suggesting a significant improvement in SNR. The average of the AF outputs corresponded well to the conventional averages, indicating that little or no systematic bias was introduced, but individual filtered responses sometimes showed significant deviations. The production of the template is critical to the success of this method, and in this respect the assumption of signal invariance is unfortunate. There are a number of other difficulties with this approach, but these will be discussed in section 4.5 after the theory of adaptive filters has been presented.

## 3.4 Other Approaches.

Averaging, optimal filtering, and their variations are all unidimensional methods which seek to recover a signal added to noise on the basis of known or estimated characteristics. However they do not attempt to take into account data from other related signal sources. This promises to be useful in the context of EP processing, as it is relatively easy to obtain multi-channel recordings from a number of electrodes distributed over the scalp. The main difficulty is that as each electrode is located at a different site the responses vary, reflecting differences in underlying cerebral activity, tissue structure, transmission characteristics etc. Thus it is not a simple case of multi-source derivation of one signal, but rather a multi-source derivation of a large number of different, but correlated signals. The simplest way of obtaining noise reduction by the use of multichannel derivations is to take combinations of the activity in each channel. The use of bipolar derivations can be an effective way of attenuating common activity in a given signal channel. When the spatial distribution of EPs are known, they can sometimes be enhanced by this method, as shown by Jeffreys and discussed in chapter 2. Incautious use of bipolar derivations can however lead to distorted estimates of the EP signals as will become apparent in chapter 7.

Hjorth [84] has described a novel way of determining the local activity near an electrode, known as 'source-derivation'. Based on a theory by Laplace, the mean curvature of the potential field distribution is zero in an area if it contains no independent sources. Thus the source derivation replaces standard derivations, which yield the activity of an electrode with respect to some reference potential, with essentially the mean potential gradient directed at each electrode

from the neighbouring ones. The result is that the contribution of the potential field under the active electrode is accentuated relative to the surrounding electrical activity. Clinical evaluations have shown that this method gives similar results to standard EEG derivations, though it often localizes the source of activity better, yielding a SNR improvement between 2 and 4. Studies of the correlations between activity of neighbouring electrodes give much smaller values than when standard derivations are used, supporting the idea that neighbouring activity is being effectively attenuated. The technique affords a gain in topographical selectivity and can be considered as a spatial filter. Essentially a weighted sum of the neighbouring activity is subtracted from each electrode, which is a linear process. Though the technique is attracting growing interest, it is not yet widely adopted, though it has recently been applied to EP studies giving encouraging results in attempts to map the scalp activity for different retinal areas stimulated [85].

There has been at least one attempt described to form a spatial average over a number of electrode locations rather than a temporal average using the sequence of trials [86]. The rationale for this is that some SNR improvement may occur as random noise in each channel is attenuated relative to the signal which may be sufficiently similar in all channels to constructively combine. In this work the motive was not so much to form an accurate estimate of the underlying signal, but rather to minimize the uncertainty in detecting its presence. Results demonstrated that this can be obtained, but this method has to be applied with care as it is quite possible for an EP signal to reverse polarity across the scalp which would obviously weaken the resulting output. It is therefore unsuitable for general use and has not been seriously espoused by others. It may be useful in certain known

situations or if it is applied in a more sophisticated manner. Other than these no attempt has been published which seeks to use multi-channel EEG data in order to obtain improved EP estimates. This clearly indicates the necessity for an extended study into this possibility.

## 3.5 Summary.

This chapter has been concerned with a description of signal averaging and its limitations when applied to the study of evoked potentials. It has been noted that suboptimal and even inaccurate results can be obtained through use of averaging procedures in certain circumstances, primarily if the signal exhibits appreciable variability in form or latency in the ensemble of responses. While procedures have been developed to overcome some of these difficulties, involving alignment of variable latency responses prior to averaging and averaging of responses that belong to homogeneous subensembles, these techniques are still unable to overcome the limitations of nonstationary EEG records and require many responses to be averaged which can be impractical or impossible in situations that do not permit this. Attempts to improve the efficacy of averaging by the use of 'optimal' filtering have achieved some success in improving the quality of EPs, but still involve averaging of moderate numbers of responses and so do not significantly alter this situation.

There is therefore considerable motivation to develop other means of improving upon averaging performance. Little work has been done on multichannel processing of data to obtain improved signal estimates by appropriate manipulation of these data, though some studies have indicated that this possibility is worth pursuing. This is the

approach that is investigated in this thesis. As discussed in the introductory chapter, adaptive filters were considered to be a promising means of gaining signal improvements by using reference activity that is correlated with the primary signal channel. The next two chapters are therefore devoted to a discussion of the theory and methods of adaptive filters, and will be followed by an account of the experiments performed to investigate this approach.

# CHAPTER FOUR

# THEORY AND REVIEW OF ADAPTIVE FILTERS.

## 4.1 Basic Adaptive Filter Theory

### 4.1.1 Introduction

In this chapter the principles and some of the main theoretical results concerning adaptive filters will be reviewed. In particular, the Widrow-Hoff Least Mean Square (LMS) adaptive algorithm will be used as a basis for examining adaptive filter operation, though extensions and developments of this algorithm and more recent approaches will also be considered. The use of certain adaptive structures in signal processing will then be considered, followed by a brief review of previous applications to certain biological signals.

Adaptive filters are not a very recent development, having been first proposed at least 25 years ago. They have been the subject of much research since then and have been useful in a number of applications in control engineering, communications and signal processing. The development of some of this research is described in [87]. Traditionally, filtering has been performed on continuous-time analogue signals by electronic circuits employing frequency-dependent elements to achieve the desired function. Pioneering work on optimal filtering strategies for stationary stochastic signals was done by Norbert Wiener in the 1940's and later extended by Kalman and Bucy to non-stationary signals [88]. Both of these approaches require a priori knowledge of the signal statistics for optimal performance to be realised, though this may be approximately achieved if estimates only

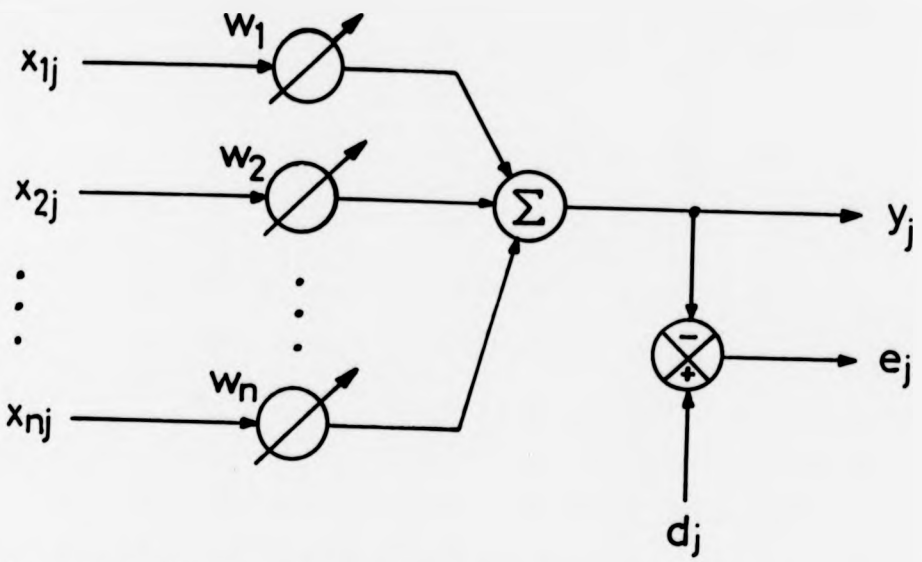
are available.

Apart from Kalman-Bucy designs, traditional filters are usually time-invariant and hence may not be suitable for signals whose properties vary in time. When the signal statistics are not known or cannot be accurately estimated, an adaptive solution may offer superior performance over Wiener filters and can approach the performance of Kalman filters with much less computation. As this is the situation that is encountered with EEG signals an investigation of this approach was pursued.

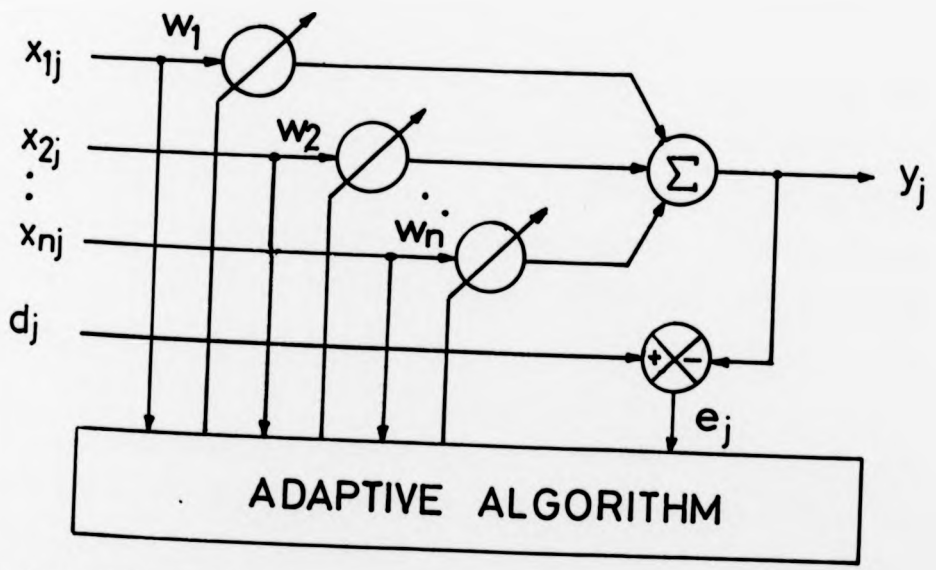
An adaptive filter may be defined to be one whose characteristics are variable and adjusted during operation by a suitable control algorithm according to some criterion of optimality. Each of these three elements, the filter structure, adaptive algorithm and performance criterion will now be discussed, though the discussion will be limited to finite-precision sampled-time systems. The reason for this is that although similar methods and results can be developed for continuous-time or analogue signals, most of the work has been done in the context of digital signal processing. The advantages of digital filters are well known, and relate to the ease with which high performance filters of predictable characteristics may be implemented on general-purpose digital computers or special-purpose hardware.

### 4.1.2 Adaptive Linear Combiner.

A simple structure that lies at the heart of many adaptive filters is the adaptive linear combiner, shown in fig. 4.1. This samples N inputs at the chosen sampling rate, multiplies each input $x_{ij}$ by a corresponding weighting coefficient $w_{ij}$, and combines the weighted




a) An adaptive linear combiner. A series of n input samples $x_{ij}$ - $x_{nj}$ are weighted by the variable coefficients $w_1$ - $w_n$ and summed to form the output $y_j$. This is compared with a desired input $d_j$ to yield an error signal $e_j$.




b) An adaptive filter based on the adaptive linear combiner. The control algorithm samples the input and output signals and adjusts the weights to obtain the desired filter function.

FIGURE 4.1

inputs to form a single output $y_j$ where j is the time index. This may be stated as follows,

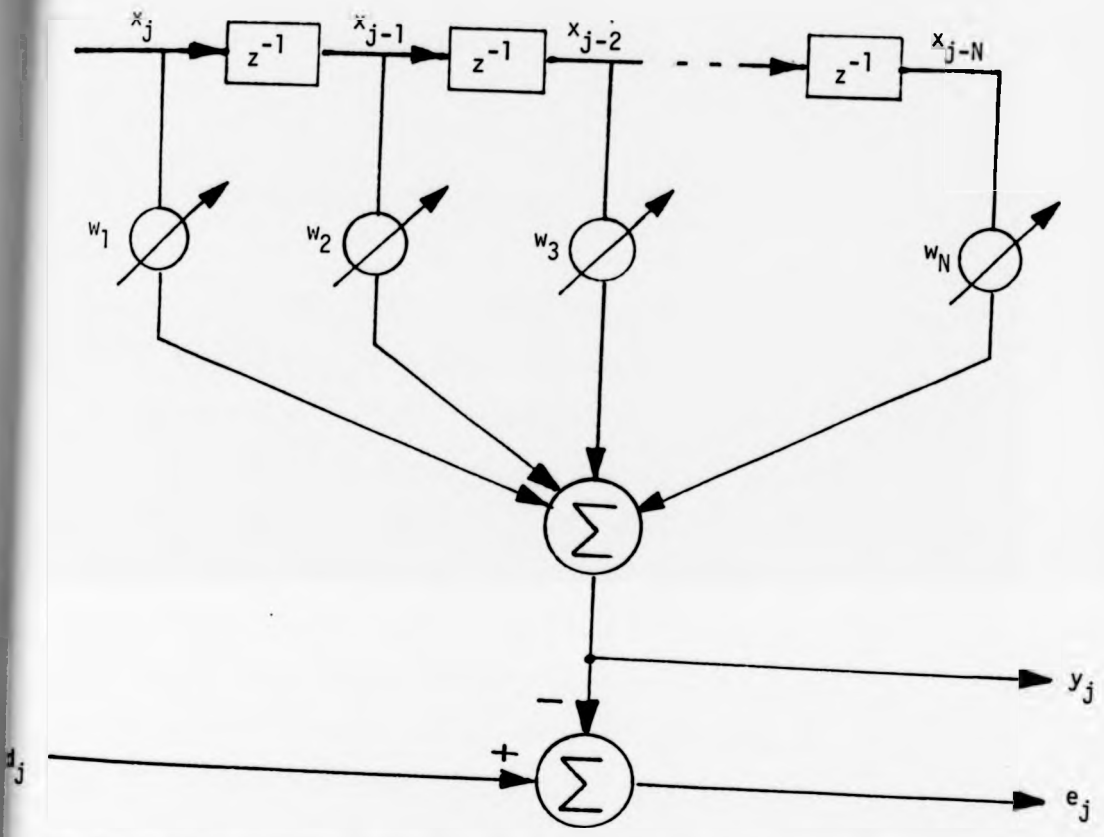
$$y_j = \sum_{i=1}^{N} x_{ij} \cdot w_{ij} \qquad (4.1)$$

The reason for applying the subscript j to the weights $\{w_{ij}\}$ is that they are permitted to vary as filtering progresses. Fig. 4.1 also shows another input, the desired signal $d_j$, which is used to train the adaptive filter. While this signal will not be available in general, (otherwise there would be no need for a filter), it is often possible to use a related signal to adequately train the adaptive filter. An error signal $e_j$ is obtained from the difference between the filter output and the desired signal at each sample instant and used to control the adaptive process:

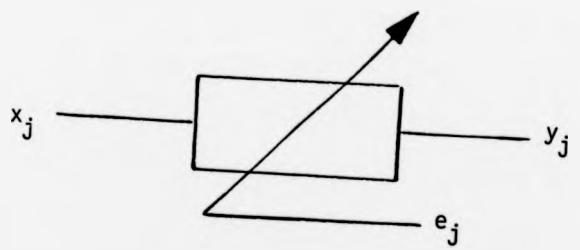
$$e_j = d_j - y_j \qquad (4.2)$$

There are a number of ways of obtaining the inputs $\{x_j\}$. These may be derived from a spatial array of sensors, for example, and combined to yield a composite signal as in adaptive array processing. In this case the weights form a spatial filter that allows the adaptive array to be optimally steered to the main signal direction, a process known as adaptive beamforming [89].

A particularly useful structure is obtained if the inputs are taken from a tapped delay line. In this case the linear combiner forms a transversal or finite impulse response (FIR) filter, and the whole structure is known as an adaptive transversal filter (ATF). This is shown schematically in fig. 4.2. It is also possible to derive adaptive structures with infinite impulse response characteristics, but a recent

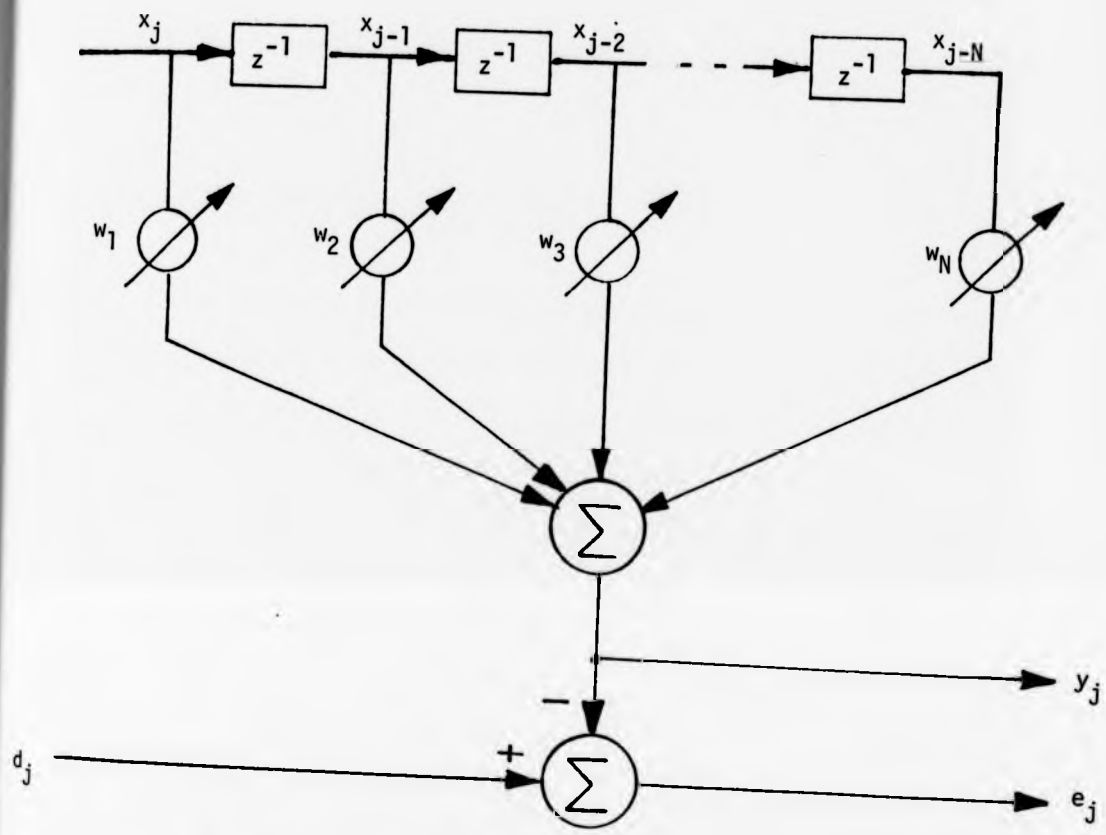



a) Diagram of an adaptive transversal filter which uses a set of successively delayed input samples as the inputs to the adaptive linear combiner.

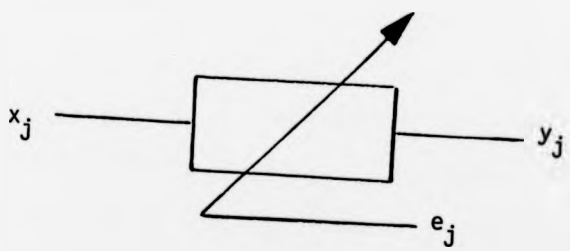



b) Symbolic representation of adaptive transversal filter.

FIGURE 4.2




a) Diagram of an adaptive transversal filter which uses a set of successively delayed input samples as the inputs to the adaptive linear combiner.




b) Symbolic representation of adaptive transversal filter.

FIGURE 4.2

review reports that these have not yet attained the level of maturity attained by the FIR algorithms, and considerable effort is required before robust algorithms with well-understood convergence properties are available [112].

### 4.1.3 Adaptive Transversal Filters.

Transversal structures are simpler to conceptualise and analyse. Two structures have been used to implement FIR filters, the direct or transversal form, and the lattice form. Each has its own particular advantages and disadvantages which will now be briefly summarised. Transversal structures are based on the tapped delay line and linear combiner, as previously shown. This structure lends itself to fairly efficient implementation, with reductions in complexity possible, e.g. by time-sharing a single fast multiplier and accumulator. The correspondence of the tap weights with the model parameters is convenient for system identification and spectral estimation, though these are not explicitly required in this application. This is not the case with the lattice filter structure shown in fig. 4.3 which uses reflection coefficients that are not simply related to the model parameters, (though they can be extracted by a suitable transformation).

Lattice filters have the advantage that the solution for a model of order N automatically produces the residual error signals for all models of lower order. This can be important for certain applications. One reason why lattice filters have come into prominence is their superior numerical properties. They are less sensitive to rounding errors and can tolerate shorter word lengths for the coefficients, which are desirable features in practical implementations [105], but they generally require more multipliers than the direct form. The lattice

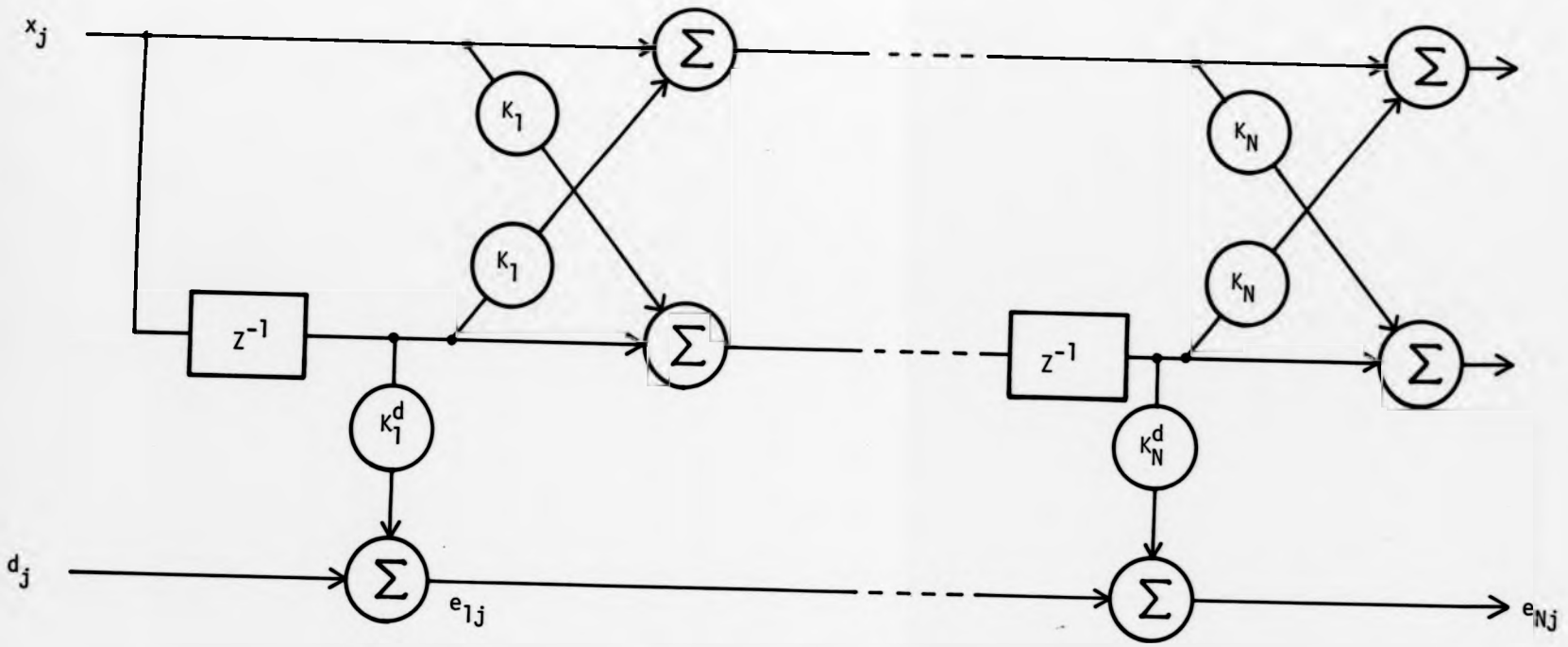



FIGURE 4.3 Adaptive filter used for joint process estimation based on a lattice structure. The coefficients $K_i$ and $K_i^d$ are adjusted by the adaptive algorithm to minimise the MSE at each lattice stage, and hence yield the solution for all filter orders simultaneously. Further details regarding lattice implementations of adaptive filters are available in the references [105] [112].

filter in fig. 4.3 requires three times as many multiplier operations as the equivalent transversal form.

As algorithms of similar complexity now exist for both lattice and transversal filters, the choice is dictated by the particular application. Here the main function is system modelling, though the model parameters are not themselves required. As it was envisaged that off-line processing would be performed using a general-purpose laboratory computer, at least in the initial stages of the investigation, the numerical advantage of lattice filters was not important. The transversal structure is simpler to implement and analyse and generally permits faster computation. It was therefore chosen for this study, though had it been necessary to construct a dedicated processor, a lattice form might well have been preferred.

The ATF can be used in a variety of situations such as noise cancelling, system identification and channel equalization. The principle behind all of these is the same and is based on the ability of the ATF to dynamically model unknown systems. The general situation is depicted in fig. 4.4 where the input and output signals, (possibly corrupted by independent random noise), to an unknown linear time-invariant system are used as the two inputs. The function of the ATF is to form an output sequence $\{y_j\}$ from the input sequence $\{x_j\}$ which closely approximates the desired signal $\{dj\}$. When convergence has been attained the filter weights characterise the unknown system.

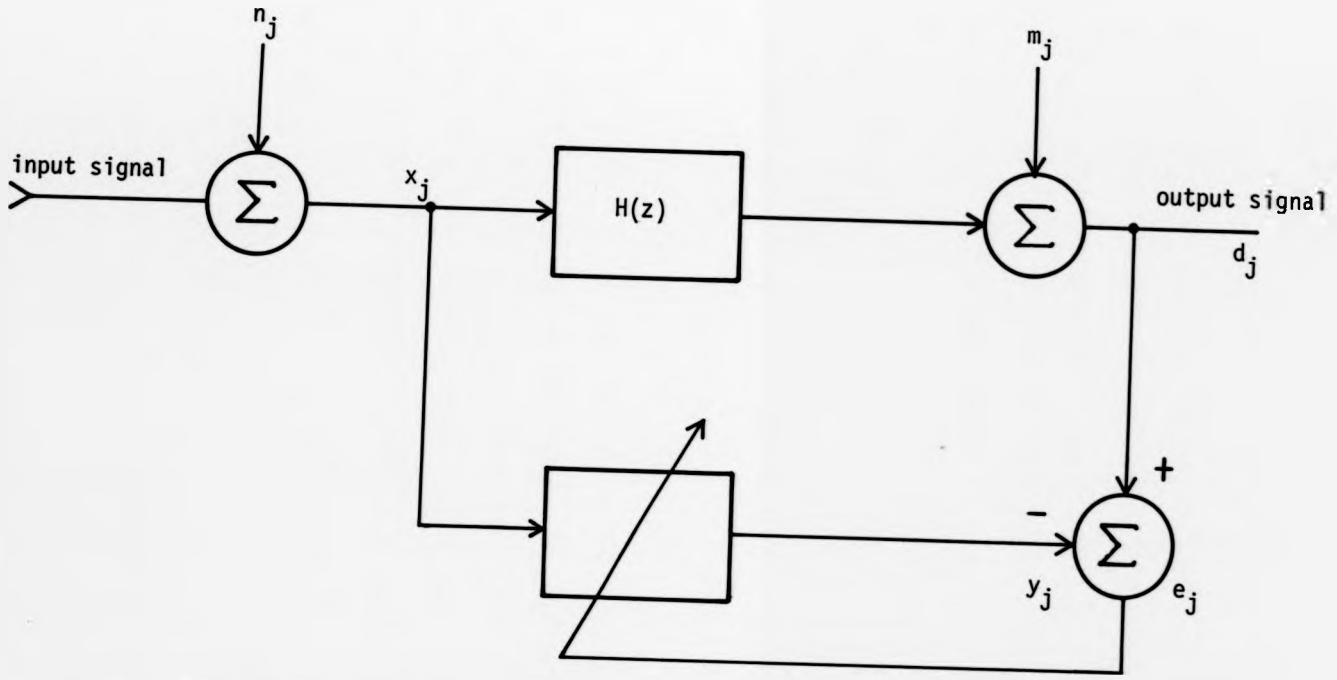



FIGURE 4.4 Use of an adaptive transversal filter to model an unknown system H(z). Input samples $x_j$ (corrupted by independent noise $n_j$) form the input to the ATF, and the output of the unknown system (also computed by independent noise $m_j$) forms the desired input $d_j$ to the ATF. The ATF attempts to model H(z) by minimizing the mean square difference signal between $d_j$ and the filter output $y_j$.

## 4.2 LMS Adaptive Filter Algorithm.

### 4.2.1 Theoretical basis.

The most obvious (and most common) criterion for optimizing filters is to minimise the mean square error (MSE), although this is not the only one that may be used [90]. Wiener's theory of optimal filters is based on this approach as this can yield the best possible linear filter for gaussian signals. Many signal processing and statistical procedures adopt this approach for the same reason and because it is mathematically tractable.

The theory of AF will now be outlined using this criterion, and will closely follow that of Widrow et al [91] [92]. Recall the basic filter structure introduced earlier which yielded an expression for the output in terms of the filter input and weights. This may be rewritten using matrix notation. Throughout this thesis the following conventions will be observed; matrices and vectors will be underlined, while scalar quantities will be printed in normal type. The transpose of a matrix will be indicated by the superscript T. Equation 4.1 may thus be rewritten:

$$y_j = \underline{X}_j^T \underline{W} = \underline{W}^T \underline{X}_j \qquad (4.3)$$

where the input vector at the j'th sampling instant is

$$\underline{X}_j^T = [x_{0j} x_{1j} \cdot \cdot \cdot \cdot x_{N-1j}] \qquad (4.4)$$

and the weight vector is

$$\underline{W}^T = [\, w_0 \; w_1 \; \cdots \; w_{N-1}] \tag{4.5}$$

The object of the adaptive algorithm is to choose weights such that the MSE signal is minimised. The MSE is formed by squaring expression (4.2) and taking expectations over the ensemble, where E[.] denotes the expectation operator, thus

$$E[e_j^2] = E[d_j^2] - 2E[d_j \underline{X}_j^T]\underline{W} + \underline{W}^T E[\underline{X}_j \underline{X}_j^T]\underline{W} \tag{4.6}$$

This may be written

$$E[e_j^2] = E[d_j^2] - 2\underline{P}^T\underline{W} + \underline{W}^T\underline{R}\underline{W} \tag{4.7}$$

where $\underline{P}$ and $\underline{R}$ are the input correlation matrices defined by

$$\underline{P} = E\,[\, d_j \, \underline{X}_j] \tag{4.8}$$

$$= E\,[\, d_j x_{0j} \;\; d_j x_{1j} \cdots \;\; d_j x_{N-1j} \,]$$

and

$$\underline{R} = E\,[\, \underline{X}_j \, \underline{X}_j^T \,] \tag{4.9}$$

$$= E\begin{bmatrix} x_{0j}x_{0j} & x_{0j}x_{1j} & \cdots & x_{0j}x_{N-1j} \\ x_{1j}x_{0j} & x_{1j}x_{1j} & \cdots & x_{1j}x_{N-1j} \\ \vdots & \vdots & & \vdots \\ x_{N-1j}x_{0j} & x_{N-1j}x_{1j} & \cdots & x_{N-1j}x_{N-1j} \end{bmatrix}$$

For stationary inputs the MSE is a quadratic function of the weights that can be pictured as a concave hyperparaboloid. This function has a unique minimum which may be found by differentiating (4.7) with respect to the weights $\{w_i\}$ to obtain the gradient $\underline{\nabla}E[e_j^2]$

$$\underline{\nabla}\, E[e_j^2] = -2\underline{P} + 2\underline{\underline{R}}\,\underline{W} \tag{4.10}$$

When this is set to zero, this results in the matrix formulation of the Wiener-Hopf equation, yielding $W^o$ as the optimal weight vector.

$$\underline{W}^o = \underline{\underline{R}}^{-1}\underline{P} \tag{4.11}$$

The development of the adaptive filter theory up to this point has been in terms of a general multi-input filter. This will now be particularised to the case of adaptive transversal filters. For these $\underline{X}_j^T$ represents the tapped delay line at sampling instant j, thus

$$\underline{X}_j^T = [\, x_{j-1} \cdots x_{j-N+1} \,] \tag{4.12}$$

In this case the correlation matrices are

$$\underline{R} = \begin{bmatrix} r_{xx}(0) & r_{xx}(1) & \cdots & r_{xx}(N-1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(N-1) & \vdots & \cdots & r_{xx}(0) \end{bmatrix} \tag{4.13}$$

and

$$\underline{p}^T = [\, r_{xd}(0) \quad r_{xd}(1) \quad \cdots \quad r_{xd}(N-1) \,] \tag{4.14}$$

where $r_{xx}(k)$ and $r_{xd}(k)$ are the discrete cross-correlation and autocorrelation matrices defined by

$$r_{xx}(k) = E[\, x_j \, x_{j+k} \,] \tag{4.15}$$

$$r_{xd}(k) = E[\, x_j \, d_{j+k} \,] \tag{4.16}$$

The equivalence of expression (4.11) with the discrete Wiener-Hopf equation is readily shown. The latter may be stated as

$$\sum_{\ell=0}^{N-1} w^o(\ell)\, r_{xx}(k-\ell) = r_{xd}(k) \tag{4.17}$$

for the finite-length, one-sided optimal filter. This may be rewritten in matrix form

$$\underline{W}^o \underline{R} = \underline{P} \tag{4.18}$$

which is clearly equivalent to (4.11).

By noting that the z-transform of the convolution sum in (4.17) is equal to the product of the z-transforms of $w^o(k)$ and $r_{xx}(k)$, (4.17) may be written

$$W^o(z).\delta_{xx}(z) = \delta_{xd}(z) \tag{4.19}$$

and hence

$$W^o(z) = \delta_{xd}(z) \; / \; \delta_{xx}(z) \tag{4.20}$$

where

$$W^o(z) = \sum_{k=-\infty}^{\infty} w^o(k).z^{-k} \tag{4.21}$$

$$\delta_{xx}(z) = \sum_{k=-\infty}^{\infty} r_{xx}(k).z^{-k} \tag{4.22}$$

$$\delta_{xd}(z) = \sum_{k=-\infty}^{\infty} r_{xd}(k).z^{-k} \tag{4.23}$$

This result will be convenient in later analysis of adaptive noise cancelling.

While it is possible to compute the exact solution from expression (4.11) this is impractical when the number of weights is large or the data rate is high as it involves N(N+1)/2 correlation measurements and the inversion of an NxN matrix. In addition this process needs to be repeated continuously if the signal statistics are varying with time. In general $\underline{R}$ and $\underline{P}$ will not be known, though direct inversion of the sample matrices can be employed when these can be estimated [93]. Recent techniques which have been developed to do this more efficiently than by direct calculation are described briefly in a

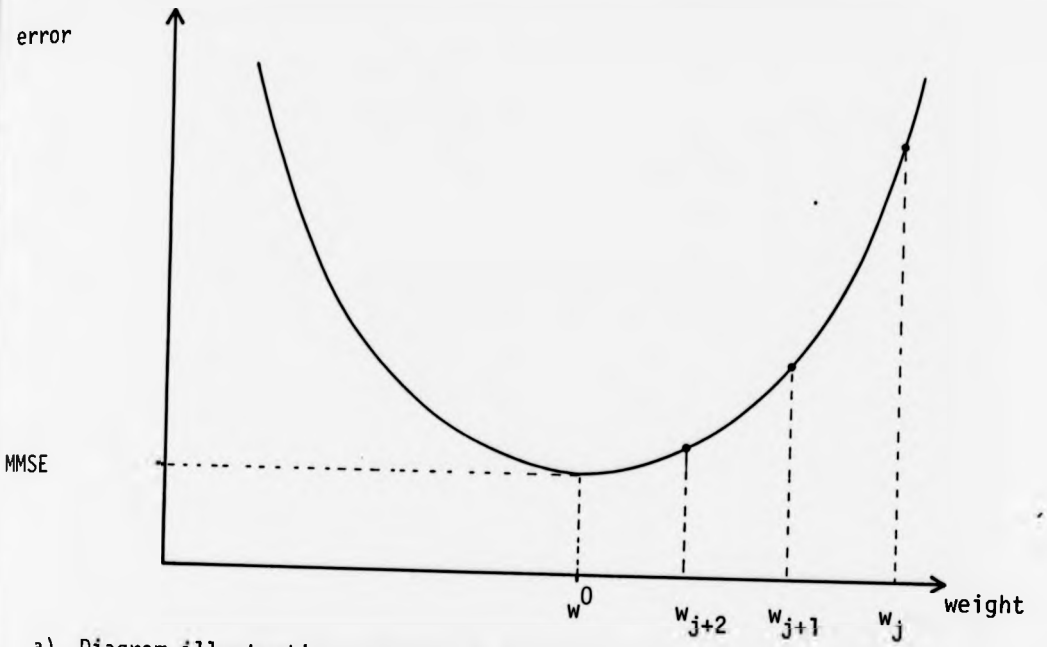
later section, but do not have the same computational advantage as the iterative methods.

The Widrow-Hoff LMS algorithm can circumvent this difficulty by providing a simple iterative procedure to achieve convergence of the filter to the Wiener solution. This algorithm does not explicitly compute the correlation matrices or matrix manipulations and only requires that the weight vector be adjusted from estimates of the gradient, in addition to computing the transversal filter output. The algorithm is based on the method of steepest descent to find the minimum of the error function. This involves adjusting the present weight vector by an amount opposite in sign and proportional to the magnitude of the instantaneous gradient, thus

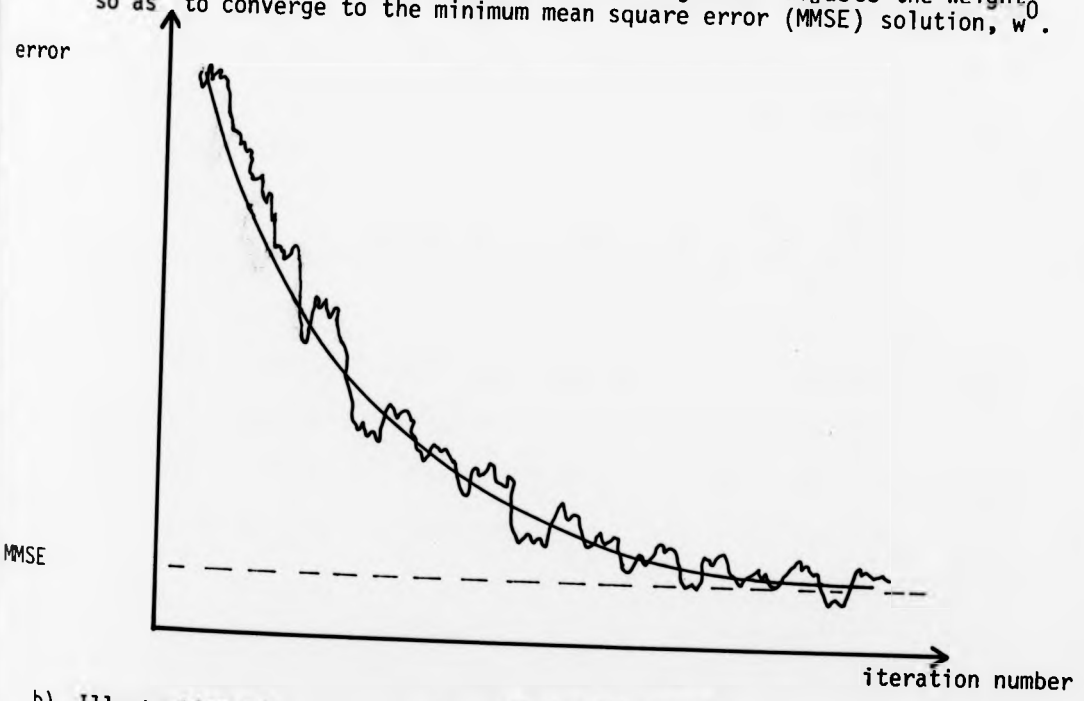
$$\underline{W}_{j+1} = \underline{W}_j - \mu \underline{\nabla}_j \tag{4.24}$$

$\mu$ is a parameter that controls the rate of convergence and stability. This is illustrated in fig. 4.5 for a one-weight filter, showing the parabolic error surface and several iterations of the weight vector as it converges to the MSE solution. The LMS algorithm uses an estimate of the gradient $\hat{\underline{\nabla}}$ based on the instantaneous error $e_j^2$ rather than the true error gradient, thus

$$\hat{\underline{\nabla}}_j = \left[ \frac{\partial e_j^2}{\partial w_0}, \ldots, \frac{\partial e_j^2}{\partial w_{N-1}} \right]_{\underline{W}=\underline{W}_j} \tag{4.25}$$




a) Diagram illustrating successive iterations of the gradient descent adaptive algorithm. The parabola represents the error function for a one weight filter. Each iteration of the algorithm adjusts the weight so as to converge to the minimum mean square error (MMSE) solution, $w^0$.




b) Illustration of the convergence of the filter to the MMSE solution as adaption progresses. The rough curve represents one such trajectory, illustrating random disturbances about the mean trajectory (smooth curve) caused by gradient estimation errors.

FIGURE 4.5

$$= 2e_j \cdot \left[ \frac{\partial e_j}{\partial w_1}, \ldots, \frac{\partial e_j}{\partial w_{N-1}} \right]^T \Big|_{\underline{W}=\underline{W}_j} \tag{4.26}$$

Using relations (4.2) and (4.3) this can be written

$$\hat{\underline{\nabla}}_j = -2e_j \underline{X}_j \tag{4.27}$$

which when applied to equation (4.12) result in the Widrow-Hoff weight update algorithm

$$\underline{W}_{j+1} = \underline{W}_j + 2\mu e_j \underline{X}_j \tag{4.28}$$

This procedure provides an unbiased estimate of the Wiener weight vector $W^O$ if the inputs are uncorrelated, though published reports indicate that even strongly correlated signals may not seriously degrade filter performance [94].

### 4.2.2 Convergence of the LMS Adaptive Filter.

The stability and convergence properties of the steepest descent gradient searching algorithm presented here have been analysed by Widrow et al [91]. A summary of the main results will now be presented, due to their importance to this investigation.

Unconditionally stable operation is obtained provided that

$$0 < \mu < 1/\lambda_{max} \tag{4.29}$$

where $\lambda_{max}$ is the largest eigenvalue of $\underline{R}$. In general $\underline{R}$ or its eigenvalues will not be known, but a sufficient condition which ensures

that (4.29) is satisfied is

$$0 < \mu < 1 / \sum_{i=1}^{N} E [x_i^2] \tag{4.30}$$

since the sum of the eigenvalues is equal to the trace of a real symmetric matrix. It is convenient in practice to use the normalised convergence coefficient $\mu'$, which ensures that stability is guaranteed provided that

$$0 < \mu' < 1 \tag{4.31}$$

$\mu'$ is then defined by

$$\mu' = \mu \sum_{i=1}^{N} E [x_i^2] \tag{4.32}$$

Another important property is the convergence time, which has been shown to be exponential. In general each weight coefficient will not converge to the optimal value at the same rate. It is possible to analyse the convergence of the AF in terms of a set of transformed weights $\{w'_j\}$ which are mutually orthogonal. These converge independently according to a first order relaxation process which constitute the basic modes of the adaptive system. The p'th mode has a time constant $\tau_p$ that is related to the p'th eigenvalue $\lambda_p$ of $\underline{R}$ as follows (for the case of slow adaption)

$$\tau_p \simeq 1 / 2\mu\lambda_p \tag{4.33}$$

where $\tau_p$ is measured in units of iteration cycles. The time constant for the MSE arising from the p'th mode is half this value, i.e.

$$\tau_p(\text{mse}) \simeq 1 / 4\mu\lambda_p \qquad (4.34)$$

This expression shows one of the major defects of this algorithm in that different modes may converge at different rates. By noting that the maximum possible value of $\mu$ is $1/\lambda_{max}$ (4.29), it may be seen that the slowest converging mode is that corresponding to the minimum eigenvalue $\lambda_{min}$, and hence overall convergence is limited by the ratio of the maximum to minimum eigenvalues. This results in a lower bound upon $\tau_p(\text{mse})$ as

$$\min \{\tau p(\text{mse})\} = \lambda_{max} / 4\,\lambda_{min} \qquad (4.35)$$

It is this difficulty that has prompted research into adaptive algorithms that are independent of eigenvalue spread. These are briefly reviewed in a later section. While it is difficult to ascertain the eigenvalues of an unknown (and possibly time-varying) system, it can be shown that the eigenvalue spread corresponds approximately to the spread in spectral power density for large N [95]. This can be useful to determine whether or not the LMS algorithm is viable in practice.

It may be helpful to consider fig. 4.6 to visualize this situation. A set of equi-valued contours of the mean square error surface are plotted as a function of the weights for a two-weight AF. In general the contours are ellipses whose major and minor axes dimensions correspond to the eigenvalues of the autocorrelation matrix, and the directions of the principal axes are parallel to the eigenvectors. It is clear that the perpendicular to the gradient of the error surface does not pass through the minimum, resulting in a suboptimal trajectory. This can seriously degrade the convergence when the ellipse dimensions are very unequal, but Widrow has argued that unless the ratio is very large ($>10$) this may not present difficulties

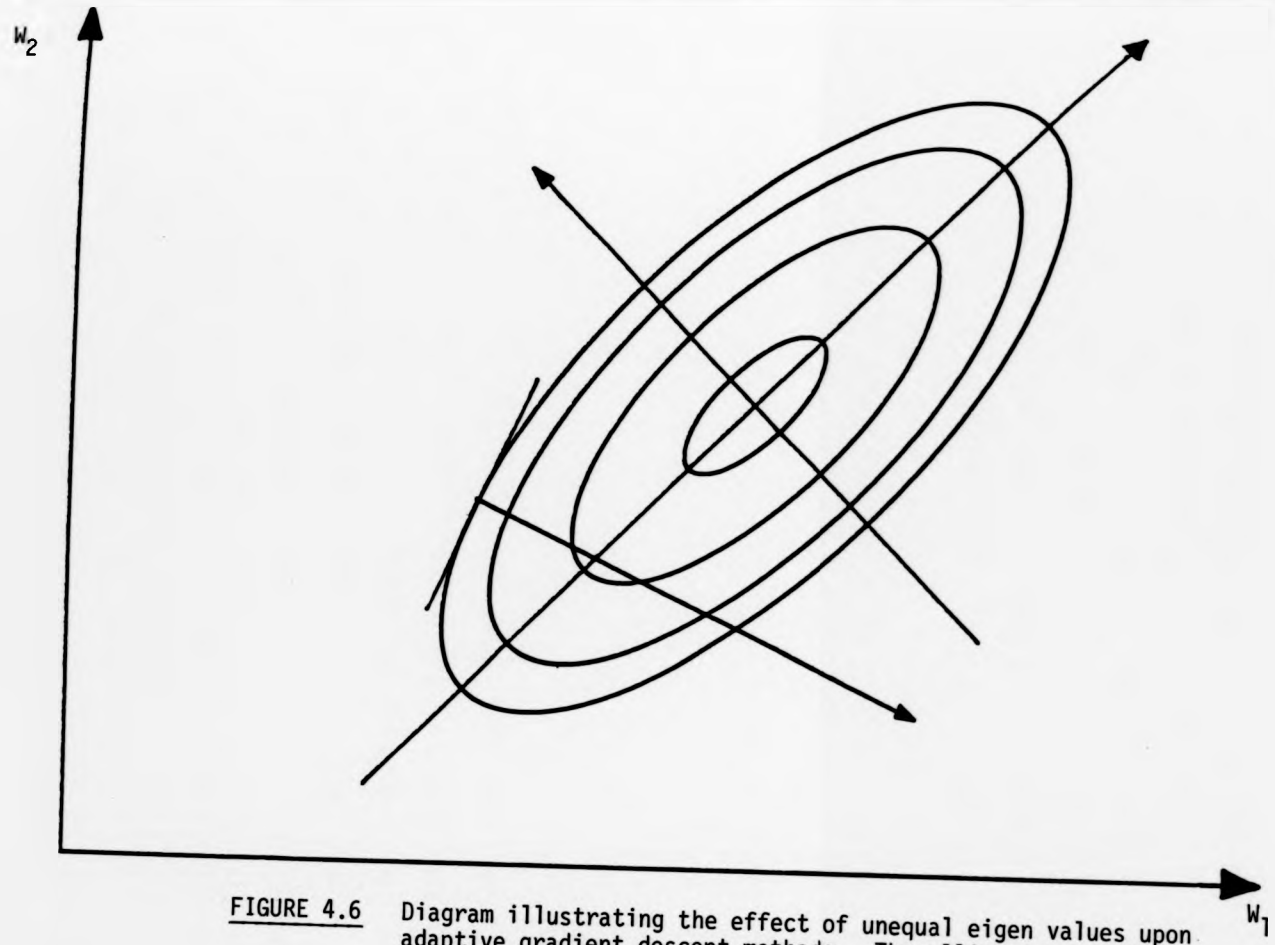



FIGURE 4.6 Diagram illustrating the effect of unequal eigen values upon adaptive gradient descent methods. The ellipses represent equal-valued MSE contours as a function of the two weights $w_1$ and $w_2$, and the principal axes show the orientation of the eigenvectors. The gradient to the error surface does not pass through the MMSE solution and hence results in slow convergence.

in practice.

An important property is the accuracy with which the filter may be expected to converge to the Wiener solution. Due to statistical errors in the samples which result in gradient estimation noise, the weight vectors rarely converge to the exact solution, resulting in an error greater than the minimum MSE. A dimensionless quantity that reflects the weight vector 'misadjustment' has been defined [87] as

$$M = \frac{\text{excess mean square error}}{\text{miminum mean square error}} \tag{4.36}$$

It is shown in [91] that the misadjustment can be related to the filter order and average eigenvalue by the expression:

$$M = \mu N \lambda_{ave} \tag{4.37}$$

where $\lambda_{ave}$ is the average eigenvalue. This may then be rewritten using expression (4.32) to yield a useful relation for the misadjustment in terms of the filter order and the time constant of the adaptive process, assuming that all the eigenvalues are equal:

$$M = N / 4 \tau_p(mse) \tag{4.38}$$

Note that $\mu'$ is exactly equal to the misadjustment M, since the sum of the diagonal elements of a real symmetric matrix is equal to the sum of the eigenvalues.

There are two practical points regarding adaptive transversal filters that will now be briefly mentioned before going on to discuss the performance characteristics. The first is a simple modification that can be made to permit data having non-zero mean to be filtered. Under this condition the filter does not normally converge to the

correct solution, and can even become unstable. Ferrara [96] has analysed this difficulty and proposed the simple expedient of augmenting the weight vector by an additional 'bias weight', and the tapped delay line with a constant to represent the dc component to be filtered out, viz:

$$\underline{X} = [\, 1 \;\; x_1 \;\; . \;\; . \;\; . \;\; x_N \,]^T \tag{4.39}$$

and

$$\underline{W} = [\, w_0 \;\; w_1 \;\; . \;\; . \;\; . \;\; w_N \,]^T \tag{4.40}$$

The effect of this is to adaptively compensate for any dc component present in the input. Further details are available in the reference [96]. This simple modification is straightforward to apply and carries with it very little overhead.

The second consideration is that of causality. The theoretical treatment outlined in the previous section required an infinite-length, two-sided (and therefore non-causal) filter, which is not physically realizable. As typical Wiener impulse responses tend to decay asymptotically to zero from the centre of the response, it is usually possible to implement an approximate solution using a finite length transversal filter. To correct for the delay that arises in the filter output due to the two-sided filter impulse response, a corresponding delay is inserted into the primary input. This results in an overall delay in the filter output, but is not important in many practical situations, such as offline processing or real-time processing using a delayed time reference. The delay required is not critical, and Widrow et al [92] suggest that a value of approximately half the filter length is satisfactory. The weight vectors will converge to the desired two-sided response centred about the chosen delay value.

### 4.2.3 Filter Performance.

The effect on filter performance of the parameters N and $\mu$ will now be discussed. To do this three aspects will be considered,

(i) filter accuracy
(ii) speed of convergence
(iii) computational accuracy

The first two are of importance in characterising the fundamental dynamic properties of the filter, whereas the last is primarily of importance as regards implementation, though this can have considerable bearing on practical designs. Filter inaccuracy can be represented by the difference between the ideal filter and the actual filter output. Inaccuracy can arise in several ways which will now be briefly discussed.

a) errors in the weight vectors due to the use of the instantaneous gradient instead of the true gradient. These are present even when the filter has attained convergence in the mean, and the weights at each sample time merely fluctuate about the true solution.
b) errors in the weight vector arising from inability to track nonstationary events in the data (if present)
c) insufficient weights to adequately represent the Wiener solution, and
d) lack of precision in storing the filter weights and performing the filter computations.

The first is a feature of the LMS gradient search technique and represents the lower limit in accuracy obtainable by this method, though arbitrarily small misadjustment error can be achieved by using a

sufficiently slow adaption rate (equation 4.35). The errors due to case b) can be minimised by ensuring that adaption occurs rapidly. This is in conflict with the previous requirement and a compromise is required. This will be discussed presently. c) and d) need not present a problem provided that sufficient attention is paid to these in the implementation.

The choice of N is governed by the need to represent the Wiener response with sufficient accuracy, though the minimum value to achieve this is clearly preferable. As the solution is generally not known a priori, this choice is often made heuristically, guided by knowledge of the system and past experience. The number of weights is given by the ratio of the sampling frequency to the desired frequency resolution [91]. This is easily shown as the Fourier transform of the impulse response yields N/2 amplitude coefficients which must span the frequency range up to the Nyquist limit. No analytic expressions have been published to select the optimum value of N for any given situation.

When N has been selected, the misadjustment is determined by the value of $\mu$. Widrow suggests that a suitable strategy is to choose $\mu$ to obtain a misadjustment of approximately 10%, which is likely to be acceptable for many practical applications [87]. When filtering short data records, it may be more desirable to have fairly rapid adaption to obtain efficient filtering and reduced processing time. In practice the initial converging phase in the output can be avoided by applying the filter a second time using the converged weight vector. This can only be done if the data are stationary and off-line processing of the signals is employed.

If the data are non-stationary to some degree, the choice of $\mu$ is no longer arbitrary and must be made to achieve satisfactory tracking performance. Nonstationary systems are notoriously difficult to analyse but an initial attempt has been made by Widrow et al [91] using a simple form of non-stationarity. They concluded from a theoretical and experimental standpoint that the best compromise between misadjustment and mistracking errors was to make these equal. This may be used to determine the value of $\mu$ if the nature of the nonstationarity is known, but a heuristic approach is likely to be necessary in practice.

The precision with which the filter coefficients are stored and the computations performed are important. These considerations have been examined in some detail by Caraiscos and Liu [97] who derived expressions for the contribution to the MSE arising from the computational process for both fixed point and floating point implementation. In general the precision of the arithmetic and weight storage is more important than that of the data. This can be seen intuitively by considering the basic algorithm (4.28). If $\mu$ is small, it is possible for the update term to be less than the quantization interval, which will prevent convergence from being attained. For this reason the weights need to be stored with greater precision than the filter computation itself requires. If 8 bit precision is sufficient to represent the data (as it is in many applications) much higher precision is required for the weight storage and computation, in the range 16-24 bits [112]. The precise value will depend upon the value of $\mu$ used. The use of single precision floating point operations, typically employing a 24 bit mantissa and 8 bit exponent, is therefore considered adequate for many applications.

The computational time using a serial processor is approximately equal to 2N multiply and add operations. This can be achieved in real time by small computers at low data rates, but may not be possible for data rates corresponding to typical biological signals when the filters have many weights. This may not be a problem in off-line analysis, but hardware designs can readily be implemented should they be necessary. The memory requirements are relatively small, being approximately 2.5N storage cells for the filter weights and delay lines. In addition the code requirements are small due to the simplicity of the algorithm. These factors contribute strongly to the widespread use of this algorithm.

## 4.3 Improved LMS algorithms.

While the Widrow-Hoff method of steepest descent is probably the most widely used and understood adaptive algorithm, it may be inadequate in some circumstances due to slow convergence. Research in recent years has been directed at finding algorithms with superior convergence properties, and also at methods of achieving computational reductions. The former is particularly important for applications to non-stationary systems. Algorithms which offer computational savings are of most value to real-time systems where the data rate is high and filter order fairly large. As it was envisaged that the technique would be implemented on a general-purpose laboratory computer for off-line analysis this was not the most important aspect, though obviously a fast algorithm is desirable to achieve efficient data analysis. Some of these techniques will therefore be briefly discussed.

### 4.3.1 Block Filtering

Substantial computational savings are obtained if block adaptive filtering is employed as this permits costly convolutions and correlations to be performed more speedily using the FFT. Significant computational savings can be achieved over the normal LMS method for $N > 32$. This may be done by partitioning the input data sequences $\{dj\}$ and $\{x_j\}$ into blocks of length L and performing adaptive filtering upon each in turn. Clark et al [98] have shown that maintaining the filter weights constant throughout each block and updating the weights once at the end of each block is essentially equivalent to the normal LMS ATF when the inputs are stationary, though the output sequence $\{y_j\}$ is delayed by one block length. (This is usually not important except for signal processing applications that involve control operations or other real-time applications.)

It has been shown that the adaption accuracy and speed are similar to the normal LMS method, provided that the time lag between the input sequences is small compared to the block length. Because the weights are only updated once every L samples, to achieve the same convergence rate and misadjustment $\mu_B$ (the block LMS parameter $\mu$) must be equal to $\mu L$. However the stability conditions for $\mu$ and $\mu_B$ are identical, which may restrict the value of $\mu_B$ in cases when fast adaption or large block length are involved.

### 4.3.2 Frequency Domain Adaptive Filter

In the previous section the FFT was simply used to reduce the number of computations involved in the time-domain BLMS ATF. A frequency-domain ATF can be constructed with the same computational

advantage through the use of the FFT, but which is structurally different from the time-domain AF. Instead of having a weight vector in the time domain which is used to implement the transversal filter, it is possible to define the filter in the frequency domain, with the weights acting upon the Fourier coefficients of the data sequences. The complex form of the LMS adaptive algorithm must be used in this case [111]. Clark et al [99] have shown that this structure is equivalent to the time-domain AF. An example of this type of structure is shown in fig. 4.7.

There are two reasons for considering frequency-domain ATFs whose weights act on the frequency transformed data. Firstly, the use of a single weight for each frequency bin can introduce simplifications into the analysis of their performance, which is very difficult in the time-domain versions. This can arise if the input data can be considered as uncorrelated random processes, in which case each filter weight is adjusted independently of the others, permitting analytic expressions for the MSE statistics to be derived. More importantly, the convergence of these filters can be optimised by computing convergence coefficients for each weight, so that the eigenvalue spread need no longer restrict the adaption speed, since each mode can adapt at the same rate. Several schemes to implement this have been proposed as discussed in [100] and [101]. Since the filter weights are held constant throughout each block, the filter is unable to respond to rapid changes, which can be a disadvantage if continuous tracking of the system is important. In these cases sequential updating algorithms must be employed.

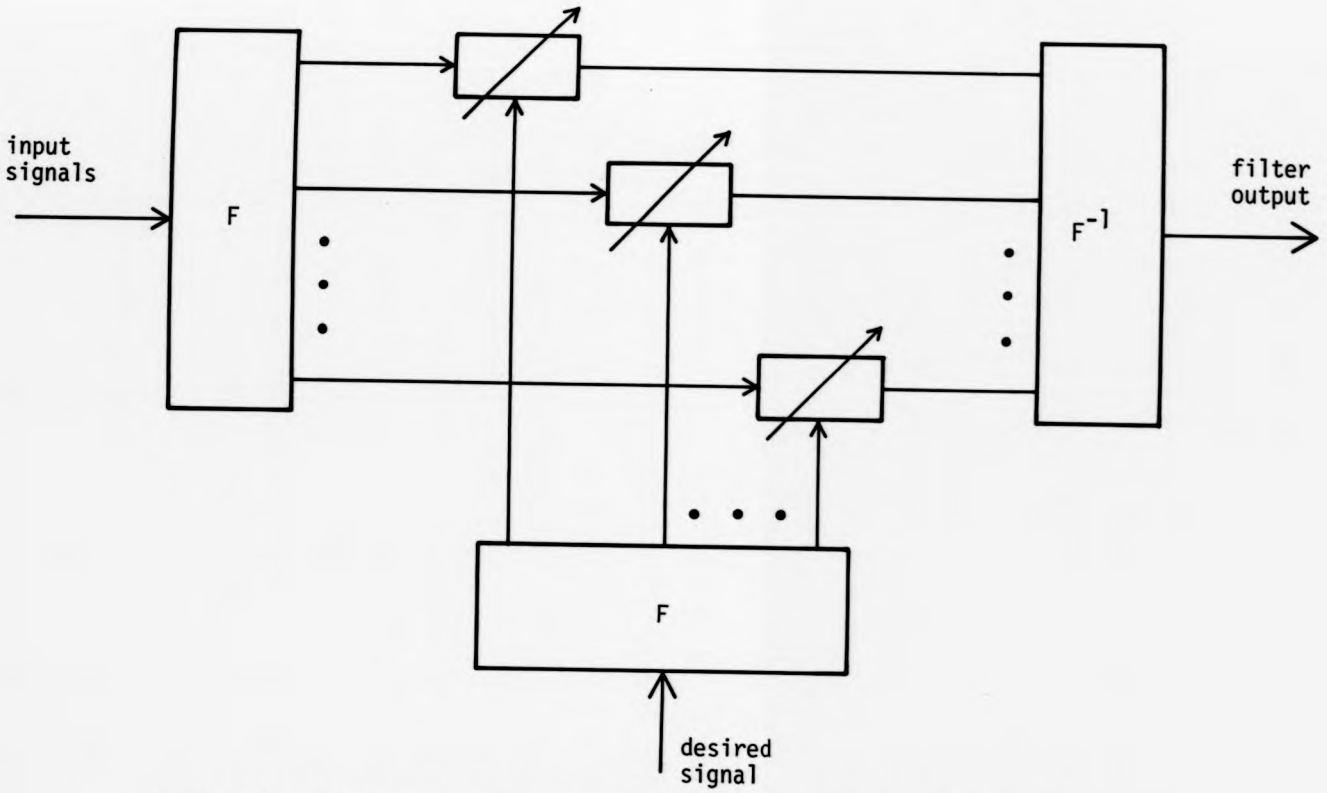



FIGURE 4.7 A frequency domain adaptive filter. The input data are transformed into the frequency domain by the Fourier Transform, and each frequency component is filtered separately before being inverse transformed to form the output sequence. F represents the Fourier Transform operation on a block of input samples, and $F^{-1}$ the inverse operation.

### 4.3.3 Time-sequenced Adaptive Filter

A modified version of the basic LMS ATF known as the Time-sequenced Adaptive filter (TSAF) was introduced by Ferrara [96] [102] to permit adaptive filtering of a class of non-stationary data to be described. It offers the possibility of effecting rapid changes in the impulse response to sudden non-stationarities, while maintaining the accuracy of slow adaption. This technique is suitable for recurring or periodically occurring signals when the recurrence moment (or period) is known a priori. In such cases the data may be considered to be composed of piece-wise stationary segments of known duration with respect to the signal recurrence moment. For each piecewise stationary segment there will be an associated error surface and optimal weight vector. By using a different weight set to form the filter output during each separate segment, it is possible to achieve the optimal time-varying filter, whereas a normal AF would converge to the best time-invariant filter. A block diagram of the time-sequenced approach is presented in fig. 4.8. Ferrara has analysed this technique and verified its validity both theoretically and by computer simulations. In appropriate situations the gain in signal improvement can be substantial, permitting finer detail to be resolved. This is gained with little extra computational overheads since only one weight set is used at any moment. The storage requirements are greater as a result of the extra weights.

### 4.3.4 Multi-reference Adaptive Filter

A powerful extension of the basic AF is the multi-reference AF, introduced by Widrow et al [92] [103]. If several linearly independent reference inputs are available containing components correlated with those in the primary input, it is possible to provide a filter structure

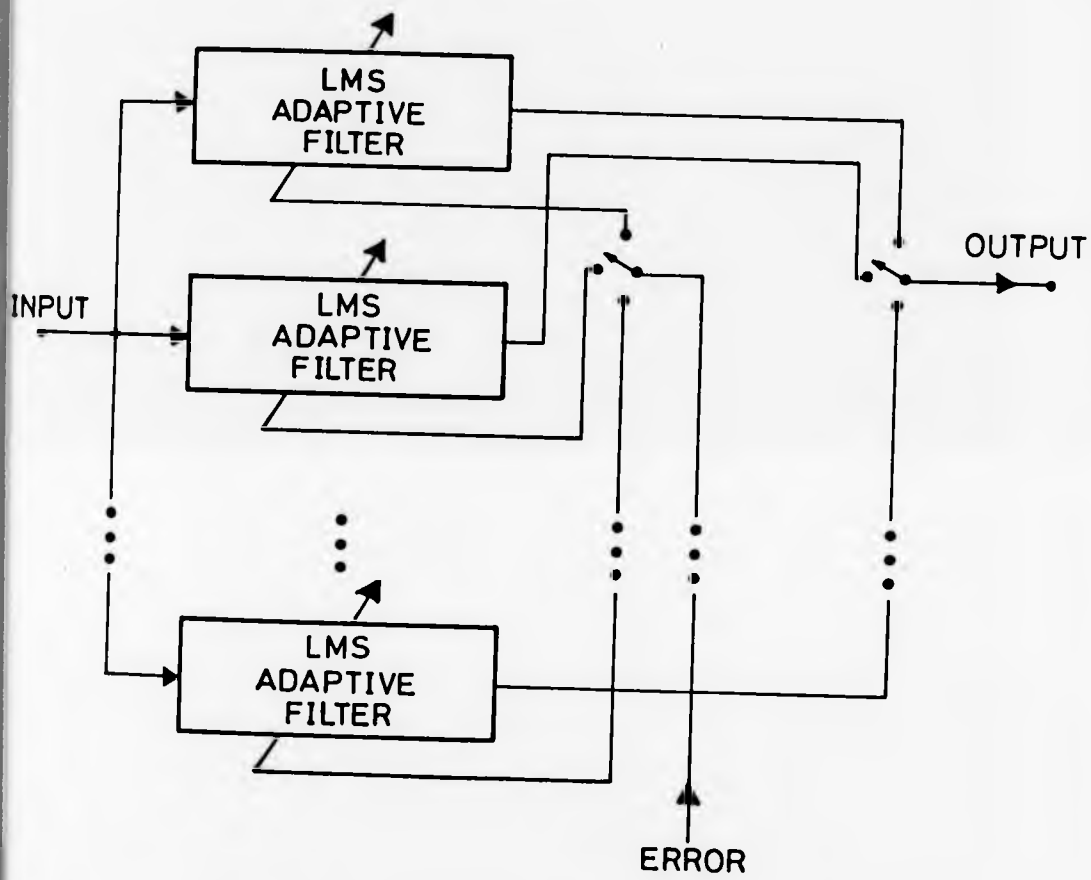



FIGURE 4.8 The time-sequenced adaptive filter may be notionally depicted as a set of separate adaptive filters which are operated in sequence, with the output taken from each in turn. An external sequencing signal controls the commutator.

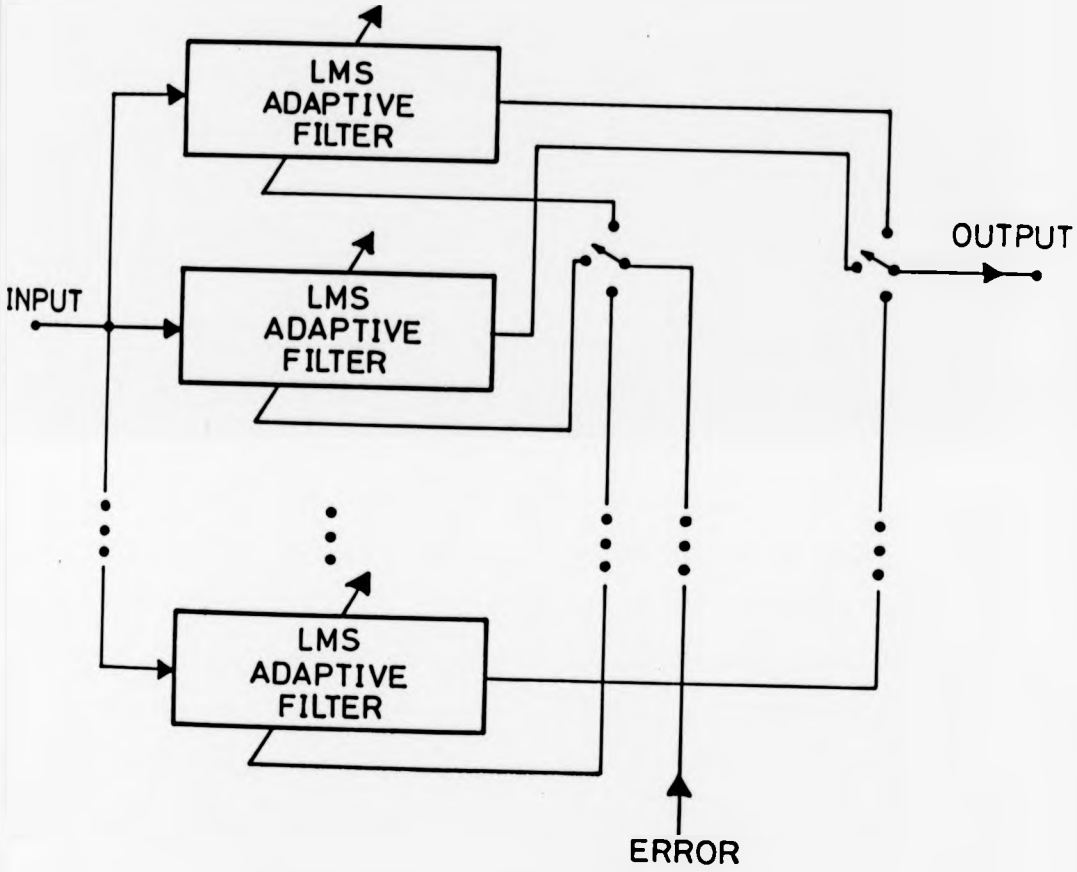



FIGURE 4.8 The time-sequenced adaptive filter may be notionally depicted as a set of separate adaptive filters which are operated in sequence, with the output taken from each in turn. An external sequencing signal controls the commutator.

that permits attenuation of these components. This is shown in fig. 4.9. In the simple two-input AF the weights are updated at each sampling instant by a computation involving the reference signal and error signal. This can be extended to the case of a multi-reference ANC by forming a combined reference output from the sum of each filter output, and using the resulting error signal to adapt the weights of each individual filter as before. These then converge to the multi-reference Wiener solution which optimises the MSE in the desired signal estimate.

