THE MODIFIED WIGNER DISTRIBUTION AND TWO

DIMENSIONAL ALLAN VARIANCE WITH

APPLICATIONS TO ACOUSTIC WELL

LOGS AND SPEECH

By

DWIGHT DAVID DAY

Bachelor of Science in Electrical Engineering
Oklahoma State University
Stillwater Oklahoma
1980

Master of Science
Oklahoma State University
Stillwater Oklahoma
1981

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfilment of the requirements
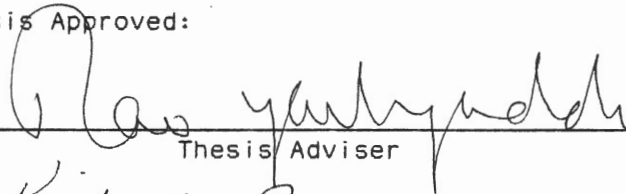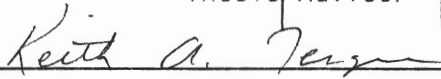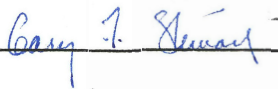for the Degree of
DOCTOR OF PHILOSOPHY
JULY 1987

THE MODIFIED WIGNER DISTRIBUTION AND TWO

DIMENSIONAL ALLAN VARIANCE WITH

APPLICATIONS TO ACOUSTIC WELL

LOGS AND SPEECH

Thesis Approved:

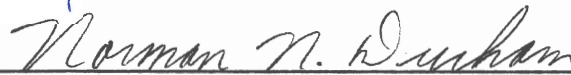Thesis Adviser

Keith a. Tergu

David F Soldan

Casy F. Steward

Norman N. Durham

Dean of the Graduate College

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.0 Motivation

Detection and identification of acoustic signals is of interest in a large range of applications. In this work we are interested in two types of acoustic signals, full wave acoustic well logs and speech. These signals, although both acoustic signals, have different properties. The acoustic well log has a narrow band spectrum, where as the speech signal has a broader spectrum. One fact that is common to both signals is that they are nonstationary. Therefore, these signals must be studied using short-time or nonstationary signal analysis. The following discussion describes the problems to be addressed for each application and presents various methods to be used in each case.

1.1 Acoustic Well Logging

To produce a full-wave acoustic log, a tool (called a sonde) suspended by a steel cable containing several conductors, is lowered into the borehole. As the sonde is drawn back up the borehole the transmitters emit bursts of acoustic waves at regular intervals of depth. These waves travel out through the borehole fluid to the formation as compressional waves. When these compressional waves strike the formation, they cause several different types of waves to occur in the formation. These different waves travel down through the formation. The waves

1

in the formation produce compressional waves in the borehole fluid, which carry the pulses back to the receivers on the sonde. The acoustic pulses are then converted to electrical signals at the receivers. The electrical signals are transmitted to the surface and recorded. Identification of the different wave types from these recorded signals is one of the major objectives of this research.

One of the uses of the acoustic log is to measure the slowness of the formation compressional wave. This is accomplished by assuming that the fastest moving wave is the formation compressional wave, and that the first signal to appear in the received waveform is the result of the formation compressional wave. The arrival of the compressional wave can be estimated as the time when the received signal exceeds some threshold value. In this way, the time required for the compressional wave to travel from the transmitter to the receiver can be measured. Once the travel time is determined, some compensation must be made for the time required for the waves to travel from the transmitter to the formation and then from the formation to the receiver. This compensated travel time is now divided by a compensated sonde length to give compressional slowness (the inverse of velocity). The compensated sonde length is slightly shorter than the actual transmitter receiver spacing; this is due to the refraction angle of the compressional wave as it reenters the formation.

The compensation factors are hard to determine and can cause considerable error. To avoid these compensation errors, two receiver tools were developed (See Figure 1). By using two receivers, the compensation for borehole travel is unnecessary. Since the trace from each receiver experiences approximately the same delay due to borehole travel, the

slowness of the compressional wave through the formation is simply the difference of arrival times at the receivers divided by the receiver spacing (See Figure 1). Even this tool is prone to error, for if this tool were to become tilted in the borehole, errors can develop. To correct for this error, a two transmitter, two receiver tool can be used. All this development of acoustic sondes is done to assure a good measurement of compressional slowness or velocity in the formation. But, what information can be gained from knowing the compressional velocity?

Cable to Recording Truck

Transmitter

Ray Path of Head Wave

Receiver 1

Receiver 2
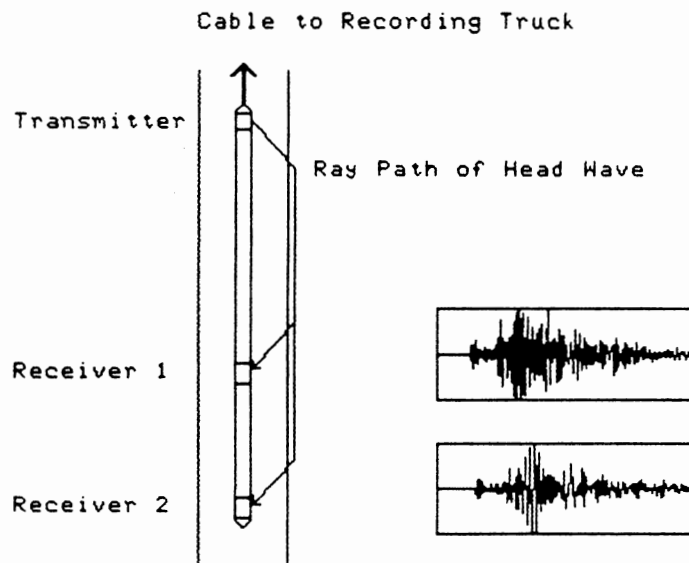
Figure 1. Schematic of Sonde and Borehole

A relationship between the compressional velocity and the formation porosity, proposed by M. R. J. Wyllie [94] and called Wyllie's "time

average formula" is given by

$$\frac{1}{V} = \frac{\Psi}{V_f} + \frac{(1-\Psi)}{V_{ma}}$$ (1.1)

where:

Ψ is the fractional porosity of rock,

V is the compressional formation velocity,

V_f is the velocity of the pore space fluid, and

V_{ma} is the velocity of the rock matrix.

This equation is now rewritten in terms of slowness ($\Delta t = 1/V$) and then solved for porosity Ψ to give

$$\Psi = \frac{\Delta t - \Delta t_{ma}}{\Delta t_f - \Delta t_{ma}}$$ (1.2)

where $\Delta t = 1/V$, $\Delta t_f = 1/V_f$, and $\Delta t_{ma} = 1/V_{ma}$. It can be seen from this equation that an "a priori" knowledge or at least an estimate of the formation make up is required before Equations (1.1) or (1.2) can be applied. The value Δt_f does not vary greatly for most borehole fluids and is assumed to be approximately 189 μsec per foot [84]. This assumption is not valid for gases. The value of Δt_{ma} can change drastically depending on the lithology of the formation about the borehole. Given some knowledge of the area and formation in question, an estimate of matrix velocity can usually be made. These estimates usually range from 167 μseconds per foot for some shales and 43.5 μseconds per foot for dolomite [84]. Recent studies show that a more accurate form of (1.1) is

$$V = (1-\Psi)^m V_{ma} + \Psi V_f$$ (1.3)

This equation, known as the "Raymer-Hunt-Gardner" equation [49], uses a value of m=2 for sandstones and 2.0 to 2.2 for carbonates. This equation agrees well with the Wyllie's formula for porosity in the range of 0.25

to 0.30.

There is more information in the wavetrain than simply the compressional velocities. With the introduction of digital waveform recorders, it is now possible to more fully analyze the wavetrain from the acoustic log and recover more of the information encoded there in.

One wave, other than the compressional, that is of interest is the shear wave. This wave is not commonly used, since it is slower than the compressional wave and its arrival is usually covered by the trailing end of the compressional wavelet. This means that more complex methods are required to identify the shear wave, especially its arrival time in the received signal.

In 1963, G. R. Pickett showed, from laboratory measurements made on core samples, some of the properties of the shear wave. First, Pickett showed that the shear wave velocity is more sensitive to porosity changes than the compressional wave. Also, the shear wave velocity obeys similar laws in its relationship to porosity as does the compressional wave. This means that, if an accurate measure of shear wave velocity could be found, the calculation of porosity using the shear wave would be less susceptible to errors. Also shown in Pickett's paper is a relationship of lithology to the ratio of the compressional to the shear velocities. This ratio (DTR) is more commonly written as the slowness of the shear wave divided by the compressional slowness, and falls into three major groups shown in Table I. Table I contains only approximations to the Pickett data [49].

A reduction in acoustic amplitudes is usually a good indicator of formation fracturing [49]. Pickett's analysis of shear wave amplitudes show them to be more sensitive to fracturing than the compressional wave, making the shear wave amplitudes very useful. This obviously is

dependent upon a consistent procedure for identifying the shear wave from the wavetrain.

TABLE I

LITHOLOGY TO DTR RELATIONSHIP

Lithology	DTR
Sandstone	1.58-1.78
Dolomite	1.8
Limestone	1.9

The remainder of the wavetrain is primarily made up of guided fluid waves. These waves include Stonely, and pseudo-Rayleigh waves. The exact usefulness of these waves is not readily apparent; however, there is speculation about using the Stonely wave for detection of fractures. The basic concept is that the energy in the Stonely wave is dissipated into a fracture or permeable formation more than into a solid formation, meaning a large drop in Stonely wave amplitudes.

Before discussing techniques to identify shear wave arrivals from the acoustic wavetrain, we first discuss some of the properties of compressional and shear waves, commonly called head waves.

The actual path of travel for the head waves is rarely as clean as that shown in Figure 1. As the waves travel down the formation, part of

the acoustic energy is radiated back into the borehole. The energy radiated back into the borehole is then reflected by the tool and reenters the formation to proceed down with the rest of the waves. These multiple reflections, along with the natural resonance of the fluid and borehole, make the head waves reverberant, which causes the spectrum of the compressional and shear wave to have sharp peaks at certain frequencies. In fact it is shown both analytically and experimentally that there is a frequency separation between the shear and compressional waves [58,73]. It is this frequency separation that is to be used in this research to identify the shear wave out of the wavetrain.

The identification of the shear wavelet is the problem to be addressed. Now, the compressional and shear waves may overlap, covering up the beginning of the shear wave. Our approach in solving this problem is to use a non-stationary time-frequency analysis known as the Wigner Distribution to analyze the wavetrains. Detection and identification of the wave arrivals is then made from this analysis.

1.2 Speech

As with the acoustic well log, an important operation in speech processing is the identification of different types of waves from a speech waveform. In the case of speech, the different waves to be identified here are voiced speech, unvoiced speech and silence. The discrimination between these three waves is an important step in many speech analysis procedures, such as speaker recognition, speech recognition and data compression. The variable nature of the speech waveform, from speaker to speaker and from word to word, complicates any such classification. Research into techniques for discriminating between these

different classes can be found in [4,29,66,68,81,92]. All these techni-
ques are based on the short-time properties of the speech waveform, such
as the short-time energy and zero crossing rate.

An example of a digitized speech waveform is shown in Figure 2. The
word spoken is 'TO'. The /t/ sound at the beginning of the utterance, is
a classic example of unvoiced speech. Unvoiced speech is produced when
the speaker forms a constriction at some point in the vocal tract and
air is then forced through this constriction at a high enough velocity
to produce turbulent flow. This turbulant flow causes the unvoiced
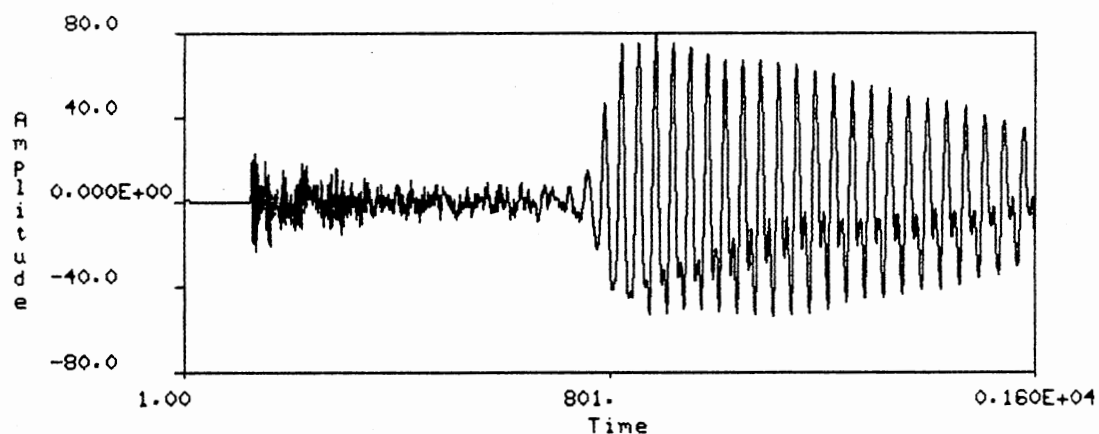
Figure 2. Speech Waveform

speech to have a high frequency spectra. Figure 3 is a plot of the
magnitude spectrum of the unvoiced segment /t/ from Figure 2, showing
the spectral nature of unvoiced speech.
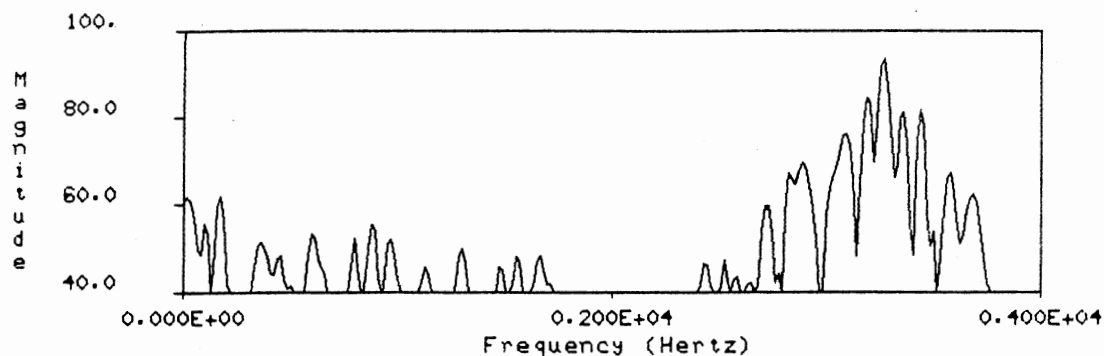
Figure 3. Unvoiced Speech Spectrum

The phoneme, /oo/, at the end of the word in Figure 2 is an example of voiced speech. Voiced speech is produced when the vocal tract is allowed to remain open and air is forced through the glottis with the tension of the vocal cords adjusted so that they produce periodic pulses of air which excite the vocal tract. This periodic excitation produces the low frequency character of voiced speech. Figure 4 is a plot of the magnitude spectrum for a segment of the voiced speech in Figure 2, demonstrating the spectral nature of voiced speech.

Another property which can be used to separate voiced and unvoiced speech is the short time energy in the speech waveform. The difference in short-time energy between voiced and unvoiced speech can be seen by observing the differences in amplitudes in Fig. 2. The unvoiced speech, produced by constricting the vocal tract, has a characteristically low amplitude and therefore low energy. On the other hand, voiced speech is produced by leaving the vocal tract open, has much higher amplitudes and thus higher energy.

The discrimination of the voiced speech, unvoiced speech and
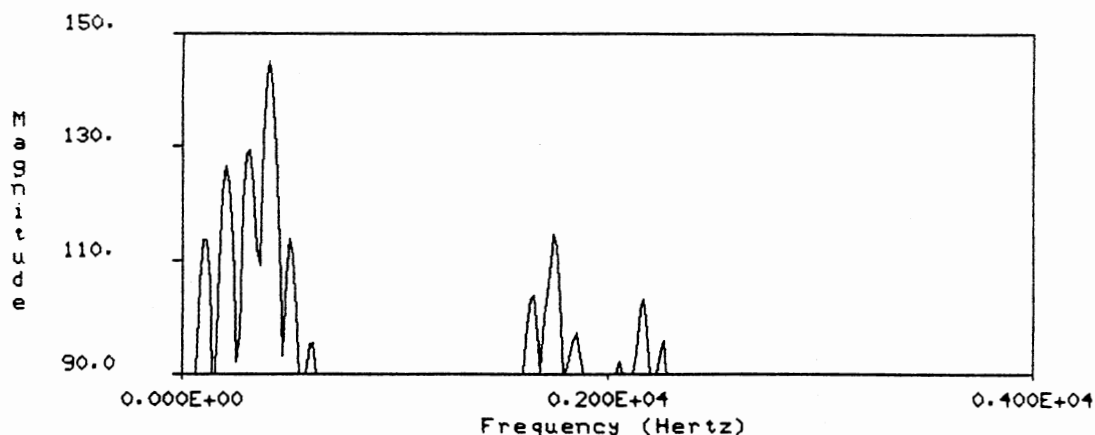
Figure 4. Voiced Speech Spectrum

silence would be simple in noise free speech. In such a case, all that would be required is to apply a threshold to detect the energy levels from a short-time energy measurement. Unfortunately, most realistic speech contains noise, and the short-time energy measurements take all frequencies into consideration. Therefore short-time energy measurements with a threshold is not the best way to discriminate between voiced speech, unvoiced speech and silence. It is therefore desirable, if possible, to get an energy measure which centralizes about the primary frequencies in the signal. Such a measure would not only aid in the detection of the beginning and ending of speech, it could also be useful in measuring the frequency shift needed to seperate voiced and unvoiced speech. The Generalized Allan Variance (GAV) is just such a measure.

The GAV is a frequency selective energy measure. The frequency selective nature of the GAV can be adjusted by varying certain parameters. By exploiting the frequency selective nature of the GAV, we are able to develop algorithms for the discrimination of voiced speech,

unvoiced speech and silence.

1.3 Literature Survey

1.3.1 Acoustic Well Logging

In Section 1.1, the use of the acoustic information in the evaluation of the formation surrounding the borehole is discussed. Many of the relationships mentioned are empirical, in that they are derived from measurement data. A considerable amount of literature exists on the analytical modeling of the acoustics of the borehole, all directed towards explaining the relationship of acoustic phenomena to rock properties and describing the other wave phenomena affecting the wavetrain.

One of the first efforts into modeling the acoustics of the borehole is presented in Biot [8]. Similar work can be found in White [85], White and Zechman [87], Peterson [63], Tsang and Radner [80] and Cheng and Toksoz [15]. All of this work centers on the production of synthetic logs and the explanation of the guided fluid waves. The frequency seperation between the compressional and shear waves, observed by Scarascia, Columbi and Cassinis [73], is confirmed analytically by Paillet [56,57,59] and Paillet and White [59].

Another important phenomena, which is not mentioned in the previous discussion, is the effect of 'slow' formations on shear waves. Slow formations are basically formations in which the shear wave velocity is less than the compressional velocity of the borehole fluid. In this case, the formation shear wave does not refract back into the borehole. Cheng and Toksoz [15] showed that for a 'slow' formation another prominent wave, the Stonely, can be used to derive the shear velocity.

With increased understanding of the acoustics and their relation-

ship to the formation, the need for better sondes came about. A sonde, proposed by Williams, Zemanek, Angona, Dennis and Caldwell [88], employs long spacings between the transmitter and receivers. The long spacings allow for more accurate recording and analysis of the wavetrains. Another tool, proposed by Zemanek, Angona, Williams, and Caldwell [95], is specifically built to excite shear waves in the formation. In this way, the recorded wavetrains are dominated by the shear wave, making shear wave logging much simpler. Some experimental sondes are built using up to 12 receivers. By recording time traces from each receiver, a two-dimensional acoustic picture of the formation can be generated.

Even with the improved tools, there is still a need for improved processing algorithms capable of separating out the various waves from the wavetrain. Many of the first efforts at recognizing the shear (S) wave centered around cross-correlation of the traces from different receivers, commonly referred to as semblence analysis. In this way, the time shift of the S-wave between the receivers can be measured. Different variations on this theme can be found in Scott and Sean [75], Aron [3], Willis and Toksoz [89], Dennis and Yang [28] and Kimball and Marzetta [44]. An overview of some of.these methods can be found in Willis and Toksoz [89]. A method similar to the cross-correlation technique is described by Ingram, Morris, Macknight and Parks [40]. This method, known as the Direct Phase Determination, uses cross-spectral techniques for the measurement of the time shift between two traces. The use of the cross-spectrum allows for the measurement of the time shift of only certain frequencies. In this way, only those frequencies dominated by the shear wave are used in the measurement of its time shift.

For the cases where more than 2 receivers are present the use of

two dimensional frequency analysis can be useful. Tanner and Koehler
[78] applied a two dimensional frequency-wavenumber (F-k) transform to
seismic data. Similarly, Parks, Morris and Ingram [62] and Parks, Mc-
Clellen and Morris [61] employ a frequency-wavenumber display (e.g. two
dimensional "Fourier" transform) of the acoustic array data to identify
the time shift of the compressional and shear wavelets. This display is
also capable of detecting other waves that are present in the wavetrain.

It should be noted that usually only 6 to 8 traces from different
transmitter receiver spacings are available. This means that standard
DFT techniques do not give sufficent frequency resolution in the spatial
dimension. For this reason, the previous algorithms developed by Parks
et al. [61,62] use Prony's method for the spatial frequency estimates.
This is why Fourier is placed in quotation marks above, to denote the
use of the specialized techniques for the transform.

Day and Yarlagadda [27] proposed the use of non-stationary time-
frequency analysis to identify the various wavelets. The non-stationary
analysis used in this approach is achieved via the Wigner Distribution.
The ability to detect the arrival of the shear wave, independent of
other signals, is not that desirable or required in the cases of acous-
tic log, but may have applications to other areas. Some of the results
in this paper are discussed further in later chapters.

1.3.2 Voiced-Unvoiced-Silence Discrimination

Many different areas of speech processing require the segmentation
and classification of the speech waveform into segments containing
voiced speech, unvoiced speech or silence [69]. The applications for
these segmentantion algorithms range from data compression, speech
coding, speech recognition, speaker recognition and others. Due to its

importance, many algorithms exist for the recognition of these different types of speech.

In section 1.2, dealing with speech, the nature of voiced speech versus unvoiced speech is discussed. It is pointed out that the two major differences between the two types of speech are their spectral character and their energy levels. Using these two facts, Rabiner and Sambur [68] developed algorithms based on the short-time properties of zero-crossing rate and short-time energy. Similar work involving these short-time properties and their approximations can be found in Drago, Molinari and Vagliani [29] and Wilpon, Rabiner and Martin [92]. Rosenthal, Schafer and Rabiner [72] use the energy in the Adaptive Delta Pulse Code Modulation (ADPCM) code words as a measure for detecting the presence of speech. The intent being that ADPCM is equally active for both voiced and unvoiced speech and inactive for silence, in this way it can balance out the large energy imbalance between phonemes. Similarly, Un and Lee [81] use the energy in the signal resulting from the delta modulation of the speech. Rabiner and Sambur [66] use a distance measure applied to the LP coeffients of the speech segment.

In all of these works, major effort is applied to the syntatics and semantics used in the segmentation of the waveform based on the measures. Atal and Rabiner [4] use a collection of these measures as part of a statistical pattern recognition system. The use of pattern recognition allows for the blending of the information from these measures, improving the decision process. Even with the pattern recognition system, some heuristic and ad hoc procedures were required to avoid undesirable interruptions in the speech.

1.3.3 Time Frequency Analysis

The occurrance of non-stationary signals is very common in nature.
In spite of this, most frequency transforms are based on the premise of
the signal being stationary. This paradox stands out prominently when
one considers the speech waveform, for the character of the speech wave-
form can change radically within a single word. This variable nature is
the reason Koenig, Dunn and Lacy [46] and Potter, Kopp and Green [65]
developed the spectrogram for the analysis of speech signals. With the
advent of digital computers and the rediscovery of the FFT, extensive
work in the analysis and synthesis of signals using the spectrogram or
more generally Short-Time Fourier Transform (STFT) techniques is now
possible. Tutorial overviews of STFT can be found in Rabiner and Schafer
[69] and Kodera, Gendrin and de Villedary [45].

An expansion of a signal in terms of a weighted sum of time shifted
and modulated Gaussian envelopes is described by Gabor [34]. The three
dimensional plot of these weighting coefficients as a function of time
and frequency can can be used as a measure of the local time-frequency
variations of the signal. These Gaussian signals, commonly called
Gabor functions, are shown to have an exact continuous time and fre-
quency form by Helstrom [38]. Helstrom also shows that the continuous
Gabor functions are proportional to the continuous time STFT analysis
procedures using Gaussian shaped windows.

The Gaussian shape of the Gabor functions is the direct result of
the minimization of the product of the time duration (ΔT) and frequency
bandwidth (ΔF). This product is minimized when it equals 1/2. This is
a very important result, since it shows that the time duration and fre-
quency bandwidth of a signal cannot be made arbitrarily small simultan-

eously. This is commonly called the Uncertainty Principle of Fourier Analysis [11].

An excellent overview on classical time-frequency analysis is presented by Boudreaux-Bartels [11]. Chapter II of Boudreaux-Bartels' dissertation is a complete tutorial on STFT, Gabor representations, the Ambiguity function and the Wigner Distribution.

1.3.4 Wigner Distribution

The Wigner Distribution (WD) can be credited to Eugene Wigner from his work in quantum mechanics [91]. It is defined for a continous signal $f(t)$ in the form

$$W_f(t,w) = \int_{-\infty}^{\infty} f(t+\tau/2)\, f^*(t-\tau/2)\, e^{-jw\tau}\, d\tau \qquad (1.4)$$

where (*) corresponds to complex conjugation and w is the usual radian frequency symbol. The WD can be used in signal processing for time frequency analysis. Towards this, a definition for discrete signals, by Classen and Mecklenbrauker [20], is given by

$$W_f(n,\theta) = \sum_{k=-\infty}^{\infty} f(n+k)\, f^*(n-k)\, e^{-j2k\theta} \qquad (1.5)$$

An alternate definintion to that of (1.5) is given below. The advantages of this new definition are discussed later. This alternate Wigner Distribution, proposed by Day and Yarlagadda [27], is defined as

$$W_f(n,\theta) = \sum_{k=-\infty}^{\infty} f(n+k+1)\, f^*(n-k)\, e^{-j2(k+1/2)\theta} \qquad (1.6)$$

This definition is referred to hereafter as the Modified Auto Wigner Distribution (MAWD).

The Wigner distribution can also be defined for two different signals. The Modified Cross Wigner Distribution (MCWD), is defined by

$$W_{f,g}(n,\theta) = \sum_{k=-\infty}^{\infty} f(n+k+1) \, g^*(n-k) \, e^{-j2(k+1/2)\theta} \qquad (1.7)$$

The bilinear nature of the Wigner distribution is the cause of three properties, which are the subject of considerable research. These three properties are 1) the possibility of negative values in the WD, 2) frequency aliasing and 3) the cross-terms that appear between spectral components.

The possible occurrence of negative values in the WD is mentioned by Wigner [91]. These negative values make an appropriate interpretation of the WD difficult, since negative values disallow the interpretation of the WD as a distribution of the signal energy or as a probablity distribution. In an effort to correct this negative value problem, Cohen and Posch [26] introduced a procedure to generate time-frequency (phase-space) distributions. Cohen showed that positive time-frequency distributions could be found for any signal, however these distribution must be either signal dependent or not be bilinear. Cohen noted in his paper that in spite of the negative values, the WD can be used as a computational tool, but care must be taken to not interpret it as an energy or probablity distribution.

The aliasing problem in the WD is similar to that encountered when sampling an analog signal, except in the case of the WD, the aliasing occurs if the signal contains frequency components greater than 1/4 the sampling frequency. This means that to use the WD to analyze a signal, it must be sampled at twice the Nyquist rate . Chan [13] and Boudreuax-Bartels [12] have both given alternate definitions in an effort to alleviate the need for over sampling. Classen and Mecklenbrauker [23] compared these alternate definitions and found each to be lacking in

certain points. The conclusion of Claasen and Mecklenbrauker's compari-
son is that if the WD is to be used and it is possible to over sample
the analog signal their definition (Equation (1.5)) is to be used, based
on its ease of computation. The MAWD allows for the use of optimal
(power of 2) FFT techniques, for the computation of its summation. This
computational efficiency is the primary reason for our choosing this
definition in our work.

The artifacts known as cross terms are present when using any def-
inition of the WD and are generated by the cross-product of the signal
with itself. In some of the first applications of the Wigner distribu-
tion, these terms were interpreted as portraying some character of the
signal. However, these artifacts have no real physical signifigance, and
in many cases can mask or accentuate the actual spectral components gen-
erating them. It is therefore important that some means of eliminating
these or at least reducing these terms be included in any application of
the WD.

One of the first efforts in reducing these cross terms, presented
by Flandrin [31], involves the filtering of the Wigner distribution with
respect to the time index. This type of filtering can reduce the magni-
tude of the cross terms, since they are generally modulated in time.
However, this technique reduces the time resolution of the WD, and time
resolution is one of the original reasons for choosing the WD.

Another area of study concerning the Wigner Distribution is the
synthesis of signals from a given WD. This synthesis is an important
step in implementing time varying signal processing. Towards this goal,
a technique for the generation of signals based on a desired WD, is
given by Boudreaux-Bartels and Parks [10].

1.3.5 Allan Variance

The Allan variance, proposed by David Allan [1] for the analysis of the frequency stability of atomic oscillators, is defined by

$$\sigma(x) = \langle (x(i)-x(i-1))^2 \rangle \tag{1.8}$$

where x is a signal and $\langle \rangle$ represents the time average over i. A modified version of the Allan Variance, proposed by Allan and Barnes [2] is given by

$$\sigma_n(x) = \langle [\sum_{I=0}^{n-1} (x(i-I) - 2 \, x(i-I-n) + x(i-I-2n))]^2 \rangle \tag{1.9}$$

The signal, x(i), in Allan's work is actually the averaged frequency, measured from the output of an oscillator using a frequency counter. The Allan variance can therefore be thought of as basically a measure applied to the signal x(i), and not strictly as a measure of frequency stability.

Most of the literature on the Allan variance is for the area of measuring frequency stability, such as Barnes et al. [5]. The primary purpose of this work is to determine the response of the Allan variance to various power law spectra. Power law spectra are important in the area of frequency stability, since these represent the primary noise spectra encountered.

Lesage and Audoin [46] discuss the statistical nature of the Allan variance and show that the Allan variance is not biased by short-time estimations. Also, Lesage and Audoin show that the variance of the measure is a function of the number of samples averaged, similar to the law of large numbers.

1.4 Chapter Overview

Chapter II of this thesis introduces the MAWD and MCWD, along with their discrete versions. The properties for each are listed, along with proofs. Chapter III describes the Allan variance and define a generalized version. Applications of the Generalized Allan Variance (GAV) to acoustic well logging and speech processing are discussed. Chapter IV begins with a brief introduction to the basics of pattern recognition. It also contains descriptions of how the MAWD and Generalized Allan Variance can be used in the analysis of the acoustic well log and speech signals, respectively. Finally, in Chapter V, the conclusions reached in this research are presented, and future areas of research are indicated.

CHAPTER II

THE MODIFIED WIGNER DISTRIBUTION

2.0 Introduction

The distribution of the frequency components of a signal over time
is an important property and is used in the analysis of a large variety
of signals. The applications of time-frequency analysis can be found in
radar, sonar, seismic prospecting, medical imaging and many others. In
each of these areas, different techniques, based on the type of signals
involved, are used. One of these techniques, the Wigner Distribution, is
a useful tool for the analysis of non-stationary signals.

The Wigner distribution, introduced by Eugene Wigner [88] and pro-
posed as a phase-space distribution for the analysis of wave functions,
can be used for time-frequency analysis of time signals (Ville,[79]).
Little research involving the WD is documented, until 1980 when Claasen
and Mecklenbrauker [20] reintroduced it and defined a discrete time
version. A large number of applications are presently under consider-
ation as possible uses of the WD.

In this chapter a modified definintion of the WD, hereafter called
the Modified Auto Wigner Distribution(MAWD), is introduced. The MAWD
maintains many of the properties of the classical WD and allows for more
efficient computation. The properties of the MAWD are derived. The
occurance of certain artifacts known as cross terms, is brought out and
techniques are described for their reduction. Finally, examples of the

21

MAWD applied to various signals are included.

2.1 Modified Wigner Distribution

The MAWD, defined by Day and Yarlagadda [27], is given by

$$W_f(n,\theta) = \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+(1/2))\theta} \qquad (2.1)$$

The advantages of this definition are to be described later, when we discuss procedures for computing the summation. Similarly, a Cross Wigner Distribution can be defined by

$$W_{f,g}(n,\theta) = \sum_{k=-\infty}^{\infty} f(n+k+1)g^*(n-k) \, e^{-j2(k+(1/2))\theta} \qquad (2.2)$$

The MCWD can also be computed more efficiently than previous definitions. Although the MCWD is clearly a more general definition and may prove applicable in later research, we shall concentrate on the MAWD, since it is more directly applicable to our work.

2.2 Some Properties of the MAWD

In the following some of the interesting properties of the MAWD are discussed.

1. Inverse Operation

The inverse of (2.1) is

$$f(n+r+1) \, f^*(n-r) = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} W_f(n,\mathcal{I}) \, e^{j2(r+1/2)\mathcal{I}} \, d\mathcal{I} \equiv A \qquad (2.3)$$

This can be shown by substituting the definintion of $W_f(n,\mathcal{I})$ into (2.3). Now,

$$A = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\mathcal{I}} \, e^{j2(r+1/2)\mathcal{I}} \, d\mathcal{I} \quad (2.4)$$

Reordering the summation and the integral gives

$$A = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \int_{-\pi/2}^{\pi/2} e^{-j2(k-r)\mathcal{I}} \, d\mathcal{I} \quad (2.5)$$

The integral in (2.5) can be expressed as

$$\int_{-\pi/2}^{\pi/2} e^{-j2(k-r)\mathcal{I}} \, d\mathcal{I} = \frac{\sin(k-r)\pi}{(k-r)} \quad , \quad k \neq r.$$

Since k and r are integers, the integral in (2.5) is zero for all $k \neq r$.

When k=r, the integral reduces to π and the result   . . . .

$A = f(n+r+1) \, f^*(n-r)$ follows.

The problem with this result is that f(n) cannot be found directly. Rather, some value must be assumed or known for the first non-zero point f(a). Unlike Classen's definition for the WD, where at least the magnitude of f(a) can be found, there is no way to recover anything more than the array $f(n+r+1) \, f^*(n-r)$.

2. Real Value Property

The MAWD is always real. That is,

$$W_f(n,\theta) = (W_f(n,\theta))^* \quad (2.6)$$

This can be proved by rewriting the summation $W_f(n,\theta)$ as,

$$W_f(n,\theta) = \sum_{k=0}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$+ \sum_{k=-\infty}^{-1} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

Substituting k=-m-1 in the summation over negative k, we have

$$W_f(n,\theta) = \sum_{k=0}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$+ \sum_{m=0}^{\infty} f(n-m) \, f^*(n+m+1) \, e^{j2(m+1/2)\theta}$$

$$= \sum_{k=0}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$+ \sum_{k=0}^{\infty} f(n-k) \, f^*(n+k+1) \, e^{j2(k+1/2)\theta}$$

$$= \sum_{k=0}^{\infty} [\, f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$+ f(n-k) \, f^*(n+k+1) \, e^{j2(k+1/2)\theta} \,]$$

The term inside the summation is a sum of a complex number and its conjugate, making the term real. From this the real valued property of the MAWD follows.

3. Symmetric and Periodic with Respect to the Frequency Variable.

The symmetry of the MAWD in the frequency domain can be character-ized as

$$W_f(n,\theta) = W_f^*(n,-\theta) \tag{2.7}$$

Using (2.1), we have

$$W_f^*(n,-\theta) = \sum_{k=-\infty}^{\infty} f^*(n+k+1) \, f(n-k) \, e^{-j2(k+1/2)(-\theta)}$$

Using $k = -m-1$,

$$W_f^*(n,-\theta) = \sum_{m=-\infty}^{\infty} f^*(n+(-m-1)+1) \, f(n-(-m-1)) \, e^{-j2(-m-1+1/2)(-\theta)}$$

Simplifying, we have

$$W_f^*(n,-\theta) = \sum_{m=-\infty}^{\infty} f^*(n-m) \, f(n+m+1) \, e^{-j2(m+1/2)\theta}$$

which is $W_f(n,\theta)$ by definintion, and the symmetry property of (2.7) follows.

The periodicty of the MAWD in the frequency domain is character- ized by

$$W_f(n,\theta+\pi m) = (-1)^m W_f(n,\theta) \tag{2.8}$$

Where m is an integer. This property can be proved by expressing

$$W_f(n,\theta+\pi m) = \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)(\theta+\pi m)}$$

$$= \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta} \, e^{-j2(k+1/2)\pi m}$$

Seperating out the terms containing m, we have

$$W_f(n,\theta+\pi m) = e^{-j2(1/2)m\pi} \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta} \, e^{-j2\pi km}$$

$$= e^{-j\pi m} \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta} \, (1)^{km}$$

$$= (-1)^m \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$= (-1)^m \, W_f(n,\theta)$$

which completes the proof.

It should be noted that the magnitude of the MAWD is periodic with period πf_s (f_s is assumed to be 1 Hz in the previous development) and not $2\pi f_s$ as is the case in the Fourier transform of discrete signals. This implies that an analog signal must be sampled at twice the Nyquist rate to avoid aliasing.

4. Time Limited Signals have Time Limited MAWD's.

Let f(n) be a time limited signal, such that

$f(n) = 0$ for $n < n_1$ and $n_2 < n$ with $n_1 < n_2$.

Then the MAWD of $f(n)$ is also time limited, with

$$W_f(n,\theta) = 0 \quad \text{for } n < n_1 \text{ and } n_2 \le n. \tag{2.9a}$$

As proof, consider the definition of the MAWD for $f(n)$,

$$W_f(n,\theta) = \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta} \tag{2.9b}$$

Defining $m = n-k$, we have

$$W_f(n,\theta) = \sum_{m=-\infty}^{\infty} f(2n-m+1) \, f^*(m) \, e^{-j2(n-m+1/2)\theta} \tag{2.9c}$$

Using the time limitedness property of $f(m)$, we can express

$$W_f(n,\theta) = \sum_{m=n_1}^{n_2} f(2n-m+1) \, f^*(m) \, e^{-j2(n-m+1/2)\theta} \tag{2.9d}$$

With $n_1 = n - i$, we have

$$W_f(n_1-i,\theta) = \sum_{m=n_1}^{n_2} f(2n_1-2i-m+1) \, f^*(m) \, e^{-j2(2n_1-2i-m+1/2)\theta} \tag{2.9e}$$

For $i > 0$, $f(2n_1-2i-m+1) = 0$ for all values of m in the range $n_1 < m < n_2$. This implies

$W_f(n_1-i,\theta) = 0$, for $i > 0$ or $n < n_1$.

Similarly, let $n = n_2+i$, then

$$W_f(n_2+i,\theta) = \sum_{m=n_1}^{n_2} f(2n_2+2i-m+1) \, f^*(m) \, e^{-j2(2n_2+2i-m+1/2)\theta} \tag{2.9f}$$

For $i \ge 0$, $f(2n_2+2i-m+1) = 0$ for $n_1 \le m \le n_2$. This implies

$W_f(n_2+i,\theta) = 0$ for $i \ge 0$, or $n \ge n_2$.

It should be noted that the limits of the MAWD is one shorter than the limits on the signal.