A very brief description of this filter is contained in [92], but lacks important detail. In particular no analysis of convergence properties or other filter parameters was presented, though a recent paper by Ferrara and Widrow [103] analysed the case of a multi-reference adaptive filter that may be used to enhance a signal in additive noise when several such channels are available. In this case each channel contains signal components whose precise waveform may vary but which are correlated with each other, albeit in unknown ways. The noise signals are uncorrelated with each other and with the signal from channel to channel. Simply adding the channels together can lead to severe signal distortion and may not provide the best possible noise cancellation, especially if the signal to noise powers vary in each channel. The multi-reference adaptive filter seeks to optimally combine these to provide a relatively undistorted signal estimate with maximum noise cancellation. Formulae were derived for the signal distortion and output noise power which showed that good estimation of the signal is possible provided that the sum of the individual channel signal to noise power ratios as a function of frequency is large compared to unity at all frequencies. The output SNR is then approximately equal to the sum of the individual channel SNR figures and distortion is small. Note

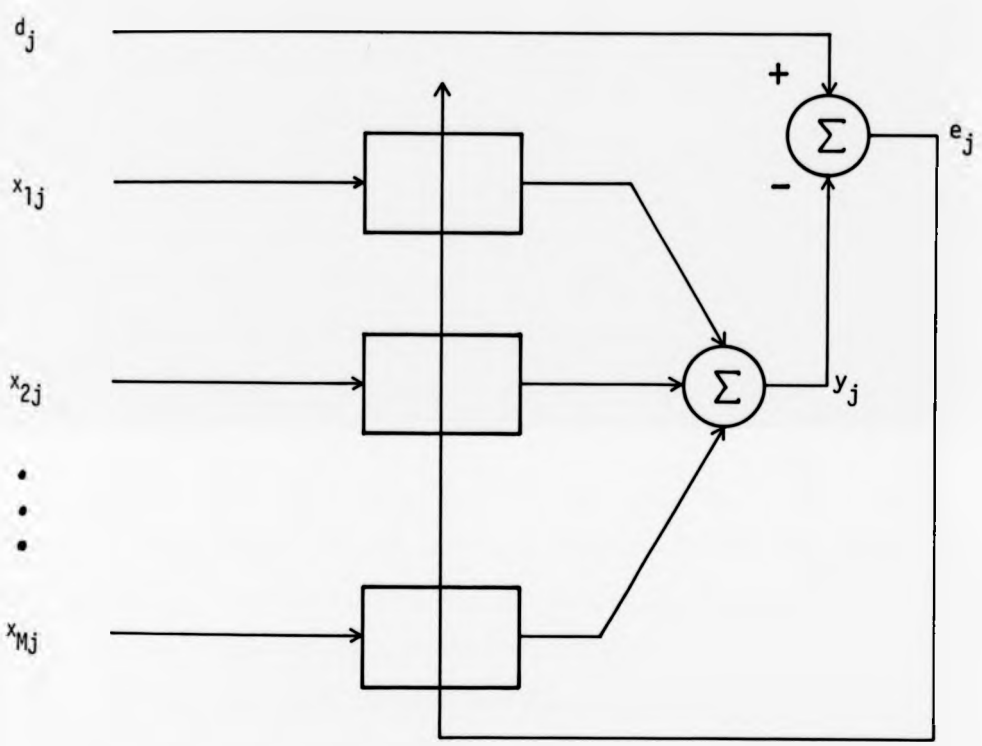



FIGURE 4.9 The multi-reference adaptive filter employs a set of adaptive transversal filters acting on each reference input $x_{mj}$ whose outputs combine to form a single output $y_j$. This is compared with the desired input $d_j$ and the error signal $e_j$ is used to adapt each filter.

that this mode of using the adaptive filter is complementary to the adaptive noise canceller to be described in the next section, which assumes that it is the noise signals that are correlated in the reference channels. The channel enhancement mode of applying an adaptive filter is illustrated in fig. 4.10.

### 4.3.5 Exact Least Squares Algorithms.

In recent years activity has focussed on adaptive algorithms which attempt to compute the exact least squares solution at each sample time. These can offer significantly faster convergence than gradient methods if the eigenvalue disparity is large (e.g. >10) and there is little uncorrelated noise present. When these conditions are fulfilled, convergence can be achieved in about 2N samples. The computational speed of these algorithms was initially rather poor, but algorithms now exist for both lattice and transversal filters which approach the computational speed of the LMS algorithm [104] [105], though they generally involve approximately three times as many multiplications. The disadvantages of these methods are that they are more difficult to analyse and implement and can involve greater computations. They may also be susceptible to instability in certain situations.

It is uncertain whether they can significantly improve upon the performance of the LMS algorithm when the data do not have high eigenvalue disparity or significant uncorrelated noise components [104], and it is not clear that they are equally effective when applied to nonstationary data. Widrow and Walach [106] have shown that for at least one form of nonstationary, (a random displacement of the optimal solution in weight vector space), it is unlikely that any algorithm can do significantly better than the LMS algorithm. For these reasons and

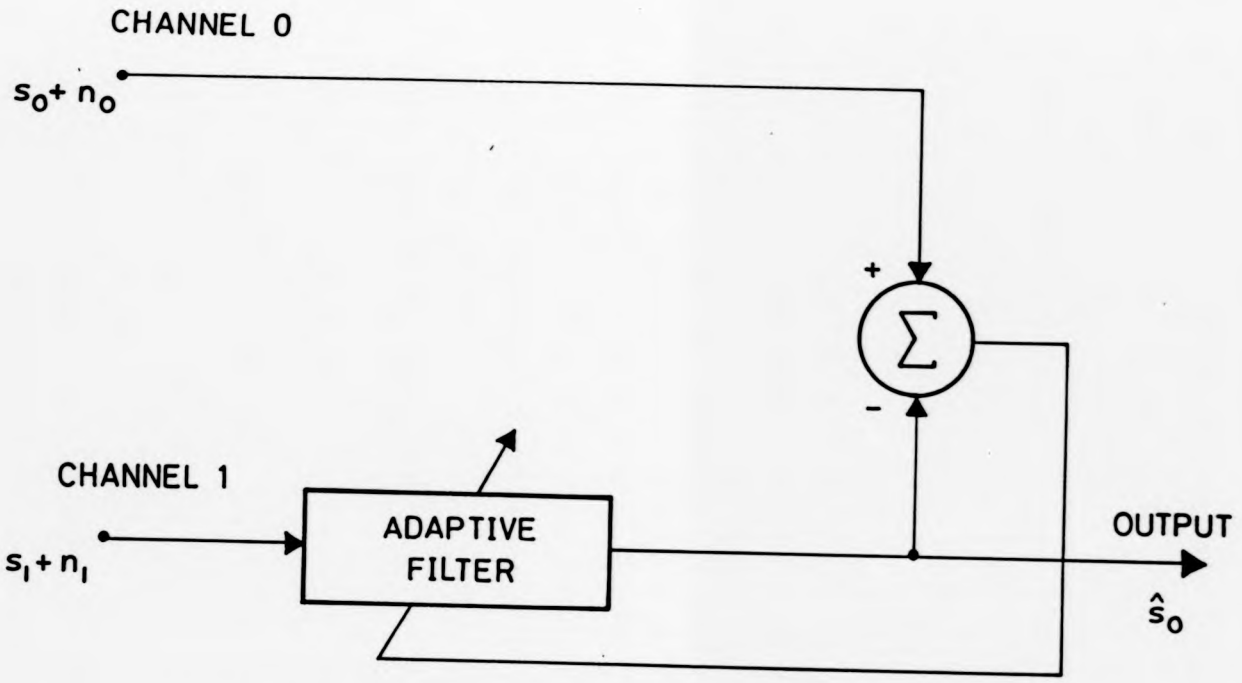



FIGURE 4.10 Use of an adaptive transversal filter in channel enhancement mode. Two channels contain signal components in independent uncorrelated noise. The adaptive filter attempts to form a best least squares estimate of $s_0$ by enhancing the correlated signal components with respect to the uncorrelated noises $n_0$ and $n_1$.

because it seemed more appropriate to investigate the possibility of adaptive filtering of evoked potentials using the most straightforward method available consideration of these was deferred until initial results became available.

## 4.4 Adaptive Noise Cancelling.

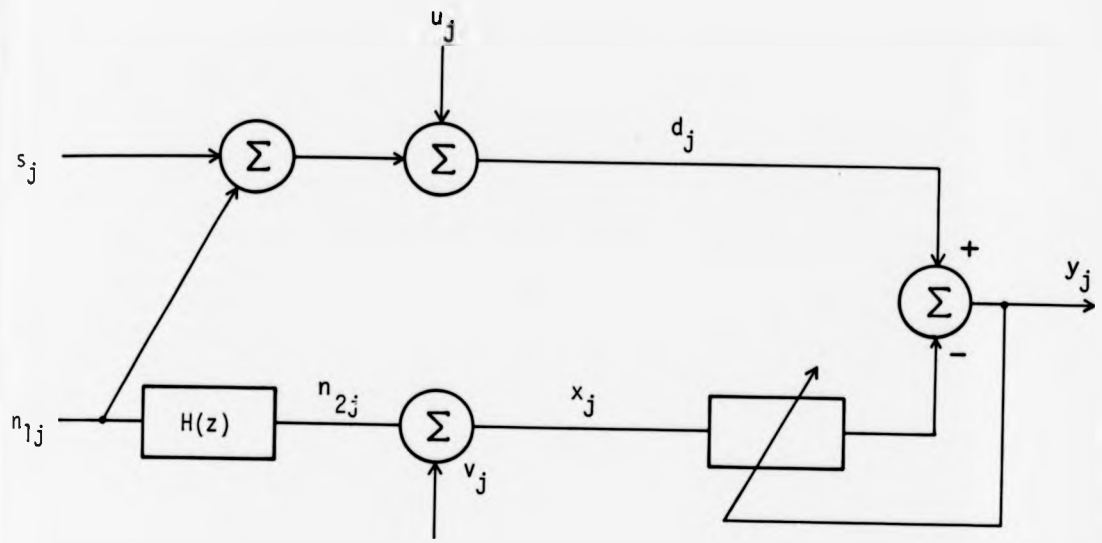
### 4.4.1 Basic principle of adaptive noise cancelling.

The particular technique to be considered here is depicted in fig. 4.11 which shows a single-channel adaptive noise canceller (ANC) comprising an ATF with 2 inputs and one output. The primary input $d_j$ contains a signal $s_j$ combined with an uncorrelated noise $n_{1j}$, and a reference input contains a noise $n_{2j}$ which is correlated with $n_{1j}$ but not $s_j$. In addition independent random noise components $u_j$ and $v_j$ are present in the primary and reference inputs respectively. It is assumed that $s_j$, $n_{1j}$, $n_{2j}$, $u_j$ and $v_j$ are statistically stationary, zero-mean random processes and that $n_{1j}$ and $n_{2j}$ are related by an unknown system function h(k) such that

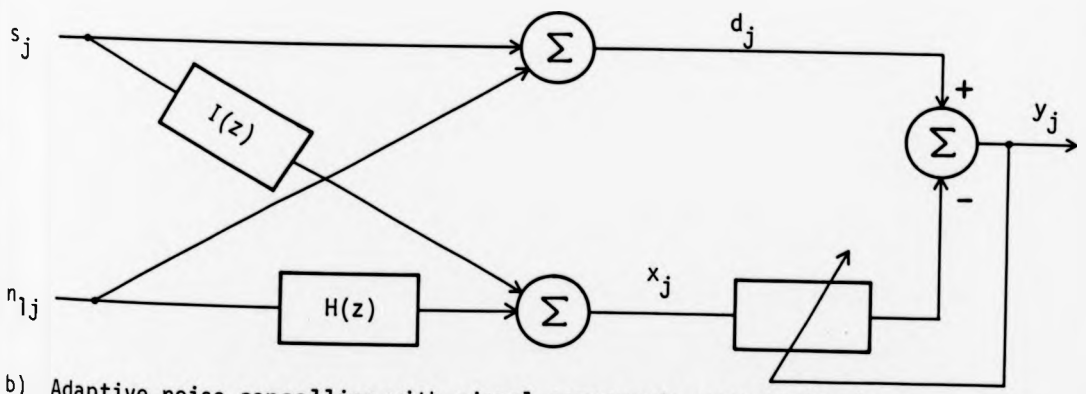
$$n_{2j} = n_{1j} * h(k) \qquad (4.41)$$

where * denotes convolution.

The object of adaptive noise cancelling is to obtain a filter function which uses the reference to optimally cancel the correlated component $n_{1j}$ from $d_j$, leaving a better estimate of $s_j$ at the output. While this can be done using a fixed filter, a priori knowledge of the different signal statistics would be required, and a filter designed to implement this cancelling function for each individual situation. Because this information is not generally available and because the




a) Adaptive noise cancelling with uncorrelated noise components $u_j$ and $v_j$ present in the filter inputs. $s_j$ represents the signal and $H(z)$ represents the transfer function relating the correlated noise components $n_{1j}$ and $n_{2j}$.




b) Adaptive noise cancelling with signal components present in the reference input. $s_j$, $n_{1j}$, $n_{2j}$ and H(z) defined above. I(z) represents the signal pathway into the reference input.

FIGURE 4.11 Use of an adaptive transversal filter to obtain noise cancellation.

appropriate filter may be estimated relatively easily by adaptive means, this approach is generally preferable, though it is not fundamental to this technique.

It will now be shown how an ATF that attempts to minimise the mean square error can yield noise cancellation in general even when the signal statistics are unknown. Assume that $u_j$ and $v_j$ are zero for simplification. The time index will be omitted in the following for brevity. The output e is then

$$e = s + n_1 - y \tag{4.42}$$

and

$$e^2 = s^2 + (n_1 - y)^2 + 2\, s\, (n_1 - y) \tag{4.43}$$

taking expectations,

$$E[\, e^2 \,] = E[\, s^2 \,] + E[\, (n_1 - y)^2 \,] \tag{4.44}$$

since s is independent of $n_1$ and assumed so for y. As $E[s^2]$ is independent of the filter, the minimum output power is given by

$$\min E[\, e^2 \,] = E[\, s^2 \,] + \min E[\, (n_1 - y)^2 \,] \tag{4.45}$$

so that as the filter converges to the minimum mean square output power, it also minimises $E[(n_1-y)^2]$. The filter output y is then the best least mean squares estimate of $n_1$, and e of s, within the constraints of the particular filter structure employed.

The usefulness of this to signal processing is evident in that if a reference is available that is strongly correlated with the noise present in a primary signal, cancellation of this noise is possible

using only this correlation, independent of their individual spectral characteristics. This will be shown more clearly in later analyses. Note that signal enhancement can also be obtained if two correlated signals are available, corrupted by uncorrelated noise. In this case $n_1$ and $n_2$ would represent the correlated signal components and s an interference component. The output y then yields the best least squares estimate of the signal $n_1$ in the primary input. These arguments can also be extended to include multiple reference inputs.

It is important to consider how the filter behaves when this simple situation no longer holds. There are four cases that can be considered

a) uncorrelated random noises present in both inputs i.e. $u_j$ and $v_j$ no longer zero
b) signal components present in the reference input
c) nonstationary signals, and
d) time-varying transfer function

It has been argued [107] that the effects of c) and d) are essentially the same as a), so analysis may be restricted to the first two. These have been analysed by Widrow et al [92], whose main results will be reproduced here due to their importance in assessing the potential usefulness of ANC in any particular situation.

### 4.4.2 Case I. Uncorrelated noises present in both inputs.

This is the most likely situation to be encountered in practice, as there are very few situations when neither input has little or no uncorrelated noise components. It is clear that uncorrelated noise in

the primary input can not be cancelled, and perfect cancellation can not be achieved if either input contains uncorrelated noise. Expressions derived by Widrow et al [92] are presented in this section that permit the performance of the ANC to be predicted in these cases. The situation to be analysed is shown in fig. 4.11(a) and described in the previous section. As $s_j$, $u_j$, $v_j$ and $n_j$ are taken to be mutually uncorrelated signals, the Wiener-Hopf response defined by equation (4.17) may be evaluated. The autospectrum and cross-spectrum of the filter inputs may be written

$$\delta_{xx}(z) = \delta_{vv}(z) + \delta_{nn}(z).|H(z)|^2 \qquad (4.46)$$

$$\delta_{xd}(z) = \delta_{nn}(z).H(z^{-1}) \qquad (4.47)$$

where $\delta_{xx}(z)$ and $\delta_{xd}(z)$ are the discrete cross-correlation and autocorrelation functions defined by (4.22) and (4.23), and $\delta_{uu}(z)$, $\delta_{vv}(z)$ and $\delta_{nn}(z)$ are the discrete crosscorrelation functions for the signals $u_j$, $v_j$ and $n_j$. The Wiener solution defined by (4.20) may thus be written

$$W^o(z) = \frac{\delta_{nn}(z).H(z^{-1})}{\delta_{vv}(z) + \delta_{nn}(z).|H(z)|^2} \qquad (4.48)$$

The performance of the ANC can be conveniently measured by the ratio of the SNR at the ANC output, $\rho_{out}(z)$, to that at the primary input, $\rho_{pri}(z)$. As the signal spectrum appears in both numerator and denominator of this expression it can be neglected, since it is unchanged by the filter. The SNR improvement may thus be written

$$\frac{\rho_{out}(z)}{\rho_{pri}(z)} = \frac{\text{primary input noise}}{\text{ANC output noise}} \tag{4.49}$$

$$= \frac{\delta_{nn}(z) + \delta_{uu}(z)}{\delta_{uu}(z) + \delta_{vv}(z).|W^{o}(z)|^{2} + \delta_{nn}(z).|[1-H(z).W^{o}(z)]|^{2}} \tag{4.50}$$

It is convenient to denote the ratio of uncorrelated to correlated noise power spectra at the primary and reference inputs by A(z) and B(z) respectively, where

$$A(z) = \frac{\delta_{uu}(z)}{\delta_{nn}(z)} \tag{4.51}$$

$$B(z) = \frac{\delta_{vv}(z)}{\delta_{nn}(z).|H(z)|^{2}} \tag{4.52}$$

The Wiener response may then be written as

$$W^{o}(z) = \frac{1}{H(z).(\,B(z) + 1)} \tag{4.53}$$

These expressions may be substituted into (4.50) to yield

$$\frac{\rho_{out}(z)}{\rho_{pri}(z)} = \frac{(\,A(z) + 1\,).(\,B(z) + 1\,)}{A(z) + A(z).B(z) + B(z)} \tag{4.54}$$

The degree of cancellation may be ascertained from this expression. When A(z) and B(z) are large, little cancellation is obtained, but increases as they approach zero. Perfect cancellation is predicted in the limit, though in practice this will be limited by considerations such as finite filter length and rounding errors. It is interesting to note that if $v_j$ alone were zero (and hence also $\delta_{vv}(z)$ and B(z) from (4.50) and (4.52)), the Wiener filter becomes equal to the

reciprocal of the channel transfer function H(z). This is no longer true when B(z) departs significantly from zero, indicating that system identification will be subject to bias in these circumstances. In practice it can be difficult to obtain an estimate of A(z) and B(z) from given data alone. In Appendix C it is shown that an equivalent expression can be derived based on the coherence function, shown below. As this function is readily estimated from given data, it proves to be a more useful method of estimating filter performance in practice.

$$\frac{\rho_{out}(z)}{\rho_{pri}(z)} = \frac{1}{1 - |\gamma|^2} \qquad (4.55)$$

where $|\gamma|^2$ is the magnitude squared coherence function defined in Appendix B. This expression is intuitively reasonable, as the coherence is a measure of the degree to which a signal may be obtained by linear operations on another signal. When the coherence is near unity, indicating highly coherent signals, good cancellation may be expected. Similarly if the coherence approaches zero, indicating essentially uncorrelated signals, no improvement will occur.

### 4.4.3 Case II. Signal components present in the reference input.

Consider fig. 4.11(b) which shows a two-input ANC as before, but also incorporating a pathway I(z) for the signal to appear at the reference input. The uncorrelated noises $u_j$ and $v_j$ will be taken to be identically zero for ease of analysis.

The unconstrained Wiener solution for this case may be expressed as

$$W^{o}(z) = \frac{\delta_{ss}(z).I(z^{-1}) + \delta_{nn}(z).H(z^{-1})}{\delta_{ss}(z).|I(z)|^2 + \delta_{nn}(z).|H(z)|^2} \qquad (4.56)$$

which converges to the solution $W^{o}(z) = 1 / H(z)$ when $I(z)$ approaches zero. It is important to know what cancellation performance may be achieved in general in order to know what levels of signal may be tolerated at the reference input, and what effect this has on the filter performance. To do this, it is necessary to find the signal and noise spectra at the ANC output. The signal spectrum is

$$(\delta_{ss}(z))_{out} = \delta_{ss}(z).| 1 - I(z).W^{o}(z) |^2 \qquad (4.57)$$

and that of the noise is

$$(\delta_{nn}(z))_{out} = \delta_{nn}(z).| 1 - H(z).W^{o}(z) |^2 \qquad (4.58)$$

These may be used to obtain the output SNR

$$\rho_{out}(z) = \frac{\delta_{nn}(z).|H(z)|^2}{\delta_{ss}(z).|I(z)|^2} \qquad (4.59)$$

Since the numerator and denominator may be identified as the noise and signal spectra at the reference input, this expression may be written

$$\rho_{out}(z) = \frac{1}{\rho_{ref}(z)} \qquad (4.60)$$

where $\rho_{ref}(z)$ is the signal to noise power density ratio (SNPDR) at the reference input. The output SNR is thus inversely proportional to the reference input SNR at all frequencies when convergence has been attained.

The extent to which the signal is distorted by partial cancellation may be determined as follows. The signal propagation path through the transversal filter is

$$- I(z).W^{o}(z) \tag{4.61}$$

When I(z) is small, this may be approximated by

$$- I(z) / H(z) \tag{4.62}$$

The spectrum of the signal propagated through the filter is thus

$$\delta_{ss}(z).| I(z)/H(z) | \tag{4.63}$$

The effect upon the primary signal of this component depends on the relative phases of these components. Maximum distortion will occur when these are completely out of phase. This worst case may be estimated by determining the ratio of the signal spectrum at the transversal filter output to that at the primary input. This may be denoted by the quantity D(z):

$$D(z) = \frac{\delta_{ss}(z).| I(z).W^{o}(z) |^2}{\delta_{ss}(z)} \tag{4.64}$$

$$= | I(z).W^{o}(z) |^2 \tag{4.65}$$

$$\simeq | I(z) / H(z) |^2 \tag{4.66}$$

It is convenient to express this in terms of the SNR at the primary and reference inputs:

$$D(z) \simeq \rho_{ref}(z) / \rho_{pri}(z) \tag{4.67}$$

indicating that low distortion occurs when the SNR is low at the reference input and high at the primary input. Equations (4.60) and (4.67) may be applied if a quantitative assessment of the effects of signal components in the reference input is required, though this may not be necessary if these are considered to be insignificant.

## 4.5 Previous Applications of Adaptive Filters to Physiological Signals.

The first application of ANC to biological signals was to the removal of mains interference in ECG signals at Stanford University in 1965 [92]. This was shown to be more effective than a notch filter as it can track the signal very closely, and possess an infinite null. This can be useful if it is known that the interfering signal drifts in frequency. When the interference is of the form of a sine wave and a reference signal exists, the filter structure can be simplified to a two-weight AF (two degrees of freedom being required to compensate for gain and phase fluctuations). When there are significant harmonics present it may be necessary to use an ATF. While this could be useful in EP methodology, it need not be required if adequate screening arrangements are employed.

Two other applications to ECG signals have been described. These are separation of donor and patient ECG signals in heart transplant operations [92] and separation of maternal and foetal ECG in surface recorded derivations [108]. In the former, an electrode inserted into the donor heart provides a reference to enable cancellation of the donor ECG during the operation. The foetal ECG application was first described by Widrow and later developed by Ferrara. The method requires selection of surface electrode sites that

will provide predominantly the MECG in one case (from chest sites), and both signals in the other (from abdominal sites). Good separation of these signals was demonstrated using a time-sequenced filter and multiple reference electrodes.

Kentie et al [109] attempted to apply Widrow's LMS ANC to electrogastrographic signals, but were unable to isolate a signal-free reference. They used a modified AF in which a band-reject filter is applied to a recorded signal to artificially produce a suitable reference signal. The attenuation of this component causes the filter to converge at the other frequencies resulting effectively in a band-pass filter. Adaption was halted after acceptable cancellation was obtained with minimum signal distortion, and the use of fixed weights after this eliminated gradient noise.

Though the results presented appeared to show that the method is effective, the methodological basis is questionable. In the first place the signal and noise spectra are not overlapping and sufficiently separated that digital filtering would seem to be entirely adequate. The authors state that conventional bandpass filtering was not acceptable because of amplitude and phase distortions, and the need to track time-varying components of the signal while being relatively insensitive to pulse-shaped motion artifacts. However the use of their filter subsequent to the adaption phase is essentially the same as a fixed FIR filter. The only advantage that may accrue is that the filter characteristics are dependent to some extent on the individual signal characteristics, but otherwise it seems unjustifiable.

The only application of AF to EP data known to the author is that of Carmon and Friedman [16] [110] who were interested in obtaining

optimally filtered estimates of single trial responses for pain studies. They used a template of the EP as the desired signal and the raw responses as their filter input. The aim of this approach was thus to form a time-varying Wiener filter to remove non-EP frequency components and pass only signal frequency components. The production of a noise-free EP template is therefore imperative, and they used W$^{O}$ody's latency correction technique on the raw data to obtain a good average EP. The AF was then repeatedly applied to the raw responses to filter out components not present in the template, though it was not clearly stated whether each response or the entire record was repeatedly filtered. Note that this is different from the approach considered in this thesis which attempts to cancel noise activity in a primary signal channel using correlated noise activity present in a reference channel.

While the results seemed initially impressive, at least for auditory and some somatosensory EPs, they were not entirely convincing for visual EPs. For instance there were peaks of the wrong polarity on occasion, and in general it seemed as though only the gross low frequency features were being passed. This may be due to excessive smoothing of the data, and is a particular hazard of the Wiener filtering approach which can yield distorted estimates of the signal while still satisfying the MSE criterion. This difficulty is discussed by de Weerd et al [113] who recommend that caution be exercised in using this approach. A criticism of this method is that it is not capable of reproducing detail in individual responses that has been averaged out in the template production, as the authors themselves concede, and so it must be used with caution in single trial studies. It is also not clear that an adaptive solution is necessary, since it may be more straightforward to compute the optimal least squares filter directly from the template and data samples.

## 4.6 Summary

In this chapter I have reviewed the concept of adaptive filtering with particular regard to FIR adaptive filters employing transversal structures. A brief derivation of the Widrow-Hoff LMS adaptive algorithm was outlined and the main analytic results governing its operation presented. In addition factors relevant to practical implementation of this filter, such as choice of parameter values and computational cost, were briefly discussed. These show that this algorithm is an attractive solution in many situations that require an adaptive approach, as it can be made unconditionally stable, has been widely studied and applied and is relatively simple to implement. Some of the recent trends in adaptive filtering were also briefly reviewed, including both modifications and developments of the basic LMS algorithm and more sophisticated techniques that attempt to compute the exact least squares solution at each sample moment. The latter are more demanding computationally and are not as intrinsically stable as the LMS algorithm, though they do offer significantly better adaption speed in certain circumstances. As it is not clear that any benefit would accrue when if these methods were applied to the signals considered in this thesis, the investigation was restricted to the use of gradient methods.

The use of AFs to perform noise cancelling has been described, and pertinent results presented that indicate the effectiveness of this method in different situations. Expressions were derived to predict the improvement in SNR when uncorrelated noise is present in either input, and these were related to the coherence function. Use of the coherence function results in a simpler and more useful expression than those previously published. Finally the use of AF in selected previous applications was reviewed. Little work has been done with EEG or EP

signals which form the main application in this thesis, and adaptive noise cancelling of EP signals is considered to be novel. The results of this research will be presented in chapters 6 and 7, but firstly methodological aspects of the experimental work will be described.

# CHAPTER FIVE

# THE GATED ADAPTIVE FILTER.

## 5.0 Introduction.

The previous chapter reviewed the concepts and theoretical aspects of adaptive filtering with particular reference to adaptive transversal filters (ATF) employing the Widrow-Hoff adaptive algorithm. Some important developments of the simple ATF were introduced, one of these being the time-sequenced adaptive filter (TSAF). In this chapter a variant of the basic AF is introduced based on similar notions to the TSAF which will be referred to as the gated adaptive filter (GAF). The principal feature of this filter is that the weight adaption is suspended during peak signal activity, as the presence of a signal transient can disturb the adaptive modelling of the cross-channel transfer function in noise cancelling applications. It will be shown in this chapter that this new filter structure is more appropriate than the unmodified adaptive filter or the time-sequenced adaptive filter when transient signals in additive noise are to be filtered, especially when fast adaption is desired. The reasons for this will be discussed and some theoretical aspects governing the gated filter approach will be considered. A comparison between the GAF and other adaptive filters is described based on a computer study using artificially generated signals. The results of this study support the hypothesis that the GAF gives superior performance than either the AF or the TSAF in these circumstances.

## 5.1 The Gated Adaptive Filter: Concepts

This filter structure is based on similar premises to those of the TSAF described in chapter 4, but is different in some important respects. Briefly, the TSAF was proposed to approximate an optimal time-varying filter when the data are not stationary, but can be considered as piece-wise stationary. The TSAF employs a set of filter weights corresponding to each stationary segment. These are selected by an external 'sequencing' signal, so that only one weight set is used at any time to generate the filter output and undergo adaption. Each weight set thus converges to the optimal solution for each discrete stationary segment. A block diagram of the TSAF is presented in fig. 4.8.

Gated adaptive filtering also employs an external signal synchronous with transient events in the input data. It differs from the TSAF in that only one weight set is employed, but the external signal controls the manner in which the filter is applied, rather than controlling the selection of the different weight sets. The control signal can, for example, change the convergence rate or even temporarily suspend adaption, during the applied gate period. This is shown diagrammatically in fig. 5.3 which compares the weight revision and selection operations for a TSAF employing two weight sets, and a GAF employing a single weight set whose adaption is suspended during signal epochs.

Consider fig. 4.11(a) which shows a common situation where the GAF may prove more effective than the ordinary AF, as described in section 4.4.2. A primary signal channel contains a periodically occurring signal $s_j$ corrupted by noise signals $n_{1j}$ and $u_j$. A reference

signal channel is also available which contains a noise signal $v_j + n_{2j}$. $n_{2j}$ is correlated with $n_{1j}$ such that

$$n_{1j} = n_{2j} * h_k \tag{5.1}$$

where $h_k$ represents a time-invariant discrete linear process. $u_j$ and $v_j$ represent noise signals uncorrelated with each other and with each of $s_j$ or $n_j$. $n_{1j}$, $n_{2j}$, $u_j$ and $v_j$ are assumed to be stationary, zero mean processes and $s_j$ is assumed to have zero mean. In chapter 4 various expressions were derived which describe the expected results of adaptive noise cancelling in this situation. These are based on the premise that the input signals are themselves uncorrelated, i.e. they have zero autocorrelation other than for zero lag. In practice random signals have some degree of colouration, represented by finite values of the autocorrelation function at non-zero lags. Studies suggest that even high values of autocorrelation do not significantly affect the theoretical results that are based on this assumption other than to decrease the convergence time of the filter [94]. This is true when the chosen adaption rate is relatively slow, but may no longer be true when adaption rate is relatively high or when the signal contains components whose energy is large compared with the noise. The reason for this is as follows. Let us assume that the filter has essentially converged to the Wiener solution. The error signal fluctuates randomly about zero due to random gradient estimation noise. A small value of the convergence coefficient corresponds to a long integration time in the error feedback loop and ensures that these fluctuations do not have much effect on the overall converged state, and hence yields a small misadjustment noise in the filter output. During the signal epoch, however, the signal itself appears in the error output when the filter is used in noise canceller mode. If the adaptive time constant is much larger than the signal duration the effect of the additional signal upon

the weights is negligible as the signal is assumed to have zero mean in this analysis. However if the adaption rate is fast it is possible for the weights to be affected by the added signal and result in divergence from the optimal solution during the signal epoch. Whether or not this will significantly affect filter accuracy is dependent on:

(i) the signal amplitude and waveform
(ii) the value of the convergence parameter $\mu$, and
(iii) the mean noise power.

This may be illustrated by means of a simple example. Consider the previously described situation and let $s_j$ represent a pulse with amplitude $\alpha$ and duration $\tau$ occurring with repetition period T. Let $n_j$ be a white noise process with variance $\sigma^2$. Let $h_k$ be unity for k=0 and zero otherwise, such that $n_{1j}$ equals $n_{2j}$, and let $u_j$ and $v_j$ be zero.

Optimal filtering will clearly take place when $W(z) = 1/H(z)$, i.e. $W(z) = 1$. This is the solution predicted by Wiener's theory, and given by equation (4.53). It will not however be attained in practice by an adaptive filter of the form considered up till now. Assuming that the filter had converged to this solution prior to the occurrence of a transient signal event, the ANC output $e_j$ would be zero, thus ensuring that no further weight revision takes place. The occurrence of the signal at time j immediately alters the weight vector by adding the term $2\mu\alpha\underline{X}_j$ to the weight vector, according to the LMS update algorithm

$$\underline{W}_{j+1} = \underline{W}_j + 2\mu\alpha\underline{X}_j \qquad (5.2)$$

This disturbance from the optimal solution results in an additional error at the next sample moment, and the accumulation of these errors leads to divergence from the optimal solution. When the signal

transient dies away the filter relaxes back to the previously attained optimal solution. The effect of the additional signal upon the filter will clearly be greater when the signal amplitude is high, when the duration of the signal is a significant fraction of the repetition period and when $\mu$ is relatively large. It will also be dependent on the signal waveform as high frequency components in the error feedback loop are likely to have less cumulative effect than low frequency components. The ensuing filter inaccuracies result in less effective cancellation and introduce signal distortion. The degree to which this can be tolerated will depend on other contributions to filter inaccuracy, such as misadjustment noise.

While it would be informative to develop this analysis in detail in order to determine how each of these factors affects filter performance, this has not been pursued here. Merely characterising the resulting errors does not necessarily help overcome the problem, and in addition it is difficult to perform this analysis as it will be dependent upon the actual signal statistics. Empirical studies were therefore undertaken to identify the conditions when these effects are likely to be significant in practice.

The obvious way to overcome this difficulty is to 'switch the filter off' during the signal epochs using a synchronous gating signal. This effectively 'freezes' the weight update procedure during the signal epochs and permits optimum cancellation to continue, provided that the noise statistics are stationary during this interval. This may be stated mathematically as follows:

$$y_j = \underline{X}_j \cdot \underline{W}_j$$

$$\underline{W}_{j+1} = \begin{cases} \underline{W}_j + 2\mu e_j \underline{X}_j & ; g_j = 1 \\ \underline{W}_j & ; g_j = 0 \end{cases} \tag{5.3}$$

where $g_j$ is a gate signal that arbitrarily takes the value of 0 or 1 to control filter adaption.

This suggestion introduces a number of benefits. Firstly it is relatively simple to implement when an external gating signal is available (such as the stimulus trigger in EP experiments). It even introduces slight computational advantages through elimination of the weight update calculations during the signal epochs. (This is of little benefit to real-time filters which must have the capacity to perform filtering at the full data rate, but does yield a slight reduction in off-line processing time.) Secondly the theoretical analysis is simplified, as there is no need to take the signal into account since the optimal solution is based purely on the statistics of the inter-signal epochs. This does not significantly affect the main results of chapter 4 except as follows:

a) while the convergence time remains the same during those portions of the data when adaption takes place, the periods when adaption is suspended obviously result in a net increase in convergence time.

b) the theory in chapter 4 was developed in terms of a zero-mean random signal. Furthermore the assumption was made that non-white noise processes do not significantly affect the results. While this is a reasonable assumption to make in many cases, later experimental results show that it is not valid when the signal is of large amplitude or when the convergence rate is high. The GAF on the other hand allows these results to be applicable in both of these cases and for any conceivable transient signal, provided that the moment of occurrence and duration of

the signal are accurately known and that the signal is essentially zero outside this interval. Steady state convergence achieved during inter-signal epochs will not be influenced by the occurrence of the signal, thus permitting optimal cancellation to take place at all times in the signal record.

c) The simplifications are most apparent in section 4.4.3, which considers adaptive noise cancelling when signal components are present in the reference input. Noise cancellation and distortion expressions were derived in that case under the assumption that there is little signal present in the reference, so that the optimal filter $W^{o}(z)$ is approximately equal to $1/H(z)$. By setting $W^{o}(z)$ equal to $1/H(z)$ in the analysis, the resulting simplifications permitted approximate expressions for the signal distortion to be readily obtained in equations (4.62), (4.63), (4.66) and (4.67). In the case of the GAF these expressions are no longer approximate but exact, as $W^{o}(z)$ is exactly equal to $1/H(z)$. Furthermore the analysis of this situation assumed that uncorrelated noises $u_j$ and $v_j$ were zero for simplification. This analysis can be extended to include the effects of these when using the GAF by substituting the expressions developed in section 4.4.2 for the optimal filter solution (4.48),(4.53) into the expressions in section 4.4.3 for the output SNR (4.57), (4.58). The resulting analysis is now tractable but yields expressions which are rather cumbersome and which do not shed further light on the adaption process compared with the simple expressions derived in section 4.4.3. They are therefore not included in this thesis.

The main disadvantage of the GAF relates to the case when $H(z)$ or the noise statistics are not time-invariant, so that maintaining the weights constant during the signal epoch no longer results in optimum

cancellation. The filter is not able to track the non-stationary data characteristics during the signal epoch and may diverge from the optimum solution. The GAF may not be appropriate in this case, though if the non-stationarity is a relatively slow process it is likely that the resulting errors may well be less than those which are caused by adaption of the signal transient. In this situation it may well be better to use a slow rate adaption rate to accommodate the changing nature of the signal statistics than to suspend adaption altogether. It is probable that optimum values exist for $\mu$ in the two intervals that will balance the errors arising from mis-tracking of the nonstationarity with those arising from signal adaption, though this has not been analysed here, and is left to a future investigation.

Note that the TSAF does not offer any substantial advantage over the normal AF in the case of fast adaption of high amplitude transient signals as it suffers from the same defects as the normal AF. While these effects can be reduced by employing small values of the convergence parameter during the signal epoch this results in much larger overall convergence times. The signal and non-signal epochs are filtered independently by the TSAF and hence convergence proceeds at different rates. Transient artifacts can also arise in the filter output when the alternate weight sets are selected if they have markedly different character. More importantly, no account is taken of signal characteristics in the interval when the alternate weight set is selected. This means that only slowly occurring non-stationarities can be adequately tracked. This point was raised by Ferrara [102] who suggested that research was needed to identify a technique which would take into account immediately preceding data characteristics. It appears that the GAF achieves this for the simple situation of repetitive transient signals analysed in this chapter. These assertions

will now be explored by means of a computer study using artificially generated data.

## 5.2 Experimental verification of the GAF.

In this section a study undertaken using computer methods is described to provide experimental verification of the assertions made concerning the proposed GAF structure. This will be used to aid comparison between this and other AF structures and to provide empirical results that will be useful in assessing the value of the modified filter. The experimental strategy is first described and will then be followed by a a discussion of the results obtained. This study was performed using the signal processing facility developed in the course of this research project, described in the following chapter.

### 5.2.1 Computer methods.

In this study it was desired to perform a quantitative comparison of the GAF with the basic AF and TSAF, using a signal buried in noise as the primary input and a correlated noise channel as the reference input. A simple signal was used, similar in form to several EPs, and a range of noise data were generated having different correlation properties. These permitted results to be obtained for a range of SNR and coherence conditions in order to assess how appropriate each filter method is for each different condition. The analysis was performed using two levels of signal amplitude and two adaption speeds as it has been previously suggested that these parameters can significantly affect normal adaptive filter operation. It was the intention that this study would provide experimental evidence regarding the use of the GAF for general applications, and hence simple forms of

signals were used rather than an accurate simulation of EP and EEG data.

An arbitrary signal shape was used, chosen to be identically zero for half of its duration, corresponding to the relative absence of signal activity in the interstimulus interval in EP data. An experimental record consisted of 64 identical signals in wideband noise, with an associated reference channel of equal length containing partially correlated wideband noise. Each signal was defined as follows over a discrete interval of M samples, with M=64 in this case:

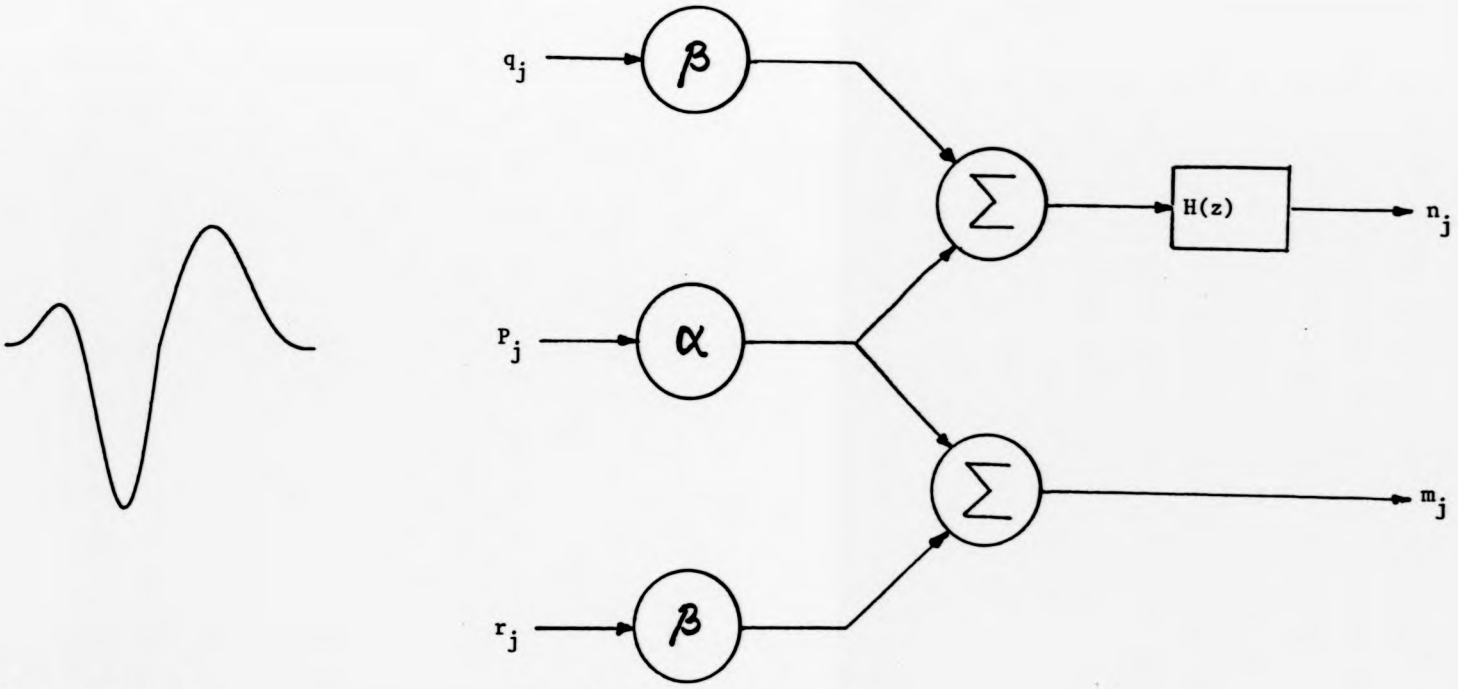
$$s_j = \begin{cases} \sin ( 8\pi j/M ) \cdot ( 1 - \cos ( 4\pi j/M ) )/2 & ; \quad 0 < j < M/4 \\ \sin ( 4\pi j/M ) \cdot \cos ( 4\pi j/M ) / 2 & ; \quad M/4 < j < M/2 \\ 0 & ; \quad M/2 < j < M \end{cases}$$

Eqn. (5.4)

This corresponds to one cycle of a sine wave of period M/4 followed by half a sine cycle of period M/2, with a cosine envelope applied to smooth out fast transitions at the start and end of the signal transient. The signal is depicted in fig. 5.1(a).

A number of artificial noise records were generated and stored in files for later use. It was considered desirable to generate approximately gaussian white noise in pairs of records comprising a range of coherence values. This would be more useful to characterise these filters in general, though the results would be expected to be relevant to EEG signals also. Specific consideration of these different filter methods to EEG data is made in chapter 8.

The procedure used to generate the noise records is now described. Three noise records comprising 64 epochs of 64 samples were separately generated by adding 12 consecutive random numbers obtained using the pseudo-random number generator function RAN supplied with DEC




a) Test Signal

b) Noise Generation Diagram

FIGURE 5.1 Details of the artificial signals used in the computer study of different adaptive filters.
a) The test signal
b) Diagram illustrating the generation of noise signals $m_j$ and $n_j$ from 3 independent random noise processes, $P_j$, $q_j$ and $r_j$. The degree of correlation between $m_j$ and $n_j$ is determined by the coefficients $\alpha$ and $\beta$. H(z) is a linear transfer function relating to $m_j$ and $n_j$.

Fortran IV [114]. This method produces approximately gaussian random noise from a uniformly distributed random sequence. An examination of the spectral characteristics showed that they were approximately white and uncorrelated with each other. Each record had zero mean and identical variance. These will be denoted by $p_j$, $q_j$ and $r_j$. Pairs of composite noise records were produced by linearly combining these, so as to have specific spectral characteristics. These are denoted by $m_j$ and $n_j$, and were generated as follows:

$$m_j = \alpha\, p_j + \beta\, r_j$$
$$n_j = (\alpha\, p_j + \beta\, q_j) * h_k \qquad (5.5)$$

where $h_k$ represents the impulse response of a linear transfer function H(z) defined later. Fig. 5.1(b) shows the noise generation procedure diagrammatically. The coefficients $\alpha$ and $\beta$ were selected from Table 5.1 to generate noise having the desired coherence, and were chosen so that the variance of the noise sequences $m_j$ and $n_j$ equalled that of the original noise records $p_j$, $q_j$ and $r_j$. The equations describing these coefficients will be developed later. A simple form for H(z) was desired so that the spectra are similar in both channels, but that more than simple subtraction is required to effect cancellation. The following response was arbitrarily chosen

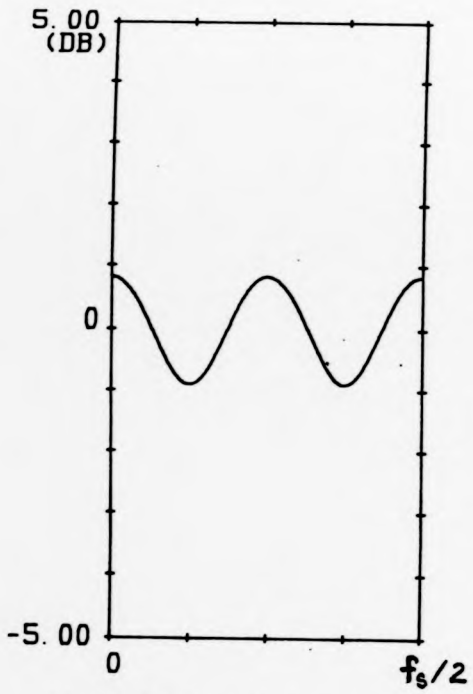
$$H(z) = 1 + 0.1\,(z^{-4}) \qquad (5.6)$$

whose transfer function is shown in fig. 5.2.

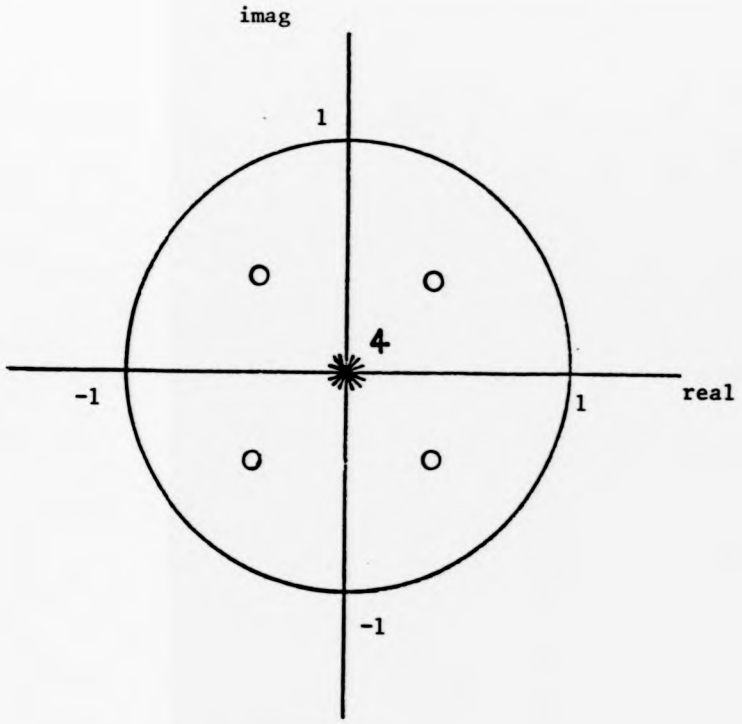
To ensure that the variance of the composite records match those of the original records, the following relation must be satisfied:

$$(\alpha^2 + \beta^2) = 1 \qquad (5.7)$$

This is easily shown by the following argument. Let the variances of




a) Magnitude response




b) Z-plane diagram

FIGURE 5.2 a) Magnitude response and b) pole-zero diagram of transfer function $H(z) = 1 + 0.1z^{-4}$ relating correlated noise signals.

the sequences $p_j$, $q_j$ and $r_j$ be equal to $\sigma^2$. Then the variances of the noise sequences $m_j$ and $n_j$ are

$$\begin{aligned}\sigma_m^2 &= E[\, m_j^2 \,] \\ &= E[\, ( \alpha p_j + \beta q_j )^2 \,] \\ &= \alpha^2 . E[\, p_j^2 \,] + \beta^2 . E[\, q_j^2 \,] + 2\alpha\beta . E[\, p_j . q_j \,] \\ &= ( \alpha^2 + \beta^2 ) \; \sigma 2 \end{aligned} \tag{5.8}$$

since $p_j$, $q_j$ and $r_j$ are independent noise processes. Similarly

$$\sigma_n^2 = ( \alpha^2 + \beta^2 ) \; \sigma^2 \tag{5.9}$$

Expressions for $\alpha$ and $\beta$ can be derived in terms of the coherence function, which is defined by

$$\begin{aligned}|\gamma(f)|^2 &= \frac{|G_{mn}(f)|^2}{G_{mm}(f) . G_{nn}(f)} \\ &= \frac{| \alpha^2 . \sigma^2 |^2}{( \beta^2 + \alpha^2 ) . \sigma^2 . ( \beta^2 + \alpha^2 ) . \sigma^2} \\ &= \frac{1}{( 1 + \beta^2/\alpha^2 )^2} \end{aligned} \tag{5.10}$$

where $G_{mn}(f)$, $G_{mm}(f)$ and $G_{nn}(f)$ are the cross-spectral and autospectral intensities of the noise signals $m_j$ and $n_j$. Thus

$$\begin{aligned}|\gamma| &= ( 1 + \beta^2/\alpha^2 )^{-1} \\ &= ( 1 + \beta^2/1-\beta^2 )^{-1} \\ &= 1 - \beta^2 \end{aligned} \tag{5.11}$$

Hence

$$\beta = \sqrt{(1 - |\gamma|)} \tag{5.12}$$

and

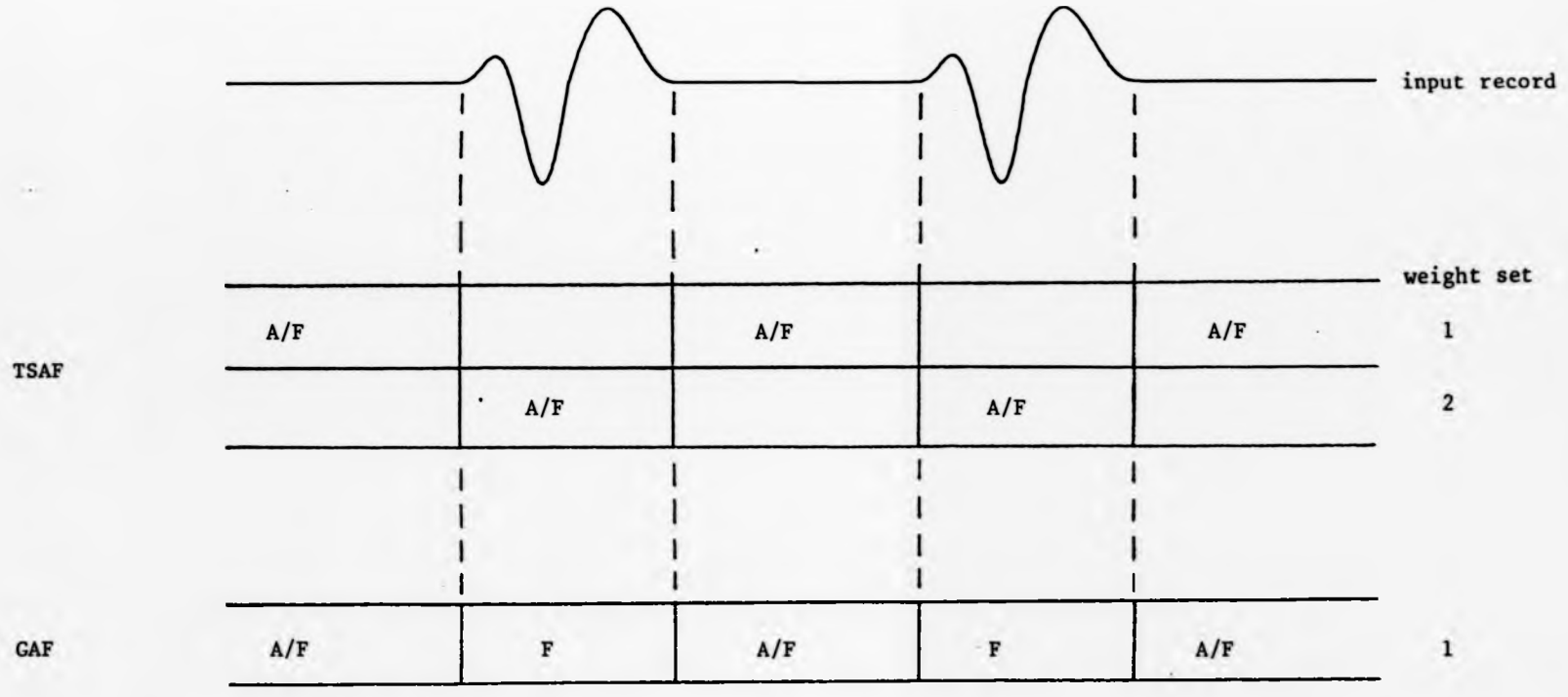
$$\alpha = \sqrt{|\gamma|} \tag{5.13}$$

These expressions were used to generate the values in Table 5.1

The noise records were used to examine the performance of each filter as follows. 64 identical signal records were added to each noise record of $m_j$ to form the primary signal channel, and $n_j$ formed the reference channel. The data were filtered by each filter i.e. the GAF, AF and TSAF. In the AF one set of weights was used throughout for adaptive filtering. In the GAF one set was used, but adaption was halted during the initial half of each signal epoch. In the TSAF two weight sets were used, corresponding to the signal and non-signal epochs. In each case a filter order of 9 was used which is sufficient to completely characterise the channel transfer function H(z). This is shown diagrammatically in fig. 5.3. (Note that in practice the weight set used for the GAF is identical to that used by the TSAF during the non-signal epoch, which saves unnecessary computation.) The mean square errors (MSE) of each filter output were computed for both signal and non-signal epochs to aid quantitative assessment. This procedure was repeated for each of the following coherence values: 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.3, 0.1. In addition results were obtained using two different levels of SNR, approximately 5 and 15 dB. Each filter was also applied using two values of $\mu'$, equal to 0.2 and 0.02, to permit the effect of adaption speed to be examined.

### 5.2.2 Results of the computer study.

Figs. 5.4 - 5.6 show typical filter outputs for selected cases. 7 records are shown in each plot, the first being the unfiltered primary input, followed by two groups of 3 records corresponding to the AF, GAF and TSAF outputs used with high and low adaption speeds. Figs. 5.7 and




A - adapt weight set

F - apply filter using weight set

FIGURE 5.3 Diagram illustrating the weight selection and update procedures for the Time-sequenced (TSAF) and Gated (GAF) adaptive filters. Whereas the TSAF employs two (or more) weight sets and selects these corresponding to different epochs in the input record, the GAF employs one weight set, but controls the manner in which it is used.

5.8 show averages of 8 epochs for each filter, adaption rate and coherence value for both SNR levels. In each case the dotted curve superimposed on the full traces is the original signal. Note that only the GAF yields perfect cancellation regardless of SNR or adaption rate when used with the high coherence data $|\gamma|^2 = 1.0$ The TSAF performs equivalently to the GAF during inter-signal epochs in this case, but is not as effective in removing noise in the signal epoch. It also introduces transient step artifacts when the different weights are selected, and the individual filtered signals are more variable than those in the GAF or AF outputs. This is particularly evident in fig. 5.4 and 5.6. To aid a quantitative comparison, the MSE during signal epochs is plotted against coherence for each condition in fig. 5.9. A normalised figure is obtained by scaling the MSE with respect to the original noise power. Thus a value of 0.5 represents a 3 dB noise reduction factor.

The results confirm the suggestion made earlier in this chapter that repetitive transient signals embedded in stationary noise can be filtered better using a gated adaptive filter, than by either of the simple AF or TSAF when fast adaption is required and especially when the signal energy is relatively high. The improvement in SNR during the signal epoch can be very significant in this circumstance, as evidenced by the MSE plots in fig. 5.9, though the GAF performed consistently better than the AF and TSAF under all the conditions of SNR, adaption rate and coherence considered. Only when low SNR data were filtered using a slow adaption speed was the improvement in performance insignificant. When fast adaption rate is employed, both the simple AF and TSAF appear to be inappropriate, as the normalised MSE figures are in considerable excess of unity in this case indicating that the filters introduce substantial noise and degrade the signal quality. This is


FIGURE 5.4 Comparison of filter outputs for the basic adaptive filter (AF), gated (GAF) and time-sequenced (TSAF) adaptive filters using artificial data. SNR = 5dB. Two adaption speeds compared. $|\gamma|^2 = 1.0$.

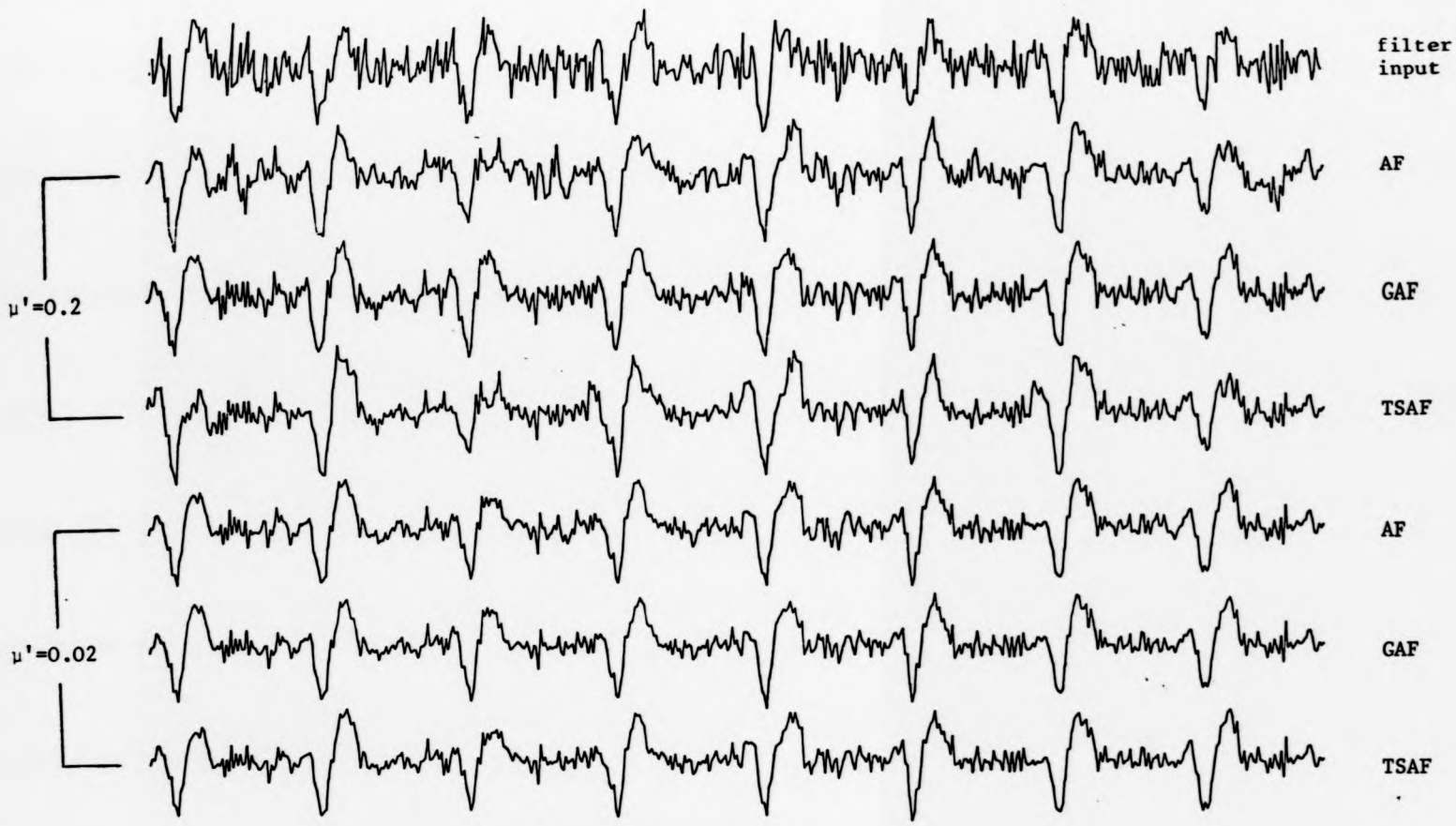



FIGURE 5.5 Comparison of three adaptive filters (AF, GAF and TSAF) using low SNR artificial data (SNR = 5dB) and moderate coherence levels ($|\gamma|^2 = 0.8$). Two adaption speeds employed.

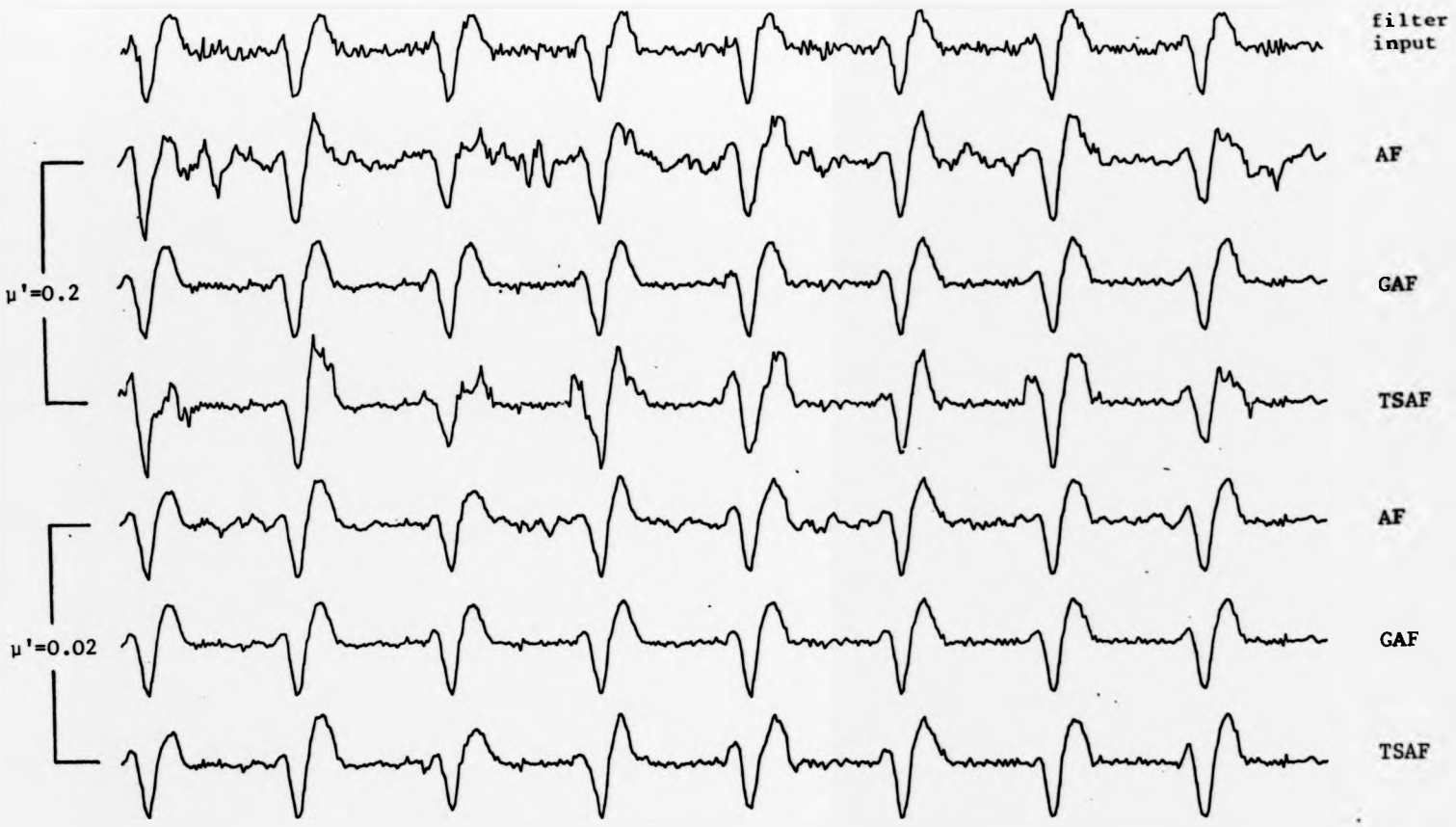



FIGURE 5.6 Comparison of three adaptive filters (AF, GAF and TSAF) using moderate SNR artificial data (SNR = 15dB) and moderate coherence levels ($|\gamma|^2 = 0.8$). Two adaption speeds employed.

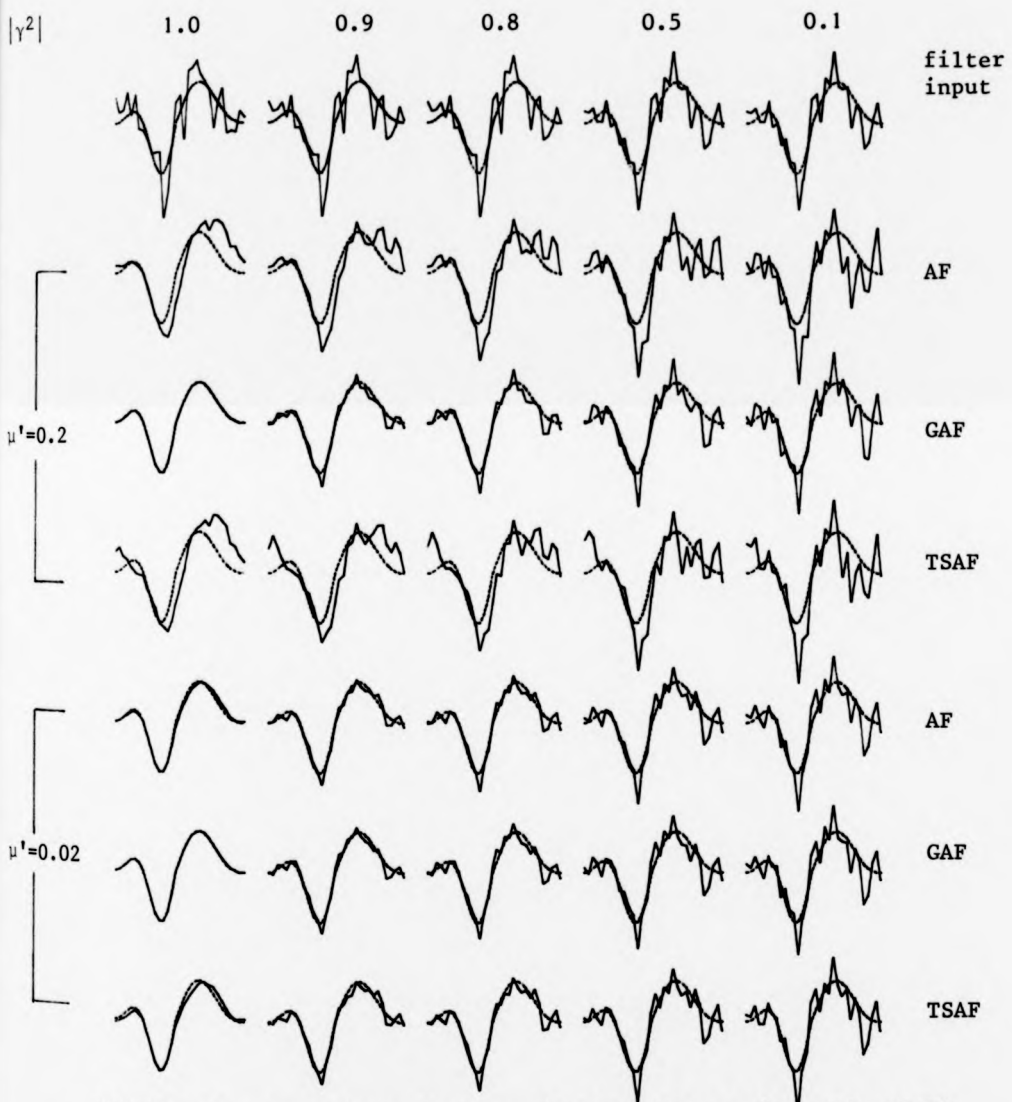



FIGURE 5.7 Comparison of three adaptive filters (AF, GAF, TSAF) using low SNR artificial data (SNR = 5dB), two adaption speeds and a range of coherence values. Averages of 8 records shown in each case.

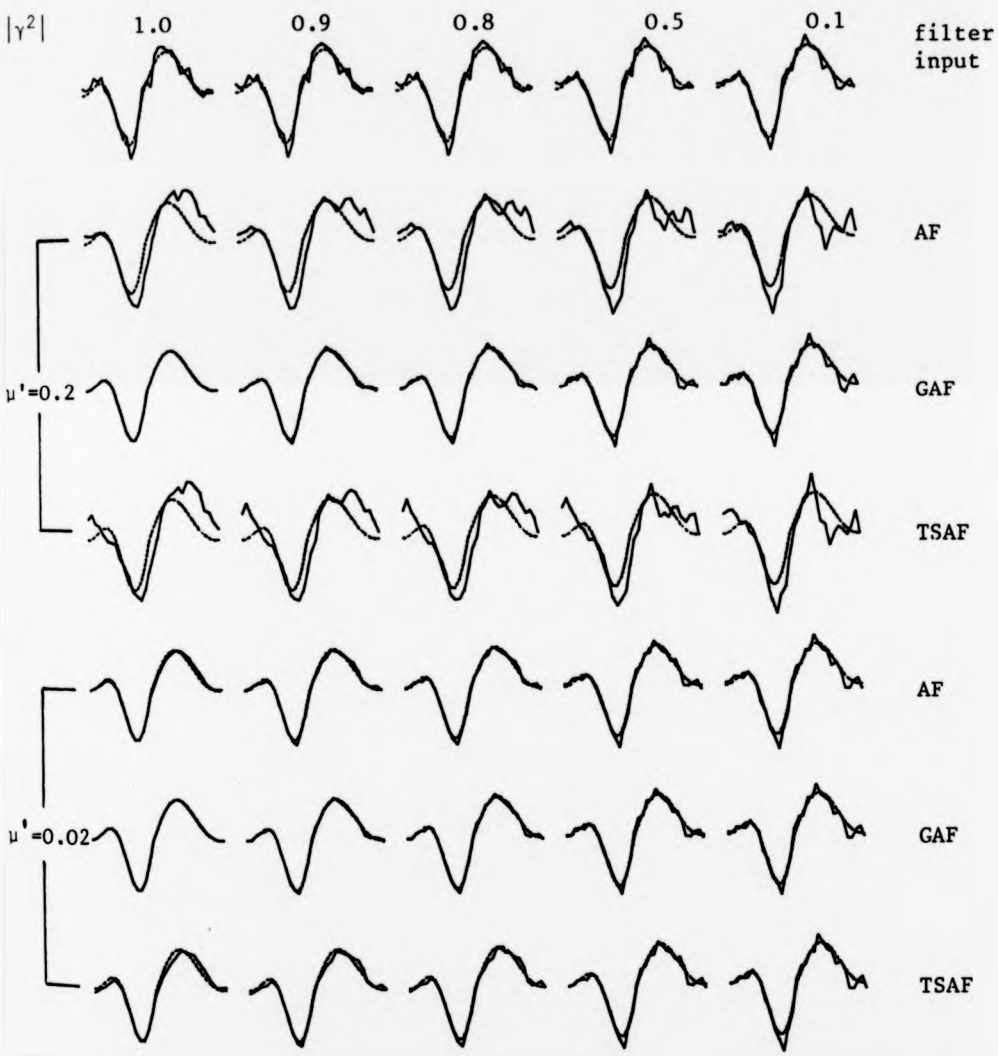



FIGURE 5.8 Comparison of three adaptive filters (AF, GAF and TSAF) using high SNR artifical data (SNR = 15dB), two adaption speeds and a range of coherence values. Averages of 8 records shown in each case.

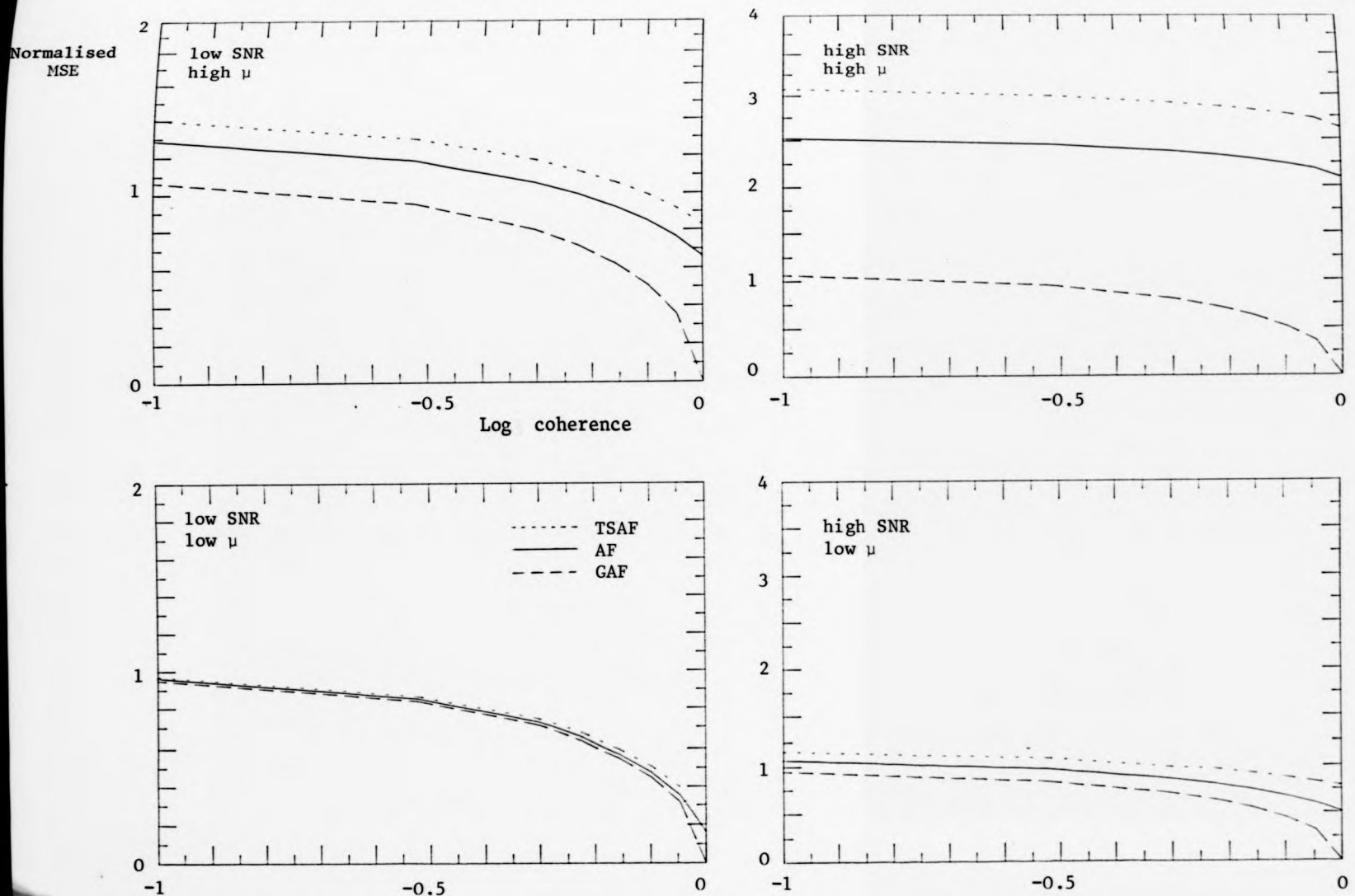



FIGURE 5.9 Normalised MSE vs log coherence for AF, GAF and TSAF filters

| $\|\gamma\|^2$ | $\|\gamma\|$ | $\alpha$ | $\beta$ |
|---|---|---|---|
| 0.0 | 0.0000 | 0.0000 | 1.0000 |
| 0.1 | 0.3162 | 0.5623 | 0.8269 |
| 0.2 | 0.4472 | 0.6687 | 0.7435 |
| 0.3 | 0.5477 | 0.7401 | 0.6725 |
| 0.4 | 0.6325 | 0.7953 | 0.6063 |
| 0.5 | 0.7071 | 0.8409 | 0.5412 |
| 0.6 | 0.7746 | 0.8801 | 0.4748 |
| 0.7 | 0.8367 | 0.9147 | 0.4042 |
| 0.8 | 0.8944 | 0.9457 | 0.3249 |
| 0.9 | 0.9487 | 0.9740 | 0.2265 |
| 1.0 | 1.0000 | 1.0000 | 0.0000 |

TABLE 5.1 Table of coefficient values for $\alpha$ and $\beta$ used to generate artificial random noise with the specified coherence values.

almost certainly due to misadjustment errors caused by adaption during the signal epoch. In contrast the MSE in the GAF records rarely exceeds unity and consistently decreases to zero as the coherence approaches unity. Perfect cancellation can thus be achieved in this circumstance, but is not achieved by either of the AF or TSAF. These results are also confirmed by visual examination of the individual filtered records and their averages. Examination of these suggests that both bias and variance errors are present. The bias is most clearly seen in the averages of the AF and TSAF outputs in figs. 5.7 and 5.8, which show a larger deviation of the signal estimate from the underlying signal. The individual responses appear to be more uniform in the case of the GAF filtered records than the others which supports the assertion that the variability in the AF and TSAF records is due to incorrectly adjusted filter weights caused by adaption during the signal epoch.

## 5.3 Summary.

A modification to the basic adaptive transversal filter has been proposed, which is based on similar premises to the time-sequenced adaptive filter. The modified filter, known as a gated adaptive filter, uses a synchronous gating signal to suspend adaption (or otherwise alter filter operation) when the presence of high amplitude signal components would otherwise cause significant disturbance of the filter weights and hence result in poor filter performance. It is shown that the GAF is more appropriate than the basic AF or TSAF for filtering repetitive transient signals when high adaption rates are employed and particularly when high energy signal components are present.

An experimental investigation of these proposals was performed using computer-generated data of specified characteristics. A wide

range of conditions were analysed, and encompass high to low correlations between the noise records, two levels of SNR and two adaption rates. The results of the study confirm the proposal that the GAF yields superior cancellation in each of these cases, compared with the unmodified AF and the TSAF. This advantage is most significant when the coherence, SNR and adaption rate are high, though consistently better results are achieved for all values of these parameters considered. The GAF was specifically developed to perform fast convergence adaptive filtering of EP data using the LMS algorithm, but these results have been obtained using a general signal waveform in wideband noise, and so are likely to be applicable to a wide range of signals. In addition the basic concepts are independent of the actual filter strategy, and apply to noise cancelling in general, regardless of the filter structure or adaptive algorithm used. A similar improvement in filter performance can therefore be expected when lattice adaptive filters or the recently developed fast convergence algorithms are employed. Before describing the main experimental work of this thesis the next chapter will detail the methodology employed in recording, analysing and processing EP data. This will form the experimental basis for the following chapters which are concerned with the use of adaptive noise cancelling in processing EP signals.

# CHAPTER SIX

# INSTRUMENTATION AND EXPERIMENTAL METHODOLOGY.

## 6.0 Introduction.

This chapter describes the experimental methodology applied to record EP signals and to perform signal analysis and processing using a laboratory minicomputer. The discussion is presented in three sections: experimental details of EP recording, followed by separate discussions of hardware and software aspects of the data analysis. In the first section a number of aspects are considered, such as the method of presenting visual stimuli, the geometry and timing of the stimulus, and the choice of electrode montage and its relation to the stimulus employed. The instrumentation required to record signals and the experimental procedure followed is described. The second section discusses the computing options available, and describes the minicomputer system established to perform this research. The last section is concerned with the software facility which was developed in order to carry out this work. Some of the possible strategies are discussed and existing software packages are surveyed. A preliminary review showed that no suitable computer program was available for this work and hence required that such a program be developed. The main strategies and design goals are described, and a brief account of the final solutions presented. This provides the necessary background to the experiments described in the following chapters of this thesis.

## 6.1 Creation of experimental data base.

### 6.1.1 Stimulus and electrode details.

The experiments undertaken in the following chapters were intended to establish whether improved estimation of VEP signals can be obtained through cancellation of activity recorded at other scalp sites, and whether adaptive noise cancelling is suitable for this purpose. To accomplish this it was necessary to create a suitable data base of EP records with which to test these propositions. A discussion of the nature of this data base and the manner in which it was created is now presented. It was clearly necessary to generate a set of EP records representing a range of typical responses from a number of scalp sites. In order to reduce the risk of obtaining results too closely tied to a particular exoeriment, data were obtained using a variety of stimuli from two normal adults. As this study was exploratory in nature this was considered adequate to obtain preliminary results regarding this approach. A much larger pool of subjects would have been required had the intention been to demonstrate the widespread validity and reliability of a particular method.

It was considered important to generate large amplitude EPs as this facilitates analysis of different processing methods. Patterned visual stimuli were used as they are known to produce large amplitude VEPs with consistent properties and for this reason are used routinely in clinical and research work [3]. A pattern of high contrast isolated square elements is particularly effective and was therefore chosen for these studies [49]. A description of the stimulus is given later. Although VEP amplitudes are generally greater when a large retinal area is stimulated, full-field stimulation can sometimes lead to reduced

amplitudes as a result of cancellation of components arising in different cortical regions [49]. For this reason half-field stimulation was employed, using either the left, right or lower half-fields. These arrangements produce EPs whose scalp distributions reflect synchronous activity in only one hemisphere or in both. If only one hemisphere is stimulated, the evoked potentials will predominate on one side of the scalp, but if both hemispheres are stimulated, EP components are likely to arise on both sides of the scalp. This enables both of these important cases to be analysed. These arrangements have been used extensively in the past in the Department of Communication and Neuroscience at Keele.

The specific stimulation and recording details for each subject are now described. To permit statistically reliable results to be obtained, a large number of responses were obtained to each experimental condition. Stimuli were presented in experimental runs of 16 presentations. 16 such runs were obtained for each stimulus condition, resulting in a total of 256 responses to each stimulus. 16 trials were found to be sufficient to obtain a reasonably noise-free average EP and avoids unduly prolonging the experiment. A fairly fast stimulation rate (approximately 1 Hz) was employed and so each run did not last much more than 16 s. This reduces the possibility of artifacts due to eyeblinks, coughing or other disturbances. As about 10 different stimuli were used for each subject, a sampling rate of 128 Hz implies a total of more than 4 million data samples for each subject, and hence necessitated the use of magnetic tape storage of the raw data. This symmetric arrangement of 16 runs and 16 stimulus presentations permits studies to be made in terms of either stimulus or experiment rank order, thereby enabling the extent of habituation to be monitored if required.

The stimulus was a high contrast pattern of regularly arranged elements, forming a checquerboard or isolated square pattern. The element size was typically 1° in a field of about 20° of visual subtense. These were presented binocularly for 25 ms, at a rate of approximately 1 Hz. This is illustrated in fig. 6.1. The use of short duration presentations has been shown to minimise adaptation effects to repetitive stimuli, permitting faster stimulation rates and hence shorter experiments [49]. Aperiodic stimulation was routinely employed as it can prevent synchrony of periodic interference components, such as alpha or a.c. mains activity. This was accomplished by inserting a random delay into the stimulation period. The advantages of aperiodic stimulation rate have been discussed previously in section 3.2.3. The stimulation details were similar for each subject though they varied slightly. They were based on previous mapping studies of the chosen subjects which had indicated whether left or right half-field stimuli were more effective in generating large amplitude EPs. The order of the stimuli and experimental conditions for each subject are shown in Table 6.1 and 6.2.

An identical electrode montage was employed for each subject. A limited number of scalp sites were available for this study in view of equipment constraints and the considerable amount of computer processing involved; the investigation was therefore restricted to a single row of electrodes transversely located over the occipital scalp. A longitudinal row was rejected as it does not permit cross-hemispheric differences elicited by the different stimuli to be examined. It is generally known that the EP waveform does not vary significantly over distances of less than 2-3 cm [49] and so to obtain reasonable coverage of the transverse scalp distribution 7 electrodes were used having a separation of 2.5 cm, on an arc 4 cm above the inion (fig. 6.1). EEG

The stimulus was a high contrast pattern of regularly arranged elements, forming a checquerboard or isolated square pattern. The element size was typically 1° in a field of about 20° of visual subtense. These were presented binocularly for 25 ms, at a rate of approximately 1 Hz. This is illustrated in fig. 6.1. The use of short duration presentations has been shown to minimise adaptation effects to repetitive stimuli, permitting faster stimulation rates and hence shorter experiments [49]. Aperiodic stimulation was routinely employed as it can prevent synchrony of periodic interference components, such as alpha or a.c. mains activity. This was accomplished by inserting a random delay into the stimulation period. The advantages of aperiodic stimulation rate have been discussed previously in section 3.2.3. The stimulation details were similar for each subject though they varied slightly. They were based on previous mapping studies of the chosen subjects which had indicated whether left or right half-field stimuli were more effective in generating large amplitude EPs. The order of the stimuli and experimental conditions for each subject are shown in Table 6.1 and 6.2.

An identical electrode montage was employed for each subject. A limited number of scalp sites were available for this study in view of equipment constraints and the considerable amount of computer processing involved; the investigation was therefore restricted to a single row of electrodes transversely located over the occipital scalp. A longitudinal row was rejected as it does not permit cross-hemispheric differences elicited by the different stimuli to be examined. It is generally known that the EP waveform does not vary significantly over distances of less than 2-3 cm [49] and so to obtain reasonable coverage of the transverse scalp distribution 7 electrodes were used having a separation of 2.5 cm, on an arc 4 cm above the inion (fig. 6.1). EEG

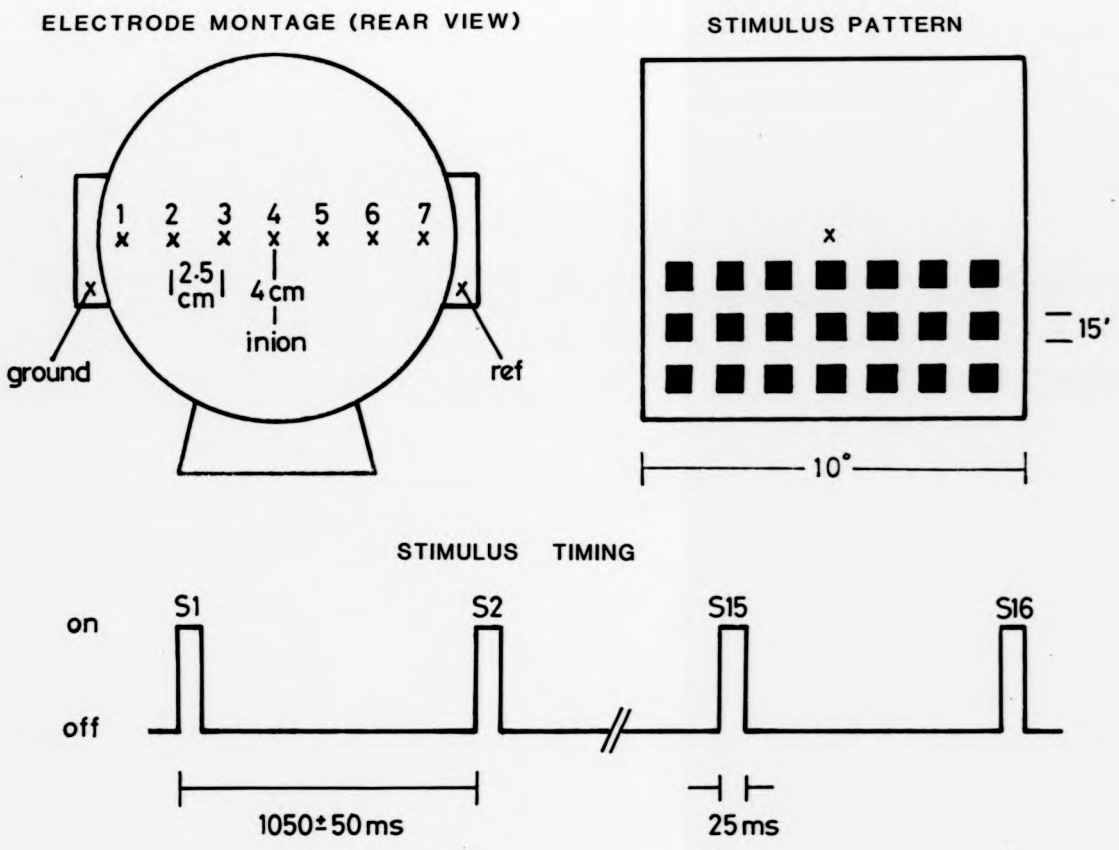



FIGURE 6.1 Details of electrode montage, stimulus pattern and stimulus timing arrangements used to record visual evoked potentials in this study.

| Series | Stimulus Duration (ms) | Period (ms) | Field | Pattern sequence |
|---|---|---|---|---|
| A | 25 | 1000 | lower | Isolated square/blank |
| B | 25 | 1000 | lower | Isolated square/outlines |
| C | 25 | 1000 | lower | Control (luminance change) |
| D | 25 | 1000 | lower | Grating/blank |
| E | 25 | 1000 | lower | Checquerboard/blank |
| F | 250 | 1000 | lower | Checquerboard/blank |
| G | 500 | 1000 | lower | Checquerboard/blank |
| H | 500 | 1000 | lower | Control (luminance change) |
| I | 25 | 1000 | right | Checquerboard/blank |
| J | 25 | 1000 | right | Control (luminance change) |
| K | 25 | 1000 | right | Checquerboard/blank |
| L | 25 | 1000 | right | Control (luminance change) |

TABLE 6.1. Experimental details of EP recording experiment for subject DAJ. Stimuli were presented in runs of 16 presentations, and there were 16 runs in each series. The stimulus conditions for each series are shown here.

| Series | Stimulus Duration (ms) | Period (ms) | Field | Pattern sequence |
|---|---|---|---|---|
| A | 25 | 1000 | lower | Isolated square/blank |
| B | 25 | 1000 | lower | Checquerboard/blank |
| C | 25 | 1000 | left | Isolated square/blank |
| D | 25 | 1000 | left | Checquerboard/blank |
| E | 25 | 1000 | left | Control (luminance change) |
| F | 25 | 1000 | lower/left | Isolated square/blank |
| G | 25 | 1000 | lower | Control (luminance change) |
| H | 500 | 1000 | upper | Checquerboard/blank |
| I | 500 | 1000 | upper | Control (luminance change) |
| J | 25 | 500 | left | Isolated square/blank |

TABLE 6.2. Experimental details of EP recording experiment for subject MJM. Stimuli were presented in runs of 16 presentations, and there were 16 runs in each series. The stimulus conditions for each series are shown here.

signals were obtained using a common reference earlobe electrode. The choice of reference site is important as it must be relatively free of EP activity. This is frequently difficult to achieve in practice but an ear reference is suitable for VEP recordings provided that upper-field stimulation is not used [115]. A single ear reference is vulnerable to ocular artifact resulting from lateral eye movements but these do not occur if visual fixation is maintained.

Experimental details are similar to those used in previous studies and described elsewhere [13]. The subject sat in a dimly lit room which was electrically and acoustically screened to minimise electrical interference and distractions. Visual stimuli were presented using a 4-field tachistoscope constructed in the department. This has 4 fields of approximate diameter 12 cm which are superimposed by an arrangement of half-silvered mirrors. Fast switching fluorescent tubes are arranged to sequentially illuminate the fields which have pattern masks inserted in front of them, so that tachistoscopic presentation of isoluminant patterns is possible. Brightness and contrast are manually presettable. Stimulus durations and rates were controlled by a Digitimer type 3290, preset by the experimenter.

### 6.1.2 Experimental procedure.

The procedure followed in recording EPs was similar to that employed in previous studies [13]. Following a warning tone, the subject fixates a mark in the centre of the visual field. An intercom. arrangement permitted communication between the experimenter and the subject and was used by the subject to indicate that he was ready for stimulus presentation. The experimenter was able to monitor 4 selected EEG channels which allowed him to initiate the experiment when the EEG

traces had settled down, and also to repeat the run if significant artifacts were detected. Instructions were given to maintain fixation throughout each run at the start of the experiment. No facility was available to monitor this, but experienced subjects were used who could be relied upon to cooperate. Large artifacts such as those caused by ocular movement were relatively easy to identify in the monitored EEG channels.

The electrode leads are connected to a headboard which routes the signals via screened cables to the EEG amplifiers. In these experiments a commercial Beckman 'TC-16' EEG amplifier was used, providing switched gain and filter combinations. A 150 Hz low pass filter was selected so as not to limit later analyses; this was the maximum cut-off frequency available on the Beckman EEG amplifiers. The amplified signals were recorded on a 14 channel Ampex CR1300A instrumentation recorder, along with stimulus timing signals and spoken identification records. These were recorded using a tape speed of 1 7/8 ips and the recorder incorporated a high order filter to attenuate frequency components above 300 Hz at this speed. Calibration signals derived from the EEG amplifier were also recorded prior to conducting the experiment. The overall instrumentation noise was estimated at about 2 $\mu$V rms.

Standard Ag/AgCl disc cup electrodes were used, attached to the scalp by means of collodion and filled with electrode gel after the skin had first been thoroughly cleaned with methanol solution to remove grease or other contamination. When the electrodes were firmly secured, the epidermis was gently abraded with the blunt end of a syringe to obtain good electrical contact. In this way impedances of 1-5 kohms were regularly obtained between each electrode and the reference

electrode. As the input impedance of the EEG machine exceeded 1 M ohm this ensured that signal loss was negligible [9]. The electrodes were removed at the end of the experiment using acetone.

Major artifacts tend to arise from gross body movements or movements of the head. It was incumbent upon the subject to adopt a comfortable position so as to avoid these during recording. As well as avoiding possible movement artifact this minimises subject fatigue, and reduces activity from neck and facial muscles which can otherwise degrade the quality of the data. Artifacts due to ocular rotation were less common, though occasional eyeblinks occurred in spite of the use of short duration recordings. These are easily identified in the data and can be used to reject contaminated runs if these are considered serious. As the experimental facility was well developed and employed adequate screening and single point grounding, no extra precautions were required to avoid a.c. mains interference.

## 6.2 Hardware aspects of the signal processing experimental facility.

### 6.2.1 Computer resources.

At the start of this research program there was no adequate computing facility for the requirements of this research. A Computer Automation Inc. CAI Alpha minicomputer was available but was routinely used in EP experiments and subsequent analysis. It had limited hardware and software features, as it supported only Fortran II andassembler languages and accessed only 16k RAM storage, which was considered to be inadequate for the requirements of this research. Central university computing facilities were available but suffered from inadequate data sampling capability, fairly heavy day-time use, limited resource

allocation and unscheduled service interruptions. It was therefore decided to set up a separate signal processing facility for use by the Device Applications Group in the Physics Department at Keele which had a similar requirement to this project. A DEC PDP 11/23 was selected as the basis for this for reasons of compatibility with existing departmental equipment, sharing of software and hardware experience within the department and externally, wide popularity of this range of minicomputers in scientific fields and some software support via the user group DECUS. A disadvantage of this choice was incompatibility with the CAI Alpha minicomputers used in the Department of Communication and Neuroscience, where the EP experiments were conducted. As off-line processing was likely to be the main activity this was not considered to be a serious disadvantage.

The PDP 11/23 was obtained in the following configuration: LSI 11/23 processor with Memory Management Unit (MMU) and Floating Point Option (FPU), 128 kbytes DRAM, twin RX02 floppy disk drives, 16 channel multiplexed analogue input to 12 bit ADC, programmable real time clock, general purpose digital I/O interface, quad serial line interface and ROM bootstrap board. Two RK05 hard disk drives were added later providing 5 MB of storage. In addition the peripherals included an LA36 DecWriter, Watanabe digital plotter and Televideo 192C VDU, all serviced via serial lines. A Gould Colorwriter digital plotter was available later to provide additional plotting capability. An analogue output card was constructed using the digital interface card to provide waveform display facilities. This unit is described in the next section.

### 6.2.2 Display Unit.

This unit was designed to provide 2 analogue output channels, suitable for driving an oscilloscope monitor or analogue plotter. To maximise display rates it was decided to use the integral oscilloscope time base, permitting 2 analogue waveforms to be repetitively displayed, synchronised by an external trigger pulse. The digital output port has a maximum data rate of 40k words/second which permits a satisfactory refresh rate of 20 Hz for a 2k word buffer. The unit was designed to permit easy software driving of the display yet with minimal hardware. As only one 16 bit port was available and 8 bit resolution was considered too poor, multiplexing of the 2 channels was necessary. Rather than setting a control bit or data word bit to designate the channel, which would require two instructions per output sample, the following solution was adopted. A control signal is used as a combined trigger/mode control, with one mode being single channel display and the other being dual (chopped) channel display. When the second mode is selected the data words are automatically routed to each channel alternately, starting with channel 1 to achieve synchrony. 16 bit DACs were available and were used for the sake of convenience, though 10 or 12 bit resolution would have been adequate. A second logic output was provided for pen control or trace brightness modulation. A block diagram of this unit and corresponding timing diagrams is shown in figs. 6.2 and 6.3.

Instrumentation had to be provided to enable the analogue signals to be sampled. As the ADC has a range of -5 to +5 volts and the instrumentation recorders had maximum outputs of approximately 1 volt rms, additional preamplification was provided to make full use of the ADC resolution. An 8 channel signal conditioner rack was built using

### 6.2.2 Display Unit.

This unit was designed to provide 2 analogue output channels, suitable for driving an oscilloscope monitor or analogue plotter. To maximise display rates it was decided to use the integral oscilloscope time base, permitting 2 analogue waveforms to be repetitively displayed, synchronised by an external trigger pulse. The digital output port has a maximum data rate of 40k words/second which permits a satisfactory refresh rate of 20 Hz for a 2k word buffer. The unit was designed to permit easy software driving of the display yet with minimal hardware. As only one 16 bit port was available and 8 bit resolution was considered too poor, multiplexing of the 2 channels was necessary. Rather than setting a control bit or data word bit to designate the channel, which would require two instructions per output sample, the following solution was adopted. A control signal is used as a combined trigger/mode control, with one mode being single channel display and the other being dual (chopped) channel display. When the second mode is selected the data words are automatically routed to each channel alternately, starting with channel 1 to achieve synchrony. 16 bit DACs were available and were used for the sake of convenience, though 10 or 12 bit resolution would have been adequate. A second logic output was provided for pen control or trace brightness modulation. A block diagram of this unit and corresponding timing diagrams is shown in figs. 6.2 and 6.3.

Instrumentation had to be provided to enable the analogue signals to be sampled. As the ADC has a range of -5 to +5 volts and the instrumentation recorders had maximum outputs of approximately 1 volt rms, additional preamplification was provided to make full use of the ADC resolution. An 8 channel signal conditioner rack was built using

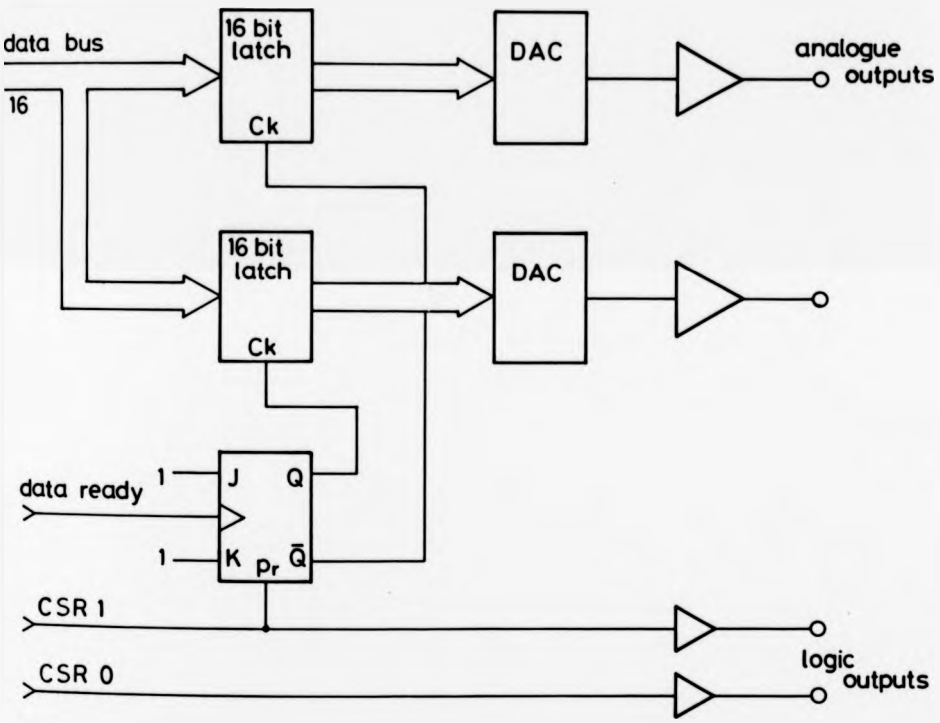



FIGURE 6.2 Block diagram of the Display Unit interfaced to the PDP-11 parallel output port to obtain dual analogue outputs.

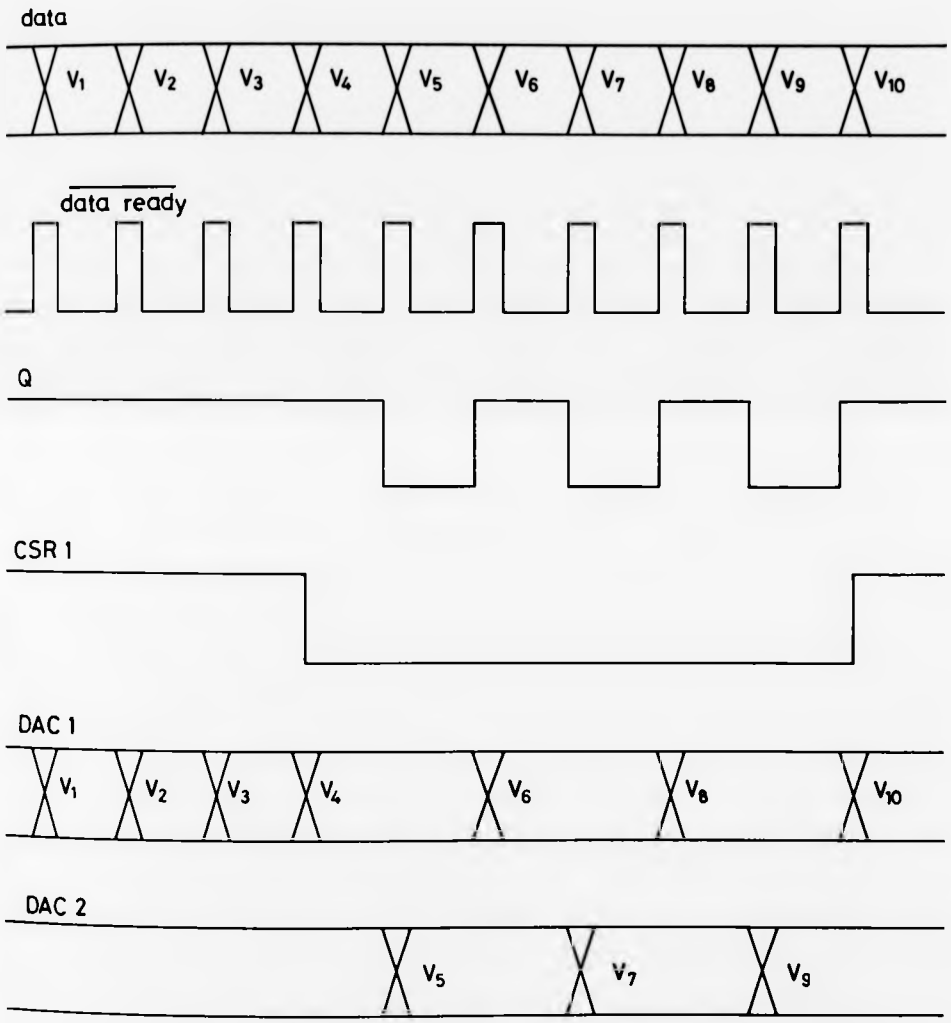



FIGURE 6.3 Timing diagram of the Display Unit (see Figure 6.2) showing single and dual output mode.

biFET op-amps in standard circuit configurations to achieve this. This unit provided switched gain in the range 1-300, anti-aliasing filters and d.c. offset controls for each channel.

### 6.2.3 Digitization of EP data records.

In order to process the EP data using the PDP 11/23, it was necessary to sample and store the raw records. To perform this efficiently a low sampling rate was needed and chosen to be 128 Hz. It is easy to verify that even with this sampling rate, 8 channels of EEG and trigger signals generate approximately 1000 samples for each second of the record. As a single run of 16 trials generates approximately 16000 samples, it was clear that higher sampling rates were impracticable. To avoid errors in sampling due to aliasing, it is well known that significant signal components must not exist above half the sampling frequency, known as the Nyquist frequency of the data. As the raw data did contain signal components beyond 64 Hz, it was necessary to filter the data prior to sampling. Although this bandwidth is a little small for all EP components to be accurately represented, it was necessary to sacrifice signal bandwidth to permit lower sampling rates to be used; examination of the EP waveforms revealed that little distortion of major features occurred.

Digitization of the data was performed by oversampling at 512 Hz followed by digital filtering and decimation to yield an effective sampling rate of 128 Hz. The advantages of this approach are that high order filters providing more than 20 dB/octave attenuation are relatively easy to implement digitally, and can be designed to have linear phase response. It is thus only necessary to employ a simple filter in each channel sufficient to prevent aliasing of frequencies

above 256 Hz, whereas the alternative approach requires high order analogue filters to be provided for each channel to attenuate signal components above 64 Hz. These are more difficult to design, and are more susceptible to phase distortion and inaccuracies arising from component tolerances.

Digital filtering and decimation was accomplished using the method suggested by Nielsen [116] which employs a 19-weight low pass FIR filter and discards every other sample to perform decimation by a factor of 2. The magnitude response of this filter is shown in fig. 6.4. This is an efficient method of performing this task as approximately half of the filter coefficients are zero, the coefficients are symmetric (thus halving the multiplications required) and they need only be applied to every alternate sample as part of the decimation algorithm. This reduces the computations by a factor of approximately 8. Decimation was accomplished by two consecutive applications of this algorithm to each data record.

## 6.3 Software aspects of the signal processing experimental facility.

As the main research activity was the development and investigation of signal processing methods, a facility to support computer studies of this nature was needed. Some care was taken to identify the needs of the project so that a suitable system could be developed. Interactive operation was considered essential as analysis procedures are in general data-dependent, often requiring the results of previous operations to determine subsequent operations. Simple extension of facilities was important to permit rapid development of new procedures. As well as requiring versatile control of processing strategy, different applications were likely to vary extensively and so

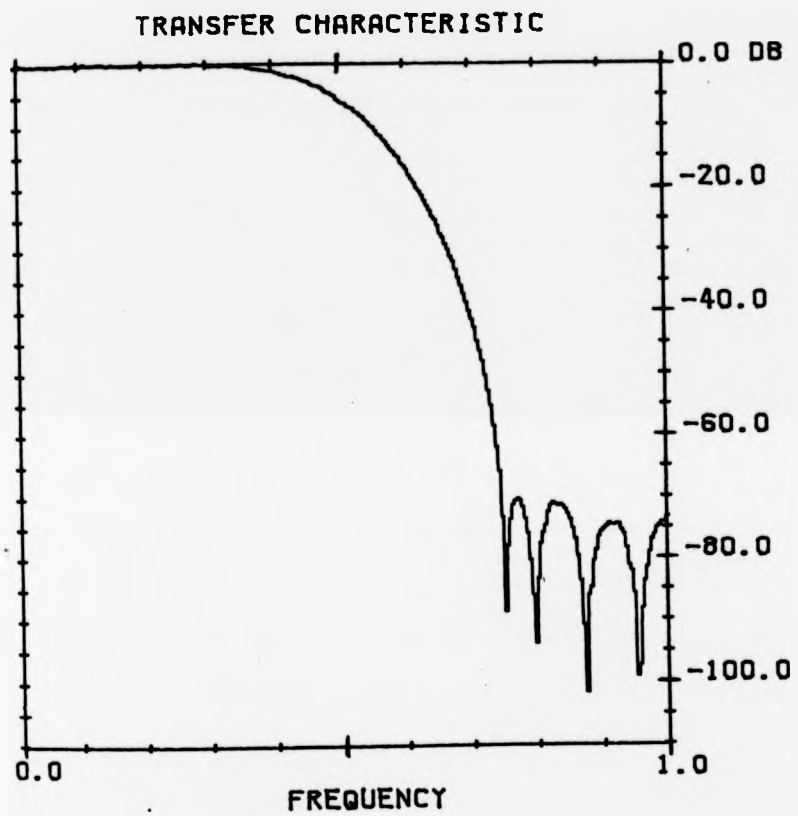



FIGURE 6.4 Magnitude response of the decimation filter based on a 19 weight FIR digital filter. Normalised frequency scale.

the facility would have to be general. As the main requirement was for a development facility, flexibility of use and ease of extension were considered more important than computational speed, which helped resolve the speed versus memory conflict in favour of efficient use of the latter. Within these general constraints, the specific tasks to be performed by the software were data acquisition, display, plotting, storage and processing, mainly of multichannel time-series data. Standard arithmetic manipulations and transformations were required, as well as a range of statistical and spectral analyses. Methods of generating artificial data and evaluating the results of processing were also required.

Although such a facility was considered to be fairly common, initial attempts to locate one in the scientific community at large proved fruitless, and the decision was made to develop one that would be tailored to these needs. Later investigations did reveal some possible contenders although by this stage program development was well under way, and the availability and delivery time could not be relied upon even if they were suitable. Some of these systems will nevertheless be included in the following discussion to illustrate some of the approaches taken and to provide a background within which the package described here may be evaluated.

### 6.3.1 Strategies in the design of a signal processing program.

There are a number of ways in which this task can be approached, and these will now be briefly examined. The first method is to create a number of self-contained analysis programs, each acting on some nominated data, performing the required operations, and presenting the results and/or saving them for later analysis. This would typically

require that a library of standard functions be constructed for input/output, display, etc. which would then be used by each individual program as needed. This is the approach taken in LAB-11 [117] and initially espoused by Hunter and Kearney [118] and others. It is also inherent in conventional high level languages extended to allow interaction with laboratory devices, such as CONVLINC [119] and CLASS [120].

There are several limitations to this approach. Firstly programs are large, as they each contain much code not directly related to the primary function of the program in order that they be self-contained. This usually necessitates the use of overlay techniques and frequently implies that task building becomes complex and time-consuming, thus hindering development. Secondly, modification of any code stored in the library can mean rebuilding all tasks making use of it. As this can occur fairly often in a developing system, maintenance of software and documentation can become tedious. The third disadvantage stems from the difficulty in achieving the required flexibility in general purpose programs of this nature. The dialogue in a menu or prompting system can become tiresome, while more sophisticated control means large overheads in each procedure. It is also difficult to perform automated analyses which have been developed interactively, especially when these cross program boundaries.

A better approach is to design a single package which contains a wide range of elementary operations which may be called up by the user in any sequence required by an analytical task. The immediate gain is flexibility, as the user has complete control over the manner in which operations are performed. If a number of elementary operations are provided, a wide range of possible analyses can be conducted

interactively. Further, modifying or adding to any of the elemental operations takes effect immediately, making available the extended functionality to all procedures requiring them. A disadvantage of this approach is that a single program containing all possible functions is necessarily large and may not be as efficient as a purpose-designed program for a given task. Another disadvantage is that in order to be applicable to different situations, generalised structures need to be employed which again are likely to be suboptimal in many situations and necessitate much more careful design. An example of the kind of question arising is in determining the proportion of memory allocated to data and to code, which will obviously depend on the anticipated usage and resources available.

A similar approach which shares many of these advantages is to have a suite of separate programs performing the elementary functions, called up by a supervisor program to provide interaction with the user. This has the advantage that each new function can be installed independently without having to be linked each time to the existing package. This obviously speeds up development, especially since the edit/compile/link cycle is only applied to the new program. This was the solution adopted by Hunter and Kearney [121] after experience with several other schemes. Unfortunately their work was proceeding at the same time as the author's and so was not available either for direct use or for direction in developing the present package. While something of this nature had been considered, it was rejected at the time as it was thought that providing shared access to common data for each program unit would be difficult using standard Fortran, and would possibly involve time-consuming file operations between each command. It might also be difficult to keep the supervisor resident between each command, which would otherwise have to be loaded each time.

The strategy described previously of having a single program containing all the required functions was therefore adopted and is similar to packages such as TOY [122] which is specifically for EP processing and SPAID [123] which is principally intended for system identification. Both of these are designed according to this general structure, and permit selection of different operations under keyboard control. TOY does not contain all the functionality that is desired for this research, and SPAID offers a number of fairly large, self-contained procedures rather than a set of basic operations. Unfortunately neither was written for a DEC machine and the language implementing TOY was not specified. SPAID was written in Fortran II and assembler and acted upon data in disk files, so would require considerable modification and extension. Neither was therefore suitable as a basis for this research work, and the decision was made to develop a suitable software facility based on a similar approach, to be implemented on a PDP 11 minicomputer.

### 6.3.2 Choice of programming language.

The choice of programming language was limited to DEC Fortran IV, since this was the only high level language available for the PDP 11 series of machines at the time of writing. Other high level languages would have been preferred and C was considered as particularly useful because of its high level structure combined with efficient low level access to machine-dependent functions. MACRO-11 was considered to be too clumsy and time-consuming to support the implementation of a package of this nature, although potentially it offers the fastest and most efficient code. One of the main advantages of using Fortran is the wide availability of software in this language. In particular the IEEE Volume of Programs for Digital Signal Processing [124] formed a useful basis for some of the spectral analysis routines as a large number of

well-documented and carefully tested programs are provided. Efficient assembly language FFT routines were available for use within these. Though these programs tended to be self-contained, it was possible to modify them for use within the program developed here. One of the main disadvantages of Fortran was its lack of structure compared to other languages such as PASCAL and C. Considerable effort was therefore taken to avoid the confusing code which could otherwise have resulted, and use was made of highly modular structure, clearly written code employing adequate comment statements, meaningful variable names, and a standard format for all user subroutines. These ideas have been suggested in texts by Kreitzberg and Schneiderman [126] and Day[127]. DEC extensions to standard ANSI Fortran were used whenever this would result in quicker program development and execution time, though it is acknowledged that this would hinder software portability.

### 6.3.3 Design requirements of the software.

The facilities that it was considered important to include in the proposed software are now described. It was necessary to provide a general means of interactive control to permit simple keyboard selection of different functions. This was most easily accomplished using mnemonic command names which were decoded to identify the function desired. Rather than employing a set of prompts for the data items (arguments) required by each command, it was convenient to include these within the same command line. Brief command mnemonics and argument references save unnecessary keyboard input and lessen user fatigue. It is common for several subroutines to require the same data items, such as sampling or plotting parameters. Rather than entering these each time it was convenient to provide global access to these so that they need not be changed once they have been assigned a suitable value. In

order to satisfy a wide range of different analysis situations it was considered important to have user-definable data structures since data arrays defined in the program are too restrictive and can be inefficient in their use of memory. Allowing the user to define these during program execution increases the complexity of the program but offers very considerable advantages in terms of flexibility and generality. The facility to define the position in memory and structure of a number of data arrays was considered invaluable.

Other facilities that were considered to be useful were the ability to store sequences of commands in files exactly as typed at the keyboard, for later execution. This can simplify use when common command sequences or even an entire analysis procedure is repeatedly required. This also permits unattended °batch° operation when time-consuming analyses are required. Logging of console dialogue was also considered important, to provide a hard copy listing of an experimental session. This also permits results of analyses etc. to be made available for later reference. Robustness of the program to run-time error conditions was essential, as these can otherwise lead to inconvenient interruptions or even loss of data. Each subroutine had therefore to be written to trap all possible errors such as floating point overflow, and a full range of error messages were needed to indicate when these occurred. These requirements are very similar to those of a simple interpretive language such as BASIC. To avoid the complexity of this approach the simplest possible syntax was devised which nevertheless implements the design requirements. More complex facilities such as expression evaluation were therefore omitted.

### 6.3.4 Implementation of the software.

Several important requirements for the software have been outlined in the previous section. A number of considerations in implementing these are discussed in this section, and will explain the nature of the final package. It has already been stated that minimal program size, flexibility of operation and ease of extension are among the most important design goals. The following sections show how these were achieved.

Efficient processing of data requires that it be resident in memory. For the case of multichannel brain recordings this involves large data arrays, which conflicts with the large program space required by a multi-purpose package. To resolve this it was decided to reserve a large data area for efficient operations on the data, and to employ overlayed command functions so that a large number of subroutines could effectively share a minimal amount of memory. Breaking system and command functions into a number of small subroutines has a number advantages, which includes modularity of design, avoidance of redundant code, clarity of program structure, simpler development and efficient use of memory by overlay techniques. These permit very large programs to be constructed, which are determined by disk capacity rather than the core memory size, as only the code for each command needs to be resident at any one time. The disadvantage is a slight reduction in execution time as non-resident code must be loaded before it can be used. Intelligent structuring of overlay regions is essential so that frequently called routines have a higher probability of remaining resident while rarely used routines (such as initialization or error message code) can be put in overlay regions which are non-resident for most of the time. Other techniques used to minimise program memory

requirements are packing of data by special techniques, use of integer or logical variable types whenever possible and storage of large amounts of text in files that are read when required (such as error messages or help information).

Two main difficulties were encountered in following this strategy. The first difficulty is that the PDP-11 family of computers have a limited addressing range of only 64 kbytes, as the 16 bit address bus addresses bytes rather than words. To provide access to more memory it is possible to use a memory management unit (MMU) which maps the 16 bit logical address space onto an 18 bit physical address space. This is required to make use of the 128 kbyte RAM available in our configuration. It is unfortunately not completely transparent as the user has to declare a variable or array as being of type VIRTUAL in Fortran IV if it is to be allocated storage in the extended memory space [125]. This is a non-standard Fortran feature which limits the portability of the package, and is not compatible with other Fortran IV statements, such as EQUIVALENCE and SUBROUTINE statements. A further disadvantage of this extension is that it is not efficiently handled by the system software, which was later discovered to set up the MMU prior to accessing each data item in extended memory, resulting in much slower access times. For this reason it was decided to reserve the full 64 kbyte of extended memory as an integer data area, which is sufficiently fast for simple operations. More complex operations would be performed by copying data into a real- or complex-valued data buffer in the base page memory for efficient processing. If this arrangement was found to be inadequate plans were made to modify the system components for improved extended memory handling, which promised substantial gains in processing speed.

The second difficulty encountered was that the Fortran compiler did not generate efficient argument transmission code to support subroutine calls. Not only was the in-line code duplicated for identical arguments, but also for identical subroutine calls. A calculation revealed that 100 subroutines having 5 arguments each would require approximately 15 kbytes of memory just for argument transmission. This is comparable with the memory available for the proposed package and is clearly excessive even with the use of overlay techniques, and so an alternative strategy was sought. As only one command subroutine needed to be resident at any one time, it was not necessary to have a separate argument transmission area allocated for each, as is required if subroutines are to be nested. This enabled considerable savings to be made by reserving a special communication area in the main keyboard control routine, common to each user subroutine. Arguments used by each command as well as globally accessible parameters were made available via this area aided by special system functions. The code requirements were thus reduced to a simple subroutine call for each command, and a communication area of less than 100 words for the arguments. Further details of this arrangement are given in Appendix A5.

Flexibility of use was achieved by supplying a wide range of elementary operations and functions. These can be called in the correct sequence to perform the desired analysis. Though this is not as efficient as purpose-designed code it did offer simpler development of new procedures. Flexibility was also achieved by the use of user-defined data structures and parameters. To permit easy extension of facilities it was decided to impose a formal structure within which compatibility of shared resources could be achieved, yet without sacrificing flexibility of approach. This was partly achieved through

the memory allocations and the argument and parameter structures described. To facilitate the addition of new commands it was decided to employ a fixed naming scheme so that the code to call up to 100 command subroutines was provided. This was intended to obviate the edit/compile cycle that would otherwise be applied to the main calling routine each time a new command was added. To aid in this process a system file was used to contain the definition of each command, and read in at the start of the program to configure the system tables. A system file containing all possible error messages was also used to provide a generalised error message handling facility which simplifies each subroutine and minimises program space. Adding a new subroutine merely required that it be written to use the communication protocol defined and linked to the existing object library to create a new task image.

The resulting facility was found to be a useful vehicle for the prototyping of different processing algorithms as well as for performing routine analyses. The final program exceeded 20,000 lines of source code and comprised approximately 150 subroutines. For this reason no attempt was made to include a full listing in this thesis. Further details may be obtained in Appendix A which contains a brief description of the program and includes an extract from the user manual, examples of the user dialogue, and various listings to show the features available.

## 6.4 Summary.

An account of the the experimental methodology in recording and analysing evoked potentials has been presented in this chapter. Equipment employed for this purpose has been described and includes a description of the computing facilities available. A large part of this chapter has been concerned with the software needs of this research

programme. A review of currently available software showed that this was inadequate for this task. The requirements and design of a suitable Fortran program for use with a PDP 11/23 minicomputer was presented, though further details of this package are contained in Appendix A. This chapter provides the necessary background for the experimental work to be described in the following chapters.

# CHAPTER SEVEN.

## PRELIMINARY CONSIDERATIONS REGARDING ADAPTIVE NOISE CANCELLING OF EPS.

### 7.0 Introduction.

In this chapter the case for using an adaptive filter to process EP data is considered. The advantages which can be gained by using adaptive filters have already been described. It was shown in chapter 3 that signal averaging and bandpass filtering of the raw data are the two most common signal processing methods applied. These achieve noise reduction by removing frequency components which lie outside the signal bandwidth, and by enhancing coherent activity with respect to incoherent activity. In this chapter experimental work is described which establishes the best manner of applying adaptive filters and what improvement in signal-to-noise ratio results. The results of applying adaptive noise cancelling to EP data are presented in the following chapter.

As discussed in Chapter 4, successful noise cancelling by adaptive transversal filters requires:

a) a primary input signal source containing the desired signal (in this case the EP) which is corrupted by additive uncorrelated noise (in this case all other signals including the EEG), and
b) one or more reference sources containing noise components that are correlated to some extent with the noise in the primary input but which contain little or no primary signal components.

The purpose of the investigations in this chapter is to determine whether or not this situation occurs in scalp-recorded evoked potentials, and to what extent the application of this technique is justified. In the next section the first of these requirements is investigated by considering typical scalp distributions of pattern evoked potentials obtained for two subjects using two particular stimulus types. This study reveals how well it is possible to identify reference electrode sites which are free from signal components. In the succeeding section a study is described which seeks to determine whether significant correlations are present in the EEG at multiple scalp sites, based on a simple linear model of signal relationships. Spectral estimation methods and in particular the coherence function will be used to characterise the correlations present in different frequency bands and to predict the signal improvement factor directly. A discussion of the results and their ramifications to the application of ANC are presented in conclusion.

### 7.1.1 Distribution of Evoked Potentials across the scalp.

In order to identify suitable ANC inputs a study was made of the scalp distribution of EPs using two normal subjects under two stimulus conditions. Had the intention been to identify general population characteristics, it would have been necessary to have used a larger pool of experimental subjects. As this study was exploratory in nature the use of a small number of subjects was considered sufficient to permit preliminary conclusions to be made, and to indicate the direction for further research.

### 7.1.2 Description of Results.

The results presented in this section primarily illustrate known spatial distribution characteristics of EPs and do not describe new findings. It was considered important to do this study rather than to refer to published results, because scalp distributions vary in different individuals. It is therefore necessary to determine the actual distribution in each experimental subject before new analyses are developed. Only those features that are of direct relevance to the aims of this thesis will be considered.

Fig. 7.1 shows 2 superimposed averages of 128 responses each, for each channel, subject and stimulus. Only the first 1 s is shown, of which the first 100 ms is the pre-stimulus epoch. Two averages are plotted in each case to indicate the repeatability of the responses. Three features of particular interest can be identified:

1) The EPs obtained for each subject have quite different morphologies. This is consistent with previous studies which conclude that significant form differences exist in the VEP for different individuals [25]. The superimposed averages support findings that intra-subject variability is generally much less than the variability across subjects.

2) While individual EP wave forms vary considerably, it is possible to infer characteristics that are more generally obtained. For example in figs. 7.1(b) and 7.1(d) both subjects show that the C1 component is predominant in the contralateral hemisphere to the one stimulated. This would be expected for a pattern presented to either left or right visual fields. In the same way, stimulation of the lower field might be expected to produce a more symmetric distribution about the midline, and

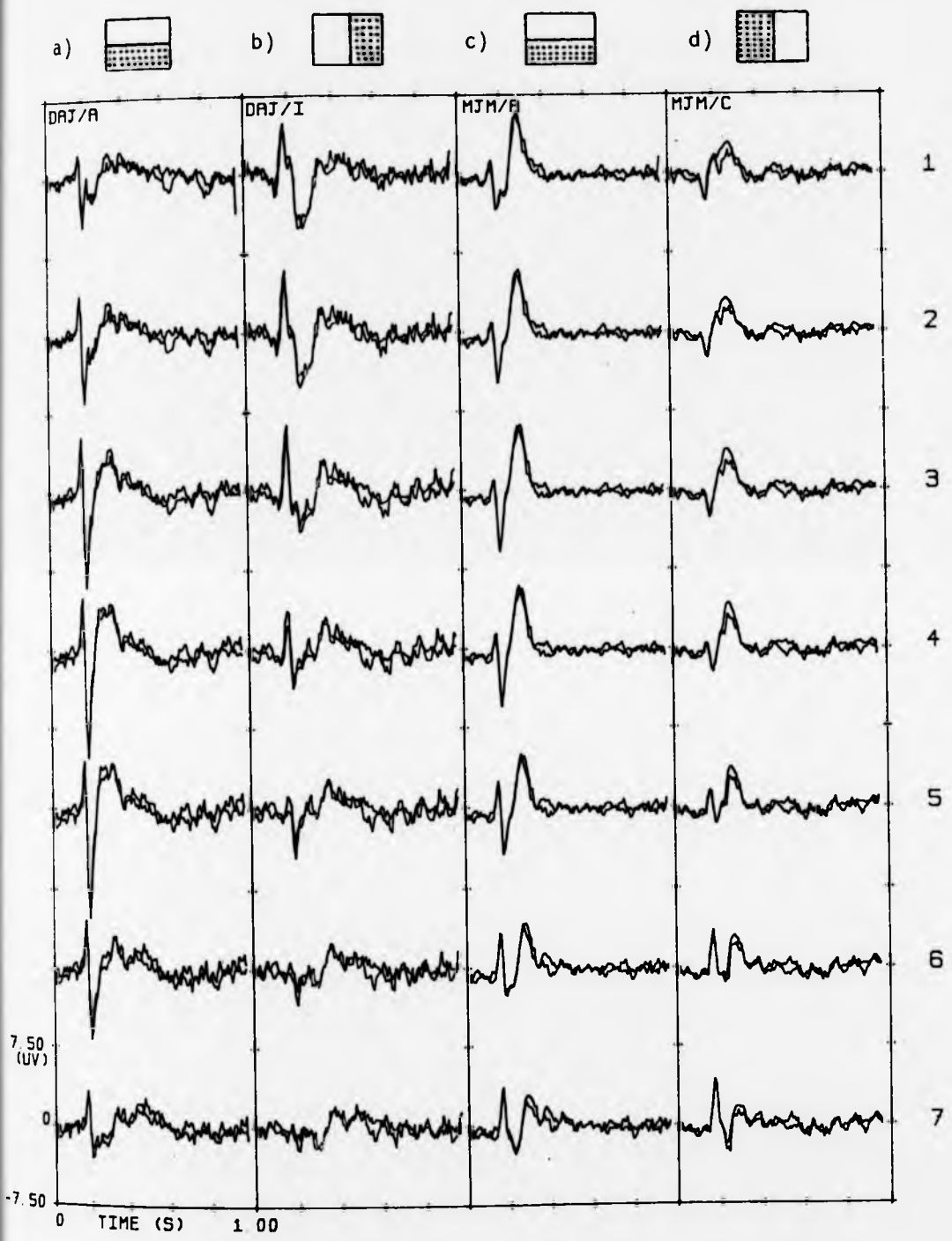



FIGURE 7.1 Average visual evoked potentials for 2 subjects and 2 stimulus types. Two averages of 128 trials shown superimposed. Electrode details in Fig.6.1.

this is observed in figs. 7.1(a) and 7.1(c). The time-courses of the main components appear to correspond well to the general pattern described in chapter 2 section 2.4.2.

3) The scalp distributions are different for each component. Thus in fig. 7.1(c) subject MJM shows a decreasing C2 and C3 amplitude as the scalp is traversed from electrodes 4 to 7, but an increasing C1 amplitude. Subject DAJ exhibits a very different scalp distribution for C1 compared with C3 in fig. 7.1(b). The significance of these results will be discussed later in this chapter.

## 7.2 Correlation properties of the EEG over the scalp.

In this section the results of a study will be described that characterise the nature of EEG correlations at several scalp sites. Fig. 7.2 shows typical raw EEG signals obtained at 7 occipital scalp sites for one experimental subject. These records were obtained with respect to an ear reference as before. From a visual examination of plots such as these it is evident that similar low and high frequency activity occurs in many of the channels. From inspection it appears that this can not be attributed to activity arising from the reference site alone. This suggests that there is a great deal of redundancy present in the multichannel EEG that conventional unidimensional methods such as averaging and filtering do not take into account. Bipolar recording does of course attempt to remove common activity present in two electrode channels by subtraction. This has the effect of enhancing spatial differences and has been used, for example, to locate the apparent sources of components. While this is a simple and generally effective method of cancelling common activity, there is no reason to suggest that simply subtracting EEG channels regardless of possible gain

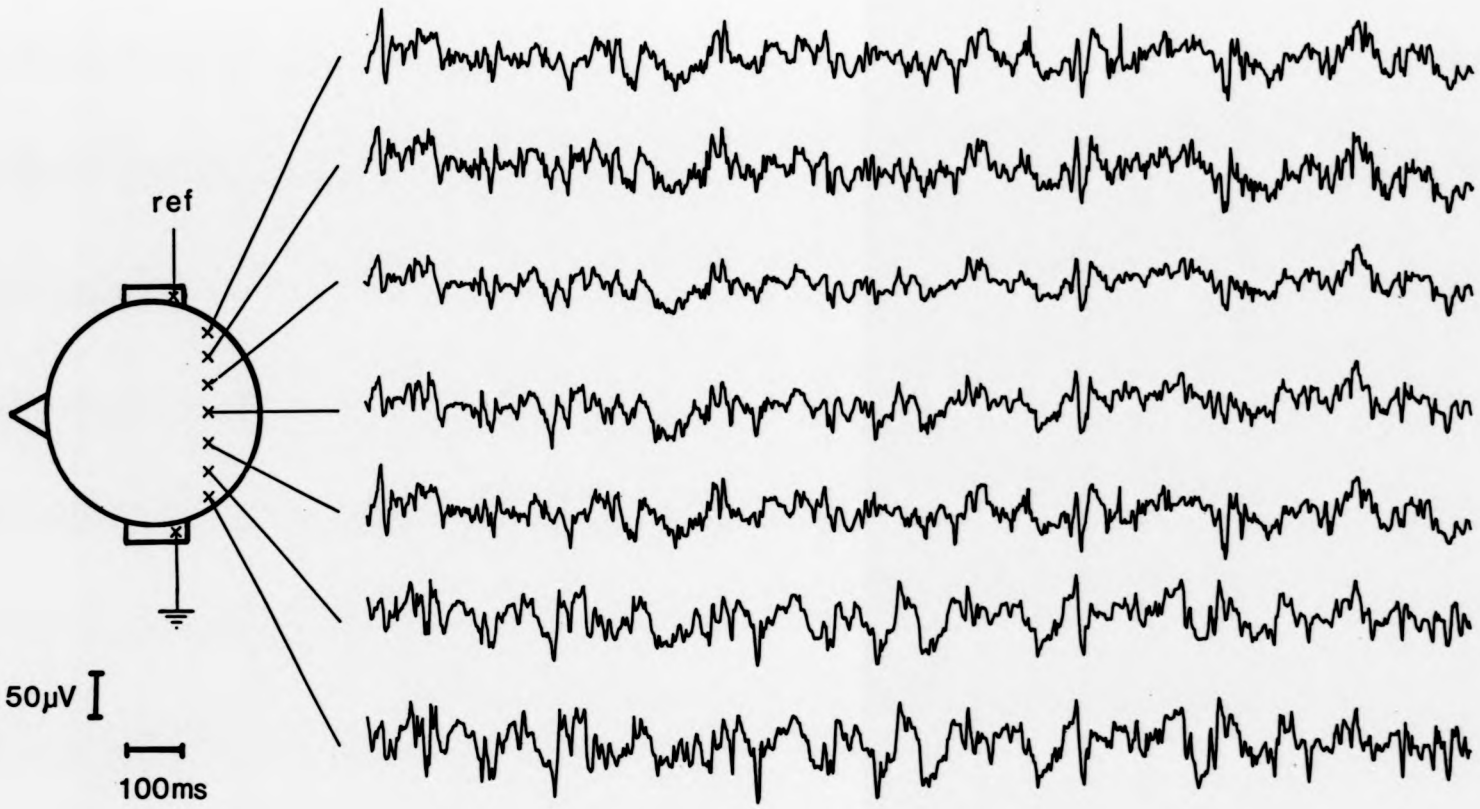



FIGURE 7.2 Typical multichannel EEG at 7 posterior scalp sites illustrating widespread similarity of activity.

and phase relationships will yield maximum signal enhancement. Used indiscriminately this can even lead to distorted EP estimates if significant EP components are present in both channels, which will become apparent later.

### 7.2.1 An EP/EEG signal model.

Fig. 7.3 shows a simple model for multichannel EP and EEG signals. It is not intended to represent the generation processes, but to be a simple descriptor of the relationship between these signals across the scalp. It will be useful in considering adaptive noise cancelling of EP data. S represents a train of dirac impulses synchronised with the stimuli. These impinge on a set of filters $h_i$ whose impulse responses correspond to the average evoked potential at each electrode site. These can be fixed or time-varying to represent deterministic or stochastic EP generation processes. G is a random noise source that feeds a corresponding set of filters $g_i$ to determine the nature of the EEG at each electrode. These might well be time-varying in general to reflect the changing composition of the EEG with time as well as scalp location. In addition there is a set of independent noise processes $N_i$ to model activity at each electrode that is independent of all other activity. This is to account for local EEG activity, electrode noise processes etc. It is apparent that this is an elementary linear model to describe EP and EEG signals. It implicitly assumes that the two are independent and that they linearly combine at each electrode to form the joint process. These assumptions are commonly made when modelling EP and EEG signals, and though not strictly true they are generally held to be sufficiently good approximations, as discussed in chapter 3.

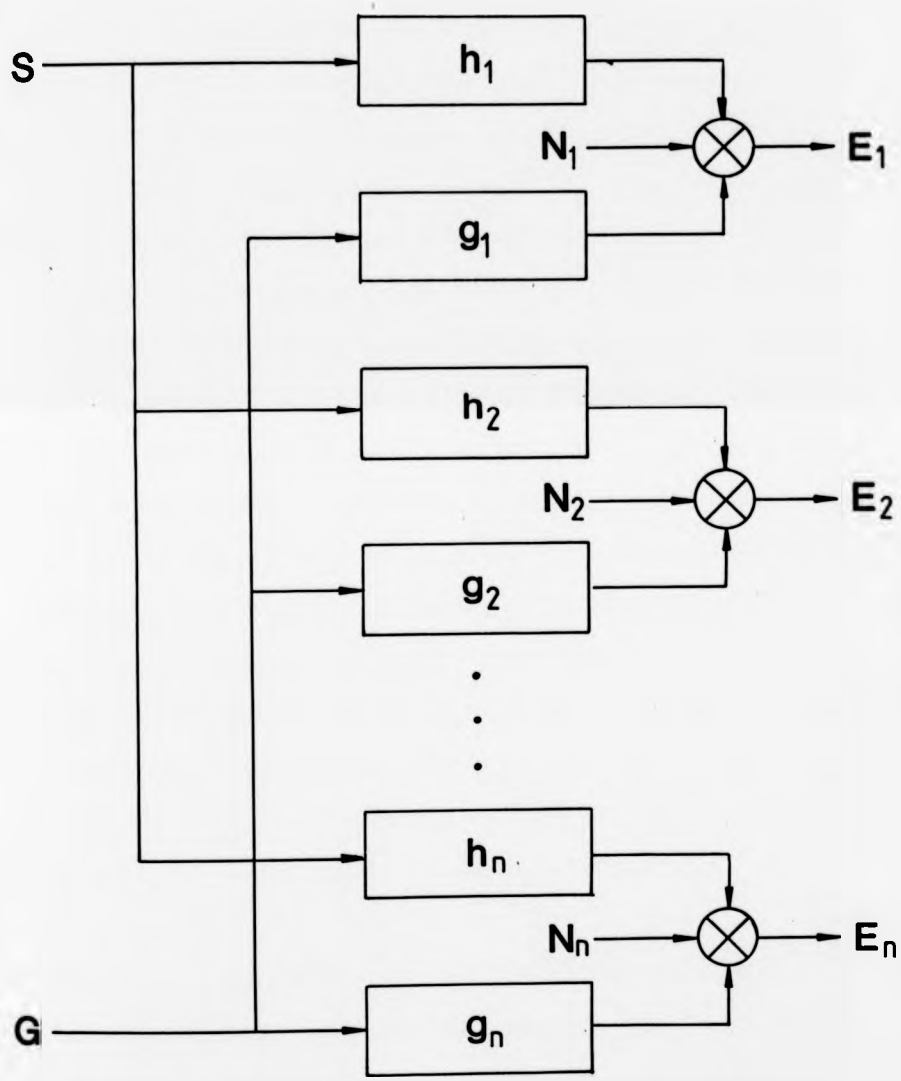



FIGURE 7.3 Simple linear model of signal relationships between EP and EEG activity at several electrode sites. (See text for an explanation of the symbols.)