5. Relationship to the Fourier Transform of $f(n)$

The relationship between the MAWD and the Fourier transform of f(n) is important. First, the Fourier transform of f(n) is [52]

$$[f(n)] = F(\theta) = \sum_{k=-\infty}^{\infty} f(k) \, e^{-jk\theta} \qquad (2.10)$$

and the inverse

$$^{-1}[F(\theta)] = f(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\mathfrak{X}) \, e^{-jn\mathfrak{X}} \, d\mathfrak{X} \qquad (2.11)$$

The MAWD of F(θ) is defined as

$$W_F(\theta,n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\theta+\mathfrak{X}) \, F^*(\theta-\mathfrak{X}) \, e^{j2(n+1/2)\mathfrak{X}} \, d\mathfrak{X} \qquad (2.12)$$

From this, we can state the interesting property that

$$W_F(\theta,n) = W_f(n,\theta) \qquad (2.13)$$

This can be shown by substituting the definition of F(θ) into the definition of $W_F(\theta,n)$.

$$W_F(\theta,n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [\sum_{m=-\infty}^{\infty} f(m) \, e^{-jm(\theta+\mathfrak{X})} \sum_{r=-\infty}^{\infty} f^*(r) \, e^{jr(\theta-\mathfrak{X})}$$

$$\cdot \, e^{-j2(n+1/2)\mathfrak{X}}] \, d\mathfrak{X}$$

Reordering the summations and the integral produces

$$W_F(\theta,n) = \sum_{m=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} f(m) \, f^*(r) \, e^{-j(m-r)\theta}$$

$$\cdot \, \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j(m+r-2n-1)\mathfrak{X}} \, d\mathfrak{X}$$

The integral in the above equation is similar to the integral in (2.5). Thus, we know that the integral is zero, except when m+r-2n-1=0, where it is 2π. This means the term under the summations is zero, except for

r=2n+1-m. The two summations can now be reduced to one. That is,

$$W_F(\theta,n) = \sum_{m=-\infty}^{\infty} f(m) \; f^*(2n+1-m) \; e^{-j(m-2n-1+m)\theta}$$

Let m = n+k+1 in the above equation. Then we have

$$W_F(\theta,n) = \sum_{k=-\infty}^{\infty} f(n+k+1) \; f^*(n-k) \; e^{-j2(k+1/2)\theta}$$

where the right hand side is the definition of $W_f(n,\theta)$, and (2.13) follows.

6. The Inverse Operation for MAWD of F(θ)

The inverse for the MAWD of F(θ) (see (2.12)) is

$$1/2 \; [\; F(\theta+\Psi) \; F^*(\theta-\Psi) - F(\theta+\Psi+\pi) \; F^*(\theta-\Psi-\pi) \;]$$

$$= \sum_{k=-\infty}^{\infty} W_F(\theta,k) \; e^{-j2(k+1/2)\Psi} \qquad \underline{=} \; B \qquad\qquad (2.14)$$

Noting (2.13), we have also

$$B = \sum_{k=-\infty}^{\infty} W_f(k,\theta) \; e^{-j2(k+1/2)\Psi}$$

Equation (2.14) can be proved by first substituting the defintion of $W_F(\theta,m)$ into (2.14). Now,

$$B = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} F(\theta+\Im) \; F^*(\theta-\Im) \; e^{j2(k+1/2)\Im} \; d\Im \qquad e^{-j2(k+1/2)\Psi}$$

Reordering the summation and integral, we have

$$B = \int_{-\pi}^{\pi} F(\theta+\Im) \; F^*(\theta-\Im) \; e^{-j(\Psi-\Im)} \; [\; \sum_{k=-\infty}^{\infty} \frac{e^{-j2(\Psi-\Im)k}}{2} \;] \; d\Im$$

Let 2(Ψ-ℑ) = λ or ℑ =Ψ-λ/2. Then we have

$$B = \int_{2\Psi+2\pi}^{2\Psi-2\pi} [\ F(\theta+\Psi-\lambda/2) \ F^*(\theta-\Psi+\lambda/2) \ e^{-j(\lambda/2)}$$

$$\frac{1}{2\pi} [\ \sum_{k=-\infty}^{\infty} e^{-j\lambda k} \ ] \ (-1/2) \ ] \ d\lambda \tag{2.14b}$$

Consider the identity, [37,52]

$$\sum_{k=-\infty}^{\infty} \delta(\lambda-2\pi k) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{jk\lambda} \tag{2.15}$$

where $\delta(\lambda)$ is the unit delta function.

Reversing the limits of integration and substituting (2.15) into (2.14b), we have

$$B = \int_{2\Psi-2\pi}^{2\Psi+2\pi} F(\theta+\Psi-\lambda/2) \ F^*(\theta-\Psi+\lambda/2) \ e^{-j(\lambda/2)}$$

$$[\ \sum_{k=-\infty}^{\infty} \delta(\lambda-2k\pi) \ ] \ (1/2) \ d\lambda$$

For any value of Ψ, there are only two values of λ for which $\lambda-2k\pi$ is equal to zero. For example, if $\Psi=\pi$, λ ranges from 3π to $-\pi$; thus only $n=0$, and $n=1$ occur in the summation. Using the sifting property of the delta function and the periodicity of $F(\theta)$, B can be expressed as

$$B = 1/2 [\ F(\theta+\Psi) \ F^*(\theta-\Psi) \ e^{-j0} + F(\theta+\Psi+\pi) \ F^*(\theta-\Psi-\pi) \ e^{-j\pi}]$$

$$= 1/2 [\ F(\theta+\Psi) \ F^*(\theta-\Psi) - F(\theta+\Psi+\pi) \ F^*(\theta-\Psi-\pi) \ ]$$

which proves (2.14).

By setting $\Psi=0$, (2.14) becomes

$$1/2 \ (F(\theta) \ F^*(\theta)) = \sum_{k=-\infty}^{\infty} W_f(k,\theta)$$

where it is assumed that $F(\theta)$ is band limited to $\pm(\pi/2)f_s$. This property is vitally important, since it establishes that the marginal fre-

quency distribution of the MAWD is equal to the spectrum of the signal f(n). The marginal frequency distribution for a time-frequency distribution is defined as the sum of the distribution, at a given frequency, over all time. The MAWD's marginal frequency distribution being equal to the spectrum, establishes a direct correspondence between the frequency variable in each.

7. The Effect of Windowing on the MAWD

If the function f(n) is not time limited, it would be impossible to calculate $W_f(n,\theta)$, as it contains an infinite summation. To estimate the MAWD at some time, we need to window the function f(n). The question is what effect does this have on our estimate of $W_f(n,\theta)$.

Let h(n) = f(n) g(n), where g(n) is a time limited window function. It can be shown that

$$W_h(n,\theta) = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} W_f(n,\mathcal{F}) \, W_g(n,\theta-\mathcal{F}) \, d\mathcal{F} \; \underline{=} \; C \tag{2.16}$$

or

$$W_h(n,\theta) = W_f(n,\theta) *_2 W_g(n,\theta)$$

where $*_2$ denotes convolution with respect to the second variable. This property can be proven by using (2.1) for $W_f(n,\mathcal{F})$ and $W_g(n,\theta-\mathcal{F})$ into (2.16). Now,

$$C = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\mathcal{F}}$$

$$\cdot \sum_{m=-\infty}^{\infty} g(n+m+1) \, g^*(n-m) \, e^{-j2(m+1/2)(\theta-\mathcal{F})} \qquad d\mathcal{F}$$

Reordering the summations and the integral gives

$$C = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, g(n+m+1) \, g^*(n-m) \, e^{-j2(m+1/2)\theta}$$

$$\cdot \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} e^{-j2(k-m)\mathcal{F}} \, d\mathcal{F}$$

The integral is zero, except when k=m. As a result, the double summation now reduces to one summation and

$$C = \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, g(n+k+1) \, g^*(n-k) \, e^{-j2(k+1/2)\theta}$$

Recalling that h(n) = f(n) g(n), we have

$$C = \sum_{k=-\infty}^{\infty} h(n+k+1) \, h^*(n-k) \, e^{-j2(k+1/2)\theta}$$

which is by definition $W_h(n,\theta)$ and (2.16) follows.

The effect of windowing on the MAWD is very similar to that found in Short-Time Fourier Analysis (STFA), One important difference is that the time resolution of the MAWD is not effected by the window length, which is not the case in STFA. To achieve good frequency resolution however, we need a good window function. To demonstrate the differences between windows, plots are generated of the MAWD for four windows that are commonly used in STFA. These plots are shown in Figures 5, 6, 7 and 8.

The plots in Figures 5,6,7 and 8 are produced by generating a complex exponential wave at a radial frequency of $(\pi/2)f_s$. The MAWD of exponential is then computed using the various windows. The conversion to dB is accomplished by taking the logrithm of the magnitude, and multiplying by ten. Ten is used instead of twenty, since the MAWD is related to the square of the Fourier transform and thus the data is already squared.

Figure 5 is the MAWD for the center point of a rectangular window. Note, that the MAWD does not decrease very fast for this window, and in fact the first side lobes are at approximately -13 dB. Figure 6 is the MAWD of a Hamming window. The Hamming window has a wider center lobe than the rectangular window. The first side lobes for the Hamming win-
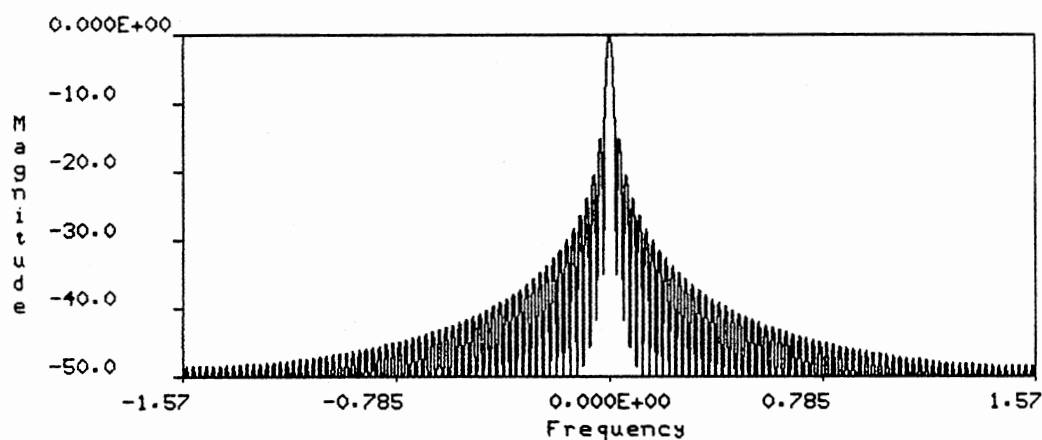
Figure 5. Rectangular Window
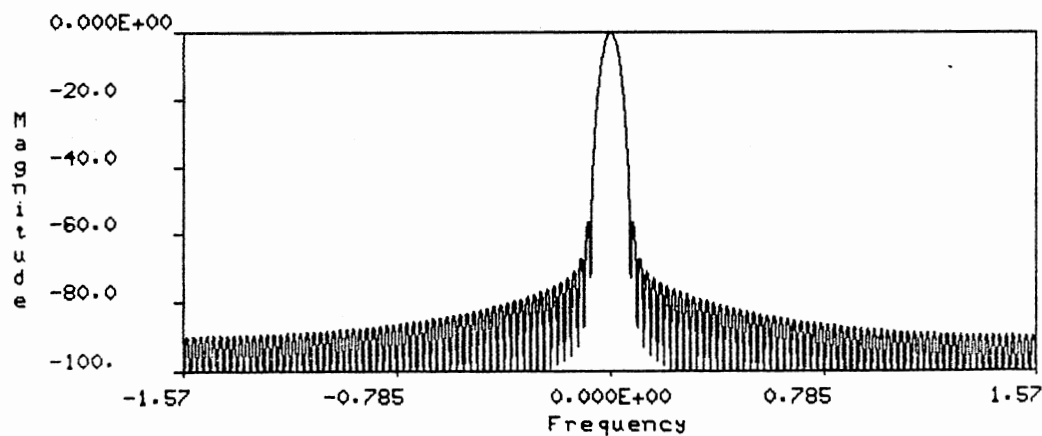
Figure 6. Hamming Window

dow are at -55 dB.

Figures 7 and 8 are for the Kaiser Window [45]. The Kaiser window can be adjusted by changing an input parameter "Beta". As can be seen from Figures 7 and 8, a change in Beta can cause considerable change in
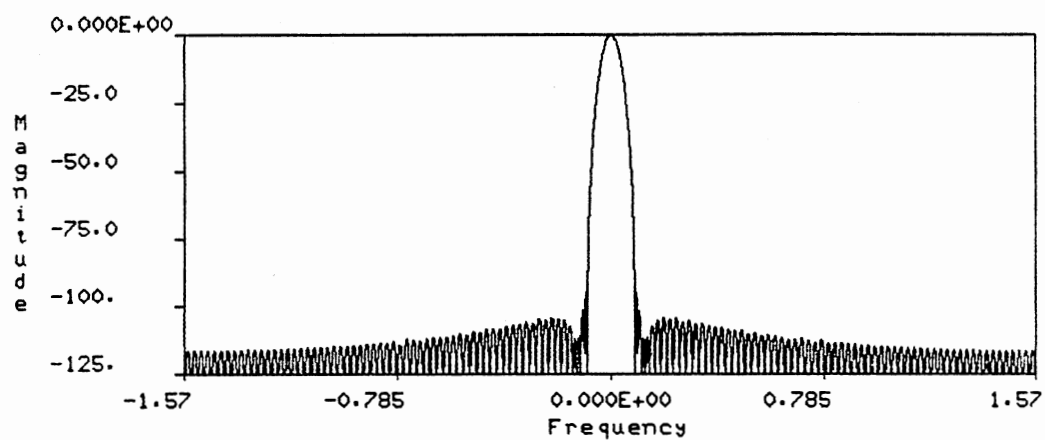
Figure 7. Kaiser Window (Beta = 5.5)
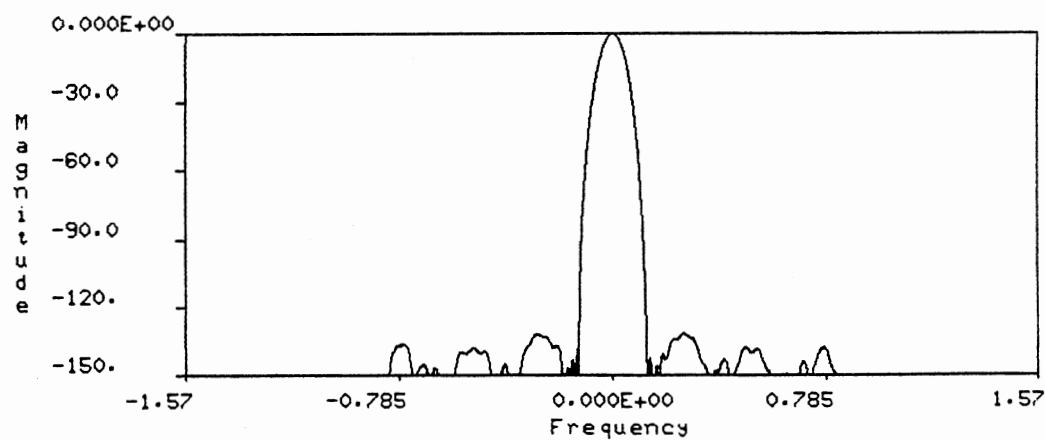
Figure 8. Kaiser Window (Beta = 8.0)

the shape of the MAWD. For Beta equal to 8.0 (Figure 8), the center lobe is rather wide, but outside this lobe the MAWD goes below our computational accuracy (This is the reason for the noisy appearance).

8. The Effect of Linear Filtering on the MAWD.

In this section, we consider the MAWD of a linearly filtered signal f(n). Linear filtering can be characterized as a convolution, given by

$$g(n) = \sum_{m=-\infty}^{\infty} f(m) \, h(n-m) \qquad (2.17a)$$

where h(n) is the unit sample response of the filter. In the transform domain,

$$G(\theta) = F(\theta) \, H(\theta) \qquad (2.17b)$$

where $G(\theta)$, $F(\theta)$ and $H(\theta)$ are the transforms of g(n), f(n) and h(n), respectively. We now state the following formula for the MAWD of g(n),

$$W_g(n,\theta) = \sum_{m=-\infty}^{\infty} W_f(m,\theta) \, W_{\bar{h}}(n-m,\theta) \qquad (2.18)$$

where $\bar{h}(n)$ is obtained from h(n) by shifting the sequence by half of a sample. In the transform domain,

$$[\bar{h}(n)] = H(w) \, e^{-j(w/2)}.$$

Equation (2.18) states that convolution in the time domain, is equivalent to the convolution, $W_f(n,T) *_1 W_{\bar{h}}(n,T)$, where $*_1$ corresponds to the convolution with respect to the first variable. The term $e^{-j(w/2)}$ is not a simple filter. For one thing it is periodic with period 4π. This means that to generate h(n), we must compute H(w) for $0 < w < 4\pi$. This can be done in two ways. The first is to add a zero between each of the points in the sequence h(n) before computing the Fourier transform. Another is to compute the Fourier transform and then simply repeat the transform for 2π to 4π. Once the transform has been computed, it is phase shifted

by $e^{-j(w/2)}$ and then inverse transformed. The output sequence is now twice the length of the original sequence and is decimated by taking only the odd numbered terms. Note that the previous operation can be thought of as shifting the sequence by half of a sample.

Equation (2.18) can be proved by first replacing the MAWD of f(n) and $\bar{h}(n)$ with the MAWD of their Fourier transforms and then replacing the MAWD by their definintions. These operations yield

$$W_g(n,\theta) = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\theta+\bar{x}) F^*(\theta-\bar{x}) e^{j2(m+1/2)\bar{x}} d\bar{x}$$

$$\cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\theta+\psi) e^{-j(\theta+\psi)/2} H^*(\theta-\psi) e^{j(\theta-\psi)/2} e^{j2(n-m+1/2)\psi} d\psi$$

Reordering the integrals and the summation and collecting the terms containing m, we have

$$W_g(n,\theta) = \frac{1}{2\pi\, 2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F(\theta+\bar{x}) H(\theta+\psi) F^*(\theta-\bar{x}) H^*(\theta-\psi) e^{j2n\psi} e^{j\bar{x}}$$

$$\cdot \sum_{m=-\infty}^{\infty} e^{j2m(\bar{x}-\psi)} d\bar{x}\, d\psi$$

The summation can now be replaced with an impulse train (See (2.15)), yielding

$$W_g(n,\theta) = \frac{1}{2\pi\, 2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F(\theta+\bar{x}) H(\theta+\psi) F^*(\theta-\bar{x}) H^*(\theta-\psi) e^{j2n\psi} e^{j\bar{x}}$$

$$\cdot 2\pi \sum_{m=-\infty}^{\infty} \delta(\bar{x}-\psi+m\pi) d\bar{x}\, d\psi$$

Assuming that F and H are band limited to $\pm(\pi/2)f_s$ and using the sifting property of the delta function , the two integrals reduce to one and we have

$$W_g(n,\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\theta+\mathcal{I}) H(\theta+\Psi) F^*(\theta-\mathcal{I}) H^*(\theta-\Psi) e^{j2(n+1/2)\Psi} d\mathcal{I}$$

Recalling the transform definition of $G(\theta)$, we have

$$W_g(n,\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} G(\theta+\mathcal{I}) G^*(\theta-\mathcal{I}) e^{j2(n+1/2)\mathcal{I}} d\mathcal{I}$$

which is the definition of the MAWD of $G(\theta)$. This completes the proof.

2.3 Some Properties of the MCWD

In the following some of the interesting properties of the Modified Cross Wigner Distribution (MCWD) (see Equation 2.2) are discussed.

1. Inverse Operation

The inverse operation of the MCWD is

$$f(n+r+1) g^*(n-r) = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} W_{f,g}(n,\mathcal{I}) e^{j2(r+1/2)\mathcal{I}} d\mathcal{I} \underline{\underline{\ }} D \qquad (2.19)$$

This can be proven by substituting (2.2) into (2.19) and simplifying. Now,

$$D = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} [\sum_{k=-\infty}^{\infty} f(n+k+1) g^*(n-k)$$

$$\cdot e^{-j2(k+1/2)\mathcal{I}} e^{j2(r+1/2)\mathcal{I}}] d\mathcal{I}$$

Reordering the summation and integral, we have

$$D = \sum_{k=-\infty}^{\infty} f(n+k+1) g^*(n-k) \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} e^{-j2(k-r)\mathcal{I}} d\mathcal{I}$$

The integral is zero, except when $k=r$ where it is equal to π. Using this property, the summation reduces to

$$D = f(n+r+1) g^*(n-r)$$

which completes the proof.

2. Periodic with Respect to the Frequency Variable.

The MCWD is periodic with respect to its frequency variable. This can be expressed as

$$(-1)^m W_{f,g}(n,\theta) = W_{f,g}(n,\theta+\pi m) \qquad (2.20)$$

From the definition

$$W_{f,g}(n,\theta+\pi m) = \sum_{k=-\infty}^{\infty} f(n+k+1) \ g^*(n-k) \ e^{-j2(k+1/2)(\theta+\pi m)}$$

Simplifying, we have

$$W_{f,g}(n,\theta+\pi m) = \sum_{k=-\infty}^{\infty} f(n+k+1) \ g^*(n-k) \ e^{-j2(k+1/2)\theta} \ e^{-j2(k+1/2)\pi m}$$

$$= e^{-j\pi m} \sum_{k=-\infty}^{\infty} f(n+k+1) \ g^*(n-k) \ e^{-j2(k+1/2)\theta} \ e^{-j2\pi km}$$

Recalling the $e^{-j\pi} = -1$ and $e^{-j2\pi} = 1$, we have

$$W_{f,g}(n,T+mp) = (-1)^m \sum_{k=-\infty}^{\infty} f(n+k+1) \ g^*(n-k) \ e^{-j2(k+1/2)\theta} (1)^{km}$$

Since k and m are integers, 1^{km} is always one and this term can be re-moved. Also, the summation is now the definition of $W_{f,g}(n,\theta)$ giving the result

$$W_{f,g}(n,\theta+\pi m) = (-1)^m W_{f,g}(n,\theta)$$

3. Relationship to the Fourier Transform.

As in the case of the MAWD, the MCWD can be shown to have a close relationship to the Fourier transform of the corresponding signals. The definition of the MCWD for the Fourier transforms of f(n) and g(n) (F(θ) and G(θ), respectively) is given by

$$W_{F,G}(\theta,n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\theta+\mathcal{I}) \, G^*(\theta-\mathcal{I}) \, e^{-j2(n+1/2)\mathcal{I}} \, d\mathcal{I} \qquad (2.21)$$

Based on (2.21) it can be shown that

$$W_{F,G}(\theta,n) = W_{f,g}(n,\theta) \qquad (2.22)$$

Equation (2.22) can be proved by expressing $F(\theta)$ and $G(\theta)$ in terms of $f(n)$ and $g(n)$. Now,

$$W_{F,G}(\theta,n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\sum_{k=-\infty}^{\infty} f(k) \, e^{-jk(\theta+\mathcal{I})} \right.$$
$$\left. \cdot \sum_{m=-\infty}^{\infty} g^*(m) \, e^{-jm(\theta-\mathcal{I})} \, e^{j2(n+1/2)\mathcal{I}} \right] d\mathcal{I}$$

Reordering the summation and integral and simplifying the expression, we have

$$W_{F,G}(\theta,n) = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f(k) \, g^*(m) \, e^{-j(k-m)\theta}$$
$$\cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j(k+m-2n-1)\mathcal{I}} \, d\mathcal{I}$$

The integral is zero, except when $k+m-2n-1=0$, when it is equal to 2π. Using this property, it follows that

$$W_{F,G}(\theta,n) = \sum_{k=-\infty}^{\infty} f(k) \, g^*(2n+1-k) \, e^{-j(k-(2n-k+1))\theta}$$

By setting $k = n+h+1$ and then simplifying, we have

$$W_{F,G}(\theta,n) = \sum_{h=-\infty}^{\infty} f(n+h+1) \, g^*(n-h) \, e^{-j2(h+1/2)\theta}$$

The right hand side is the definition of $W_{f,g}(n,\theta)$, proving (2.22).

4. The Inverse of the MAWD of the Fourier Transform

We now consider the inverse of $W_{F,G}(T,n)$. The inverse is given by

$$1/2 \ [\ F(\theta+\psi) \ G^*(\theta-\psi) - F(\theta+\psi+\pi) \ G^*(\theta-\psi-\pi) \ ]$$

$$= \sum_{k=-\infty}^{\infty} W_{F,G}(\theta,k) \ e^{-j2(k+1/2)\psi} \ \underline{\underline{=}} \quad E \qquad (2.23)$$

By using (2.22), E can also be expressed as

$$E \ = \ \sum_{k=-\infty}^{\infty} W_{f,g}(k,\theta) \ e^{-j2(k+1/2)\psi} \qquad (2.24)$$

Equation (2.23) can be proved by substituting the definition of $W_{F,G}(\theta,k)$ into (2.23). Now,

$$E \ = \ \sum_{k=-\infty}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\theta+\Sigma) \ G^*(\theta-\Sigma) \ e^{j2(k+1/2)\Sigma} \ d\Sigma \quad e^{-j2(k+1/2)\psi}$$

Reordering the summation and integral and simplifying the expression we have,

$$= \int_{-\pi}^{\pi} \left[\ F(\theta+\Sigma) \ G^*(\theta-\Sigma) \ e^{-j(\psi-\Sigma)} \right.$$

$$\left. \cdot \ \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-j2k(\psi-\Sigma)} \ \right] \ d\Sigma$$

Defining $2(\psi-\Sigma) = \lambda$, we have

$$E \ = \ \int_{2\psi-2\pi}^{2\psi+2\pi} \left[\ F(\theta+\psi-\lambda/2) \ G^*(\theta-\psi+\lambda/2) \ e^{-j\lambda/2} \right.$$

$$\left. \cdot \ [\ \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-jk\lambda} \ ] \ (1/2) \ \right] \ d\lambda$$

Using (2.15), we have

$$E = (1/2) \int_{2\Psi-2\pi}^{2\Psi+2\pi} F(\theta+\Psi-\lambda/2) \ G^*(\theta-\Psi+\lambda/2) \ e^{-j\lambda/2}$$

$$\cdot \sum_{n=-\infty}^{\infty} \delta(\lambda-2n\pi) \ d\lambda$$

For any value of Ψ, there are only two values of λ for which $\lambda-2n\pi$ will be zero. For example, if $\Psi=\pi$, λ ranges from 3π to $-\pi$; thus only n=0, and n=1 effect the summation. Using the sifting property of the delta function and the periodicity of $F(\theta)$ and $G(\theta)$, E can be expressed as

$$E = 1/2 \ [\ F(\theta+\Psi) \ G^*(\theta-\Psi) \ e^{-j0} + F(\theta+\Psi+\pi) \ G^*(\theta-\Psi-\pi) \ e^{-j\pi} \ ]$$

$$E = 1/2 \ [\ F(\theta+\Psi) \ G^*(\theta-\Psi) \ - \ F(\theta+\Psi+\pi) \ G^*(\theta-\Psi-\pi) \ ]$$

and (2.23) now follows.

An interesting result follows from (2.24) when $\Psi=0$, which is

$$F(\theta) \ G^*(\theta) = \sum_{k=-\infty}^{\infty} W_{f,g}(k,\theta)$$

where we have assumed that $F(\theta)$ and $G(\theta)$ are band limited to $\pm(\pi/2)f_s$.

5. The Effect of Windowing on the MCWD

It is impossible to numerically calculate the MCWD for non-time limited signals as the MCWD contains an infinite summation. To estimate the MCWD at some time n, as in the MAWD case, it is necessary to window the sequence about the time n. The question is how does the windowing effect our estimate of the MCWD?

Assume we have two functions, f(n) and g(n), which are windowed by two finite windowing functions, w1(n) and w2(n), respectively. Let h(n) = f(n) w1(n) and d(n) = g(n) w2(n), then the MCWD of h(n) and d(n) can be expressed as

$$W_{h,d}(n,\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_{f,g}(n,\Sigma)W_{w1,w2}(n,\theta-\Sigma) \, d\Sigma \underline{\underline{=}} F \qquad (2.25)$$

which can be proved by expressing $W_{f,g}(n,\Sigma)$ and $W_{w1,w2}(n,\theta-\Sigma)$ in terms of their time functions. Now

$$F = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\sum_{k=-\infty}^{\infty} f(n+k+1) \, g^*(n-k) \, e^{-j2(k+1/2)\Sigma} \right.$$

$$\left. \cdot \sum_{m=-\infty}^{\infty} w1(n+m+1) \, w2^*(n-m) \, e^{-j2(m+1/2)(\theta-\Sigma)} \right] \, d\Sigma$$

Rearranging the integral and summations and simplifying, we have

$$F = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f(n+k+1) \, g^*(n-m) \, w1(n+k+1) \, w2^*(n-m)$$

$$\cdot e^{-j2(m+1/2)(\theta)} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j2(k-m)\Sigma} \, d\Sigma$$

The integral is zero, except for k-m=0, when it is equal to 2π. This property reduces the double summation into one giving

$$F = \sum_{k=-\infty}^{\infty} f(n+k+1) \, w1(n+k+1) \, g^*(n-k) \, w2^*(n-k) \, e^{-j2(k+1/2)\theta}$$

Replacing $f(n) \, w1(n)$ by $h(n)$ and $g^*(n) \, w2^*(n)$ by $d^*(n)$ gives

$$F = \sum_{k=-\infty}^{\infty} h(n+k+1) \, d^*(n-k) \, e^{-j2(k+1/2)\theta}$$

which is by definition the $W_{h,d}(n,\theta)$. This completes the proof.

2.4 A Discrete Version of the MAWD

In this section, we define the discrete MAWD (DMAWD). The DMAWD is obtained by substituting $\theta = (m+1/2)\Sigma$, where $\Sigma = \pi/2N$ into the definition of the MAWD and then truncate the summation. This is given by

$$W_f(n,m) = \sum_{k=-N}^{N-1} f(n+k+1) \; f^*(n-k) \; e^{-j2(k+1/2)(m+1/2)\mathcal{I}} \qquad (2.26)$$

where $\mathcal{I} = \pi/2N$. Note that this definition can also be considered as a sliding rectangular window ($w(n)$) applied to the function $f(n)$,

$$w(k) \quad = \quad \begin{cases} 1 & n-N < k \le n+N \\ 0 & \text{otherwise.} \end{cases}$$

2.4.1 Properties of the DMAWD

The properties of the DMAWD are very similar to those of the MAWD. A list of some of these properties, with proofs, is included to show the validity of the DMAWD.

1. Inverse Operation

The inverse of the DMAWD is given by

$$f(n+r+1) \; f^*(n-r) = \frac{1}{2N} \sum_{m=-N}^{N-1} W_f(n,m) \; e^{j2(r+1/2)(m+1/2)\mathcal{I}} \; \equiv \; G \quad (2.27)$$

This can be proved by substituting the definition of $W_f(n,k)$ into (2.27) and simplifying. Now,

$$G \; = \; \frac{1}{2N} \sum_{m=-N}^{N-1} \; \sum_{k=-N}^{N-1} \; f(n+k+1) \; f^*(n-k) \; e^{-j2(k+1/2)(m+1/2)\mathcal{I}}$$

$$\cdot \; e^{j2(r+1/2)(m+1/2)\mathcal{I}}$$

Reordering the summations and simplifying the expression we have

$$G \; = \; \sum_{k=-N}^{N-1} \; f(n+k+1) \; f^*(n-k) \; e^{-j(k-r)\mathcal{I}} \; \frac{1}{2N} \sum_{m=-N}^{N-1} \; e^{-j2(k-r)m\mathcal{I}} \qquad (2.28)$$

Considering only the second summation and using $m = i-N$, it follows that

$$\sum_{m=-N}^{N-1} e^{-j2(k-r)m\,\mathcal{I}} = \sum_{i=0}^{2N-1} e^{-j2(k-r)(i-N)\mathcal{I}}$$

$$= e^{-j2N(k-r)\mathcal{I}} \sum_{i=0}^{2N-1} e^{-j2(k-r)i\mathcal{I}}$$

$$= e^{-j2N(k-r)\mathcal{I}} \frac{1-e^{-j2(k-r)2N\mathcal{I}}}{1-e^{-j2(k-r)\mathcal{I}}}$$

Recalling that $\mathcal{I}=\pi/2N$, the summation becomes

$$\sum_{m=-N}^{N-1} e^{-j2(k-r)m\,\mathcal{I}} = \begin{cases} 2N & \text{for } k-r = m2N \\ 0 & \text{otherwise} \end{cases}$$

We use this result to simplify (2.28). For the simple case of $-N \leq r \leq N-1$, $k = r$ only one time. Using this in (2.28), we have

$$G = \frac{1}{2N} (f(n+r+1)\, f^*(n-r))\, 2N = f(n+r+1)\, f^*(n-r)$$

which completes the proof.

2. Real Valued

The DMAWD is always real, as can be characterized by

$$W_f(n,m) = (W_f(n,m))^* \qquad (2.29)$$

This can be proved by expressing $W_f(n,m)$ in terms of two sums, one with $k\geq0$ and another with $k < 0$. That is,

$$W_f(n,m) = \sum_{k=0}^{N-1} f(n+k+1)\, f^*(n-k)\, e^{-j2(k+1/2)(m+1/2)\mathcal{I}}$$

$$+ \sum_{k=-N}^{-1} f(n+k+1)\, f^*(n-k)\, e^{-j2(k+1/2)(m+1/2)\mathcal{I}}$$

Using $k = -i-1$ in the second summation, we have

$$W_f(n,m) = \sum_{k=0}^{N-1} f(n+k+1)\, f^*(n-k)\, e^{-j2(k+1/2)(m+1/2)\mathcal{I}}$$

$$+ \sum_{i=0}^{N-1} f(n+(-i-1)+1)\, f^*(n-(-i-1))\, e^{-j2(-i-1+1/2)(m+1/2)\mathcal{I}}$$

Now the two sums can be written as one, giving

$$W_f(n,m) = \sum_{k=0}^{N-1} [\; f(n+k+1) \; f^*(n-k) \; e^{-j2(k+1/2)(m+1/2)\mathcal{I}}$$

$$+ \; f(n-i) \; f^*(n+i+1) \; e^{j2(i+1/2)(m+1/2)\mathcal{I}} \;]$$

The term inside the brackets is the sum of complex conjugates and is strictly real. Therefore,

$$W_f(n,m) = 2 \sum_{k=0}^{N-1} \text{Re}[\; f(n+k+1) \; f^*(n-k) \; e^{-j2(k+1/2)(m+1/2)\mathcal{I}} \;]$$

which completes the proof.

3. Symmetry and Periodicity in the Frequency Domain

The symmetry in the frequency domain of the DMAWD can be character-ized by

$$W_f(n,m) = W_{f*}(n,-m-1) \qquad\qquad (2.30)$$

This can be proved by expressing,

$$W_{f*}(n,-m-1) = \sum_{k=-N}^{N-1} f^*(n+k+1)f(n-k) \; e^{-j2(k+1/2)(-m-1+1/2)\mathcal{I}}$$

$$= \sum_{k=-N}^{N-1} f^*(n+k+1)f(n-k) \; e^{-j2(k+1/2)(-m-1+1/2)\mathcal{I}}$$

$$= [\; \sum_{k=-N}^{N-1} f(n+k+1)f^*(n-k) \; e^{-j2(k+1/2)(m+1/2)\mathcal{I}} \;]$$

$$= [\; W_f(n,m) \;]^*$$

Recalling (2.29), (2.30) now follows.

The periodicity of the DMAWD can be expressed in terms of

$$(-1)^i \; W_f(n,m) = W_f(n,m+2Ni) \qquad\qquad (2.31)$$

This can be proved by expressing

$$W_f(n, m+2Ni) = \sum_{k=-N}^{N-1} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)(m+2Ni+1/2)\mathcal{T}}$$

$$= \sum_{k=-N}^{N-1} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)(m+1/2)\mathcal{T}}$$

$$\cdot \; e^{-j4Ni(k+1/2)\mathcal{T}}$$

$$= \sum_{k=-N}^{N-1} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)(m+1/2)\mathcal{T}} \; e^{-j2\pi ik} \; e^{-j\pi i}$$

Recalling that $e^{-j2\pi}$ equals 1 and $e^{-j\pi}$ equals -1, we have

$$W_f(n, m+2Ni) = (-1)^i \sum_{k=-N}^{N-1} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)(m+1/2)\mathcal{T}}$$

Recalling the definition of $W_f(n,m)$, (2.31) now follows.

As with the case of the MAWD, the DMAWD's magnitude is periodic with period πf_s.

4. Relationship to the DFT

The DFT, at some time n, can be defined as

$$F_n(m) = \sum_{k=n-N+1}^{n+N} f(k) \, e^{-jkm\mathcal{T}} \tag{2.32}$$

where $\mathcal{T} = \pi/2N$ and $-2N \leq m \leq 2N-1$. Equation (2.32) is a modified version of the standard definition (See Oppenhiem and Schafer [52]). However, (2.32) can be computed by first padding the sequence with N zeros at the beginning and end, computing the classical DFT for 4n points and then multiplying by $e^{-jm\pi}$. The preceding operation produces $F_n(m)$ for m = 0 to 4N-1, which is the same as m = -2N to 2N-1, since $F_n(m)$ is periodic. Based on this definition of the DFT, we can now define the DMAWD for the DFT of a signal as

$$W_{F_n}(m, i) = \frac{1}{4N} \sum_{r=-2N}^{2N-1} F_n(m+r+1) \, F_n^*(m-r) \, e^{j2(i+1/2)(r+1/2)\mathcal{T}} \tag{2.33}$$

Note that the summation is of length 4N, this is a consequence of the
zero padding.

It would be best to give a certain amount of insight into what is
actually being calculated here. For the case of the MAWD, we showed that
the spectrum of f(k) could be recovered, to within a constant, from the
MAWD (See (2.14)). The only difference is, f(k) is zero for k<n-N+1 and
n+N<k. Previously we have written $W_f(n,m)$ and simply assumed that the
window is centered about n.To be specific we now write $W_{f_n}(i,m)$, indi-
cating that n is the center of the window. For the limits on the summa-
tions, we need to consider the following: 1) i>n, 2) i=n and 3) i<n.

$$W_{f_n}(i,m) = \sum_{k=-N}^{N-1} f_n(i+k+1) \, f_n^*(i-k) \, e^{-j2(k+1/2)(m+1/2)\mathcal{I}}$$

Since $f_n(k)$ is assumed to be time limited, the limits on the summation
can be changed. If i is greater than n then the two terms i+k+1 and i-k
are greater than n+N before k goes through its full range. To find the
range of k, set i+k+1 and i-k equal to n+N and solve for the upper and
lower bounds respectively. That is, i+k+1 = n+N and i-k = n+N, which
implies that k = n-i+N-1 and k = -n+i-N. If i is equal to n, then k goes
through its full range and the definition is unchanged. When i is less
than n, i+k+1 and i-k are equal to n-N+1, before k goes through its full
range. To find the proper limits of k, set i+k+1 and i-k equal to n-N+1
and solve for k. That is, i+k+1=n-N+1 and i-k=n-N+1, which implies that
k = n-i-N and k = i-n+N-1. Considering all three cases, the definition
of $W_{f_n}(i,m)$ can be written as

$$
W_{f_n}(i,m) = \begin{cases}
\displaystyle\sum_{k=i-n-N}^{n-i+N-1} f(i+k+1)\, f^*(i-k)\, Y_{2N}^{km} & \text{for } i>n \\[2em]
\displaystyle\sum_{k=-N}^{N-1} f(i+k+1)\, f^*(i-k)\, Y_{2N}^{km} & \text{for } i=n \\[2em]
\displaystyle\sum_{k=i-n-N}^{n-i+N-1} f(i+k+1)\, f^*(i-k)\, Y_{2N}^{km} & \text{for } i>n
\end{cases}
$$

where $Y_{2N}^{km} = e^{-j2\pi km/(2N)}$ The previous definition can be consolidated into the form

$$
W_{f_n}(i,m) = \sum_{h=|i-n|-N}^{N-1-|i-n|} f(i+h+1)\, f^*(i-h)\, e^{-j2(h+1/2)(m+1/2)\mathcal{I}} \tag{2.34}
$$

Using (2.34), we can state the following property.

$$
W_{f_n}(i,m) = W_{F_n}(m,i) \tag{2.35}
$$

Equation (2.35) can be shown by writing out the right hand side and substituting in the definition of $F_n(m)$.

$$
W_{F_n}(m,i) = \frac{1}{4N} \sum_{r=-2N}^{2N-1} \sum_{k=n-N+1}^{n+N} f(k)\, e^{-jk(m+r+1)\mathcal{I}}
$$

$$
\cdot \sum_{h=n-N+1}^{n+N} f^*(h)\, e^{jh(m-r)\mathcal{I}}\; e^{j2(r+1/2)(i+1/2)\mathcal{I}}
$$

Reordering the summations and sorting out the terms containing r, we have

$$
W_{F_n}(m,i) = \sum_{k=n-N+1}^{n+N} \sum_{k=n-N+1}^{n+N} f(k)\, f^*(h)\, e^{-j(km+k+hm-i-1/2)\mathcal{I}}
$$

$$
\cdot \frac{1}{4N} \sum_{r=-2N}^{2N-1} e^{-jr(k+h-2i-1)\mathcal{I}}
$$

Let the summation over r be represented as

$$A(p) = \frac{1}{4N} \sum_{r=-2N}^{2N-1} e^{-jpr\Im}$$

where $p = k+h-2i-1$. Using $r=c-2N$, we have

$$A(p) = \frac{1}{4N} \sum_{c=0}^{4N-1} e^{-jp(c-2N)\Im}$$

$$= \frac{1}{4N} e^{jp2N} \sum_{c=0}^{4N-1} e^{-jpc\Im}$$

which can be expressed in the closed form,

$$A(p) = \frac{1}{4N} e^{jp\Im} \frac{1-e^{-j2\pi p}}{1-e^{-j(p/2N)\pi}}$$

where we have used $\Im=\pi/2N$. The function $A(p)$ is therefore zero, except when p is an integer multiple of 4N, in which case, $A(p)$ is 1. Since $n+2N > k$, h, $i > n-2N+1$, and $p = k+h+2i-1$, then the only integer multiple of 4N that p is be equal to is 0.