### 7.2.2 Spectral analysis of the EEG.

The first task has been to determine the spectral properties of the EEG at the selected scalp sites. Firstly the auto-spectra of both raw responses and average EPs are obtained using the discrete Fourier transform (DFT). These indicate the main frequency components that describe these signals. The cross-spectra are also computed for all electrode pair combinations to indicate the degree of correlated activity present in each pair. This information is best represented using the coherence function, which permits direct prediction of noise cancelling performance. The procedures used will be outlined briefly, followed by the results and discussion. Note that the main purpose of these analyses was to obtain a general picture of spectral composition, sufficient for the identification of gross features. Conventional spectral analysis techniques employing the DFT are sufficient for this purpose. Finer frequency resolution can often be obtained from short time records by modern spectral methods such as autoregressive modelling techniques, but it is clear from these results alone that considerable spectral overlap exists between the EP and EEG even if the limitations of the DFT are taken into account. This is sufficient to show that the EP cannot be extracted from the EEG on the basis of spectral occupancy alone.

The spectral estimation was performed using the methods suggested by Welch [128] and by Carter, Knapp and Nuttal [129], which are described in Appendix B. The data record following each stimulus was divided into two equal-length epochs, one containing the EP and the other essentially unstimulated EEG activity. The EP was considered to be approximately contained within the first 500 ms, and this is borne out by the averages shown in fig. 7.1. Since it is the correlated EEG

components that are of chief concern, only the unstimulated epochs were analysed, as the presence of the EP in the record following the stimulus would have affected the results. Each segment had the d.c. and linear trend removed prior to further processing as these can result in erroneous spectral estimates. A full cosine Hanning window was applied to reduce the effects of spectral leakage. This is necessary to avoid errors in the coherence function [135]. The DFT of each segment was computed using the fast Fourier transform (FFT) and the average of the moduli obtained. The average cross-spectrum was computed using a similar procedure though the phase was retained in these calculations. Several other spectral functions including the coherence and the cross-phase were obtained from these as described in Appendix B. The auto-spectral power densities at each electrode are shown for both subjects together with the corresponding AEP spectra in fig. 7.4. Each plot of the raw EEG spectra is based on an average of 128 spectra, and the EP spectrum is based on the average of 128 consecutive EPs.

A number of observations can be made from these data. Firstly there are clear differences in the spectral structure of the EEG for each subject though these are relatively consistent for each subject. There are several distinctive features, e.g. subject DAJ shows a pronounced peak in the spectrum at approximately 10 Hz for all 7 electrode sites, though there is a subsidiary peak at about 3 Hz. Activity is reduced in the 16 - 32 Hz band. Subject MJM has a similarly bimodal EEG spectrum, but with broader peaks whose relative amplitudes change over the scalp. These peaks centred at approximately 3 and 10 Hz correspond to the delta and alpha bands respectively. It is possible that low frequency artifacts also contribute to this activity, though an examination of the spectral power density of test signals recorded on the instrumentation recorder failed to reveal significant components in

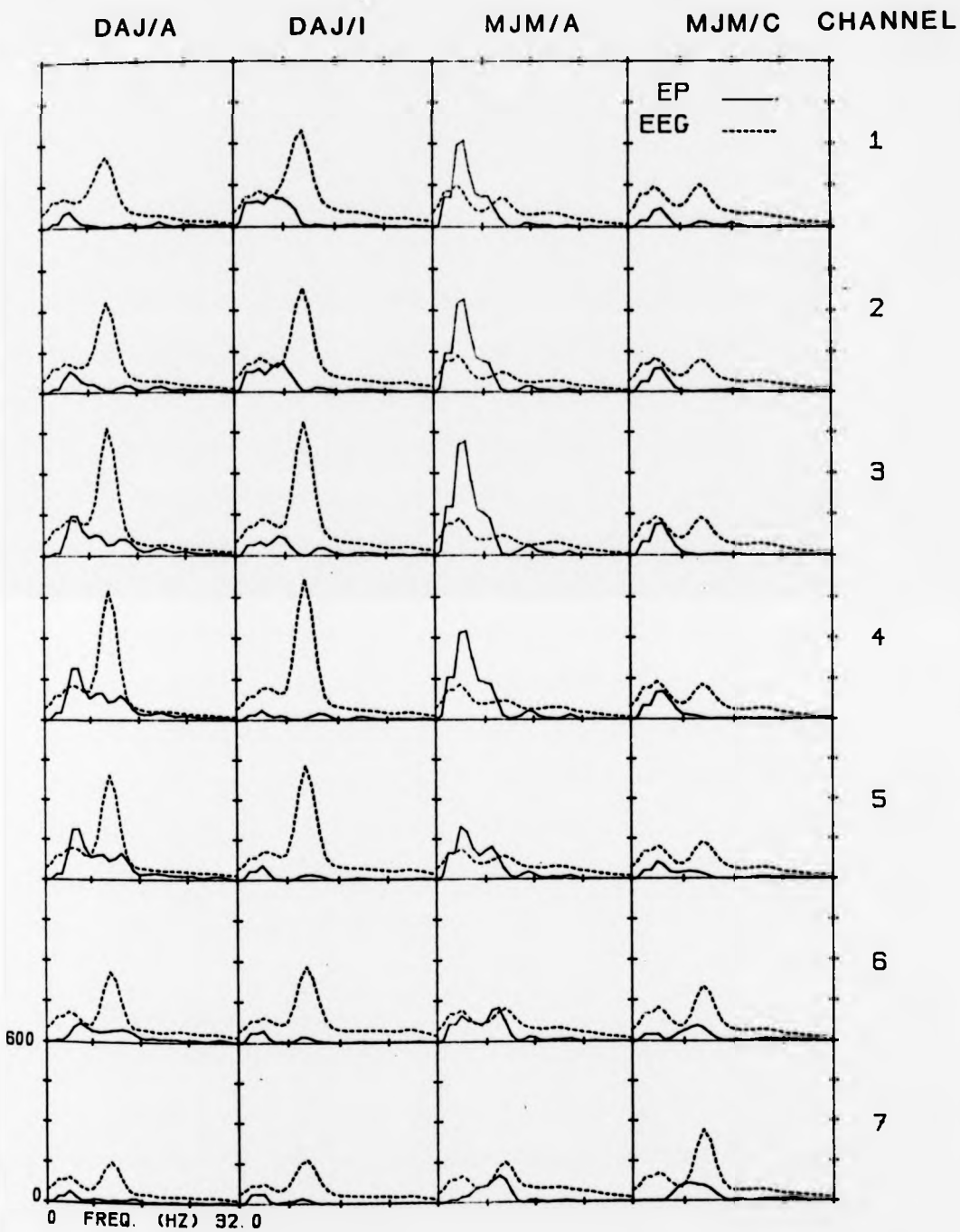



FIGURE 7.4 Sample EP and EEG spectral intensities for different subjects, stimulus conditions and electrode locations, showing overlap of gross spectral features (Arbitrary units of spectral intensity).

the signal bandwidth, so instrumentation noise may be discounted. Motion-induced artifacts due to muscle tremor, breathing or pulsing blood flow are also unlikely to be significant in the bandwidth 1-3 Hz. This indicates that neural activity is primarily responsible for activity at these frequencies.

The cross-power spectra are shown in figs. 7.5 and 7.6 for subject DAJ and in figs. 7.7 and 7.8 for subject MJM. The corresponding cross-phase spectra are also shown in figs. 7.9 - 7.11. The cross-power spectrum is a measure of the spectral power at each frequency that is due to correlated activity in both channels. The coherence function is more useful for determining cross-channel correlations, but it is important to compare this function with the constituent spectra to avoid errors in interpretation. Note that the cross-phase function is approximately zero in general, though significant phase differences are present in these data for certain electrode combinations. Studies of the phase differences across the scalp have reported these to be higher for longitudinal rather than transverse orientations [31] [32] [131] and were generally not significant for closely separated occipito-occipital sites, which is in accord with these data.

### 7.2.3 Coherence analysis.

A most useful measure that jointly expresses the auto- and cross-spectral information is the magnitude-squared coherence function (MSC), which will be referred to as the coherence in the remainder of this thesis. It is represented by $|\gamma|^2$, and obtained from the auto- and cross-spectral power in the manner described in Appendix B. The coherence provides a normalised estimate of the correlation between two channels at each frequency on a scale 0 to 1 rather like the correlation

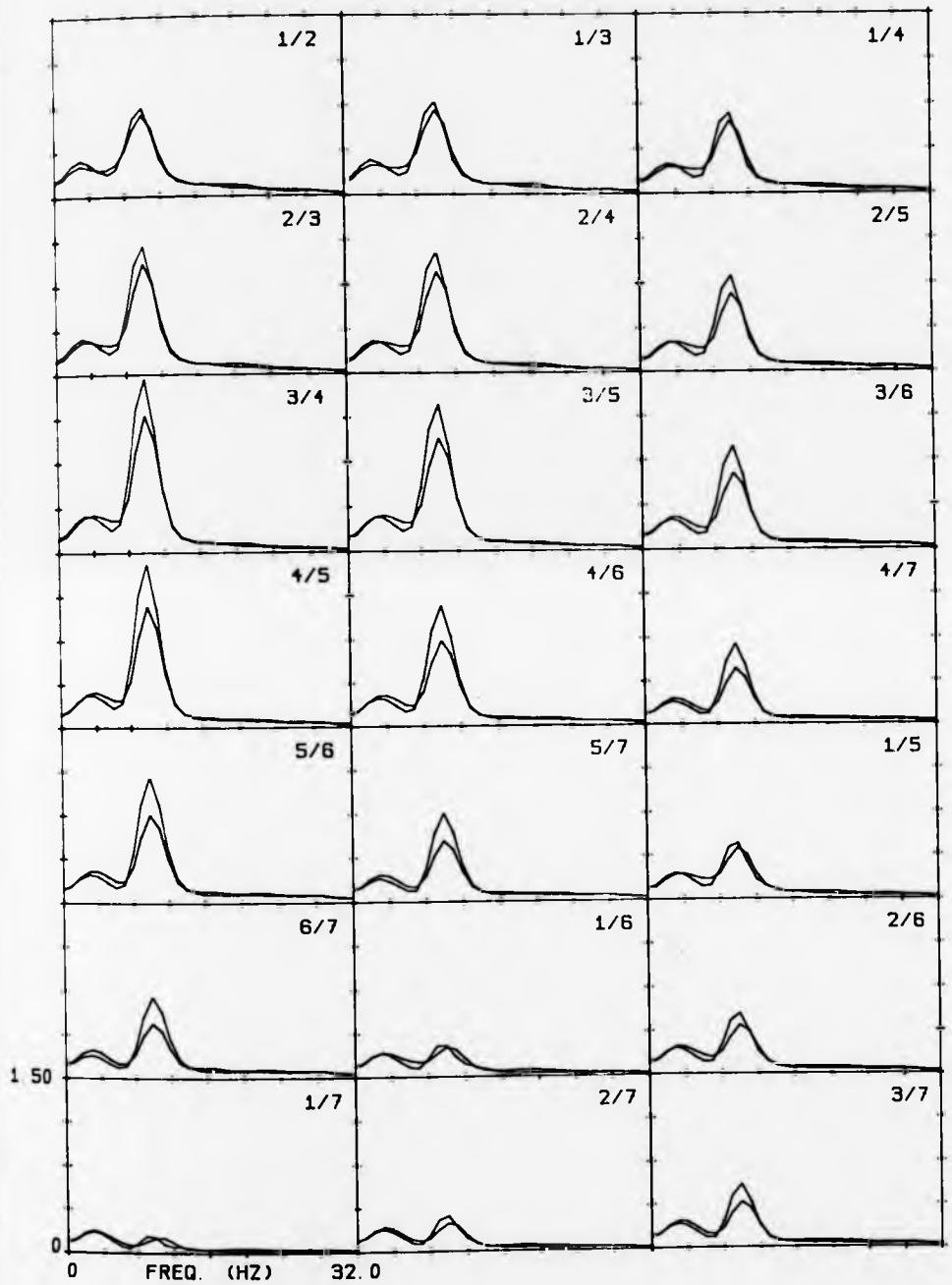



FIGURE 7.5 Sample EEG cross-spectral intensities for different electrode pairings. Subject DAJ, stimulus A. Two plots superimposed in each case. Arbitrary units of cross-spectral intensity.

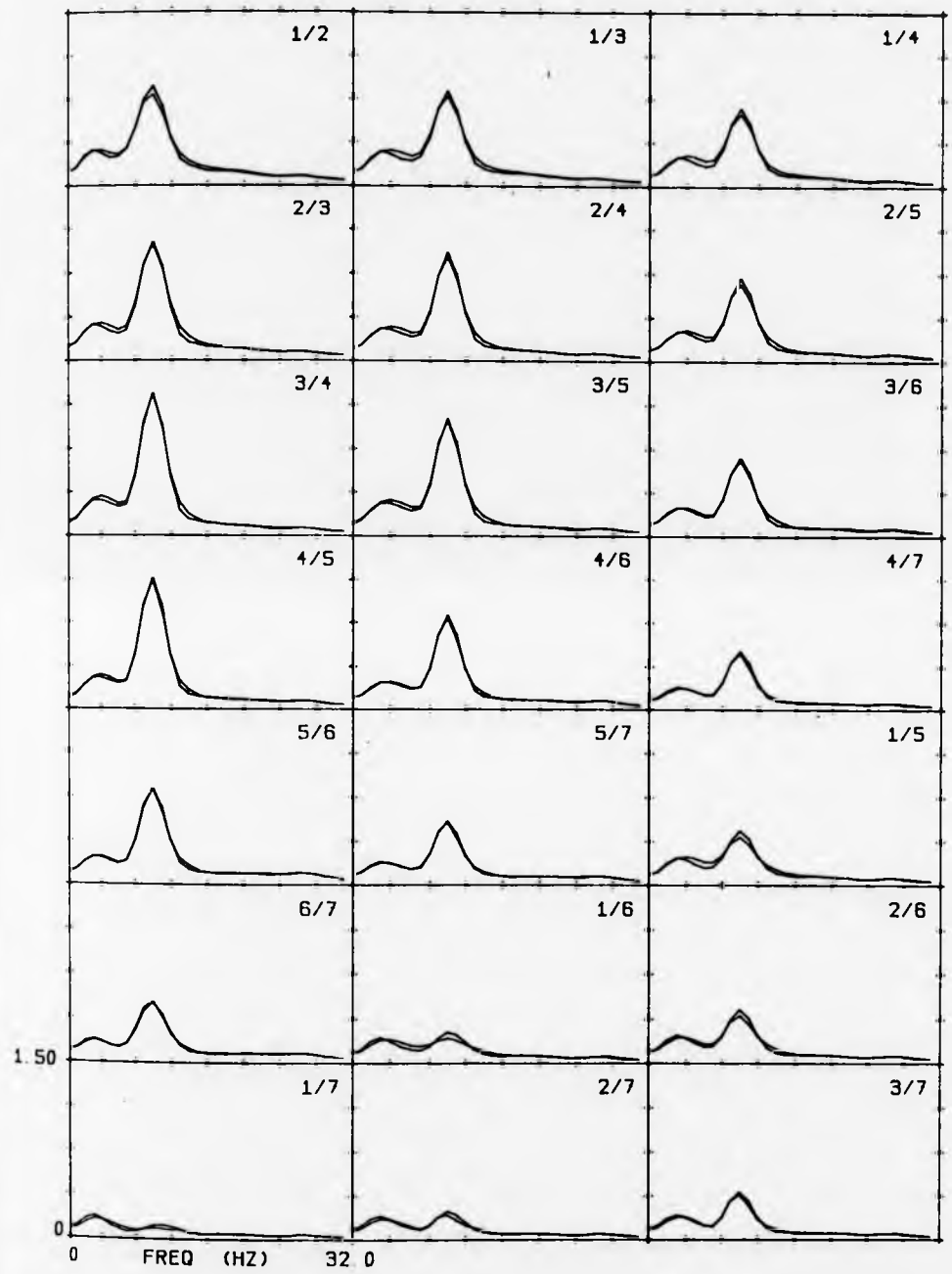



FIGURE 7.6 Sample EEG cross-spectral intensities for different electrode pairings. Subject DAJ, stimulus I. Two plots superimposed in each case. Arbitrary units of cross-spectral intensity.

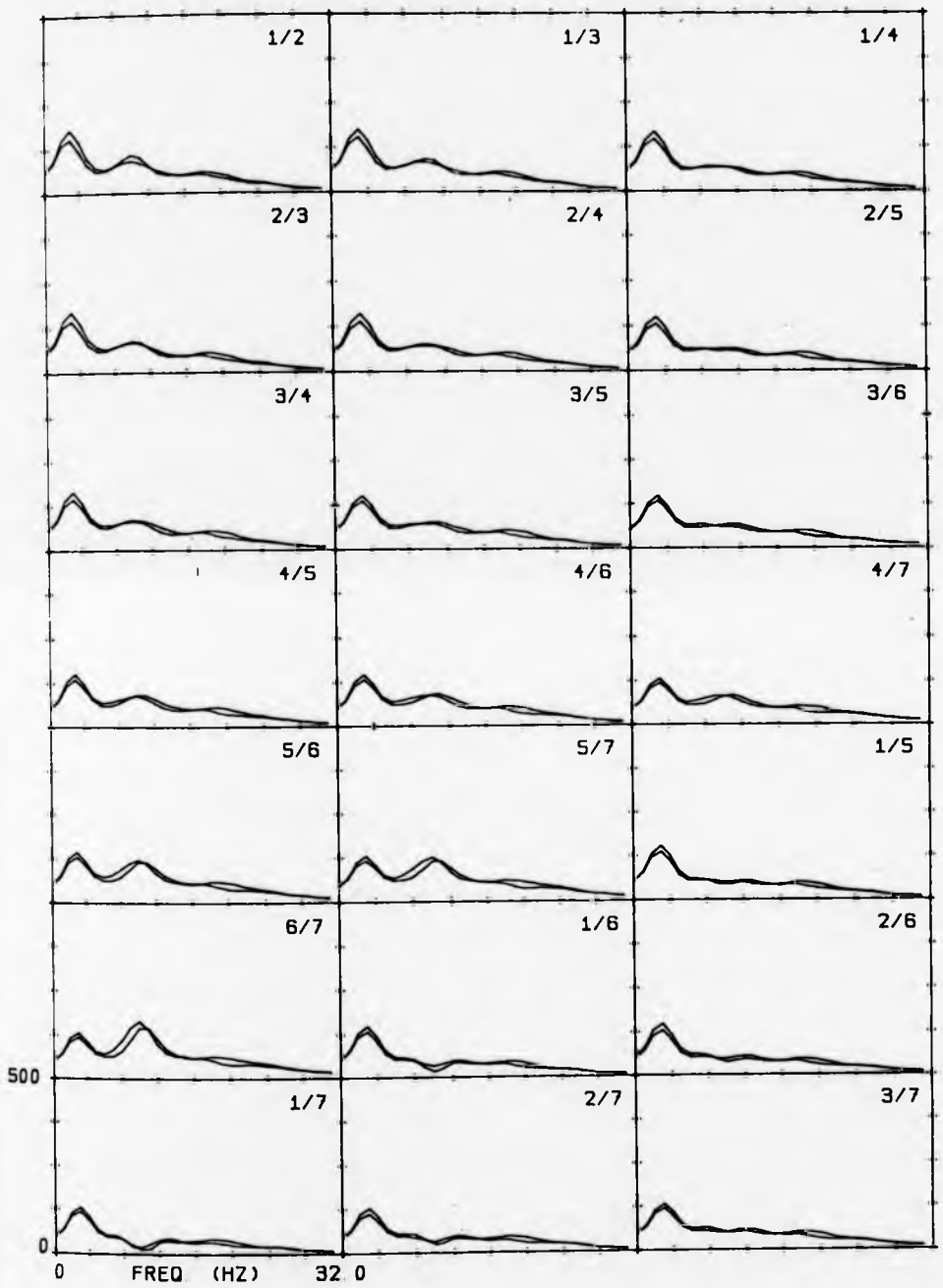



FIGURE 7.7 Sample EEG cross-spectral intensities for different electrode pairings. Subject MJM, stimulus A. Two plots superimposed in each case. Arbitrary units of cross-spectral intensity.

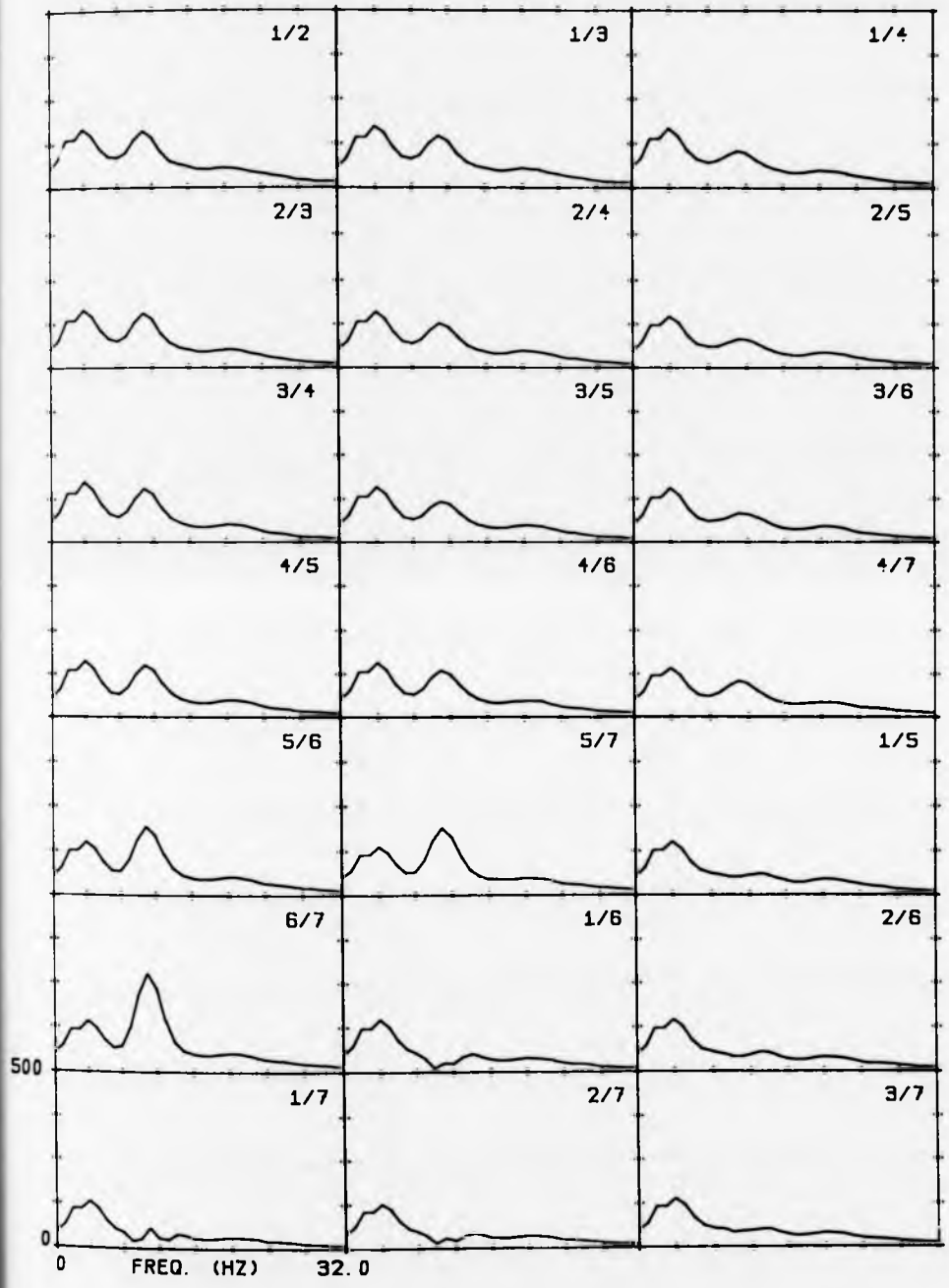



FIGURE 7.8 Sample EEG cross-spectral intensities for different electrode pairings. Subject MJM, stimulus C. Arbitrary units of spectral intensity.

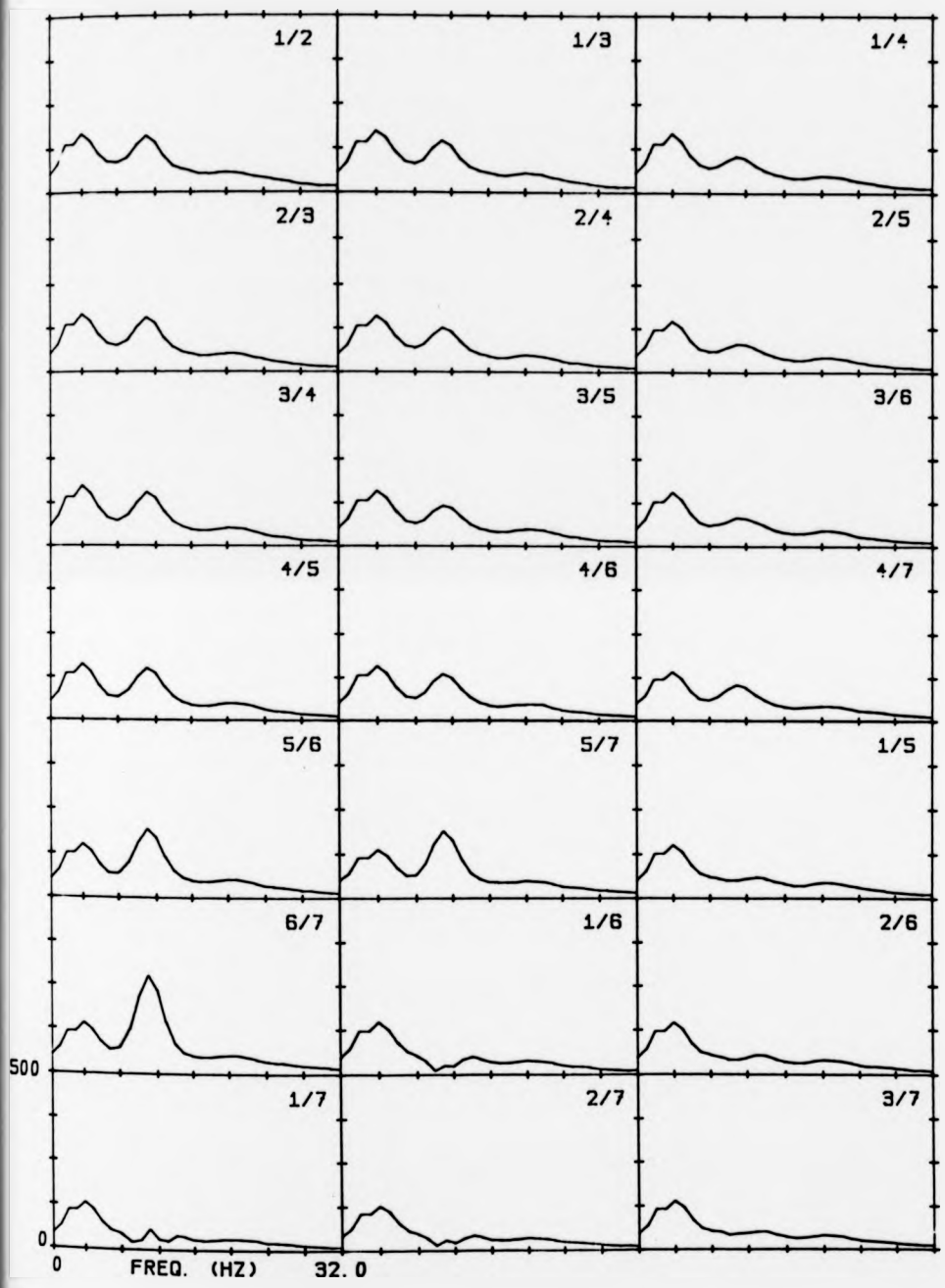



FIGURE 7.8 Sample EEG cross-spectral intensities for different electrode pairings. Subject MJM, stimulus C. Arbitrary units of spectral intensity.

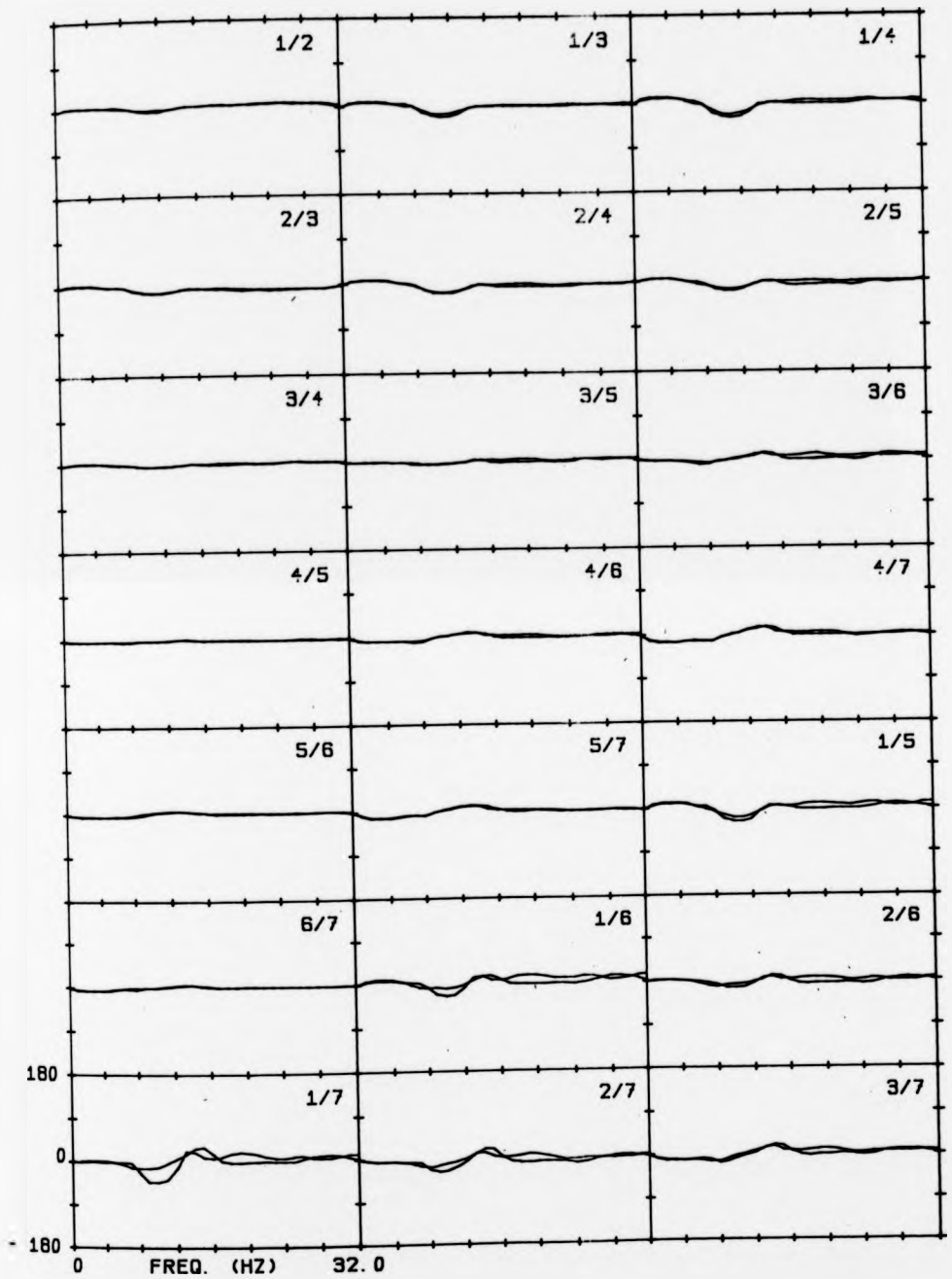



FIGURE 7.9 Sample EEG cross-phase functions for different electrode pairings. Subject DAJ, stimulus A. Two plots superimposed in each case.

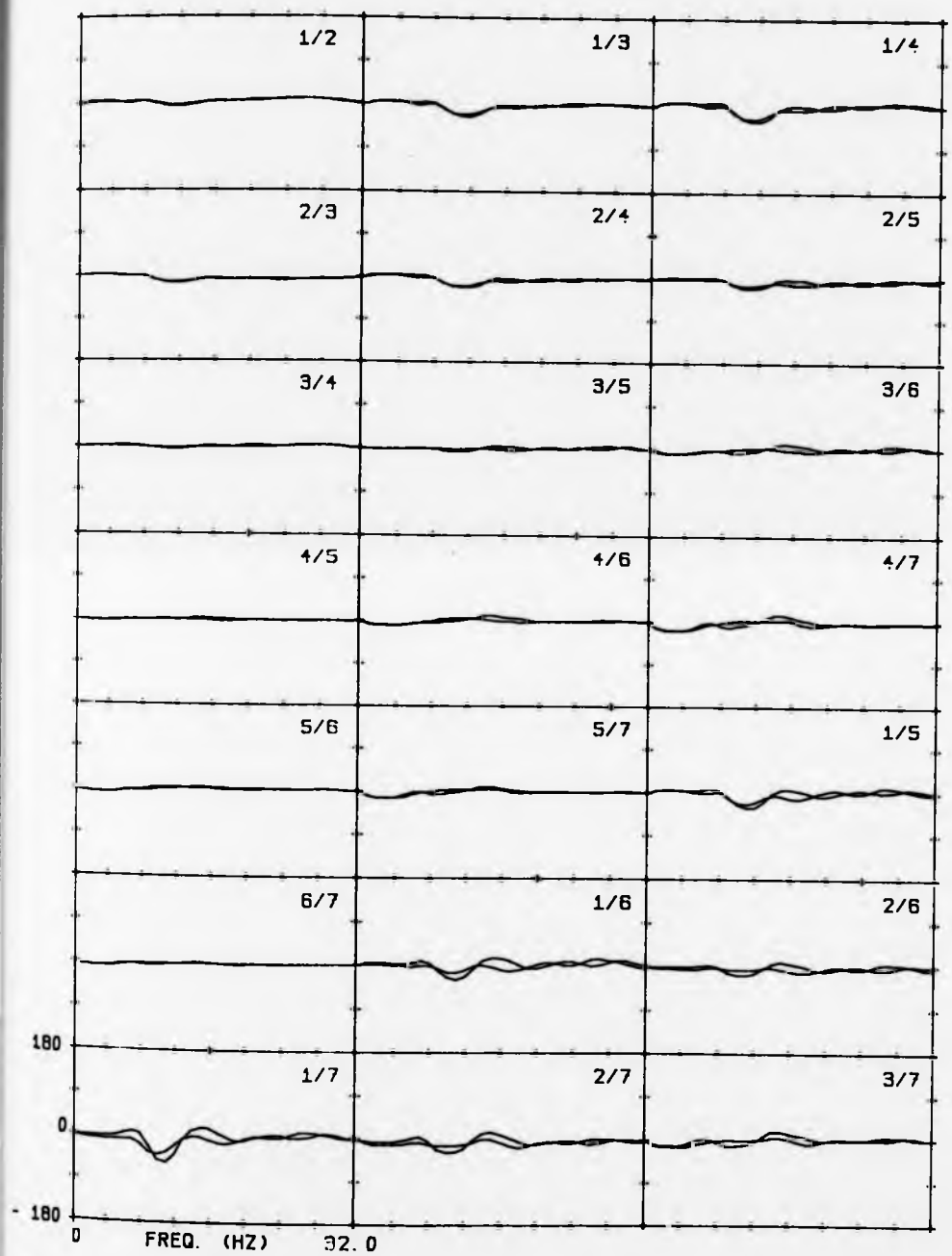



FIGURE 7.10 Sample EEG cross-phase functions for different electrode pairings. Subject DAJ, stimulus I. Two plots superimposed in each case.

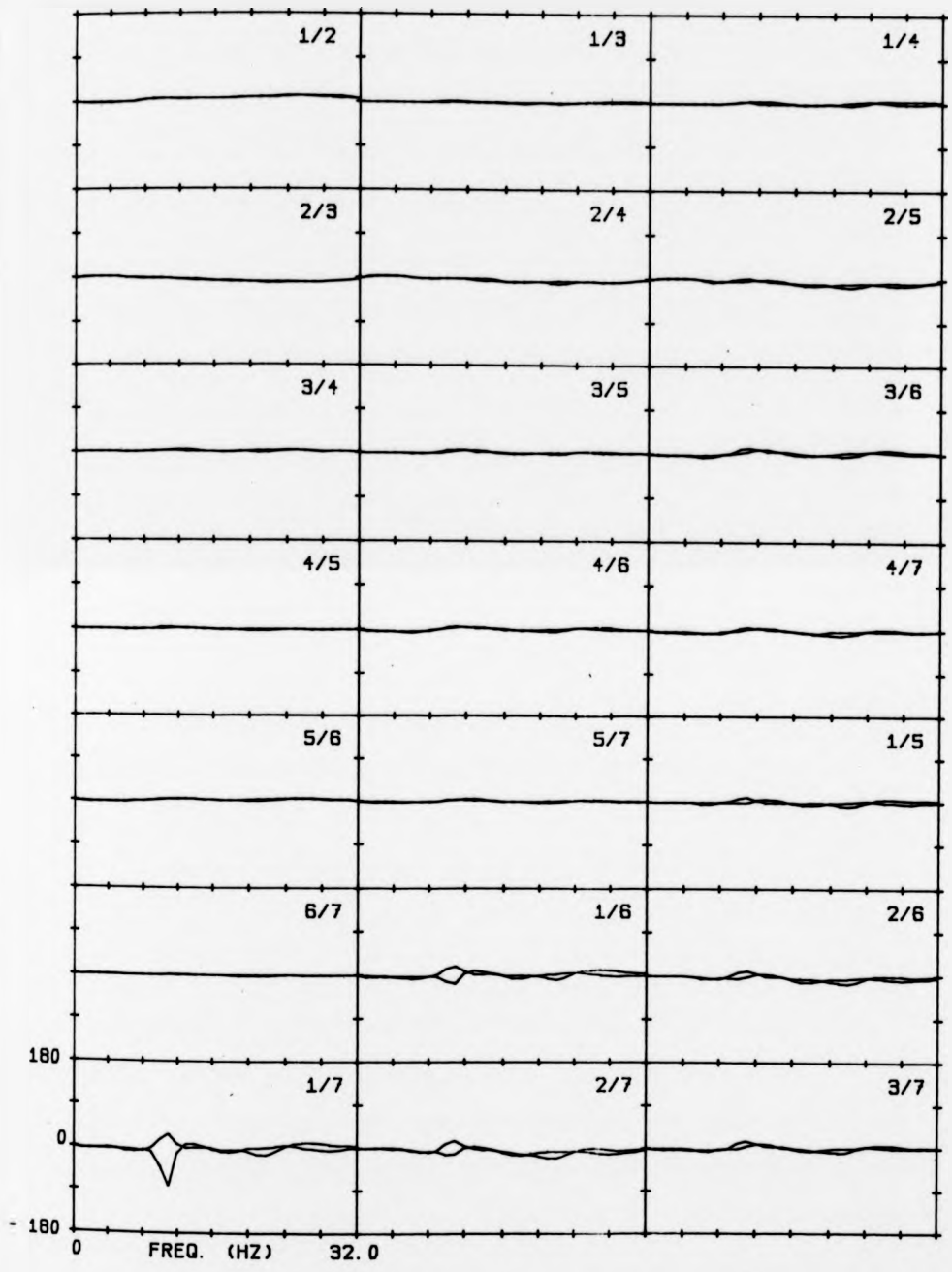



FIGURE 7.11 Sample EEG cross-phase functions for different electrode pairings. Subject MJM, stimulus A. Two plots superimposed in each case.

coefficient, and is thus independent of the actual signal powers at each frequency and of any linear transformation applied to either data channel. Its properties are well-known and described in texts on signal analysis such as [132]. The most important of these are briefly summarised here. A coherence of unity indicates that the signals in each channel are perfectly correlated at that frequency, and a value of zero indicates no correlation. Intermediate values are due to a mixture of correlated and uncorrelated processes at each input, which can be either deterministic or stochastic. These can arise from

a) a non-linear process relating the two inputs (and hence introducing non-linearly related components)
b) other inputs in the system, or
c) extraneous noise.

Further details of the coherence function, its statistical properties and methods of estimating it are presented in Appendix B.

Consideration will now be made of the features of interest in the coherence function computed for the experimental data shown in figs. 7.12 - 7.15. Firstly there is a general trend for the coherence to decrease as electrode separation increases. This is consistently observed for all the data analysed and indicates that this is a general result. The coherence functions computed for electrodes of a given separation do not differ widely and can clearly be classified according to this criterion. They do however exhibit small variations depending on actual scalp location; for example these occur when electrodes are placed close to the reference site.

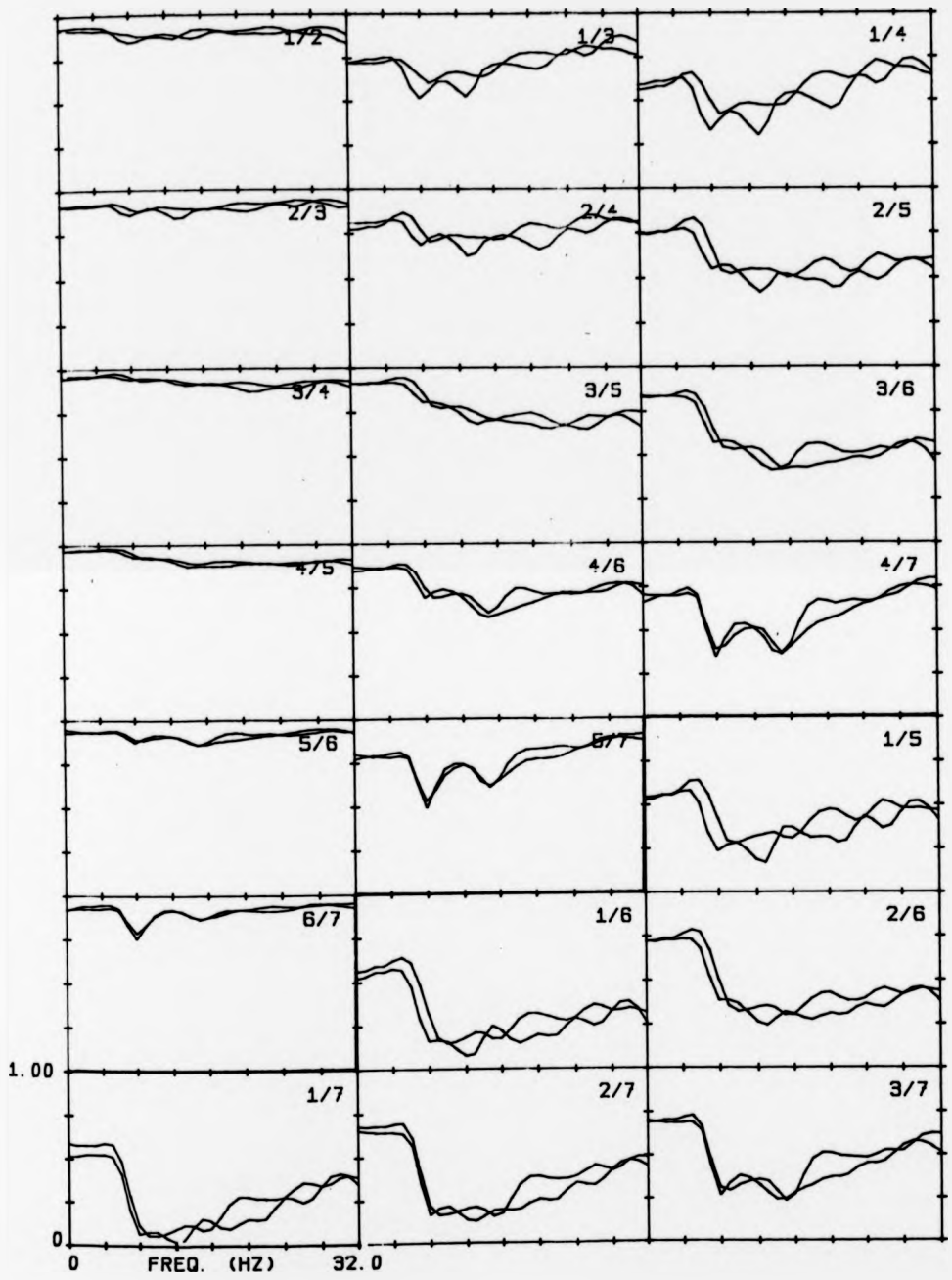



FIGURE 7.12 Sample coherence functions for different electrode pairings. Subject DAJ, stimulus A. Two plots superimposed.

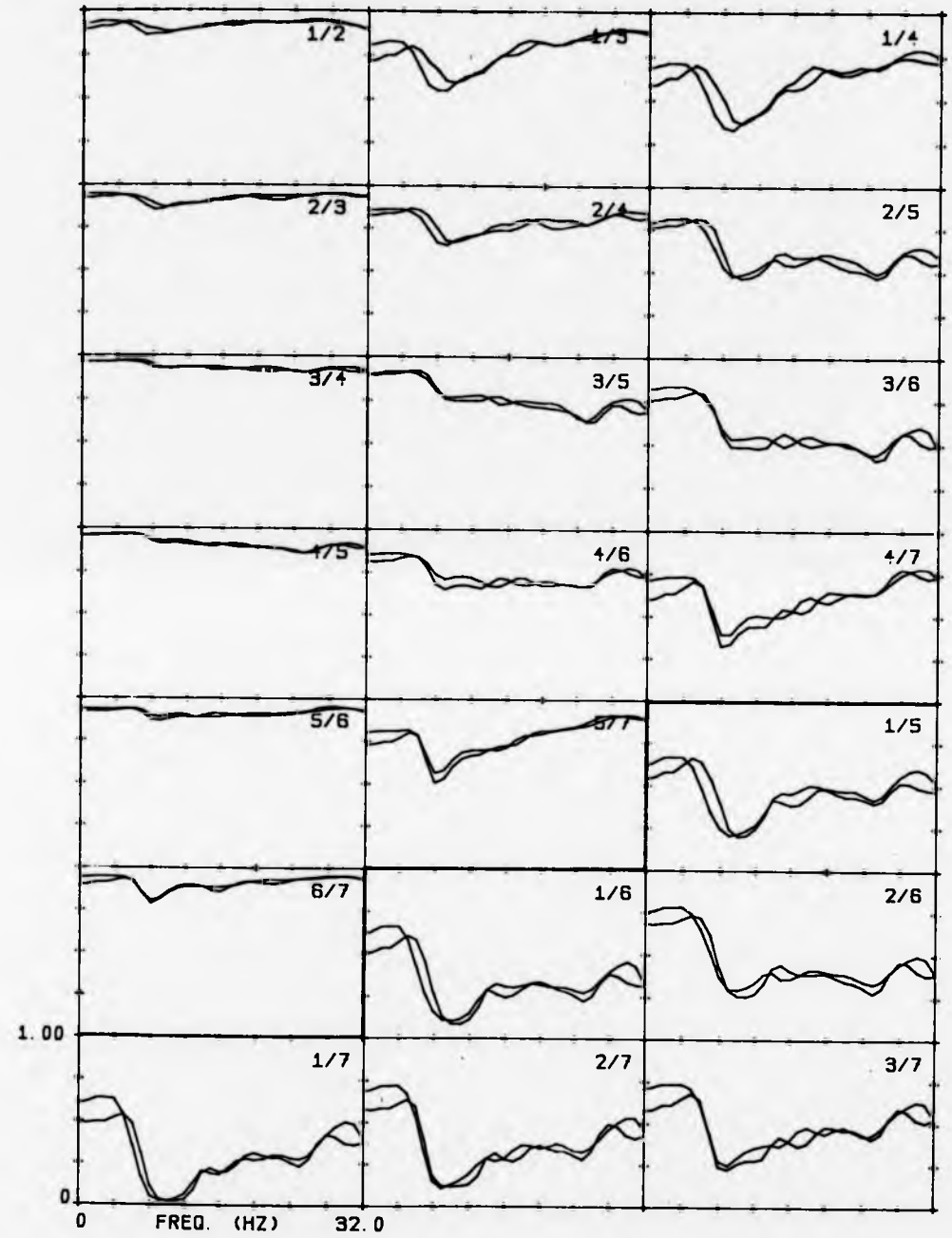



FIGURE 7.13 Sample coherence functions for different electrode pairings. Subject DAJ, stimulus I. Two plots superimposed.

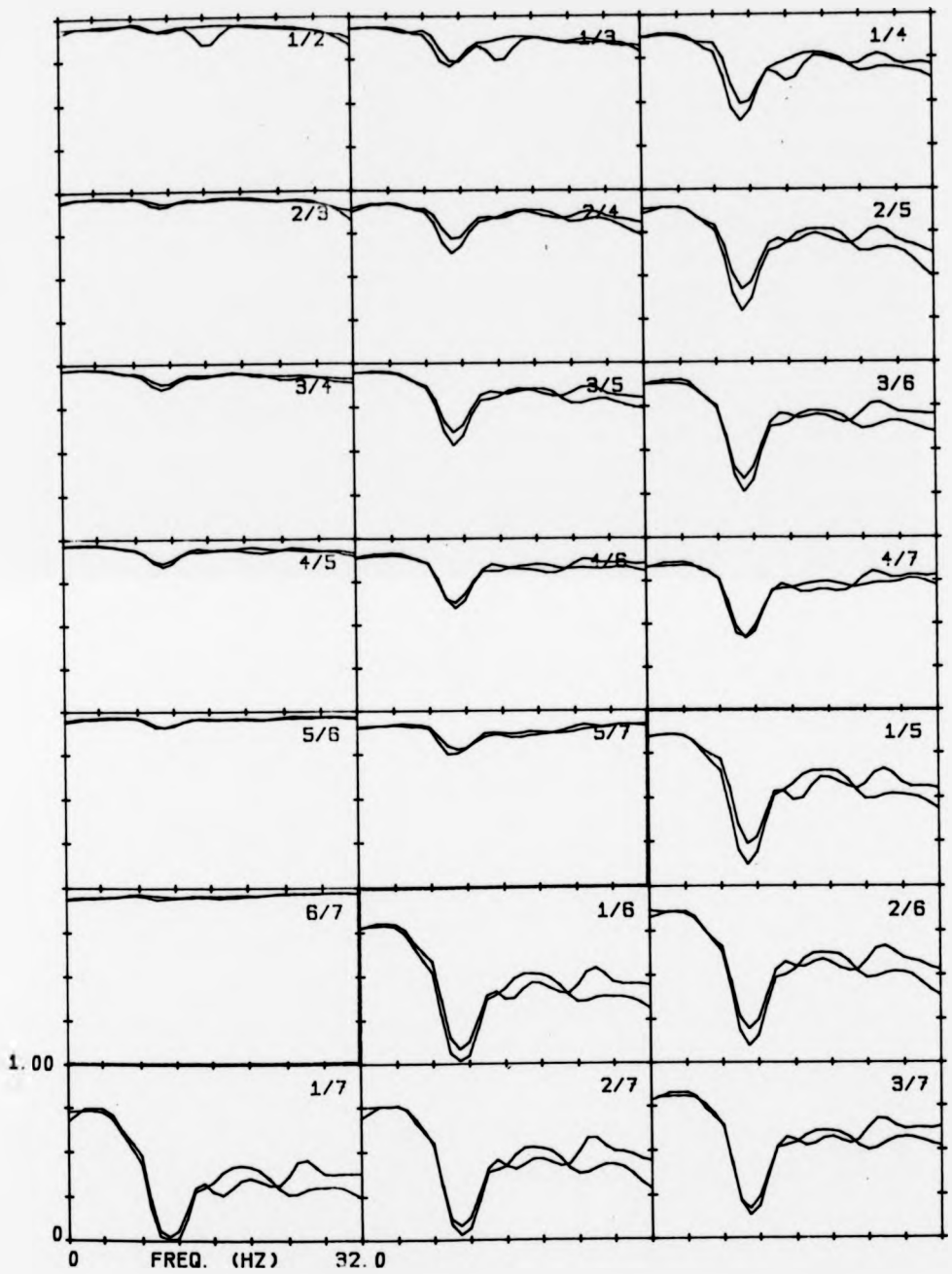



FIGURE 7.14 Sample coherence functions for different electrode pairings. Subject MJM, stimulus A. Two plots superimposed.

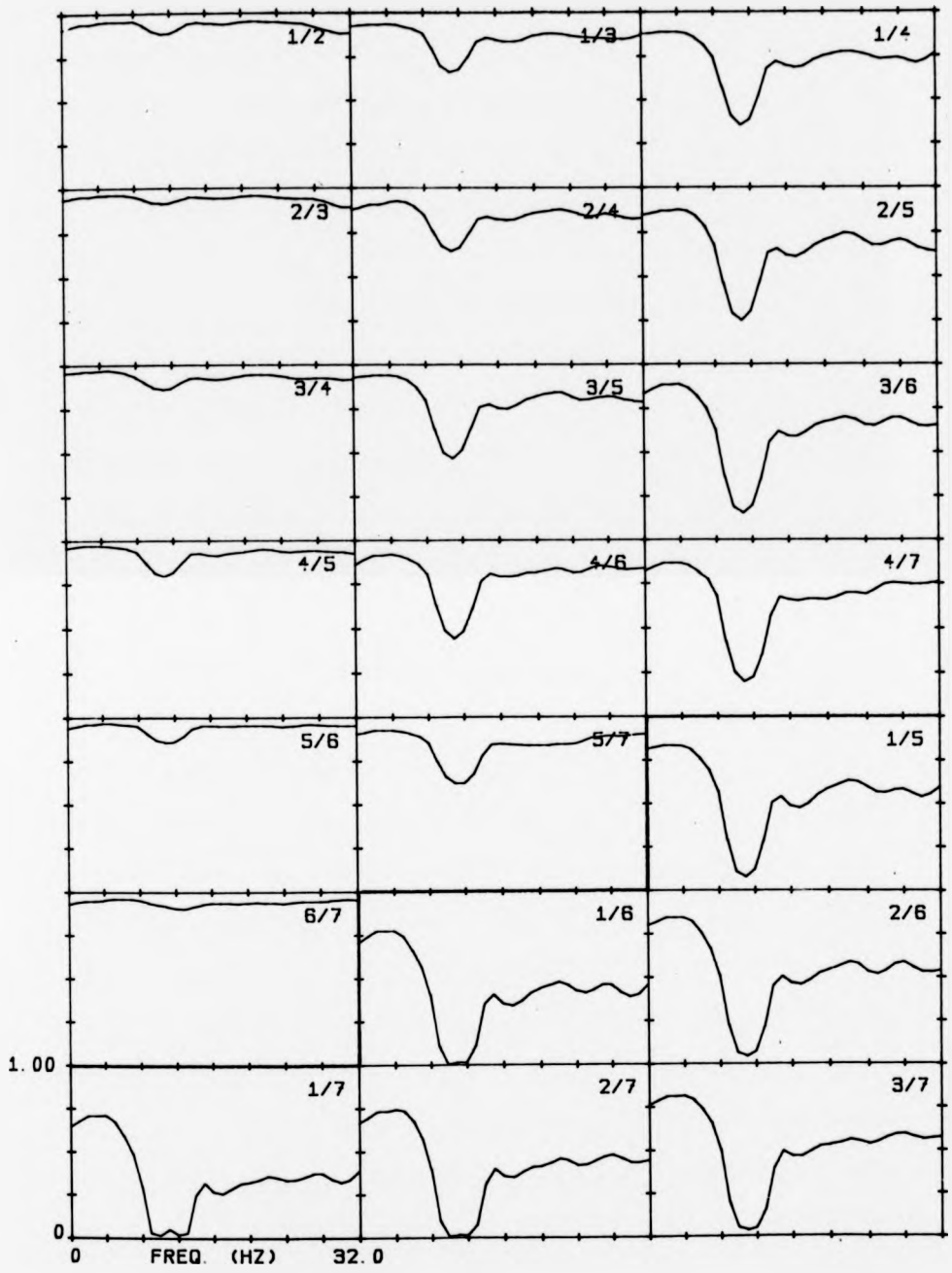



FIGURE 7.15 Sample coherence functions for different electrode pairings. Subject MJM, stimulus C.

An important feature in these data is a significantly reduced coherence occurring at and beyond the alpha frequencies (8-16 Hz) for all other than the closest electrode separations. This decrease in coherence is strongly dependent on electrode separation and is very marked for medium to large separations (5 - 15 cm), when it can approach zero. The extent of this dip is fairly broad for subject DAJ, whose coherence gently rises again in the region 16 - 32 Hz, whereas subject MJM exhibits a much narrower dip of width about 6 Hz centred at 10 Hz. The coherence is fairly high in the low frequency band 0 - 4 Hz where it shows least variation with electrode separation, while the frequencies above about 16 Hz show an intermediate dependence of coherence with electrode separation. The interpretation of these results will be presented in the following section.

### 7.2.4 Discussion of Results.

The correspondence of the major dip in coherence with the alpha band indicates that activity in this band is not well correlated over the scalp, even though alpha activity features prominently at all the occipital scalp sites considered. This may be seen clearly by comparing the cross-spectra with the corresponding autospectra in figs. 7.16 and 7.17, which shows a gradual attenuation of this component with increased electrode separation. This result is in good agreement with the assertions of Lopes da Silva and Storm van Leeuwen [133] and others that the alpha rhythm is not generated in a central, synchronising source (e.g. in the thalamus) but in small aggregates of neurons (networks) which are loosely coupled over the cortex. Each network generates activity in the alpha frequency band which can influence neighbouring networks, but the influence decreases with increasing scalp separation.

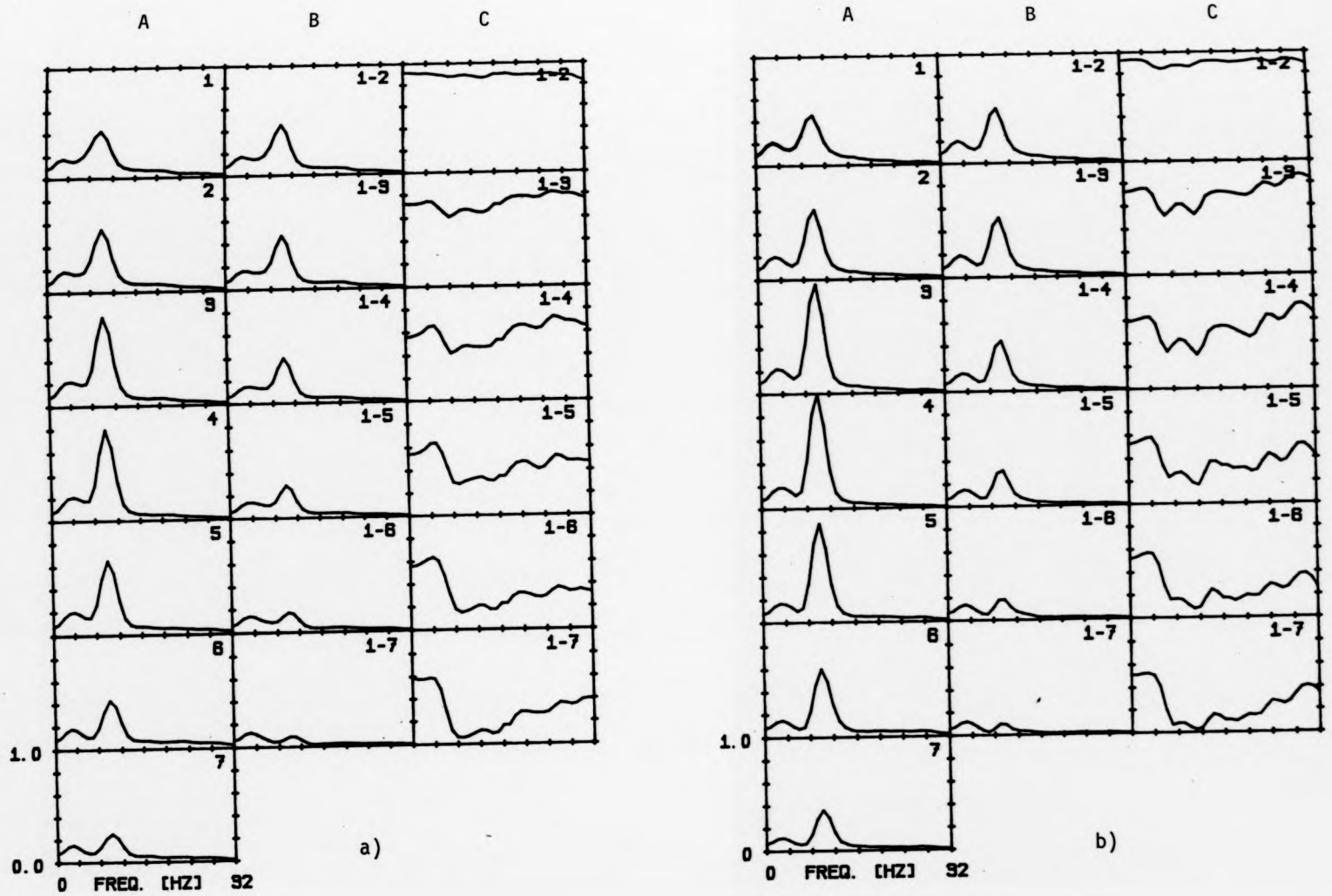



FIGURE 7.16 (A) Auto-spectra, (B) cross-spectra and (C) coherence functions computed for selected electrode combinations. Subject DAJ, a) stimulus A, b) stimulus I. Arbitrary units of spectral intensity.

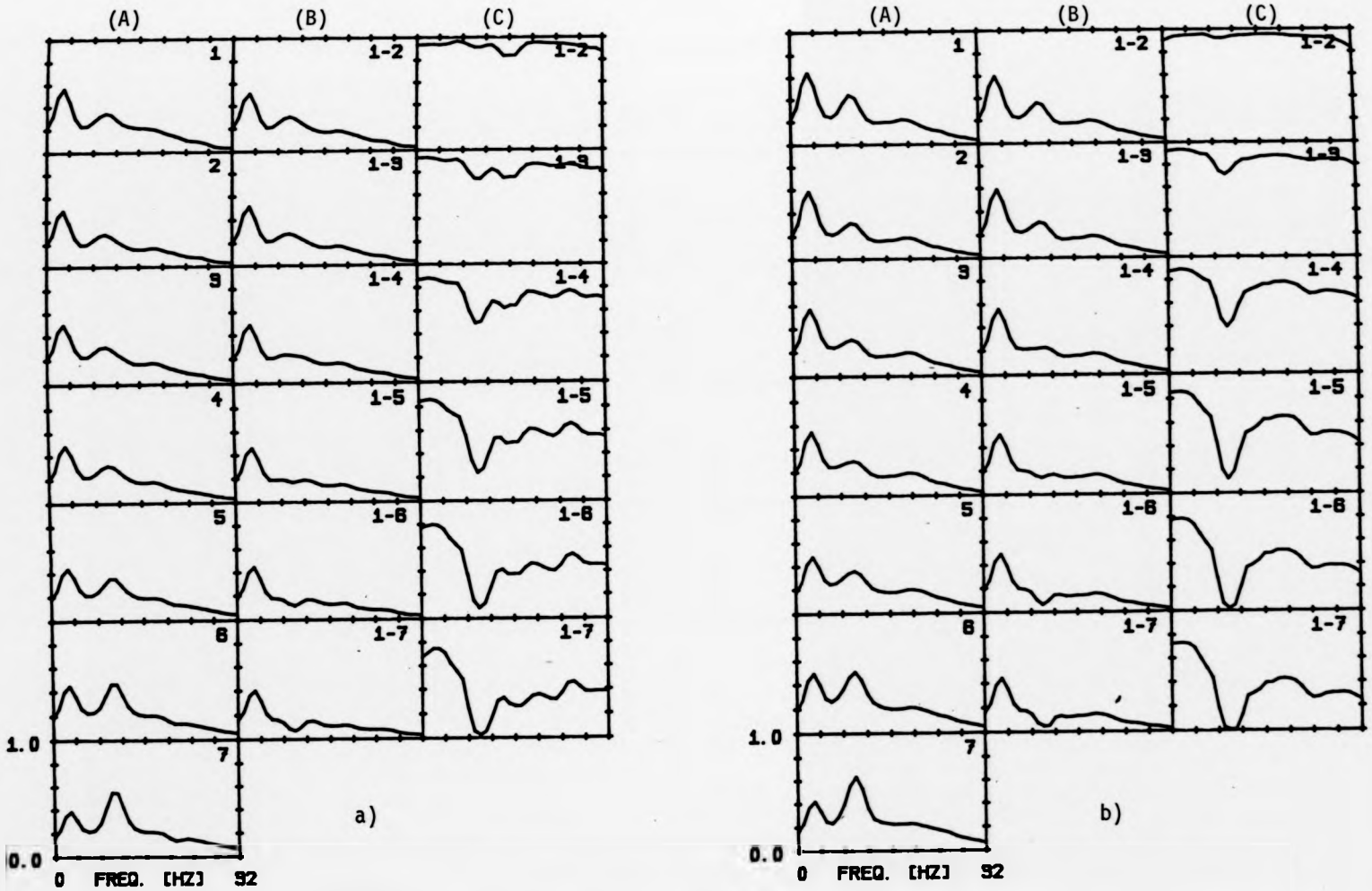



FIGURE 7.17 (A) Auto spectra, (B) cross-spectra and (C) coherence functions computed for selected electrode combinations. Subject MJM, a) stimulus A, b) stimulus I. Arbitrary units of spectral intensity.

The consistently high coherence values in the frequency range 0 - 6 Hz suggest that this activity is globally present over the scalp sites considered. This may be partly attributable to artifacts caused by motion, tape recorder transport irregularities etc. but is more likely to be caused by correlated neural activity for the reasons previously suggested. The slight decrease in coherence that occurs as electrode separation increases indicates a progressive decrease in the level of shared activity and/or a relative increase in local EEG or random noise activity. At frequencies above about 16 Hz the EEG activity is diminishing, and so poorer coherence might be expected due to the relatively greater influence of independent random noise. This is indeed observed in these results though significant correlations are still present at these frequencies.

These results can be used to predict the effectiveness of cancelling the correlated activity in different electrode pairs as the coherence function permits an assessment of the potential value of particular electrode choices; high coherence values indicate that good cancellation can be achieved. A coherence value of 0.5 leads to an improvement in SNR of 3dB, using expression B.12 in Appendix B. Much below this value, the returns are unlikely to justify use of this technique, though use of several reference electrodes may yield significant improvement even when the coherence is poor [103]. This will only occur if the correlated components in each reference are themselves relatively uncorrelated with each other, as otherwise there is little added information that may be used to gain further noise reduction.

## 7.3 Ramifications for adaptive noise cancelling of EPs.

The results of this study indicate that ANC may yield some improvement when applied in the manner originally proposed. Initially it had been thought that the EEG was well-correlated over extensive scalp areas, which is shown not to be the case. At the separations required to obtain an EP-free reference input the correlations are moderately good, but tend to be poor at alpha frequencies. Significant reduction of this activity may therefore not be possible by this method. Useful noise reduction is predicted at other frequencies, particularly at frequencies below about 6 Hz, which could lead to effective enhancement of low-frequency EP components such as the P300 component, though consideration of this or other slow wave phenomena has not been pursued in this thesis.