Using the function $A(k+h+2i-1)$ in the summation, we can write

$$W_{F_n}(m,i) = \sum_{k=n-N+1}^{n+N} \sum_{h=n-N+1}^{n+N} f(k) \, f^*(h) \, e^{-j(km+k-hm-i-1/2)\Im} \, A(k+h-2i-1)$$

Since $A(k+h-2i-1)$ is zero, except for $k+h-2i-1$ being equal to zero, we can reduce the summations over k and h to one summation over k. We begin solving for the range on the new summation by first solving for k in terms of h, i and n, which yields,

$k = 2i + 1 - h$

The range of i is broken into the three cases of 1) $i>n$, 2) $i=n$, and 3) $i<n$.

For the case of $i > n$, the range of k is set by the upper bound on h. Thus $h = n+N$ implies that $2i+1-n-N$ is the lower bound on k or, in other words $2i+1-n-N < k < n+N$. The DMAWD can now be expressed as

$$W_{F_n}(m,i) = \sum_{k=2i+1-n-N}^{n+N} f(2i+1-k)\, f^*(k)\, e^{-j((2i+1-k)(m+1)\,-km-i-1/2)\mathcal{I}}$$

Let $k = i-h$, which implies $2i+1-k = i+h+1$ and

$$W_{F_n}(m,i) = \sum_{h=i-n-N}^{n-i+N+1} f(i+h+1)\, f^*(i-h)\, e^{-j((i+h+1)(m+1)-(i-h)m-i-1/2))\mathcal{I}}$$

$$W_{F_n}(m,i) = \sum_{h=i-n-N}^{n+N-i+1} f(i+h+1)\, f^*(i-h)\, e^{-j2(h+1/2)(m+1/2)\mathcal{I}}$$

For $i=n$, we can go through a similar argument to establish the bounds to be the same as those for $i>n$. Substituting $i=n$ into the above equation we get

$$W_{F_n}(m,i) = \sum_{h=-N}^{N-1} f(i+h+1)\, f^*(i-h)\, e^{-j2(h+1/2)(m+1/2)\mathcal{I}}$$

For the case of $i < n$, the range of k is set by the lower bound on h. Thus, $h = n-N+1$ implies that $2i+1-n+N-1$ is the upper bound on k or, in other words, $n-N < k < n+N$. The DMAWD can now be expressed as

$$W_{F_n}(m,i) = \sum_{k=i-n-N}^{2i+1-n+N-1} f(2i+1-k)\, f^*(k)\, e^{-j((2i+1-k)(m+1)-km-i-1/2)\mathcal{I}}$$

Using $k=i-h$, which implies $2i+1-k = i+h+1$ and

$$W_{F_n}(m,i) = \sum_{h=n-i-N}^{i-n+N-1} f(i+h+1)\, f^*(i-h)\, e^{-j2(h+1/2)(m+1/2)\mathcal{I}}$$

All three cases can be consolidated into the following form

$$W_{F_n}(m,i) = \sum_{h=|i-n|-N}^{N-1-|i-n|} f(i+h+1)\, f^*(i-h)\, e^{-j2(h+1/2)(m+1/2)\mathcal{I}}$$

The right hand side of the above equation is simply the definition of $W_{f_n}(i,m)$. This establishes (2.35).

To recover the spectrum from $W_{F_n}(i,m)$ or $W_{f_n}(m,i)$, we need an inverse operation for $W_{F_n}(m,i)$. The inverse is given by

$$(1/2) \; [\; F_n(m+p+1) \; F_n^*(m-p) - F_n(m+p+2N+1) \; F_n^*(m-p-2N) \;]$$

$$= \sum_{k=-N}^{N-1} W_{F_n}(k,m) \; e^{-j2(k+1/2)(p+1/2)\mp} \qquad = H \qquad (2.36)$$

Using (2.35) in (2.36), we can express (2.36) in the form

$$H = \sum_{k=-N}^{N-1} W_{f_n}(m,k) \; e^{-j2(k+1/2)(p+1/2)\mp} \qquad\qquad (2.37)$$

Equation (2.37) can be established by substituting the definition of $W_{F_n}(m,k)$ into (2.37) and simplifying, we have

$$H = \sum_{k=-N}^{N-1} \frac{1}{4N} \sum_{r=-2N}^{2N-1} F_n(m+r+1) \; F_n^*(m-r) \; e^{j2(k+1/2)(r+1/2)\mp}$$

$$\cdot \; e^{-j2(k+1/2)(p+1/2)\mp}$$

Reordering the summations and collecting terms containing k, we have

$$H = \sum_{r=-2N}^{2N-1} F_n(m+r+1) \; F_n^*(m-r) \; e^{j(2(r+1/2)(1/2)-2(p+1/2)(1/2))\mp}$$

$$\cdot \; \frac{1}{4N} \sum_{k=-N}^{N-1} e^{j2(r+1/2-p-1/2)k\mp}$$

$$= \sum_{r=-2N}^{2N-1} F_n(m+r+1) \; F_n^*(m-r) \; e^{-j(p-r)\mp} \; \frac{1}{4N} \sum_{k=-N}^{N-1} e^{-j2(p-r)k\mp}$$

Considering the summation over k from the previous equation and using h=p-r, we have

$$A(h) = \frac{1}{4N} \sum_{k=-N}^{N-1} e^{-j2kh\mp}$$

A closed form solution for the function A(h) can be written as

$$A(h) = \frac{1}{4N} \; e^{j2Nh\mp} \; \frac{1-e^{-j2h2N\mp}}{1-e^{-j2h\mp}}$$

Recalling that $\bar{J}=\pi/2N$, we have

$$= \frac{1}{4N} e^{jh\pi} \frac{1-e^{-j2\pi h}}{1-e^{-j(2\pi/2N)h}}$$

Using the arguement as before, A(h) is zero for all h, except when h is an integer multiple of 2N, for which case A(h) is equal to 1/2. Using A(p-r) in the expression for H, we have

$$H = \sum_{r=-2N}^{2N-1} F_n(m+r+1) \; F_n^*(m-r) \; e^{-j(p-r)\bar{J}} \; A(p-r)$$

For $-2N < p < 2N-1$, there are only two values of r for which A(p-r) is non-zero. To determine these values of r for p-r = i2N, the values of p are divided into two ranges. If $0 < p < 2N-1$, then r = p and r=p-2N is in the range of r, making p-r=0 and 2N respectively. Now,

$$H = (1/2)[\quad F_n(m+p+1) \; F_n^*(m-p) \; e^{j(p-p)\bar{J}}$$
$$+ \quad F_n(m+p-2N+1) \; F_n^*(m-p+2N) \; e^{j(p-p-2N)\bar{J}} \quad]$$

If $-2N < p \leq 0$, then r=p and r=p+2N is in the range of r and making p-r = 0 and 2N respectively. Now,

$$H = (1/2)[\quad F_n(m+p+1) \; F_n^*(m-p) \; e^{j(p-p)\bar{J}}$$
$$+ F_n(m+p-2N+1) \; F_n^*(m-p+2N) \; e^{j(p-p-2N)\bar{J}} \quad]$$

The terms in each of the above equations is the product of three periodic functions, with period 4N, and thus we can r=p-2N in the first equation with r=p+2N with no effect on the results. Recalling that $e^{j\pi}$= -1, then the summation reduces to the sum of only two terms, and is given by

$$H = (1/2) [F_n(m+p+1) \; F_n^*(m-p) - F_n(m+p+2N+1) \; F_n^*(m-p-2N)]$$

which establishes the inverse relationship of (2.36) and (2.37).

It should be noted that in standard use, all the terms of $W_{f_n}(i,m)$

would not be computed, rather only the case of n=i would be computed. In other words, the term $W_{f_n}(n,m)$ is computed for all values of n and the cases where the window is not centered at the time indices are ignored.

2.5 A Discrete Version of the MCWD

The DMCWD is defined by

$$W_{f,g}(n,m) = \sum_{k=-N}^{N-1} f(n+k+1) \, g^*(n-k) \, e^{-j2(k+1/2)(m+1/2)\mathcal{I}} \qquad (2.38)$$

where $\mathcal{I}=\pi/2N$.

2.5.1 Properties of the DMCWD

The following gives a list of some interesting properties of the DMCWD. The proofs for the DMCWD are similar to that for the DMAWD and are therefore omitted here.

1. Inverse of the DMCWD is given by

$$f(n+r+1) \, g^*(n-r) = \sum_{m=-N}^{N-1} W_{f,g}(n,m) \, e^{j2(r+1/2)(m+1/2)\mathcal{I}}$$

2. The DMCWD is periodic with respect to its frequency variable and can be characterized as

$$(-1)^i \, W_{f,g}(n,m) = W_{f,g}(n,m+2Ni)$$

3. The relationship of the DMCWD to the DFT of the signals can be stated as

$$W_{f,g}(n,m) = W_{F_n,G_n}(m,n)$$

and

$$(1/2) \, [\, F_n(m+r+1) \, G_n^*(m-r) - F_n(m+r+1+2N) \, G_n^*(m-r-2N) \,] =$$

$$\frac{1}{4N} \sum_{k=-N}^{N-1} W_{F_n,G_n}(m,k) \, e^{j2(k+1/2)(r+1/2)\mathcal{I}}$$

2.6 Computation of the DMAWD

Algorithms for computing the DMAWD and the DMCWD, using the FFT, are now derived.

First,

$$W_f(n,m) = \sum_{k=-N}^{N-1} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)(m+1/2)\mathcal{I}}$$

$$= e^{-jm\mathcal{I}} \, e^{-j(1/2)\mathcal{I}} \sum_{k=-N}^{2N-1} f(n+k+1) \, f^*(n-k) \, e^{-jk\mathcal{I}} \, e^{-j2km\mathcal{I}}$$

Second, using k=i-N, we have

$$W_f(n,m) = e^{-jm\mathcal{I}} \, e^{-j(1/2)\mathcal{I}} \sum_{i=0}^{2N-1} f(n+i-N+1) \, f^*(n-i+N)$$

$$e^{-ji\mathcal{I}} \, e^{jN\mathcal{I}} \, e^{-j2im\mathcal{I}} \, e^{j2Nm\mathcal{I}}$$

$$= e^{-j(m+1/2)\mathcal{I}} \, e^{j2(m+1/2)N\mathcal{I}} \sum_{i=0}^{2N-1} (f(n+i-N+1) \, f^*(n-i+N) \, e^{-ji\mathcal{I}}) \, e^{-j2im\mathcal{I}}$$

Third, the summation over i is basically a 2N point DFT of the sequence $f(n+i-N+1) \, f^*(n-i+N) \, e^{-ji\mathcal{I}}$, i = 0,1,...,2N-1. If 2N is a power of 2, then the FFT can be used to compute it.

Clearly, the above approach can be applied to the DMCWD. Thus, we can write

$$W_{f,g}(n,m)$$

$$= e^{-j(m+1/2)\mathcal{I}} \, e^{j2(m+1/2)N\mathcal{I}} \sum_{i=0}^{2N-1} (f(n+i-N+1) \, g^*(n-i+N) \, e^{-ji\mathcal{I}}) \, e^{-j2im\mathcal{I}}$$

The summation is a 2N point DFT. Again, if 2N is a power of 2, then the FFT can be used to compute it.

The definition of the WD, as given by Claasen and Mecklenbrauker (See (1.5) and is given below for easy reference), can also be imple-

mented in a similar fashion. Now,

$$W_f(n,\theta) = \sum_{k=-N}^{N} f(n+k) \, f^*(n-k) \, e^{-j2k\theta} \qquad (2.39)$$

where we have assumed that the signal has been windowed. Using k=i-N, we have

$$W_f(n,\theta) = \sum_{k=0}^{2N} f(n+i-N) \, f^*(n-i+N) \, e^{-j2k\theta}$$

substituting $\theta = m\bar{x}$, with $\bar{x} = \pi/(2N+1)$, we have

$$W_f(n,m) = \sum_{i=0}^{2N} f(n+i-N) \, f^*(n-i+N) \, e^{-j2im\bar{x}} \, e^{j2Nm\bar{x}}$$

$$= (-1)^m \, e^{-jm\bar{x}} \sum_{i=0}^{2N} f(n+i-N) \, f^*(n-i+N) \, e^{-j2im\bar{x}}$$

The summation is now a 2N+1 point DFT. The length of the DFT, being odd, makes it less efficient to compute.

From the previous derivation, we can see that the MAWD is some what more efficient to compute. The real point is that the MAWD is a valid discrete version of the WD, has some useful properties, and has an efficient algorithm for computation.

2.7 The Bilinear Nature of the MAWD

In Wigner's paper, where he introduced the Wigner Distribution, he stated that there are some problems that are intrinsic to the distribution. The basic source of many of these problems arise due to the bilinear nature of the Wigner Distribution. This nature is best described by considering the MAWD of a sequence s(n), where s(n) can be expressed as the sum of two sequences f(n) and g(n). We begin by writing the definition of the MAWD of s(n).

$$W_s(n, \ominus) = \sum_{k=-\infty}^{\infty} s(n+k+1) \, s^*(n-k) \, e^{-j2(k+1/2)\theta}$$

Substituting $f(n)+g(n)$ for $s(n)$ yields

$$W_s(n, \theta) = \sum_{k=-\infty}^{\infty} [\, f(n+k+1)+g(n+k+1) \,] \, [\, f(n-k)+g(n-k) \,]^* \, e^{-j2(k+1/2)\theta}$$

$$= \sum_{k=-\infty}^{\infty} [\, f(n+k+1) \, f^*(n-k) + f(n+k+1) \, g^*(n-k) + g(n+k+1) \, f^*(n-k)$$

$$+ \, g(n+k+1) \, g^*(n-k) \,] \quad e^{-j2(k+1/2)\theta}$$

$$= \sum_{k=-\infty}^{\infty} f(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$+ \sum_{k=-\infty}^{\infty} f(n+k+1) \, g^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$+ \sum_{k=-\infty}^{\infty} g(n+k+1) \, f^*(n-k) \, e^{-j2(k+1/2)\theta}$$

$$+ \sum_{k=-\infty}^{\infty} g(n+k+1) \, g^*(n-k) \quad e^{-j2(k+1/2)\theta}$$

Recalling the definition of $W_f(n,\theta)$ and $W_{f,g}(n,\theta)$, we have

$$W_{f+g}(n,\theta) = W_f(n,\ominus) + W_g(n,\theta) + W_{f,g}(n,\theta) + W_{g,f}(n,\ominus) \qquad (2.40)$$

The Modified Cross Wigner Distributions, $W_{f,g}$ and $W_{g,f}$, in (2.40) are

responsible for the artifacts commonly called cross terms. These cross

terms can cause large erroneous peaks to occur in the WD. It is there-

fore important that we investigate techniques to remove, or at least

reduce these terms, before attempting to use the WD.

The effect of cross terms can be seen from the simple example of a

cosine wave. Let $f(n) = \cos(wn)$, then the MAWD is given by

$$W_f(n,\theta) = \sum_{k=-\infty}^{\infty} \cos(w(n+k+1)) \cos^*(w(n-k)) e^{-j2(k+1/2)\theta}$$

$$= \sum_{k=-\infty}^{\infty} (1/4)(e^{jw(n+k+1)} + e^{-jw(n+k+1)})(e^{-jw(n-k)} + e^{jw(n-k)})$$

$$\cdot e^{-j(2k+1)\theta}$$

$$= \sum_{k=-\infty}^{\infty} (1/4) e^{jw(n+k+1)} e^{-jw(n-k)} e^{-j(2k+1)\theta}$$

$$+ \sum_{k=-\infty}^{\infty} (1/4) e^{jw(n+k+1)} e^{jw(n-k)} e^{-j(2k+1)\theta}$$

$$+ \sum_{k=-\infty}^{\infty} (1/4) e^{-jw(n+k+1)} e^{-jw(n-k)} e^{-j(2k+1)\theta}$$

$$+ \sum_{k=-\infty}^{\infty} (1/4) e^{-jw(n+k+1)} e^{jw(n-k)} e^{-j(2k+1)\theta}$$

$$= \sum_{k=-\infty}^{\infty} (1/4) e^{j(2k+1)(w-\theta)}$$

$$+ (1/4) e^{j(2n+1)w} \sum_{k=-\infty}^{\infty} e^{-j(2k+1)\theta}$$

$$+ (1/4) e^{-j(2n+1)w} \sum_{k=\infty}^{\infty} e^{-j(2k+1)\theta}$$

$$+ \sum_{k=-\infty}^{\infty} (1/4) e^{-j(2k+1)(w-\theta)}$$

Noting Equation (2.15), we have

$$\sum_{k=-\infty}^{\infty} e^{-j2k\theta} = \sum_{k=-\infty}^{\infty} \delta(\theta - \pi k)$$

Now,

$$W_f(n,\theta) = (1/4) \sum_{i=-\infty}^{\infty} [\, \delta(w-\theta-i\pi) + \delta(\theta+i\pi-w) + 2\,\delta(\theta+i\pi)\cos((2n+1)w)\,]$$

Note that the contribution of the cross terms, i.e. the components at θ=iπ, have a peak magnitude twice that of the actual components, which are located at θ=±w+iπ (i = ...,-1,0,1,...).

From this simple example, it can be seen that cross terms can produce erroneous artifacts in the time-frequency plane. In the case of the cosine wave, the cross terms are the result of the positive and negative exponentials required to produce the real valued cosine wave. In the next section, the concept of an analytical signal is introduced. The analytical signal does not have any negative frequencies and therefore cross terms are reduced.

2.8 Analytical Signals

An analytical signal is basically any signal, f(t), for which F(θ) = 0 for θ < 0 where F(θ) is the Fourier transform of f(t). This simple restriction on the Fourier transform of the signal has some important implications. The first implication is that the signal is always complex, which is of litte consequence to the MAWD, since it is defined for a complex sequence. However, the absence of negative frequencies reduces the number of cross terms present in the WD. This is caused by the shear reduction in the number of components in the signal.

Another important fact about analytical signals is that if an analytical signal is sampled, it need to only be sampled at half the Nyquist rate to avoid aliasing. In other words, if the signal f(t) is analytical and F(θ) = 0 for θ > f_c then the signal need only be sampled at a sampling frequency of f_c as opposed to $2f_c$ for a real signal. This second implication is extremely important for the WD, since it can be used to reduce aliasing in the WD.

A discrete time signal cannot be analytical in the strict sense, since its spectrum is periodic and therefore cannot be zero for all negative frequencies. For discrete time signals, we require that the Fourier transform be zero for $(2i-1)\pi f_s < \theta < (2i)\pi f_s$, where i is an integer. Using this definition, we can state the following. Given a discrete analytical sequence g(n), there is an analytical signal f(t) such that f(nT) = g(n), where T is the sampling interval (inverse of the sampling frequency). In other words, the sequence g(n) is equivalent to a sampled version of an analytical signal.

Now, it is not likely that an analytical signal is going to be produced and sampled just to relieve the aliasing problems in the WD. However, an analytical signal can be constructed from the real sampled waveform. This can be achieved in two ways. The most common one is to transform the data then construct the transform of the analytical signal using the following formula [6].

$$F_a(\theta) = \begin{cases} 2\,F(\theta) & \text{for } 0 < \theta \\ F(0) & \text{for } 0 = \theta \\ 0 & \text{for } \theta < 0 \end{cases}$$

where $F(\theta)$ is the Fourier transform of the input signal f(n) and $F_a(\theta)$ is the Fourier transform of the analytical signal. In this work, this approach, i.e., the Fourier transform method, is used to produce the analytical form of a signal.

Another way in which the analytical form of f(n) can be produced is to note that for a given real signal $f_r(n)$, the analytical form, $f_a(n)$ is given by

$$f_a(n) = f_r(n) + j \bar{f}_r(n)$$

where $\bar{f}_r(n)$ denotes the Hilbert transform of $f_r(n)$ [6].

Now that we have the analytical form of the signal, consider what the WD of this signal is like. The produced analytical signal has the same sampling rate as the real signal used to produce it. However an analytical signal need only be sampled at half the rate to avoid aliasing. Thus the analytical signal, produced from the real signal, is sampled at twice the required rate and its WD is not be aliased.

The previous statement that the analytical form of the signal is oversampled is true, however this somewhat ignores an important point about the WD of an analytical signal. This point can be brought out by recalling Equation (2.14), which is repeated here.

$$1/2 \; [F(\theta+\Psi) \; F^*(\theta-\Psi)-F(\theta+\Psi+\pi) \; F^*(\theta-\Psi-\pi)] = \sum_{k=-\infty}^{\infty} W_F(\theta,k) \; e^{-j2(k+1/2)\Psi} \quad (2.14)$$

It should be noted that the aliasing that we are trying to remove is not the same type of aliasing as that seen in signals which have been under sampled, for in that case, information has been lost and cannot be recovered. Rather the aliasing to be removed is that caused by components between $(\pi/2)f_s$ and πf_s, which are still with in the allowable range of the sampling theory. This aliasing is the result of the second term on the left hand side of (2.14). Considering the important case of $\Psi=0$, the left hand side of (2.14) reduces to

$$1/2 \; [\; F(\theta) \; F^*(\theta) - F(\theta+\pi) \; F^*(\theta-\pi) \;]$$

Now if $0 \leq \theta < \pi$, then at least one of the terms $\theta+\pi$ or $\theta-\pi$ is in the range $(2i-1)\pi$ to $2i\pi$ and if the signal is analytical, the second term is zero. This means that the output of the transform is valid for

$0 \leq \theta < \pi.$

As an example of how the analytical form of a signal can reduce cross terms and aliasing, we now derive the analytical form of a cosine wave and its MAWD. To find the analytical form of a cosine wave, recall that the Hilbert transform of a cosine is a sine wave. Therefore, if

$f(n) = \cos(wn)$

then, the analytical form of $f(n)$ is given by

$f_a(n) = \cos(wn) + j \sin(wn) = e^{jwn}$

The MAWD of $f_a(n)$ is given by

$$W_{f_a}(n,\theta) = \sum_{k=-\infty}^{\infty} e^{jw(n+k+1)} \; e^{-jw(n-k)} \; e^{-j2(k+1/2)\theta}$$

$$= \sum_{k=-\infty}^{\infty} e^{jw(2k+1)} \; e^{-j(2k+1)\theta}$$

$$= \sum_{k=-\infty}^{\infty} e^{j(w-\theta)} \; e^{j2k(w-\theta)}$$

Recalling (2.41), we have

$$W_{f_a}(n,\theta) = \sum_{k=-\infty}^{\infty} \delta(w-\theta-\pi k) \; (-1)^k$$

From this example, we see that the terms at $\theta=0$, present in the MAWD of the cosine, have been removed.

2.9 Time Filtering

Converting a real signal to its analytical form can alleviate some of the cross terms seen in the WD of a signal. However, cross terms occur between positive frequency components, the same as between positive and negative frequency components. Thus, another technique besides converting the signal to its analytical form is required to combat cross

terms.

Consider the nature of the cross terms. These terms are artifacts that occur in the time-frequency plane of the WD. However, they do not represent energy in the signal. Since the WD averaged over a sufficient amount of time at a given frequency approximates the energy in the signal at that frequency, then the cross terms tend towards zero when averaged. This brings about the following algorithm, proposed by Flandrin [30], for the reduction of cross terms. Flandrin's approach is to compute the terms that make up the ambiguity plane, apply what he called a seperable window to this plane and then compute the Fourier transform along one axis.

The ambiguity plane in Flandrin's work is a two dimensional array made up of the values of $f(n+k) f^*(n-k)$. For the case of the MAWD the ambiguity plane is comprised of the values $f(n+k+1) f^*(n-k)$. The discrete WD and the DMAWD can be computed by windowing their respective ambiguity planes in the direction of the k axis and then computing the Fourier transform along the k axis.

To better understand Flandrin's algorithm, we define the Psuedo MAWD (PMAWD), given by

$$\overline{W}_f(n,\theta) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} Q(n-m,k) \, f(m+k+1) \, f^*(m-k) \, e^{-j2k\theta} \qquad (2.42)$$

where $Q(n,k)$ is an arbitrary two dimensional window function. Note that under the summation in (2.42), all of the terms which make up the ambiguity plane are present. The window, $Q(n,k)$, is what determines the portion of the ambiguity plane that is considered. By choosing $Q(n,k) = t(n)$, with

$$t(n) = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{otherwise} \end{cases}$$

this PMAWD is equal to the MAWD. Similarly, if $Q(n,k) = t(n)\,w(k)$, with

$$w(k) = \begin{cases} 1 & \text{for } -N \leq k \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

then the Psuedo MAWD becomes the estimate of the DMAWD, $W_{f_n}(n,m)$, which we described in Section 2.4. Flandrin proposed a psuedo Wigner distribution, similar to that of (2.41), except he used continuous time signals.

The window $Q(n,k) = t(n)\,w(k)$, is basically the same as the seperable window proposed by Flandrin. The advantage of having a seperable window is that now we can choose windows which provide the appropriate degree of smoothing with respect to both the time and frequency axis. The effect of windows on the MAWD are discussed in Section 2.1. We now consider the effect of windows applied along the time axis of the PMAWD.

Note that in the example of the cosine wave, the time modulation of the cross term is a function of the frequency of the cosine wave. A general theory explaining the character of cross terms is given by Berry [6]. From this characterization of cross terms, it is shown that the frequency of the time modulation for the cross terms is dependent on the frequency seperation of the components generating these terms. Thus to consistently resolve components seperated in the frequency domain by D (1/seconds), time filtering over at least 1/D (seconds) is required.

At this point it would be instructive to stop and reflect on exactly what the last result means. In the case of the STFT, the fre-

quency resolution is dependent upon the length of the transform. Thus,
to gain high resolution in the frequency domain, meant increasing the
length on the transforms computed. Increasing the length of the trans-
form in turn reduced the time resolution. The same trade off now exists
in the WD.

In the previous paragraphs, we pointed out the trade offs associat-
ed with setting the length of the time window for the PMAWD. We have not
discussed the question of what shape this window should be. We are not
going to get too involved with this question, for it has already been
covered by others ([35,52,67]) much more extensively than what we can
hope to include at this time. It should be pointed out that a multitude
of windows exist, with the choice being dependent upon the application.
This is considered a reasonable area for further research.

2.10 Examples of the MAWD Applied to Signals

This section demonstrates the response of the MAWD, to various
types of signals. Three examples are included. The first example, a
cosine wave, is used to exemplify the effect of converting the signal to
its analytical form can have on the MAWD. The second waveform is the sum
of two cosine waves and is intended to demonstrate the use of time
filtering to reduce cross terms. The third signal is a time trace from
the acoustic well log, which demonstrates the complexity of the MAWD. In
the first two examples and in the following discussions a sampling rate
of 1 Hertz is assumed.

The following MAWD's are computed using a 128 point sliding Ham-
ming window, which is padded with 64 leading and trailing zeros. The
padded zeros make the transform 256 points long. The padding with lead-

ing and trailing zeros is required, since the MAWD reverses the sequence about its center point and then cross multiply it with the original. By centering the windowed data before the transform, proper windowing is insured.

In Section 2.7, a cosine wave is used as an example of cross terms and how the analytical form of a signal can be used to reduce cross terms. Thus, the first example chosen is a 0.125 Hz cosine wave. Figure 9 contains a plot of the waveform and Figure 10 is the MAWD of the cosine wave.

From Figure 10 we can see the character of the data returned from the DMAWD. Note the large response at zero and 0.5 Hz, which are the cross terms generated by the positive and negative frequencies. Also of interest is the negative response for the negative frequency component (i.e. at 0.375 Hz). The negative response is to be expected, based on Equation (2.14). In the next figure, Figure 11, we have computed the analytical form of the cosine wave and then plotted the MAWD of the new signal. Hereafter, we refer to the MAWD of the analytical form of a sig- nal as the Analytical MAWD. Note the disappearance of the response at zero. Even more important is the fact that is no repsonse for the upper frequencies. This means that components in the frequencies range of 0.25 to 0.5 Hz can now be resolved, provided we convert to the analyti- cal form first.

The next example is the sum of two cosines at frequencies of 0.15 and 0.35 Hz. Figure 12 contains a plot of the waveform. A plot of the analytical MAWD appears in Figure 13. The important feature that should be noted about Figure 13, is the oscillating components at 0.25 Hz. The oscillating components are the cross terms generated between the two
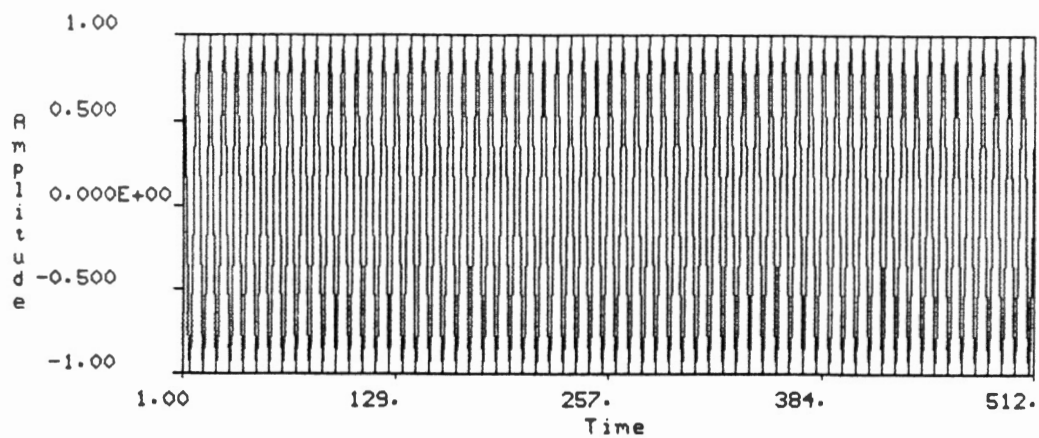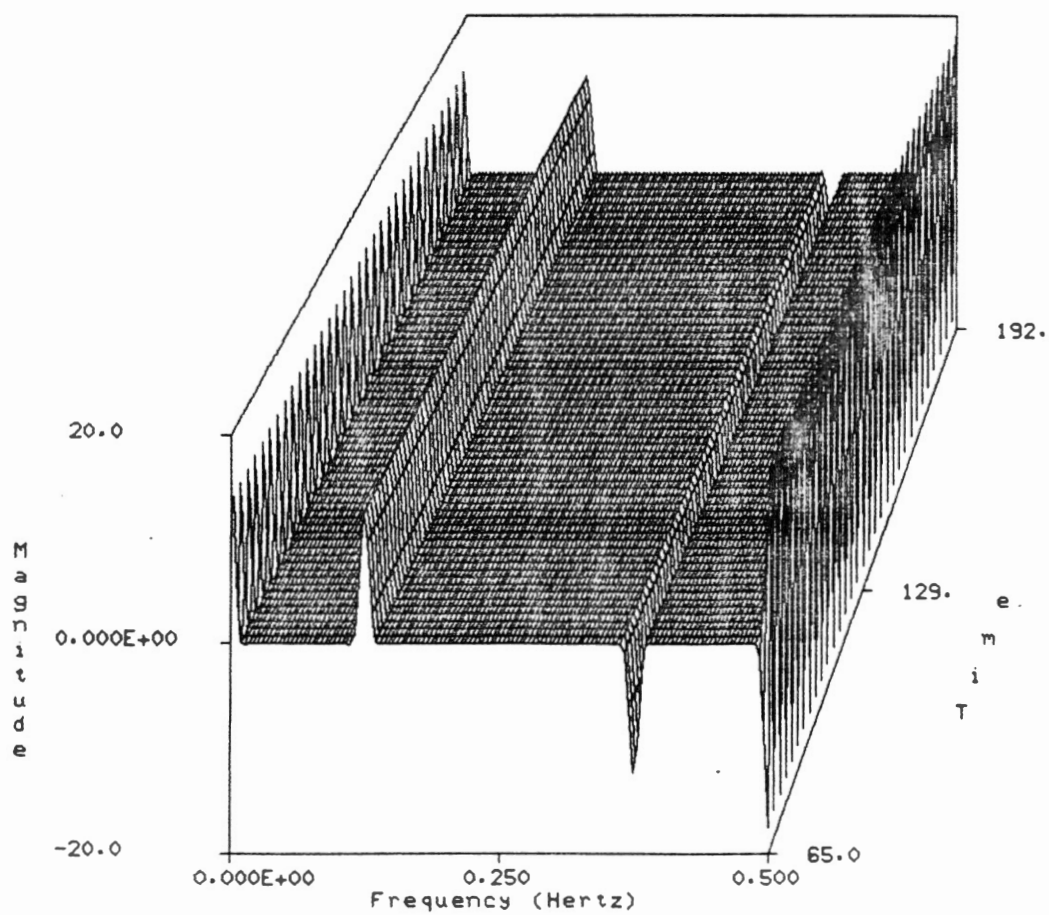
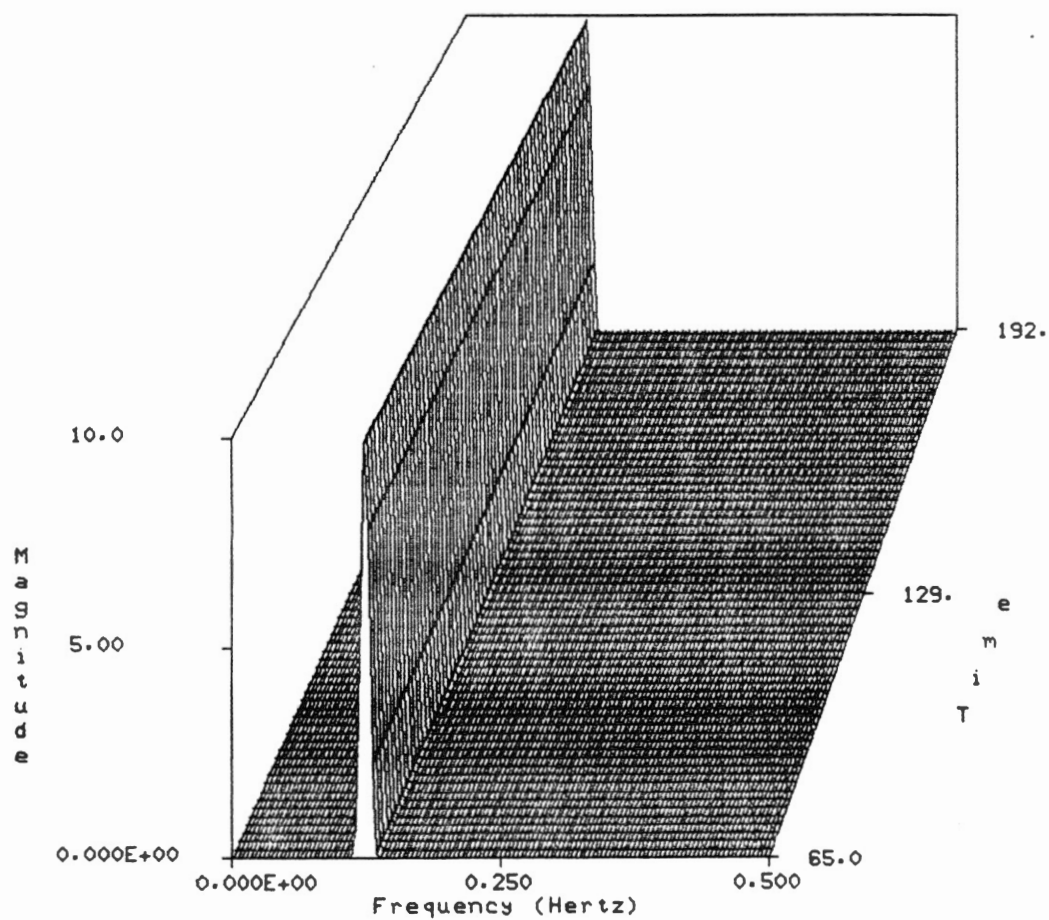Figure 9. Cosine Wave (0.25 Hz)

Figure 10. MAWD of Cosine

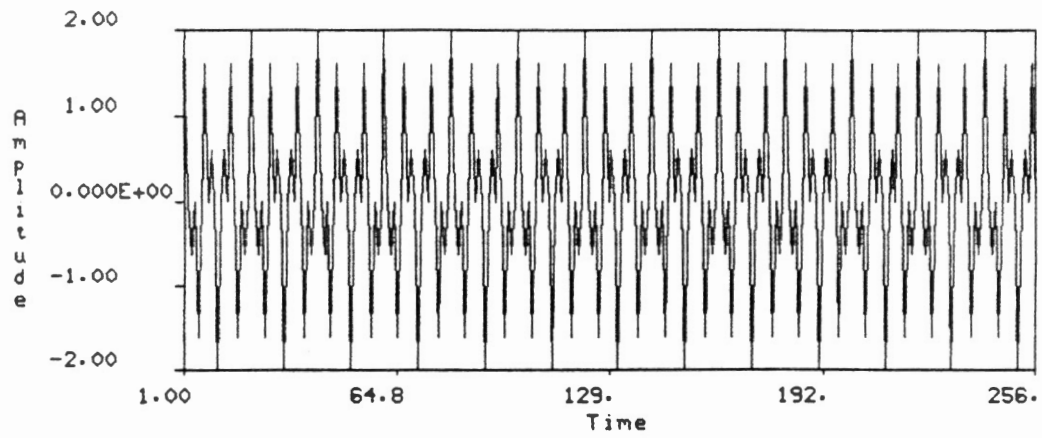Figure 11. MAWD of Analytical Cosine
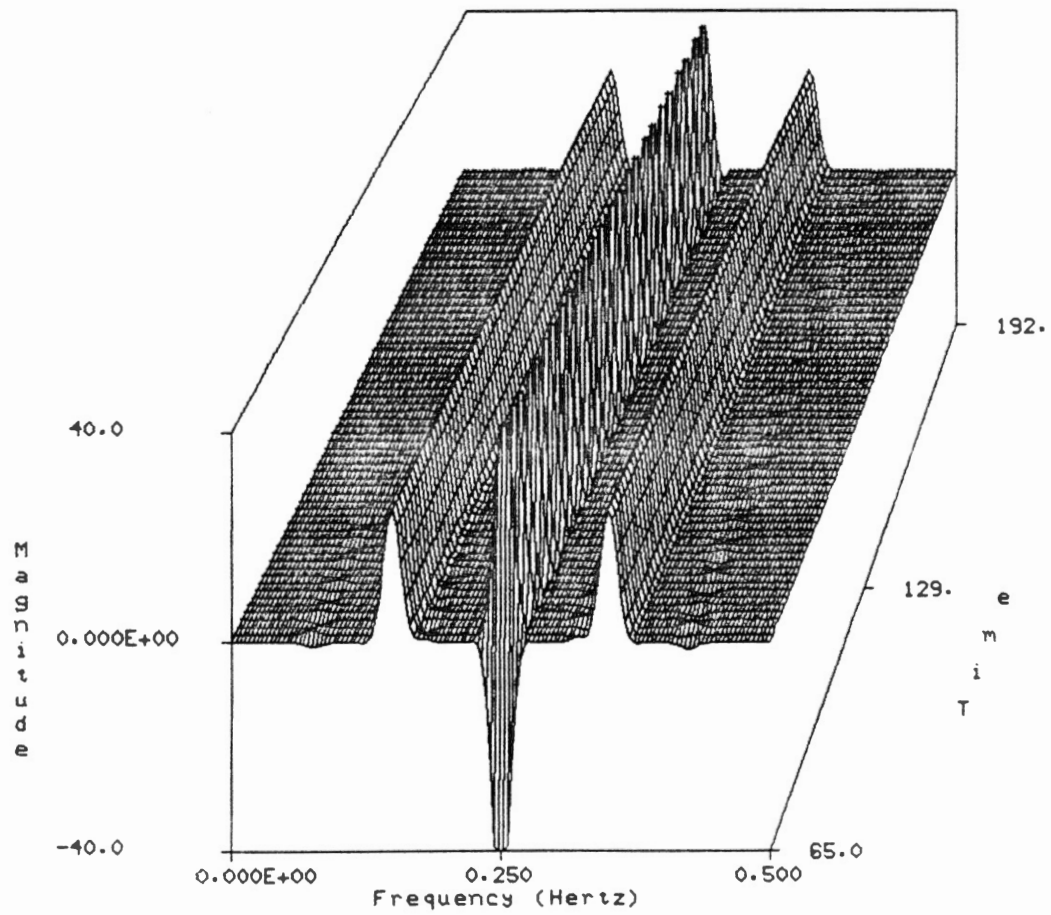
Figure 12. Sum of Two Cosine Waves.

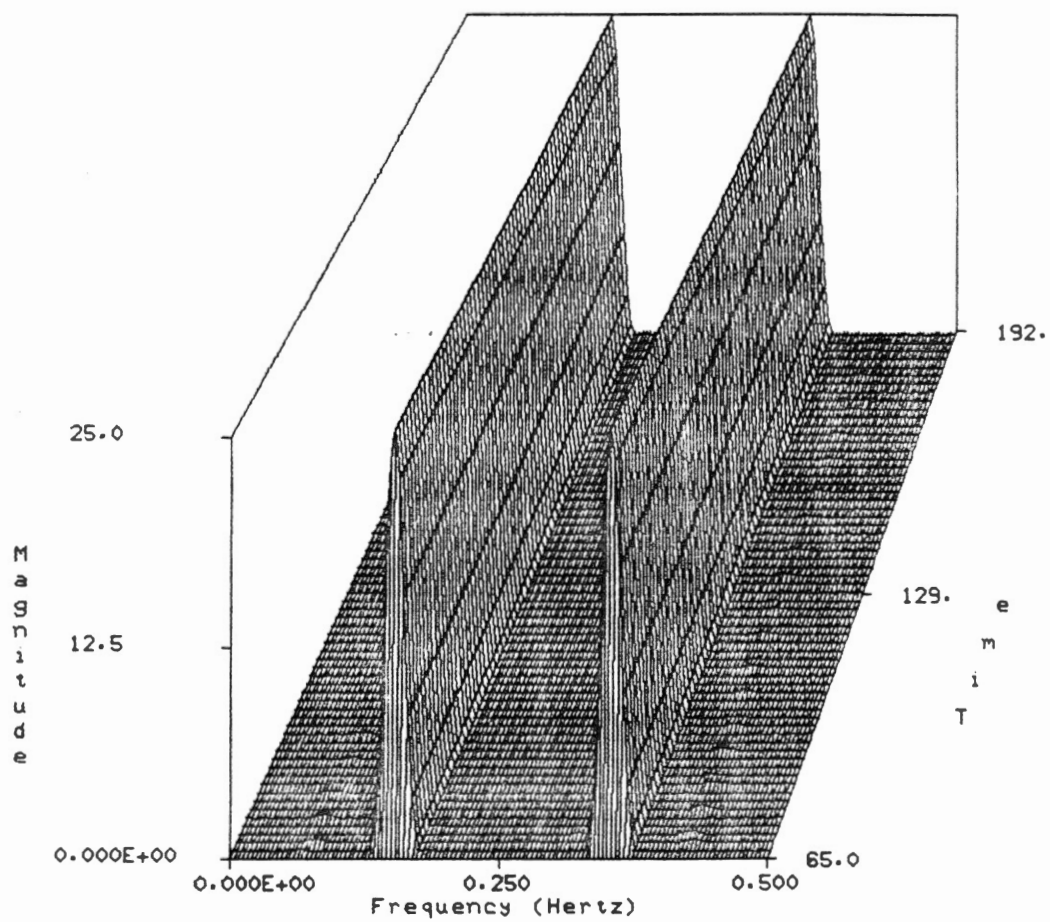Figure 13. Analytical MAWD of Two Cosines

Figure 14. Time Filtered Analytical MAWD of Two Cosines

cosines and requires time filtering to remove.

Figure 14 is a time filtered version of the analytical MAWD in Figure 13. The filter is a mean filter, with its length equal to the samples. The cross-terms are virtually eliminated in this example, however this is some what of a contrived situation in that we know the frequencies of our two components explicitly. Thus, in general, we can not expect quite this dramatic of a decrease in the cross terms.

The third example is an actual trace from an acoustic well log. The 500 point trace is plotted in Figure 15 and the analytical MAWD is displayed in Figure 16. Figure 17 is the time filtered version of the Figure 16. The filter is a sixteen point mean filter, the length of which is chosen somewhat arbitrarily. The primary reason for including this example is to introduce the complexity of the MAWD for a real world signal. We are also concerned with analyzing this type of signal in Chapter 4.

2.11 Conclusions

In this chapter, a discrete version of the WD, which allows for more efficient computation, is introduced. This new definition, called the MAWD, is shown to retain many of the properties of the classical WD. The MCWD, DMAWD and DMCWD are also introduced along with some of the basic properties of each. Problems associated with the bilinear nature of the WD are noted. One common problem noted is the occurance of artifacts, commonly called cross terms, in the time-frequency plane. Two techniques for the reduction of cross-terms are discussed. Examples demonstrating many of the MAWD's properties and problems are included. In Chapter IV, the MAWD is used in the detection and recognition of
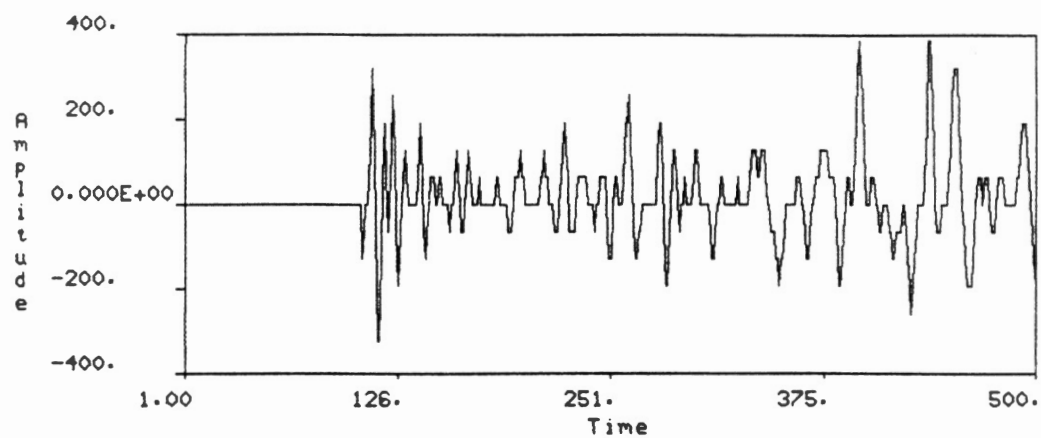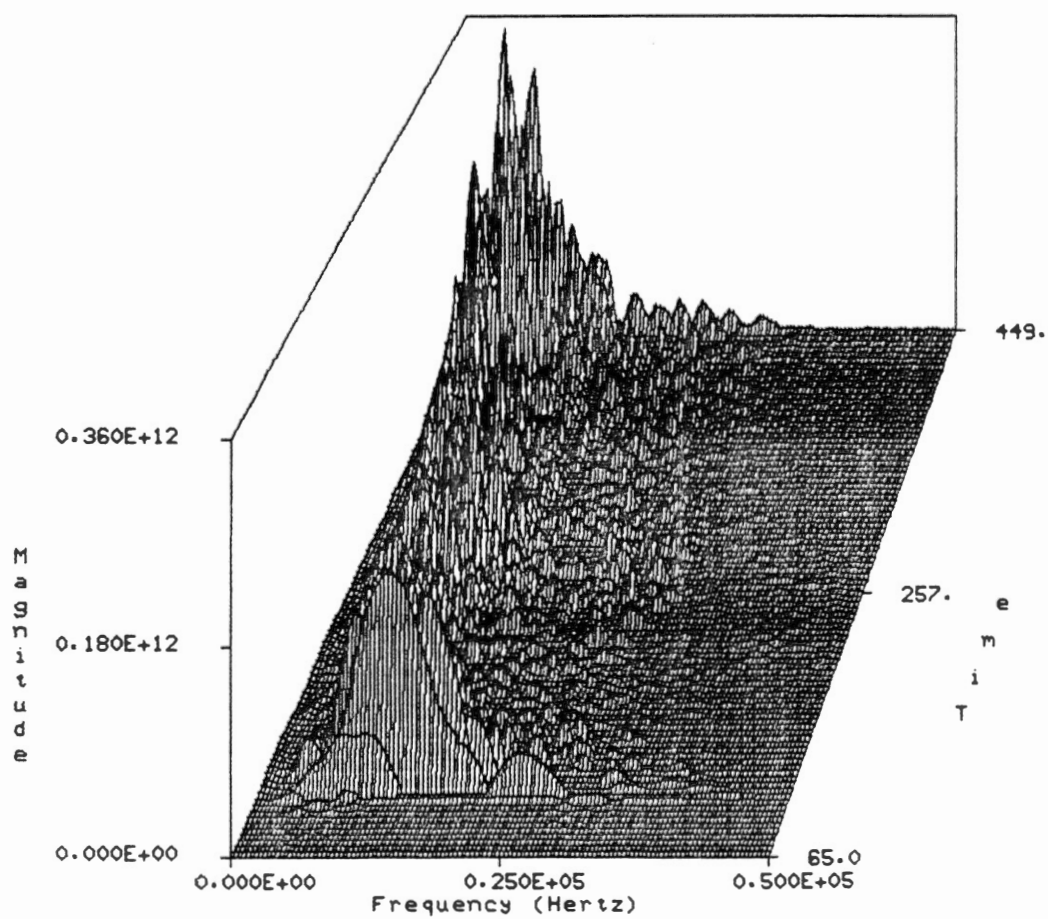
Figure 15. Acoustic Well Log Trace

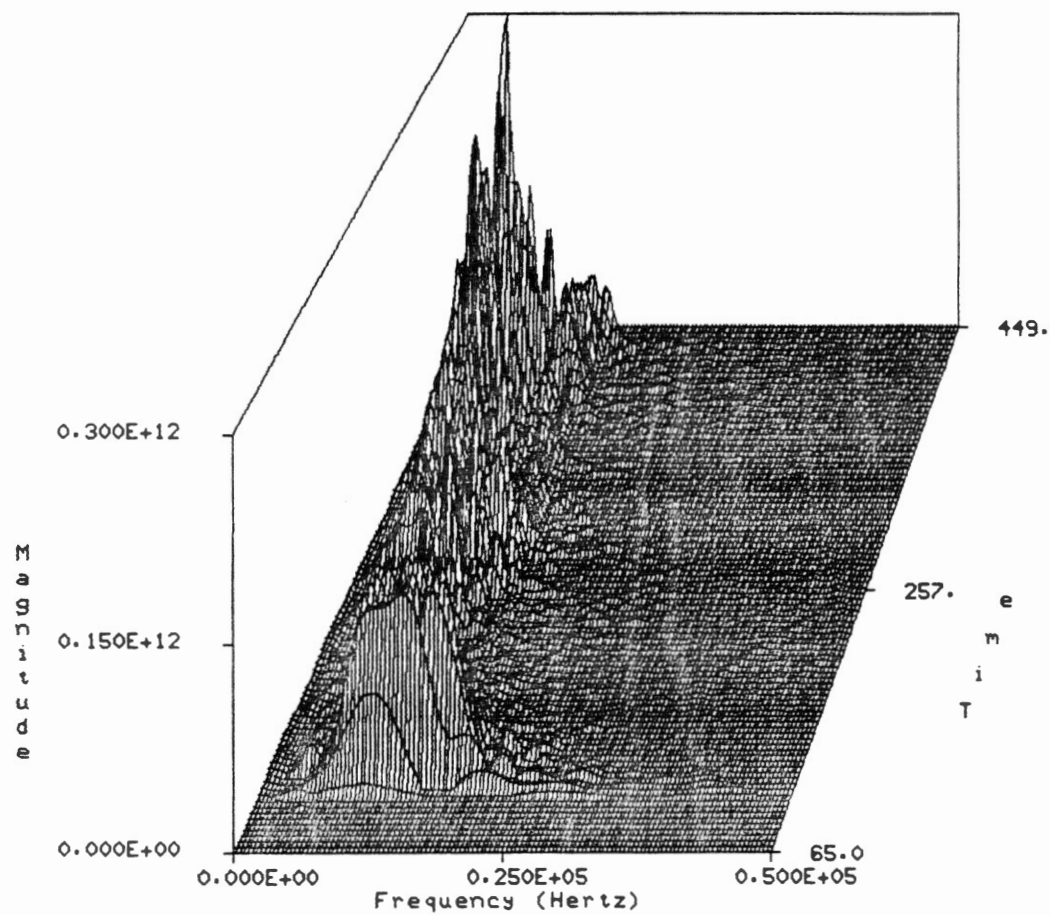Figure 16. Analytical MAWD of Acoustic Trace

Figure 17. Time Filtered Analytical MAWD of Acoustic Trace

shear waves in the time trace of the acoustic well log.

CHAPTER III

THE GENERALIZED ALLAN VARIANCE

3.0 Introduction

The classification of a speech segment as either voiced speech, unvoiced speech or silence is an important step in many speech processing algorithms. In Chapter I, various techniques to perform this classification were described. These techniques are all based on short-time measures applied to the speech waveform. Two of the more common measures are the short-time energy and the zero crossing rate.

The short-time energy, defined quite simply as the short-time average of the square of the speech samples, is given by

$$E_N(i) = \frac{1}{N} \sum_{k=i}^{i+N-1} x(k)^2 \qquad (3.1)$$

The zero crossing rate is given by

$$ZC_N(i) = \frac{1}{2} \sum_{k=i}^{i+N-1} |\,sgn(x(k)) - sgn(x(k-1))\,| \qquad (3.2)$$

$$\text{where } sgn(a) = \begin{cases} 1 & \text{if } a \geq 0 \\ -1 & \text{if } a < 0 \end{cases}$$

These measures are based on simple concepts. The short-time energy, $E_N(i)$, is an approximation of the energy in a signal on a short-time basis and the zero crossing rate, $ZC_N(i)$, is a commonly used psuedo fremeasure [67].

The short-time energy and the zero crossing rate are commonly used together to distinguish speech from silence and voiced speech from unvoiced speech. Voiced speech tends to have large amplitudes and low frequency content, causing the short-time energy to increase and the zero crossing rate to decrease. Unvoiced speech, having lower amplitudes and higher frequency content, tends to have lower short-time energies, and higher zero crossing rates. Silence has much lower energy and erratic zero crossing rates. For the noise free case, the previously mentioned characteristics of the short-time energy and the zero crossing rate can be used to detect and classify speech. However, in the presence of noise, the short-time energy and the zero crossing rate begin to falter.

The Allan variance is a frequency selective energy measure, which can be used to produce a set of measures capable of detecting not only the energy in a signal, but allowing for a coarse mapping of how that energy is distributed in the frequency domain. In this chapter, the character of these measures and how they can be used to classify a segment of the speech waveform as either voiced speech, unvoiced speech or silence is discussed. The Allan variance is also used for the detection of the compressional wavelet in the acoustic well log.

3.1 The Generalized Allan Variance

The Allan Variance (AV), proposed by David Allan [1] for the measurement of frequency stability of atomic oscillators, is given by

$$\sigma(x) = < [x(i) - x(i-1)]^2 > \qquad (3.3)$$

where x is some signal and $< >$ denotes the time average. The Modified Allan Variance (MAV) (proposed by Allan [2]) is given by

$$\sigma_n(x) = < [\ \frac{1}{n}\ \sum_{k=1}^{n}\ (\ x(i) - 2x(i-n) + x(i-2n)\ )\ ]^2 > \qquad (3.4)$$

Note that in the Modified Allan Variance, a parameter (n) is added to the definition. This parameter can be used to adjust the frequency character of the measure.

The signal, analyzed in Allan's work, is actually the average frequency measured from the output of an oscillator using a frequency counter. The AV is therefore a measure applied to the signal x(i) and not strictly a frequency stability measure. The AV can therefore be used to analyze any signal.

Before discussing the nature of the AV or the MAV, we define the Generalized Allan Variance (GAV), as

$$\sigma_{n,m}(x) = < [\ \sum_{k=0}^{m}\ (-1)^k\ \binom{m}{k}\ \frac{1}{n}\ \sum_{h=0}^{n-1}\ x(i-h-k\ n)\ ]^2 > \qquad (3.5)$$

where $\binom{m}{k}$ denotes the binomial coefficients. The Allan variance (AV) and the Modified Allan variance (MAV) can be obtained from the above generalized definition by setting n=1,m=1 and n=1, m=2, respectively.

3.2 The GAV in Terms of the Spectrum

Before using any measure, it is important to obtain a good understanding of that measures character. Now, in the case of the GAV, this understanding is best obtained by examining the GAV's spectral character. We can express the GAV in terms of the spectrum of the input signal, by considering the GAV as the response of a system. Figure 18 schematically describes the process used to derive the GAV.
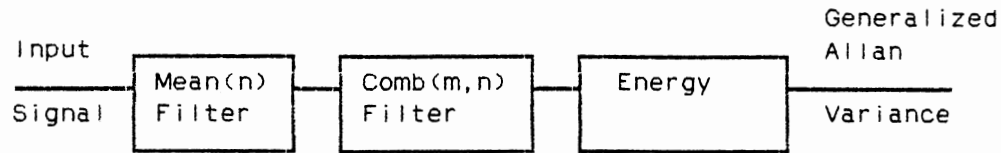
Figure 18. Generation of the GAV

The inner sum in (3.5) can be considered as an n-point mean filter applied to the input signal x and is represented by the first block in Figure 18. The next sum (over k) is a comb filter of differing order and delay dependent on the parameters n and m and is represented by the second block in Figure 18. Finally, the square and time average operator computes the energy for the output of the first two summations.

Let $z_{n,m}(i)$ represent the signal after the first two stages in Figure 18. This signal can be expressed as the convolution of the input, $x(i)$, with the following functions,

$$m_n(i) = \begin{cases} 1/n & \text{for } 0 \leq i \leq n-1 \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

and

$$d_n(i) = \begin{cases} 1 & \text{for } i = 0 \\ -1 & \text{for } i = n \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

Note that the comb filter actually requires m convolutions of the function d_n with the signal to obtain the proper filtering. Thus,

$$z_{n,m}(i) = x(i) * m_n(i) * d_n(i) * \ldots * d_n(i) \tag{3.8}$$

or in the transform domain

$$Z_{n,m}(\theta) = X(\theta) \, M_n(\theta) \, (\, D_n(\theta) \,)^m \tag{3.9}$$

where $Z_{n,m}(\theta)$ is the Fourier transform of the signal $z_{n,m}(i)$,

$X(\theta)$ is the Fourier transform of the input signal $x(i)$,

$M_n(\theta)$ is the Fourier transform of $m_n(i)$ and

$D_n(\theta)$ is the Fourier transform of $d_n(i)$.

The GAV is the energy in the signal $z_{n,m}$. That is,

$$\sigma_{n,m}(x) = \langle \; [\, z_{n,m}(i) \,]^2 \; \rangle \tag{3.10}$$

Applying Parseval's theorem, Equation (3.10) can be expressed as

$$\sigma_{n,m}(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} | \, Z_{n,m}(\mathcal{F}) \, |^2 \; d\mathcal{F} \tag{3.11}$$

Using the following expressions,

$$|M_n(\theta)|^2 = | \; [m_n(i)] \; |^2 = \frac{\sin^2(n\theta/2)}{n^2 \, \sin^2(\theta/2)} \tag{3.12}$$

$$|D_n(\theta)|^2 = | \; [d_n(i)] \; |^2 = \sin^2(n\theta/2) \tag{3.13}$$

and (3.9), we can express

$$|Z_{n,m}(\theta)|^2 = |X(\theta)|^2 \, \frac{\sin^{2m+2}(n\theta/2)}{n^2 \, \sin^2(\theta/2)} \tag{3.14}$$

The GAV can now be expressed as,

$$\sigma_{n,m}(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} | \, X(\theta) \, |^2 \, \frac{\sin^{2m+2}(n\theta/2)}{n^2 \, \sin^2(\theta/2)} \; d\theta \tag{3.15}$$

Equation (3.15) points out that the GAV can be considered to be a frequency selective energy measure, with the frequency band being set by the parameters n and m.

The frequency character of the GAV (and therefore the AV and MAV) is determined by the kernel function in (3.15), which is given in Equa-

tion (3.16). The shape of this function is dependent on the parameters

$$K_{n,m}(\theta) = \frac{\sin^{2m+2}(n\theta/2)}{n^2 \sin^2(\theta/2)} \qquad (3.16)$$

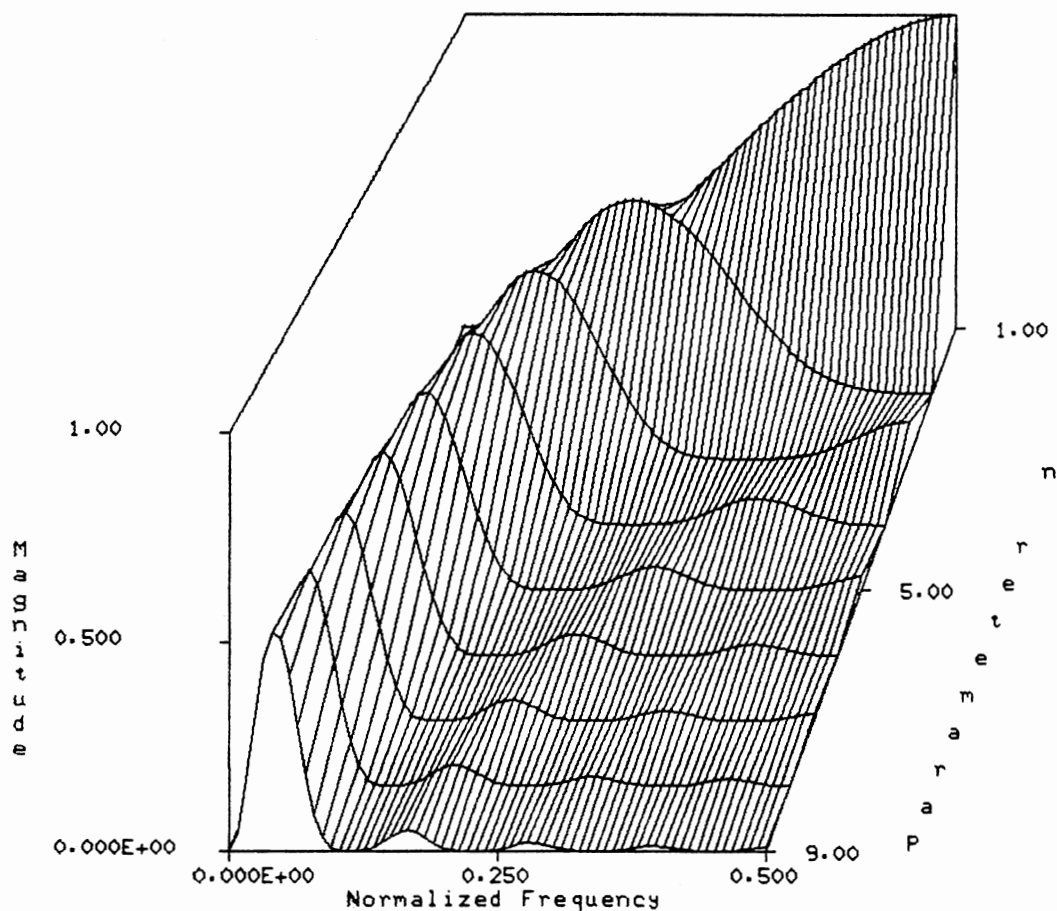n and m. The question is "How is the shape of the kernel effected by n

Figure 19. Kernels for GAV (m=1)

and m?" Surface plots showing the shape of the kernels for n = 1 to 9

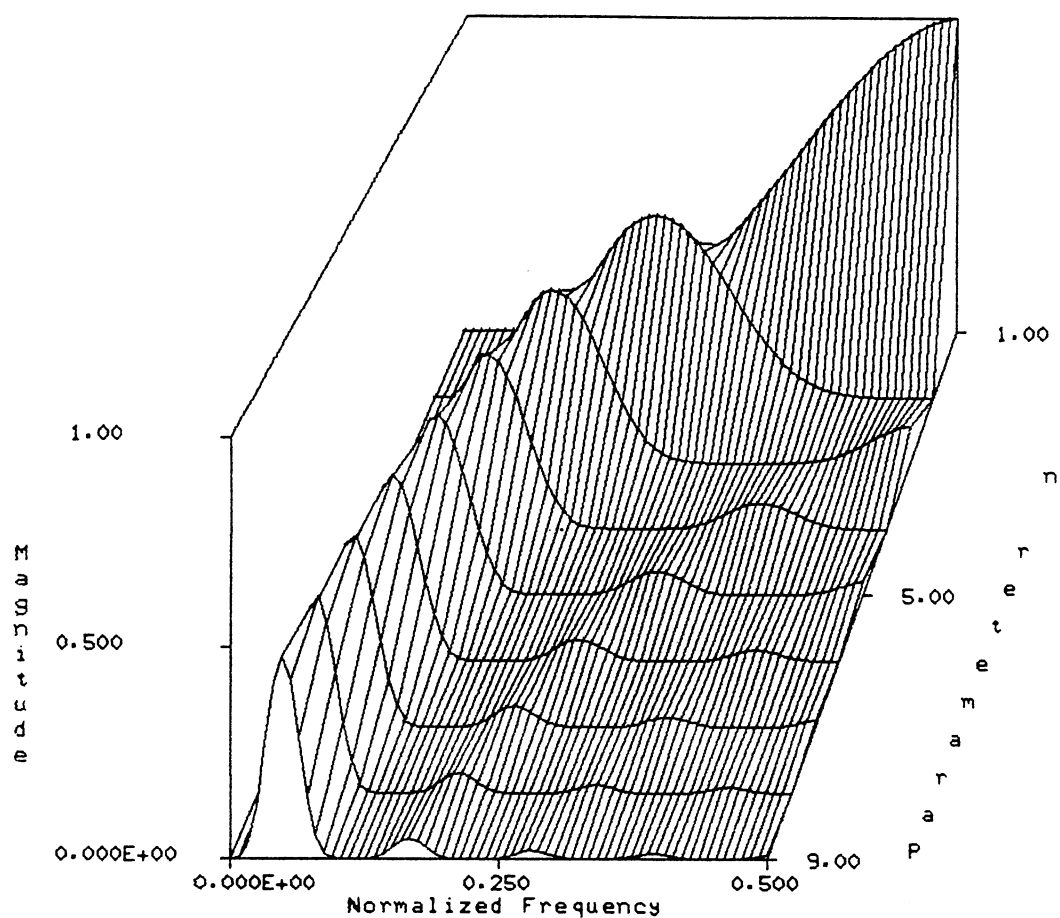are displayed in Figures 19 and 20 for m = 1 and m = 2, respectively.

Figure 20. Kernels of GAV (m=2)

Figures 19 and 20 are actually collections of kernel functions and the actual kernel for a given setting of n and m, is found by slicing the appropriate surface (determined by m) parallel to the frequency axis at the appropriate value of n. The surface plots are used here to aid in the understanding of how the parameter n effects the kernel. A more traditional way might have been to plot them as families of curves. By studying Figures 19 and 20, it can be seen that as n increases the center frequency and the bandwidth decreases. Also, as m increases, the

bandwidth of the kernel decreases. Another important point about these kernels is that they are a restricted set. In other words, the center and the bandwidth of the GAV kernel cannot be chosen arbitrarily, but rather must be chosen from a set of possible kernels. This fact somewhat limits the application of the GAV.

As with many measures, the GAV cannot be computed directly, but can only be estimated. The next section gives ways in which we can estimate the GAV.

3.3 Estimating the GAV

If x(i) is not known analytically, it is impossible to calculate the GAV, because of the time average operator. Thus, we can only estimate the GAV and must therefore concentrate on how this can best be accomplished. A common approximation is to simply replace the time average operator with a short-time average. A better way might be to consider the GAV as the output of the system in Figure 21.

Input ————— Signal │ Mean(n) Filter │—│ Comb(m,n) Filter │—│ Squarer │—│ Low-pass Filter │ Estimated ————— GAV
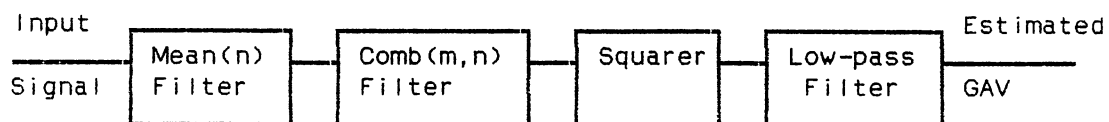
Figure 21. GAV Estimator

The reason for replacing the time average operator with a low-pass filter is that the time average operator can be thought of as an infini-

tesimally narrow band low pass filter. For the case where the time average is replaced with a short-time average, the basic idea is to use a mean filter for the low-pass operation. The next step is the selection of an appropriate filter, having the proper time and frequency response.

We can represent the system in Figure 21 in an equation form, given by

$$\bar{\sigma}_{n,m}(x,h) = \sum_{i=-\infty}^{\infty} W(i-h) \left[\sum_{k=0}^{m} (-1)^k \binom{m}{k} \frac{1}{n} \sum_{p=0}^{n-1} x(i-p-k\,n) \right]^2 \qquad (3.17)$$

Equation (3.17) is very similar to (3.5), except that the time average operator is replaced by a summation over i and a window function. In this way, we can introduce any type of a low-pass filter we choose, by selecting the appropriate window. Another important change is in the introduction of a time variable. Since the Estimated GAV (EGAV) is the output of a filter, the EGAV is a function of time, rather than a single value. The EGAV, therefore, represents the short-time character of the signal. Thus, two important facts which need to be considered in the selection of a window are the window's time response and its frequency response.

Determining the length and the time response of the window to be used in computing the EGAV requires some type of an "a priori" knowledge about how quickly the spectral character of the signal is expected to change. This decision is therefore heavily application dependent. The next two sections discuss applications of the EGAV and describe how the window lengths are chosen. The pass band and stop band attenuations are primarily dependent upon the shape of the window. For illustrative purposes, two windows, namely the rectangular and the Hamming windows are discussed [37].

The rectangular window is given by .

$$W_r(i) = \begin{cases} 1 & \text{for } 0 \leq i \leq N-1 \\ 0 & \text{otherwise} \end{cases} \qquad (3.18)$$

This window is commonly used in theoretical arguements and in applications where minimal computer power is available. Figure 22 is a plot of the frequency response of the rectangular window. The Hamming window, another commonly used window, is given by

$$W_h(i) = \begin{cases} 0.54 - 0.46 * \cos(2\pi i/(N-1)) & \text{for } 0 \leq i \leq N-1 \\ 0 & \text{otherwise} \end{cases} \qquad (3.19)$$

Figure 23 contains a plot of the frequency response of the Hamming window. Note that these plots are for the frequency response of the windows and should not be confused with the Wigner Distribution of the same windows, displayed in Figures 5 and 6.
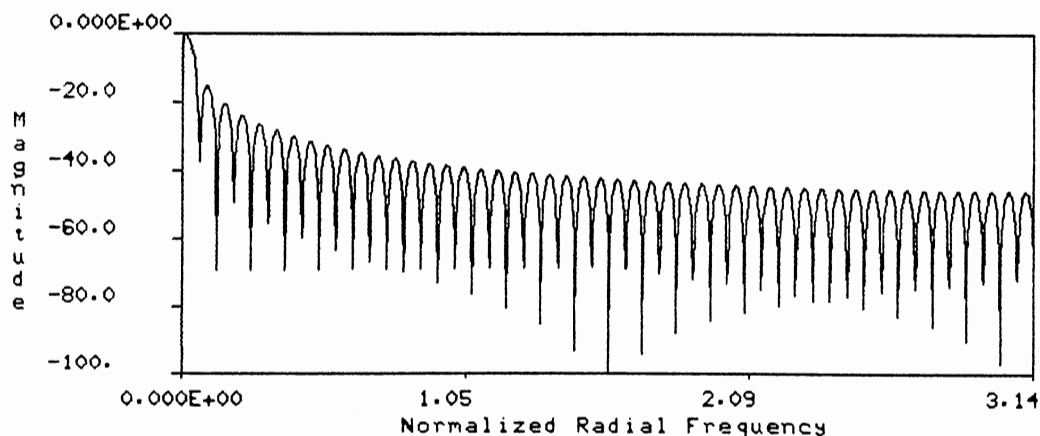

Figure 22. Frequency Response of Rectangular Window

From Figures 22 and 23, we can see that the rectangular window has
a bandwidth of approximately half that of the Hamming window. However,
the attenuation of the Hamming window is considerably better in the stop
band. The bandwidth and attenuations of a window are properties that
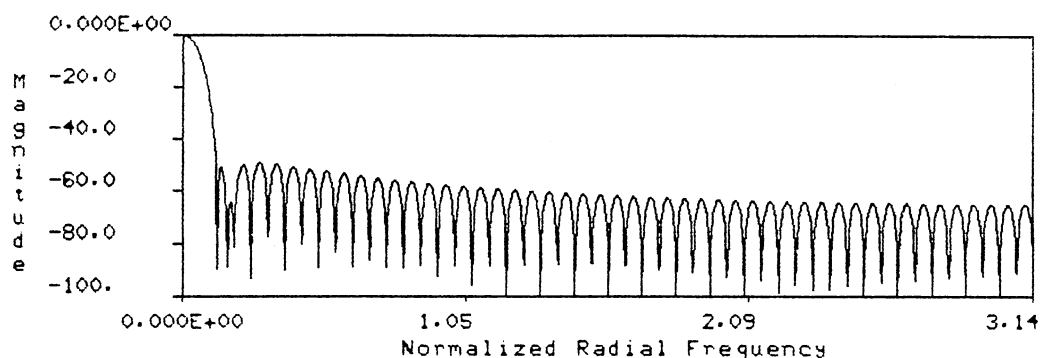must be chosen dependent upon application, and computational power

Figure 23. Frequency Response of Hamming Window

available. In our examples, we are not directly concerned with the com-
putational load. We are also not concerned with an exact design of the
window. Therefore we use a Hamming window when computing the EGAV. For
a more extensive treatment of windows and their properties, the inter-
ested reader should refer to [35] and [67].

In the next few sections, we discuss applications of the GAV. The
first application is in the area of acoustic well logging, where the GAV
is employed in the detection of the compressional wavelet. The second
application is in the area of speech processing for the classification
of speech as voiced, unvoiced, and silence.

3.4 Detection of Signals via the GAV

The first application for which we use the GAV is in the area of
acoustic well logging. In this application, the GAV is used to detect
the arrival of the compressional wave. In the introduction, it is
pointed out that the formation compressional wave is the first wave to
effect the recorded signal and that it is commonly been detected via
thresholding. However, other signals, such as noise, may be present and
can result in false detections. We propose to use the GAV to improve the
detection of the compressional arrival.

In Section 1.3.1, it is stated that a frequency seperation exists
between the compressional and shear wavelets. It is also proposed that
this seperation could be used to identify these waves. For the detection
of the shear wave, the frequency seperation is important, since its
arrival can be masked by the compressional wave. For the compressional
wave, there are no other acoustic waves present to interfere, however,
noise may be present to mask its arrival. To improve the detection of
the compressional wavelet, the EGAV is used to measure the energy in
the frequencies where the compressional wave is dominant. The primary
task is therefore to find a setting for the parameters m and n for the
EGAV, such that the kernel reasonably matches the spectral character of
the compressional wave.

The spectral character of the compressional wave is not generally
predictable, and can be some what dependent on the compressional velo-
city [56]. This brings about the question of how can we set the para-
meters, if we do not know the frequency character of our signal? The
answer is that we can set a range over which the compressional wave may
vary, based on the spectral character of the acoustic source. From

Paillet's simplified model [56], the compressional and shear waves can be thought of as the output of a linear system excited by the acoustic source pulse. In this way, it can be conjectured that the frequency range of the compressional wave must be in the frequency range of the source.

The data for testing our algorithms, consists of a set of real data collected from an oil well in Oklahoma. This data is provided by Amoco Research of Tulsa. An exact spectral characterization of the source is not available. We must therefore estimate the spectrum of the source by computing the spectrum of the full acoustic trace. By using the full trace, we should get the full range of frequencies produces by the source, and therefore the full frequency range of the compressional wave. Now any additional information about the compressional wave's frequency range can be used to improve the selection of a kernel. However, it may prove valuable not to select too narrow a band of frequencies. In this way, we are allowed a certain variance in our estimate of the compressional waves frequency content.