The coherence results restrict the choice of reference inputs to those which are relatively closely separated and which have markedly different EP contributions. Data from the two subjects indicate that this is not always possible and some distortion of the EP may be unavoidable. It is usually possible to satisfy these conflicting conditions for one or two components, for example left/right half-field stimulation appears to produce asymmetric transverse distributions for C1. This suggests that electrode pairs 3/5, 3/6 or 3/7 may be suitable for subject DAJ, and electrode pairs 3/7, 4/7 and 5/7 for subject MJM in this case. This demonstrates clearly the compromise that has to be made between noise reduction and signal distortion. A range of possible inputs are indicated in Table 7.1 with the corresponding signal improvement factor. Electrode pair 3/5 (DAJ) exhibits fairly high coherence at all frequencies, leading to a predicted improvement in SNR of 3 - 14 dB, but accompanied by some cancellation of EP components.

| Subject/ Stimulus | Electrode pair | EP component to be enhanced | Coherence | | Improvement in SNR (dB) | | approximate dynamic range |
|---|---|---|---|---|---|---|---|
| | | | min. | max. | min. | max. | |
| DAJ/I | 3/5 | C1 | 0.62 | 0.93 | 2.6 | 14.3 | 21 |
| | 3/6 | | 0.40 | 0.80 | 1.7 | 5 | 9 |
| | 3/7 | | 0.25 | 0.75 | 1.3 | 6 | 16 |
| DAJ/A | 6/7 | C2 | 0.75 | 0.95 | 6 | 13 | 15-21 |
| | 5/7 | | 0.50 | 0.90 | 3 | 10 | 15-21 |
| | 4/7 | | 0.35 | 0.75 | 1.5 | 6 | 15-21 |
| MJM/C | 4/7 | C1 | 0.20 | 0.85 | 1.3 | 6.7 | 15-30 |
| | 5/7 | | 0.65 | 0.90 | 2.8 | 10 | 16-36 |
| | 3/7 | | 0.10 | 0.80 | 1.1 | 5 | 16-25 |
| MJM/C | 3/6 | C1 | 0.15 | 0.85 | 1.2 | 6.7 | 16-25 |
| | 4/6 | | 0.45 | 0.90 | 1.8 | 10 | 15-30 |
| MJM/A | 4/6 | C2 | 0.60 | 0.90 | 2.5 | 10 | 7 |
| | 4/1 | | 0.35 | 0.85 | 1.5 | 6.7 | 18 |

TABLE 7.1 Possible electrode choices for adaptive noise cancelling of EP data. The table shows the maximum and minimum coherence estimates and hence the expected improvement in SNR. The eigen value spread is approximately equal to the dynamic range, which is estimated from the spectral intensity plots.

| Subject/ Stimulus | Electrode pair | EP component to be enhanced | Coherence | | Improvement in SNR (dB) | | approximate dynamic range |
|---|---|---|---|---|---|---|---|
| | | | min. | max. | min. | max. | |
| DAJ/I | 3/5 | C1 | 0.62 | 0.93 | 2.6 | 14.3 | 21 |
| | 3/6 | | 0.40 | 0.80 | 1.7 | 5 | 9 |
| | 3/7 | | 0.25 | 0.75 | 1.3 | 6 | 16 |
| DAJ/A | 6/7 | C2 | 0.75 | 0.95 | 6 | 13 | 15-21 |
| | 5/7 | | 0.50 | 0.90 | 3 | 10 | 15-21 |
| | 4/7 | | 0.35 | 0.75 | 1.5 | 6 | 15-21 |
| MJM/C | 4/7 | C1 | 0.20 | 0.85 | 1.3 | 6.7 | 15-30 |
| | 5/7 | | 0.65 | 0.90 | 2.8 | 10 | 16-36 |
| | 3/7 | | 0.10 | 0.80 | 1.1 | 5 | 16-25 |
| MJM/C | 3/6 | C1 | 0.15 | 0.85 | 1.2 | 6.7 | 16-25 |
| | 4/6 | | 0.45 | 0.90 | 1.8 | 10 | 15-30 |
| MJM/A | 4/6 | C2 | 0.60 | 0.90 | 2.5 | 10 | 7 |
| | 4/1 | | 0.35 | 0.85 | 1.5 | 6.7 | 18 |

TABLE 7.1 Possible electrode choices for adaptive noise cancelling of EP data. The table shows the maximum and minimum coherence estimates and hence the expected improvement in SNR. The eigen value spread is approximately equal to the dynamic range, which is estimated from the spectral intensity plots.

Electrode pairs 3/6 and 3/7 introduce less distortion but offer less noise reduction. In the next chapter results will be presented showing the application of an ANC using some of these electrode combinations.

## 7.4 Summary.

In this chapter evidence has been presented that justifies the application of adaptive noise cancelling in selected situations to evoked potential data. Methods have been described to enable identification of suitable electrode sites for this purpose. Studies of two subjects have shown that it is difficult to derive reference electrodes which simultaneously have high EEG correlations (necessary for effective cancellation) and negligible EP contributions (necessary for distortion-free cancellation. A compromise must therefore be made between noise cancellation and signal fidelity in these cases. The application of an ANC to specific EP data records will be pursued in the next chapter. This will also consider practical questions regarding the proper application of ANC to the data such as the selection of the appropriate filter from the alternatives available, selection of suitable filter parameters and the procedures necessary to precondition the signal. The conclusions regarding the effectiveness of ANC will be summarised in the final chapter.

# CHAPTER EIGHT

# THE APPLICATION OF ADAPTIVE NOISE CANCELLING TO EVOKED POTENTIAL DATA.

## 8.0 Introduction

In the previous chapters the case for using an ANC to process EP data has been presented. In particular it was demonstrated that signal enhancement can be expected provided that electrode choice is made carefully on the basis of the scalp topography of the EP to various stimuli. In this chapter a series of studies is described that attempts to identify the best way of applying this technique in practice. Firstly consideration is given to the choice of filter to be used from the range of alternatives available. This is followed by a discussion of filter parameter selection, though a more detailed investigation of the effects of different parameter values is presented in a later section. As the presence of d.c. offsets in the signal can degrade filter performance, a study was made to identify an effective way of overcoming this difficulty. Following this a comparison is made of gated adaptive filtering (described in chapter 5) with the basic AF using EP data. The results support the use of a GAF rather than the basic AF for EEG signals when fast adaption is employed.

Having established the best way of applying the ANC, results of filtering a wide range of data records are presented. The effectiveness of the method is discussed and compared with elementary processing operations. Both qualitative and quantitative conclusions are drawn which show that considerable improvement in SNR is attained compared with the unfiltered records, and especially so when they contain sporadic alpha activity. The results are variable when severe

contamination of alpha activity is present, and the method may require further development. The ramifications of this work and recommendations regarding adoption of this method in research and clinical settings are discussed in the concluding chapter.

## 8.1 Choice of Adaptive Filter.

Some of the factors governing the choice of AF have already been described in Chapter 4. These include the required filter characteristics (such as filter order and adaption properties), signal properties (such as dynamic range and stationarity) and implementation aspects (such as processing speed and cost). They will be referred to in the following discussion with little further elaboration.

The LMS AF was used as the basis for these investigations for a number of reasons. As discussed previously this algorithm was attractive due to its well-established properties such as unconditional stability, ease of implementation and computational efficiency. These were important in a research project which intended to explore new applications and had therefore to employ robust techniques. The simplicity of implementation had the advantage of faster development, and thus allowed initial results to be obtained more quickly. The computational speed is superior to most other methods and this permitted faster processing and evaluation. Only the frequency-domain methods are faster, but there are restrictions on their use and significant computational savings are only achieved for moderate to high filter orders. A further consideration regarding the choice of the LMS AF was that it is easily implemented in a hardware processor were the investigation to support adoption of this technique.

One of the most important factors influencing filter choice is the degree to which the signal statistics are stationary. If the time-variation is relatively slow, then it is possible for the adaptive filter to track these, albeit with some delay. Previous studies have reported that the EEG may be considered to be stationary over short intervals of about 12-25 s, as discussed in Chapter 2. While this is borne out by visual inspection of these EEG data, it seems clear that alpha activity is rather non-stationary as it can appear in spindles or bursts of activity that last only a few seconds. In addition it is relatively uncorrelated over the scalp, and is best regarded as being uncorrelated noise. For these reasons it is not easy to predict how well the technique will perform when applied to EEG data, and an experimental investigation is required.

For the most useful electrode-stimulus combinations indicated in Table 7.1 the dynamic range (defined as the ratio of maximum to minimum signal power over the chosen bandwidth) is in the range 10-40, and hence the eigenvalue spread for the EEG is in this range also, as noted in chapter 7. Though this is a little high for most effective filtering using the LMS algorithm, it is nevertheless worth pursuing since reasonably fast convergence is still possible. This can be seen by considering equation 4.38 and substituting 0.2 for M, which yields 40 sample periods as the average time constant for a filter of order 32. This corresponds to a time constant of about 0.3 s for data sampled at 128 Hz. Some improvement in adaption speed might be possible through the use of faster converging algorithms, though it is not certain that significant gains would be achieved if the degree of nonstationarity were high. The initial preference therefore lay with the LMS algorithm.

Having selected this algorithm, a number of choices remained to be made. One of these regarded the possible use of FFT procedures to obtain faster processing. This was not pursued for two reasons. Firstly the use of the FFT is only indicated when the filter order is moderately large, e.g. $N > 32$ [99]. This was unlikely to be very significant in the current application as later studies showed that suitable values for N lay in the range 16 - 64. Secondly, the use of FFT procedures implicitly demands that block processing be employed (i.e. the filter weights are held constant throughout each block, and updated at the end of each block). Though this has been shown to be equivalent to continuous adaptive filtering when using stationary data, it has not been justified for nonstationary data nor when the ability to track time-varying parameters continuously is required. More seriously, block processing cannot achieve fast adaption compared with direct methods because the stability criterion is more restrictive by a factor equal to the block length L (section 4.3.1). Direct computation of the transversal filter output and weight update calculations was therefore performed.

The remaining choices regarding filter structure are part of the study of this thesis. Comparison of ANC performance using AF, TSAF and GAF structures are presented in a later section. Some studies are first described which were intended to establish the optimum manner of use. This included the selection of parameter values and appropriate baseline removal methods.

## 8.2 Initial Selection of Filter Parameter Values.

The factors affecting the settings of the LMS filter parameters and the effect of these upon filter performance have been discussed in

some detail in chapter 4. The two main choices concern the filter order N, which must be large enough to filter the lowest frequency components, and the convergence parameter $\mu'$, which must be chosen to obtain accurate and fast convergence without the risk of instability arising. In this section I shall describe how initial values for these parameters were chosen. A more detailed consideration of the effects of different values is included within the discussion of different filter types in sections 8.5.1 and 8.5.2. As the transversal filter implements a moving average model of the desired data sequence based on the reference data sequence, the filter must be large enough to accommodate all significant crosscorrelation coefficients if it is to filter all desired low frequency components. This means that the filter order must correspond approximately to the effective duration of the cross-correlation function.

A useful way to determine the initial value of this parameter is to compute the sample transfer function from typical data records, and obtain the resulting impulse response from the inverse Fourier Transform. The transfer function approximates the Wiener filter response for stationary data

$$W^{o}(f) = | G_{xy}(f) | / G_{xx}(f) \qquad (8.1)$$

and may be computed using the procedures described in Appendix B. By this method normalised impulse responses corresponding to the cross-channel transfer function were obtained for several 1 s records for both subjects. These are shown in figs. 8.1 and 8.2 for selected electrode combinations. Fig. 8.3 shows the impulse response obtained by using averages of 8 spectra. The plots in figs. 8.1 and 8.2 exhibit greater statistical variability due to the use of short data records, but were obtained to indicate the degree to which the transfer function might vary from record to record. In general the individual impulse

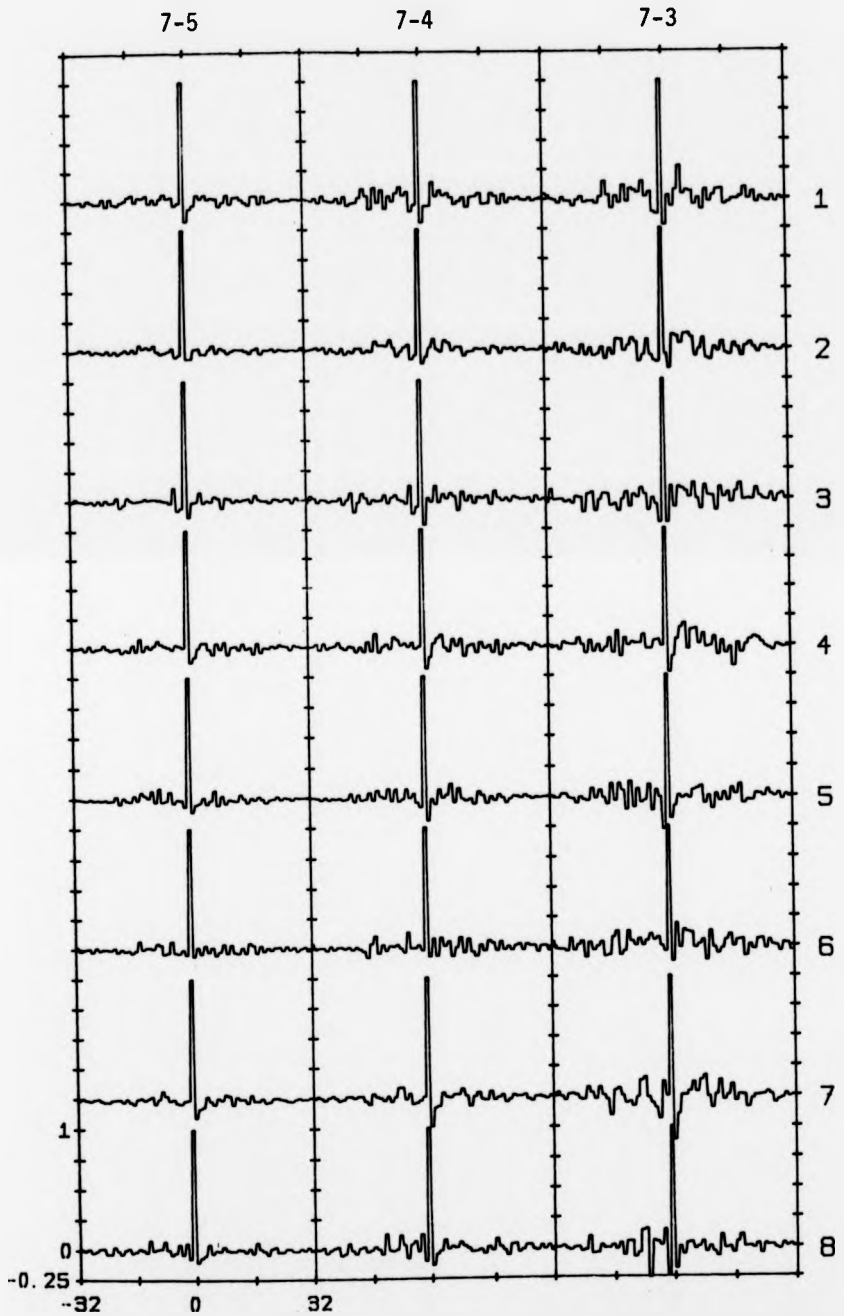



FIGURE 8.1 Normalised impulse responses representing the transfer function between the correlated EEG components in the indicated channels. 8 responses shown, computed for each of 8 1s records. Lag times are in units of the sampling period T=7.8ms. Subject MJM, stimulus C.

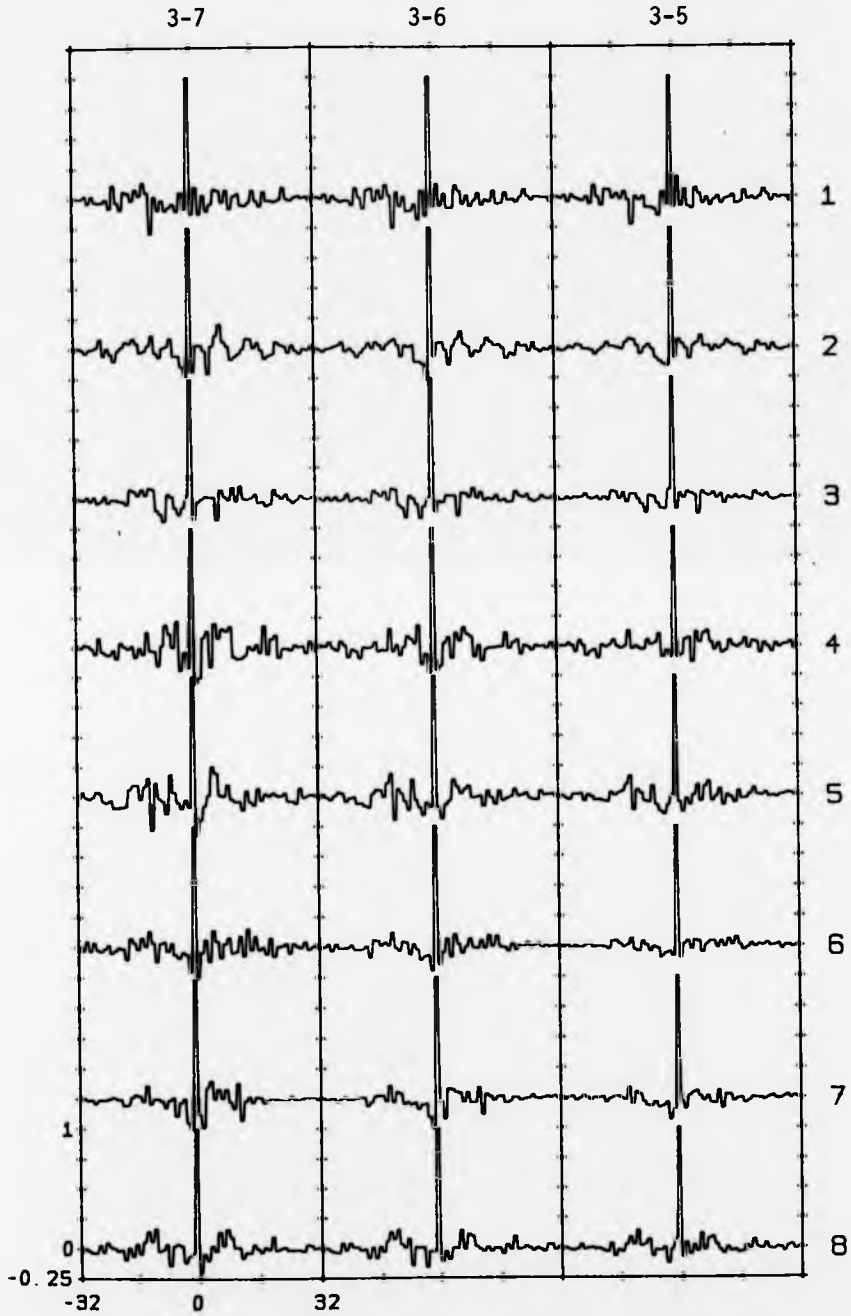



FIGURE 8.2 Normalised impulse responses representing the transfer function between the correlated EEG components in the indicated channels. 8 responses shown, computed for each of 8 1s records. Lag times are in units of the sampling period T=7.8ms. Subject DAJ, stimulus I.

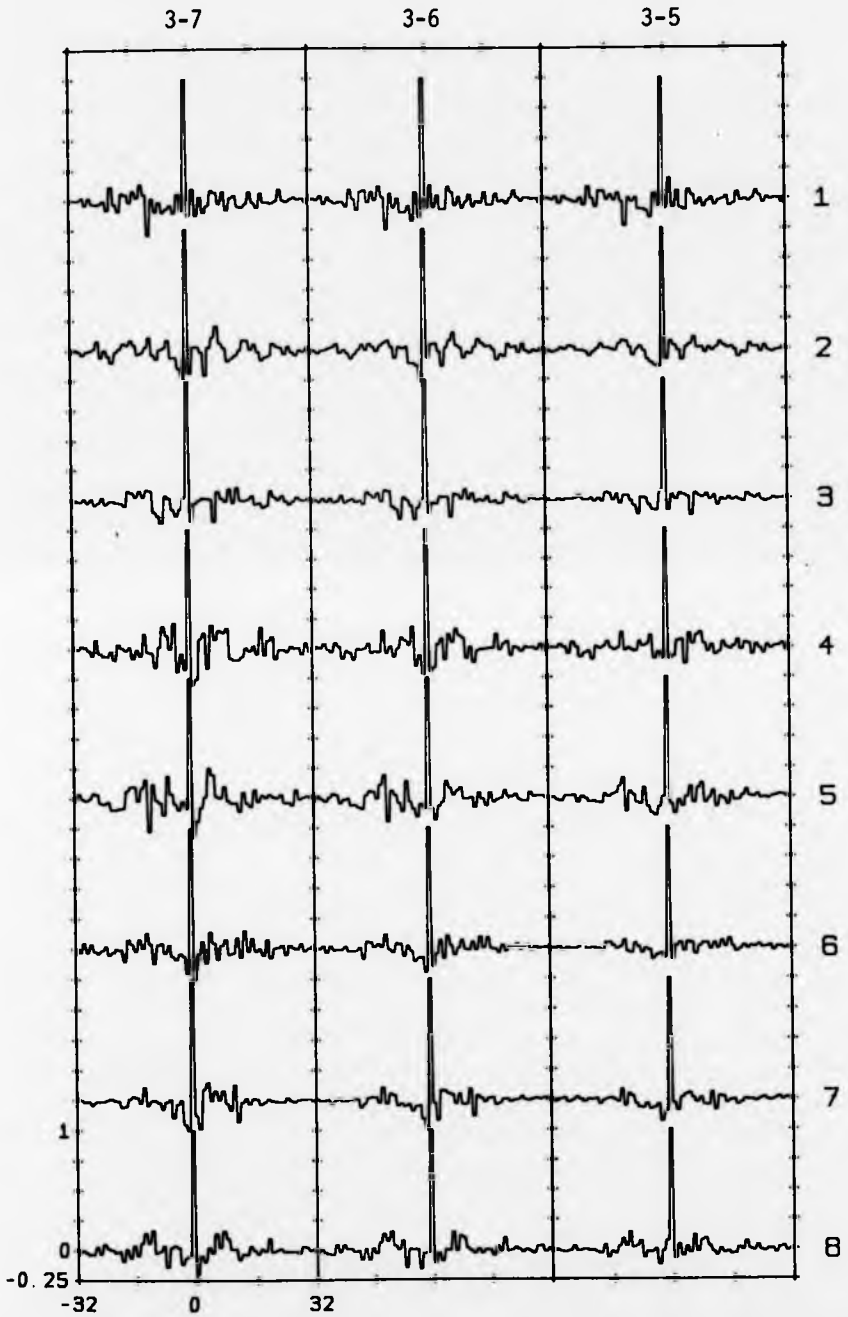



FIGURE 8.2 Normalised impulse responses representing the transfer function between the correlated EEG components in the indicated channels. 8 responses shown, computed for each of 8 1s records. Lag times are in units of the sampling period T=7.8ms. Subject DAJ, stimulus I.

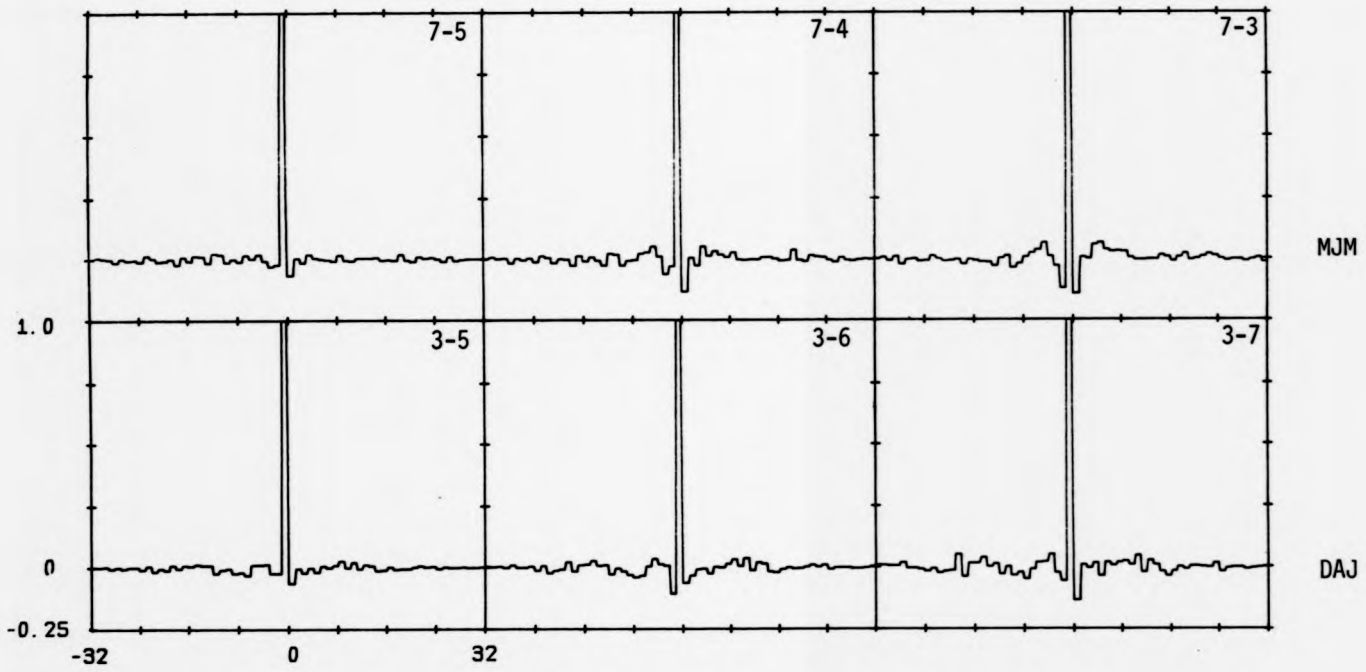



FIGURE 8.3 Normalised impulse responses, computed for an 8s record for subjects DAJ and MJM, and for the indicated channel pairs. Lag times in units of 7.8 ms.

responses are similar to the average, but occasional differences are evident as might be expected. For the purpose of setting the filter order it is likely that values of N in the range 32-64 are satisfactory.

The value of $\mu'$ was also obtained experimentally, as it is difficult to determine suitable values in practice using the expressions derived in chapter 4. As it was anticipated that fairly rapid convergence would be required in order to track time-varying signal properties, a fairly high value of $\mu'$ is needed. Values of the normalised convergence parameter $\mu'$ were chosen in the range 0.2 to 0.02. The upper value was experimentally found to be the largest that could be routinely employed without the risk of instability arising during adaption. An order of magnitude range was considered to be sufficient to characterise the range of useful values.

## 8.3 Baseline drift removal.

It has been noted in chapter 4 that d.c. components present in either record can seriously affect filter accuracy and can even lead to a divergent solution. A number of procedures that can be employed to avoid these effects will now be discussed, and results of a quantitative comparison presented.

For off-line analysis purposes it is possible to precondition the data before applying the AF, (though this is not possible if continuous on-line filtering is to be performed). Several methods were considered, such as removing the linear trend and mean value from the entire record. The latter is only partly effective as it cannot satisfactorily deal with a fluctuating baseline, which commonly occurs. A more effective way of dealing with this eventuality is to remove the

d.c. component and linear trend from successive short segments of the data record. This was attempted using 1 s data segments, and resulted in much smaller baseline deviation about the mean. The results are considerably better than when no preconditioning measures were used, and largely overcame the problem. This method was therefore used in some of the early work. Experience has shown however that certain difficulties arise through use of this method. Because each data segment is processed independently, it is possible for discontinuities to arise at the segment boundaries. This is illustrated using an artificial signal in fig. 8.4. Not only does this method cause artificial discontinuities but it also is not suitable for on-line processing. Another approach was therefore pursued.

The solution proposed by Ferrara [96] is to employ an additional tap weight in the adaptive filter that adaptively tracks and removes the running mean. The bias weight causes the filter to act as a simple notch filter at d.c., having the following form:

$$G(z) = \frac{z - 1}{z - b} \qquad (8.2)$$

The parameter b is related to the convergence coefficient $\mu$ [92]. The magnitude response of this filter is shown in fig. 8.5 for two values of the parameter b. Widrow et al [92] reported that effective d.c. removal is also obtained when the bias weight is used in combination with the other weights, although this has not been rigorously justified. Though this modification is simple to apply and involves little extra computational overhead, later results show that this is still not the most satisfactory solution, primarily because the magnitude response is inadequate to effectively remove low frequency baseline components. The use of the bias tap should therefore be restricted to cases where it is known that the d.c. component does not vary significantly.

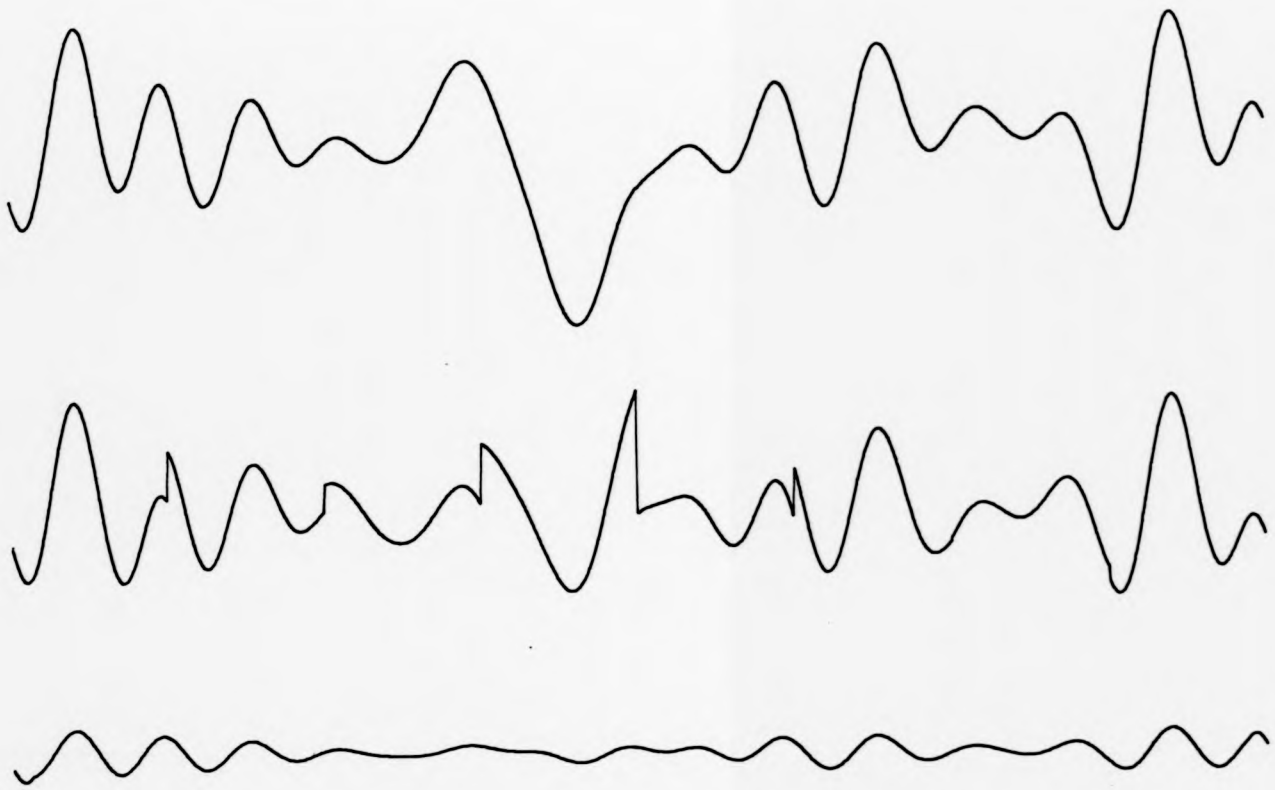


FIGURE 8.4 Comparison of two baseline removal methods on a computer-generated segment of noise of bandwidth 0-2 Hz.
a) original record
b) removal of dc and linear trend from each 1s segment
c) application of a 2nd order recursive high pass filter with M=33.

The best solution, (adopted in most of the following studies), is to employ a simple recursive high pass filter prior to the AF. This has the advantage of being readily incorporated within a continuous filter process, and may be configured optimally to the data requirements by use of an appropriate filter characteristic. The filter used in practice is similar to that described in Lynn [134]. It belongs to a class of digital filters of moving-average form that can be implemented recursively, with consequent computational savings. A second-order low-pass filter was used, having the following transfer function:

$$G(z) = \frac{(1 - z^{-M})^2}{(1 - z^{-1})^2} \tag{8.3}$$

This has a simple recurrence relation, which is given by

$$y(n) = 2y(n-1) - y(n-2) + x(n) - 2x(n-M) + x(n-2M) \tag{8.4}$$

The high pass output is formed by simply subtracting the low pass output from the input signal, with a delay in the input of M samples to offset the delay in the moving average filter. The resulting filter has zero-phase response. Its magnitude response is shown in fig. 8.5 for two values of M. The value of M=33 was experimentally found to be effective in removing low frequency baseline components without significantly distorting the EP.

Each of these baseline removal methods were applied to 8 data records that were subsequently filtered with an ANC (with N=32 and $\mu'$=0.2). The plots in figs. 8.6 and 8.7 are two records that show the worst case differences obtained in this study. In addition averaged responses for each record are shown in fig. 8.8 with the grand average of 64 subtracted responses superimposed to aid comparison. In figs. 8.6 and 8.7 the subtracted input signals are paired with the corresponding ANC output, and four pairs are plotted in each figure as follows:

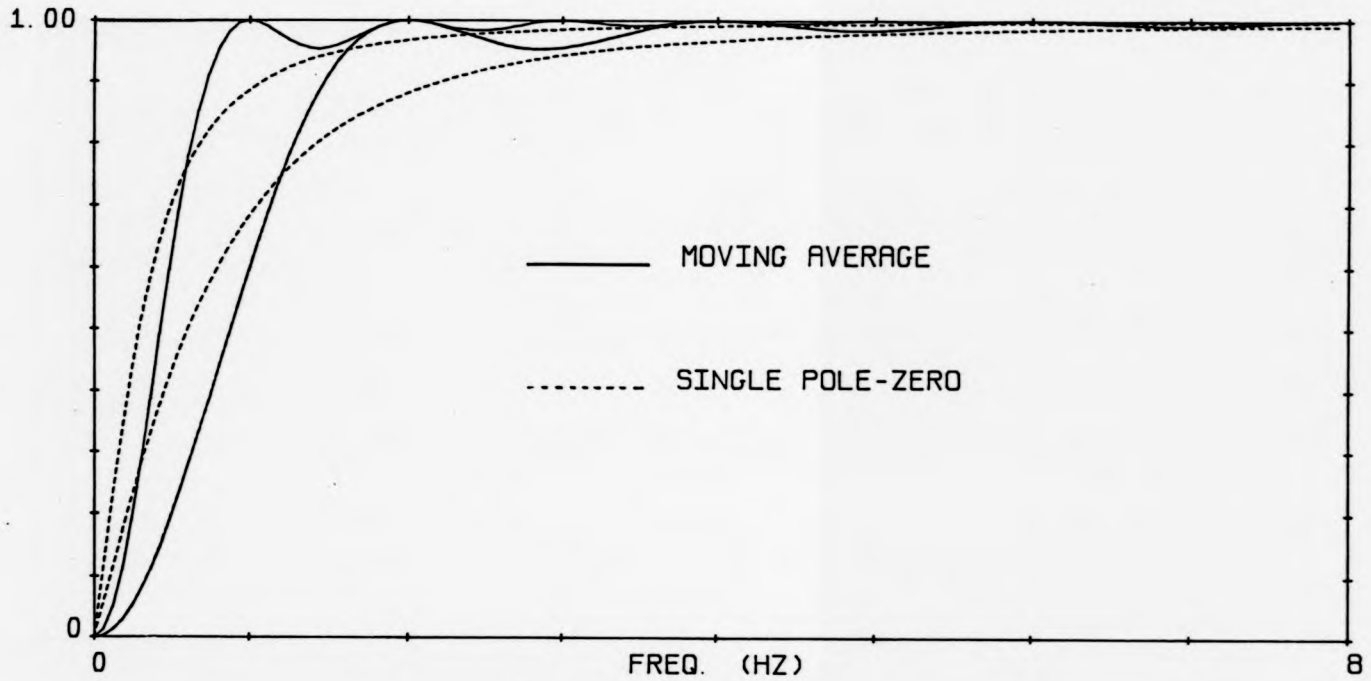



FIGURE 8.5 Magnitude response characteristics of recursive digital high pass filters
a) 2nd order moving average filter of order M=65 and M=33
b) 1st order notch filter with zero at z=1 and pole at b=0.95 and b=0.9.

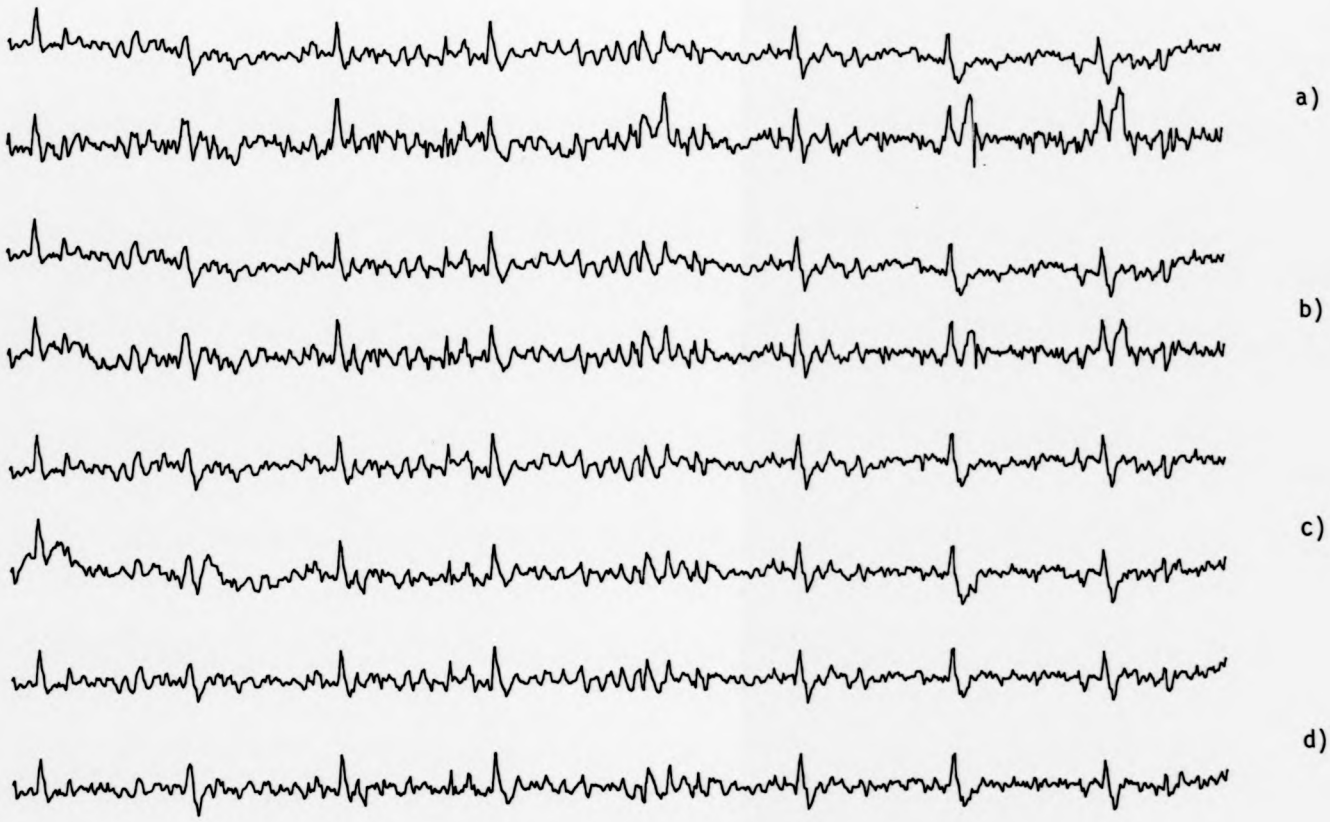



FIGURE 8.6 Comparison of different baseline removal methods, showing input signal and output of ANC for one record, with
a) no baseline removal
b) bias tap employed in ANC
c) linear trend removal of 1s segments
d) 2nd order recursive digital filter.

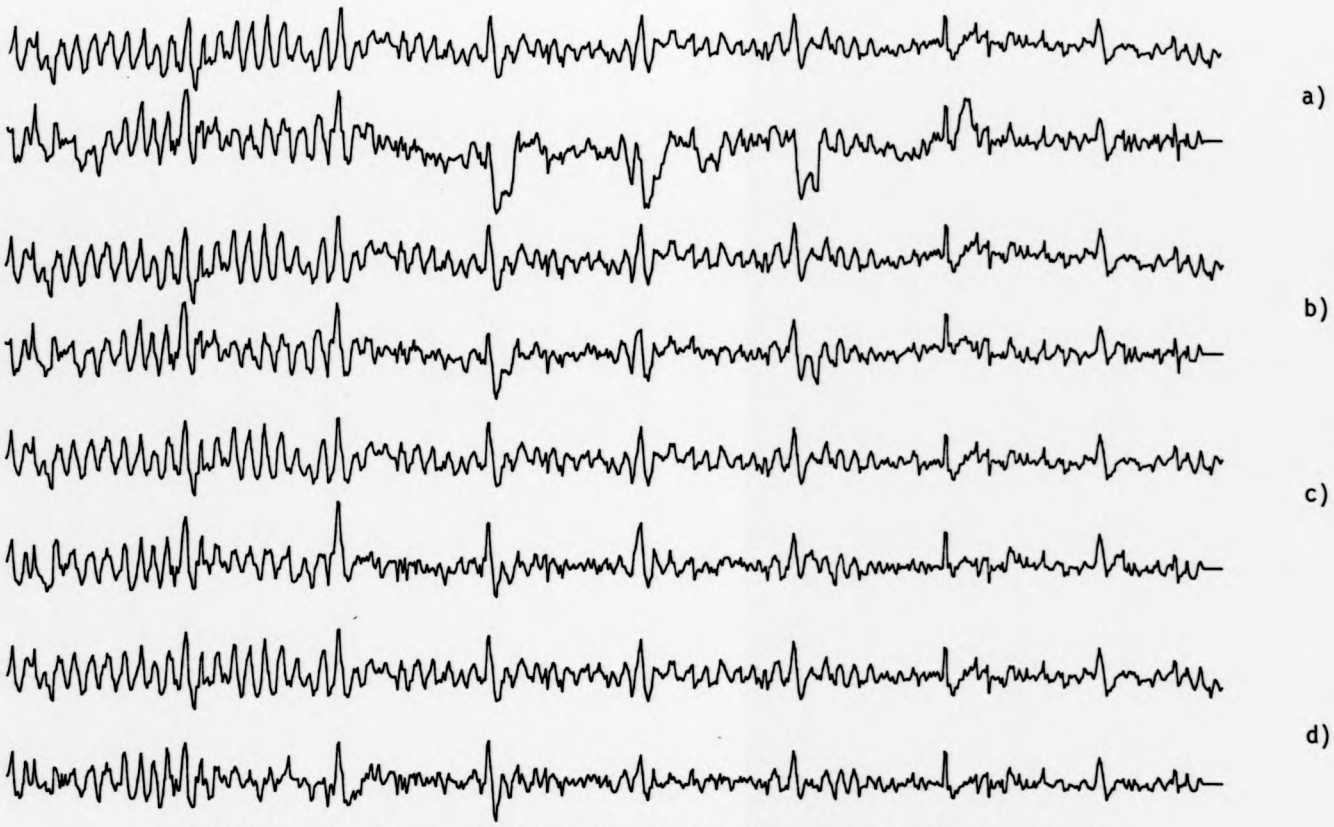



FIGURE 8.7 Comparison of different baseline removal methods, showing input signal and output of ANC for one record, with
a) no baseline removal
b) bias tap employed in ANC
c) linear trend removal of ls segments
d) 2nd order recursive digital filter.

(i) no baseline removal

(ii) bias tap employed in the adaptive filter

(iii) linear trend removal applied to each 1 s segment of data, and

(iv) second order high pass recursive filter, with M=33

Examination of these yields the following conclusions:

a) baseline fluctuations seriously degrade the effectiveness of adaptive noise cancelling. The effect of local d.c. offsets was to distort individual EPs and obtain poorer noise cancellation in the inter-signal data segments, with the likelihood of artifactual activity arising from poorly adapted weights. This is most clearly seen in the averaged responses of the filter outputs. The use of some baseline removal method was therefore clearly necessary.

b) The different methods that were considered achieve similar results, though the use of a bias tap gave poorer results in general. It was also more susceptible to instability, which occurred in one record following an unusually large transient offset.

c) the linear trend removal method was about as effective as the recursive notch filter, though the main peaks of the averaged response show greater amplitude variations than the recursively filtered averaged responses.

The use of the high pass filter thus appeared to give the best results most consistently. This was verified by measuring the rms disparity between each average response and the grand average, and by measuring the rms amplitude of the inter-signal segments for each record. These results are shown in Table 8.1 and support the conclusion that this method was the best of the alternatives considered.

(a)

| record | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | overall average |
|---|---|---|---|---|---|---|---|---|---|
| detrending method | | | | | | | | | |
| (A) | 60 | 66 | 90 | 44 | 111 | 53 | 71 | 85 | 78.4 |
| (B) | 46 | 55 | 50 | 34 | 79 | 48 | 64 | 55 | 55.3 |
| (C) | 41 | 50 | 35 | 32 | 65 | 47 | 55 | 49 | 47.8 |
| (D) | 40 | 49 | 31 | 33 | 64 | 47 | 54 | 49 | 47.0 |

(b)

| record | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | overall average |
|---|---|---|---|---|---|---|---|---|---|
| detrending method | | | | | | | | | |
| (A) | 132 | 110 | 1004 | 92 | 116 | 83 | 96 | 135 | 369.7 |
| (B) | 71 | 80 | 321 | 55 | 75 | 55 | 78 | 74 | 131.2 |
| (C) | 54 | 64 | 38 | 41 | 79 | 53 | 58 | 59 | 57.0 |
| (D) | 48 | 63 | 31 | 36 | 72 | 51 | 51 | 55 | 52.4 |

TABLE 8.1. RMS values of (a) inter-signal epochs and of (b) disparity between single EPs and average EP in 8 records after adaptive noise cancelling, following the use of different detrending methods: (A) no detrending (B) bias tap in AF (C) linear detrending of 1s segments (D) recursive digital filter.

## 8.4 Comparison of Adaptive Filtering and Gated Adaptive Filtering.

In chapter 5 the rationale for employing a modified filter, known as the gated adaptive filter (GAF) was presented. The analysis suggested that improved performance is obtained over both the basic AF and the time-sequenced AF if the SNR is relatively poor and/or fast convergence is required. These proposals were first tested and verified using artificial data generated by computer. In this section some results are presented of a comparison between AF and GAF using a selection of EP data that represent a broad range of activity. Four particular classifications of EP activity were considered:

a) high SNR, representing the best quality data obtained, with good EP and low-level EEG interference
b) moderate SNR, with levels of interference due to EMG and alpha activity representative of typical data records
c) poor SNR, with relatively high EMG activity, and
d) poor SNR, with relatively high alpha activity.

Four randomly selected records having these characteristics were processed. The results of the comparison are shown in figs. 8.9 - 8.11. In each case N was chosen to be 32 and $\mu'$ to be one of 0.2, 0.05 and 0.02 respectively. The average EPs for each of these cases are shown in fig. 8.12 where the GAF output is shown dotted and the AF output as a solid line.

These results confirm the general conclusion that the GAF is more suitable than the basic AF when high adaption rates are required. Examination of the filtered records in each case reveals that the EP is larger in the GAF output than the AF output, particularly for the larger

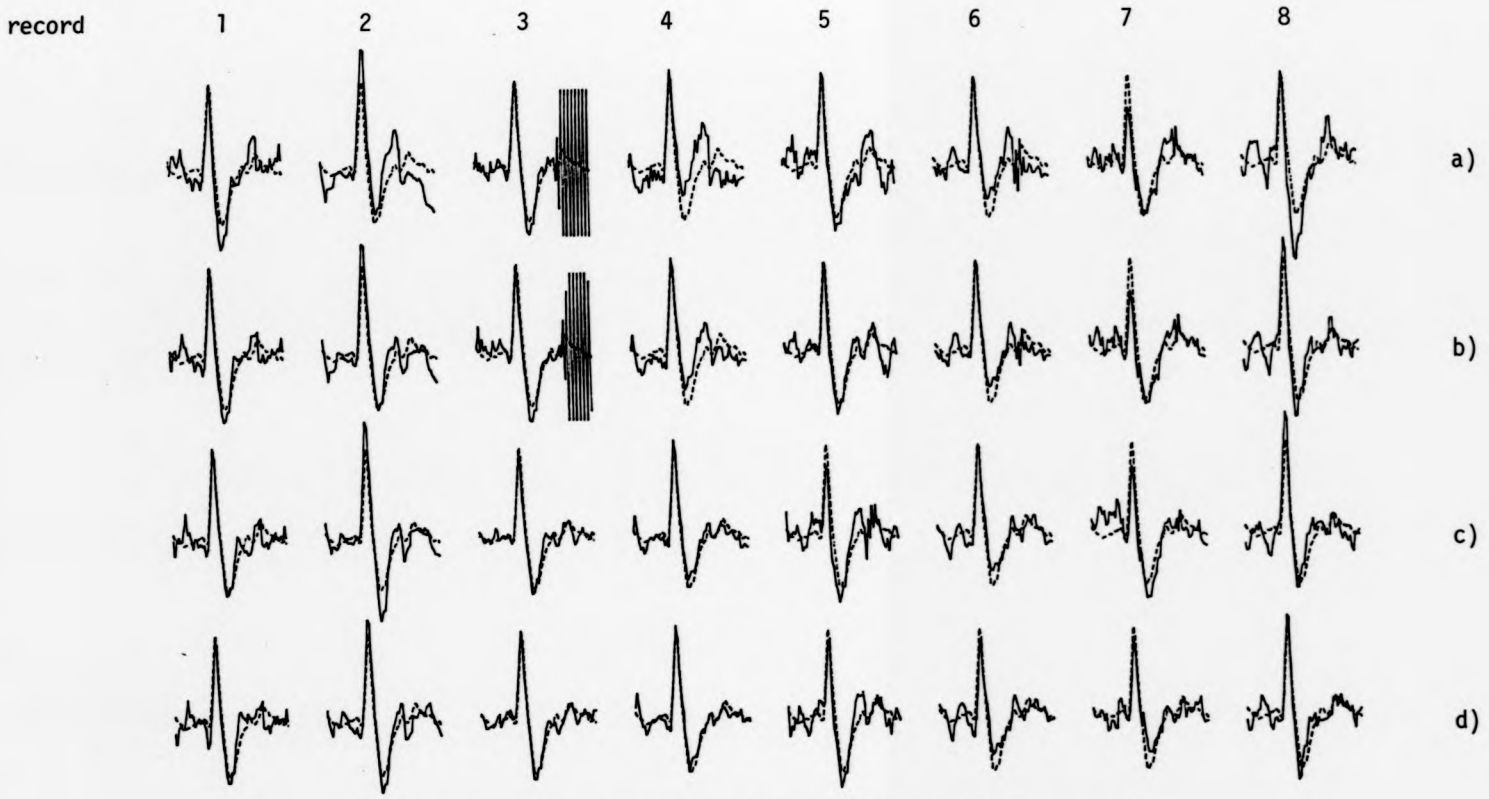



FIGURE 8.8 Comparison of different baseline removal methods. Averages of 8 responses shown for 8 consecutive records filtered using an ANC with
a) no baseline removal
b) bias tap employed in ANC
c) linear trend removal of 1s segments
d) 2nd order recursive digital filter.

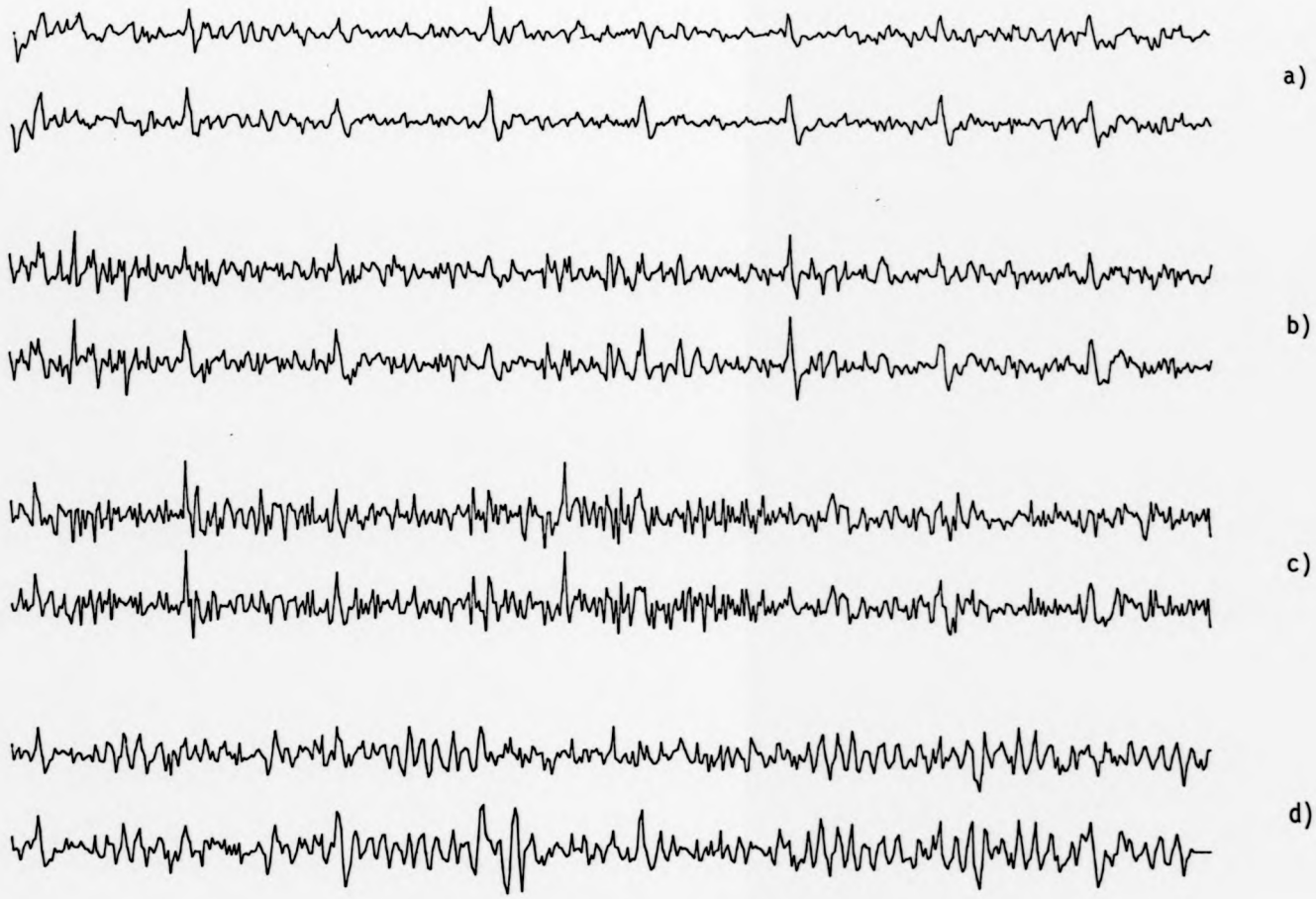



FIGURE 8.9 Comparison of AF and GAF applied to 4 different EP records, with fast adaption ($\mu'$=0.2, N=32)
a) low level EEG activity
b) moderate EEG activity
c) high level EMG activity
d) high level alpha activity

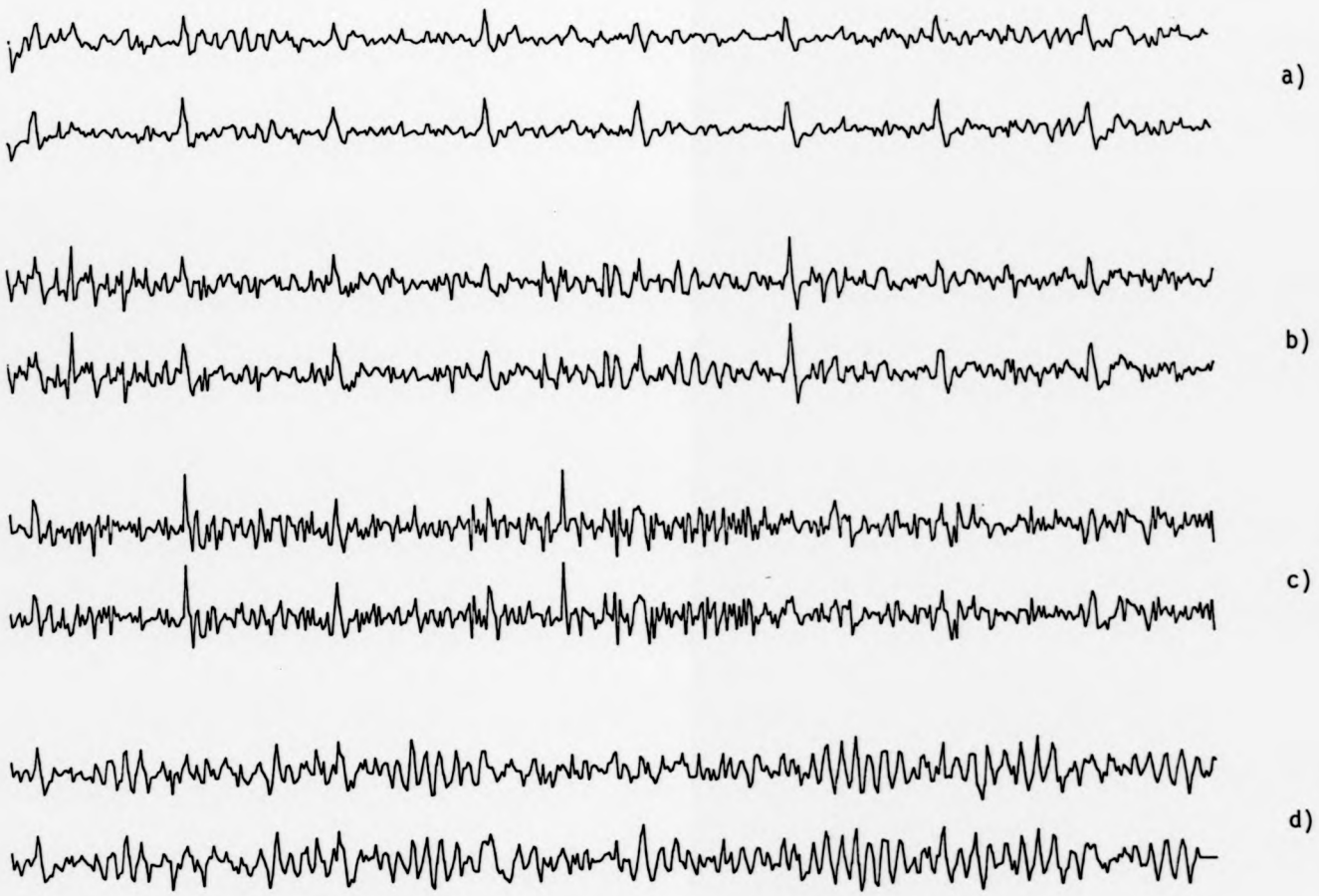



FIGURE 8.10 Comparison of AF and GAF applied to 4 different EP records, with moderately fast adaption ($\mu'$=0.05, N=32)
a) low level EEG activity
b) moderate EEG activity
c) high level EMG activity
d) high level alpha activity

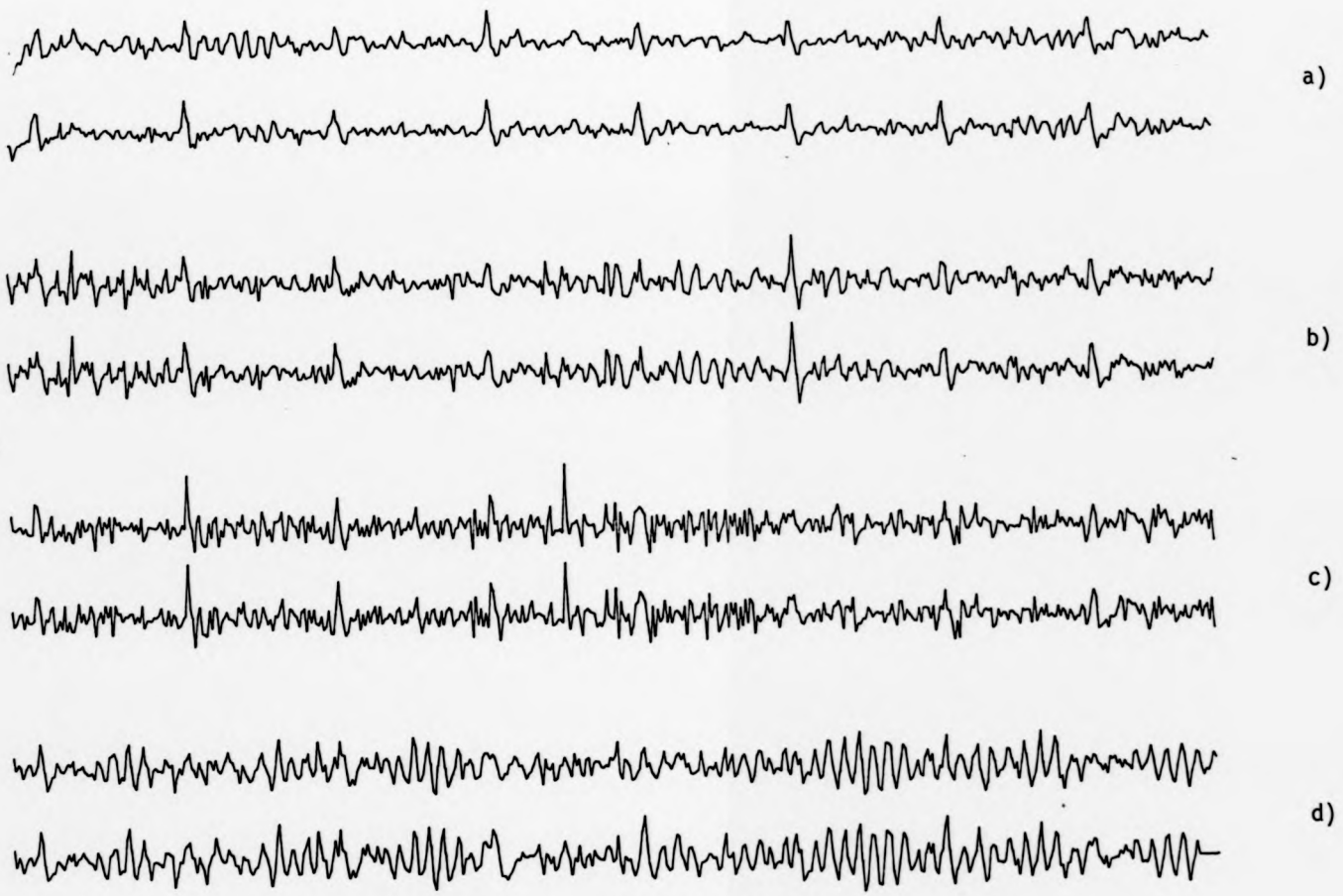



FIGURE 8.11 Comparison of AF and GAF applied to 4 different EP records, with moderately slow adaption ($\mu'$=0.02, N=32)
a) low level EEG activity
b) moderate EEG activity
c) high level EMG activity
d) high level alpha activity

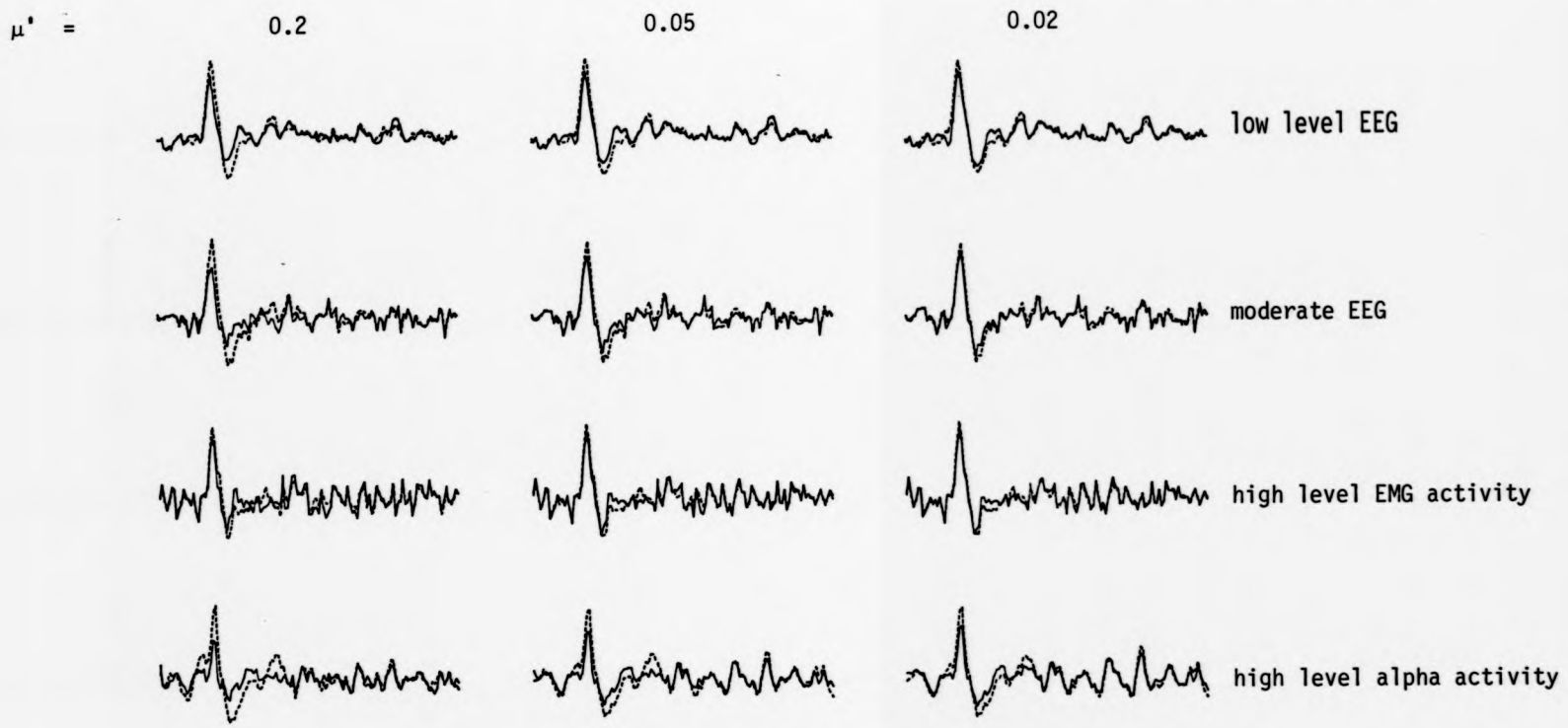



FIGURE 8.12 Comparison of AF and GAF average outputs following filtering of 4 EP records using 3 adaption rates. (AF outputs shown solid, GAF outputs shown broken.)

values of $\mu'$, though the filters perform similarly otherwise. This is most clearly seen in the averaged responses shown superimposed for the two filters in fig. 8.12. In each case the average EP amplitude is smaller using the AF than using the GAF, which supports the hypothesis that attenuation of the EP occurs due to adaption in the EP epoch. The GAF AEPs on the other hand seem to be relatively unaffected by the value of $\mu'$. This is particularly evident in the later components of the EP. The activity during the inter-signal segments of the data is very similar for both filter types. As well as causing less distortion of the individual EPs in the GAF records, it appears that noise cancellation can be more effective in the inter-signal interval. This is most clearly seen in the first pair of records in fig. 8.9. Similar results obtained with subject DAJ also indicate that better overall cancellation is obtained using the GAF than the AF.

## 8.5 Adaptive noise cancelling of EPs using a gated adaptive filter.

In the preceding sections of this chapter it has been established that a gated adaptive filter is likely to be more appropriate than the basic adaptive filter for cancelling correlated EEG signals in EP data. Filtering can be conveniently accomplished using the LMS adaptive algorithm, and previous results indicate the approximate parameter values to be used. A study identified suitable methods of removing undesired baseline trends from the data prior to filtering.

In this section the results of filtering several different records from each subject are analysed to determine the effectiveness of filtering EP data by this means. The study includes a comparison of the effects of different parameter values, as these are difficult to specify

a priori. The results of these approaches are presented for direct visual evaluation with corresponding quantitative measures in some cases. Two particular cases are investigated in detail. These are:

a) data records from subject DAJ and experiment I, using electrodes 3 and 5 as the primary and reference inputs respectively, (discussed in section 8.6.1) and
b) data records from subject MJM and experiment C, using electrodes 6/4 and 7/4 (discussed in section 8.6.2).

The data set includes records that encompass a wide range of activity types. These cases were selected on the basis of the investigation in chapter 7, which suggested the most useful electrode/stimulus combinations.

### 8.5.1 Subject DAJ.

Early exploratory work showed that good results were obtained with a filter order of about 32 and a high value of $\mu'$, about 0.2. Higher values of $\mu'$ led to the occurrence of instability in the filter and this value was considered a safe maximum. Each of these choices is now considered in more detail.