Figure 24 is an example of a real acoustic well log trace and Figure 25 is the spectrum of the full trace. From Figure 25, we can see that the principle frequency components in the acoustic trace range from approximately 1 KHz to 20KHz. Figures 26 and 27 are the kernals for m equal to 1 and n equal to 2 and 3, respectively. By comparing Figures 25, 26 and 27, we can see that the spectrum of the compressional wave is best matched by the kernel for n equal to 3. The application of the GAV to an acoustic pulse is demonstrated in Chapter 4, along with thresholding techniques to detect the arrival of the compressional wave.
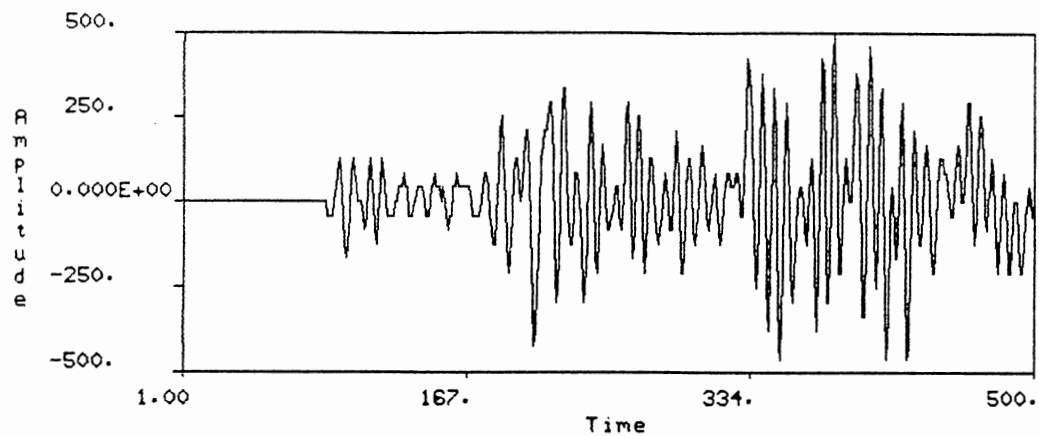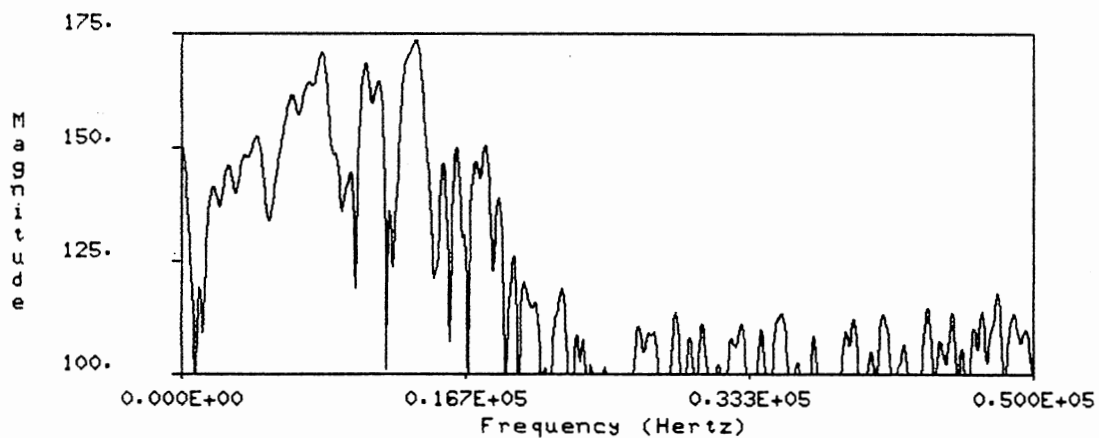
Figure 24. Acoustic Well Log Trace

Figure 25. Spectrum of the Acoustic Well Log Trace

It can be seen from the previous discussion, that the selection of parameters m and n is not an exact science. In fact, the introduction of any additional knowledge or information could effect our choice of parameters. The preceding discussion must therefore be considered as an ad hoc approach of how we can select the parameters. A more exact method

might first estimate the envelop of the spectrum (possibly by averaging the spectra of multiple traces) and then attempt to match this smoothed spectrum with a kernel function using a closeness of fit measure.
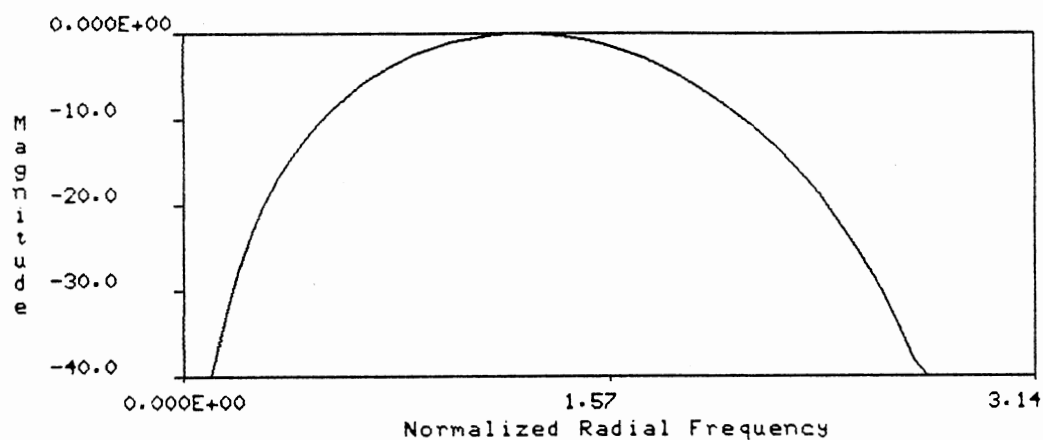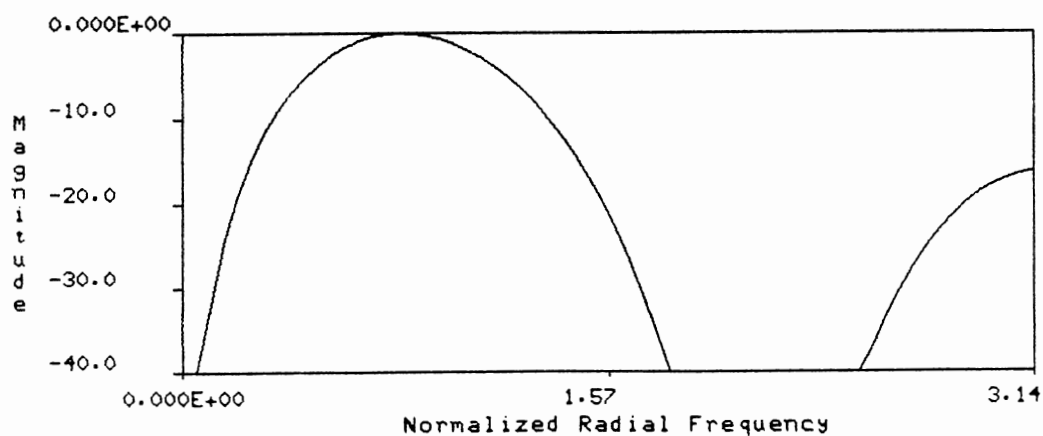
Figure 26. Kernel for m = 1 and n = 2

Figure 27. Kernel for m = 1 and n = 3

3.5 The Two Dimensional Allan Variance

Speech is commonly referred to as a short-time stationary process. The term, short-time stationary, implies that over short intervals of time the spectral character of the waveform is not expected to change significantly, which is basically true within a phoneme. However during the transition from one phoneme to another, extensive changes in the spectral character can occur. This is most pronounced in the transition from a voiced phoneme to an unvoiced phoneme. It is therefore reasonable to consider the use of the GAV for the detection of these different types of speech segments.

The variable spectral nature of speech makes an accurate selection of the parameters m and n rather difficult. One solution to this problem is to fix the parameter m and compute the EGAV for several values of n. The result is an array, which is referred to hereafter as the Two Dimensional Generalized Allan Variance (TDGAV). The TDGAV represents a coarse mapping of the signals energy distribution in the time-frequency plane.

The basic idea of the TDGAV is to measure the energy in different frequency bands and display how they change with respect to time. The interesting aspect about the TDGAV is that the frequency bands from which we can select are some what limited. This limitation is the reason that the TDGAV is referred to as a coarse mapping of the time-frequency plane. To understand the mapping, see Section 3.2 where the character of the kernels is discussed. For example, consider the case of m and n equal to 1. From Figure 19, it can be seen that the kernel ranges from approximately 0.5 Hz to 0.25 Hz (a one Hertz sampling rate is assumed). Thus, the n = 1 column in the TDGAV is the energy in the

upper frequencies. Similarly the n = 2 column represents the energy in approximately the 0.125 to 0.25 Hz frequency range. As we move to larger values of n the center frequency of the kernel moves down and its band-width decreases. This relationship of n to frequency bands is the coarse mapping we referred to earlier.

Before applying the TDGAV to the speech waveform, some approxima-tions need to be made. The first approximation involves the selection of the parameter m. Referring back to Section 3.2 and Figures 19 and 20, it can be seen that as the parameter n increases the kernels begin to overlap each other in frequency. This is not a problem in the previous application, since only one kernel is used. However, now that multiple kernels are used and their outputs are being compared, more independent measures are needed. This can be accomplished by allowing the parameter m to adjust the width of the kernel based on the value of n. In other words, m is set equal to n. Setting n = m for the GAV, we have the Al-ternate Allan Variance (AAV).

The kernels for the AAV can be found from Equation (3.15), since the AAV is only the GAV with m = n. Figure 28 gives a surface plot of the kernels of the AAV for $1 \leq n \leq 7$. It can be seen from Figure 28, that the frequency overlap between the kernels is greatly reduced. For speech analysis, the Two Dimensional Alternate Allan Variance (TDAAV) is used.

It was stated in Chapter I that voiced and unvoiced speech are different in two ways. One way is in terms of the spectral character, and another way is in the energy levels of the speech waveform. Unvoic-ed speech has lower amplitudes than voiced speech and therefore much lower energies. The difference in energy between voiced and unvoiced
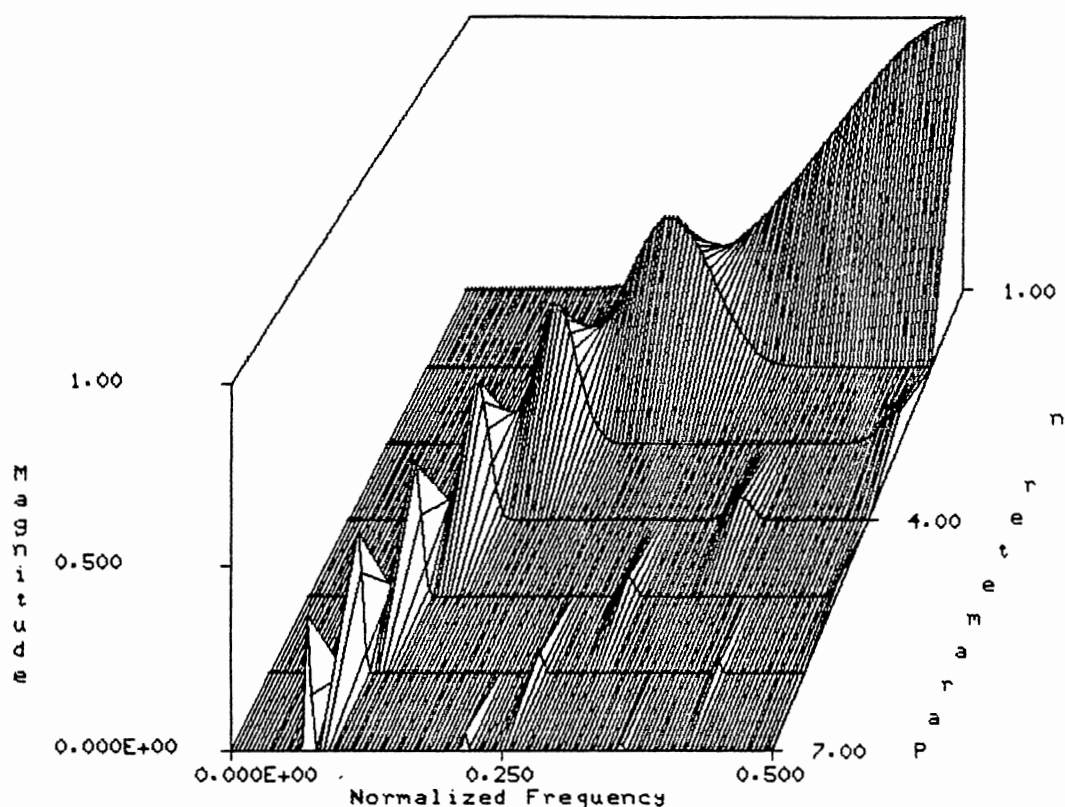
Figure 28. Kernels of the AAV

speech has some interesting implications in the TDAAV. Since the TDAAV

maps the frequency in the range of 2 to 4 KHz (an 8 KHz sampling rate

is assumed for speech) into the n = 1 column and the spectrum of un-

voiced speech lies primarily in that range, the response of the TDAAV

for n = 1 is dominated by unvoiced speech. The TDAAV of n = 2 and above

responds to voiced speech and since the amplitude for voiced speech

tends to be larger than unvoiced, these upper values overshadow the n=1

response. This difference in response can complicate the detection and

recognition of unvoiced speech segments.

The extreme difference in response is due in part to the non-linear squaring operation, which over emphasizes the peak values. The problem of disproportionate response is encountered in the measurement of the short-time energy and can be some what alleviated by replacing the square with an absolute value. The same approximation is therefore made for the TDAAV.

Including the previously mentioned changes, the TDAAV is now given by

$$\bar{\sigma}_n(x,n) = \sum_{i=-\infty}^{\infty} W(i-n) \left| \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{1}{n} \sum_{p=0}^{n-1} x(i-p-k\ n) \right| \qquad (3.20)$$

where $W(n)$ is a window function, similar to those described in Section 3.3. It is hard to determine the exact consequence of these last changes on the TDAAV. The last change is ad hoc in nature and therefore the effect of this change must be judged based on the response of the new TDAAV to speech and upon how well speech can be classified from the TDAAV.

The next section contains examples of the GAV applied to acoustic well log traces and the TDGAV and TDAAV applied to speech waveforms.

3.6 Examples

In this section, the GAV and TDGAV are applied to some example waveforms. The first example is an acoustic well log trace and the second is an acoustic well log corrupted by noise. The GAV is used to detect the arrival of the compressional wave in both of these examples. The third example is a speech sample containing voiced and unvoiced speech. The TDGAV and TDAAV are computed for the speech waveform for various parameter settings.

In Section 3.4, we did a comparison of the spectrum of an acoustic
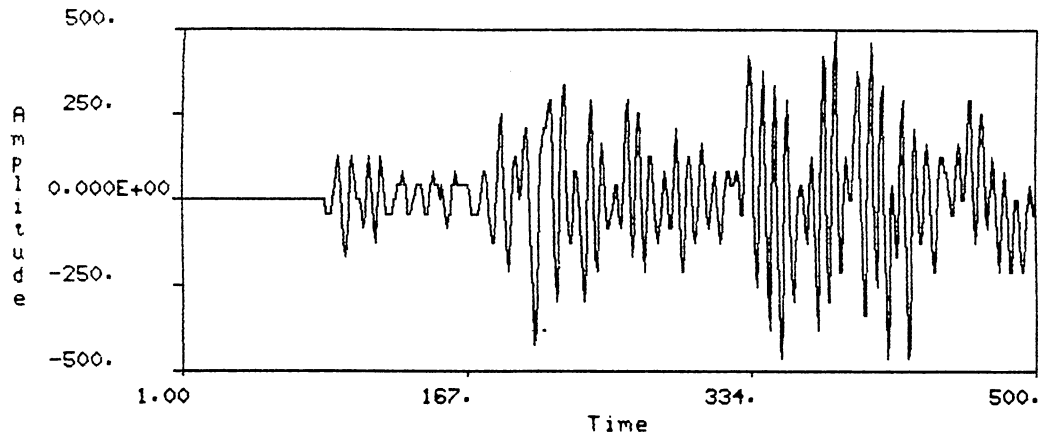

Figure 29. Acoustic Time Trace
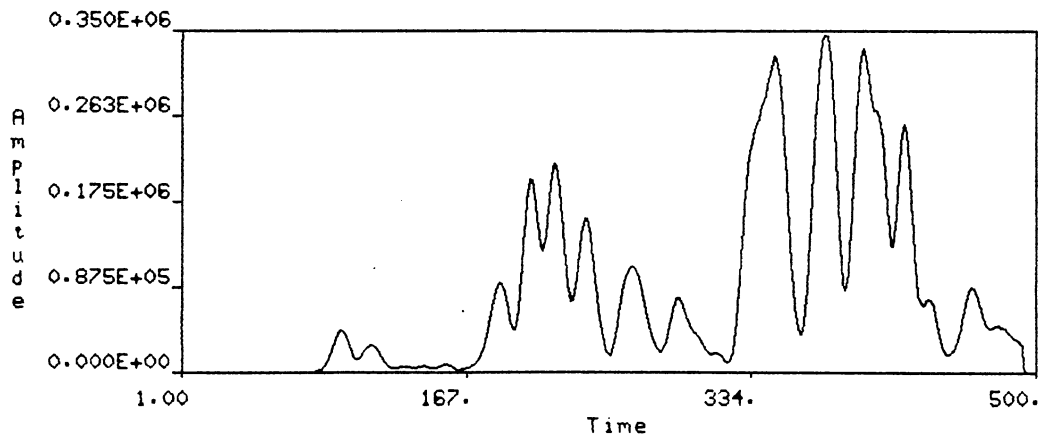
Figure 30. GAV (n=3,m=1) of Acoustic Time Trace

well log time trace, with the kernels of the GAV. From this comparison,

m=1 and n=3 are chosen as the parameters for the GAV in the detection of

the compressional wave. Our first example demonstrates the response of the GAV to an acoustic well log time trace. Figure 29 is a plot of an example acoustic time trace. The GAV (n=3,m=1) of the signal is plot-

Figure 31. Noisy Acoustic Time Trace

Figure 32. GAV (n=3,m=1) of Noisy Acoustic Time Trace

ted in Figure 30. The energy estimate in the GAV is computed using a 16 point Hamming window. The length of the window is chosen based on an estimate of the source pulse duration.
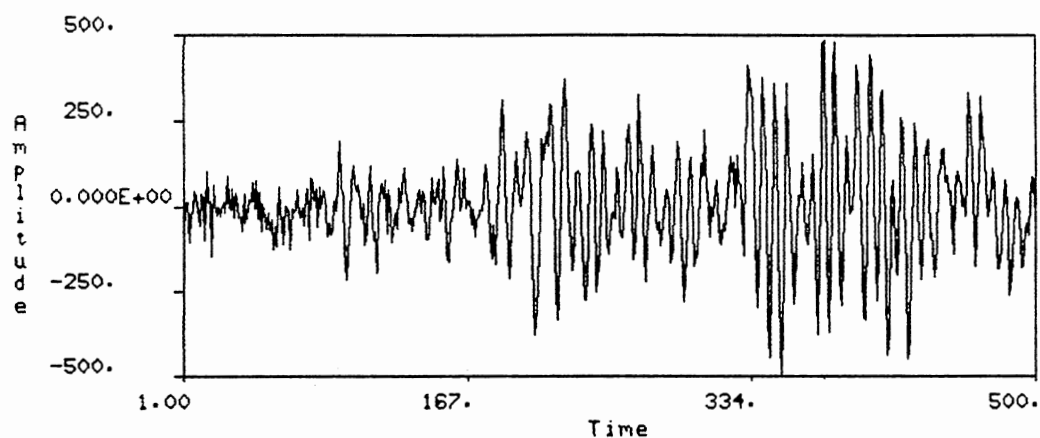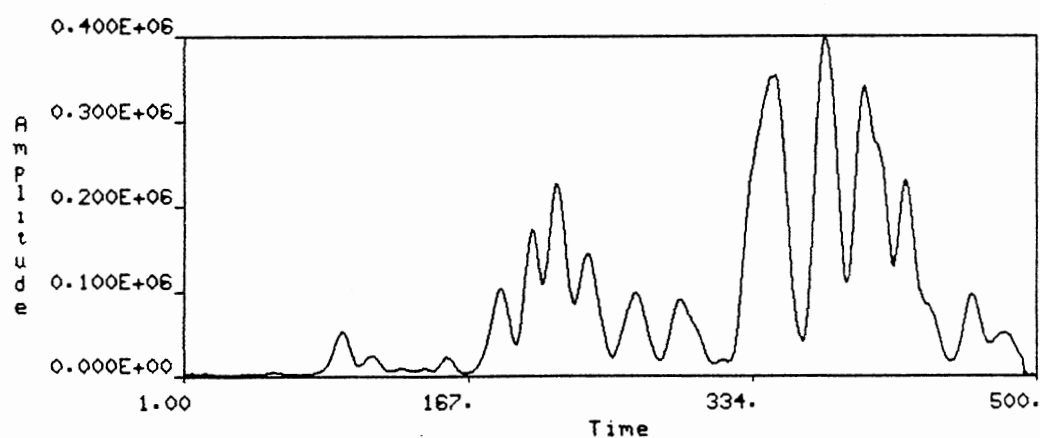
In this example, the arrival of the compressional wave can be seen quite easily in both the waveform, and the GAV. In the next example, Gaussian noise is added to the time trace giving a signal to noise ratio of 10 dB. The noisy trace is displayed in Figure 31. Figure 32 contains a plot of the GAV. Note how the major peaks are not heavily effected by the noise.

Noise is, in general, not a large problem for the acoustic well log, and this example is some what of a contrived situation. This is especially true concerning the use of Gaussian noise. For in the acoustic log, the primary noise is road noise, caused by the tool bouncing against the borehole wall or some similar phenomena [84]. This causes the noise to have a narrow band character. This example does however serve as a demonstration of how the GAV can be used to detect a signal.

The next example is a short segment of speech. The word spoken is 'to' and is chosen because of the good contrast between the unvoiced 't' phoneme at the beginning of the word and the voiced 'oo' at the end. A plot of the waveform is displayed in Figure 33.

The TDGAV, with m equal to 1 and 2, is computed for the example speech segment, using a 64 point Hamming window. The window length is chosen to be about the same as the maximum pitch period.

Figures 34, 35, 36 and 37 are surface plots of the various forms of the TDGAV applied to the speech signal given in fig. 33. Figure 34 is a surface plot of the TDGAV for m=1 and Figure 35 is the TDGAV for m=2. From these figures, we can see the difference in the TDGAV's response to
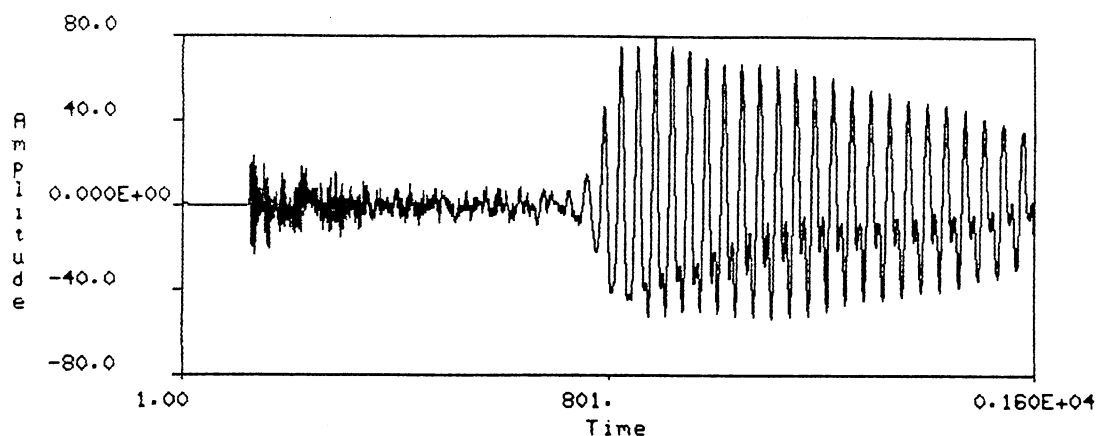
Figure 33. Example Speech Waveform

the unvoiced phoneme 't', which peaks for n=1, and the voiced 'oo', which peaks at n about 11. These plots also demonstate how as the parameter n increases, the overlap of the kernels causes a blurring in the TDGAV. This is the motivation for using the Alternate Allan Variance when analyzing speech.

Figure 36 contains a surface plot of the TDAAV using the square operator. Note how the peak for the body of the 'oo' phoneme has a much narrower peak, than in the case of the TDGAV. However there seems to be an artifact of some type occuring primarily at the begining of a phoneme. The character of these artifacts is not quite clear. However, we believe they arise due the long length of the unit sample response associated with large values of n. Consider the fact that the AAV consists of a mean filter followed by a comb filter. The length of the filters unit sample response is, in the case of the AAV, equal to the n times n+1, where n is the parameter n. This means that for n = 16, the unit sample response of the AAV would be 272 points long, much longer
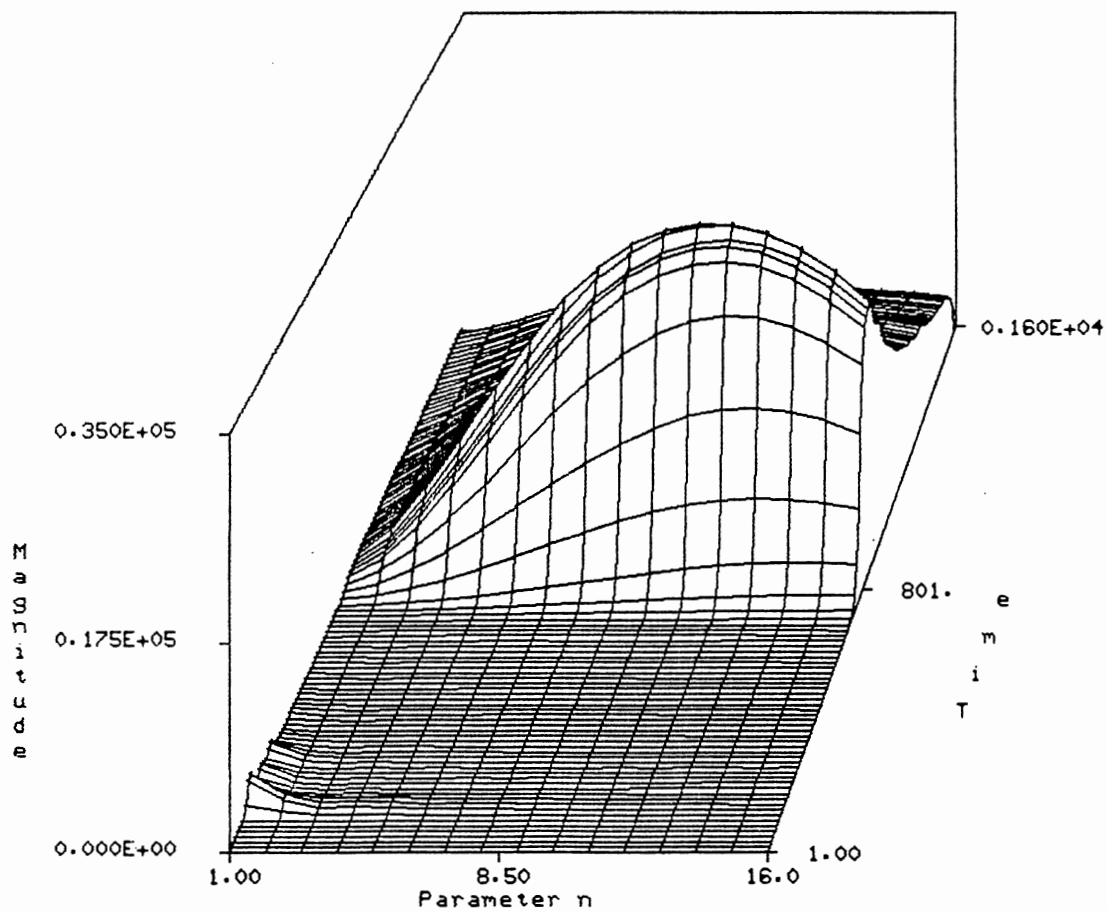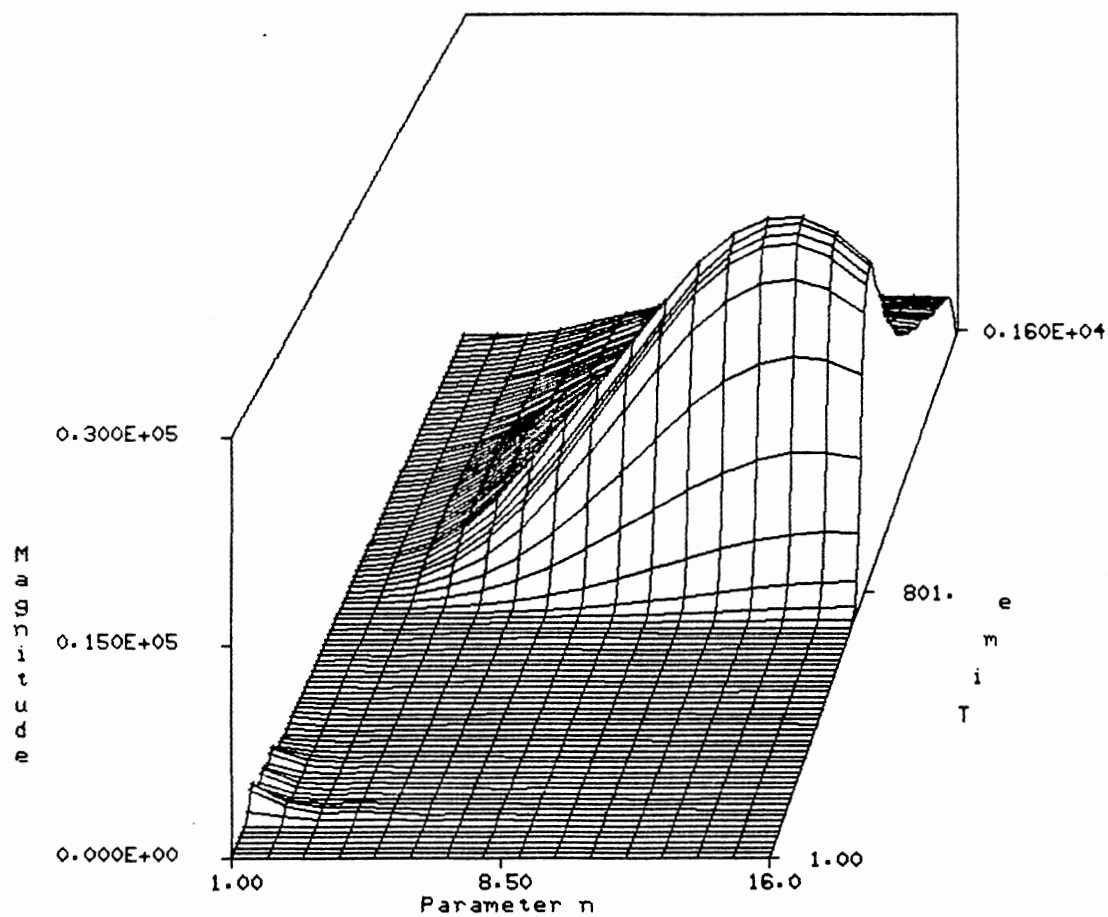
Figure 34. The TDGAV (m=1) of Speech Sample
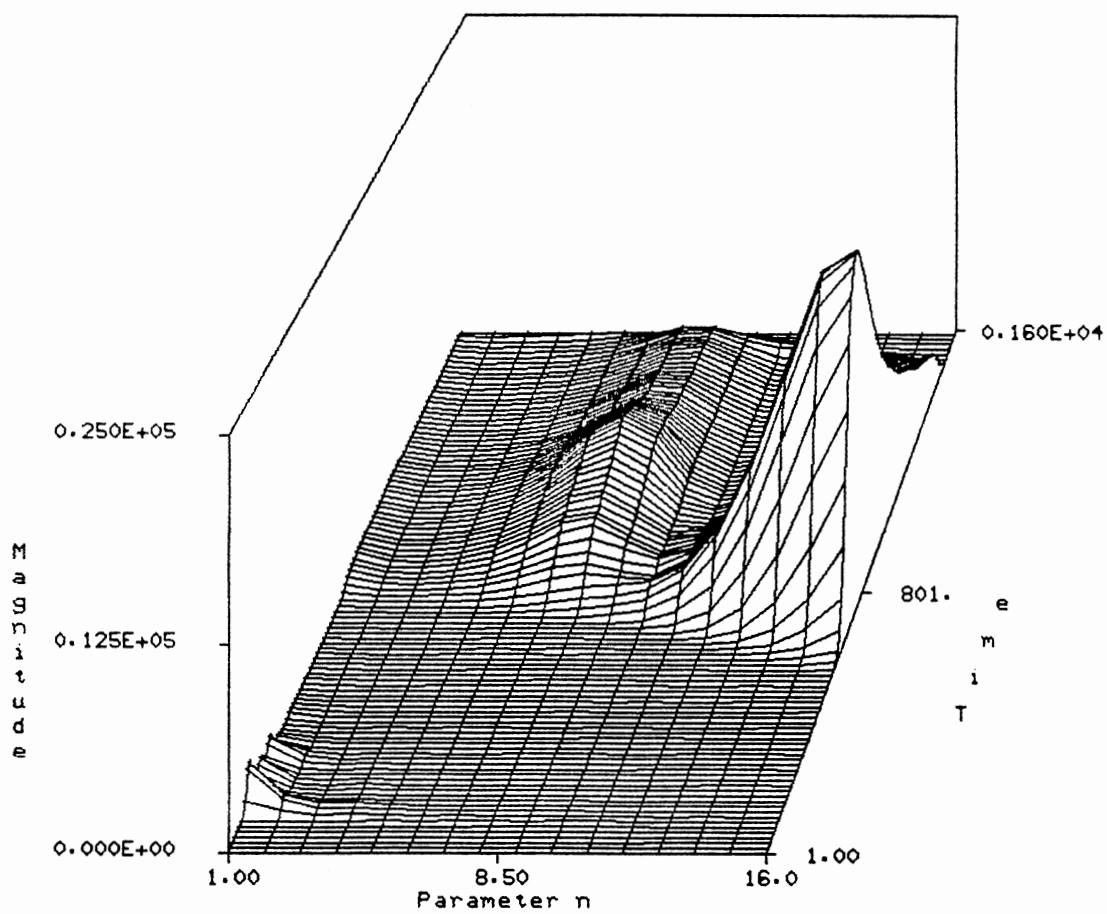
Figure 35. The TDGAV (m=2) of Speech Sample

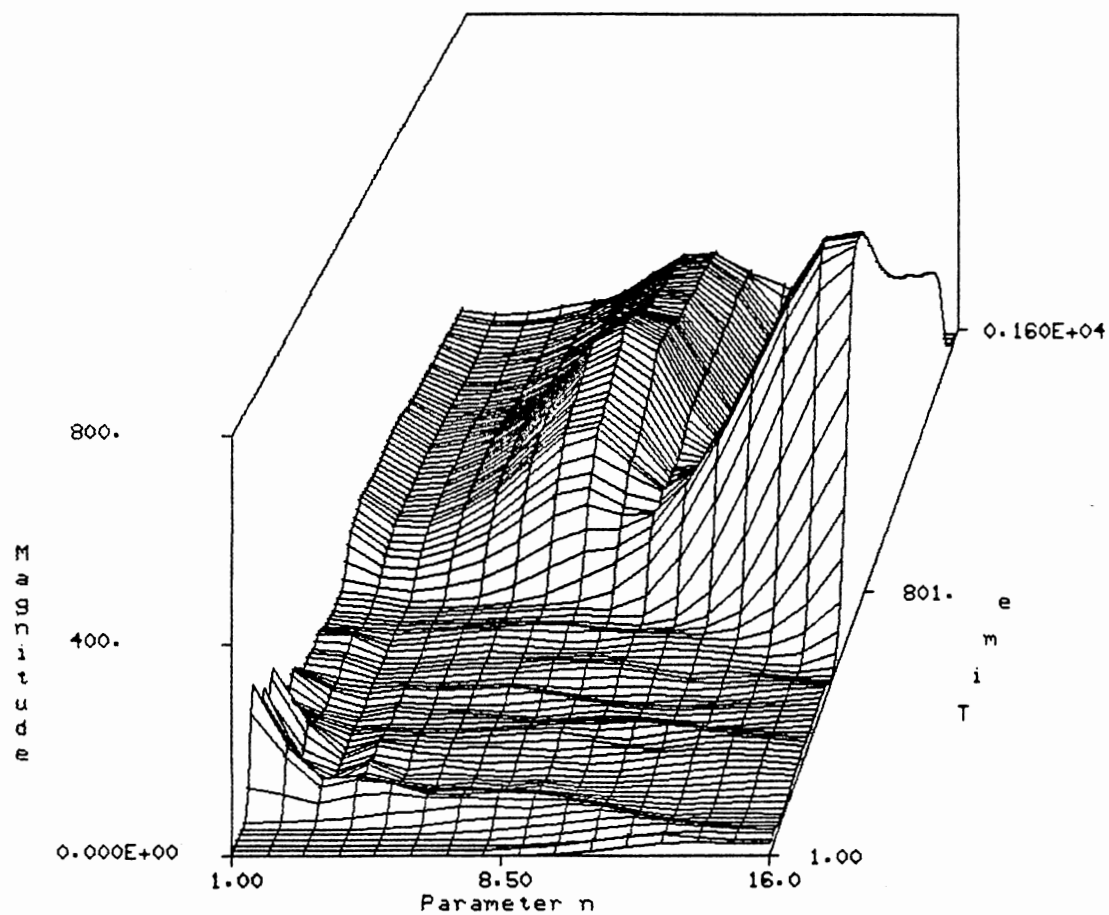Figure 36. The TDAAV of Speech Sample Using Square

Figure 37. The TDAAV of Speech Sample Using Absolute Value

than our energy estimating window. From this we can see that it may be necessary to restrict the range of n.

Figure 37 is a surface plot of the TDAAV using an absolute value operator. In this plot, we demonstrate how changing from the square to the absolute value can improve the response to unvoiced speech. This is best seen in the response of the n = 1 column for the phoneme 't', which now has an amplitude of about half that of the voiced 'oo'. This is in comparison with about one fifth for the earlier cases.

The evaluation of how effective these variations on the TDGAV are in the classification of speech is be addressed in the next chapter.

3.7 Conclusions

In this chapter, we introduced the Generalized Allan Variance. It is pointed out that the basic character of the GAV is that of a frequency selective energy measure. The frequency character of the GAV is shown to be dependent upon a kernel function, which in turn is dependent upon the setting of two parameters, m and n. Example kernel functions are generated and displayed. Also discussed are techniques for estimating the GAV for a general signal. These techniques are based on estimating the short-time energy using properly chosen window functions.

By computing the GAV for a range of values of n, a computationally efficient coarse mapping of the signals energy distribution in the time frequency plane is obtained. By producing a two dimensional array of the various GAV's, each column representing the GAV for a different value of n, we are able to display surface plots of the TDGAV. Applications for which the GAV and the TDAAV could be used, are discussed. The GAV is applied to acoustic well logs for the detection of the arrival of the

compressional wave. Examples of both the GAV applied to an acoustic well log trace and the TDAAV applied to a speech waveform are included.

In Chapter IV, we discuss further how the TDAAV can be used to detect and classify speech segments as voiced speech, unvoiced speech or silence.

CHAPTER IV

DETECTION AND RECOGNITION OF SIGNALS

4.0 Introduction

Recall that our purpose in this work is to detect the presence of a signal. In the case of the acoustic well log, we are interested in detecting the arrival of the compressional and shear wavelets. This is because acoustical properties of the surrounding formation are encoded into the time traces of the acoustic well log in the form of travel times for compressional and shear waves. For speech, we are interested in detecting the presence of voiced speech, unvoiced speech or silence. Properly making this classification is an essential step in many speech processing schemes, such as data compression, speech coding, speech recognition, speaker recognition and others. Now, in both acoustic well logging and speech processing, we are interested in detecting a change in frequency content, which exists between the various classes of signals. For this reason, Chapters 2 and 3 dealt with the development of two different types of time-frequency representations of a signal. We now address the question of how these representations can be used to detect and recognize the various signals.

The algorithms for recognizing signals begin with the computation of the Modified Auto Wigner Distribution (MAWD) or the Two Dimensional Alternate Allan Variance (TDAAV) for the signal, then features are extracted from the time-frequency representation at each point in time.

A feature is basically any measurable quantity, such as mean value, which can be used to recognize a class of signals. The features are used together in a classifier to detect and recognize the signal of interest. This process is schematically laid out in Figure 38. The first block in Figure 38 consists of computing one of the time-frequency representations described in Chapters 2 and 3. This chapter now deals with the remaining operations of feature extraction and classification.

Signals | Time Frequency Representation | Feature Extraction | Classifer | Decision
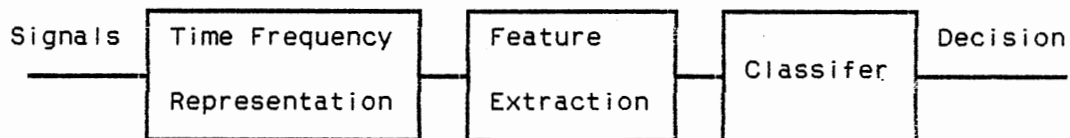
Figure 38. Signal Recognition System.

Section 4.1 introduces some of the basic concepts of pattern recognition. Sections 4.2 and 4.3 deal with the topics of feature selection and classifier design for application to acoustic well logging and speech, respectively.

4.1 Pattern Recognition Basics

The primary goal of a pattern recognition system is to associate an input pattern with a given population. Processes performing this association can be categorized into three major methodologies: heuristic, syntactic and mathematical. Heuristic methods are based on human intuition and experience. Systems designed using this principle consist of

a set of ad hoc procedures developed for specialized recognition tasks.
The performance of these systems tends to be dependent upon the clever-
ness of the system designer and therefore little can be said concerning
general principles in this area. Syntactic systems are based on the con-
cept of detecting primitive elements and recognizing the pattern by
establishing relationship grammars for the various elements. Extensive
studies into the various grammar structures have been described in the
literature [79]. However, these systems, like the heuristic systems,
require cleverness on the part of the designer to identify the proper
primitive elements. Mathematical methods are based on classification
rules which are stated and derived in a mathematical framework. The
most common mathematical method is based on the statistical models of
the populations and is known as statistical pattern recognition (SPR).
It should be noted that real systems commonly use concepts from each of
these methodologies and can rarely be considered as purely heuristic,
syntactic, or mathematical. We will centralize our attention upon math-
ematical methods, for two reasons. The first is that these systems lend
themselves to analysis. Secondly, mathematical classifiers tend to make
good baseline systems. A mathematical or statistical classifier is many
times the first process used, with other processes, such as syntactic
grammars being used to analyze the results.

A generic pattern recognition (PR) system is shown in Figure 39.
This system is the same as the one in Figure 38, with the exception of
the time-frequency block. However, if we consider the time-frequency
block as part of the feature extraction, then the two systems are the
same.

```
Signals │ Feature  │   ┌──────────┐ │ Decisions
─────────│          ├───│Classifier│─│──────────
         │Extraction│   └──────────┘
```
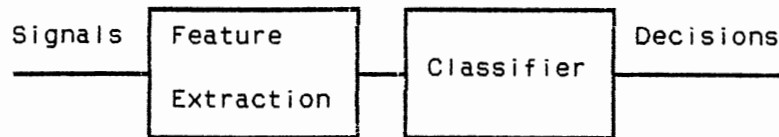
Figure 39. Generic PR System

The first operation in our generic system is the features extraction. Now, the derivation of the features to be extracted from the signal is by far the least scientific part of statistical pattern recognition. Possibly the best way to describe the type of reasoning used in the development of features is to go through this process for an example. Duda and Hart [30] give such an example, in which they present the imaginary problem of classifying lumber as either ash or birch from digitized images of the wood. The features proposed for this classification are the average intensity and a measure of the grain prominence (derived from some type of measurement of the edges in the image). These features were based on the observations that birch is lighter than ash and that ash has a more prominent grain pattern. This brings out the important point that features are largely derived from the designer's observations. Because of this, features tend to be highly application dependent. In the next two sections, features are developed for two different applications.

Once a set of features has been proposed, some type of testing needs to be done to determine which of these features will be required to discriminate between the various classes. This process is called feature selection. If there is a theoretical background for a feature, then it may be possible to show a separation of the various classes analyt-

ically, although this is rarely possible. If we cannot verify our observation, then a training set is required for the selection of features. A training set is a collection of signals, which are considered to be representative of the classes we wish to recognize, with each sample having been identified "a priori" as belonging to a certain class.

Once a training set has been gathered, which is considered to represent the patterns that will be encountered, the proposed features are then computed for the samples. The result is a large amount of multi-dimensional data, which must now be processed in an effort to determine which features possess the greatest ability to discriminate between the different classes. Many different techniques have been proposed for this purpose, all of which seek to provide understanding into the distribution of the data. For the case of only two features, a simple graphical technique called a crossplot can be used. The crossplot of two features is obtained by placing a class identifier for each sample on an x,y grid; the x,y location of each identifier is a function of the features for that sample. Figure 40 is an example of a crossplot, in which two classes are marked by the ones and twos. By inspection, we can see how the various classes group in the feature space. In this work, we will be involved with sets of only one or two features, making this one of our primary techniques for analyzing our data. For cases where more features are required, graphical techniques must be replaced with more advanced algorithms. Cluster seeking, probability distribution estimation and orthogonal decomposition of the data are some of the general classes of algorithms developed for this operation. Extensive treatment of the various algorithms can be found in references [30, 33, and 79].
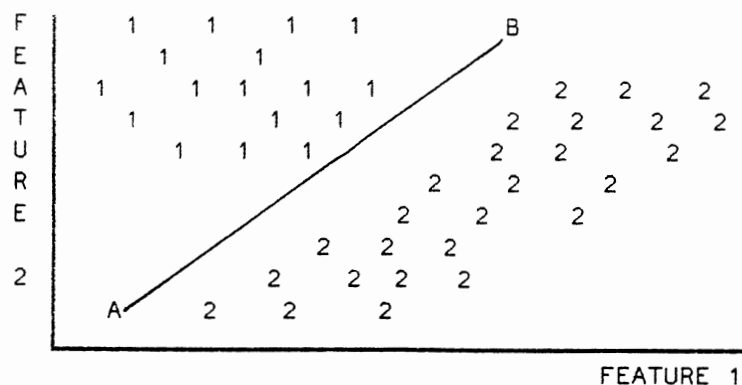
Figure 40. Example Crossplot

The last step in our PR system is the classifier. The purpose of the classifer is to associate a set of features measured from a signal with the appropriate class. In another sense, the job of the classifier is to divide up the feature space into regions, with each region corresponding to a certain class. From our two feature example in Figure 40, we can see the line AB perfectly divides the two classes. In an ideal situation, the feature space is divided up such that the features for a sample always fall into the appropriate region. However, this is not always possible in a real situation.

One of the easiest ways to divide up the feature space is by linear functions. The line AB, used to divide the classes in Figure 40, is a classic example. For this reason, linear and piecewise linear discriminate functions are common techniques for developing classifiers. As with feature selection, when the number of features is greater than two, the selection of linear functions is more complicated. Fisher's linear discriminate functions, which are based upon the minimization of a criterion function [33], is a technique for developing these discriminate

functions. Another way to set the regions in the feature space might be to choose them such that they minimize the probability of error; or, if certain errors are more costly than others, we may need to minimize the average cost of errors. The Bayes classifier, based on Bayes law in statistics, is a very common classifier, which will minimize the probability of error. However the Bayes classifier requires that the conditional probabilities be known or estimated from the training set. For this reason, we restrict our work to the development of linear or piecewise linear classifiers.

In the next section, we develop features for the detection of the shear wave arrival in the MAWD of the acoustic trace. This detection requires only one feature, however the classifer must employ some contextual information in its decisions. In Section 4.3, features are developed for the classification of speech segments. A classifer is designed and tested on speech samples.

4.2 Shear Wave Detection via the MAWD

In Section 1.1, we introduced the basic concepts of acoustic well logging and pointed out that important information about the surrounding formation is encoded into the wavetrain in the form of acoustic wave travel times. The compressional wave is commonly detected by assuming it will be the first acoustic phenomenon to affect the received signal. In this way, the first point in the signal to exceed a set threshold is considered to be the beginning of the compressional wave and from this detection, compressional travel times can be measured. Another acoustic wave useful in the evaluation of formations is the shear wave. However, the shear wave cannot be detected by thresholding, since the trailing

end of the compressional wave generally obscures the arrival of the shear wave. We therefore need to detect the shear wave based on some property other than the magnitude of the signal. In Section 1.1, we stated that the compressional and shear waves have different frequency characters. Our intent is to make use of the MAWD to detect this frequency seperation and there by recognize the shear wave's arrival in the wavetrain.

From our initial experiments with the MAWD, we found that although a frequency separation exists between the shear and compressional wavelets, it is of such a small magnitude and unpredictable nature that it is very difficult to detect. We are therefore forced to look for different features that could be used to identify the wave. In searching for a better feature we noted that linear scaling did not allow many of the details of the MAWD to be seen. To correct for this we began by setting any zero or negative values of the MAWD to a small positive number (1.e-30). Next we compute the log of the MAWD. We refer to the resulting data as the log MAWD. A surface plot of the log MAWD, for the example acoustic well log trace in Section 2.10, is shown in Figure 41. A very important characteristic, seen in this plot, is that at the arrival of the shear wave the MAWD seems to spread out along the frequency axis. Note that in between the arrivals, the Log MAWD contains sharp peaks. These peaks may correspond to the natural response of the borehole to the various waves. Arrows are included, marking the arrivals of the compressional and shear waves, to aid the reader in identifying the spreading effect, to which we refer. Also, noticeable in this plot is another location, later in time, where the Log MAWD spreads out in the frequency domain. This occurrence is attributed to the arrival of one of the guid-

ed fluid waves in the wavetrain. Note that the frequency range of the plot is reduced to those frequencies where the shear and compressional waves dominate.
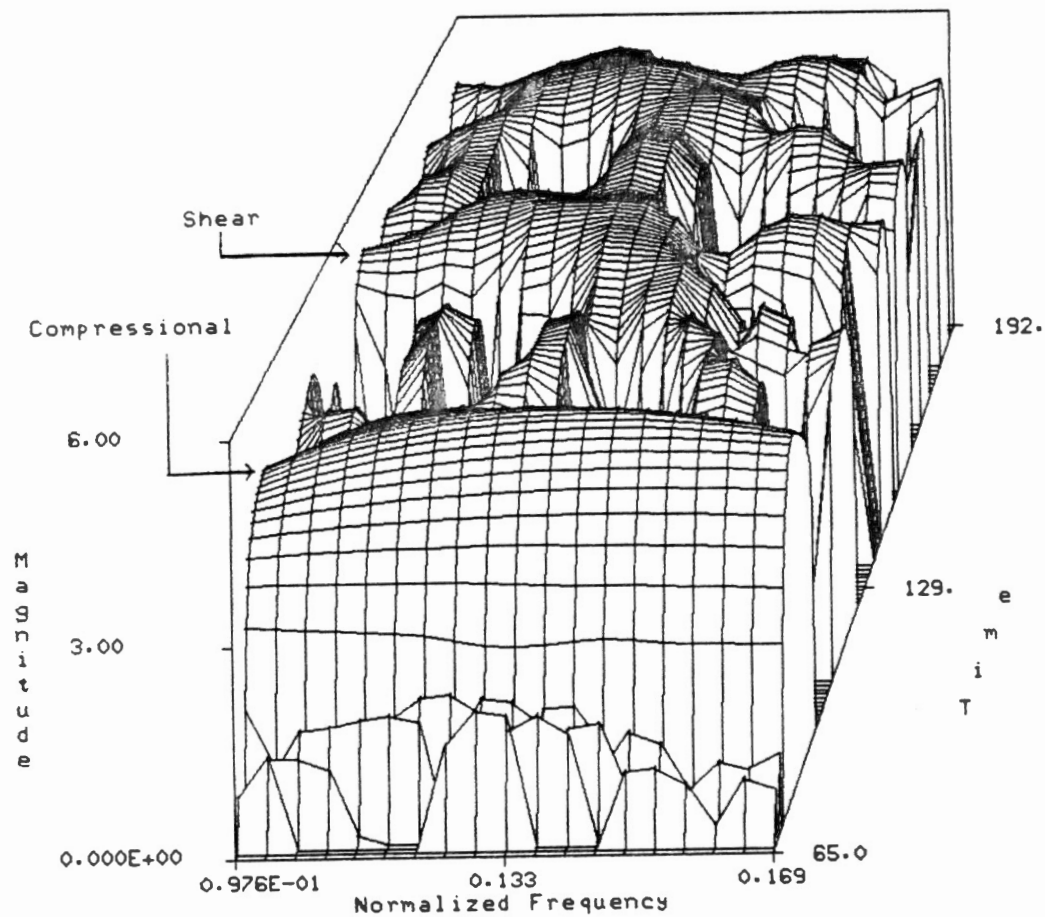

Figure 41. Log MAWD of an Acoustic Trace.

Based on these observations, the feature we use is a measure of how flat the log MAWD is over a given frequency range. This feature, referr-

ed to hereafter as the flatness feature, is defined by

$$Flat(n) = [\sum_{i=n_1}^{n_2} | \overline{W}_f(n,i) - \mu |] \qquad (4.1)$$

where $\overline{W}_f(n,i)$ represents the log MAWD of f,

μ is the mean value of $\overline{W}(n,i)$ for $n_1 \leq i \leq n_2$ and

n_1, n_2 are the indices of the desired frequency range.

Note that the value of this feature decreases if the log MAWD is flat

and increases if it is not. To explain what this feature is trying to

detect, let us consider the MAWD of a step function. A step function is

given by

$$u(n) = \begin{bmatrix} 1 & \text{for } n \geq 0 \\ 0 & \text{otherwise} \end{bmatrix} \qquad (4.2)$$

and its MAWD is given by

$$W_u(n,\theta) = \begin{bmatrix} 0 & \text{for } n < 0 \\ \dfrac{\sin((2n+2)\theta)}{\sin(\theta)} & \text{for } n \geq 0 \end{bmatrix} \qquad (4.3)$$

Considering $W_u(0,\theta)$, we can see that the MAWD is rather wide band in

character. Now, an extremely simple, but intuitively helpful, model of

the shear wave would be a frequency modulated square pulse. In this way,

the arrival of the shear wave can be considered similar in nature to the

leading edge of the step function. Thus, the spreading of the MAWD at

the shear wave is to be expected.

The nature of this feature is still rather intuitive and therefore

let us illustrate its response with an example. Figure 42a contains a

plot of a 500 point acoustic well log trace (from an actual well log)

and the flatness feature computed for each point in time is plotted in

Figure 42b. In the trace we can see the arrival of the shear wave by

visual inspection. Also apparent in Figure 42b is the large drop in the flatness measure which accompanies the shear wave arrival. Note both of these scale are arbirtary and strictly dependent on the acoustic sensor and the definition of the feature. To aid the reader, both the compressional and shear wave arrivals are marked in Figures 42 and 43.
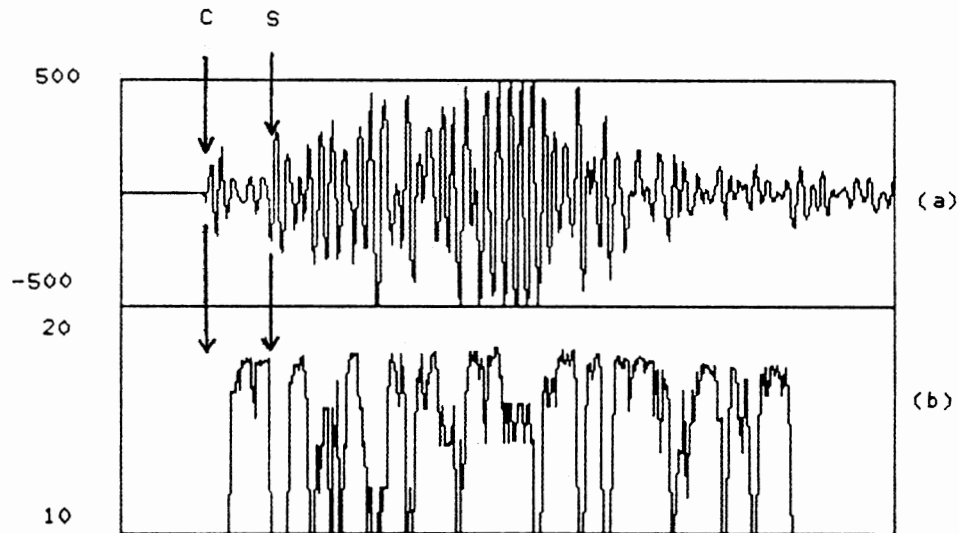
Figure 42. a) Time Trace for Ten Foot Spacing
b) Flatness Measure of Trace

The trace in Figure 42a is for a ten foot transmitter-receiver spacing. The next trace, given in Figure 43a, is for basically the same depth point but having a twelve foot spacing. The effect on the magnitude of the shear wave is immense. Plots of the time trace and it flatness measure for the twelve foot spacing are shown in Figures 43a and 43b, respectively. In this case, the shear wave cannot be identified as

distinctly as in Figure 42, however, the flatness measure shows a very distinct change.
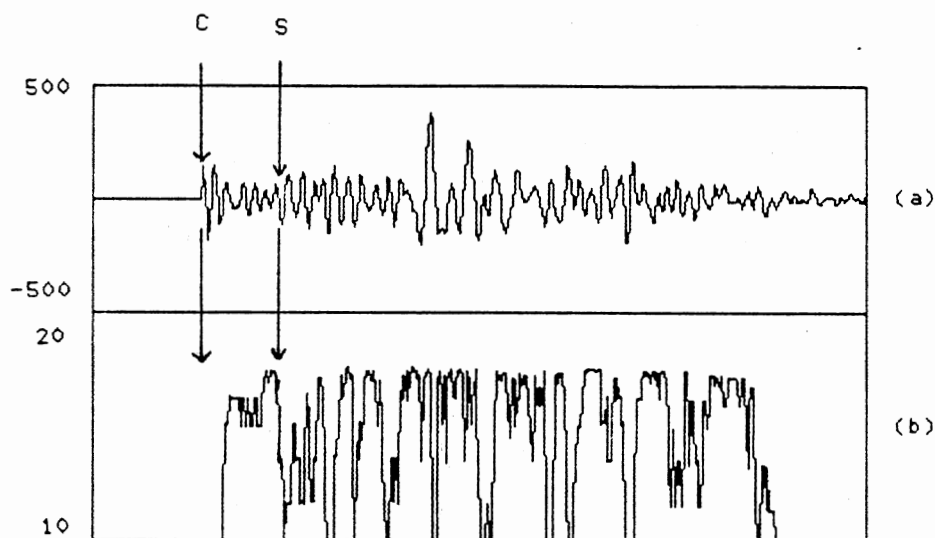
Figure 43. a) Time Trace for Twelve Foot Spacing
b) Flatness Measure of Trace

A collection of well log traces taken from an actual well log are displayed in Figure 44. These traces were recorded from an acoustic sonde which has transmitter receiver spacings of eight and ten feet. The first trace is from the eight foot spacing, and the second trace is from the ten foot spacing at the same depth. The rest of the traces can be divided up into depth point pairs in a similar fashion. Figure 45 is the flatness feature computed for the 128 point DMAWD of the same traces. The frequency range, over which the flatness is computed, is from index 25 to 43, corresponding to a range of 0.0976 f_s to 0.168 f_s. The reason

Figure 44. Acoustic Traces

Figure 45. Flatness Feature of Traces

Figure 46. Filtered Features

for these figures being included is to demonstrate that the flatness feature is consistent in its response.

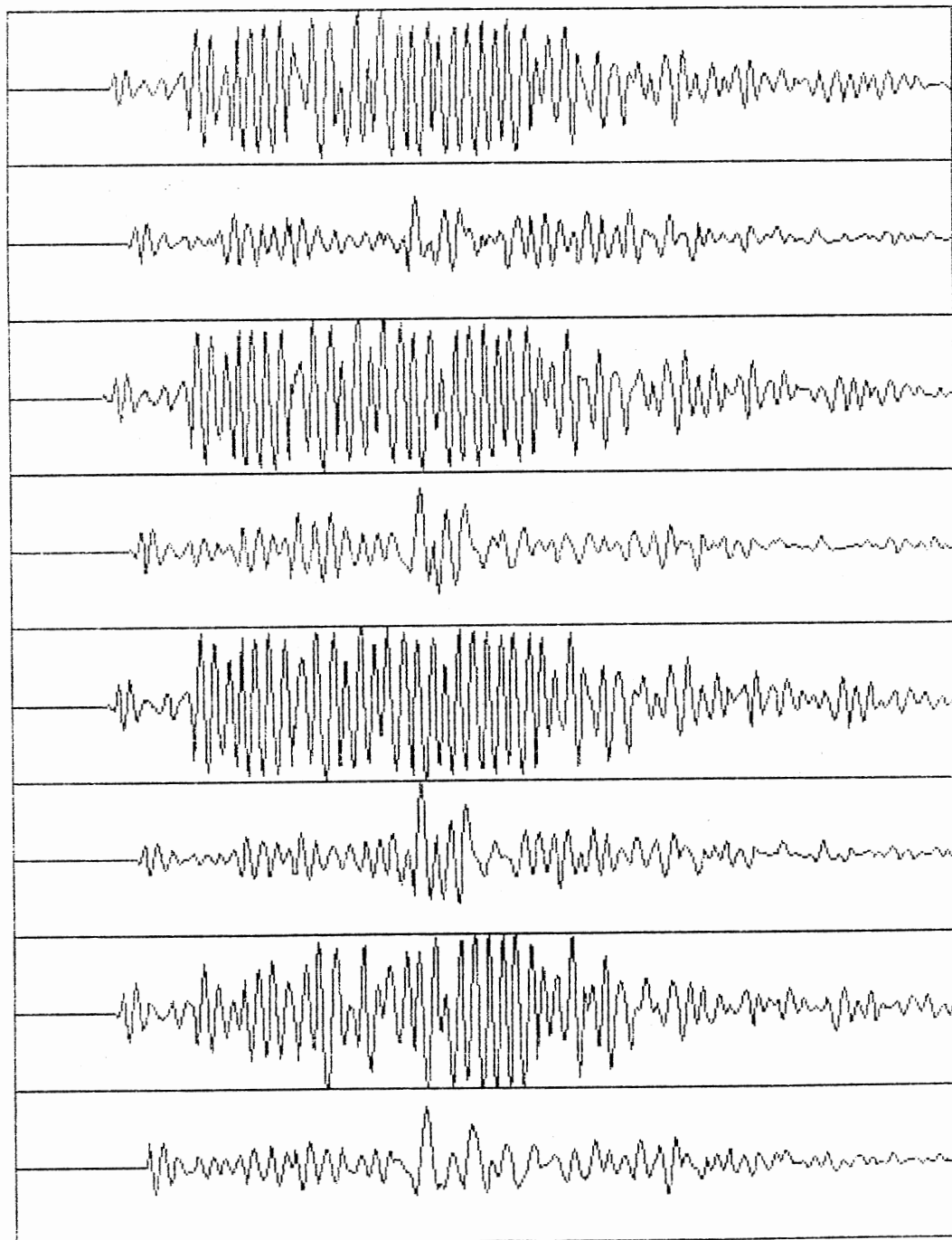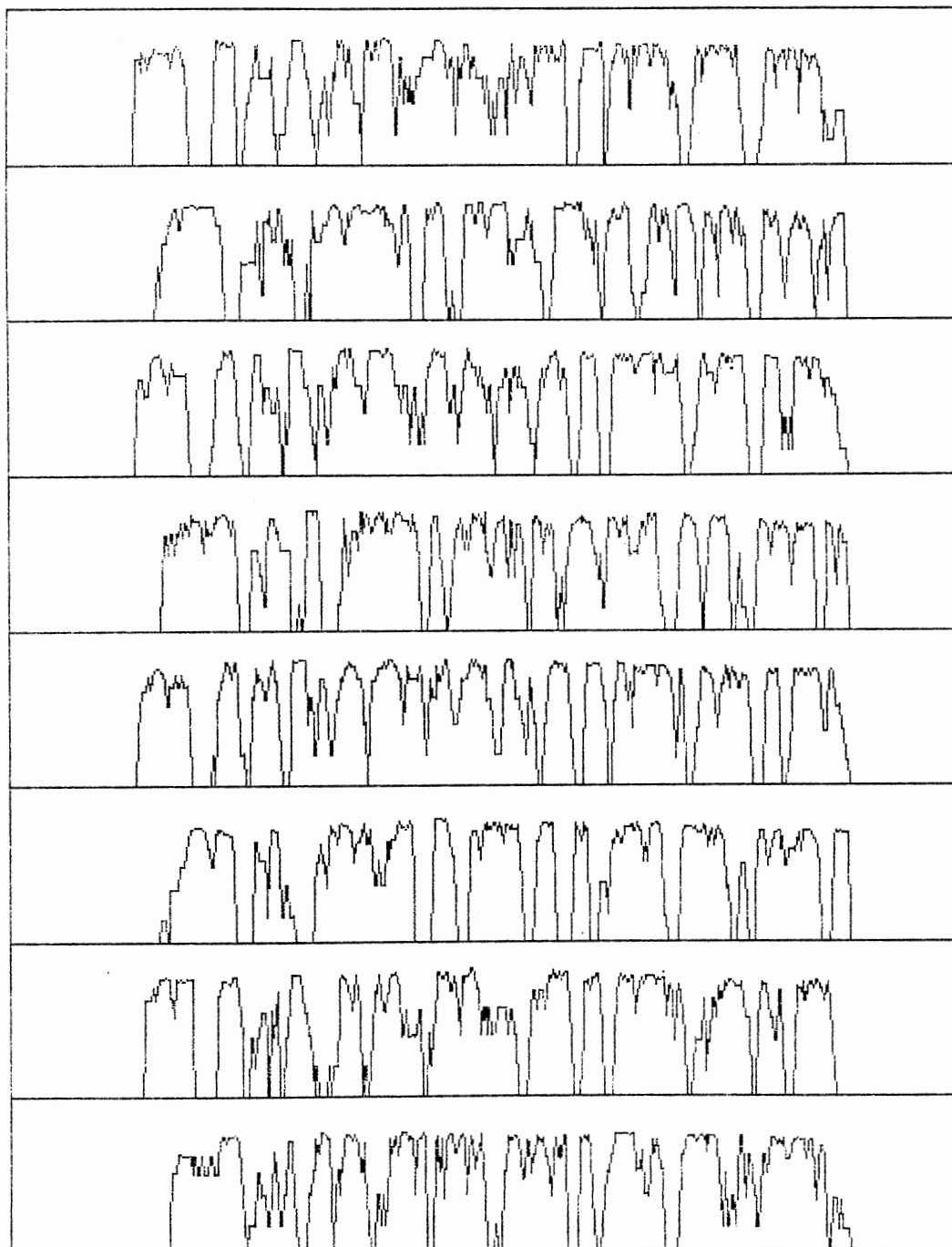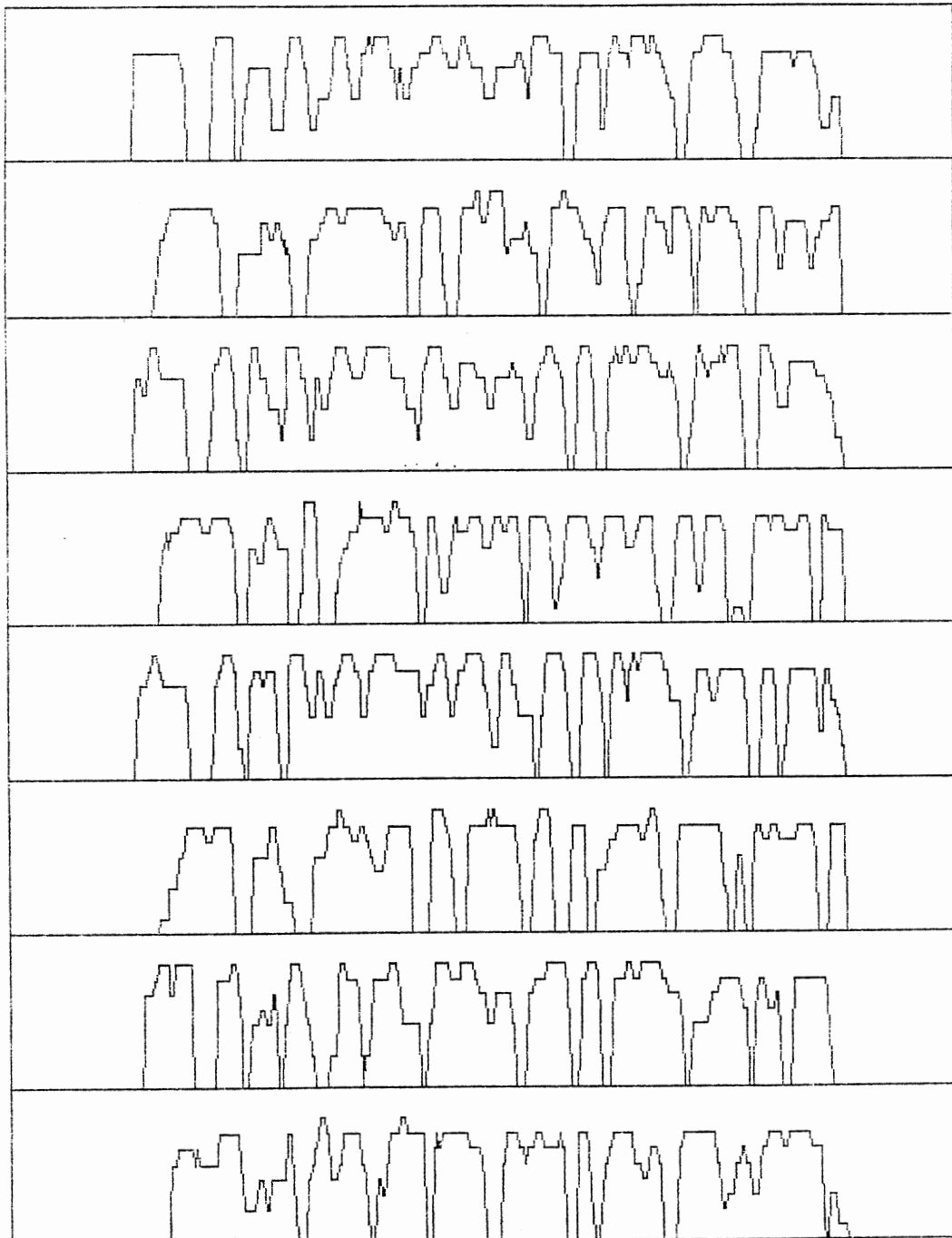In some cases the feature contains some noise. To decrease the effect of this noise while still allowing for the sharp arrivals to be detected, a median filter is used on the feature. The length of the filter was chosen to be 5 points long, strictly on the observed effect of the filter on the data. Figure 46 shows the features after filtering.

The next step is to develop a detection procedure for detecting the shear wave arrival based on the flatness measure. Since these traces are virtually noise free, the arrival of the compressional wave is detected by thresholding the time trace. Recall that a sliding 128 point Hamming window is applied to the data, prior to computing the MAWD. The time index of the MAWD is equivalent to the center of the window. Thus, the first point in time, for which we can compute the MAWD is 64 points into time trace. Thus, we have not used the MAWD to recognize the arrival of the compressional wave, since the compressional wave frequently arrives before the 64'th time sample. Having detected the compressional wave, we now begin searching for the shear arrival. This search begins at the compressional arrival plus an offset. The offset is included to insure that we are past the effects of the compressional wavelet. We then move forward in time until the flatness descends below a threshold. In this work, we have used 14 but is, in general, dependent upon the energy level of the signal. The time at which the flatness descends below the threshold is considered the shear wave arrival. This procedure can only determine the arrival of the various waves to within the sampling interval of the recorder. Note that cross-spectral techniques have been developed which estimate the travel time for sub-sample units of time. The

waveforms in this work are sampled at a rate of 100 KHz, which means our estimates are in 10 μ second increments. Now, before we begin applying this procedure to data, a brief description of the training set is given.

Our training set consists of a collection of time traces from an actual well log, supplied to us by Amoco Research of Tulsa. The tool used has two transmitters and two receivers. From these different transmitters and receivers, traces for an eight, ten and twelve foot spacing can be obtained. The traces in Figure 44 are examples of these traces. Also, supplied with the time traces are the compressional and shear wave logs computed using thresholding and cross-spectral techniques. These logs are displayed in Figure 47.

The first step is to compute the flatness measure of the well log traces. Then the shear wave arrival is detected using the search procedure outlined earlier. Now the traces are taken at different depth points, four traces per depth. Thus, to derive the shear wave velocities from the shear wave arrivals, the difference in arrival times for two traces at the same depth point must be computed. Then the spacings between the two receivers is divided by the travel time, giving the velocity. The way the sonde is set up, the traces come out in pairs, the first pair are for a transmitter and receiver spacings of eight and ten feet and the second pair are for the ten and twelve foot spacings. First, the set of traces representing the eight and ten foot spacings are sorted out. Then the difference in the shear wave arrival time is found between each pair of traces. The resulting transit times (in μseconds) are plotted in figure 48.

Comparing the shear log obtained using cross-spectral techniques
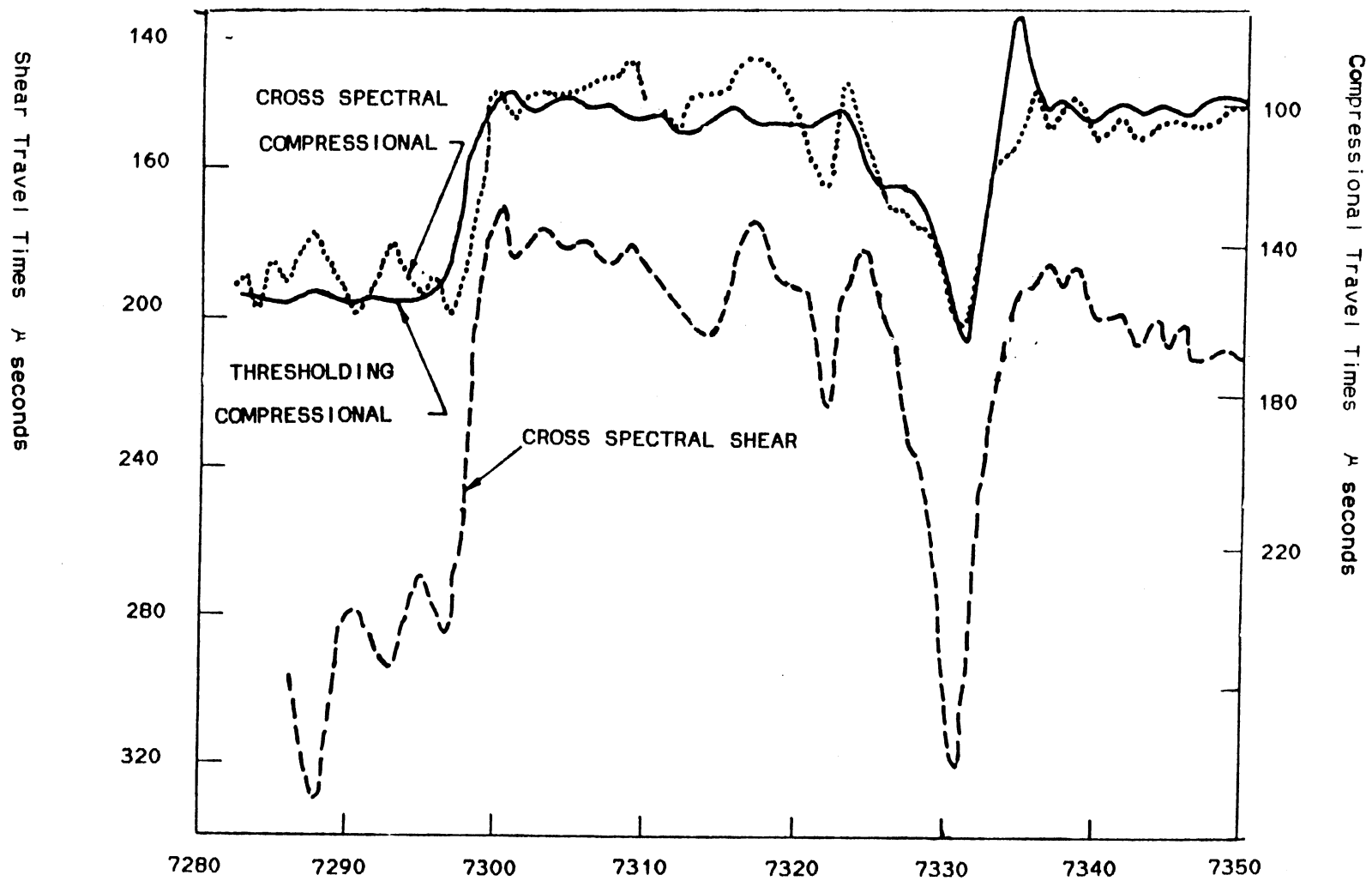
Figure 47. Compressional and Shear Log of Actual Well Log Data

and the log obtained from the flatness measure, good agreement can be seen between the shear transit times derived from each method. This can be observed, by noting the first three primary beds in the log. The first bed is between depths 7290 and 7295 and has a long transit time. The second bed is relatively long, going between 7295 and 7325, and has a transit time of approximately 200 μ seconds. Finally, a third bed, at 7330, can be seen having a long transit time. This third bed is an excellent marker for correlating the logs.
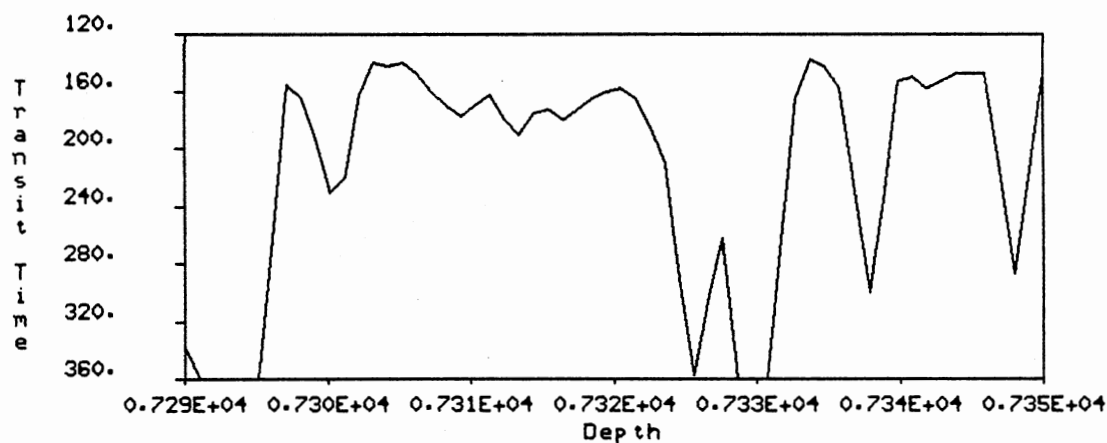
Figure 48. Shear Log of Actual Well Log Data Using
the MAWD and Flatness

However, some points can be found where the flatness appears to be insufficient to recognize the shear wave, such as the points between 7335 and 7350, where the shear transit time drops below 240 μ seconds. This does not correspond well with the cross-spectral log, which is consistent at approximately 200 μ seconds. This implies that a more com-

plete set of features may be required to properly recognize the shear wave. It should also be noted that for the slow formations, where the shear transit time becomes greater than approximately 250 μ seconds, the flatness log seem to record a much larger shear transit time than does the cross spectral shear log. It could be that the flatness measure is missing the shear arrival in these slow formations, since the shear wave tends to drop in amplitude for these formations. It may be that instead of detecting the shear wave, we are detecting the arrival of one of the guided fluid waves. It is also possible, since we do not know what type of processing has been done to the log in Figure 47, that the cross-spectral log is in error. However to address these problems would require a trained geophysicist and will therefore not be addressed at this time.

It is hard to make a definite statement as to which technique, the flatness measure or the cross spectral technique, is the most effective. To properly compare the two techniques would require more data and more knowledge of the borehole acoustics. However, we can see that the flatness measure does respond to the arrival of the various waves in the wavetrain (See Figures 42 and 43) and is therefore an alternate technique for measuring the shear wave transit time.

In the next section, we move on to the task of developing features and classifiers, based on the TDAAV, for speech signals.

4.3 Classification of Speech based on the TDAAV

In Section 1.2, it was pointed out that voiced and unvoiced speech have very different spectral character. Section 3.5 showed how the TDAAV can be used to measure a signal's energy in various frequency bands. We

bring these two properties together to develop algorithms for the classification of speech segments.

In Section 3.5, the response of the TDGAV to a sample speech waveform is demonstrated and it is noted that there is a difference in response for voiced and unvoiced speech. Unvoiced speech tends to have a large portion of its energy concentrated in the frequencies greater than 2 KHz. On the other hand, the primary components of voiced speech occur below 2 KHz. The goal of the algorithm is therefore to compare the energy in these two frequency bands, namely (0-2KHz) and (2-4KHz). Three features, which are to be applied to the TDAAV, are proposed for the purpose of classifying speech. The first feature is a measure of the signal's short-time energy level, while the second and third features measure how the signal's energy is distributed in the frequency domain.

Before describing the features used in classifying speech, a brief review of the TDAAV is given. Also, the selection of certain parameters is discussed. Recall from Section 3.2, that the GAV is basically the energy in a filtered version of the input signal. The parameters, which adjust the filtering applied to the input, are used to select the frequency bands from which the energy is computed. The TDAAV is derived by computing the GAV for a range of these parameters at each point in time. At the end of Section 3.5, two different expressions for the TDAAV were introduced. The first, uses the square operation to compute the energy in the filtered signal, and is given by

$$\sigma_n(x,h) = \sum_{i=-\infty}^{\infty} W(i-h) \left(\sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{1}{n} \sum_{p=0}^{n-1} x(i-p-k\,n) \right)^2 \qquad (4.4)$$

If the square operation is replaced by an absolute value, then

$$\bar{\sigma}_n(x,h) = \sum_{i=-\infty}^{\infty} W(i-h) \left| \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{1}{n} \sum_{p=0}^{n-1} x(i-p-k\ n) \right| \qquad (4.5)$$

The second definition is hereafter called the Two Dimensional Psuedo Allan Variance (TDPAV), and must be considered as an approximation of the TDAAV. Now, each of these definitions requires the selection of a window, for estimating the energy. In this work, a 64 point Hamming window. The frequency response of the Hamming window was described in Section 3.3. The length of the window corresponds to a time interval of eight milliseconds, assuming an 8 KHz sampling rate, and is longer than the pitch period of most speech. Finally, the range of the parameter n needs to be set, for which the TDAAV is to be computed. The range chosen was $1 \le n \le 11$. The maximum of eleven, for the parameter n, is chosen so that the kernels of the TDAAV cannot go below 250 Hertz. In this way, the TDAAV avoids possible problems with 60 Hertz noise and its third harmonic. Having set all the various parameters, features need to be developed that can be extracted from the TDAAV and TDPAV to discriminate between the different classes of speech. Since the TDPAV is an approximation of the TDAAV, the same features can be used in both cases.