#### 8.5.1.1 The effect of $\mu'$ upon ANC performance

The first study was intended to establish suitable values for the convergence parameter $\mu'$. N was set to 32 and $\mu'$ values of 0.2, 0.05 and 0.02 were used with a GAF. 8 consecutive data records were filtered under each of these conditions and the results for four of these are shown in figs. 8.13 - 8.16 with the filter inputs and subtracted inputs. D.c. offsets had been removed by the method of

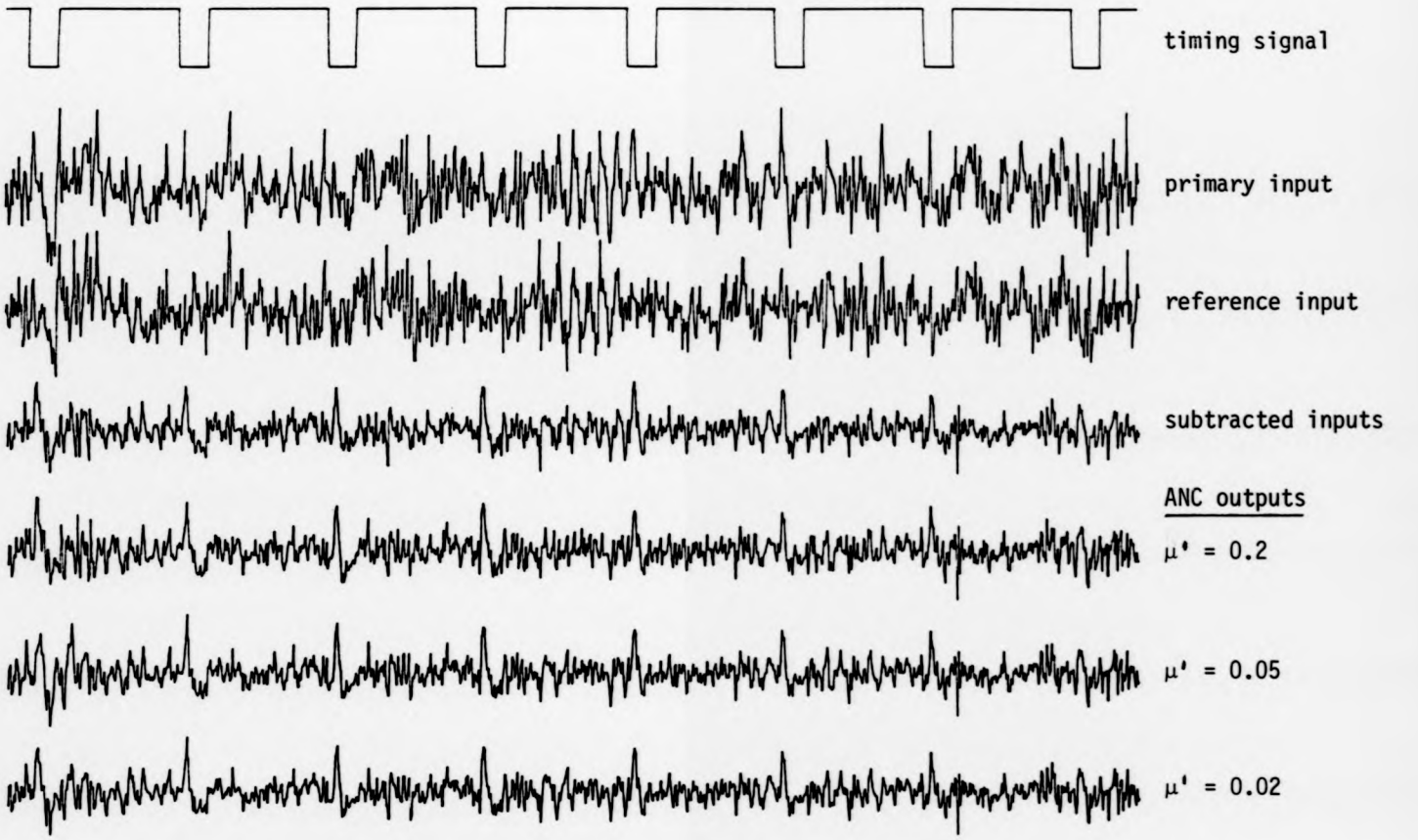



FIGURE 8.13 A comparison of the effect of convergence rate upon adaptive noise cancelling of EP data.
(Data from subject DAJ/I, channels 3 and 5, N=32)

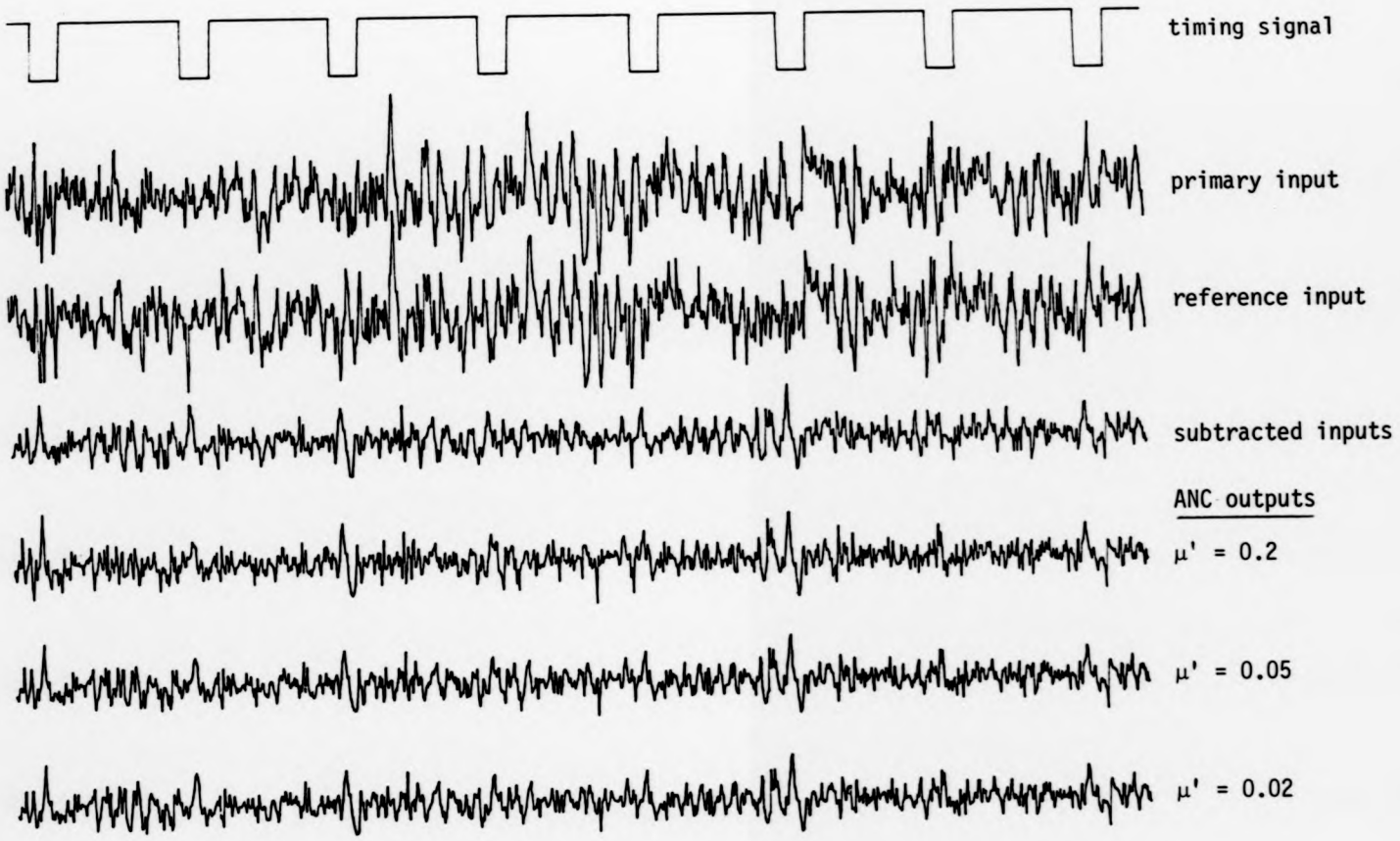



FIGURE 8.14 A comparison of the effect of convergence rate upon adaptive noise cancelling of EP data.
(Data from subject DAJ/I, channels 3 and 5, N=32)

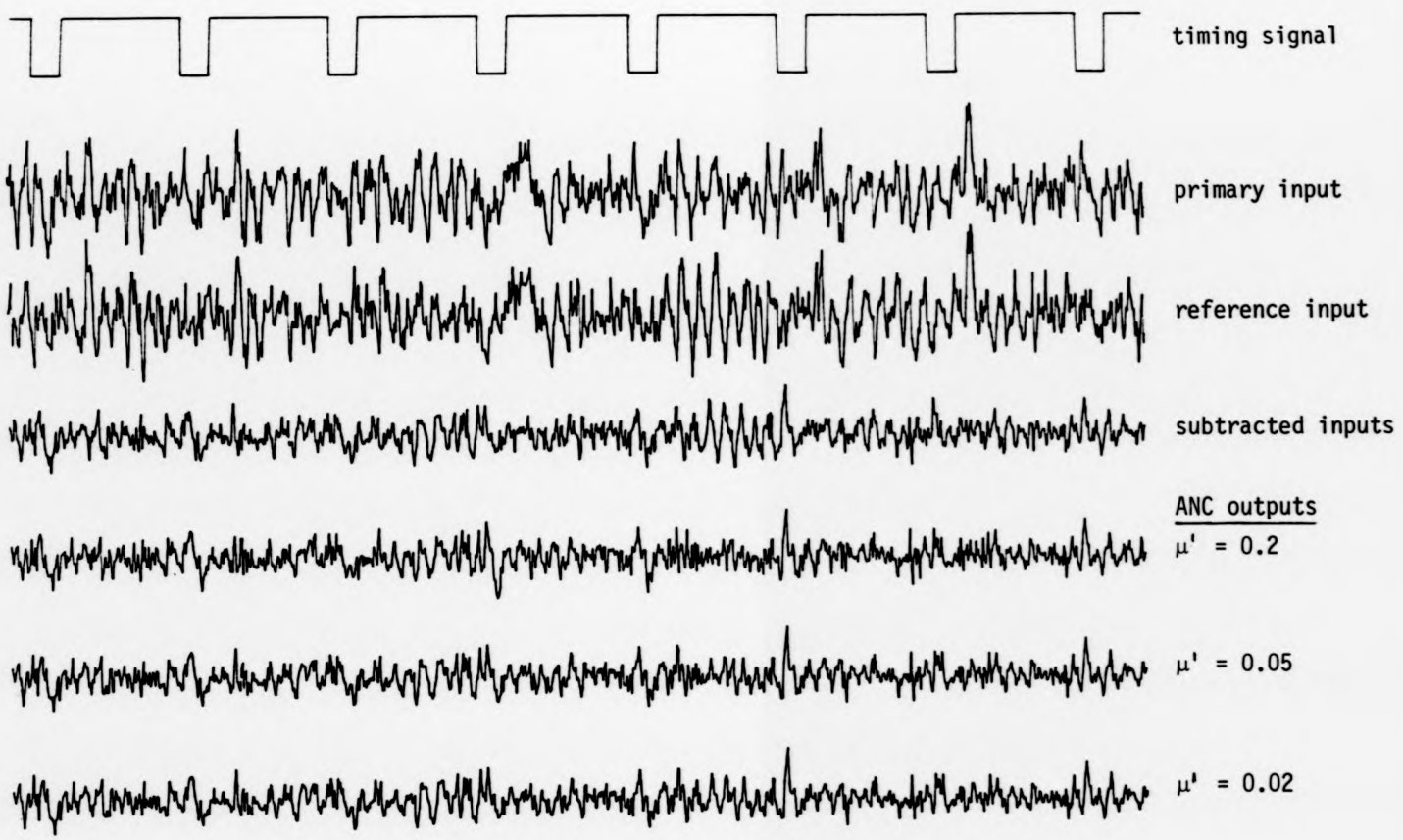



FIGURE 8.15 A comparison of the effect of convergence rate upon adaptive noise cancelling of EP data.
(Data from subject DAJ/I, channels 3 and 5, N=32)

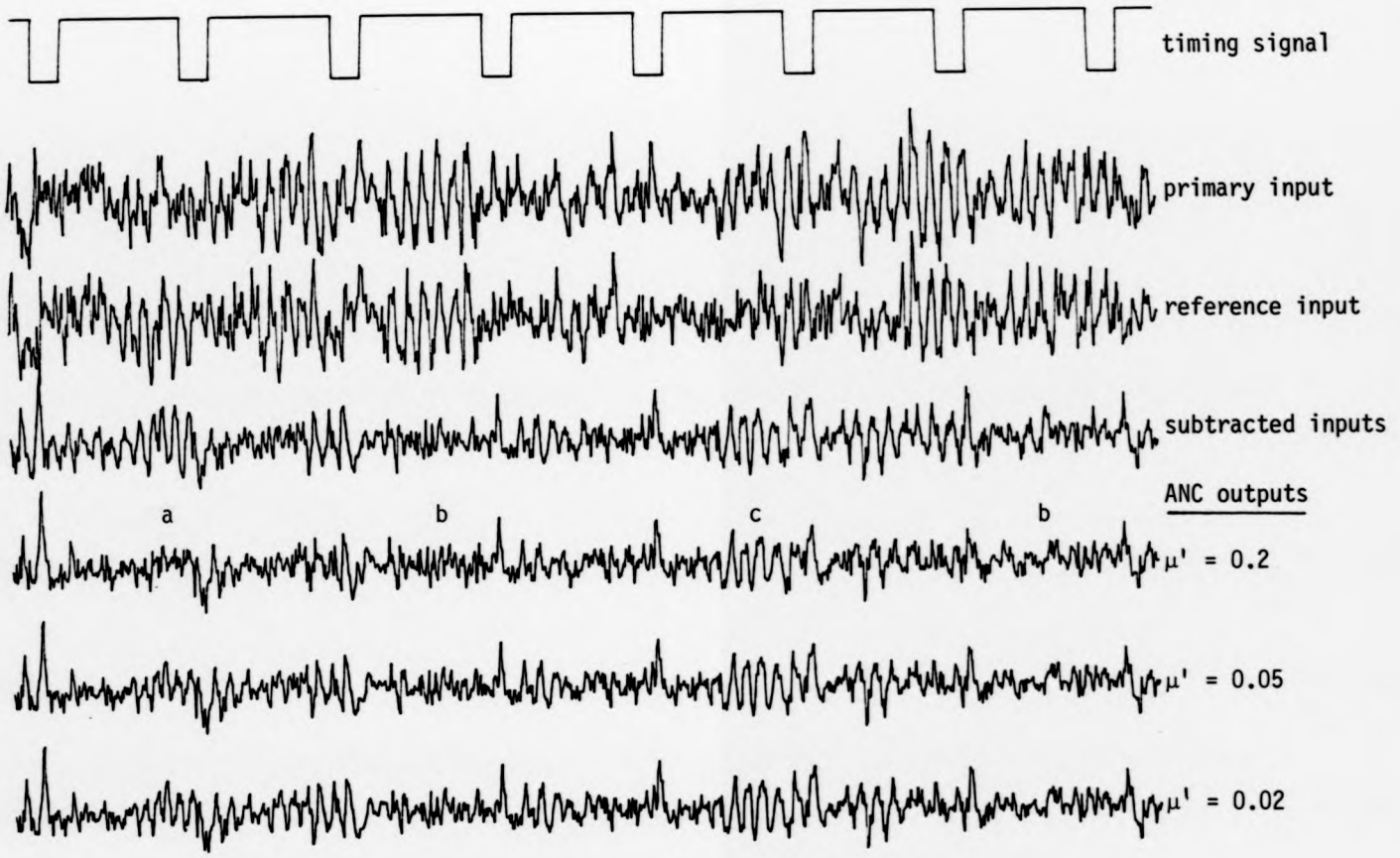



FIGURE 8.16 A comparison of the effect of convergence rate upon adaptive noise cancelling of EP data.
(Data from subject DAJ/I, channels 3 and 5, N=32)
Note filter performance when alpha is present in a) the reference alone, b) both inputs, c) the primary input.


FIGURE 8.17 A comparison of ANC outputs performed using different convergence rates. 8 responses averaged in each of 7 records with grand average of 56 responses.
a) subtracted inputs
b) ANC output, $\mu' = 0.2$
c) ANC output, $\mu' = 0.05$
d) ANC output, $\mu' = 0.02$

linear detrending described in section 8.3 prior to filtering. Three filter outputs for two 8 s records are shown in each plot corresponding to $\mu'$ values of 0.2, 0.05 and 0.02. The averages of each of 7 records are shown in fig. 8.17 with the grand average of all 7 records.

Visual examination of the results suggests that the value of $\mu'$ has little effect upon cancellation other than to provide greater cancellation of alpha activity when $\mu'$ is large. This is particularly evident in fig. 8.16. The other EEG components appear to be little affected by this choice. The averages in fig. 8.17 are very similar and there is no evidence of significant differences between them, apart from the greater alpha attenuation in records 2, 5 and 6 for the case when $\mu'$=0.2. This indicates that the ANC is not systematically distorting the EP signals.

These results show that much EEG activity may indeed be considered stationary, and hence justifies the use of the LMS algorithm in this application. Alpha activity however is exceptional in that it exhibits rapid variations which are not so easily tracked, though substantial attenuation of this activity is also achieved when sufficiently fast adaption is employed. Better cancellation of this activity may take place by using faster converging filters, though this will be discussed more fully later.

**8.5.1.2 The effect of filter order upon ANC performance.**

The second study was initiated to investigate the effect of filter order upon filter performance. The same 8 data records were filtered using $\mu'$=0.2 and using various values of N - 64, 32, 16, 8, and 4. Typical results may be seen in figs. 8.18 - 8.21. The averages of

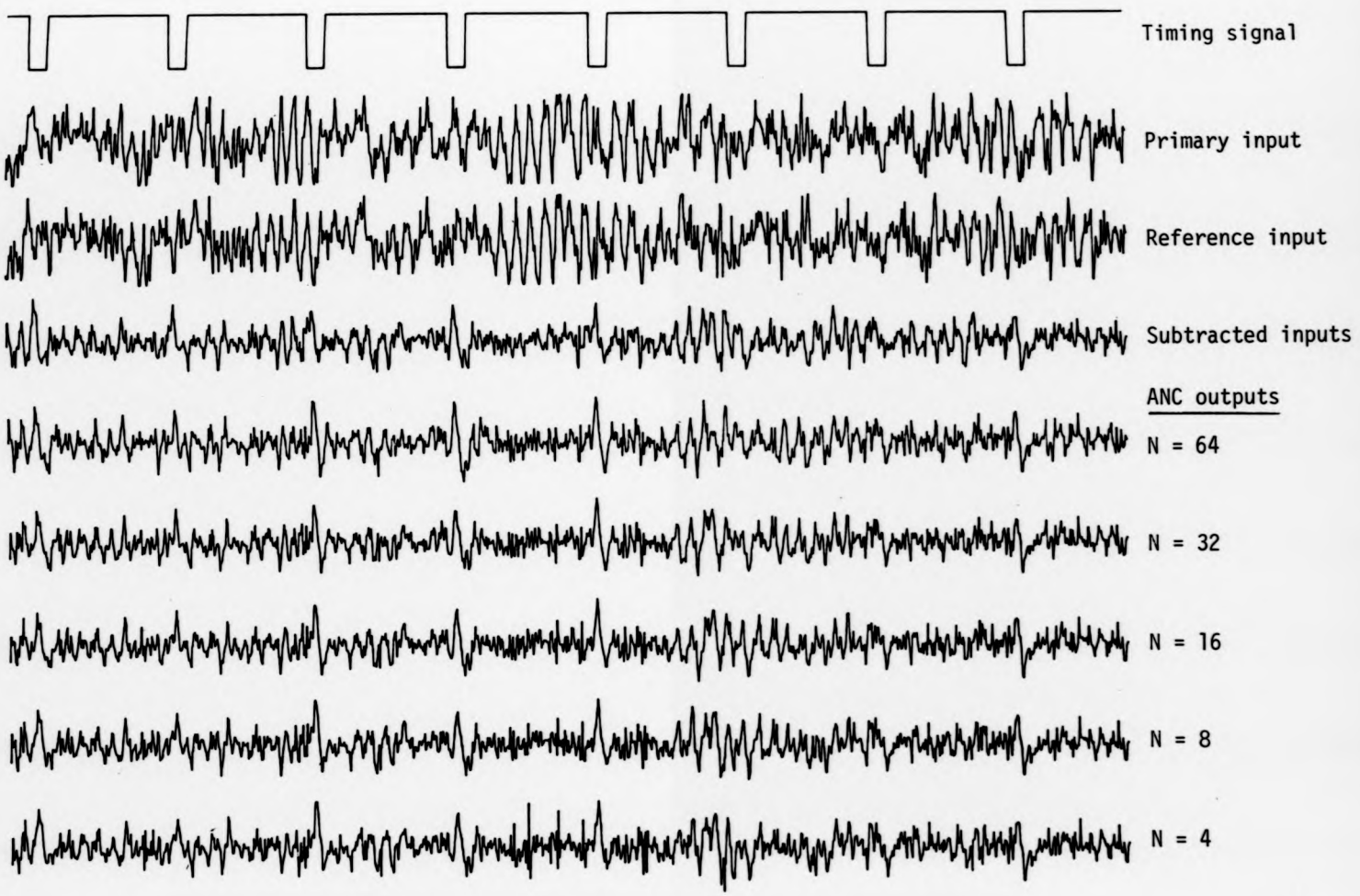



FIGURE 8.18 A comparison of the effect of filter order upon adaptive noise cancelling of EP data. (Data from subject DAJ/I, channels 3 and 5, $\mu' = 0.2$)

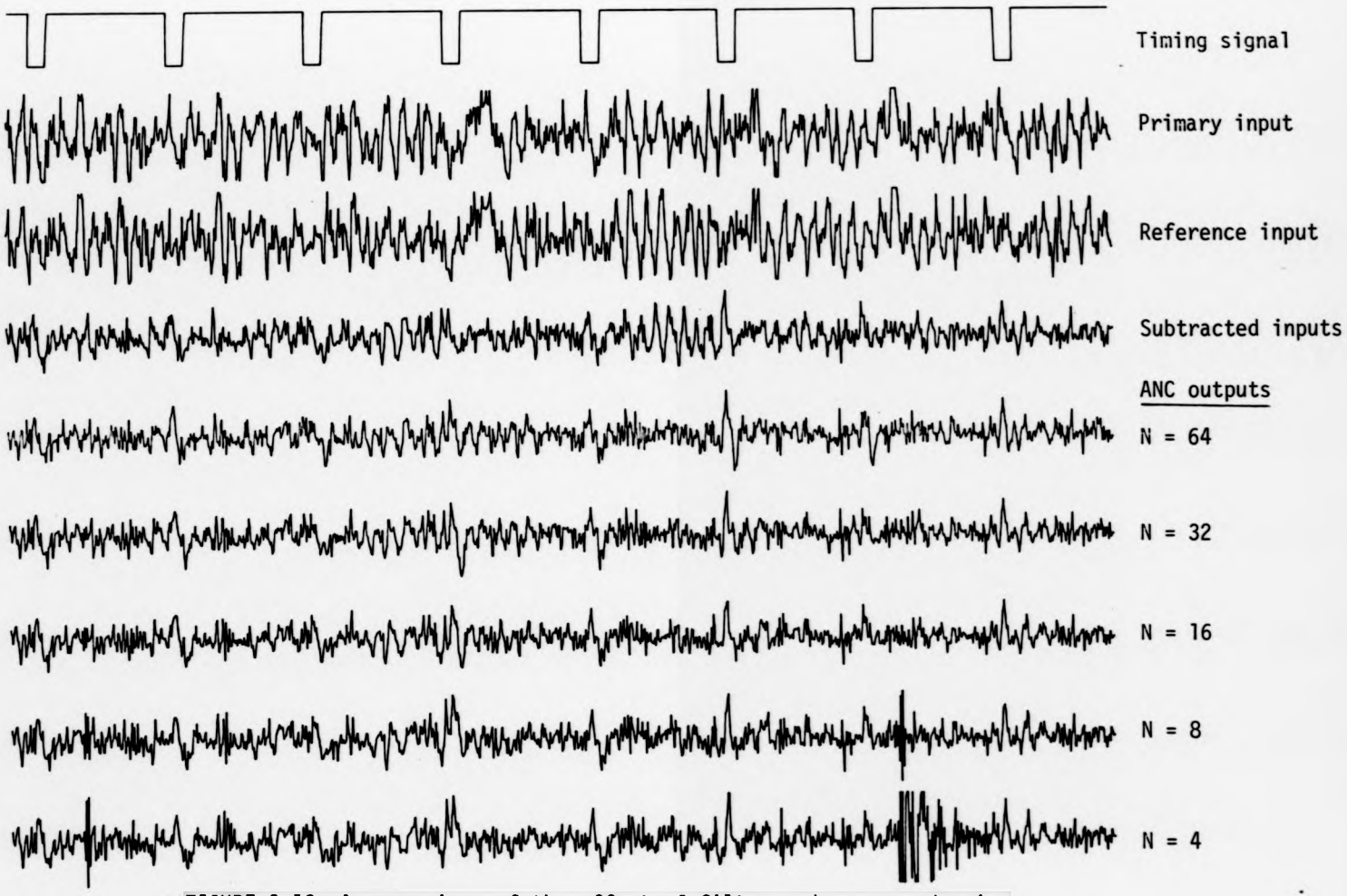



FIGURE 8.19 A comparison of the effect of filter order upon adaptive noise cancelling of EP data. (Data from subject DAJ/I, channels 3 and 5, $\mu' = 0.2$) Note transient instability due to artifacts in input channels.

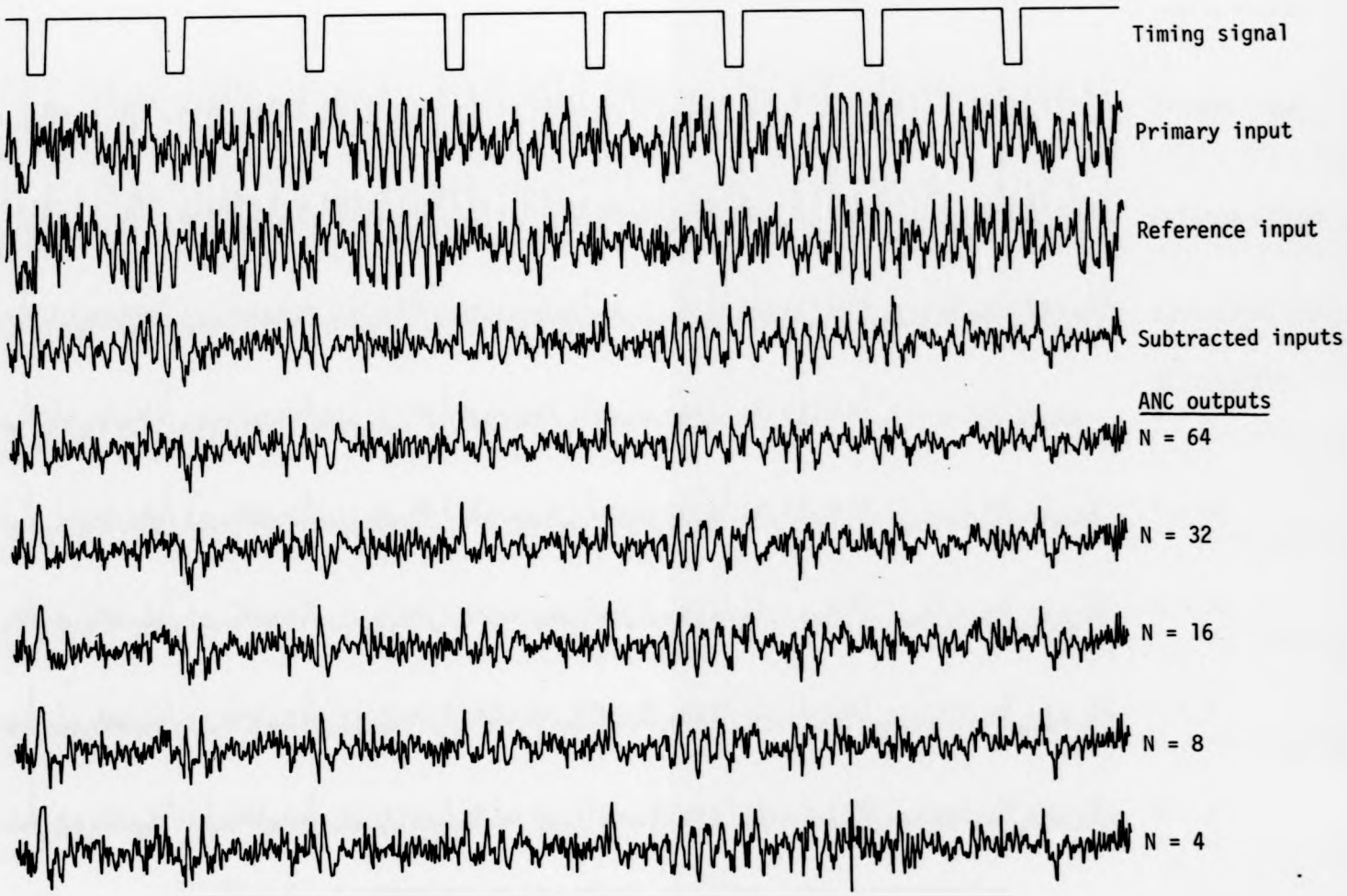



FIGURE 8.20 A comparison of the effect of filter order upon adaptive noise cancelling of EP data. (Data from subject DAJ/I, channels 3 and 5, $\mu' = 0.2$)

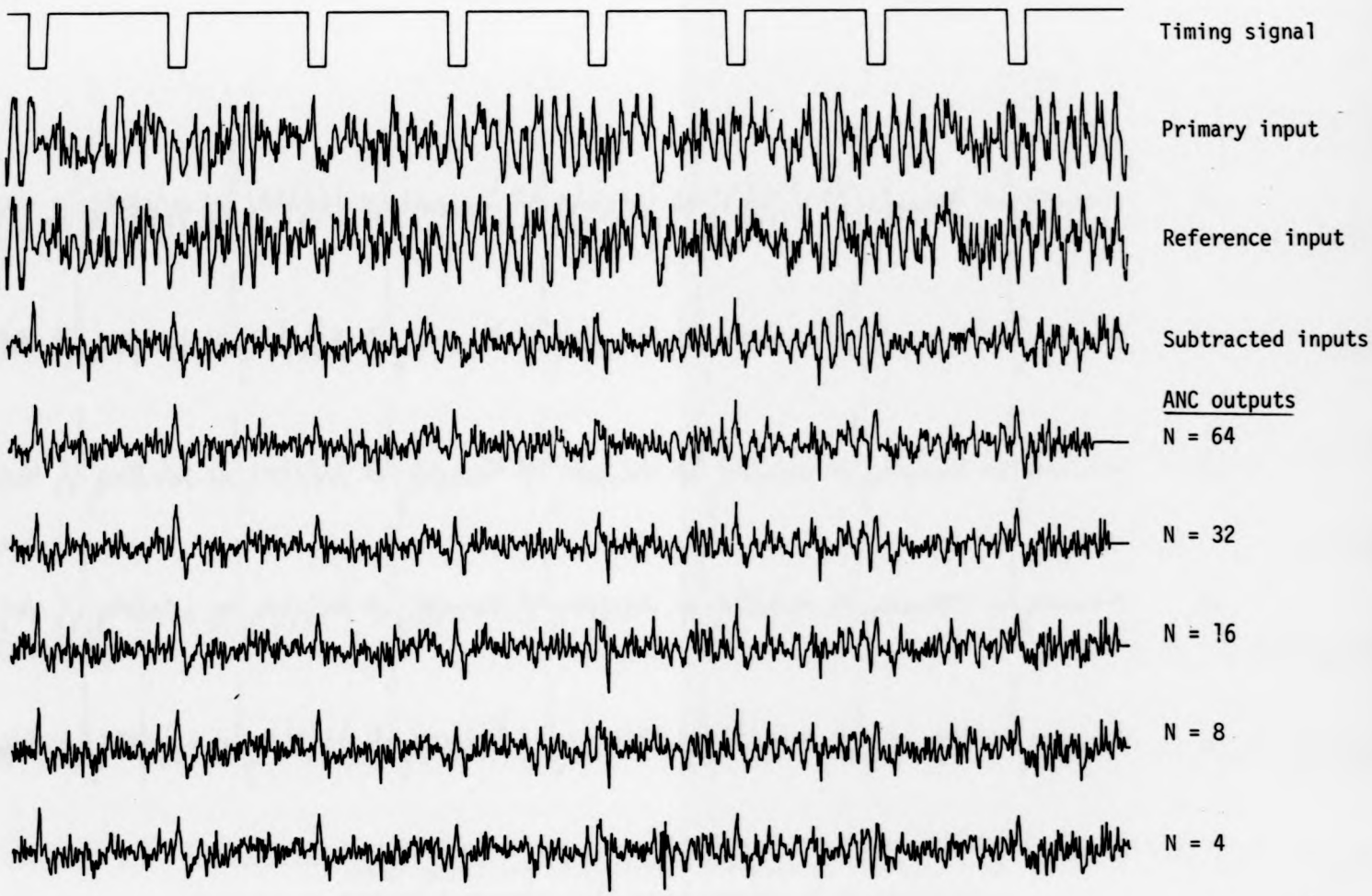



FIGURE 8.21 A comparison of the effect of filter order upon adaptive noise cancelling of EP data. (Data from subject DAJ/I, channels 3 and 5, $\mu' = 0.2$)

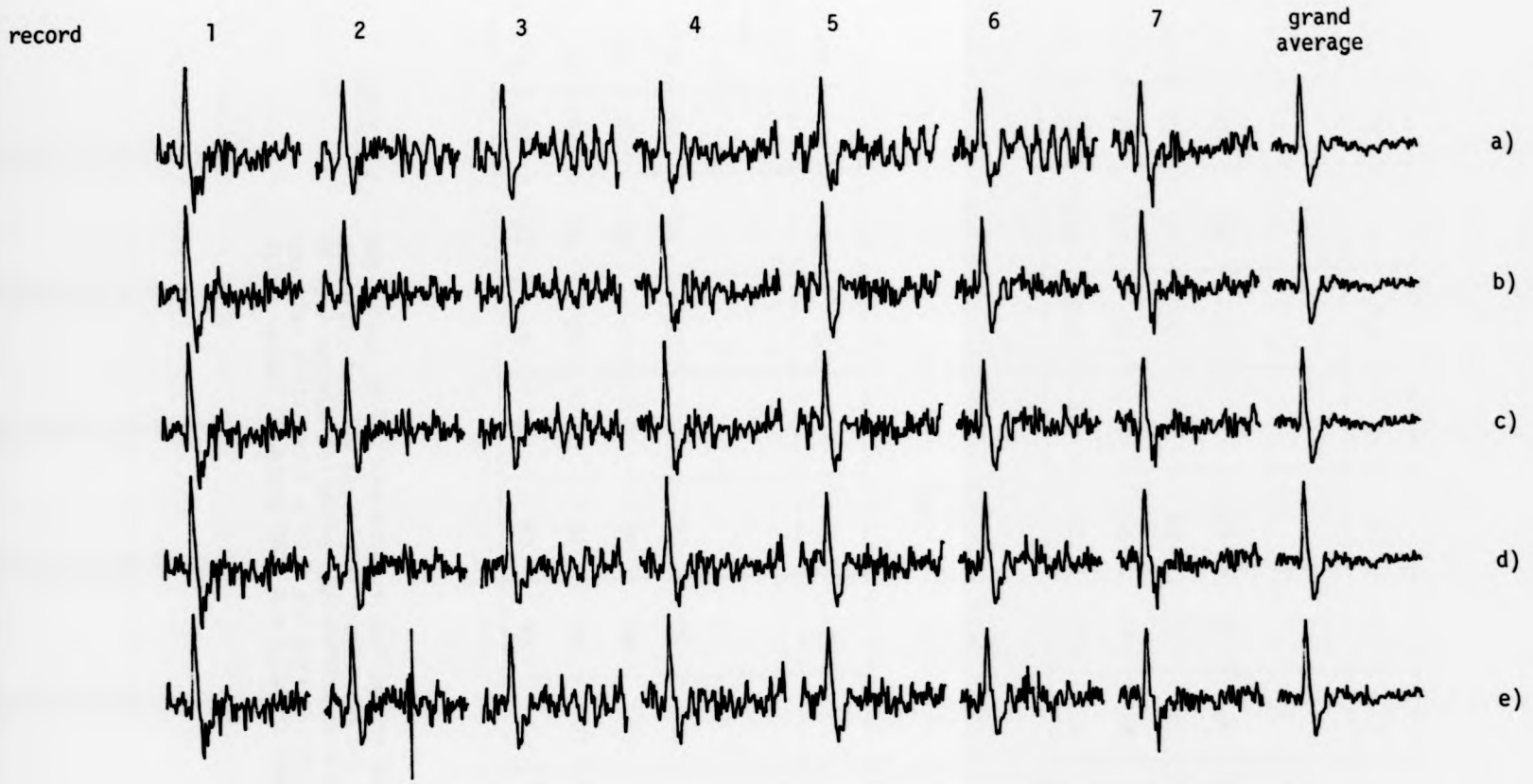



FIGURE 8.22 A comparison of the effect of filter order upon adaptive noise cancelling of EP data. 8 responses averaged from each of 7 records, with grand average of 56 responses.
a) subtracted inputs
b) ANC output, N = 64
c) N = 32
d) N = 16
e) N = 8

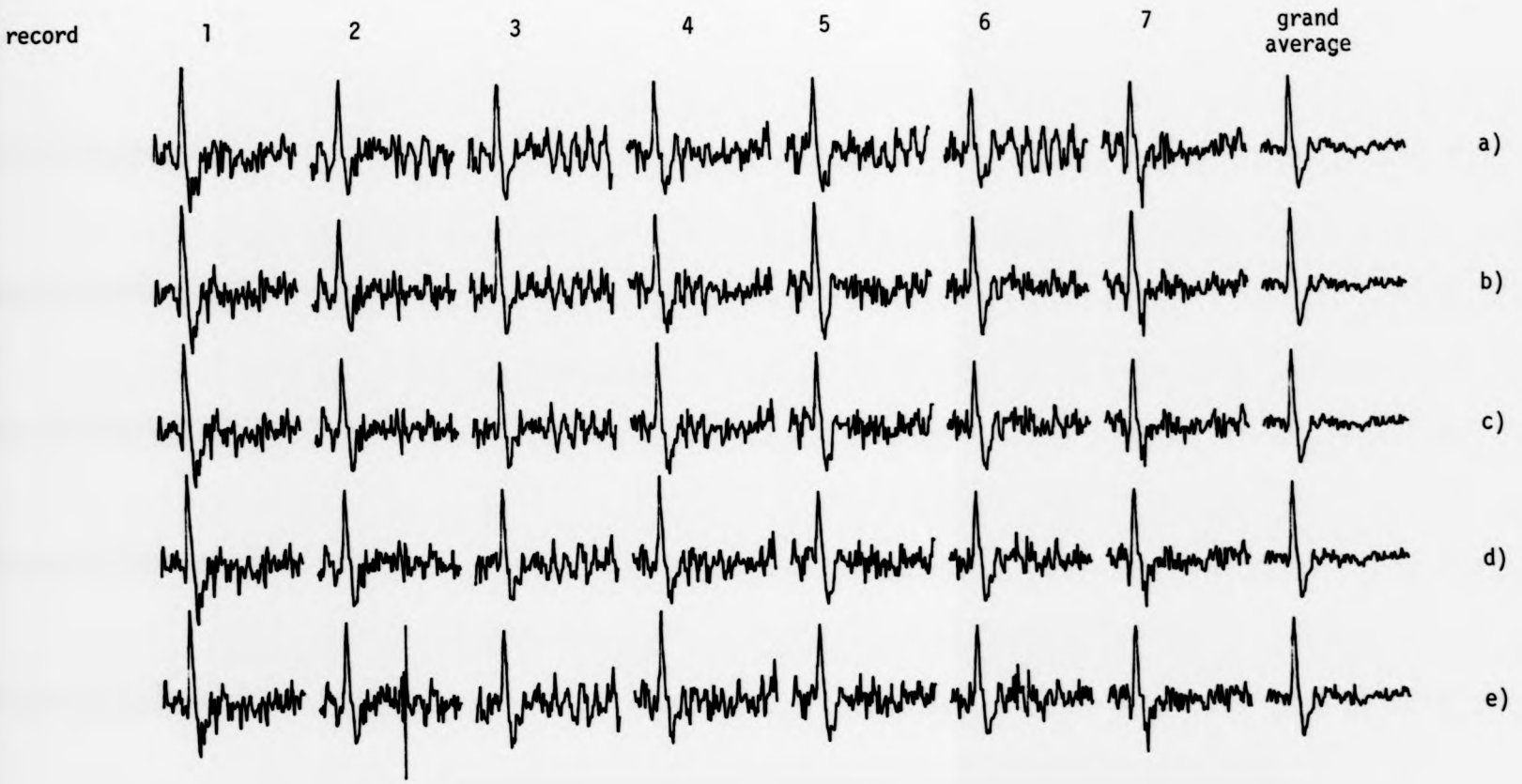



FIGURE 8.22 A comparison of the effect of filter order upon adaptive noise cancelling of EP data. 8 responses averaged from each of 7 records, with grand average of 56 responses.
a) subtracted inputs
b) ANC output, N = 64
c) N = 32
d) N = 16
e) N = 8

(a)

| record | 1 | 2 | 3 | 4 | 5 | 6 | 7 | overall average |
|---|---|---|---|---|---|---|---|---|
| filter order | | | | | | | | |
| 64 | 94 | 80 | 99 | 92 | 93 | 77 | 80 | 88.2 |
| 32 | 93 | 79 | 99 | 95 | 93 | 77 | 81 | 88.5 |
| 16 | 96 | 80 | 98 | 96 | 94 | 77 | 85 | 89.8 |
| 8 | 96 | 149 | 103 | 104 | 94 | 81 | 87 | 104.0 |

(b)

| record | 1 | 2 | 3 | 4 | 5 | 6 | 7 | overall average |
|---|---|---|---|---|---|---|---|---|
| filter order | | | | | | | | |
| 64 | 96 | 86 | 87 | 99 | 92 | 86 | 86 | 90.5 |
| 32 | 95 | 86 | 93 | 96 | 90 | 84 | 85 | 90.1 |
| 16 | 94 | 85 | 90 | 96 | 94 | 84 | 84 | 89.8 |
| 8 | 95 | 87 | 92 | 98 | 95 | 94 | 83 | 92.3 |

TABLE 8.2. RMS values of (a) the inter-signal epochs and (b) disparity between the signal epochs and the average EP in 7 records following application of a gated adaptive filter with $\mu'=0.2$ and various values of filter order. Subject DAJ, Run I, electrodes 3/5.

each of 7 records are shown in fig. 8.22. with the grand average of all 7 records. Note the disappearance of the signal at the end of the traces in fig. 8.21 which is due to the delay in the primary input of the adaptive filter, and is not a serious artifact. Artifactual high frequency activity can however be observed, for example in fig. 8.19 for the cases of N=8 and N=4. These are examples of the transient instability that can arise when the signal inputs significantly exceed their average values, and in this case may have been caused by an eyeblink artifact that had not been removed prior to filtering. This illustrates that artifact removal must be employed if routine application of adaptive noise cancelling is to be undertaken.

Examination of the individual traces shows that there is no single value of N that will provide overall best performance for these data. In an attempt to quantify these results the rms amplitude was computed for the inter-signal epochs, presented in Table 8.2. Figures are also presented for the rms deviation between each response and the average of 64 responses. These confirm the conclusion drawn from visual inspection that no single value of N in the range 16-64 gives optimum noise cancellation with minimum signal distortion. The lower values of N=4 and N=8 result in a filter which is more susceptible to transient instability, as borne out by fig. 8.19, though this was primarily due to a large amplitude artifact in the filter inputs as explained previously.

The inability to determine an optimum value of N is not completely unexpected, as variation in the statistical properties of the signal from record to record clearly occurs. As the differences are not very significant, it would seem that any of the values in the range 16-64 are admissible. In general slightly more smoothing is effected by the use of higher values of N, though smaller values have the advantage

of shorter computation times. In each case the average responses are not significantly different. These results supported the use of N=32 for the following ANC studies.

Use of an ANC has been compared in these records with bipolar subtraction of the filter inputs, as this is the simplest form of canceller possible, and permits a comparison with the general adaptive noise canceller. The results indicate that much of the cancellation is achieved through removal of coherent activity present in both channels, which may be due to common neural sources under each electrode, or activity arising at the reference site, or both. This is to be expected in the case considered here as the cross-phase plots in fig. 7.10 show no significant phase differences existing between correlated activity in channels 3 and 5. This would not be true of other electrode pairs which could be used to obtain EP enhancement in other situations; in these cases subtraction can lead to serious errors due to significant phase differences between the correlated activity in each channel. These results provide good evidence that the ANC is operating as desired.

Alpha activity was shown in the previous chapter to be poorly correlated in these experimental subjects for moderate scalp separations. Nevertheless substantially better cancellation of this activity is achieved by the ANC than by subtraction when it is present in the reference input alone provided that sufficiently fast adaption is employed, as shown in fig. 8.19. This is possible because the filter attenuates all uncorrelated activity in the reference input to minimise the output signal power. This may be seen in fig. 8.15. Note that in this case attenuation of activity is not based on cancellation of correlated components in each channel, as it is when alpha occurs in both inputs. However it is not possible to cancel alpha that is present

in the primary input alone by this method.

### 8.5.2 Subject MJM.

A more detailed study was performed on MJM/C data. As previously described 4 records were selected that typified the range of data records, and specifically included low noise data and EMG- and alpha-contaminated records. The results of applying a GAF ANC to each of these records is shown in figs. 8.23 - 8.26. In each plot the subtracted signal and the ANC output are shown for N=16, 32 and 64 and for $\mu'$=0.2 and 0.05. In addition the average of each of these signals is shown in fig. 8.27, with the grand average response of all 4 records.

The main conclusions drawn in the previous section are generally supported by these data. The ANC is clearly effective in enhancing the individual EPs in the background EEG and performs similarly to subtraction of the inputs in many cases. This is to be expected as the cross-phase plots in fig. 7.11 show that insignificant phase differences are present between the primary and reference inputs in this case. The close similarity between these outputs is apparent in the small difference signals that may be seen in fig. 8.29, which indicates that the ANC is performing correctly. However the disparity in fig. 8.28 has a considerable alpha component suggesting that this activity is effectively removed by the ANC. This is significant as it has been shown in chapter 7 that EP spectral components lie in the alpha band, and removal of this activity could not have been accomplished by conventional filtering without seriously distorting the EP. The effect of parameter values upon filter performance is also similar to the previous study.

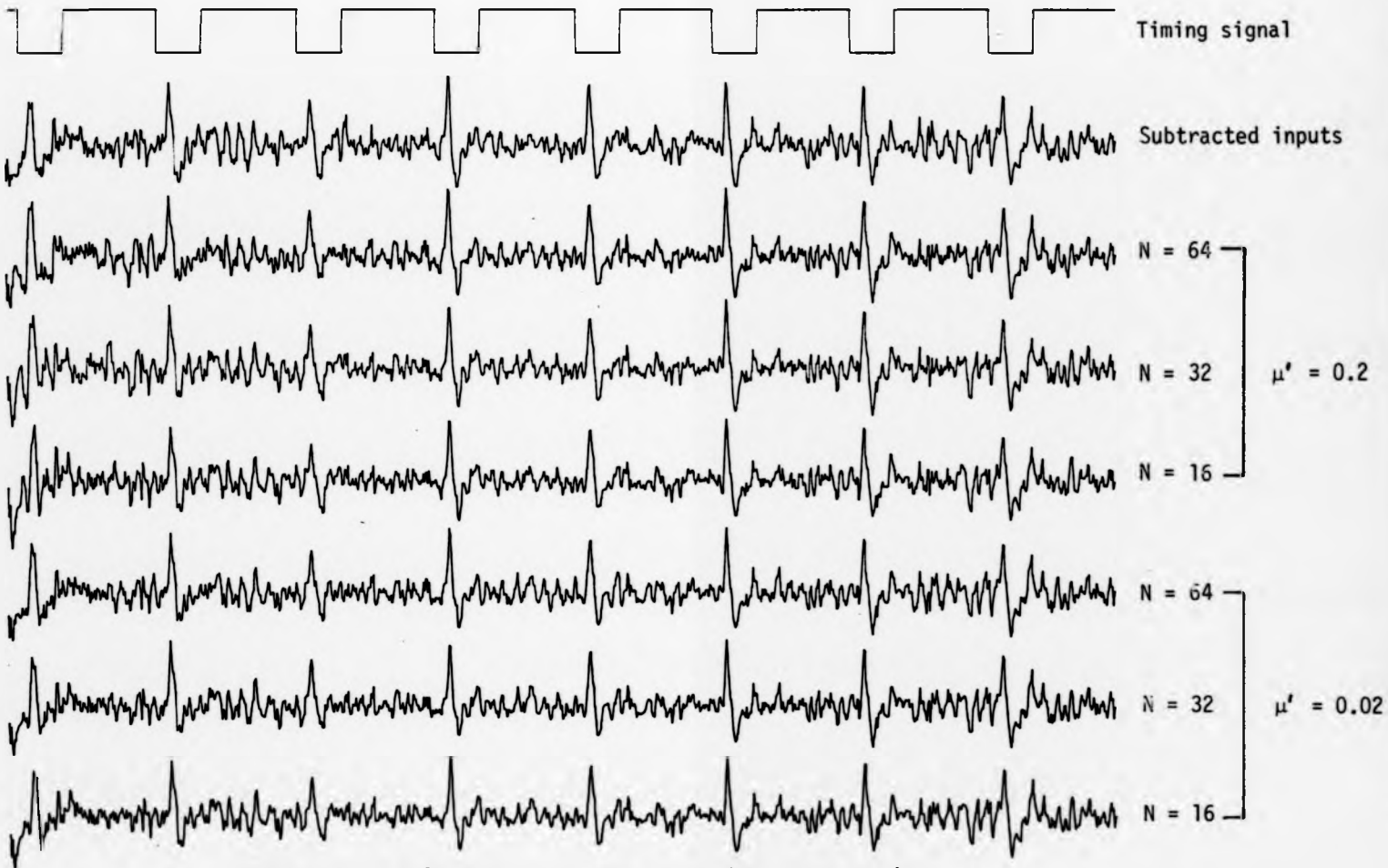



FIGURE 8.23 Application of an ANC to EP data representing low level EEG activity. (Data from subject MJM/C; N = 64, 32, 16; $\mu'$ = 0.2, 0.05)

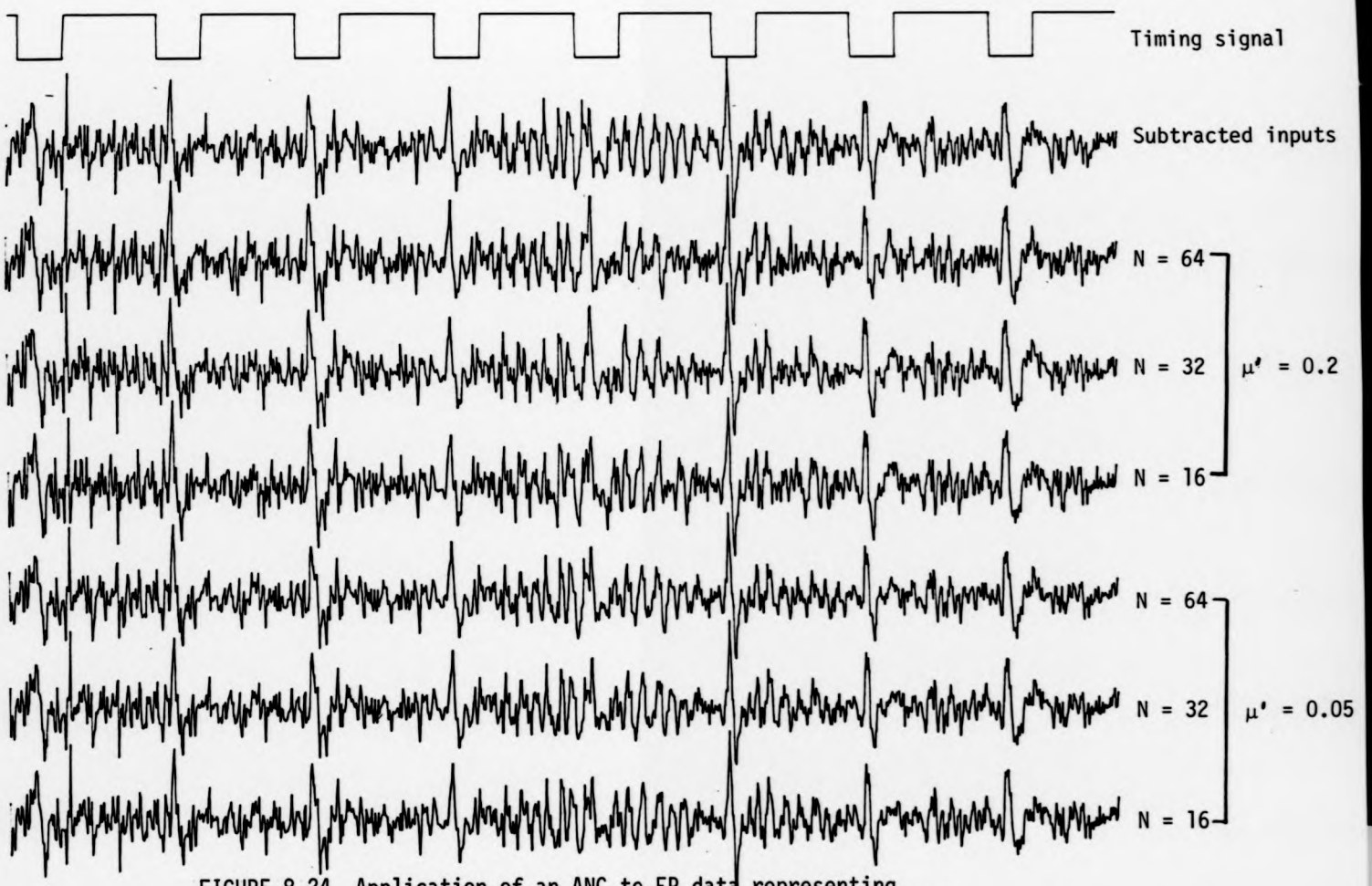



FIGURE 8.24 Application of an ANC to EP data representing moderate EEG activity. (Data from subject MJM/C; N = 64, 32, 16; $\mu^1$ = 0.2, 0.05)

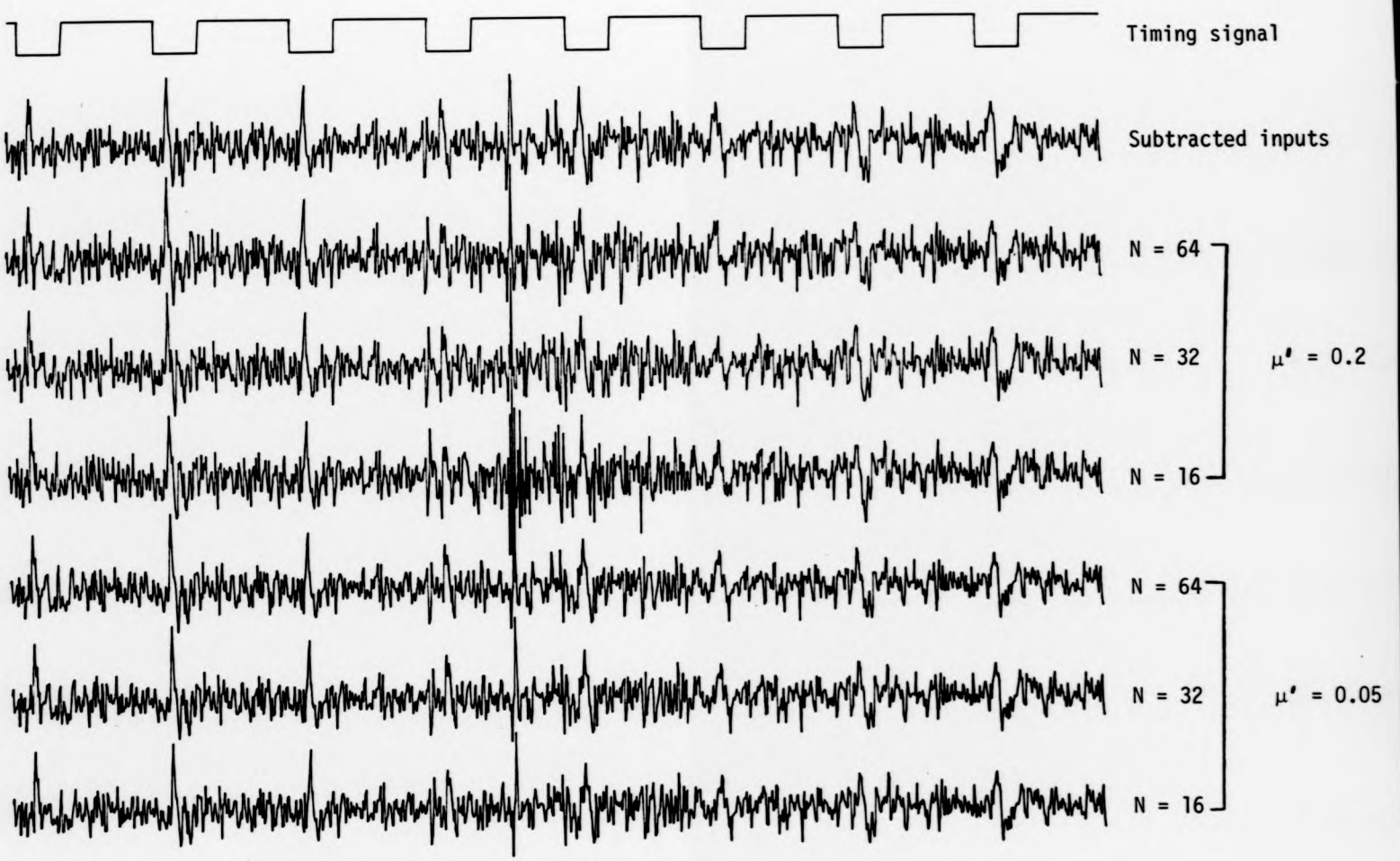



FIGURE 8.25 Application of an ANC to EP data representing high level EMG interference. (Data from subject MJM/C; N = 64, 32, 16; $\mu'$ = 0.2, 0.05)

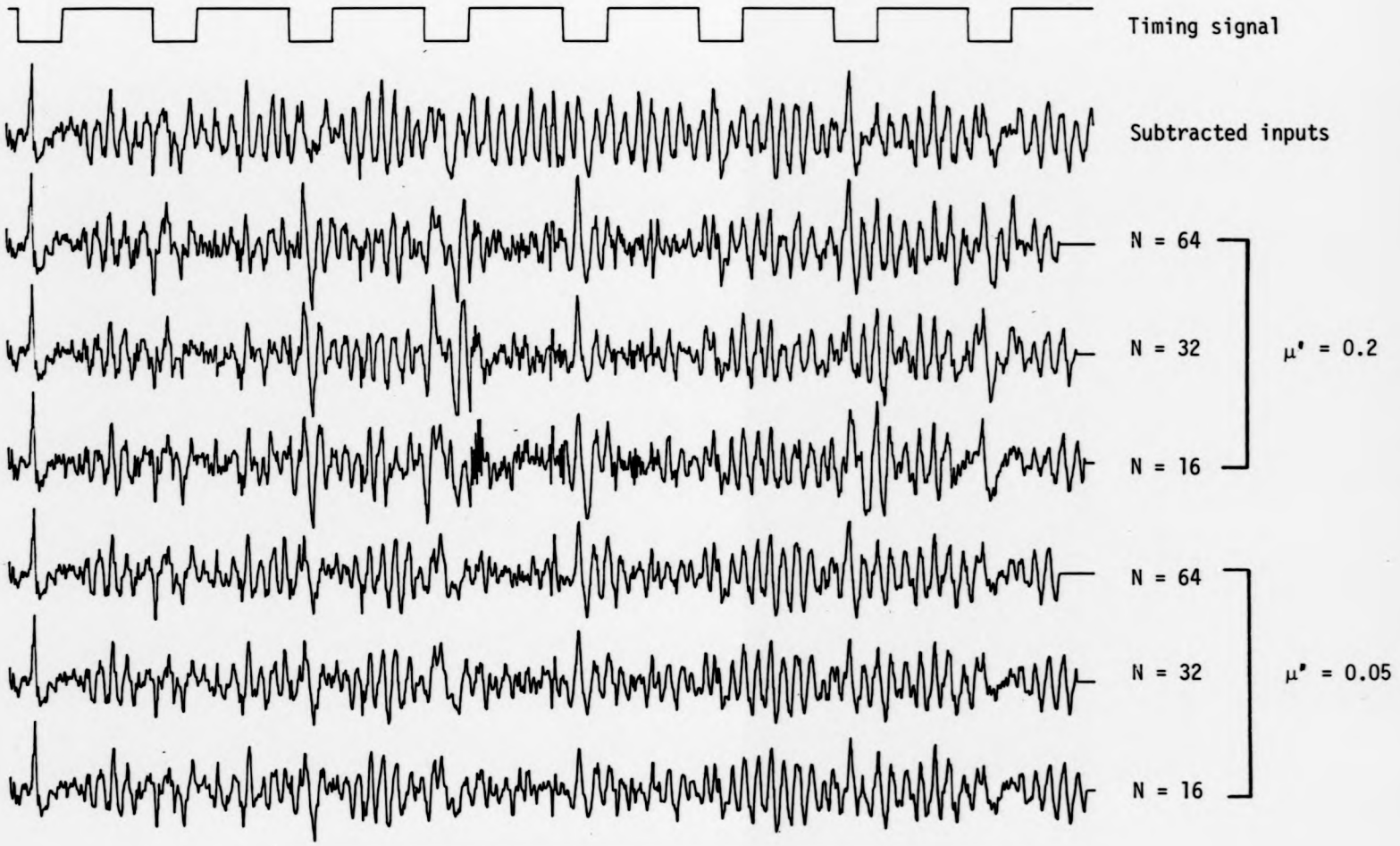



FIGURE 8.26 Application of an ANC to EP data representing severe alpha contamination. (Data from subject MJM/C; N = 64, 32, 16; $\mu'$ = 0.2, 0.05)

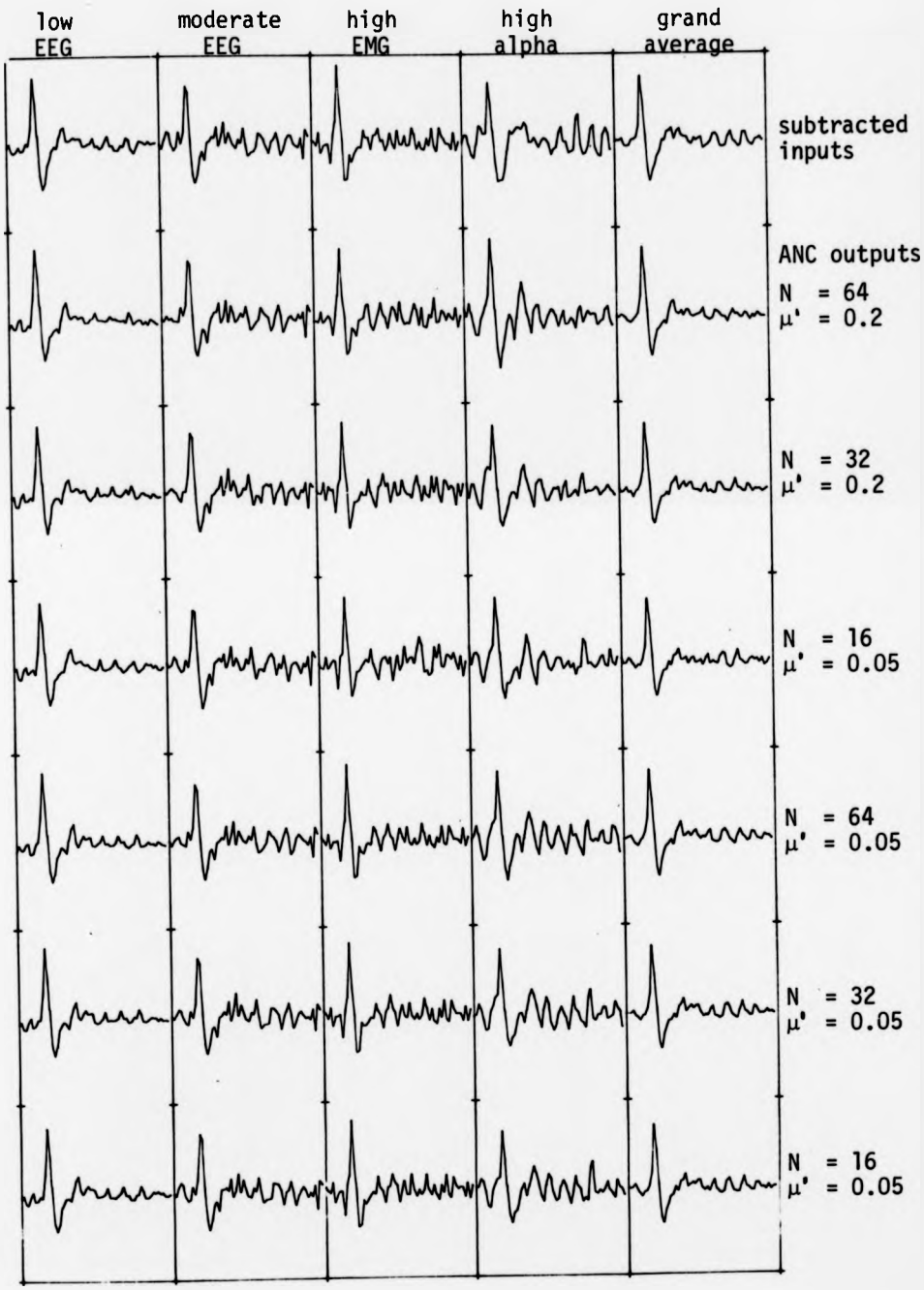



FIGURE 8.27 Average EPs obtained following application of an ANC with different parameter values to 4 data records.

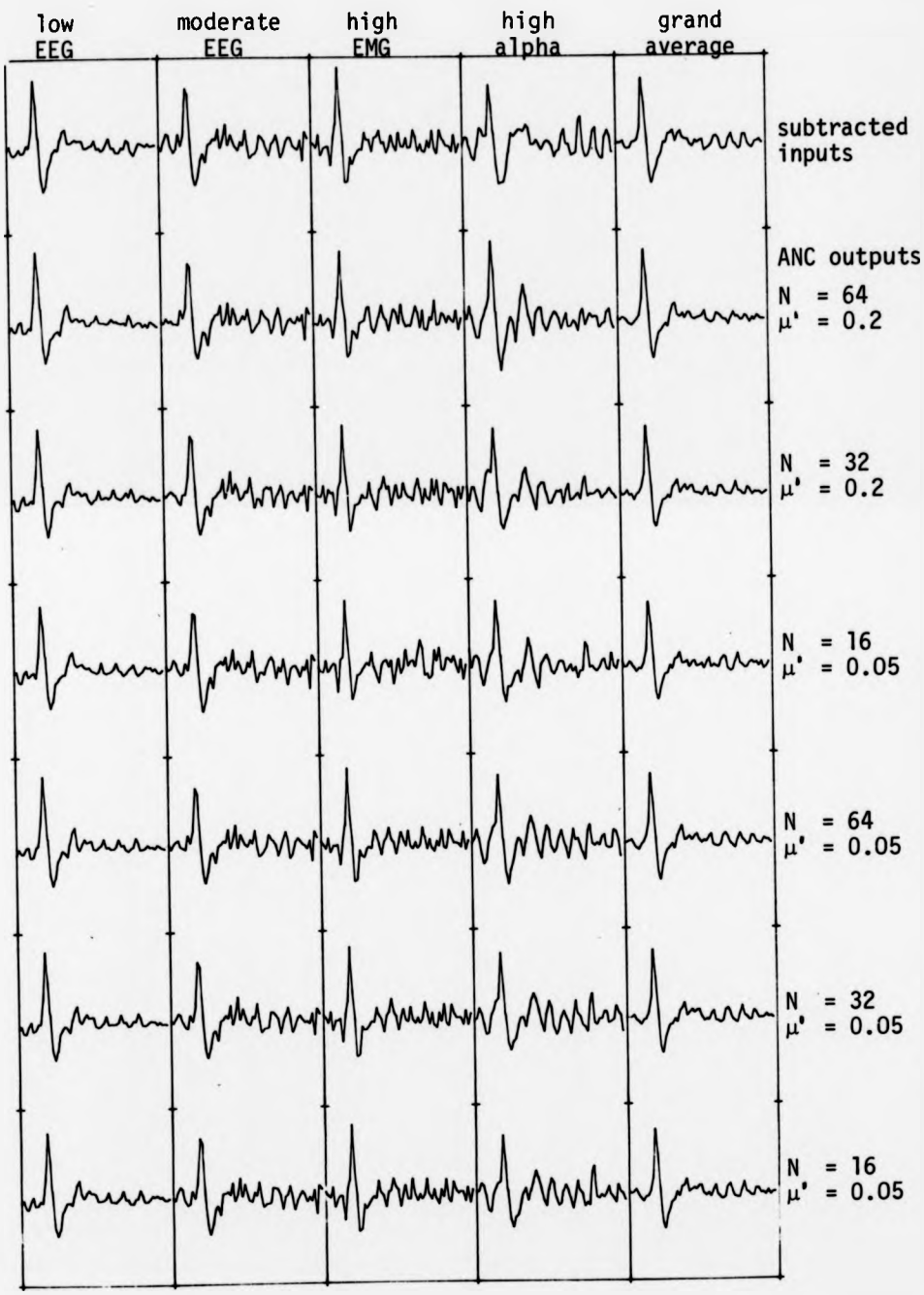



FIGURE 8.27 Average EPs obtained following application of an ANC with different parameter values to 4 data records.

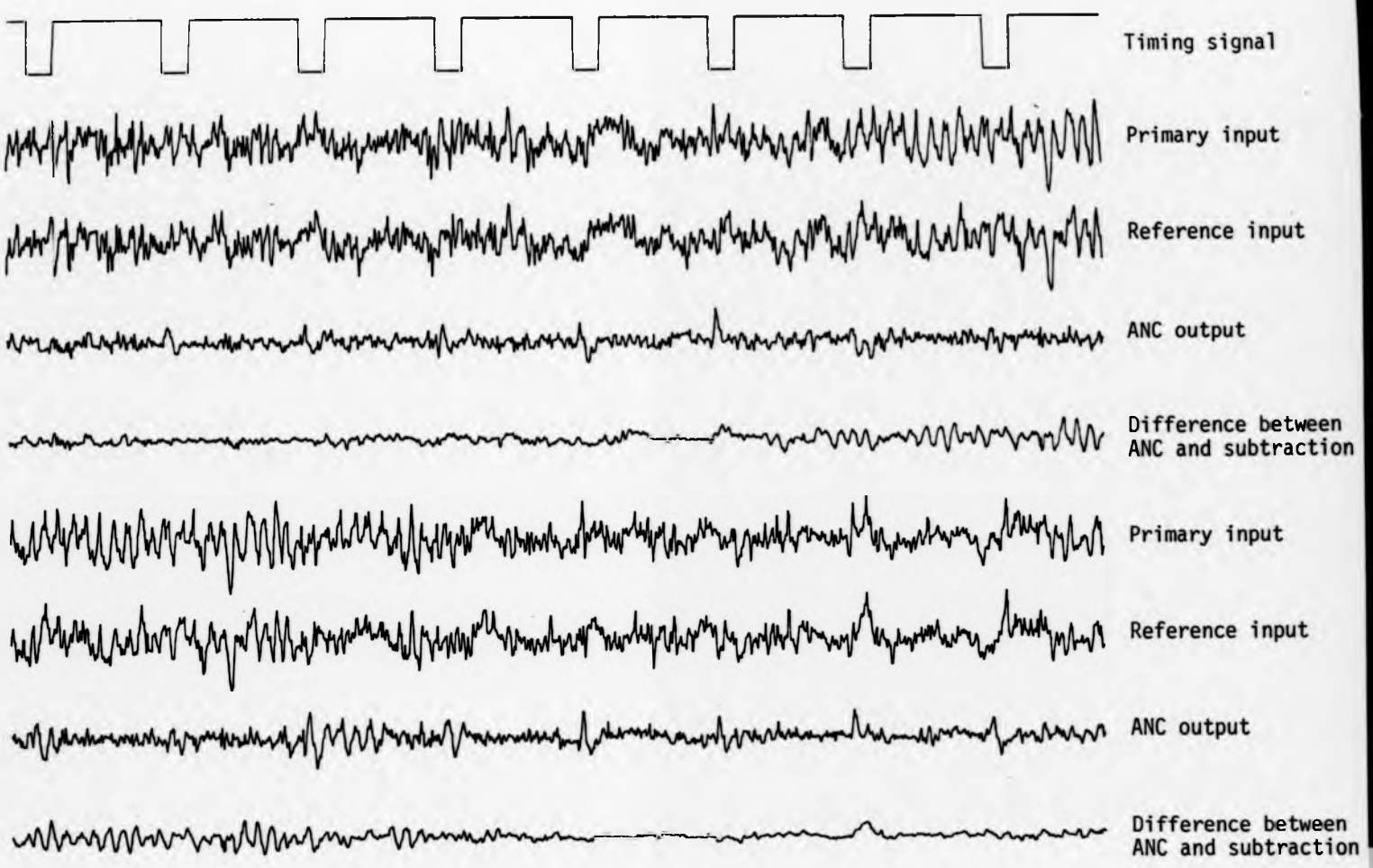



FIGURE 8.28 Comparison of adaptive noise cancelling of EP data with subtraction of filter inputs. Results shown for 2 records, illustrating reduction of alpha activity. (Data from subject MJM/C, N = 32, $\mu' = 0.2$)

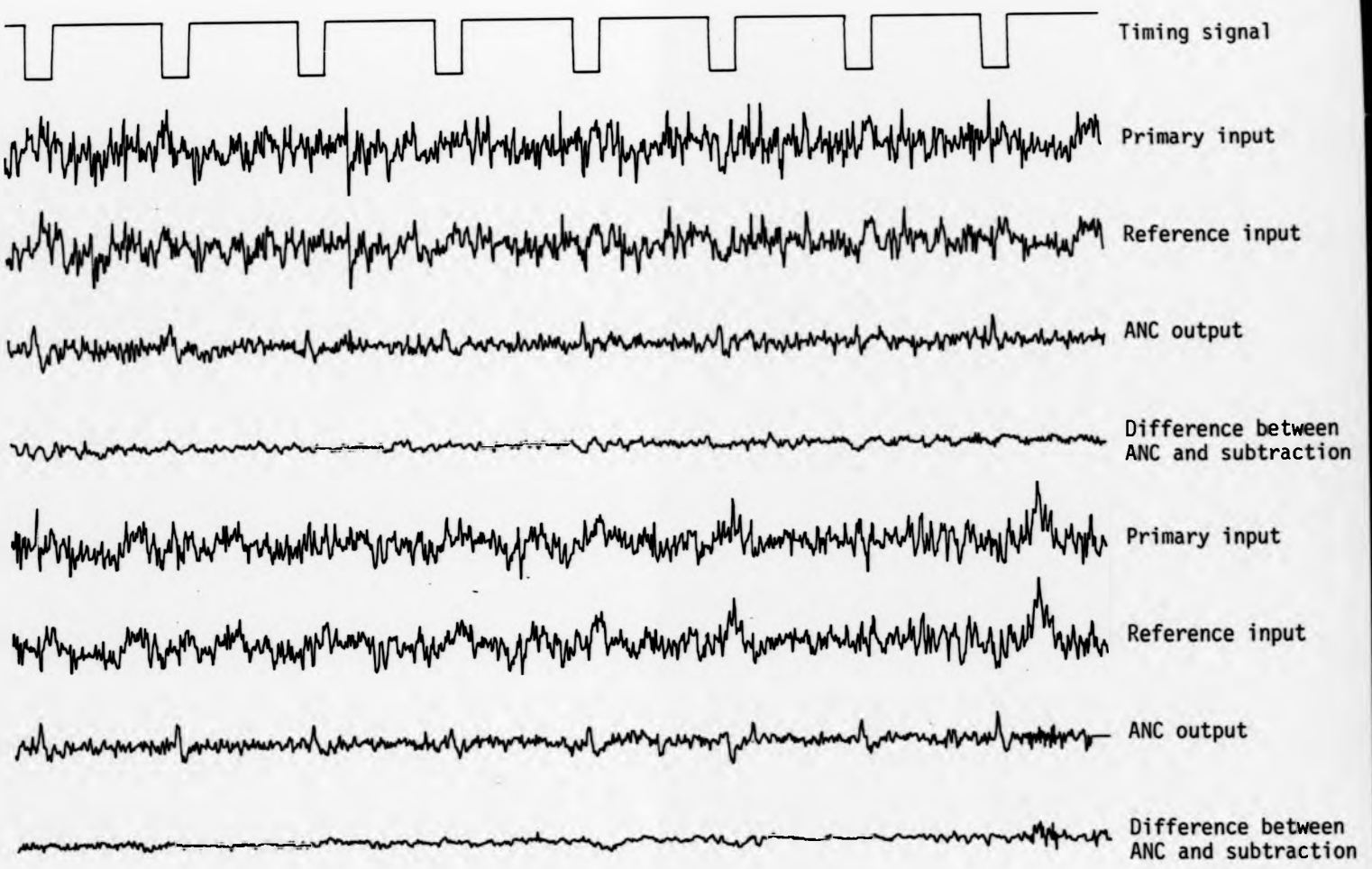



FIGURE 8.29 Comparison of adaptive noise cancelling of EP data with subtraction of filter inputs. Results shown for 2 records, illustrating similarity of responses. (Data from subject MJM/C, N = 32, $\mu' = 0.2$)

The choice of N does not seem to be very critical in the range 16-64 when moderate convergence rates are used, though when fast convergence is employed slightly better alpha cancellation is achieved by use of higher filter orders with less risk of instability. Varying the value of $\mu'$ has a greater effect upon canceller performance than varying the filter order, though the main result is better cancellation of alpha activity when high values of $\mu'$ are used. Effective cancellation is obtained in the first and second records (figs. 8.23 and 8.24) representing good and moderate signal quality, which clearly shows the possibility of extracting single EP responses by this method. The results are similar to subtraction of the inputs as expected, though there is evidence that better cancellation is obtained than subtraction when $N>16$ and $\mu'=0.2$. The results are similar in the third record (contaminated by muscle activity), though the average is slightly worse than when direct subtraction is performed. This may be due to mistracking of the high frequency muscle activity, and may possibly be avoided by the use of faster convergence filters. Note the transient instability that arises in this case when $N=16$ and $\mu'=0.2$ caused by an eyeblink in the data. The other filter outputs do not appear to be seriously affected by this artifact.

The fourth record (fig. 8.26) shows an example of particularly persistent, large amplitude alpha activity which almost completely obscures the individual EPS when subtraction of the two inputs is performed. Inspection of the ANC filter outputs shows that significant reduction of this activity has taken place, particularly when fast adaption is employed, though not all the alpha is removed. From an examination of the filter outputs in fig. 8.28 it may be seen that cancellation of this activity takes place when it occurs in both inputs or in the reference input alone, and that the possibility of detecting

The choice of N does not seem to be very critical in the range 16-64 when moderate convergence rates are used, though when fast convergence is employed slightly better alpha cancellation is achieved by use of higher filter orders with less risk of instability. Varying the value of $\mu'$ has a greater effect upon canceller performance than varying the filter order, though the main result is better cancellation of alpha activity when high values of $\mu'$ are used. Effective cancellation is obtained in the first and second records (figs. 8.23 and 8.24) representing good and moderate signal quality, which clearly shows the possibility of extracting single EP responses by this method. The results are similar to subtraction of the inputs as expected, though there is evidence that better cancellation is obtained than subtraction when $N>16$ and $\mu'=0.2$. The results are similar in the third record (contaminated by muscle activity), though the average is slightly worse than when direct subtraction is performed. This may be due to mistracking of the high frequency muscle activity, and may possibly be avoided by the use of faster convergence filters. Note the transient instability that arises in this case when $N=16$ and $\mu'=0.2$ caused by an eyeblink in the data. The other filter outputs do not appear to be seriously affected by this artifact.

The fourth record (fig. 8.26) shows an example of particularly persistent, large amplitude alpha activity which almost completely obscures the individual EPS when subtraction of the two inputs is performed. Inspection of the ANC filter outputs shows that significant reduction of this activity has taken place, particularly when fast adaption is employed, though not all the alpha is removed. From an examination of the filter outputs in fig. 8.28 it may be seen that cancellation of this activity takes place when it occurs in both inputs or in the reference input alone, and that the possibility of detecting

the individual responses is considerably improved. More careful analysis however shows that this may be at the risk of distortion in the individual responses. For example in fig. 8.26 there seem to be large variations in the filtered single responses. While it is difficult to place reliance upon these alone, the plots in fig. 8.27 show some distortion in the average response, which is not seen in the first three cases, where the averages bear close resemblance to each other. The distortion appears to be added rhythmic activity in the averages following the main EP transient. This may be due to the suspension of weight adaption during the signal epoch which results in filter weights incorrectly tracking the input signals if these are changing during this period. There are also significant variations in the filter outputs to each of the different parameter values, particularly in the case of fast adaption. While some variation would be expected, the degree to which this occurs suggests that different cancellation models are being used at times, and that these depend upon the filter order. These results cast some doubt on the value of this method to accurately determine response waveshape in the single trials when applied in the manner described. Nevertheless it is clear that identification of the single responses has been considerably improved, and these difficulties may be overcome by further development.

## 8.6 Summary and conclusions.

The initial sections of this chapter were concerned with the determination of adaptive filter strategy, which involved selection of a suitable adaptive algorithm and associated filter parameters. A discussion of these suggested that an initial exploration was appropriate using the LMS adaptive algorithm, as frequency domain filters are inappropriate for fast adaptive filtering and it was

uncertain whether there would be any advantage in using other, more complex algorithms. This was followed by a consideration of filter parameter values as no straightforward means exist at present to specify these. It has been suggested that the filter order could be estimated from the impulse response representing the transfer function for the correlated channels, and this was subsequently validated by experimental application of the filter to EP data. Methods of removing fluctuating baseline components from the data prior to adaptive filtering were discussed. A study of several different methods showed that a simple recursive digital filter is adequate and can be readily incorporated within an on-line filter. A study was also performed to compare the effectiveness of the gated adaptive filter with an ordinary adaptive filter when processing EP data. The results confirmed that this technique is appropriate for filtering EP data when fast adaption is necessary, and this approach was followed in all subsequent work.

It has been shown in this study that the ANC is effective in cancelling EEG activity present in one channel using a reference channel containing correlated activity. This technique yields enhancement of the embedded EPs with no evidence of significant distortion, though EP components present in the reference channel are removed from the primary channel. Electrode choice was made on the basis of previously obtained scalp distributions to minimise the possibility of this occurring. Data from two subjects were analysed and gave similar results, though there were some differences between them that may be attributed to differences in the individuals' EEG. Subject DAJ tended to have a fairly uniform EEG characteristic with only occasional short duration, low amplitude alpha activity, while subject MJM displayed a much wider range of EP signal quality and permitted the effects of these to be studied. In every case considered the phase relationship between correlated EEG

activity was not significantly different from zero, and so one would expect the ANC to be as effective as direct subtraction of the signal channels, provided that there were no amplitude differences between each channel. This was indeed observed and provides experimental verification for the correct operation of the filter. However in a number of records the ANC performed significantly better than subtraction and clearly shows the value of the technique. This approach is likely to offer substantial improvements when significant amplitude or phase differences exist in the data, as shown to exist in chapter 7. In general ANC performed better than subtraction for subject DAJ, though the improvement was small. MJM yielded similar improvement in some records (particularly those reflecting moderate EEG activity), but performed slightly poorer when high frequency muscle noise was significantly present.

Though it was not expected that alpha would be significantly removed, as it is poorly correlated over the scalp sites used, substantial cancellation did in fact occur. There are two ways in which this can occur. When alpha is present in both channels the filter assumes these to be correlated and attempts to remove it from the primary input in the usual manner. However the filter can also remove alpha which is present in the reference lead alone by attenuating all uncorrelated activity so that it does not appear in the output. Both of these means are used to effect alpha cancellation, though some residual alpha activity can be present due to the finite adaption time of the ANC. It is not possible to cancel alpha activity that is present in the primary input alone. These results are very significant as they demonstrate the ability of the noise canceller to remove activity in the signal bandwidth without distorting the signal itself. This cannot be achieved by conventional filters applied to a single data channel and

demonstrates the potential value of this technique for EP processing.

While short duration alpha bursts were often effectively removed by the ANC, the results were more variable when persistent, large amplitude alpha was present. Though a great deal of this activity was removed or attenuated, there is evidence that the individual EPs suffered distortion. Artifactual activity was present in the averages, and the response variability appeared to be higher than for the subtracted inputs. These results cast doubt on the reliability of the technique as applied in these circumstances, nevertheless individual responses are more easily detected by this method. It is possible that use of the gated filter structure contributes to unreliable estimation of the signal in this case. Disabling the coefficient updates during the EP epoch may no longer lead to improved filtering over the ordinary AF if the EEG is rapidly changing in character during this interval. It is not clear what the difference in performance might be if the ordinary AF were used, nor whether distortion of the EP would still be unacceptable, but a fuller investigation of this was not pursued due to the limited time available. The LMS algorithm was able to provide sufficiently fast adaption rates for most EEG components though the maximum adaption rate was required to achieve cancellation of alpha activity. Use of faster converging algorithms may lead to better tracking and hence cancellation of this activity, but with the possibility of greater misadjustment and distortion.

The overall conclusion is that ANC has been shown to be very useful in obtaining improved estimates of the EP, particularly of the single responses, though the average responses have also been enhanced in several cases. The method cannot be applied routinely to every EP regardless of scalp location and is restricted to situations where

correlated activity is present in a reference channel that does not contain significant EP components. This requires a study to be made of scalp distributions prior to filtering. Though these results have been obtained using selected VEPs recorded at particular electrode sites, it is expected that this technique will also be appropriate when used with other electrode sites and stimulus modalities. This possibility is discussed further in the next chapter which considers the results of this study in the wider context of the aims of signal processing of EPS in the research and clinical environments. Some of the methodological approaches taken in this investigation are also reviewed, and recommendations are made regarding these to aid future investigations. The chapter concludes with suggestions for further research in this field.

# CHAPTER NINE.

# CONCLUSIONS.

## 9.1 Summary.

In this thesis a study has been described which was performed to gauge the suitability of adaptive filters to noise cancelling in EP data records. The overall problem was first placed in context by presenting a discussion of the nature of EP signals, both in terms of their general properties and main applications in assessing sensory pathway function, and then in terms of their signal characteristics. The common difficulties in deriving noise-free estimates of these signals were illustrated. These difficulties are frequently circumvented by the use of coherent signal averaging, a procedure which is not always wholly applicable. The limitations of signal averaging have been described and a discussion of alternative or modified procedures presented. In general these do not provide a routine means of enhancing signal activity nor do they enable detailed investigation of single trial responses for which alternative signal processing strategies must be sought.

Adaptive filters were described, based on a transversal filter structure operating under the control of the Widrow-Hoff LMS adaptive algorithm. The theory of operation of these filters was reviewed and pertinent results derived or otherwise presented that describe their operating characteristics. A large part of this theory section was concerned with the noise cancelling mode of adaptive filtering which is of particular relevance to this thesis. Several expressions were presented which indicate the potential advantages and which permit an

assessment to be made of the validity of this approach in individual situations. A development of the basic adaptive filter has been introduced which carries less risk of distorting transient signals under conditions of fast adaption. A theoretical justification for this was presented and computer studies were described using artificial signals which show that the gated filter approach is better than either the basic adaptive filter or time-sequenced adaptive filter when fast adaption is required. This study was later extended using EP signals, and the results confirmed that routine use of the gated adaptive filter was appropriate for filtering the EP signals considered in this thesis.

The experimental work required a general-purpose software facility to implement advanced signal analysis and processing, and to develop new filter strategies. Such a facility was not available at the start of this research and had to be developed as part of this research project. The hardware and software options available are described in detail together with a number of other methodological points prior to the main experimental chapters of this thesis.

Experimental investigation of EP signal enhancement by the method of adaptive noise cancelling was performed in two stages. The first stage was concerned with the justification of this approach, and involved a detailed analysis of the properties of a range of EP and EEG signals, concentrating in particular upon their scalp distribution properties. This confirmed the initial hypothesis that significant correlations are present in the EEG scalp record and this justifies the use of a canceller based on correlated reference sources. It was discovered that the correlation properties over posterior scalp locations are not identical in all frequency bands, and in particular were poorer for midband EEG frequencies (8-16 Hz). A general decrease

in correlation was observed with increasing electrode separation, and this indicates that the reference sources must be closely spaced to the desired site of EP activity for optimum cancellation of the correlated EEG activity. A determination of the EP scalp distributions showed that it is possible to obtain reference sites close to the primary signal source which permit useful noise cancellation to be obtained.

The second stage of the investigation was concerned with the experimental validation of this approach using an adaptive noise cancelling to process typical EP signals from previously obtained recordings. A number of practical aspects of this approach were first examined, and included the selection of a suitable adaptive filter and the determination of the parameters of the filter - specifically the filter order and convergence rate. These are not easily specified a priori when the statistics of the signals are unknown. Experimental investigation of the effect of these upon filter behaviour was therefore an important part of this study, though the results showed that the choice of convergence coefficient was more critical to the correct operation of the canceller than the filter order. A suitable value for the latter lay in the range 16-64, and many studies were based on the intermediate value of 32. The convergence parameter is easier to determine if the desired convergence rate is known, but this is not always possible to specify beforehand and was experimentally derived in this study. Fast adaption was found to give best cancellation of nonstationary features of EEG activity, and a value of 0.2 was typically used for the normalised coefficient. It was necessary to develop an effective way of removing baseline fluctuations prior to adaptive filtering, and a comparison of several possible methods showed that a simple recursive digital filter could accomplish this satisfactorily as part of an on-line, continuous filter process.

The results of applying the ANC to a wide range of EEG activity in two experimental subjects showed that very significant reduction of activity is obtained by removal of correlated components in reference EEG sources, which in many cases permits the individual responses of these data to be made available for inspection. The correlated EEG activity primarily consists of in-phase components that may be attributed to either reference activity or a common EEG substrate for the scalp sites used in this study. The results are analysed in detail in the previous chapter, though the following discussion summarises the main conditions and presents these in the context of the overall strategy for EP signal processing.

## 9.2 Discussion of results.

The experimental results in the previous chapter have demonstrated that significant and often very substantial improvements in EP recordings can be obtained by using a correlated reference source to attenuate background EEG activity. This provides clear evidence that this investigation has been well justified. In many of the records the individual responses are clearly visible where previously they were hopelessly submerged in background activity. On the basis of these results, studies of single response properties are clearly possible, and may permit response variability, for example, to be studied in detail. While the correlated activity happened to be largely in-phase components of both channels in these studies, whether due to reference activity or common EEG sources underlying each electrode, this is unlikely to be the case for all possible subjects and electrode combinations, and hence the technique may prove to be a valuable tool in general EP methodology.

There are a number of assumptions that are implicit in the use of an ANC which will now be discussed. Foremost among these is the basic assumption that the EP signal is additively combined with the background (unevoked) EEG. While this assumption has not been specifically tested in the course of this investigation, it is clear that this method itself provides a useful means of doing so. Inspection of the individual records appears to show, in many cases at least, that the EPs are fairly consistent in terms of their amplitude and overall morphology. Their amplitude is often much larger than the background EEG, so it is difficult to reconcile this observation with the assertion of Basar et al [73] [130] that the EP is simply due to the stimulus acting upon existing EEG activity such that the different frequency components are temporarily brought into synchrony. This is particularly apparent in fig. 8.23 which shows very large single EPs in low amplitude EEG. The individual EPs appear to be consistent in form and bear close similarity to the average EP. While this result has of course been obtained for the special case of low amplitude EEG activity, it is probable that the same generation mechanisms are responsible for the EP in the other cases. It appears then that the EP primarily represents an additive disturbance component to the ongoing EEG signals, which justifies the assumption made earlier in performing this study.

The use of the simple LMS algorithm also appears to have been justified from an examination of the filter outputs, as the correlated components of different EEG channels exhibit fairly uniform properties over time. This is supported by the similarity of results obtained when different convergence rates are employed. Apart from nonstationary alpha activity, there is little difference between these results. The ANC solution also converges to the differential solution in many cases which is maintained throughout each record, as shown in fig. 8.29. This

provides good evidence for the stability of the filter solution in these circumstances. Nonstationary alpha activity can present some difficulties if it is persistently present and is of large amplitude, though episodic alpha activity is generally effectively attenuated. Fast adaption is necessary to track the rapid variations of alpha signal properties. Alpha activity appears to take place fairly independently in each channel, and though some cancellation is obtained when short-term correlations occur, or when alpha appears in the reference source alone, it is not possible to cancel alpha activity that takes place in the primary channel alone. Use of several reference sources may be useful in this regard as there is greater probability of alpha activity occurring in one of these to provide cancellation opportunities. The restricted time available for this study prevented an exploration of this possibility, though some initial results are shown in Appendix D and appear to indicate that this may be a worthwhile possibility for a future investigation.

A difficulty with the use of adaptive filters is the lack of a clearly defined theoretical framework regarding practical application, and the selection of suitable filter values necessitated a heuristic approach. The studies in this thesis have shown that successful noise cancelling can be accomplished using an estimate of filter order obtained by the spectral method suggested in Chapter 8. This is likely to be adequate for practical application of these filters, though it may not be appropriate if very nonstationary activity is present. Further research in this area is clearly desirable.

Adaptive noise cancelling of EP signals has been shown to be effective provided that a reference source is available that does not contain significant EP components of interest. Application to routine

EP processing is therefore restricted to cases where the scalp distributions of EPs are known or can be readily estimated, and may well have most value in research applications that involve detailed studies using a relatively limited number of long-term subjects. It is recommended that other techniques be routinely employed together with ANC, to guard against the possibility of spurious results. Experience gained in this study favours the use of dedicated instrumentation to perform the filtering operations, as general purpose computing equipment is rather slow for routine use, though it does offer the potential of detailed analysis of filter performance. Hardware designs are relatively straightforward to construct, and can offer real-time processing capability without incurring great expense. A discussion of hardware implementations is outside the scope of this thesis, and the reader is referred to the text by Cowan and Grant [112] for an introduction to this topic.

As has been noted in Chapter 8, adaptive filtering under the condition of fast adaption can be susceptible to transient instability if the data contain significant artifactual components. For routine use of the technique it is clearly necessary to employ artifact rejection procedures. Detection of these is relatively easy to perform, and can be based on amplitude or rate of change criteria. Some artifacts such as those due to eye blinks or eye movements can be avoided altogether by the use of linear nulling techniques [21] [22]. If this is not possible it may be necessary to halt adaptive operations for the duration of the detected artifact, much as the gated filter suspends weight revision from taking place during the signal epoch. This obviously requires further investigation if the technique is to be reliably used in routine recording.

## 9.3 Suggestions for further research.

There are three particular aspects of this investigation which could be fruitfully explored further. The first concerns the underlying concepts employed in this study, which are the assumptions that EP generation processes are independent of the spontaneous unevoked EEG activity such that the post-stimulus record can be considered as the linear summation of the two, and that EEG activity at multiple scalp sites contains components that are related by linear transformations and hence permit cancellation of these. While these assumptions may reasonably be made in an initial exploration of this approach, a deeper study would be desirable. In particular it would be beneficial to to know more clearly the properties of the EEG signals, in terms of their correlations and higher order functions in order to decide whether linear methods alone are likely to offer sufficient improvement of signals, or whether nonlinear methods are worth investigating. It would also be useful to have a model of the signal in terms of spatial and temporal properties to permit the benefits of noise cancelling to be assessed.

The second way in which this study could be usefully extended would be to increase the range of subjects and stimuli used, and to employ a greater number of electrode sites. While this would involve a considerable degree of effort, it is clear that routine use of a technique such as this must be properly validated under the range of different conditions that are likely to be faced. If it were contemplated to use this technique in the clinical setting, a study of clinical subjects would be required to ensure that the underlying basis of the technique was applicable in these subjects also. The extension of the study to include a wider range of electrodes and possibly other

sensory EPs is one that could be fruitfully performed without major departure from the strategy followed in this thesis. Investigation of the use of multiple-reference adaptive filters in particular may well yield profitable results. The linear array of electrodes used in this study was not suitable for this, and a two-dimensional array of electrodes would ideally be required. A further advantage of this arrangement is that it would permit a comparison with methods such as the source derivation technique of Hjorth [84]. A difficulty with multi-reference adaptive filters is that the theory governing their operation is not well developed, and a detailed investigation of their performance characteristics would first be required prior to applying these. This would be a rewarding area of research as this situation is commonly encountered with other physiological signals [48].

The third avenue of further research concerns the use of alternative adaptive filter methods, particularly the fast converging algorithms [104] [105]. Though noise cancelling using the simple LMS algorithm has been shown to be effective, and convergence rates close to the maximum attainable have been employed, it may well be that a further increase in adaption rate would yield better cancellation of nonstationary activity such as alpha or muscle activity. Adaptive filtering of nonstationary signals such as speech has been satisfactorily accomplished by these methods. Further development of the gated adaptive filter approach would be interesting to pursue, primarily to establish whether it is suitable in nonstationary environments. It would also be interesting to consider the suggestion made in Chapter 5 of employing different adaption rates during the different phases of the signal record. Other ways of using adaptive filters may also be worth investigating, such as the use of the channel enhancement mode. This may be beneficial in cases when correlated EP

signals are present over extensive scalp areas while the EEG itself is poorly correlated, and may form a useful complementary technique to adaptive noise cancelling when the scalp distributions do not permit use of this method. Finally, use of techniques such as these requires an investigation of artifact removal methods. These have already been developed to some extent to permit routine computer analysis of EEG signals, but would need to be evaluated in the context of these particular techniques.

In conclusion then, adaptive techniques are a promising alternative to traditional techniques of signal enhancement based on signal averaging. Some of the possibilities in applying these have been demonstrated in this thesis, though further research into their application to other sensory EPs and to the development of other filter strategies is likely to be promising, in particular the multi-reference and fast converging algorithms. It is considered that techniques such as these may well hold the key to routine analysis of single trial EPs in the future.

## APPENDIX A1 EXTRACT FROM THE DISPAC USER MANUAL

DISPAC MANUAL
Introduction

### 1.0 INTRODUCTION.

This manual describes the DISPAC program (for Digital Signal Processing Package) written in the Department of Physics, University of Keele. DISPAC is a modular software package which was developed to enable rapid evaluation of digital signal processing algorithms using one main structured program rather than a number of specific ones which often entails duplication, inconvenience and redundancy. It is conceived primarily as a development tool and so simple alteration of features was considered important. Speed of execution was sacrificed where memory savings would result as it was envisaged that a multiplicity of function was preferable. Major features are:

* interactive control by simple keyboard commands
* display, plotting and sampling functions available
* disk storage of data files, command files and parameter files
* creation, listing and execution of command files
* 32k integer data array space
* flexible user-definition and referencing of data arrays
* structured for general applications
* easily expanded or altered functions
* comprehensive data analysis and processing functions provided

It is comprised of one main program which communicates with the user, accepts simple commands and executes them by calling the appropriate subroutines. A number of subroutines are provided but the facility to easily add more commands was considered important. The basic design allows both system and user functions to be invoked using

the same command structure.

The implementation is biased towards our specific application and the particular constraints imposed by our computing resources, nevertheless much consideratiom was given to the generality of the solution to allow for wider application. The program is thus suitable for many applications requiring interactive control of computing operations, but is ideally suited to one involving mathematical manipulation of fairly large data records.

The program is written almost entirely in DIGITAL Fortran IV though a number of device-dependent routines are written in MACRO-11, the DIGITAL macro assembler. No apology is made for the use of non-ANSI Fortran IV features, which was dictated by reason of program development speed and efficiency of local implementation.

The program was implemented on a PDP 11/23 minicomputer with 128k bytes of memory and was written with the object of reserving the main 64k for program code, system tables and working space, whilst utilising the 64k of memory outside the normal addressing range for large data arrays. Peripherals include disk drives, hard copy printer, digital plotter, analogue to digital converter, real time clock and digital interface, though none of these is a requirement with the possible exception of the disk drive.

## 2.0 RUNNING DISPAC ON A PDP 11/23

As supplied DISPAC requires a PDP 11/23 minicomputer with 64k words of fast access memory running under the RT-11 operating system. At least one fast system device is needed such as floppy diskette or disk drive, and a console terminal such as a VDU. Additionally a hard copy printer, digital plotter, ADC card, digital interface card and second disk drive are ideally required.

The main program is called DISPAC.SAV and this should reside on the system device with the supporting system data files COMFIL.DAT and MESFIL.DAT and a basic RT-11 system (e.g. Single job monitor, SWAP.SYS, DIR.SYS, DUP.SYS, EDIT. Data files may be stored on the default device DK: which may be a second disk drive or the system device.

The program is started by typing:

```
R DISPAC
```

after the monitor prompt '.' . After a short delay the program clears the screen and displays the message:

```
DIGITAL SIGNAL PROCESSING PACKAGE   V2.0   JUL 1983
>
```

The '>' is the prompt for a command to be typed in, and appears whenever further keyboard input is awaited. This will be discussed in the next section. If this message does not appear, one of the following

messages should appear, which indicate an incorrectly prepared system disk:

```
STOP -- No command file COMFIL.DAT
STOP -- No message file MESFIL.DAT
STOP -- Bad command spec
KMON-F-Not found
```

In addition there are a number of system error conditions that may occur causing program failure. Insufficient memory is the most likely condition resulting in the following messages:

```
?ERR 62 FORTRAN start fail
?ERR 64 Virtual array initialization failure
```

If the following message appears, there is some inconsistency in COMFIL.DAT which may indicate that some command is improperly defined, though the program will proceed ignoring the error:

```
*** Warning - Bad argument type
```

DISPAC requires at least one (and preferably two) fast mass-access storage devices such as floppy diskette drive. The system device is the one containing the program file and system files, and the other is useful for storing data records. In the current implementation a VDU is used as the interactive console device and a DECwriter is used to obtain hard-copy output. For certain commands output may be routed to either, depending on the listing device selected. A digital parallel output

port is used for displaying waveforms on an oscilloscope. When in display mode, complete scans are generated repeatedly while the program is waiting for a command, and this is disabled while the command is executed. This requires special hardware to convert the digital signals into analogue form and to trigger the display scope at the start of each sweep. A Watanabe digital plotter is serviced via a serial line. The plotter commands controlling the pen motion to plot waveforms, axes etc. are generated by subroutines written for the purpose. The program was initially written to use the Watanabe DIGI-PLOT, but can also drive the Gould Colorwriter. There is provision for sampling up to 16 channels of analogue data via the ADC and multiplexer at rates up to 5 kHz, suitable for most bio-electric signals. These samples are stored im memory in a previously defined array and may be be stored on disk for subsequent analysis, or after further processing. All the peripherals require certain options to be selected, and these are set up by the appropriate parameters. It is incumbent on the user to ensure that these are correctly set up, otherwise spurious operation may result, though some elementary checks are performed in some cases.

## 2.1 COMMAND SYNTAX

Commands are typed following the prompt '>', though RT-11 permits input to be keyed in before it is required, a feature that should be used with caution. Characters wrongly typed may be corrected by pressing RUBOUT to delete previous characters, or CONTROL+U to delete the entire line, though both of these only act on the current line before RETURN is pressed which sends the line to the computer.

Instead of typing a sequence of commands repeatedly, it is possible to create a file containing these exactly as would be typed. The disk file is given an appropriate name, and has the extension .CMD to identify it. This facility is useful when the program is first started to select initial options or functions, and for command sequences that are repeatedly issued. An error that occurs during execution of such a command file will cause further commands to be ignored and return control to the direct keyboard mode, unless this option has been disabled. Operations that result in an unusual condition for any reason cause a message to be displayed on the console. There are 5 classes of message:

a) information - execution continues
b) warning - execution continues
c) error - current command failed
d) abort - current command was aborted by the user
e) fatal error - program is aborted

If it is desired to suspend output to the console for examination, pressing CONTROL+S achieves this. Listing continues by pressing CONTROL+Q. Output to the console during the current command can be skipped by pressing CONTROL+O at any time, and normal output will be resumed when the prompt for the next command is issued, or by pressing CONTROL+O again. The program is terminated by the STP command. The normal double CONTROL+C abort facility has been disabled to permit certain commands only to be aborted, but without stopping the program. Most commands can be programmed to make use of this facility.

Each command is initiated by typing its code name e.g. MAP followed by any arguments it may require, followed by RETURN. One or more spaces separate the command codes and arguments. The arguments may be of different types, depending on the command, and a check is made to ensure that a valid argument is entered. At present up to 6 may be permitted though some commands may use none. In addition some commands allow one or more arguments to be omitted if a specified default is acceptable. This can simplify use by minimising key strokes, though care must be taken to ensure that the intended action results. A warning message is issued whenever default action is taken. The purpose of arguments is to link the desired operands to the subroutine. This mechanism permits general commands to be written which allow the user to provide specific data to be used in each case. Examples will be given in succeeding sections to demonstrate usage. There are at present 4 basic categories of argument to cover most situations. These are as follows:

'I' represents an integer (or fixed-point) constant. It can take any value in the range -32768 to 32767. No decimal point is permitted. e.g. 271

'R' represents a real (or floating-point) constant. It is entered as a decimal number with an integer and fractional part and can optionally include an exponent as in scientific notation, though no embedded blanks are permitted e.g. 0.023 or 2.3E-2

'L' represents a literal (or character) string. These are delimited by

single quotes, e.g. 'EXPERIMENTAL RESULTS'.

Note 1. An error remaining from a previous revision requires that 3 character command or parameter names which use the RAD50 code must not have quotes when used as arguments to the following commands: H,AP,SP,LP,DP,INC,DEC,DAC.
In the present version only file names require quotes.
The final quote can be omitted if it is followed by RETURN.

'V' represents a vector or array of vectors
These will be discussed more fully in the next section.

It is important to note the distinction that for arguments of type I,R, and L the user supplies a value, be it numerical or literal, whereas for V arguments the user references a previously defined array or vector.

When the H (Help) command is issued, a list of helpful information for each command is produced. Each line contains:

a) a number (1-99) specifying the subroutine that performs the command
b) a code name which the user types to initiate the command
c) an argument list specifier eg. VIR1
   which describes which arguments are required, and
d) a brief help message to aid the user.

The subroutine number is not needed for normal useage, but is helpful in identifying which subroutine performs a certain task, or in

determining which subroutines are free to be added to the system. The argument list specifier is important, in that it must be understood for correct program use. In many cases the argument type will be obvious, and the user will quickly become familiarised with the system without having to consult the help information. In this example 3 arguments are required: a vector reference, an integer constant and a real constant. If the argument type is in lower case, this denotes an optional argument which may be omitted, such as the literal constant in the previous example. The help text will usually explain what default action results from omitting an argument, but the exact description of each command should be referred to, at least initially. Note that if two optional arguments are specified there are three valid entries viz. none, the first or both. It is not possible to enter only the second as it will be interpreted as the first because of the order of entry, unless the subroutine can deduce this in special cases.

Also the symbol '?' is used to denote an unknown argument type which cannot be specified explicitly but whose type will be apparent from the context. An example of this is the SP (Set Parameter) command which enables the user to assign a new value to a currently defined parameter. The format of the command is SP <parameter name> <value> but as the command is general, the type of the value will obviously depend on the parameter used which cannot be set beforehand. Thus SP has argument list specifier 'N?' and two examples of use might be:

>SP NUM 16

```
>SP TXT 'PLOT NO.  4'
```

## 2.2 MEMORY PARTITIONING AND DATA ARRAYS

Central to the design of DISPAC is the capability of handling large data records efficiently in core memory. Without dwelling on the issue, a brief explanation may prove helpful. There is a finite limit to the range of memory that may be directly accessed by any computer. For the PDP 11 this is 32k words which is quite adequate for many applications. However for the kind of application envisaged, processing is simplified if much more memory is available than that available in a 32k word system once system and program code are included.

The PDP 11 allows access to memory outside this range via a 'memory management unit', not as efficiently as the basic 32k word store but nevertheless much more efficiently than disk. The resultant solution was to reserve the basic 32k word store for program code, system tables and data arrays used in complex processing where fast access is most important, and use another 32k word store primarily as a large repository for data during processing, though also available for direct manipulation.

Another basic design consideration was flexible data structuring. This is because different situations require different sizes, structures and numbers of arrays. To avoid having to work within the constraints of a predefined structure, the solution devised was to allow the user to define his own arrays. The user then decides his own requirements,

allocates storage for each array he needs and supplies the corresponding structure. Currently there is provision for defining up to 26 arrays or vectors, using the letters A-Z as a simple reference name. The range of each dimension is also specified as is the absolute starting location in data memory. Once defined in this way the user need not be concerned with where the data resides in absolute memory, but always refers to the data using the letter code.

Because of the intended application, the definition was chosen to simplify use. The vector is taken as the basic data structure, and most commands operate in terms of notional vectors. This is convenient for the digital representation of a time series. Up to 3 dimensions are available, but the 3rd is always taken to index the vector elements. The remaining 2 may be used to define an ensemble or set of ensembles of vectors. For instance, the 2nd dimension may be used to index the analogue channel source, and the 1st dimension indexes ensembles of vectors, perhaps different experimental runs.

The general reference of any element in an array is X<i:><j><.k> where angle brackets enclose optional subscripts and

X - represents an array name A-Z

i - represents the subscript of the 1st dimension

j - represents the subscript of the 2nd dimension

k - represents the subscript of the 3rd dimension

If a subscript is omitted it is taken as 1, but if no subscripts are given the array name references the entire array as one linear

vector. However if the array letter is followed by an asterisk, this denotes that the command must reference each vector consecutively, known as 'auto-stepping'. These features allow the user the option of treating the array as one large data record or as an ensemble of independent vectors. For instance a filter command would properly act on each individual vector separately, whereas storing a data array on disk might be most conveniently performed using one disk file rather than a number of files.

Thus,

| | |
|---|---|
| B1:2.3 | specifies the 3rd element of the 2nd vector in the 1st ensemble |
| B2.3 | does the same (subscripts taken to be 1 if omitted) |
| B2: | specifies the subensemble of vectors in the 2nd dimension as one linear vector |
| B2 | specifies the 2nd vector in the 1st subensemble |
| B | specifies the entire array as one linear vector |
| B* | specifies each vector in the array to be accessed sequentially |

Note 1. As well as user-defined arrays A-Z, there are special 'system arrays' called $ and ◇. $ spans the entire data memory range, currently in 256 blocks containing 128 words. It is a useful mechanism for referencing any word in memory, and particularly when allocating storage to user arrays The notional blocks are a useful aid to managing memory allocation. $ should not be deallocated or redefined in any way (unless specifically desired), but ◇ is free for normal use, though it is likely that future revisions will have a special function reserved for it.

Note 2: As indexing in Fortran uses integers, there is a limit of 32767

on the size of an array. This means that the full 32k of memory available can not be accessed by this means, and any attempt to do so will result in a system error. Although $ is given as spanning this region, it also suffers from the same restriction, and accessing beyond data item $256.126 is liable to lead to erroneous operation. The TAB command which gives the size of an array has been specially doctored to yield 32767 instead of 32768.

Note 3: Arrays may be allocated overlapping storage, as this can be useful when used with care e.g. referencing a particular region of memory using differently structured arrays which span the same region, or to allow an area of memory to be shared by two different arrays which do not conflict in their useage.

Note 4: The order of subscript progression is the same as that for FORTRAN when considering the array storage, i.e. an array defined as I3:2.4 would be stored in memory in the following order:

I1:1.1 I1:1.2 I1:1.3 I1:1.4 I1:2.1 I1:2.2 I1:2.3 I1:2.4
I2:1.1 I2:1.2 I2:1.3 I2:1.4 I2:2.1 I2:2.2 I2:2.3 I2:2.4
I3:1.1 I3:1.2 I3:1.3 I3:1.4 I3:2.1 I3:2.2 I3:2.3 I3:2.4

Note 5: Care must be taken with commands that operate using two or more arrays to ensure that the intended operation results
e.g. for LPF which low pass filters the 1st vector and stores the result in the 2nd, the command

>LPF I* J*

would filter each vector in I and store the output in the corresponding vector in J, though this command would fail if there were a different

number of vectors in I from J. This command is powerful as it saves typing:

```
>LPF I1 J1

>LPF I2 J2

    etc.
```

This operation is quite different from

```
>LPF I J
```

which filters the array I as one linear vector and stores the result in J.

## 2.3 PARAMETERS

Parameters aid the use of DISPAC by allowing user control of various data elements that help define the processing environment or simplify the execution of commands. They are given a symbolic name and may be examined and modified. However they are really auxiliary data as they are usually not the main object of processing. Typically they are single elements e.g. integer constants or character strings. As with arguments, they may be of type V, I, R or L. Their main purpose is to communicate with the subroutines performing the various commands which often require subsidiary data in order to function properly. This data could be passed using arguments except that:

a) arguments are intended to be used when the value changes frequently and so must be specifiesd each time, whereas parameters generally are not subject to frequent change, thus saving unnecessary typing.

b) apart from arrays, arguments specify constants used by the command and thus only give data to the subroutine, while parameters refer to locations in memory that can be used to pass or receive data from subroutines.

c) only 6 arguments are available at present, so there needs to be a mechanism to enable more than 6 data items to be passed to the subroutine if required.

d) parameters are handled more efficiently since they do not need to be decoded as arguments do, and the subroutine operates on them directly.

The user has complete freedom in defining arrays and allocating storage, but subroutines which require parameters have certain locations specifically reserved for their use, and they will use that location independently of whether a parameter name has been correctly assigned to that location or not. In other words, the user does not have freedom to assign parameters as and where he wishes, but is merely setting up a mechanism to enable him to interact with the subroutine via a pre-defined location. The name is purely for convenience and may be chosen at any time to most appropriately describe the parameter, but the location, type and range of values is determined by the subroutine. In general use, it will only be required to examine and/or modify parameter

values to suit the needs of the moment.

There are commands to allocate a particular parameter cell a name, deallocate it, assign a value, and examine its value. However there are another two commands which simplify the use of parameters considerably. As it is envisaged that sets of parameters will be required fairly routinely, to obviate the need to define each of them explicitly each time the program is restarted, which would be tedious, it is possible to store an entire set of parameters, both definitions and values in force, in a named disk file so that the whole set may be read in at any time. This is also useful if the user has a number of different paradigms, each of which requires a different set of parameter definitions, as each can be read in at any time and replace the current set. These files are stored on the system disk in a binary file with a .PAR extension and consequently cannot be examined using the editor, unlike MESFIL.DAT and COMFIL.DAT. Alternatively a command file which allocates sets or subsets of parameters may be executed if this is more convenient.

## 3.0 BASIC COMMAND SET

### 3.1 HELP - H 1

This lists a line of help information for the given command or for all commands if none is specified. The format is a number (1-99) which

specifies the subroutine performing the command, the 3-character name used to call the command, the arguments expected, and a brief help text following a colon. This information appears in the same form as in CONFIL.DAT e.g.

```
>H H

1 H 1 :List Help information (named command only if specified)
```

## 3.2 ALLOCATE ARRAY - AL ??

This command is used to allocate a section of memory to an array and to define its structure. The first argument is a general vector reference where the subscripts define the maximum permissible range. The second argument specifies the starting position in memory to be that of the given array. A convenient way of doing this is to use the array $ which spans memory in blocks of 128 words e.g.

```
>AL I16:4.128 $10
```

creates an array of vectors of length 128, with 16 ensembles and 4 subensembles starting at block 10 i.e. word 1281. It is convenient but not necessary to specify a block boundary as the start, e.g. $10.65 would set the start location in the middle of block 10.

Note 1. Arrays may overlap in storage if this is desired, though care should be exercised if this is done.

Note 2. If the array is already defined, the previous definition will

be replaced by the new one, but a warning message will be issued.

Note 3. The data values spanned by the array are not themselves initialised, the command merely imposes a formal structure to reference a particular region of memory.

### 3.3 DEALLOCATE ARRAY - DE V

The definitions in force for the named array are removed, and further reference to this array by any command will cause it to fail.

### 3.4 DATA TABLE - TAB

This command lists the current definitions for all arrays. The start location (in words), dimension ranges and size of each array are shown, including the special purpose arrays ◊ and $. Arrays which are not defined have zero entries. Array sizes of 32768 are truncated to 32767.

### 3.5 MEMORY MAP - MAP v

A symbolic representation of data memory allocation is presented in terms of blocks. The array letter identifies blocks partially containing the named array (or all arrays if no argument is given). If more than one array is allocated storage in any block this is indicated by the symbol '+'. Unused blocks remain blank.

## 3.6 ALLOCATE PARAMETER - AP LLI?

This command is used to supply a name (and optionally an initial value) for a particular parameter required by a command. The name is entered first, then the parameter type, then the location in the parameter area optionally followed by an initial value. A null or zero value is the default. e.g

>AP IO I 2 6

names the 2nd integer parameter location IO and assigns it the value 6

Note 1. Several system functions are controlled by parameters, so these should not be altered and only unused cells or those used by subroutines should be allocated. In particular the 1st 10 integer cells are reserved for use by DISPAC.

Note 2. While integer and real constants use 1 parameter cell, obviously character strings depend on their length, and so the string location should be chosen to avoid overlapping previous strings. Also when entering character string values, the maximum possible length should not be exceeded as other strings following the desired one may be corrupted. This length will be obvious from the context or may be found by examining the current parameter definitions e.g. file names are restricted to 15 characters which is the largest valid file name. A null(0) character is appended at the end of each string so that the total length is one greater than the maximum number of characters.

Note 3. Due to the way in which parameters are stored, type V are

stored in the same area as type I parameters. As these are not often required, and anyway should only be allocated in accordance with the subroutine requirements, this should not be a problem. See the system description for more details of this.

### 3.7 SET PARAMETER - SP L?

Assign a value to the given parameter. The 1st argument is the name and the second is the value. e.g.

```
>SP IO 7
```

assigns 7 to IO.

### 3.8 DEALLOCATE PARAMETER - DP L

The current definition for the named parameter is removed, thus freeing that parameter cell for redefinition. This does not affect the value of the actual location, which will continue to be used by any subroutine accessing it, regardless of whether or not it is defined or correctly assigned. This command is useful if a parameter is no longer required in the list, or if it is to be renamed.

### 3.9 LIST PARAMETER - LP L

The current definition, location and value of the given parameter

is listed. If none is specified, the entire set of parameters is listed. e.g.

```
>LP IO
IO (I 2) = 7
```

### 3.10 WRITE PARAMETERS TO DISK - WRP L

The entire set of parameter definitions and values is stored on disk in a binary file with the given name and extension .PAR .e.g.

```
>WRP 'FILE5'
```

Note 1. If a file of the same name exists, it will be overwritten.
Note 2. The first 20 integer cells are reserved for system use and are not stored on disk file, even if they are changed.

### 3.11 READ PARAMETERS FROM FILE - RDP L

The named file is read and replaces the current definitions and values of all parameters e.g.

```
>RDP 'JOHN'
```

### 3.12 EXECUTE COMMAND FILE - EX L

The commands contained in the named ASCII file are executed in

sequence. Action resulting if an error occurs in this sequence depends on the value of the parameter TRP as follows:

TRP = 1 - proceed with the remaining commands

TRP = 2 - abort remaining command stream on file

Note. The display is switched off while these commands are being executed.

## 3.13 CREATE COMMAND FILE - CR L

Allows the user to create a new file with the given name and extension .CMD to contain a stream of commands as would be typed on the keyboard. These are entered following the prompt '*', and the file is closed when RETURN is pressed immediately after a '*' e.g.

```
>CR 'DO'
*H
*AL I16.128 $1
*MAP
*
```

Note. If a file of the same name exists it will be overwritten.

## 3.14 LIST FILE - LF L

Lists the named ASCII file having extension .CMD. This allows the contents of a command file to be examined e.g.

```
>LF 'DO'
H
AL I16.128 $1
MAP
```

### 3.15 LIST ARRAY - L v

The named array is listed as one large block of data. If no array is specified, the currently displayed vector is listed.

### 3.16 STOP - STP

This is the normal program exit command. Control returns to the monitor

### 3.17 DISPLAY ARRAY - D V

This command causes the named vector or array to be displayed on the CRO screen. It also defines the default argument for many commands which commonly act on the display.
Note. At present only vectors of size equal to a power of 2 are correctly displayed.

### 3.18 COPY ARRAY - CPY VV

Copy all the data in the 1st array to the 2nd. The sizes of each must agree. e.g.

>CPY I J

3.19 WRITE ARRAY TO DISK - W v1

Store the given array on disk file in binary form using the given name. If a name is given, the extension is automatically updated following storage and this name is stored in parameter DWF, the default write file. If no name is given this default name is used. If no array is named the displayed vector is used.

Note 1. Any existing file, or set of files, of the same name is liable to be overwritten.

3.20 READ DATA FROM DISK - R V1

The data in the named data file is read into the given array. A check is made to ensure that the size agrees with that stored on disk. If no filename is supplied the default stored in parameter DRF is used. The file extension is automatically incremented (if it is numerical with 3 digits) following successful use of this command.

## APPENDIX A2 LIST OF CURRENT DISPAC COMMANDS

### S Y S T E M

```
H    l      :List help information [named command only if specified]
EX   L      :Execute the commands in the named command file
CR   L      :Create a command file
LF   L      :List the named ASCII file
STP         :Stop the program
DMP         :Dump system tables
ARG vvirl   :Dump LINK information
VDU         :Send all output to the VDU
PRI         :Send all output to the printer
REM         :Treat this command as a remark
BEL         :Ring the TTY bell
```

### P A R A M E T E R S

```
AP   LLI?   :Allocate a parameter given name, type, location, [value]
SP   L?     :Set the named parameter to the given value
DP   L      :Deallocate the named parameter
LP   l      :List the values of all parameters [or just the named one]
WRP  L      :Write all parameter information to named file (drive 0)
RDP  L      :Read all parameter information from named file (drive 0)
INC  Lr     :Increment named parameter by given constant [default 1]
DEC  Lr     :Decrement named parameter by given constant [default 1]
```

### D A T A  S T R U C T U R E S

```
AL   ??     :Allocate space for named array given dimensions and start
DE   V      :Deallocate storage for the named array
TAB         :List the table for data arrays showing dimensions and start
MAP  v      :List a map of storage allocation [just the named array if given]
TRA  V      :Transpose the 2D array of vectors in storage and update table
```

### S A M P L I N G ,  P L O T T I N G  A N D  D I S P L A Y

```
L    v      :List the named array
SAM  V      :Sample ADC and store in array
P    v      :Plot vector [default display]
D    V      :Display vector on channel 1 (sets default vector)
DIS  Vi     :Display vector [section size opt.] using <- -> HOME RETURN
D2   V      :Display vector on channel 2 of display
XY   ?      :Control the plotter directly using the standard ASCII commands
DAC  Vl     :Output vector to DACs once or repetitively [ONE/REP]. C C aborts
AX   LIIll  :Plot axes using T/LIN/LOG, NDX, NDY, [y-axis text], [BOX]
G    ?      :Drive the GOULD Colourwriter directly
V    Vii    :Plot a crude graph on the VDU like D. Optional max/min limits.
```

### V E C T O R  A R I T H M E T I C

```
SC   Rvi    :Scale the vector [default display ] by constant, opt. clip value
SH   Iv     :Shift the vector [default display] by a constant
AD   Vv     :Add the 1st vector to the second [default display vector]
SUB  Vv     :Subtract 1st vector from 2nd vector [default display]
MUL  VVi    :Multiply 1st vector by 2nd vector, scaling down by given constant
DIV  VVi    :Divide 1st vector by 2nd vector, scaling up by given constant
```

## D A T A  M A N I P U L A T I O N

```
SET VI     :Set all the elements of the vector to the given value
FIL        :Fill memory with random data
ZER        :Zero out virtual memory
C   v      :Clear the vector [default display] to zero
CPY VV     :Copy the 1st array to the 2nd
W   vl     :Write the array to disk 1 using the default name if none given
R   Vl     :Read array from disk file using the default name if none given
NG1 V      :Noise generator. B/width, random □ seeds & freq. steps in params
SIN VRir   :Generate sine of given period, [amplitude and phase]
DCM VV     :Decimate the 1st vector, output to the 2nd (of half size)
SEG VV     :Segment 1st array into 2nd, no. of segments given by 2nd vector
REV v      :Reverse the order of the elements in the [display] vector
NG2 Vrrii  :Gaussian random (not white) noise. [St. dev, mean, random seeds]
NG  Vrrr   :Stationary noise generator given freq range and freq steps in Hz
```

## S I M P L E  S T A T I S T I C A L  O P E R A T I O N S

```
DC  v      :Remove DC mean from vector [default display]
DT  v      :Remove linear trend and DC mean from vector [default display]
WIN v      :Window the vector using cosine bell [default display]
HST VVi    :Form frequency histogram of 1st vector in 2nd vector [given range]
S   vv     :Simple statistics - min, max, mean, sum, standard dev, rms power
SNS VV     :Signal/noise power ratio given noisy signal and signal vectors
SNR VV     :Signal to noise routine, using mse and modified loss function
AV  VVv    :Form average [and st. dev.] of data array in given vectors
```

## R E A L  B U F F E R

```
GET V      :Get real and imaginary (scaled) vectors from complex buffer
PUT V      :Put real and imaginary contiguous vectors into complex buffer
SQR        :Square the complex buffer
LOG        :Compute the log of the complex buffer
LC  ii     :List real buffer given size, or start and size [defaults 1,2048]
CLR ii     :Clear real buffer given size, or start and size [defaults 1,2048]
ST  Vi     :Store in the vector the scaled data buffer contents given start
LD  Vi     :Load the scaled vector into real data buffer given start [=1]
```

## V A R I O U S  F I L T E R S

```
INI iii    :Initialising TVF [given bins, subensembles, smoothing factor]
TVF VVv    :Time Varying Filtering given input, o/p filter [and average] vectors
TF  vv     :Transversal filter using same weights (+WO) as AF. Also param N.
LPF vr     :1st order recursive low-pass filter [given vector and corner freq.]
MON VViii  :Monostable, given i/p and o/p vectors, [and edge,delay,duration]
LEV Vviii  :Level detector, given i/p [& o/p vectors, thresh. hyst. & o/p amp]
GLP RR     :Generate FIR lowpass coeffs. Need centre and width frequencies
PZN vr     :Pole-zero notch filter. Require vector and pole position (R).
MA1 VVi    :Moving average recursive LPF. Needs i/p o/p [and filter size]
```

## S P E C T R A L  A R I T H M E T I C

```
FFT        :Compute forward FFT of complex buffer
INV        :Compute inverse FFT of complex buffer
SPE        :Obtain spectrum of FFT in complex buffer
SXY VViill:Compute joint spectra for X and Y [size, %ovrlp, window, lin/circ]
LPS Vv     :Compute log power spectrum[s] for last SXY data
ACF Vv     :Compute autocorrelation[s] for last SXY data
COH Vvvv   :Compute coherence, [cross-spectrum, phase & |transfer fn.|]
CCF Vv     :Compute crosscorrelation [& impulse response] for SXY data
NTC VI     :Normalised transformed magnitude coherence. Need no of averages
LIP Vvvr   :Get the linear power spectra Gxx, Gyy and Gxy using SCl to scale
EIG V      :Compute eigenvalues of correlation matrix of V and list
```

# A D A P T I V E  F I L T E R S

```
TRF        :Compute filter output Y given signal and weight vectors
UPD        :Update weight vector given E, U, signal and weight vectors
ATF VVVv   :Basic LMS adaptive filter given X, D, Y [and E] + parameters
AF  VVVv   :Simple adaptive filter. Need X,D,Y,[E] + params N,U,UO,DEL
TAF VVVVvl:Time-sequenced AF given X, D, Y, Control vector, [E], [CON]
LSL VVVV   :Least squares lattice AF, normalized, pre-windowed, X D Y E form
LAT VVVVVl:LSL filter, needs X D, gives e r ex. Use DEB to list all
MAF VVVVvl:Multichannel AF. Uses X D Y E [C] [CON]. D in array X.
```

KEY:-
- I - integer constant
- R - real constant
- L - literal string constant
- V - vector
- ? - unspecified argument type; depends on context

i, r, l and v denote optional arguments which if omitted are assigned default values.

## APPENDIX A3 LIST OF PARAMETERS

| Name | Type | Cell | Value | Function | Used in command |
|---|---|---|---|---|---|
| ST1 | (I | 11) = | 1 | Start address (1) | D D2 D1S |
| SZ1 | (I | 12) = | 256 | Vector size (1) | |
| ST2 | (I | 13) = | 1 | Start address (2) | |
| SZ2 | (I | 14) = | 1 | Vector size (2) | |
| SWP | (I | 15) = | 256 | Sweep size | |
| DM | (I | 16) = | 0 | Display mode | |
| TRP | (I | 19) = | 2 | Error trap | EX |
| ECH | (I | 20) = | 1 | Echo/No echo | |
| IX | (I | 21) = | 0 | Initial plot position | P AX |
| IY | (I | 22) = | 2200 | | |
| LX | (I | 23) = | 3300 | Page limits | |
| HX | (I | 24) = | 2600 | | |
| XP | (I | 25) = | 256 | Plot limits | |
| AP | (I | 26) = | 400 | | |
| INC | (I | 27) = | 2 | Pen increment step | |
| SEQ | (I | 28) = | -1 | Automatic plot sequence | |
| LT | (I | 29) = | 0 | Line type | |
| AS | (I | 30) = | 10 | Character size | |
| INX | (I | 31) = | 300 | Auto-plot increments | |
| INY | (I | 32) = | -200 | | |
| ORX | (I | 33) = | 0 | Auto-plot origin | |
| ORY | (I | 34) = | 2200 | | |
| RUN | (I | 36) = | 16 | No. of runs | S |
| NCH | (I | 37) = | 4 | No. of channels | |
| BIN | (I | 38) = | 256 | No. of bins | |
| SAM | (I | 39) = | 4000 | Sampling period | |
| SM | (I | 40) = | 0 | Sampling mode | |
| POV | (I | 41) = | 0 | % overlap | SXY |
| NFT | (I | 42) = | 256 | FFT size (augmented) | |
| LL | (I | 43) = | 0 | No. of subensembles | TVF 1N1 |
| ISM | (I | 44) = | -1 | Smoothing factor | |
| RS1 | (I | 45) = | 0 | Random seed (1) | NG NG1 NG2 |
| RS2 | (I | 46) = | 0 | Random seed(2) | |
| NFS | (I | 47) = | 1 | No. of frequency steps | |
| PLO | (I | 48) = | 128 | Period of lowest frequency | |
| PHI | (I | 49) = | 2 | Period of highest frequency | |
| RNG | (I | 50) = | 32767 | Histogram amplitude range | HST |
| I | (I | 51) = | 0 | Loop counters | AF ATF TAF MAF |
| J | (I | 52) = | 0 | | |
| K | (I | 53) = | 0 | | |
| L | (I | 54) = | 0 | | |
| N | (I | 55) = | 16 | Filter order | |
| DEL | (I | 56) = | 9 | Delay parameter | |
| KEY | (I | 57) = | 0 | Data protection key | SXY LPS LIP COH ACF CCF NTC |
| SW1 | (I | 58) = | 3 | Detrending switch | |
| NNN | (I | 59) = | 256 | Segment size | |
| EDG | (I | 60) = | 100 | Edge polarity and size | MON |
| DLY | (I | 61) = | 0 | Output pulse delay | |
| DUR | (I | 62) = | 1 | Output pulse duration | |
| DAC | (I | 63) = | 0 | DAC output control | DAC |
| THR | (I | 64) = | 0 | Schmitt input threshold | LEV |
| HYS | (I | 65) = | 0 | Hysteresis | |
| AMP | (I | 66) = | 1000 | Output pulse amplitude | |

| Name | Type | Cell | | Value | Function | Used in command |
|---|---|---|---|---|---|---|
| YS | (R | 1) | = | 0.200000E-01 | Plot scale factor | P |
| SC1 | (R | 2) | = | 1.00000 | Scale factors | Type conversions |
| SC2 | (R | 3) | = | 1.00000 | | |
| U | (R | 4) | = | 0.000000 | Convergence coefficient | AF ATF TAF MAF |
| Y | (R | 5) | = | 0.000000 | Filter output sample | |
| E | (R | 6) | = | 0.000000 | Error output sample | |
| PWR | (R | 7) | = | 0.000000 | Input rms value | |
| SML | (R | 8) | = | 0.100000E-29 | Smallest real value | SXY |
| FC | (R | 9) | = | 50.0000 | Cut-off frequency | LPF |
| U0 | (R | 10) | = | 0.000000 | Bias tap convergence | AF |
| SF | (R | 11) | = | 1.00000 | Spectrum scale factor | SXY LIP LPS |
| LMD | (R | 12) | = | 1.00000 | Lattice filter parameter | LSL |
| DWF | (L101) | | = | OUTPUT.007 | Default output file name | W R |
| DRF | (L117) | | = | DATA.001 | Default input file name | |
| PLT | (I100) | | = | 0 | Select plotter device | P |

## APPENDIX A4 EXAMPLE OF CONSOLE DIALOGUE
(User input is underlined)

```
        DIGITAL SIGNAL PROCESSING PACKAGE     V2.0     JUL 1983

>REM - THIS IS AN EXAMPLE OF TYPICAL CONSOLE INTERACTION.

>RDP 'RW

>ZER

>AL S2.32 $1

>AL N16.64 $2

>AL D16.64 $10

>AL Y16.64 $18

>AL E16.64 $26

>MAP

  256 Blocks of 128 words

BLOCK |  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
------------------------------------------------------------------
   0+ |  S  N  N  N  N  N  N  N  N  D  D  D  D  D  D  D  D  Y  Y  Y
  20+ |  Y  Y  Y  Y  Y  E  E  E  E  E  E  E  E
  40+ |
  60+ |
  80+ |
 100+ |
 120+ |
 140+ |
 160+ |
 180+ |
 200+ |
 220+ |
 240+ |

>TAB

NAME ADDR DIM1 DIM2 DIM3  SIZE  |  NAME ADDR DIM1 DIM2 DIM3  SIZE
---- ---- ---- ---- ----  ----  |  ---- ---- ---- ---- ----  ----
  ◇    0    0    0    0     0   |    N  129    1   16   64  1024
  A    0    0    0    0     0   |    O    0    0    0    0     0
  B    0    0    0    0     0   |    P    0    0    0    0     0
  C    0    0    0    0     0   |    Q    0    0    0    0     0
  D 1153    1   16   64  1024   |    R    0    0    0    0     0
  E 3201    1   16   64  1024   |    S    1    1    2   32    64
  F    0    0    0    0     0   |    T    0    0    0    0     0
  G    0    0    0    0     0   |    U    0    0    0    0     0
  H    0    0    0    0     0   |    V    0    0    0    0     0
  I    0    0    0    0     0   |    W    0    0    0    0     0
  J    0    0    0    0     0   |    X    0    0    0    0     0
  K    0    0    0    0     0   |    Y 2177    1   16   64  1024
  L    0    0    0    0     0   |    Z    0    0    0    0     0
  M    0    0    0    0     0   |    $    1    1  256  128 32767
```

```
>REM - GENERATE ARTIFICIAL SIGNAL AND NOISE

>SIN S1 16 1000

>NG2 N 500

>L S
   382    707    923   1000    923    707    382      0   -382   -707
  -923  -1000   -923   -707   -382      0    382    707    923   1000
   923    707    382      0   -382   -707   -923  -1000   -923   -707
  -382      0      0      0      0      0      0      0      0      0
     0      0      0      0      0      0      0      0      0      0
     0      0      0      0      0      0      0      0      0      0
     0      0      0      0

>S N

MINIMUM        MAXIMUM      MEAN          SUM        STANDARD DEV   RMS POWER
  -1817           1512  0.611328       626.000          498.261       498.018

>CPY N D

>AD S D*

>SNS D* S
S/N POWER RATIO = 0.839339      =-0.760625      DB
S/N POWER RATIO =  1.05321      = 0.225160      DB
S/N POWER RATIO = 0.910180      =-0.408728      DB
S/N POWER RATIO = 0.843725      =-0.737993      DB
S/N POWER RATIO = 0.860628      =-0.651848      DB
S/N POWER RATIO =  1.33198      =  1.24499      DB
S/N POWER RATIO = 0.801855      =-0.959042      DB
S/N POWER RATIO = 0.920236      =-0.361007      DB
S/N POWER RATIO =  1.52875      =  1.84336      DB
S/N POWER RATIO = 0.988969      =-0.481718E-01  DB
S/N POWER RATIO =  1.08067      = 0.336930      DB
S/N POWER RATIO = 0.879942      =-0.555461      DB
S/N POWER RATIO =  1.49436      =  1.74455      DB
S/N POWER RATIO = 0.877833      =-0.565881      DB
S/N POWER RATIO =  1.49294      =  1.74041      DB
S/N POWER RATIO = 0.985145      =-0.649988E-01  DB

>W S 'SIGNAL

>W N 'REF
W
>D 'PRI

>REM - NOW PERFORM ADAPTIVE FILTERING OF DATA

>AL 016.64 E1.17

>SP N 32

>SP DEL 16

>SP U .2

>AF N D Y E
```

```
>SNS O* S
S/N POWER RATIO =  1.21046     = 0.829492     DB
S/N POWER RATIO =  1.61850     =  2.09112     DB
S/N POWER RATIO =  5.34677     =  7.28091     DB
S/N POWER RATIO =  2.83534     =  4.52605     DB
S/N POWER RATIO =  5.47091     =  7.38060     DB
S/N POWER RATIO =  4.71634     =  6.73605     DB
S/N POWER RATIO =  5.16057     =  7.12698     DB
S/N POWER RATIO =  3.71109     =  5.69502     DB
S/N POWER RATIO =  6.87492     =  8.37267     DB
S/N POWER RATIO =  2.86942     =  4.57795     DB
S/N POWER RATIO =  4.50683     =  6.53871     DB
S/N POWER RATIO =  3.93519     =  5.94966     DB
S/N POWER RATIO =  4.00530     =  6.02635     DB
S/N POWER RATIO =  2.90414     =  4.63018     DB
S/N POWER RATIO =  12.1023     =  10.8287     DB
S/N POWER RATIO =  9.12117     =  9.60050     DB

>REM - REPEAT USING A GATED ADAPTIVE FILTER

>REM - CREATE CONTROL VECTOR

>AL C16.64 $41

>AL Z2.32 C

>SET Z1 -1

>SET Z2 1

>CPY C1 C*

>CLR


>C 0

>TAF N D Y C E


>SNS O* S
S/N POWER RATIO = 0.927657     =-0.326127     DB
S/N POWER RATIO =  2.53787     =  4.04469     DB
S/N POWER RATIO =  4.17633     =  6.20795     DB
S/N POWER RATIO =  11.6804     =  10.6746     DB
S/N POWER RATIO =  16.6410     =  12.2118     DB
S/N POWER RATIO =  36.6270     =  15.6380     DB
S/N POWER RATIO =  41.6238     =  16.1934     DB
S/N POWER RATIO =  123.602     =  20.9203     DB
S/N POWER RATIO =  187.160     =  22.7221     DB
S/N POWER RATIO =  274.702     =  24.3886     DB
S/N POWER RATIO =  294.321     =  24.6882     DB
S/N POWER RATIO =  827.891     =  29.1797     DB
S/N POWER RATIO =  5388.27     =  37.3145     DB
S/N POWER RATIO =  4975.60     =  36.9685     DB
S/N POWER RATIO =  9501.56     =  39.7780     DB
S/N POWER RATIO =  16876.0     =  42.2727     DB

>W O 'OUTPUT

>STP

STOP --  DISPAC
```

## APPENDIX A5. SUBROUTINE CALLS AND ARGUMENT TRANSMISSION.

The block diagrams contained in this Appendix show the basic structure of the main DISPAC code, and the form of each subroutine implementing a command function. Each command subroutine receives coded information pertaining to the arguments that have been entered at the keyboard after each command. Because DEC Fortran IV does not handle argument transmission to subroutines efficiently, the following scheme to do this is implemented. Each subroutine has a COMMON block containing variables and arrays used to pass arguments and parameters. The main section of DISPAC decodes the typed arguments according to the command definition, and after checking these for validity puts them into the link variables. The Fortran code for the COMMON block is as follows:

```
VIRTUAL IDATA(32767)
COMMON BLOCK /LINK/ IARG, IERR, NOP, JSUB, IPARM, RPARM, LPARM
LOGICAL*1 LPARM(200)
INTEGER IARG(2,6), IPARM(100)
REAL ARG(6), RPARM(25)
EQUIVALENCE (IARG, ARG)
```

The variables and arrays have the following function:-

IDATA is an integer array that spans the entire 32k words of data memory. All operations on data arrays must therefore use IDATA.