The first feature is very simple in form, and is simply the maximum TDAAV for that point in time (Equation 4.4). Since the TDAAV is a measure of the signals energy, the peak value

$$Peak\_value(h) = \max_{i=1}^{11} (\sigma_i(x,h)) \qquad (4.6)$$

should in some way represent the energy in the speech waveform. This is true, except for the fact that this energy is only for some unknown band of frequencies. This feature is used primarily in detecting speech versus silence.

The second feature is the ratio of the TDAAV for n = 1 and the sum

of the remaining TDAAV's (n = 2 to 11). This feature is given by

$$\text{Sum\_Ratio}(\sigma_n(x,h)) = \frac{\sigma_1(x,h)}{\sum_{i=2}^{11} \sigma_i(x,h)} \tag{4.7}$$

This feature compares the response of the AAV for two different ranges

of n. Referring back to Figure 28 in Section 3.2, it can be seen that

the AAV for n=1 is a measure of the energy in the upper frequencies (2

to 4KHz). Similarly, the AAV's for n = 2 to 11 measure the energy in the

lower frequencies of approximately 250 Hz to 2 KHz. In this way, we can

see that the ratio feature is a comparison of the energy in these two

bands and should be useful in separating voiced from unvoiced speech. It

should be noted that speech is assumed to have been sampled at 8000

Hertz. This assumption is important, since the range of n for the

sum_ratio and the frequency response of the TDAAV is directly related to

the sampling rate.

The third feature is similar to the ratio feature, being the ratio

of the TDAAV for n equal 1 and the peak TDAAV for n equal to 2 to 11

(Equation 4.8). Testing is now required to determine which combination

of features and definition of the TDAAV, has the greatest ability to

classify speech.

$$\text{Peak\_Ratio}(\sigma_n(x,h)) = \frac{\sigma_1(x,h)}{\underset{i=2}{\text{Max}} \, \sigma_i(x,h)} \tag{4.8}$$

The training set used in the development of our classification

algorithms consists of five sentences. These sentences are chosen be-

cause they have extremely low noise levels, allowing noise to be added

later so that the robustness of the algorithms can be tested. The sex of the speaker, pitch and words spoken are given in Table II. The pitch levels are broadly divided into three classes, low, moderate and high. Note that the pitch levels are dependent upon the sex of the speaker, with the moderate female voice having a pitch close to that of a high male. The pitch levels given in Table II are determined by listening to the samples. From Table II, it can be seen that this set of waveforms represent a wide range of phonemes and pitch periods.

TABLE II

SPEECH TRAINING SET

Id	Sex of Speaker	Relative Pitch	Words spoken
1	Female	High	The pipes began to rust while new
2	Female	Moderate	Add the sum to the product of these three
3	Male	Low	Oak is strong and also gives shade
4	Male	High	Thieves who rob friends deserve jail
5	Male	Moderate	Cats and dogs each hate the other.

A feature file is created for each of the sentences in the test set. Feature files are two dimensional arrays, with the second dimension (the columns) representing time and the first dimension (the rows) rep-

resenting the various features computed for the speech waveform at that point in time. The first feature is a number which identifies the speech for that point in time as either silence (0), voiced speech (1), unvoiced speech (2) or unknown (3). The TDAAV and the TDPAV are computed and features extracted for each waveform in the training set. The features are then placed in the appropriate location in the feature file. The resulting feature files therefore represent the response of different features to voiced speech, unvoiced speech and silence. Figure 49 is a crossplot of the peak_value versus the sum_ratio for the TDAAV applied to the training set. Figure 50 is a crossplot of the peak_value versus the peak_ratio for the TDAAV. Figure 51 is a crossplot of the peak_value versus the sum_ratio for the TDPAV and Figure 52 is a crossplot of the peak_value versus the peak_ratio for the TDPAV applied to the training set. Note that the peak_value axis in each of the cross plots is log scaled. Log scaling is used to streach out the data, allowing the reader to see the separation of the classes. If log scaling is not used, the silence and unvoiced speech tend to cluster along the sum_ratio or peak_ratio axis.

No combination of features perfectly separate the different classes of speech. However, each set shows the potential of separating the classes. Our choice of features is based on the following observations from the cross plots. The first is that the TDPAV seems to not separate the voiced and unvoiced speech as well as the TDAAV. In Section 3.5, it was stated that the reason for using an absolute value in the TDPAV is to equalize the response to voiced and unvoiced speech. However, this approximation tends to equalize the response between different bands (different values of n). This equalization tends to therefore reduce

the ratio for unvoiced speech and there by reduces the power of the
TDPAV to discriminate between voiced and unvoiced segments.

Another observation is based on Figures 49 and 50, where it appears
that for the voiced speech, the peak_ratio does not cluster next to the
peak_value axis as well as the sum_ratio. Also the peak_ratio seems to
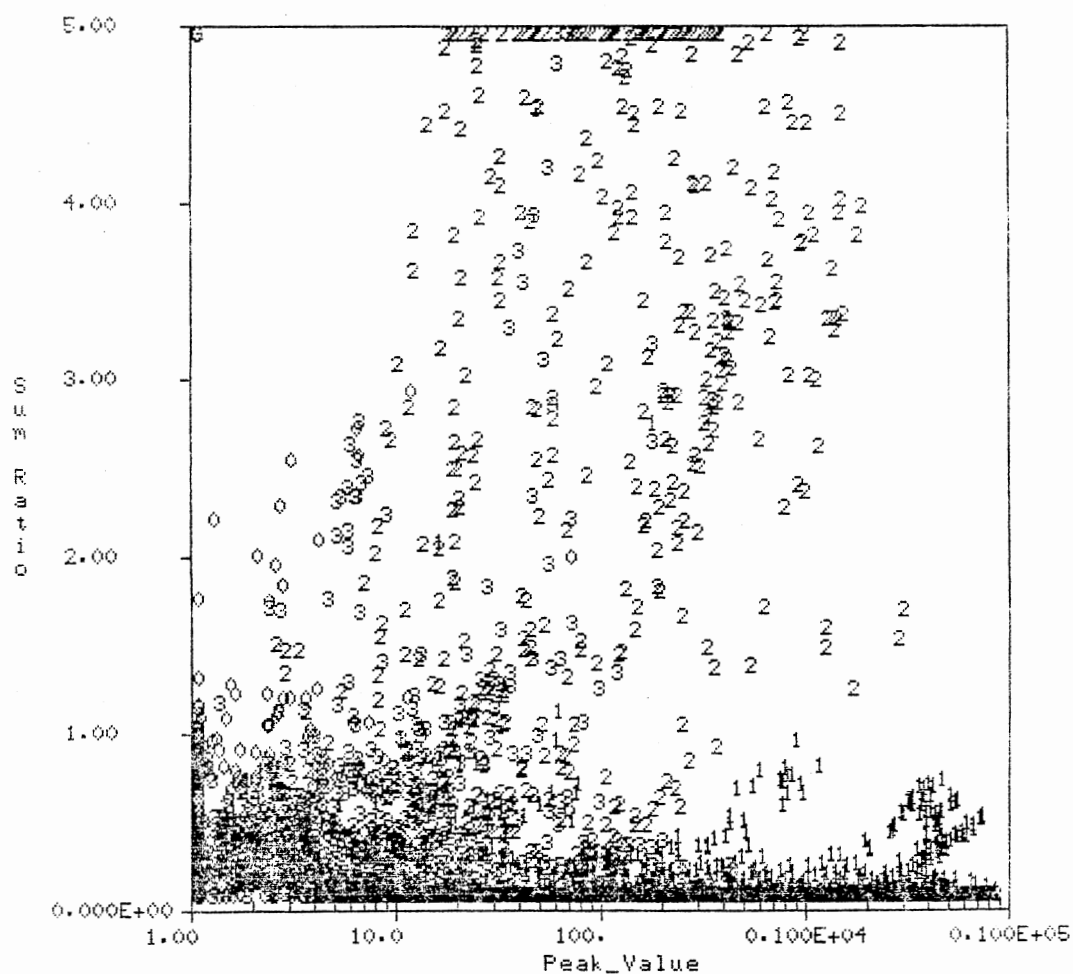be more erratic in the case of silence. The reason for this could be

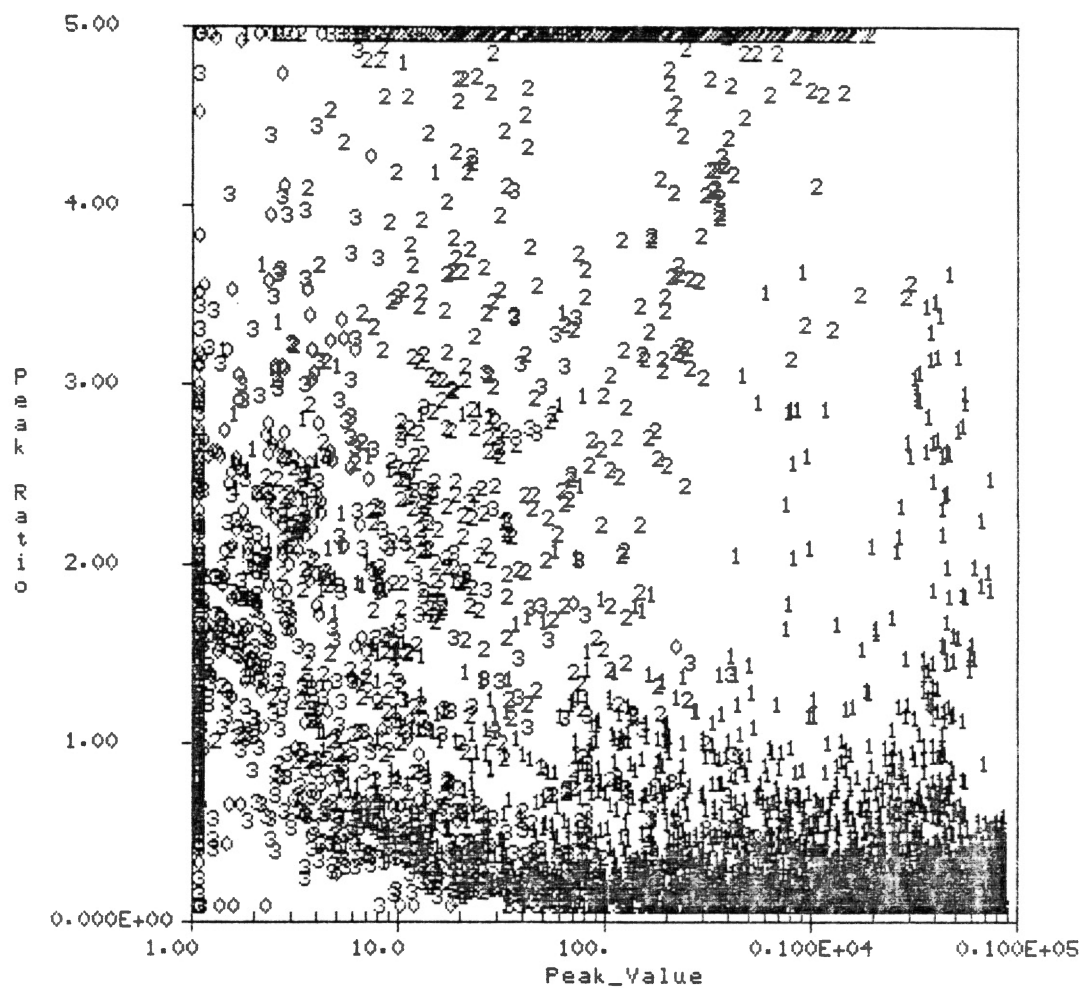Figure 49. Peak_value versus the Sum_Ratio
for the TDAAV

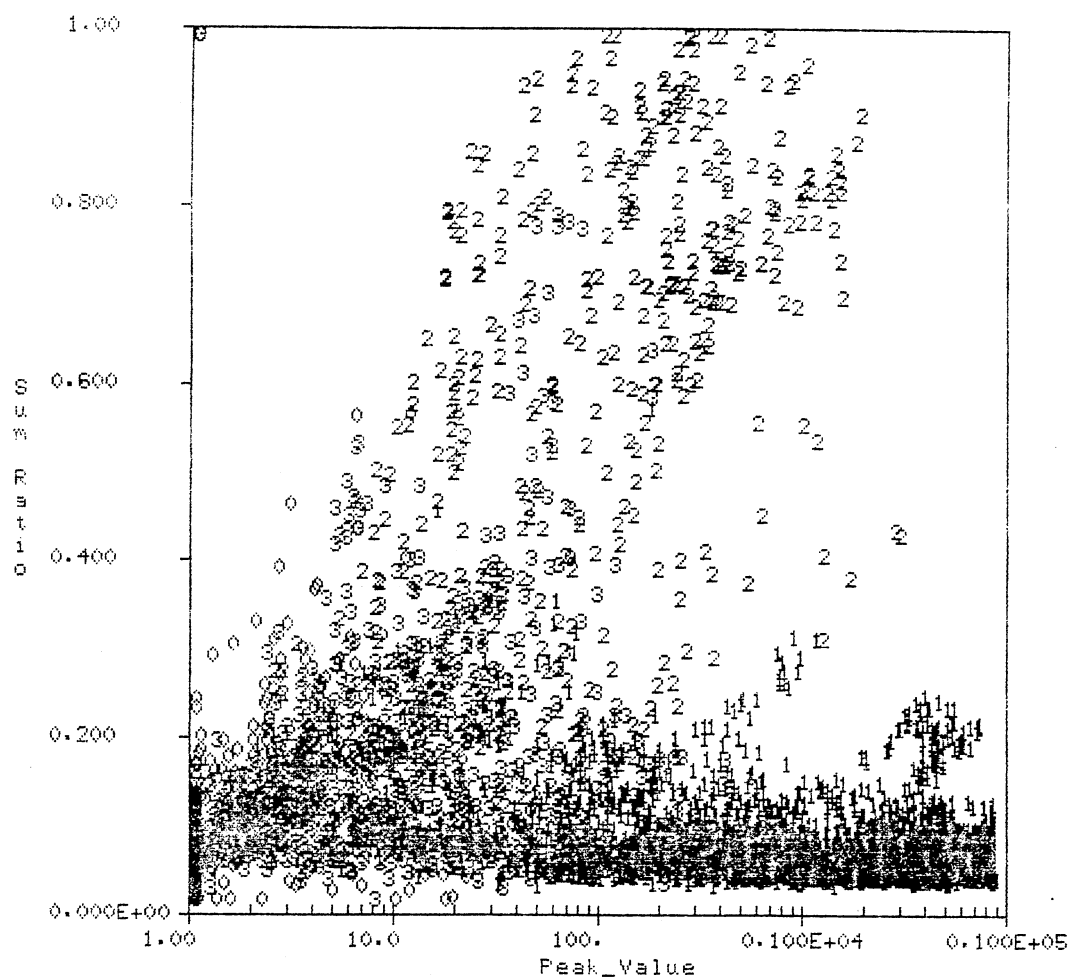Figure 50. Peak_value versus the Peak_Ratio
for the TDAAV
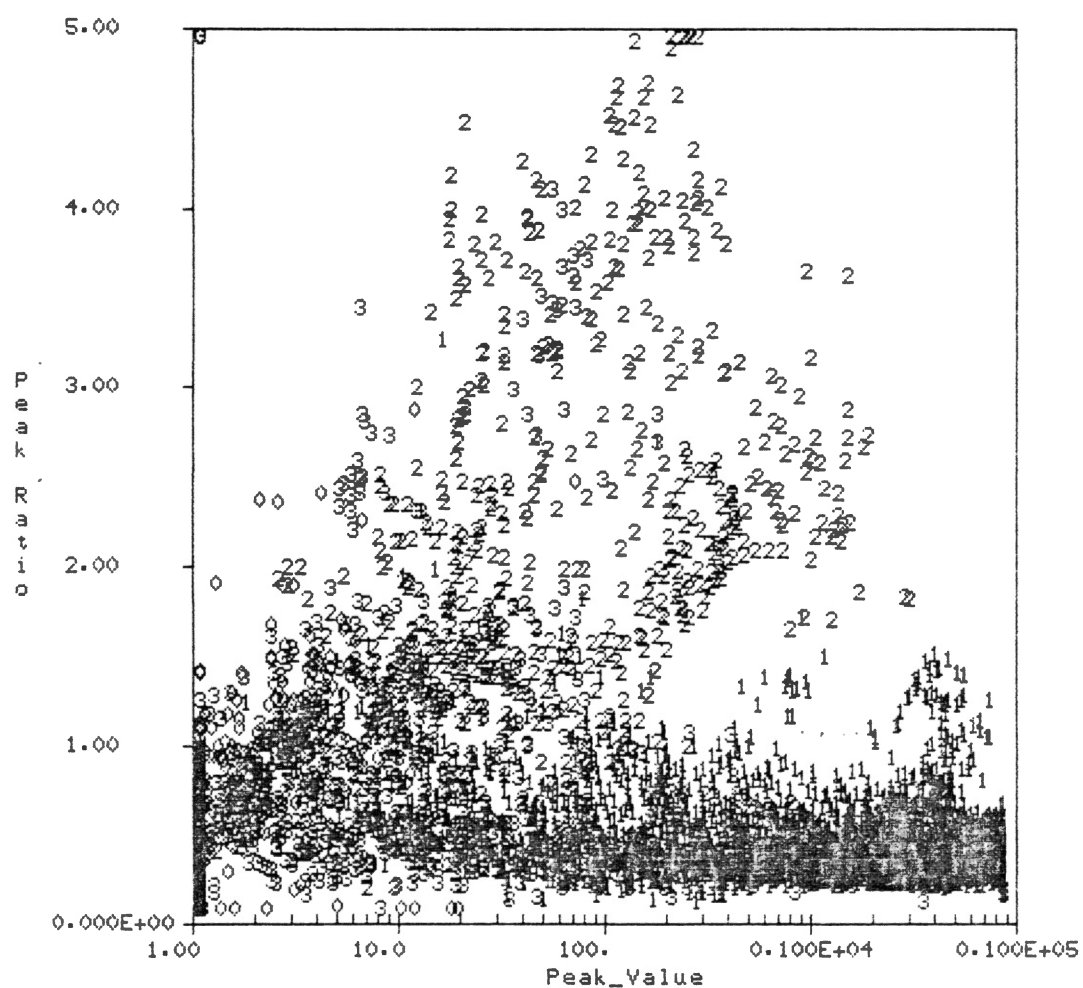
Figure 51. Peak_value versus the Sum_Ratio
for the TDPAV

Figure 52. Peak_value versus the Peak_Ratio
for the TDPAV

that frequency kernels of the TDAAV become narrower as we move towards

higher values of n. As the kernels become narrower, the response of the

TDAAV to a wide band signal tends to spread out between different kern-

els, and decrease in amplitude. Thus, two signals may have similar

energy, but if one of the signals is of lower frequency content, they

will have different peak TDAAVs. The sum of the TDAAVs can in some ways

offset this difference in the kernels.

Based on the two previous observations, we direct our attention

towards the TDAAV and the Sum_ratio. This does not mean that the other

combinations are not effective, but simply that they are not addressed

at this time. If computational load proves to be a problem in some

applications, the TDPAV may prove useful. Having chosen features, the

next step is to develop a classifier. However, before beginning on the

design of a classifier, some discussion of the distribution of the

features and what effect noise has on these distributions is desired.

In the cross plot of Figure 49, it is hard to distinguish the var-

ious classes. However, it can be seen that the unvoiced speech tends to-

ward the upper part of the plot, voiced speech tends toward the lower

part of the plot and silence tends towards the left portions of the

plot. A more accurate description of these regions is included later,

when we discuss the classifier. Figure 59 graphically depicts each of

these regions. To display the distribution of the various classes more

effectively, each class is sorted into a separate feature file. Then,

crossplots of each individual class are included. From the plots of

these separate classes, a much better view of the distributions can be

seen. Figure 53 contains a crossplot of the silence samples, Figure 54

is a crossplot of the voiced samples, and Figure 55 is a crossplot of

unvoiced samples. The unvoiced speech does not cluster as well as the

other classes (See Figure 55). This may be due to the large range of

energy levels exhibited by unvoiced speech. Also, some samples, which

should be classified as voiced speech or silence, may be improperly

classified as unvoiced speech in the training set.

Figure 53. Peak_Value Versus Sum_Ratio for
Silence
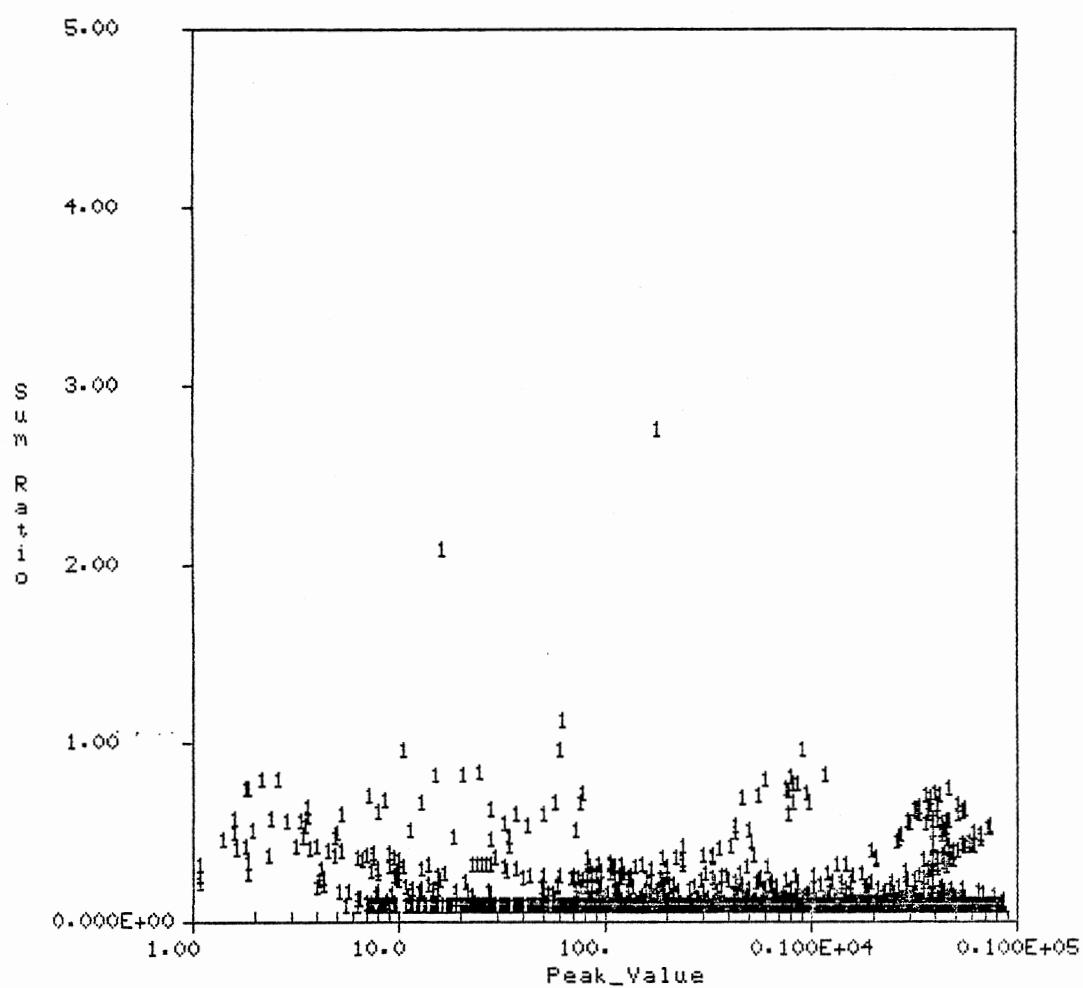
Figure 54. Peak_Value Versus Sum_Ratio for
Voiced Speech
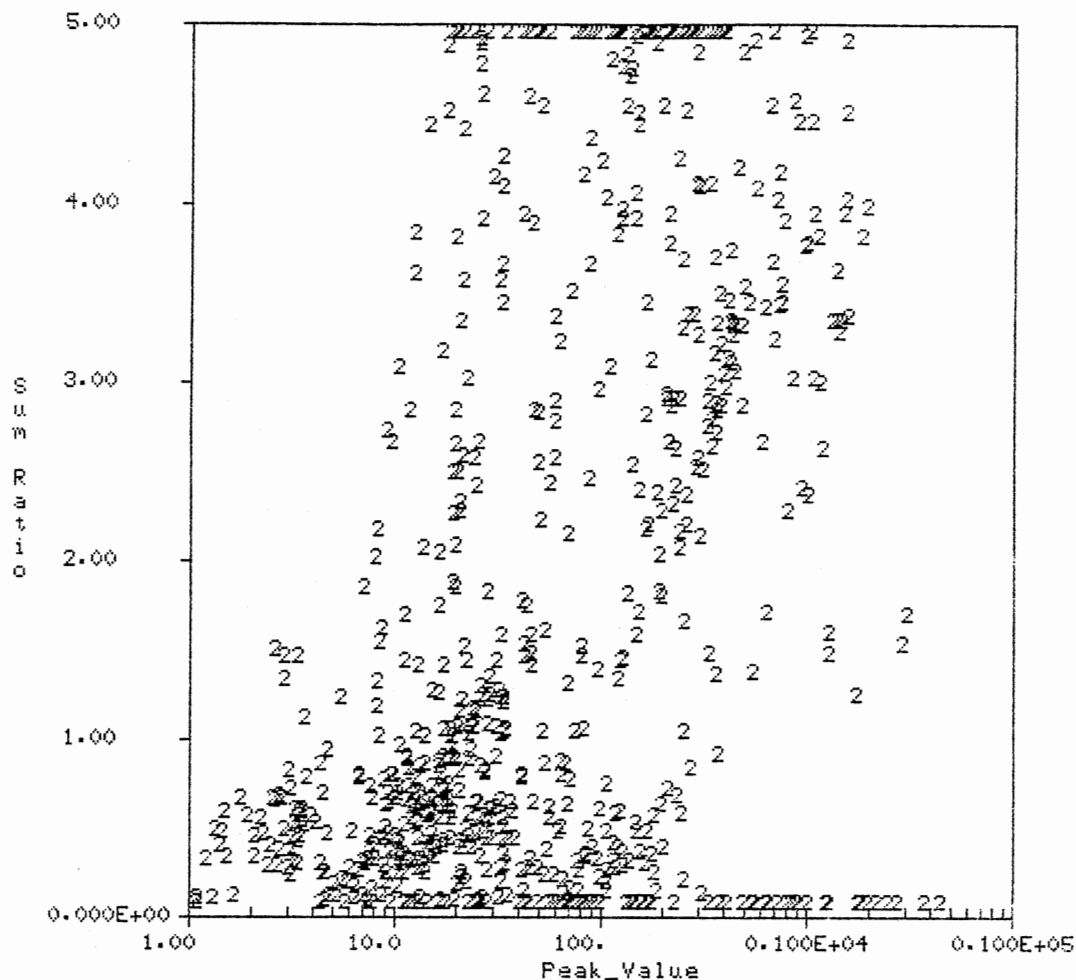
Figure 55. Peak_Value Versus Sum_Ratio for
Unvoiced Speech

Now, one of the reasons for choosing the sentences in Table II as

a test set, is their high Signal to Noise Ratio (SNR). This means that

Figure 53 through 55 could be some what misleading, since the speech

is virtually noise free. Thus, before designing a classifier, we need

to check the effect of noise on the features. Guassian noise (giving

an SNR of 20 dB) is therefore added to each of the waveforms. The TDAAV

is computed for each of the waveforms and features extracted. These

features are added to the feature files and each class sorted out. Cross plots of the features for the silence, voiced speech and unvoiced speech are included in Figures 56, 57 and 58, respectively. By separating out the classes and plotting them separately, we can see that the classes still separate. However, the distribution of the features has changed considerably.

In spite of the effect of noise, we can still see that the classes separate in our features space. It should therefore be possible to divide up the feature space in such a way as to classify the speech segments to their appropriate classes. One of the simpler ways to divide the feature space is by linear function. Now, in our case, we are working in a log-linear plane, which means the actual lines in the feature space are exponential in character. This however is not that important, for in actuality all we are doing is redefining the peak_value feature to be the logarithm of the peak_value.

The classifier will be based upon three linear discriminate functions. These functions were chosen by first estimating lines which appeared to separate the classes, estimating the parameters of those lines and then testing the resulting classifier against the training set. The results of multiple attempts of the above procedure are shown graphically in Figure 59, along with a crossplot of the features for the silence samples for both the low noise and the 20 dB case. The regions shown in Figure 59 can be expressed in the form of three decision functions, given by

$$d_1 = \text{Peak\_value} - 20. \tag{4.9a}$$

$$d_2 = \text{Sum\_ratio} - 0.2*\text{Log}_{10}(\text{Peak\_value}) - 0.146 \tag{4.9b}$$

$$d_3 = \text{Sum\_ratio} + 0.75*\text{Log}_{10}(\text{Peak\_value}) - 3.0 \tag{4.9c}$$
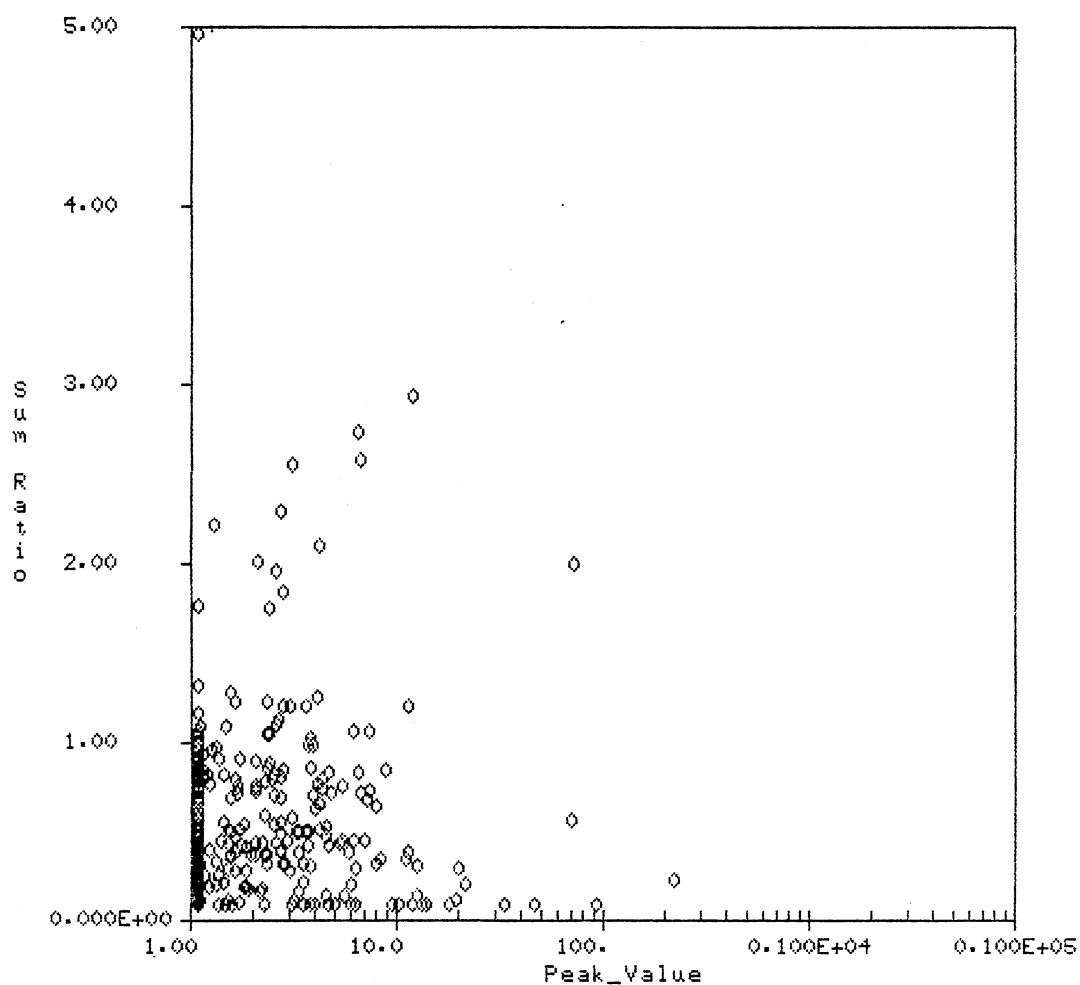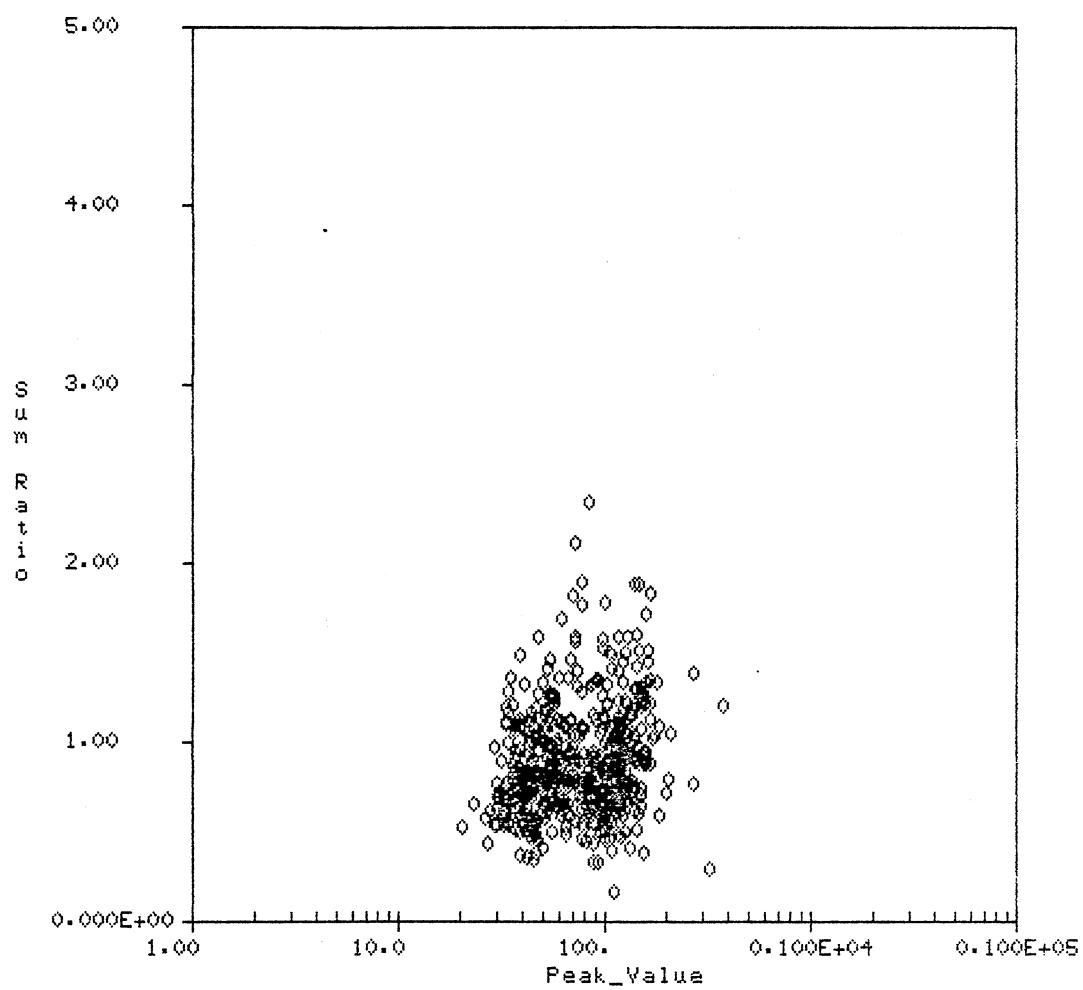
Figure 56. Peak_Value Versus Sum_Ratio for
Noisy Silence
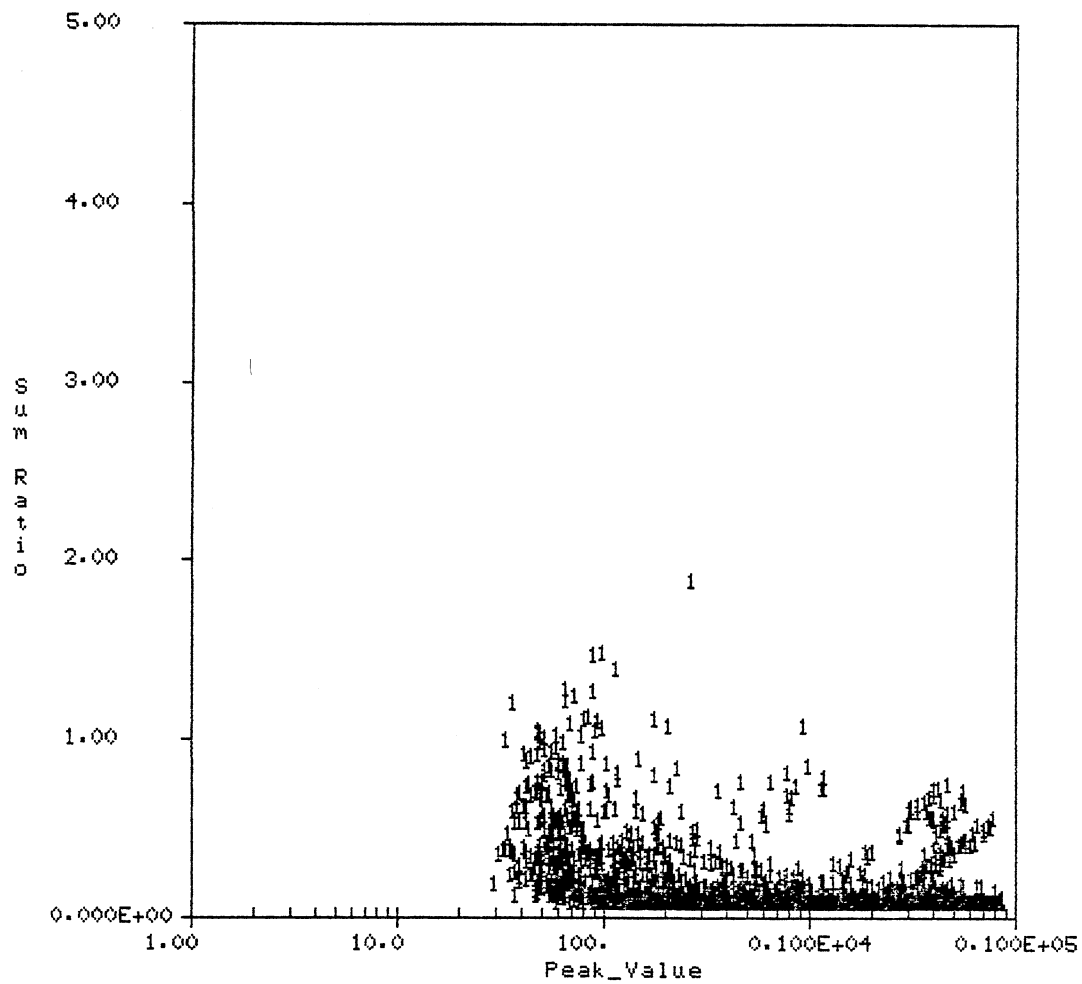
Figure 57. Peak_Value Versus Sum_Ratio for
Noisy Voiced Speech
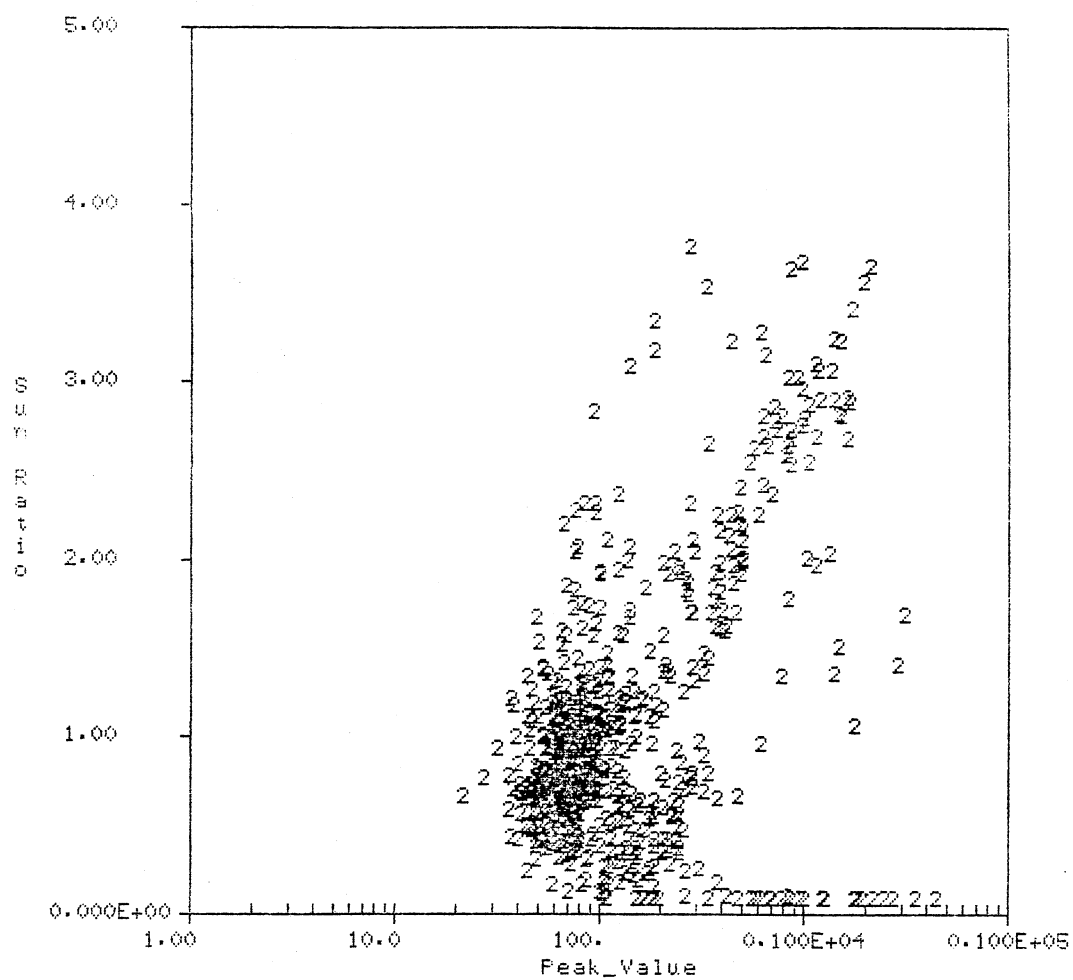
Figure 58. Peak_Value Versus Sum_Ratio for
Unvoiced Noisy Speech
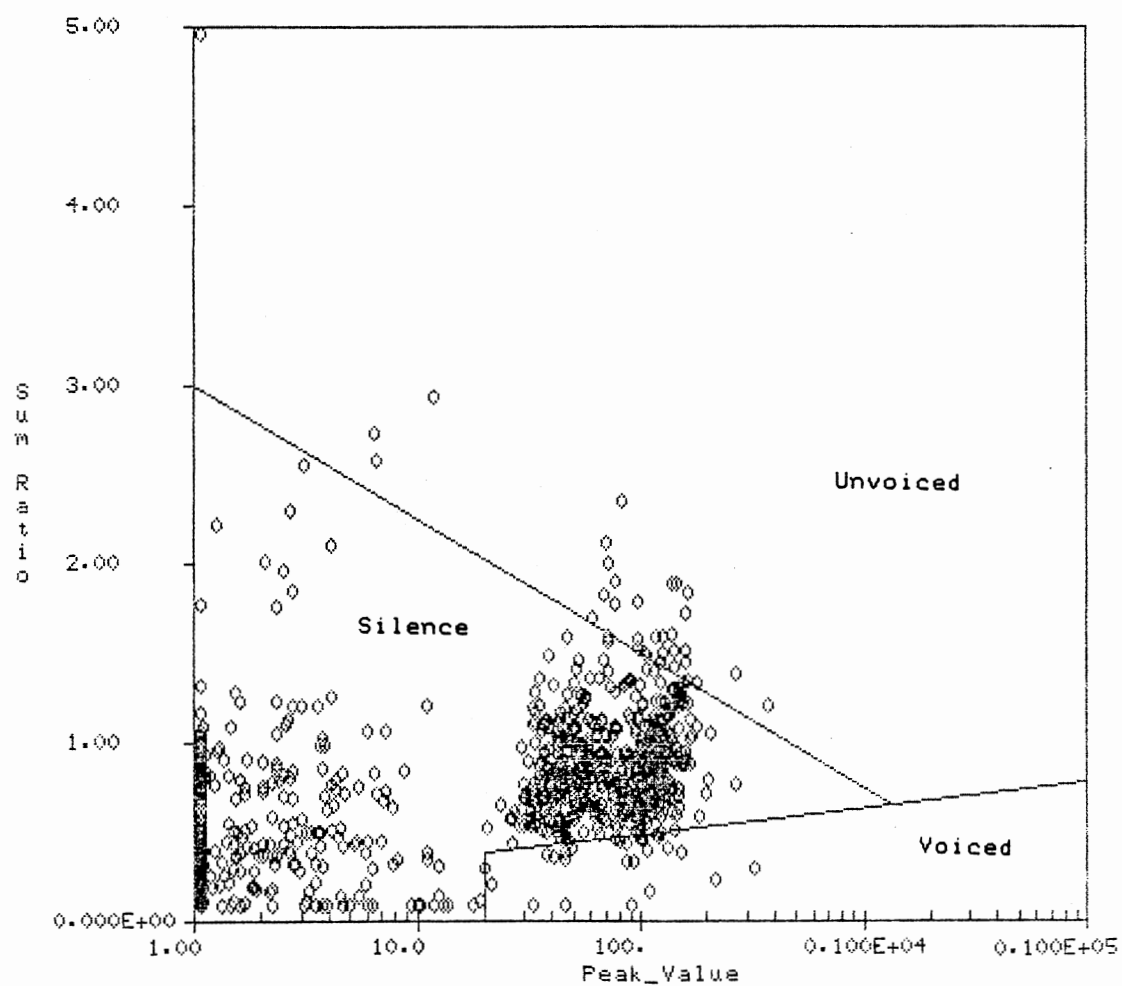
Figure 59. Crossplot with Decision Boundaries