IARG, ARG, JSUB and NOP are set by the command line decoder to pass arguments to the command subroutine.

JSUB is the number of the subroutine that has been called.

NOP is the number of arguments that have actually been supplied,

ARG and IARG are used to pass up to 6 arguments. The n'th argument is passed as follows: Real constants are stored in ARG(n); Integer constants are stored in IARG(1,n); String constants are referred

## APPENDIX A5. SUBROUTINE CALLS AND ARGUMENT TRANSMISSION.

The block diagrams contained in this Appendix show the basic structure of the main DISPAC code, and the form of each subroutine implementing a command function. Each command subroutine receives coded information pertaining to the arguments that have been entered at the keyboard after each command. Because DEC Fortran IV does not handle argument transmission to subroutines efficiently, the following scheme to do this is implemented. Each subroutine has a COMMON block containing variables and arrays used to pass arguments and parameters. The main section of DISPAC decodes the typed arguments according to the command definition, and after checking these for validity puts them into the link variables. The Fortran code for the COMMON block is as follows:

```
VIRTUAL IDATA(32767)
COMMON BLOCK /LINK/ IARG, IERR, NOP, JSUB, IPARM, RPARM, LPARM
LOGICAL*1 LPARM(200)
INTEGER IARG(2,6), IPARM(100)
REAL ARG(6), RPARM(25)
EQUIVALENCE (IARG, ARG)
```

The variables and arrays have the following function:-

IDATA is an integer array that spans the entire 32k words of data memory. All operations on data arrays must therefore use IDATA.

IARG, ARG, JSUB and NOP are set by the command line decoder to pass arguments to the command subroutine.

JSUB is the number of the subroutine that has been called.

NOP is the number of arguments that have actually been supplied,

ARG and IARG are used to pass up to 6 arguments. The n'th argument is passed as follows: Real constants are stored in ARG(n); Integer constants are stored in IARG(1,n); String constants are referred

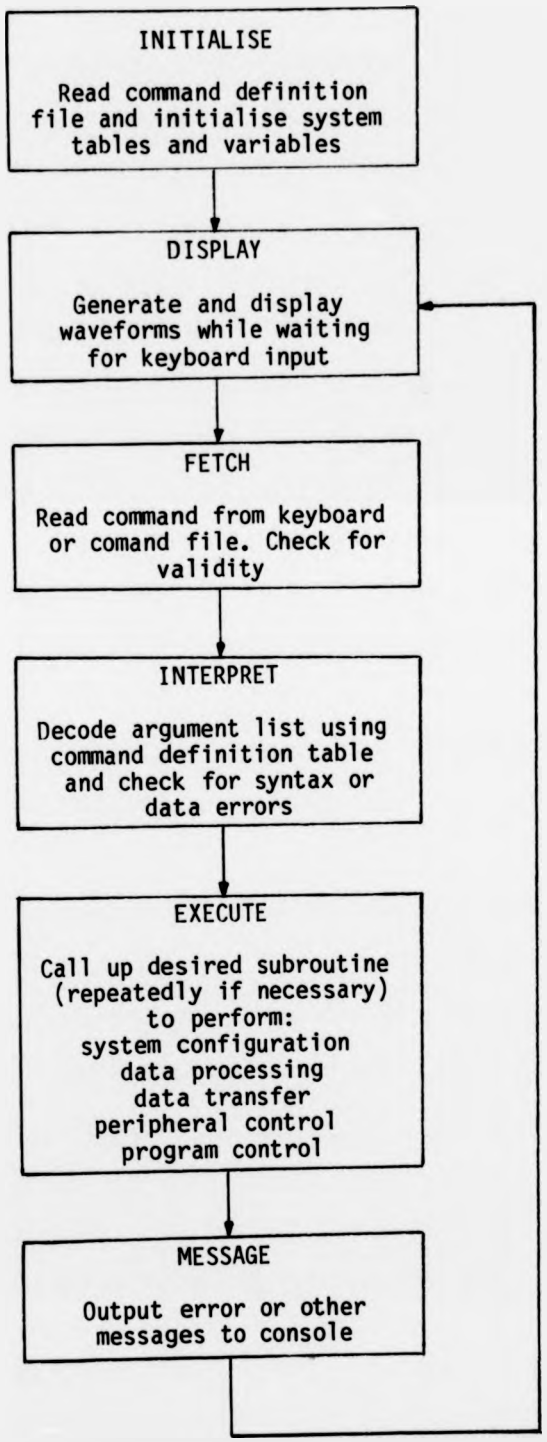
to by a pointer IARG(1,n) to denote their position in the command line and their size IARG(2,n); Data vectors are referred to by their base address in IARG(1,n) and size in IARG(2,n) to denote their location in IDATA.

IERR returns an error code to the main calling program on completion of the command.

The memory required to transmit arguments is thus limited to 13 words.

Parameters are made available via the arrays IPARM, RPARM and LPARM, which can store up to 100 integer parameters, 25 real parameters and 200 characters in the current implementation. Each cell is allocated to hold a particular parameter, thus ST1 which points to the start of the display vector is stored in cell 11 of IPARM. The subroutine must access the appropriate parameter cell, and this is conveniently done by assigning a local variable to the parameter cell using an EQUIVALENCE statement. The subroutine may then use the parameter cell as required, either to obtain the parameter value or to update it. An example of this useage can be seen in Appendix A6 which lists a typical command subroutine.

**BLOCK DIAGRAM OF THE MAIN STRUCTURE OF DISPAC.**



## GENERAL STRUCTURE OF COMMAND SUBROUTINES.

A. Subroutine title and initial description of function.

B. Common blocks to provide access to global variables and arrays.
COMMON BLOCK LINK - tables and variables through which arguments and parameters are made available. Present in every subroutine.
COMMON BLOCK DATA - represents the region of memory available to the user subroutine for real or complex arithmetic operations. Present only in subroutines that require to use this area.
COMMON BLOCK DECARG - contains the results of decoding a particular argument. Used only when the subroutine itself must decode an argument.
COMMON BLOCK TABLES - contains command definition table, parameter definition table, command decoder syntax tables, data definition tables. Only present when the subroutine requires access to these tables.
COMMON BLOCK LIMITS - contains the current sizes of each table. Only present when the subroutine requires these.

C. Fetch the arguments entered at the keyboard; these are either constants or pointers to data items in memory. Take default action if any are omitted. Check that sufficient arguments have been entered and that they are valid. Link local variables to parameter table and check that they are valid.

D. Perform the main subroutine code, whether accessing files controlling peripherals, altering system tables or manipulating data in memory.

E. Set error flag if an error occurs in any of the above procedures. The error number indicates both the severity of the error (and hence corresponding action) and the message in the error file that is to be output.

F. Return to the main calling program.

## APPENDIX A6 EXAMPLE OF A COMMAND SUBROUTINE (ADAPTIVE FILTER)

```
      SUBROUTINE SUB61(IDATA)
C
C     **********************************
C     *       SUBROUTINE AF            *
C     **********************************
C
C     LMS adaptive filter subroutine. Simplest and fastest algorithm.
C
C     The following statements define the argument and parameter
C     link areas. Use EQUIVALENCE statements to refer to parameter
C     cells by meaningful names. The COMMON /DATA/ block is the
C     real buffer which is used to perform the filter operations.
C     Provision is made for up to 512 weights. TDL is the tapped
C     delay line and DEL implements the delay in the primary input.
C
      VIRTUAL IDATA(32767)
      COMMON /LINK/ IARG, IERR, NOP, JSUB, IPARM, RPARM, LPARM
      LOGICAL*1 LPARM(200)
      INTEGER IARG(2,6), IPARM(100)
      REAL ARG(6), RPARM(25)
      EQUIVALENCE (ARG,IARG)
C
      EQUIVALENCE (IPARM(6),ICC)
      EQUIVALENCE (IPARM(51),I), (IPARM(52),J)
      EQUIVALENCE (IPARM(55),N), (IPARM(56),IDEL)
      EQUIVALENCE (RPARM(7),PWR), (RPARM(10),U0)
      EQUIVALENCE (RPARM(4),U),  (RPARM(5),Y),  (RPARM(6),E)
C
      COMMON/DATA/ W0, W, TDL, DEL
      DIMENSION TDL(512), W(512), DEL(512)
C
C     Check that at least 3 arguments are entered.
C     Set pointers to the start addresses of the data vectors.
C     Check the sizes of each to make sure they agree.
C     Check validity of parameter values.
C
      IF(NOP .LT. 3) GOTO 97
      IX=IARG(1,1)-1
      ID=IARG(1,2)-1
      IY=IARG(1,3)-1
      IE=IARG(1,4)-1
      N1=IARG(2,1)
      IF((IARG(2,2) .NE. N1) .OR. (IARG(2,3) .NE. N1) .OR.
     ◇ ((NOP .EQ. 4) .AND. (IARG(2,4) .NE. N1))) GOTO 99
      IF((N .LE. 0) .OR. (IDEL .LT. 0) .OR. (I .LT. 0)
     ◇   .OR. (U0 .LT. 0.0)
     ◇   .OR. (U .GT. 1.0) .OR. (U .LT. 0.0)) GOTO 98
      IF((N .GT. 512) .OR. (IDEL .GE. N)) GOTO 98
C
C     Compute basic constants and power of X vector
C
      PWR=0.0
      DO 50 I=IX+1,IX+N1
      PWR=PWR+FLOAT(IDATA(I))**2
50    CONTINUE
      PWR=PWR/N1
      IF(PWR .EQ. 0) GOTO 96
      U1=U0*2.0
      U2=U*2.0/(PWR*N)
```

```
C
C       Set the NEXT label to skip storing the E values if not required
C
        ASSIGN 400 TO NEXT
        IF(NOP .EQ. 3) ASSIGN 500 TO NEXT
C
C       Now process the vectors in one loop containing a basic
C       transversal filter computation and filter weight update process.
C       Implement 2 circular buffers TDL and DEL in real arrays as the
C       heart of the process.
C       Note that in both these cases the TDL is traversed in the
C       opposite direction from W and to implement this efficiently
C       2 loops are used with the first decreasing IT (for TDL) from I
C       to 1 and the next from N to I+1. IW indexes W. I and I2 index
C       the output and input samples of DEL.
C
        DO 500 J=1,N1
        IF(ICC .NE. 0) GOTO 95
        I=I+1
        IF(I .GT. N) I=1
        IT=I
        TDL(I)=IDATA(IX+J)
C
C       This is the transversal filter computation for one sample time
C
        Y=W0
        DO 100 IW=1,I
        Y=Y+TDL(IT)*W(IW)
        IT=IT-1
100     CONTINUE
        IT=N
        IF(I .EQ. N) GOTO 190
        DO 150 IW=I+1,N
        Y=Y+TDL(IT)*W(IW)
        IT=IT-1
150     CONTINUE
C
C       This is the LMS update algorithm
C
190     I2=I+IDEL
        IF(I2 .GT. N) I2=I2-N
        DEL(I2)=IDATA(ID+J)
        E=DEL(I)-Y
        UE2=U2*E
        W0=W0+U1*E
        DO 200 IW=1,I
        W(IW)=W(IW)+UE2*TDL(IT)
        IT=IT-1
200     CONTINUE
        IF(I .EQ. N) GOTO 300
        IT=N
        DO 250 IW=I+1,N
        W(IW)=W(IW)+UE2*TDL(IT)
        IT=IT-1
250     CONTINUE
```

```
C
C       Now store the filter output (and error) values
C       If integer overflow would occur, truncate the values to a
C       magnitude of 32767 and set the error condition appropriately.
C
300     IF(ABS(Y) .LE. 32767.0) GOTO 310
        IERR="561
        Y=SIGN(32767.0,Y)
310     IDATA(IY+J)=Y
        GOTO NEXT
400     IF(ABS(E) .LE. 32767.0) GOTO 320
        IERR="561
        E=SIGN(32767.0,E)
320     IDATA(IE+J)=E
500     CONTINUE
        RETURN
C
C       Filtering complete. Return to main calling routine.
C
C       Error conditions
C
95      IERR="743
        RETURN
96      IERR="612
        RETURN
97      IERR="12
        RETURN
98      IERR="322
        RETURN
99      IERR="422
        RETURN
        END
```

# APPENDIX B.

## ESTIMATION OF COHERENCE AND SPECTRAL FUNCTIONS.

The magnitude squared coherence (MSC) function is defined to be

$$|\gamma(f)|^2 = \frac{|G_{xy}(f)|^2}{G_{xx}(f).G_{yy}(f)} \qquad (B.1)$$

where $G_{xx}(f)$ and $G_{yy}(f)$ are the autospectra and $G_{xy}(f)$ the cross-spectrum of two zero-mean, stationary stochastic processes $x(t)$ and $y(t)$. Current techniques for estimating the MSC generally use the FFT to obtain estimates of the spectral functions from a finite data record. These are conveniently obtained by segmenting a large record into N equal length (overlapping or disjoint) sections, each of which is sampled at T equally-spaced time intervals. The spectra are computed for each segment and averaged to obtain smoothed estimates, a procedure necessary to reduce the bias [135]. These are used to obtain the estimate of the MSC as follows:

$$|\hat{\gamma}(f)|^2 = \frac{|\hat{G}_{xy}(f)|^2}{\hat{G}_{xx}(f).\hat{G}_{yy}(f)} \qquad (B.2)$$

The advantage of this approach is that it may be efficiently performed by small general-purpose computers with limited store, as the DFT is not taken of the entire record, as would be required if smoothing of the spectral estimates was performed by averaging over neighbouring frequencies. The alternative approach to spectral estimation, that of computing the correlogram and transforming truncated versions of these give essentially the same results, but is computationally less efficient.

There are several ways of proceeding that depend upon the application. Two conflicting requirements are present in estimating the spectra of a given data set. Good frequency resolution is obtained when T is large, but this reduces the number of segments available for averaging, resulting in greater bias and variance. Lower bias can be achieved by using smaller values of T, but this is at the expense of frequency resolution.

This conflict can be alleviated to some extent by overlapping the segments, which permits more averages to be taken whilst maintaining the same frequency resolution. This is only effective up to a point, as the computational overheads increase rapidly while the improvement declines due to greater correlation between segments. Carter, Knapp and Nuttal [129] have shown theoretically and experimentally that for Hanning windowed segments there is little point in employing overlapping beyond about 50%.

The statistical properties of the MSC estimator have been described by Carter et al [129] who derived expressions for the bias and variance that result from the use of the segmented estimation procedure. Two approximations useful for N in the range N > 32 are reproduced here. These assume ideal windowing and no overlapping of zero-mean, stationary gaussian processes.

$$\text{Bias } B(|\gamma|^2) = (1 - |\gamma|^2)^2 / N \tag{B.3}$$

$$\text{Var } V(|\gamma|^2) = \begin{cases} 1/N2 & ; \ |\gamma|^2 = 0 \\ 2|\gamma|^2 \cdot (1 - |\gamma|^2)^2 / N; & 0 < |\gamma|^2 < 1 \end{cases} \tag{B.4}$$

These show clearly that the bias and variance of the estimation error are dependent on both the number of segments N, and the true MSC value. Computed values of these are listed in Table B.1 for a range of coherence values and segment sizes.

(a)

| $\|\gamma\|^2$ | 32 | 64 | 128 |
|---|---|---|---|
| 0.0 | 0.0313 | 0.0156 | 0.0078 |
| 0.1 | 0.0253 | 0.0127 | 0.0063 |
| 0.2 | 0.0200 | 0.0100 | 0.0050 |
| 0.3 | 0.0153 | 0.0077 | 0.0038 |
| 0.4 | 0.0112 | 0.0056 | 0.0028 |
| 0.5 | 0.0078 | 0.0039 | 0.0020 |
| 0.6 | 0.0050 | 0.0025 | 0.0012 |
| 0.7 | 0.0028 | 0.0014 | 0.0007 |
| 0.8 | 0.0012 | 0.0006 | 0.0003 |
| 0.9 | 0.0003 | 0.0002 | 0.0001 |
| 1.0 | 0.0000 | 0.0000 | 0.0000 |

(b)

| $\|\gamma\|^2$ | 32 | 64 | 128 |
|---|---|---|---|
| 0.0 | 0.0010 | 0.0002 | 0.0001 |
| 0.1 | 0.0051 | 0.0025 | 0.0013 |
| 0.2 | 0.0080 | 0.0040 | 0.0020 |
| 0.3 | 0.0092 | 0.0046 | 0.0023 |
| 0.4 | 0.0090 | 0.0045 | 0.0022 |
| 0.5 | 0.0078 | 0.0039 | 0.0020 |
| 0.6 | 0.0060 | 0.0030 | 0.0015 |
| 0.7 | 0.0039 | 0.0020 | 0.0010 |
| 0.8 | 0.0020 | 0.0010 | 0.0005 |
| 0.9 | 0.0006 | 0.0003 | 0.0001 |
| 1.0 | 0.0000 | 0.0000 | 0.0000 |

TABLE B.1. Computed values of (a) bias and (b) variance for the magnitude-squared coherence function for a range of coherence values and three different ensemble sizes N = 32, 64 and 128. Computations based on the approximate expressions given in the text.

Details of the computational procedure will now be related. N equally spaced samples of data records $\{x_j\}$ and $\{y_j\}$, j=0,...,N-1, may be partitioned into K segments of length L, with the starting point of each segment separated by D samples. The degree of overlap is then given by (1-D/L) X 100%, and the resulting sequences may be denoted by

$$\begin{aligned} \{\underline{x}_1(j)\} &= \{x(j)\} \qquad j = 0, \ldots, L-1 \\ \{\underline{x}_2(j)\} &= \{x(j+D)\} \\ &\vdots \\ \{\underline{x}_k(j)\} &= \{x(j+(K-1)D)\} \end{aligned} \tag{B.5}$$

The segment length L must be a power of two for most efficient computation of the DFT using the FFT, though each segment can be artificially extended to this length by appending zero-valued samples.

It is common practice to multiply each segment by a suitable window function, (not including any zero-padded extension), in order to reduce 'leakage' of spectral energy into adjacent frequencies. This phenomenon is well-known and described in texts on spectral analysis. A simple window such as the Hanning window given by

$$W(j) = 0.5 \left[ 1 - \cos (2\pi j/L-1) \right] , \; j = 0, \ldots, L-1 \tag{B.6}$$

reduces the effect of sidelobes in the spectrum that arise from the use of a finite time record. The spectrum has to be scaled to counteract the energy loss caused by windowing, according to Parseval's Theorem. This loss of signal energy may largely be avoided by overlapping the segments. The data segments $\{\underline{x}_k(j)\}$ are weighted by the window function W(j) and the DFT obtained as follows:

$$X_k(n) = 1/L \sum_{j=0}^{L-1} x_k(j).W(j).e^{-2\pi ijn/L} \tag{B.7}$$

where $i = \sqrt{(-1)}$. The DFT is obtained similarly for $\underline{y}_k(j)$ and denoted by $Y_k(n)$:

$$Y_k(n) = 1/L \sum_{j=0}^{L-1} y_k(j).W(j).e^{-2\pi ijn/L} \quad (B.8)$$

The spectral estimate is the average of these periodograms i.e.

$$\hat{G}_x(f_n) = 1/K \sum_{k=1}^{K} (L/U).|x_k(n)|^2 \quad (B.9)$$

where

$$f_n = n/L \; ; \; n = 0, \ldots, L/2 \quad (B.10)$$

is the frequency, and

$$U = 1/L \sum_{j=0}^{L-1} W(j)^2 \quad (B.11)$$

is the factor by which the mean signal power is reduced by the window. The cross-spectrum may be found using a similar procedure, i.e.

$$\hat{G}_{xy}(f_n) = L/KU \sum_{k=1}^{K} X_k(n)^* \cdot Y_k(n) \quad (B.12)$$

where * indicates the complex conjugate function. These are then substituted into equation (B.2) to obtain the estimate of the MSC.

There can be difficulties in estimating the cross-spectrum (and hence also the coherence) if the data records exhibit a rapid phase change with frequency, which corresponds to a time delay between the correlated data. This causes the MSC estimate to be significantly overestimated. When this is likely to occur it is recommended [136] that the data records be pre-aligned in time. Other procedures have been suggested to minimise the bias arising from the use of a small number of short data records. Thus Stearns [137] recommends averaging of the coherences obtained over of the data segments. This procedure was first proposed by Tick [138] who suggested several ways in which the averaging could be performed. Experimental results seem to indicate that this is effective when bias is likely to be significant, but

otherwise can give poorer results. For this reason and because these procedures are still under investigation, they are best avoided unless clearly necessary.

# APPENDIX C.

## Use of the Coherence function to predict ANC performance.

It is straightforward to obtain a measure of the SNR improvement that can be expected by the use of an adaptive noise canceller (ANC). In this appendix I shall show that the magnitude squared coherence (MSC) is useful in this context. Fig. C.1 shows the situation to be considered. Let c(t) be a continuous-time random signal which represents a correlated signal present in two channels $x_1(t)$ and $x_2(t)$. In each case c(t) has been passed through a linear filter, ($H_1(f)$ and $H_2(f)$ respectively). Random noise signals $u_1(t)$ and $u_2(t)$, uncorrelated with each other and with c(t), are also present in each signal channel. Each of the signals c(t), $u_1(t)$ and $u_2(t)$ are assumed to be stationary with zero mean, and to have power spectral densities C(f), $U_1(f)$ and $U_2(f)$ respectively.

Widrow et al [92] have considered the SNR improvement that would be expected if an ANC were used to cancel the correlated signal from one channel by using the other as a reference. This situation is depicted in fig. C.2 and is essentially the same as that described in Chapter 4, with the main signal omitted for the purpose of this analysis. This is entirely valid as it has previously been shown that a signal present only in the primary input appears unmodified at the ANC output, provided that it is uncorrelated with each of c(t), $u_1(t)$ and $u_2(t)$. The SNR can then be expressed in terms of the ratio of noise power present in the ANC output to that in the primary input. It has been shown [92] that the ratio of the noise powers is:

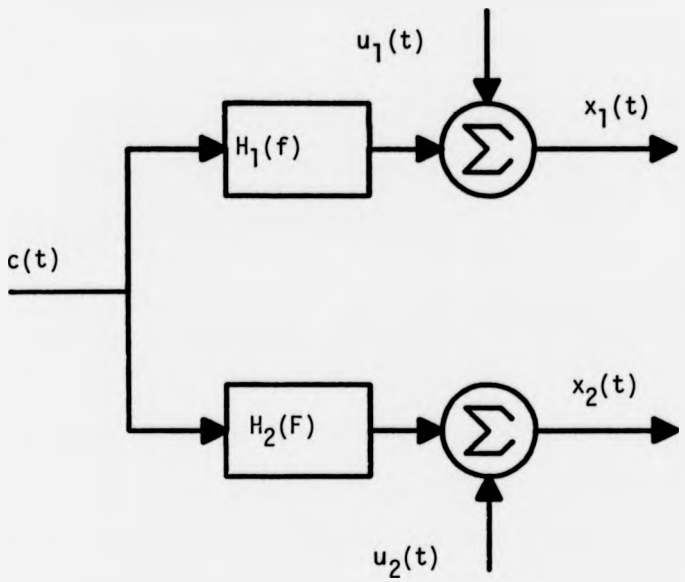



FIGURE C.1 Derivation of two partially correlated signal channels $x_1(t)$ and $x_2(t)$

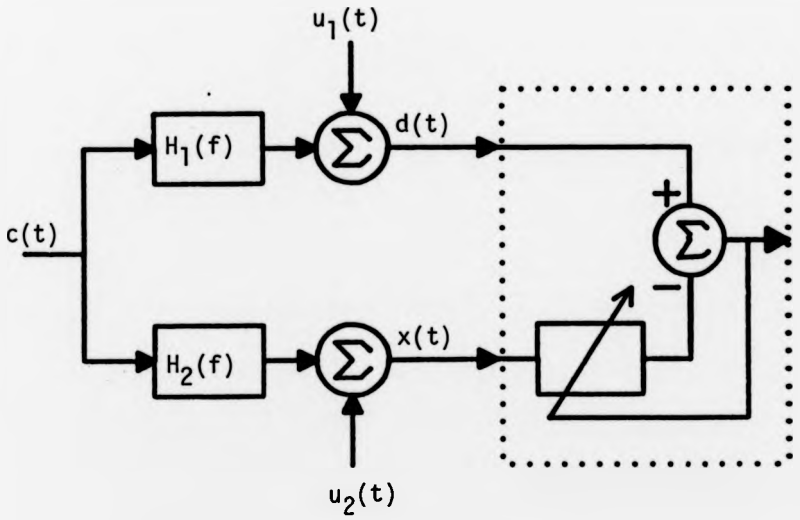



FIGURE C.2 Cancellation of correlated component in d(t) using reference signal x(t)

$$(C.1) \qquad P(z) = \frac{\rho_{out}(z)}{\rho_{pri}(z)} = \frac{[\ A(z) + 1\ ].[\ B(z) + 1\ ]}{A(z) + A(z).B(z) + B(z)}$$

where A(z) and B(z) are the ratio of the uncorrelated to the correlated power densities of each signal channel. Rewriting this expression in terms of frequency and rearranging gives

$$(C.2) \qquad P(f) = \frac{1}{1 - [\ (\ A(f) + 1\ ).(\ B(f) + 1\ )\ ]^{-1}}$$

Though useful as an analytical result, this expression is difficult to apply in practice since A(f) and B(f) are generally unknown and not easily estimated. This may be seen by considering these functions in terms of the signal spectra:

$$(C.3) \qquad A(f) = \frac{U_1(f)}{C(f).|H_1(f)|^2}$$

$$(C.4) \qquad B(f) = \frac{U_2(f)}{C(f).|H_2(f)|^2}$$

The spectral functions of each signal channel can be readily obtained, viz:

$$(C.5) \qquad X_1(f) = C(f).|H_1(f)|^2 + U_1(f)$$

$$(C.6) \qquad X_2(f) = C(f).|H_2(f)|^2 + U_2(f)$$

and

$$(C.7) \qquad X_{12}(f) = C(f).H_1(f).H_2(f)^*$$

It is not possible to estimate A(f) and B(f) from these, without knowing some of the underlying signal spectral functions C(f), $U_1(f)$ and $U_2(f)$ or the filter responses $H_1(f)$ and $H_2(f)$. It will now be shown that P(f) can be obtained by means of the MSC. From the definition, the coherence between the channels $x_1(t)$ and $x_2(t)$ can be written,

$$(C.8) \qquad |\gamma|^2(f) = \frac{C(f).|H_1(f).H_2(f)^*|^2}{[\ C(f).|H_1(f)|^2 + U_1(f)\ ].[\ C(f).|H_2(f)|^2 + U_2(f)\ ]}$$

Using the definitions of A(f) and B(f) this may be written

$$(C.9) \qquad |\gamma|^2(f) = \frac{C^2(f).|H_1(f).H_2(f)^*|^2}{C^2(f).|H_1(f)|^2.|H_2(f)|^2.(\ 1 + A(f)\ ).(\ 1 + B(f)\ )}$$

$$(C.10) \qquad = \frac{1}{(\ 1 + A(f)\ ).(\ 1 + B(f)\ )}$$

Substituting (C.10) into (C.2) yields the result

$$(C.11) \qquad \frac{\rho_{out}(f)}{\rho_{pri}(f)} = \frac{1}{1 - |\gamma(f)|^2}$$

provided that $|\gamma(f)|^2$ is not equal to unity at any frequency.

# APPENDIX D.

## A PRELIMINARY INVESTIGATION OF A MULTIPLE-REFERENCE ANC.

The results obtained in Chapter 8 using a single-reference adaptive noise canceller were sufficiently promising to support further investigation using multiple reference sources. This approach (described in Chapter 4, section 4.3.4) allows better noise cancellation to be obtained if each reference source contains activity correlated with the primary channel but uncorrelated with each other. As time was restricted for this study no attempt was made to analyse the mutual correlation properties of the reference channels. It was considered that a preliminary exploration of this approach would be sufficient to indicate whether further research would be worth pursuing. The investigation was therefore limited to two test cases, based on the recorded EPs obtained previously.

In case (i) 8 records from subject DAJ/I were filtered using a single reference GAF (using each of channels 5 and 6 as reference), and a multiple reference GAF (MGAF) (using both channels 5 and 6 as reference). The primary input was channel 3 as before, and N was set to 32 and $\mu'$ to 0.2. The results of applying each of these filters to two records are shown in fig. D.1 and compared with simple subtraction of the two inputs. The averages of 8 records are shown in fig. D.2, with the +1 standard deviation limit shown superimposed as a measure of the variability of the individual responses. In case (ii) four records from subject MJM/C were filtered using a single reference GAF (with channel 4 as reference) and a MGAF (with channels 1, 2, 3 and 4 as references). The primary input was channel 6. N was 32 for the GAF and 16 for the MGAF, and $\mu'$ was 0.2. The results of filtering a typical record are

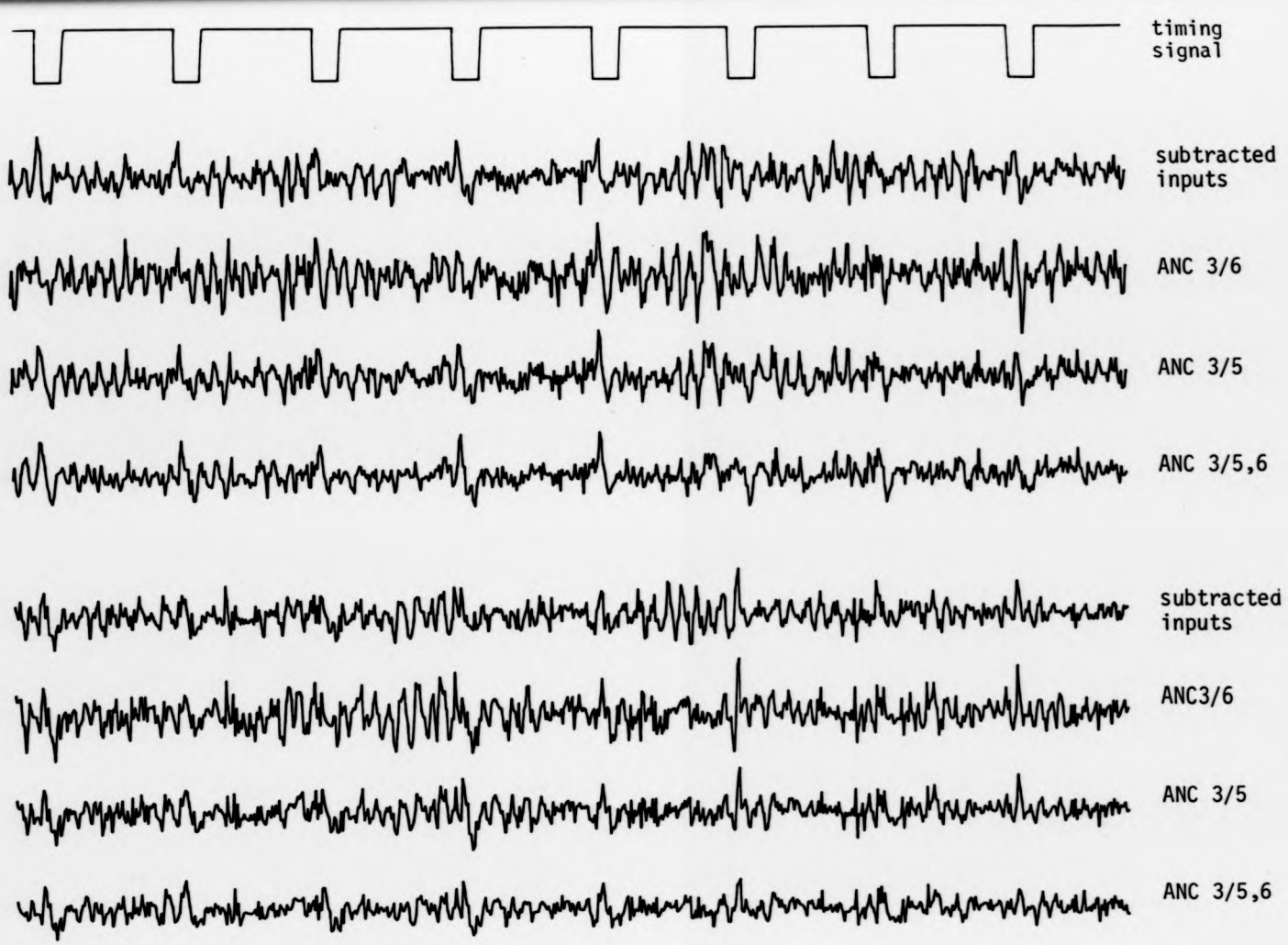



Figure D1

Comparison of two-reference ANC with one-reference ANC and with subtraction of inputs. Two records shown here. Subject DAJ Run I, primary input channel 3, reference inputs 6 and 7.

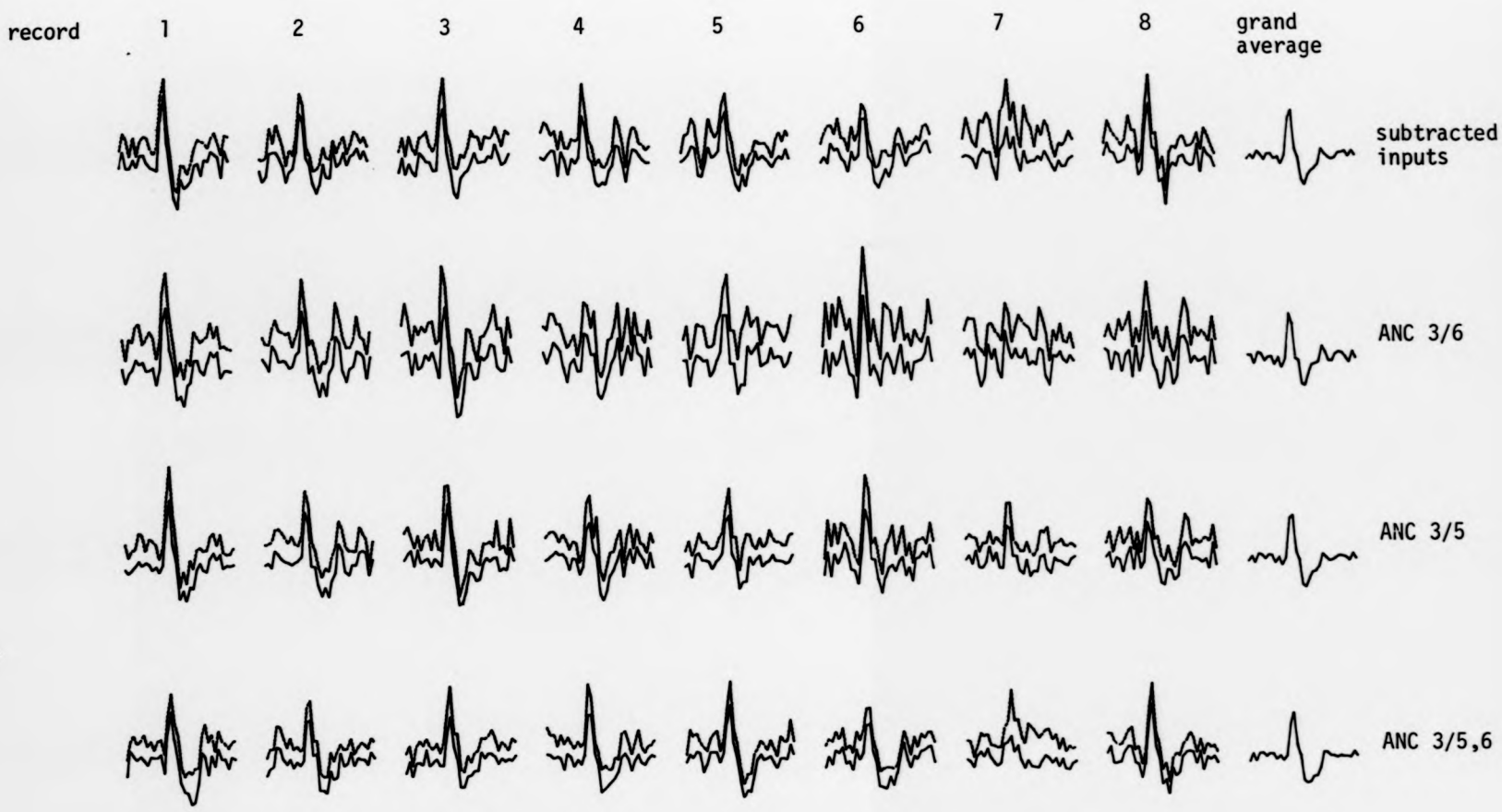



Figure D2

Comparison of two-reference ANC with one-reference ANC and with subtraction of inputs. Averages of 8 filtered records shown here, with grand average of 64 responses. Upper confidence limit (1 standard deviation) shown dotted to indicate variability of responses. Subject DAJ run I.

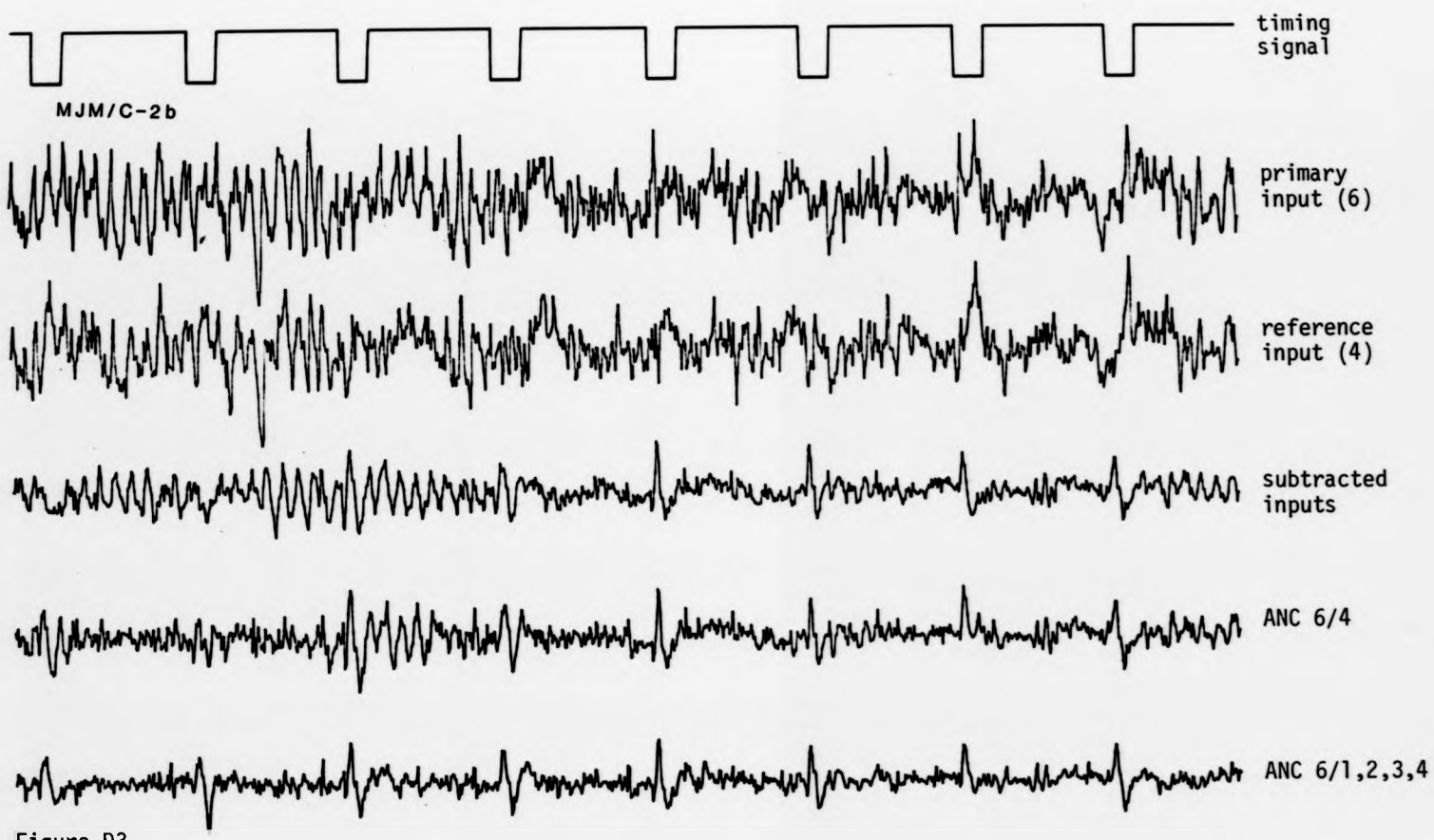



Figure D3

Comparison of multi-reference ANC, single-reference ANC and subtracted inputs. One record shown, subject MJM Run C. Primary input 6, reference inputs 1, 2, 3 and 4.

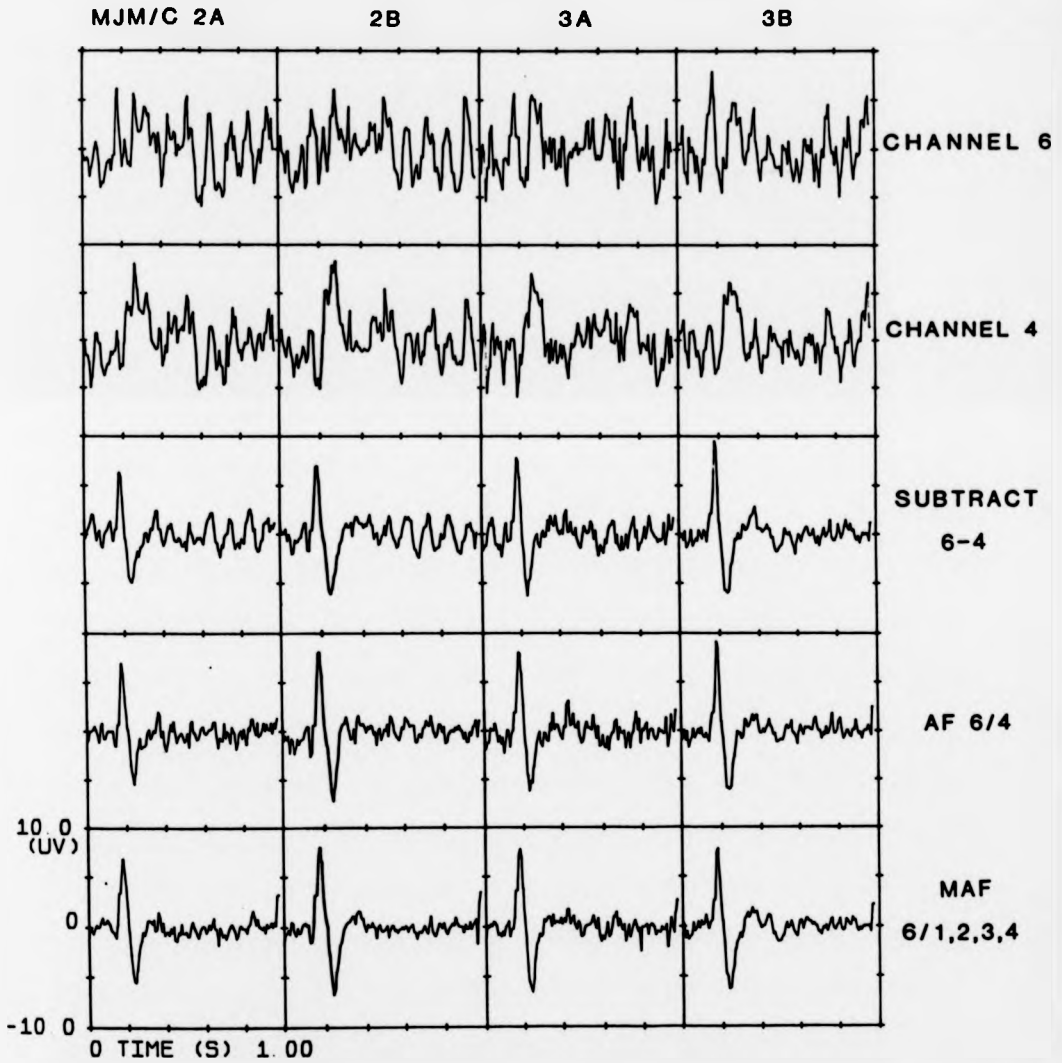



Figure D4

Comparison of multi-reference ANC, single-reference ANC and subtracted inputs. Averages of 4 records shown here. Note the reduction of general noise activity, and alpha activity in particular.

shown in fig. D.3, and the averages of 4 such records are shown in fig. D.4.

The results are encouraging and show that greater cancellation of correlated activity is indeed attainable through use of multiple reference channels. In the first case, channel 6 by itself does not provide as good cancellation of the primary correlated activity as channel 5, and might not be expected to provide additional cancellation when used with channel 5 in the MGAF. The results however reveal that this is not the case and in several records the MGAF performs better than either of the single reference ANCs. The averages in fig. D.2 show no evidence for distortion of the EP signals through use of the MGAF, though there does not appear to have been much improvement compared with the alternative approaches, other than some reduction in alpha activity. These results are confirmed in the second case though the use of four reference channels appears to give better cancellation. Alpha activity in particular is markedly reduced in two records (as shown by the averages in fig. D.4) though general EEG activity is also attenuated. This appears to have been achieved without any signal distortion occurring. The results of this preliminary study therefore confirm the suggestion made in this thesis that improved signal estimation is made possible by the use of adaptive noise cancelling, and it may be concluded that the use of multiple reference channels is an important extension of the single reference canceller described in this thesis. These proposals nevertheless need to be investigated thoroughly using a larger number of subjects and stimulus conditions, though it is expected that the use of a two-dimensional array of electrodes will lead to even better results than those obtained here using a single row of transverse electrodes, as reference electrodes can be chosen which have high correlations with the primary channel.

## REFERENCES.


1. Dawson G.D. Cerebral Responses to Electrical Stimulation of Peripheral Nerve in Man. J. Neurol. Neurosurg. Psychiat., **10** 134 (1947)

2. Regan D. Evoked Potentials in Psychology, Sensory Physiology & Clinical Medicine. Chapman & Hall, London (1972)

3. Barber C. (ed). Evoked Potentials. Proceedings of an International Evoked Potentials Symposium held in Nottingham, England. MTP Press Ltd. (1980)

4. Desmedt, J.E. (ed). Visual Evoked Potentials in Man: New Developments. Clarendon Press, Oxford. (1977)

5. Chiappa K.H. Evoked Potentials in Clinical Medicine. Raven Press, New York (1983)

6. Callaway E., Tueting P. and Koslow S.H. (ed). Event Related Brain Potentials in Man. Academic Press, London (1978)

7. Childers D.G. Evoked Responses: Electrogenesis, Models, Methodology and Wavefront Reconstruction and Tracking Analysis. Proc. IEEE, **65** 611-625 (1977)

8. Berger H. Ueber das Elektrenkephalogramm des Menschen. Arch. Psychiat. Nervenkr., **87** 527-570 (1929) English translation in : Gloor P. (ed). Hans Berger on the Electroencephalogram of Man. Electroenceph. Clin. Neurophysiol. Suppl., **28** 37-73 (1969)

9. Bickford, R.G., Jacobson J.L. and Cody T.R. Nature of Average Evoked Potentials to Sound and Other Stimuli in Man Ann. N.Y. Acad. Sci., **112** 204-223 (1964)

10. Cooper R., Osselton J.W. and Shaw J.C. EEG Technology. Butterworth, London (1974)

11. Cooper R., Winter A.L., Crow H.J. and Walter W.G. Comparison of Subcortical, Cortical and Scalp Activity using Chronically Indwelling Electrodes in Man. Electroenceph. Clin. Neurophysiol., **18** 217-228 (1965)

12. Abraham K. and Ajmone-Marson C. Patterns of Cortical Discharges and their Relation to Routine Scalp Electroencephalography. Electroenceph. Clin. Neurophysiol., **10** 447-461 (1958)

13. Jeffreys D.A. and Axford J.G. Source Locations of Pattern-Specific Components of Human Evoked Potentials. I. Component of Striate Cortical Origin. Exp. Brain Res., **16** 1-21(1972)

14. Hillyard S.A. and Picton T.W. Event-Related Brain Potentials and Selective Information Processing. In: J.E.Desmedt (ed). Progress in Clinical Neurophysiology VI. Basel, Karger, 1-50 (1978)

15. Regan D. Evoked Potential Studies of Visual Perception. Canad. J. Psychol./Rev. Canad. Psychol, **35** 77-111 (1981)

16. Carmon A., Friedman Y., Coger R. and Kenton B. Single Trial Analysis of Evoked Potentials to Noxious Thermal Stimulation in Man. Pain, **8** 21-32 (1980)

17. Vidal J.D. Real Time Detection of Brain Events in EEG. Proc. IEEE, **65** 633-641 (1977)

18. Wastell D.G. and Kleinman D. Potentiation of the Habituation of Human Brain Potentials Biological Psychology, **10** 21-29 (1980)

19. Harding G.F.A. The Visual Evoked Response. Adv. Ophthal., **28** 2 (1974)

20. Harden A., Picton-Robinson N., Bradshaw K. and Pampiglione G. Ten Years' Experience of ERG/VEP/EEG Studies on Visual Disorders in Paediatrics. In: Barber C. (ed.). Evoked Potentials. MTP Press Ltd, 257-266 (1980)

21. Fortgens C. and de Bruin M.P. Removal of Eye Movement and ECG Artifacts from the Non-Cephalic Reference EEG. Electroenceph. Clin. Neurophysiol., **56** 90-96 (1983)

22. Barlow J.S. and Remond A. Eye Movement Artifact Nulling in EEGs by Multichannel On-Line EOG Subtraction. Electroenceph. Clin. Neurophysiol., **52** 418-423 (1981)

23. O'Donnel R.D., Berkhout J. and Adey W.R. Contamination of Scalp EEG Spectrum during Contraction of Cranio-Facial Muscles. Electroenceph. Clin. Neurophysiol., **37** 145-151 (1974)

24. De Weerd J.P.C.M. and Kap J.I. Spectro-temporal Representations and Time-varying Spectra of Evoked Potentials. Biol. Cyb., **41** 101-117 (1981)

25. Childers D.G., Doyle T.C., Brinck A.G. and Perry Jnr. N.W. Ensemble Characteristics of the Human Visual Evoked Response: Periodic and Random Stimulation. IEEE Trans., **BME-19** 408-415 (1972)

26. Aunon J.I. and McGillem C.D. Detection and Processing of Individual Components in the VEP Psychophysiology, **16** 71 (1979).

27. Kooi K.A. Fundamentals of Electroencephalography. Harper & Row, New York (1971)

28. Tyner F.S. Knott J.R. and Mayer Jnr, W.B. Fundamentals of EEG Technology. Vol 1: Basic Concepts and Methods. Raven Press, New York (1983)

29. Storm van Leeuwen W. α colleagues Proposal for an EEG Terminology. Electroenceph. Clin. Neurophysiol., **20** 306 (1966)

30. Jasper H.H. The Ten-Twenty System of the International Federation Electroenceph. Clin. Neurophysiol., **10** 371-375 (1958).

31. Walter D.O., Rhodes J.M., Brown D. and Adey W.R. Comprehensive Spectral Analysis of Human EEG Generators in Posterior Cerebral Regions. Electroenceph. Clin. Neurophysiol., **20** 224-237 (1966).

32. Dumermuth G. and Fluhler H. Some Modern Aspects in Numerical Spectrum Analysis of Multichannel Electroencephalographic Data.

Med. and Biol. Eng., 5 319-331·(1967).

33. Dumermuth G. Fundamentals of Spectral Analysis in Electroencephalography. In: Remond A. (ed). EEG Informatics. A Didactic Review of Methods and Applications of EEG Data Processing. Elsevier/North Holland Biomedical Press, Amsterdam (1977).

34. McEwen J.A. and Anderson G.B. Modeling the Stationarity and Gaussianity of Spontaneous Electroencephalographic Activity. IEEE Trans., BME-22 361-369 (1975).

35. Kawabata N. Test of Statistical Stability of the Electroencephalogram. Biol. Cyb., 22 235-238 (1976).

36. Cohen B.A. Stationarity of the Human Electroencephalogram. Med. and Biol. Eng. and Comp., 15 513-518 (1977).

37. Bode H.W. and Shannon C.E. A Simplified Derivation of Linear Least Squares Smoothing and Prediction Theory. Proc. IRE., 38 417-425 (1950).

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

39. Bendat J.S. Mathematical Analysis of Average Response Values for Non-stationary Data. IEEE Trans., BME-11 72-81 (1964).

40. Cotman C.W. and McGaugh J.L. Behavioral Neuroscience. An Introduction. Academic Press, London, (1980).

41. Woody C.D. Characterization of an Adaptive Filter for the Analysis of Variable Latency Neuroelectric Signals. Med. and Biol. Eng., 5 539-553 (1967).

42. Gasser Th., Mocks J. and Verleger R. SELAVCO: A Method to Deal with Trial-to-Trial Variability of Evoked Potentials. Electroenceph. Clin. Neurophysiol., 55 717-723 (1983).

43. Pfurtscheller G. and Cooper R. Selective Averaging of the Intracerebral Click Evoked Responses in Man: An Improved Method of Measuring Latencies and Amplitudes. EEG and Clin. Neurophysiol., 38 187-190 (1975).

44. Sayers B.McA., Beagley H.A. and Riha J. Pattern Analysis of Auditory Evoked EEG Potentials. Audiology, 18 1-16 (1979)

45. De Weerd J.P.C.M. and Martens W.L.J. Theory and Practice of a Posteriori "Wiener" Filtering of Average Evoked Potentials. Biol. Cyb., 30 81-94 (1978)

46. Wastell D.G. and Kleinman D. N1-P2 Correlates of Reaction Time at the Single Trial Level. Electroenceph. Clin. Neurophysiol., 48 191-196 (1980)

47. Vaughan Jr. H.G. The Relationship of Brain Activity to Scalp Recordings of ERPs In: Donchin E. and Lindsley D.B. (eds.). Average Evoked Potentials. Methods, Results and Evaluations. NASA SP-191 (1969)

48. Weiss P.L., Hunter I.W. and Kearney R.E. Reduction of Physiological Signal Contamination Using Linear Filter Identification. Med. and Biol. Eng. and Comp., 21 521-524 (1983)

49. MacKay D.M. and Jeffreys D.A. Visual Evoked Potentials in Man and Visual Perception. In: Jung R. (ed). Handbook of Sensory Physiology. Springer, Heidelberg and New York, 7/3B 647-678 (1973)

50. McCool J.M. and Widrow B. Principles and Applications of Adaptive Filters: A Tutorial Review. IEE Conference Publication 144 (1976)

51. Bourne J.R. (ed.) CRC Critical Reviews in Bioengineering. 5 Issue 4 (1981)

52. Bourne J.R. (ed.) CRC Critical Reviews in Bioengineering. 6 Issue 3 (1981)

53. Brazier M.A.B. Evoked Responses Recorded from the Depths of the Human Brain. Ann. NY Acad. Sci., 112 33-59 (1964)

54. Wastell D.G. Statistical Detection of Individual Evoked Responses: An Evaluation of Woody's Adaptive Filter. Electroenceph. Clin. Neurophysiol., 42 835-839 (1977)

55. Aunon J.I. and Sencaj R.W. Comparison of Different Techniques for Processing EPs. Med. and Biol. Eng. and Comp., 16 642-650 (1978)

56. McGillem C.D. and Aunon J.I. Measurements of Signal Components in Single Visually Evoked Brain Potentials. IEEE Trans., BME-24 232-241 (1977)

57. Sjontoft E. A Deconvolution Procedure for use in Extracting Information in Average Evoked Response EEG Signals. IEEE Trans., BME-27 227-230 (1980)

58. McGillem C.D., Pomalaza C.A. and Aunon J.I. Method of Computing Enhanced Average Evoked Brain Potentials. IEEE Frontiers of Engineering in Health Care, 2 134-136 (1980)

59. Ciganek L. Variability of the Human Visual Evoked Potential: Normative Data. Electroenceph. Clin. Neurophysiol., 27 35-42 (1969)

60. Twyman R., Lastimosa A.C.B. and Lastimosa A.C.B. Constrained Deconvolution: Enhanced Resolution of Evoked Potentials. IEEE Frontiers of Engineering in Health Care, 3 297-301 (1981)

61. Steeger G.H., Hermann O. and Spreng M. Some Improvements in the Measurement of Variable Latency Acoustically Evoked Potentials in Human EEG. IEEE Trans., BME-30 295-303 (1983)

62. Auerbach V.H. and Haber F. Two Novel Ways of Averaging waveforms by Fourier Analysis. The Representative and the Synchronous Averages. J. Franklin Inst., 297 169-178 (1974)

63. Rodriguez M.A., Williams R.H. and Carlow T.J. Signal Delay and Waveform Estimation using Unwrapped Phase Averaging. IEEE Trans., ASSP-29 508-513 (1981)

64. Ruchkin D.S. Analysis of Nonhomogeneous Sequences of Evoked Potentials. Exp. Neur., **20** 275-284 (1968)

65. John E.R., Ruchkin D.S. and Vidal J.J. Measurement of Event-Related Potentials In: Callaway E., Tueting P. and Koslow S.H. (eds). Event-Related Brain Potentials in Man. Academic Press, London, 93-138 (1978)

66. Cummins K.L. Evoked Potential Processing: Applications of Linear Estimation and Detection Theory. IEEE Frontiers of Engineering in Health Care, 3 187-191 (1981)

67. Ruchkin D.S. An Analysis of Average Response Computations Based upon Aperiodic Stimuli. IEEE Trans., **BME-12** 87-94 (1965)

68. Picton T.W. and Hillyard S.A. Human Auditory Evoked Potentials. II. Effects of Attention. Electroenceph. Clin. Neurophysiol., **36** 191 (1974)

69. Remond A. and Lesevre N. Variations in AVEP as a Function of the Alpha Rhythm. Electroenceph. Clin. Neurophysiol. Suppl., **26** 42-52 (1967)

70. Pfurtscheller G., Aranibar A. and Moresch H. Amplitude of EPs and Degree of Event-Related Desunchronisation (ERD) During Photic Stimulation. Electroenceph. Clin. Neurophysiol., **47** 21-30 (1979)

71. Salomon G. and Barfod J. A New Concept of Vertex ERA and EEG Analysis Applyimg Inverse Filtering. Acta Otolaryngol., **83** 200-210 (1977)

72. Jervis B.W., Nichols M.J., Johnston T.E., Allen E. and Hudson N.R. A Fundamental Investigation of the Composition of Auditory Evoked Potentials. IEEE Trans., **BME-30** 43-50 (1983)

73. Basar E., Gonder A. and Ungan P. Comparative Frequency Analysis of Single EEG-Evoked Potential Records. J. Biomed. Eng., 2 9-14 (1980)

74. De Weerd J.P.C. Facts and Fancies about a Posteriori "Wiener" Filtering. IEEE Trans., **BME-28** 252-257 (1981)

75. Walter D.O. A Posteriori "Wiener Filtering" of Average Evoked Potentials. Electroenceph. Clin. Neurophysiol. suppl., **27** 61-70 (1969)

76. Doyle D.J. Some Comments on the use of Wiener Filtering for the Estimation of Evoked Potentials. Electroenceph. Clin. Neurophysiol., **38** 533-534 (1975)

77. Aunon J.I. and McGillem C.D. High Frequency Components in the Spectrum of the Visual Evoked Response. J. Bioengineering, 1 157-164 (1977)

78. Strackee J. and Cerri S.A. Some statistical Aspects of Digital Wiener Filtering and Detection of Prescribed Frequency Components in Time Averaging of Biological Signals. Biol. Cyb., **28** 55-61 (1977)

79. Carlton E.H. and Katz S. Is Wiener Filtering and Effective Method of Improving Evoked Potential Estimation" IEEE Trans., **BME-27** 187-192 (1980)

80. Woestenburg J.C., Verbaten M.N., Sjouw W.P.B. and Slangen J.L. A Statistical "Wiener" Filter. Zeist Proceedings.

81. Wastell D.G. When Wiener Filtering is Less than Optimal: An Illustrative Application to the Brain Stem Evoked Potential. Electroenceph. Clin. Neurophysiol., **51** 678-682 (1981)

82. Hartwell J.W. and Erwin C.W. Evoked Potential Analysis: On-Line Signal Optimalization Using Mini-Computer. Electroenceph. Clin. Neurophysiol., **41** 416-421 (1976)

83. De Weerd J.P.C.M. A Posteriori Time-Varying Filtering of Averaged Evoked Potentials. I. Introduction and Conceptual Basis. Biol. Cyb., **41** 211-222 (1981)

84. Hjorth B. An Online Transformation of EEG Scalp Potentials into Orthogonal Source Derivations. Electroenceph. Clin. Neurophysiol., **39** 526-530 (1975)

85. MacKay D.M. Source Density Mapping of Human Visual Receptive Fields Using Scalp Electrodes. Exp. Brain Res., **54** 579-581 (1984)

86. Horstfehr H.O. The Spatial Sum of Evoked Responses to a Single Stimulus (light or sound) Recorded from the Human Scalp. Elecroenceph. and Clin. Neurophysiol. Suppl., **26** 139-146 (1967).

87. Widrow B. Adaptive Filters I: Fundamentals. Stanford Electronics Lab., Stanford University, Rep. SU-SEL-66-126 (1966)

88. Young P. Recursive Approaches to Time Series Analysis. I.M.A. **10** 209-224 (1974)

89. Widrow B., Mantey P.E., Griffiths L.J. and Goode B.B. Adaptive Antenna Systems. Proc. IEEE, **55** 2143-2159 (1967)

90. Walach E. and Widrow B. The Least Mean Fourth (LMF) Adaptive Algorithm and its Family IEEE Trans., **IT-30** part 1 275-283 (1984)

91. Widrow B., McCool J., Larimore M.G. and Johnson Jnr, C.R. Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter. Proc. IEEE, **64** 1151-1162 (1976)

92. Widrow B., Glover J.R., McCool J.M., Kaunitz J., Williams C.S, Hearn R.H., Zeidler J.R., Dong E. and Goodlin R.C. Adaptive Noise Cancelling: Principles and Applications Proc. IEEE, **63** 1692-1716 (1975)

93. Reed I.S., Mallett J.D. and Brennan L.E. Rapid Convergence Rate in Adaptive Arrays. IEEE Trans., **AES-10** 853-863 (1974)

94. Bershad N.J. and Qu L.Z. LMS Adaptation with Correlated Data: A Scalar Example. IEEE Trans., **ASSP-32** 695-700 (1984)

95. Ekstrom M.P. A Spectral Characterisation of the Ill-Conditioning in Numerical Deconvolution. IEEE Trans., **AU-21** 344-348 (1973)

96. Ferrara E.R. The Time-Sequenced Adaptive Filter. PhD Dissertation, Stanford University, Stanford, CA (1977)

97. Caraiscos C. and Liu B. A Roundoff Error Analysis of the LMS Adaptive Algorithm. IEEE Trans., **ASSP-32** 34-41 (1984)

98. Clark G.A., Mitra S.K. and Parker S.R. Block Implementation of Adaptive Digital Filters. IEEE Trans., **ASSP-29** 744-752 (1981)

99. Clark G.A., Parker S.R. and Mitra S.K A Unified Approach to Time and Frequency Domain Realization of FIR Adaptive Digital Filters. IEEE Trans., **ASSP-31** 1073-1083 (1983)

100. Narayan S.S., Peterson A.M. and Narasimha M.J. Transformation Domain LMS Algorithm. IEEE Trans., **ASSP-31** 609-622 (1983)

101. Ogue J.C., Saito T. and Hoshiko Y. A Fast Convergence Frequency Domain Adaptive Filter. IEEE Trans., **ASSP-31** 1312-1314 (1983).

102. Ferrara E.R. and Widrow B. The Time-Sequenced Adaptive Filter. IEEE Trans., **ASSP-29** 679-683 (1981)

103. Ferrara E.R. and Widrow B. Multichannel Adaptive Filtering for Signal Enhancement. IEEE Trans., **ASSP-29** 766-770 (1981)

104. Cioffi J.M. and Kailath T. Fast Recursive-Least-Squares Transversal Filters for Adaptive Filtering. IEEE Trans., **ASSP-32** 304-337 (1984)

105. Friedlander B. Lattice Filters for Adaptive Processing. Proc. IEEE, **70** 829-867 (1982)

106. Widrow B. and Walach E. On the Statistical Efficiency of the LMS Algorithm with Nonstationary Inputs. IEEE Trans., **IT-30**, part 1 211-221 (1984)

107. Anderson B.D.O. and Johnstone R.M. Adaptive Systems and Time-Varying Plants. Int. J. Control, **37** 367-377 (1983)

108. Ferrara E.R. and Widrow B. Fetal Electrocardiogram Enhancement by Time-Sequenced Adaptive Filtering. IEEE Trans., **BME-29**, 458-460 (1982)

109. Kentie M.A., van der Schee E.J., Grashuis J.L. and Smout A.J.P.M. Adaptive Filtering of Canine Electrogastrographic Signals. Part I: System Design. Med. and Biol. Eng. and Comp., **19** 759-764 (1981) Adaptive Filtering of Canine Electrogastrographic Signals. Part II: Filter Performance. Med. and Biol. Eng. and Comp., **19** 765-769 (1981)

110. Friedman Y. and Carmon A. Analysis of Single Cerebral Evoked Responses by Adaptive Filtering. Proc. San Diego Biomed. Symp., **17** 233-238 (1977)

111. Widrow B., McCool J. and Ball M. The Complex LMS Algorithm. Proc. IEEE, **63** 719-720 (1975)

112. Cowan C.F.N. and Grant P.M. (eds) Adaptive Filters. Prentice-Hall, Englewood Cliffs, New Jersey (1985)

113. De Weerd J.P.C.M. and Kap J.I. A Posteriori Time-Varying Filtering of Averaged Evoked Potentials. II. Mathematical and Computational Aspects. Biol. Cyb., **41** 223-234 (1981)

114. Scientific Subroutines Reference Manual. Digital Equipment Corporation, Massachussetts, (1978)

115. Sokol S. Visually Evoked Potentials: Theory, Techniques and Clinical Applications. Survey of Ophthalmology, **21** 18-44 (1976)

116. Nielsen G. Digital Processing of Analogue Signals. Part 2. Electronic Engineering, 53-64 (Dec. 1978)

117. LAB-11 - Laboratory Applications-11. Library Software Product Description 15.35.3. Digital Equipment Corporation, Maynard, Massachussetts.

118. Kearney R.E. and Hunter I.W. APS - A Package for the Analysis of Physiological Systems. Proc. 14th Canadian DECUS Symposium, **17** 885-890 (1981)

119. Tomkins W.J. CONVLINC - A New Laboratory Computer Language. Proc. DECUS 2, 557-560 (1975)

120. Piner S.D. and East L.V. CLASS - A High Level Interactive Laboratory Software System. Proc. DECUS 2, 181-186 (1975)

121. Hunter I.W. and Kearney R.E. NEXUS - A Command Line Decoder and Task Despatcher for Systems and Signal Analysis. Proc. 15th Canadian DECUS Symposium, **8** 829-834 (1982)

122. Thomas C.W. and Welch A.J. A Computer System for Interactive Processing and Evaluation of Light Stimulated Evoked Responses. Comp. Biomed. Res., **6** 530-543 (1973)

123. Billings S.A., Sterling M.J.H. and Batey D.J. SPAID - An Interactive Data Analysis Package and its Application to the Identification of an Electric Arc Furnace Control System. Random Signal Analysis, IEE Conf. Publ. 159, 161-175 (1977)

124. Digital Signal Processing Committee (eds) Programs for Digital Signal Processing. IEEE Press, New York (1979)

125. PDP-11 FORTRAN Language Reference Manual. Digital Equipment Corporation, Maynard, Massachussetts (1977)

126. Kreitzberg C.B. and Schneiderman B. The Elements of Fortran Style: Techniques for Effective Programming. Harcourt Brace Jovanovich, New York (1972)

127. Day A.C. Fortran Techniques with Special Reference to Non-Numerical Applications. Cambridge University Press, (1972)

128. Welch P.D. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodiograms. IEEE Trans., **AU-15**, 70-73 (1967)

129. Carter G.C., Knapp C.H. and Nuttall A.H. Estimation of the Magnitude-Squared Coherence Function via Overlapped Fast Fourier Transform Processing. IEEE Trans., **AU-21** 337-344 (1973)

130. Basar E. EEG-Brain Dynamics. Relation Between EEG and Brain Evoked Potentials. Elsevier/North Holland, Amsterdam (1980)

131. Pocock P.V. The Spatial and Temporal Distributions of Alpha Activity and their Modification During Motor Preparation. In: Pfurtscheller G., Buser P., Lopes Da Silva F.H. and Petsche H.(eds). Rhythmic EEG Activities and Cortical Functioning, Elsevier/North Holland, Amsterdam, 135-149, (1980)

132. Bendat J.S. and Piersol A.G. Random Data: Analysis and Measurement Procedures. Wiley Interscience, New York, (1971)

133. Lopes Da Silva F.H. and Storm Van Leeuwen W. The Cortical Source of the Alpha Rhythm. Neurosci. Lett., **6** 237-241 (1977)

134. Lynn P.A. Online Digital Filters for Biological Signals: Some Fast Designs for a Small Computer. Med. and Biol. Eng. and Comp., **15** 534-540 (1977)

135. Carter G.C. and Knapp C.H. Coherence and its Estimation via the Partitioned Modified Chirp-Z Transform. IEEE Trans., **ASSP-23** 257-264 (1975)

136. Carter G.C. Bias in Magnitude-Squared Coherence Estimation due to Misalignment. IEEE Trans., **ASSP-28** 97-99 (1980)

137. Stearns S.D. Test of Coherence Unbiasing Methods. IEEE Trans., **ASSP-29** 321-323 (1981)

138. Tick L.J. Estimation of Coherency. In: Harris B. Spectral Analysis of Time Series, Wiley, (1967)