The different classes of speech can then be recognized from these function by the following rule.

$$
Class = \begin{cases}
\text{Voiced} & \text{,for } d_1 > 0.0 \text{ and } d_2 < 0.0 \\
\text{Unvoiced} & \text{,for } d_2 > 0.0 \text{ and } d_3 > 0.0 \\
\text{Silence} & \text{, otherwise}
\end{cases} \qquad (4.10)
$$

This classifier is applied to the various sentences in our training set and error rates computed in each case. The error rate is simply the percentage of time locations where the classifier disagrees with the classification done manually. The results of these tests are shown in Table III. The error rates may seem rather large, however where these errors occur may be of more importance than the percentages. For if these errors occur during the transitions between phonemes, a large portion of them could be corrected by some type of syntactic processing of the output.

TABLE III

SPEECH CLASSIFICATION ERRORS

Sentence ID	Noise Low	Noise 20 dB
1	5.8	9.4
2	14.5	15.3
3	5.6	8.3
4	8.4	9.7
5	9.7	12.7

In an effort to display how the different segments of speech were classified, sentence three and the classifier results are plotted in Figures 60a and 60b, respectively. A classification signal is simply a time signal, set equal to 0 for silence, 1 for voiced speech and 2 for unvoiced speech. Figure 60c is a plot of the classification signal, used in the training of the classifiers. Note that a new class is included in the classification signal of Figure 60c. This class, where the classification equals 4, is for a voiced fricative, which is part of the word 'IS'. Figure 61 is similar to figure 60, except the classifier results in Figure 61b are from the classifier applied to sentence three with 20 dB SNR. The noise free speech waveform is included in Figure 61a, to allow the reader to more easily identify the various speech segments. Sentence three is chosen because it shows some excellent examples of the types of error that could occur.

The words in sentence three are 'OAK IS STRONG AND ALSO GIVES SHADE'. By studying the plots of Figures 60 and 61, we can see that the voiced segments are easily identified, however the unvoiced segments did not classify well at all. In order to exemplify the errors that arose, two words have been marked in Figures 60 and 61. The first word or utterance marked off in sentence three is the word 'OAK'. This utterance begins with a strong voiced segment and ends with an unvoiced stop. The classifier identifies the voiced segment well. On the other hand, the unvoiced stop is classified as silence. Now, an unvoiced stop is produced by closing the vocal tract, building up pressure and then releasing the air pressure. In this way, the unvoiced stop is of low energy except for a quick burst at the end. Since it is of such low energy, contextual information is probably required to detect the presence of
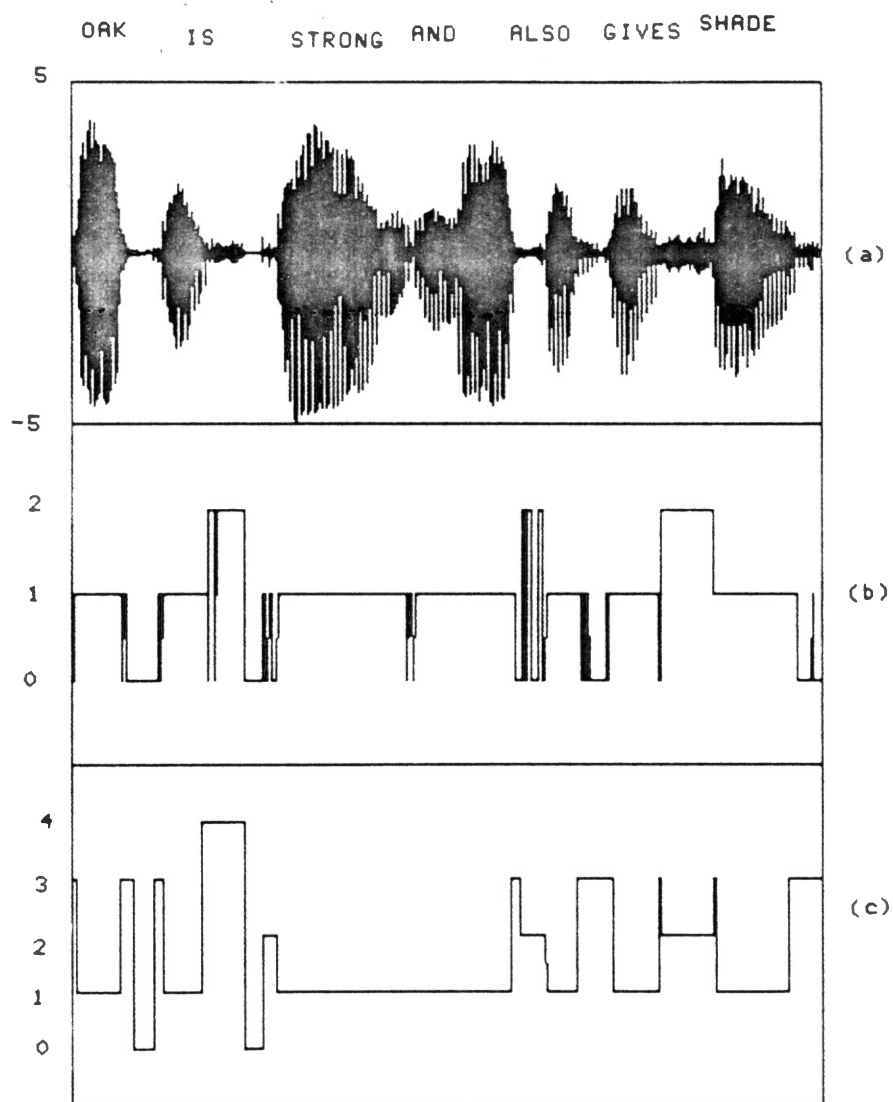
Figure 60. Sentence Three and Classifier Results

OAK   IS   STRONG   AND   ALSO   GIVES SHADE
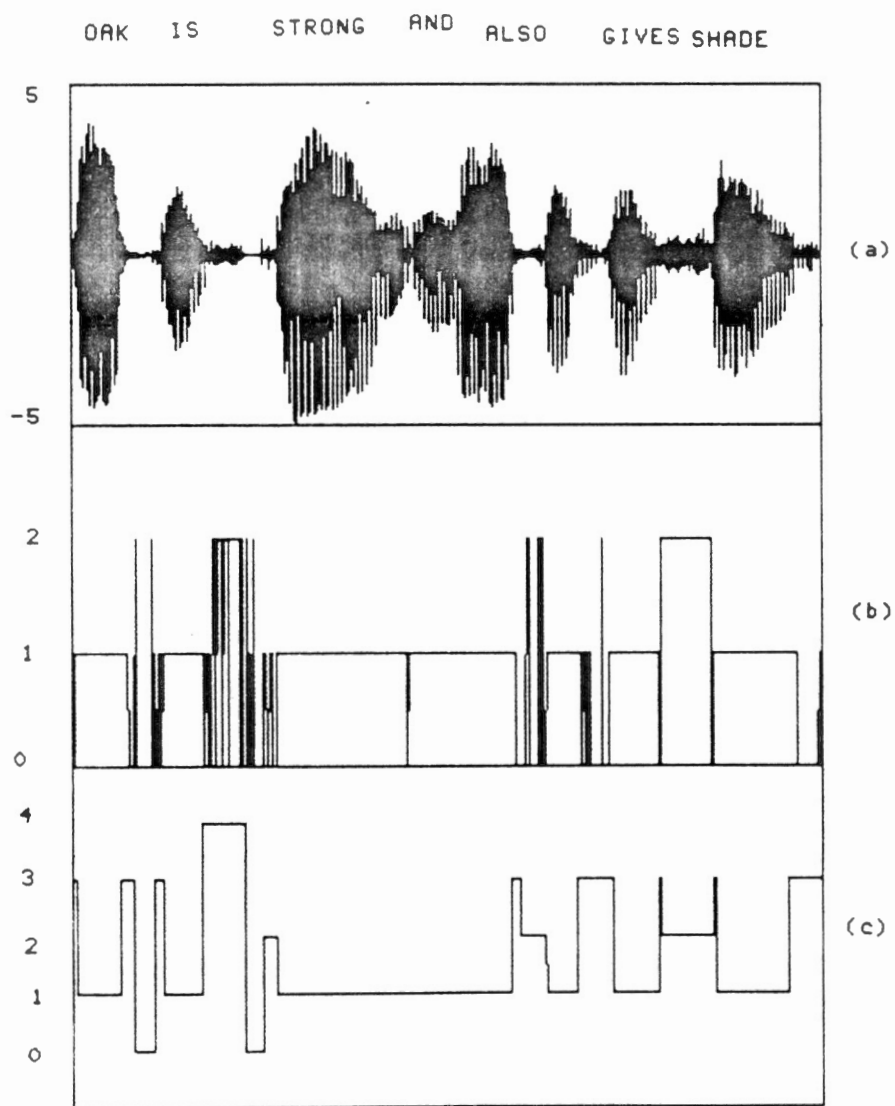
Figure 61. Sentence Three and Classifier Results
for Noisy Sentence Three Waveform

unvoiced segments such as these. For example, the algorithm used by
Rabiner and Sambur [68] employed short-time energy to detect the voiced
segment. Then assuming that the segments preceding and following the
voiced segment are probably unvoiced, the zero crossing rate of these
regions is analyzed in an effort to detect these weak phonemes. A sim-
ilar procedure, applied to our features, could improve the detection
of phonemes, such as the unvoiced stop.

Now the last word in the sentence is 'SHADE'. This word begins with
an unvoiced fricative /sh/. This phoneme contains more energy than most
of the other unvoiced segments in this example, and has therefore been
classified quite distinctly as unvoiced in both the low noise and noisy
samples. Note that the transition to the voiced segment is classified as
silence. This region, like the unvoiced stop, requires additional pro-
cessing to to be classified correctly.

Finally, we need to consider the voiced fricative. This phoneme is
produced by constricting the vocal tract, as is commonly done in unvoic-
ed speech, and then forcing air through the glottis, causing the vocal
cords to vibrate periodically as in the case of voiced speech. In this
way, the phoneme has a spectral character similar to both voiced and un-
voiced speech. The primary energy in the waveform lies above two KHz,
and it is therefore expected that the segment would appear as unvoiced
to our algorithms. Note that this exception was trapped in the earlier
test of the classifier and not included in the computation of the error
rates.

4.5 Conclusions

In this chapter, applications of the MAWD and TDAAV for the detec-

tion and recognition of various acoustic signals are presented. The MAWD

is applied to the analysis of time traces from acoustic well logs, for

the detection of the shear wave arrival. The TDAAV is applied to speech

for the classification of speech as voiced, unvoiced or silence.

A feature, which measures the flatness of the MAWD, is defined for

detecting the frequency spreading associated with the arrival of waves

in the MAWD. This feature is applied to a collection of actual acoustic

well log traces, and a log produced from the results. This log is then

compared to a log derived from cross spectral techniques.

Three different features are defined for the analysis of the TDAAV.

One of these features is a measure of the energy in the signal, and the

other features measure the distribution of that energy in the frequency

domain. These features are computed for a collection of waveforms, which

has both low noise and noisy samples. From the distribution observed

for these features, a classification scheme is developed and applied to

example waveforms.

CHAPTER V

CONCLUSION

5.1 Results

The purpose of this research is to develop time-frequency analysis techniques and demonstrate their application to acoustic signals. The two analysis techniques investigated are the Wigner Distribution and the Two Dimensional Alternate Allan Variance (TDAAV).

For the case of the Wigner Distribution, a discrete version of the WD, which allows for more efficient computation, is introduced. This new definition, called the Modified Auto Wigner Distribution (MAWD), is shown to retain many of the properties of the classical WD. The Modified Cross Wigner Distribution (MCWD), Discrete MAWD (DMAWD) and Discrete MCWD (DMCWD) are also introduced, along with some of the basic properties of each. Problems associated with the bilinear nature of the WD are noted. One common problem noted is the occurance of artifacts, commonly called cross terms, in the time-frequency plane. Two techniques for the reduction of these cross-terms are discussed. The primary advantage of our Modified definition is the ability to compute the summation, in the WD, using FFT techniques.

A Generalized Allan Variance (GAV) is introduced. It is pointed out that the basic character of the GAV is that of a frequency selective energy measure. The frequency character of the GAV is shown to be dependent upon a kernel function, which in turn is dependent upon the

setting of two parameters, called m and n. Example kernel functions are generated and displayed. Also discussed are techniques for estimating the GAV for a general signal. These techniques are based on estimating the short-time energy using properly chosen window functions. By computing the GAV for a range of values of n, a computationally efficient coarse mapping of the signal's energy distribution in the time frequency plane is obtained. By producing a two dimensional array of the various GAV's, each column representing the GAV for a different value of n, we are able to display surface plots of the TDGAV. The primary advantage of the TDGAV and TDAAV is their computational simplicity; they do however provide a good characterization of the time-frequency nature of a signal.

Applications of the Wigner Distribution and TDAAV are presented in Chapter IV. The Wigner Distribution is applied to the analysis of time traces from acoustic well logs, for the detection of the shear wave arrival. The TDAAV is applied to speech for the classification of speech as voiced, unvoiced or silence. Features and classification schemes are developed for each application and were tested against real data. The primary purpose being to demonstrate the utility of both the MAWD and the TDAAV.

5.2 Future Work

Several different areas are introduced in this work and each area can be considered as open for more research. This section contains some thoughts of what we consider as possible extensions of this research.

In the area of the MAWD, further work is needed into the description of cross-terms. These descriptions should be geared towards the

development of algorithms for reducing the effects of cross terms. A study of the effect of time windows on the MAWD and how they relate to the reduction of cross-terms is also needed.

In the case of the TDAAV, research is needed into alternate ways of defining the parameters m and n. It is noted that the Alternate Allan Variance (AAV) provides good separation between the various frequency kernels, however the length of the mean and comb filters become some what excessive. An alternative to the AAV could be to set the parameter m (the order of the comb filter) and then use only those values of n which provide good separation of the kernels. An example might be to set m equal to one and then compute the General Allan Variance (GAV) for n equal to one, two, four, eight and twelve. Of course, application of the GAV to other types of signals is another possible area of research.

The features proposed for the analysis of both the MAWD and the TDAAV are by no means the only useable features and other features should be considered. Also, the use of syntactic information in the classification of both the shear wave arrival in the acoutic well log and the voiced-unvoiced decision in speech should be considered.

Another area which could prove interesting, is the reconstruction of the various wavelets in the wavetrain. This can be achieved by first detecting the beginning and body of the wavelet in the MAWD, segmenting out that portion of the MAWD and then reconstruct the wavelet from this portion of the MAWD. A synthesis technique, for constructing a signal based upon a least squares match to a given WD, has been proposed by Boudreaux-Bartels and Parks [10].

REFERENCES

1. Allan D.W., "Statisitcs of Atomic Frequency Standards," Proc. of the IEEE, Vol. 54, no 2, pp. 221-230, February 1966.

2. Allan D.W. and Barnes J.A., "A Modified Allan Variance with Increased Oscillator Characterization Ability," Proc. 35th Ann. Freq. Control Symposium, pp. 470-475, USAERADCOM, Ft. Monmouth, NJ, May 1981.

3. Aron J., Murray J. and Seeman B., "Formation Compressional and Shear Interval-Transit-Time Logging by Means of Long Spacings and Digital Techniques, 53rd Annual Conference and Exhibition of SPE of AIME, SPE 7446, 1978.

4. Atal B.S. and Rabiner L.R., "A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Applications to Speech Recognition," IEEE Transactions on ASSP, Vol. ASSP-24, No. 3, pp. 201-212, June 1976.

5. Barnes J.A., Chi A.R., Cutler L.S., Healey D.J., Leeson D.B., McGunigal T.E., Mullen J.A., Smith W.L., Syndor R.L., Vessot R.F.C. and Winkler G.M.R., "Characterization of Frequency Stability," IEEE Trans. on Instrumentation and Measurement, Vol. IM-20, No. 2, pp. 105-120, May 1977.

6. Bellanger M., Digital Processing of Signals, John Wiley and Sons, New York, New York, 1984.

7. Berry M.V., "Semi-Classical Mechanics in Phase Space: A Study of Wigner's Function," Phil. Trans. Roy. Soc., Vol. 287, pp. 237-271, 1977.

8. Biot M.A., "Propagation of Elastic Waves in a Cylinderical Bore Containing a Fluid," Journal of Applied Physics, Vol. 23, No. 9, pp. 997-1005, September 1952.

9. Bouachache B. and Flandrin P., "Wigner-Ville Analysis of Time-Varying Signals," Proc. IEEE ICASSP, pp. 1329-1332, Paris, France, 1982.

10. Boudreaux-Bartels G.F. and Parks T.W., "Time Varying Filtering and Signal Estimation Using Wigner Distribution Synthesis Techniques," IEEE Transactions on ASSP, Vol. ASSP-34, No. 3, pp. 442-451, June 1986.

11. Boudreaux-Bartels G.F., "Time-Frequency Signal Processing Algori-

thms: Analysis and Synthesis using Wigner Distributions," Ph.D. dissertation, Rice University, Houston, TX, December 1983.

12. Boudreaux-Bartels G.F. and Parks T.W., "Reducing Aliasing in the Wigner Distribution Using Implicit Spline Interpolation," Proc. IEEE ICASSP, pp. 1483-1441, Boston, Mass., April 1983.

13. Chan David S., "A Non-Aliased Discrete-Time Wigner Distribution for Time Frequency Signal Analysis," Proc. IEEE ICASSP, pp. 1333-1336, Paris, France, 1982.

14. Chen S.T. and Willen D.E., "Shear Wave Logging in Slow Formations," SPWLA, 25th Annual Logging Symposium, Paper DD, New Orleans, Louisiana, June 1984.

15. Cheng C.H. and Toksoz M.N., "Determination of Shear Wave Velocities in 'Slow' Formations," SPWLA, 24th Annual Logging Symposium, Paper V, Calgary, Alberta,.Canada, June 1983.

16. Cheng C.H., Toksoz M.N. and Tubman K.M., "Determination of Shear Wave Velocities in 'Soft' Formations," 53rd SEG Conference, Las Vegas, Nevada, 1983.

17. Cheng C.H., Toksoz M.N. and Willis M.E., "Determination of In Situ Attenuation from Full Waveform Acoustic Logs,", J. Geophys. Research, Vol. 87, pp. 5477-5484, July 10,1982.

18. Cheng C.H. and Toksoz M.N., "Elastic Wave Propagation in Fluid-Filled Borehole and Synthetic Acoustic Logs," Geophysics, Vol. 46, No. 7, pp. 1042-1053, July 1981.

19. Claasen T.A.C.M. and Mecklenbrauker W.F.G., "The Wigner Distribution - A Tool for Time-Frequency Signal Analysis Part I: Continuous-Time Signals," Philips Journal of Research, Vol. 35, No. 3, pp. 217-250, 1980.

20. Claasen T.A.C.M. and Mecklenbrauker W.F.G., "The Wigner Distribution - A Tool for Time-Frequency Signal Analysis Part II: Discrete-Time Signals," Philips Journal of Research, Vol. 35, No. 4/5, pp. 276-300, 1980.

21. Claasen T.A.C.M. and Mecklenbrauker W.F.G., "The Wigner Distribution - A Tool for Time-Frequency Signal Analysis Part III: Relations with other Time-Frequency Signal Transformations," Philips Journal of Research, Vol. 35, No. 6, pp. 372-389, 1980.

22. Claasen T.A.C.M. and Mecklenbrauker W.F.G., "Time-Frequency Signal Analysis by Means of the Wigner Distribution," Proc. IEEE ICASSP, pp. 69-72, 1981.

23. Claasen T.A.C.M. and Mecklenbrauker W.F.G., "The Aliasing Problem in Discrete-Time Wigner Distribution," IEEE Transactions on ASSP, Vol. ASSP-31, No. 5, pp. 1067-1072, October, 1983.

24. Claasen T.A.C.M. and Mecklenbrauker W.F.G., "On the Time-Frequency Discrimination of Energy Distributions: Can They Look Sharper than Heisenberg?" Proc. IEEE ICASSP, Paper 41B.7, San Diego, CA, 1984.

25. Cohen L., "Generalized Phase-Space Distribution Functions," J. of Math Physics, Vol. 7, pp. 781-786, 1966.

26. Cohen L. and Posch T.E., "Positive Time-Frequency Distribution Functions," IEEE Trans. on ASSP, Vol. ASSP-33, No. 1, pp. 31-38, Feb. 1985.

27. Day D.D. and Yarlagadda R., "The Modified Discrete Wigner Distribution and Its Application to Acoustic Well Logging," Acoustical Imaging, Vol. 13, pp. 293-313.

28. Dennis J.R. and Yang S.Y., "Real Time Shear Wave Logging," SPWLA 25th Annual Logging Symposium, Paper EEE, New Orleans, Louisiana, June 1984.

29. Drago P.G., Molinari A.M. and Vagliani F.C., "Digital Dynamic Speech Detectors," IEEE Transaction on Communications, Vol. COM-26, No. 1, pp. 140-145, January 1978.

30. Duda R.O. and Hart P.E., Pattern Classification and Scene Analysis, John Wiley and Sons, New York, New York, 1973.

31. Flandrin P., "Some Features of Time Frequency Representation of Multicomponent Signals," Proc. IEEE ICASSP, Paper 41B.4, San Diego, CA, 1984.

32. Flandrin P. and Martin W., "Psuedo-Wigner Estimators," IEEE ASSP Spectrum Estimation Workshop II, pp. 181-185, Tampa, Fl, 1983.

33. Fukunaga K., Introduction to Statistical Pattern Recognition, Academic Press, New York, New York, 1972.

34. Gabor D., "Theory of Communication," J. IEE (London), Vol. 93(III), pp. 429-457, November 1946.

35. Gonzalez R.C. and Wintz P., Digital Image Processing, Addison-Wesley Publishing Co., Reading Massachusetts, 1977.

36. Gregory A.R., "Fluid Saturation Effects on Dynamic Elastic Properties of Sedimentary Rocks," Geophysics, Vol. 41, pp. 895-913, October 1976.

37. Harris F.J., "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," Proceedings of the IEEE, Vol. 66, No. 1, pp. 51-83, January 1978.

38. Helstrom C.W., "An Expansion of a Signal in Gaussian Elementary Signals," IEEE Transactions on Information Theory, Vol. IT-12, pp. 81-82, 1966.

39. Hsu P., _Fourier Analysis_, Unitech Divison, Associated Educational Services Corporation, New York, New York, 1967.

40. Ingram J.D., Morris C.F., MacKnight E.E., and Parks T.W., "Shear Velocity Logs Using Direct Phase Determination:" _51st SEG Conference_, Los Angeles, California, 1981.

41. Jacobson L. and Wechsler H., "The Composite Pseudo Wigner Distribution (CPWD): A Computable and Versatile Approximation to the Wigner Distribution (WD)," _Proc. IEEE ICASSP_, pp. 254-256, Boston, MA, April 1983.

42. Janse C.P. and Kaizer A.J.M., "Time-Frequency Distribution of Loudspeakers: The Application of the Wigner Distribution," _J. Audio Eng. Soc._, 31(4), pp. 198-223, April 1983.

43. Janssen A.J.E.M., "Gabor Representation and Wigner Distribtion of Signals," _Proc. IEEE ICASSP_, Paper 41B.2, San Diego, CA, 1984.

44. Kimball C. and Marzetta T., "Semblance Processing of Borehole Acoustic Array Data," _Geophysics_, Vol. 49, pp. 274-281, March 1984.

45. Kodera K., Gendrin R. and de Villedary C., "Analysis of Time Varying Signals with Small BT Values," _IEEE Transactions on ASSP_, Vol. ASSP-62, No. 1, pp. 64-76, February 1978.

46. Koenig R., Dunn H.K. and Lacy L.Y., "The Sound Spectrograph," _Journal of the Acoustical Society of America_, Vol. 18, pp. 19-49, 1946.

47. Kou F.F. and Kaiser J.F., _System Analysis by Digital Computer_, John Wiley and Sons Inc., New York, New York, 1966.

48. Lesage P. and Audoin C., "Estimation of the Two Sample Variance with a Limited Number of Data," _Proc. 31st Annual Frequency Control Symposium_, pp. 311-318, 1977.

49. Leslie H.D. and Mons F., "Sonic Waveform Analysis: Applications," _SPWLA_, Twenty-third Annual Logging Symposium, Paper GG, Corpus Christi, Texas, 1982.

50. Martin W., "Measuring the Degree of Non-stationarity by Using the Wigner-Vile Spectrum," _Proc. IEEE ICASSP_, Paper 41B.3, San Diego, CA, 1984.

51. Martin W., "Time-Frequency Analysis of Random Signals," _Proc. IEEE ICASSP_, pp. 1325-1328, Paris, France, 1982.

52. Martin W. and Flandrin P., "Wigner-Ville Spectral Analysis of Non-stationary Processes," _IEEE Transactions on ASSP_, Vol. ASSP-33, No. 6, pp. 1461-1470, December 1985.

53. Minear J.W. and Fletcher C.R., "Full-Wave Acoustic Logging," <u>SPWLA</u>, 24th Annual Logging Symposium, Paper EE, Calagary, Alberta, Canada, June 1983.

54. Oppenheim A.V. and Schafer R.W., <u>Digital Signal Processing</u>, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1975.

55. Page C.H., "Instantaneous Power Spectra," <u>Journal of Applied Physics</u>, Vol. 23, No. 1, pp. 103-106, January 1952.

56. Paillet F.L., "Acoustic Propagation in the Vicinity of Fractures Which Intersect a Fluid Filled Borehole," <u>SPWLA</u>, 21st Annual Logging Symposium, Paper DD, July 1980.

57. Paillet F.L., "Frequency and Scale Effects in the Optimization of Acoustic Waveform Logs," <u>SPWLA</u>, 24th Annual Logging Symposium, Paper V, Calgary, Alberta, Canada, June 1983.

58. Paillet F.L., "Predicting the Frequency Content of Acoustic Waveforms obtained in Boreholes," <u>SPWLA</u>, 22nd Annual Logging Symposium, Paper SS, Mexico City, Mexico, June 1981.

59. Paillet F.L. and White J.E., "Acoustic Modes of Propagation in the Borehole and their Relationship to Rock Properties," <u>Geophysics</u>, Vol. 47, No. 8, pp. 1215-1228, August 1982.

60. Papoulis A., <u>Signal Analysis</u>, McGraw-Hill, New York, New York, 1977.

61. Parks T.W., McClellan J.H. and Morris C.F., "Algorithms for Full Waveform Sonic Logging," <u>Proc. 2nd ASSP Workshop on Spectral Estimation</u>, Tampa, Florida, pp 186-191, 1983.

62. Parks T.W., Morris C.F. and Ingram J.D., "Velocity Estimation from Short-Time Temporal and Spatial Frequency Estimates," <u>Proc. IEEE ICASSP</u>, pp. 399-402, Paris, France, 1982.

63. Peterson E.W., "Acoustic Wave Propagation along a Fluid-Filled Cylinder," <u>Journal of Applied Physics</u>, Vol. 45, No. 8, pp. 3340-3350, August 1974.

64. Pickett G.R., "Acoustic Character Logs and their Applications in Formation Evaluation," <u>Journal of Petroleum Technology</u>, pp. 659-667, June 1963.

65. Potter R.K., Kopp G.A. and Green H.G., <u>Visible Speech</u>, D. Van Nostrand Co., New York, 1947, Republished by Dover Publications, 1966.

66. Rabiner L.R. and Sambur M.R., "Application of an LPC Distance Measure to the Voiced-Unvoiced-Silence Detection Problem," <u>IEEE Transactions on ASSP</u>, Vol. ASSP-25, pp. 338-343, August 1977.

67. Rabiner L.R. and Sambur M.R.," Some Preliminary Experiments in the Recognition of Connected Digits," <u>IEEE Transactions on ASSP</u>,

Vol. ASSP-24, pp. 170-182, April 1976.

68. Rabiner L.R. and Sambur M.R., "An Algorithm for Determining the End-point of Isolated Utterances," Bell System Technical Journal, Vol. 54, No. 2, pp. 297-315, Feburary 1975.

69. Rabiner L.R. and Schafer R.W., Digital Processing of Speech Signals, Prentice-Hall, 1978.

70. Rihaczek A.W., "Signal Energy Distribution in Time and Frequency," IEEE Transactions on Information Theory, Vol. IT-14, No. 3, pp. 369-374, May 1968.

71. Robinson E.A. and Trietel S., Geophysical Signal Analysis, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1980.

72. Rosenthal L.H., Schafer R.W. and Rabiner L.R., "An Algorithm for Locating the Beginning and End of an Utterance Using ADPCM Coded Speech," Bell System Technical Journal, Vol. 53, No. 6, pp. 1127-1135, July-August 1974.

73. Scarascia S., Columbi B. and Cassinis R., "Some Experiments on Transverse Waves," Geophysical Prospecting, Vol. 24, pp. 249-568.

74. Scheibner D., "A Ray Model for Head Waves in a Fluid-Filled Bore-hole," Master Thesis, Rice University, Department of Electrical Engineering, April 1982.

75. Scott J.H. and Sena J., "Acoustic Logging for Mining Applications," SPWLA, 15th Annual Logging Symposium, Paper L, June 1974.

76. Siegfried R.W. and Castagne J.P., "Full Waveform Sonic Logging Techniques," SPWLA, 23rd Annual Logging Symposium, Paper I, July 1982.

77. Sommers G.C. and Broding R.A., "Continuous Velocity Logging," Geophysics, Vol. 17, pp. 598-614, 1952.

78. Tanner M.T. and Koehler F., "Velocity Spectra-Digital Computer Derivation and Applications of Velocity Functions," Geophysics, Vol. 34, pp. 859-881, 1969.

79. Tou J.T. and Gonzalez R.C., Pattern Recognition Principles, Addison-Wesley, Reading, Massachusetts, 1974.

80. Tsang L. and Radner D., "Numerical Evaluation of the Transient Acoustic Waveform Due to a Point source in a Fluid-Filled Bore-hole," Geophysics, Vol. 44, No. 10, pp. 1706-1720, October 1979.

81. Un C.K. and Lee H.H., "Voiced/Unvoiced/Silence Discrimination of Speech by Delta Modulation," IEEE Transactions on ASSP, Vol. ASSP-28, No. 4, pp. 398-407, August 1980.

82. Ville J.,"Theorie et Applications de la Notion de Signal Analytique," Cables et Transmission, 2 A(1), pp. 61-74, Translated into English by I. Selin, Rand Corp. REport T-92, August 1, 1958 (1948).

83. Vogel C.B., "A Seismic Velocity Logging Method," Geophysics, Vol. 17, pp. 586-597, 1952

84. Well Logging and Interpretation Techniques, Dresser Atlas, Dresser Industries Inc., 1982.

85. White J.E., Seismic Waves, McGraw-Hill, New York, New York, 1965.

86. White J.E., "Elastic Waves Along a Cylindrical Bore," Geophysics, Vol. 28, No. 3, pp. 327-333, June 1962.

87. White J.E. and Zechman R.E., "Computed Response of an Acoustic Logging Tool," Geophysics, Vol. 33, No. 2, pp. 302-310, April 1968.

88. Williams D.M., Zemanek J., Angona F.A., Dennis C.L. and Caldwell R.L., "The Long Spaced Acoustic Logging Tool," SPWLA, 25th Annual Logging Symposium, Paper T, New Orleans, Louisiana, June 1984.

89. Willis M.E. and Toksoz M.N., "Automatic P and S Velocity Determination from Full Waveform Digital Acoustic Logs," Geophysics, Vol. 48, pp. 1631-1644, December 1983.

90. Willis M.E. and Toksoz M.N., "Velocity and Attenuation Determination from Digital Full Wave Sonic Logs," Presented at 49th SEG Conference, New Orleans, 1979.

91. Wigner E., "On the Quantum Correction for Thermodymanic Equilibrium," Physics Review, Vol. 40, pp. 749-759, 1932.

92. Wilpon J.G., Rabiner L.R. and Martin T., "An Improved Word Detection Algorithm for Telephone-Quality Speech Incorporating Both Syntactic and Semantic Constraints," Bell System Technical Journal, Vol. 63, No. 3, pp. 479-497, March 1984.

93. Winbow, G.A., "How to seperate Compressional and Shear Arrivals in a Sonic Log," Presented at 50th SEG Conference, Houston, Texas, 1980.

94. Wyllie M.R.J., Gardner G.H.F. and Gregory A.R., "Some Phenomena Pertinent to Velocity Logging," Journal of Petroleum Technology, pp. 629-636, July 1961.

95. Zemanek J., Angona F.A., Williams D.M. and Caldwell R.L., "Continous Acoustic Shear Wave Logging", SPWLA, 25th Annual Logging Symposium, Paper U, New Orleans, Louisiana, June 1984.

VITA

Dwight David Day

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE MODIFIED WIGNER DISTRIBUTION AND TWO DIMENSIONAL ALLAN VARIANCE WITH APPLICATIONS TO ACOUSTIC WELL LOGS AND SPEECH

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Perry, Oklahoma, April 29, 1957, the son of F. Dwight Day and Ida Mae Day.

Education: Graduated from Mulhall-Orlando High School, Orlando, Oklahoma, in May 1975; received the Bachelor of Science degree in Electrical Engineering from Oklahoma State University in May of 1980; received the Master of Science degree in Electrical Engineering from Oklahoma State University in December 1981; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1987.

Professional Experience: Engineering Intern, Magnetic Peripherals Inc., Oklahoma City, Oklahoma, June 1980 to August 1980; Graduate Teaching Assistantship, School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, Oklahoma; September 1980 to December 1980; Graduate Research Scholarship Magnetic Peripherals Inc., Oklahoma City, Oklahoma, January 1981 to May 1981; Engineering Intern, Magnetic Peripherals Inc., Oklahoma City, Oklahoma, June 1981 to August 1981; Graduate Research Scholarship, Magnetic Peripherals Inc., Oklahoma City, Oklahoma, September 1981 to December 1981; Instructor/ Research Associate, School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, Oklahoma, January 1982 to August 1983; Systems Engineer, Texas Instruments Inc., Dallas, Texas, August 1983 to August 1985; Research Associate, School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, Oklahoma, August 1985 to May 1987.

Awards/Affiliations: Member of Eta Kappa Nu, Institute of Electrical and Electronics Engineers and Society of Professional Well Log Analysts.