A WALSH VOCODER

A WALSH VOCODER

By

STEPHEN MILUNOVIC, B.ENG.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Engineering

McMaster University

January 1975

MASTER OF ENGINEERING (1975)
(Electrical Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: A Walsh Vocoder

AUTHOR: Stephen Milunovic, B.Eng. (McMaster University)

SUPERVISORS: Dr. J. Anderson
Dr. R. DeBuda

NUMBER OF PAGES: viii, 149

ABSTRACT

The concepts of the analogue channel vocoder and digital FFT vocoder are utilized in the analysis of a Walsh vocoder. The FFT and Walsh vocoders are simulated in Fortran on a CDC6400 computer. The simulations contain a simplified cepstrum pitch detector to simulate a hardware pitch detector. The Walsh vocoder simulation is identical to the FFT vocoder simulation with the Fast Walsh Transform operation replacing the Fast Fourier Transform operations (FWT and FFT respectively) in the channel signal calculations and transforms into the time domain.

Sentences of different lengths and speakers are processed by the simulated vocoders and the resultant synthesized speech is analyzed for comparative quality and intelligibility.

An unsuccessful attempt is made to effect a vocoder which does not require a pitch detector with the calculation of the time invariant sequency power spectrum replacing the channel signal calculations in the vocoder simulations. Long term frequency and sequency envelopes are used to shape the frequency and sequency power spectra derived from the transmitted time invariant sequency power spectrum in the FFT and Walsh vocoder simulations respectively.

ACKNOWLEDGEMENTS

To Drs. J. Anderson and R. DeBuda I would like to extend my gratitude for their extended guidance and support in their supervisory capacities. I would also thank Dr. S. S. Haykin, Director of the Communications Research Laboratory, for the use of the computer facility without which this research could not have been possible, and Dr. A. R. Elliott who inspired this research. To my colleagues C. Hawkes and J. Bodie I am indebted for their altruistic aid.

To my family I am indeed grateful for their benefaction and sacrifices during my trials on the road from posteriority.

A special note of thanks to my sister Maryana who typed this manuscript, and to the National Research Council of Canada who funded this research.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I: AN INTRODUCTION

Digital techniques of voice communication have much to offer from a transmission standpoint. Digital transmissions are relatively insensitive to noise, crosstalk, distortion and faded signals can be readily regenerated without introducing cumulative degradation. A high degree of security can be obtained with digital encryption in which a pseudo-random binary sequence is used to scramble the digital signal. Modern integrated circuit technology makes all such equipment physically small and inexpensive.

Low bit rate digital speech transmission systems are required to transmit secure speech over transmission links whose bandwidth is only adequate for analog transmission. Many existing links such as switched analog telephone networks, mobile radio, high frequency radio etc., only permit transmission rates in the order of 2.4 K bit/s. (In some cases, 4.8 K bit/s bit rates maintain compatibility with multiplexed digital data links).

The most common digital speech system currently employed in commercial telephony has a bit rate of 64 K bit/s which requires high bandwidth channels such as satellite links, microwave relay stations, and optical pipes. These

new avenues of digital communications can be more efficiently utilized with low bit rate digital speech transmissions by multiplexing many transmissions into one channel.

In order to determine the lower limit of the bit rate necessary to transmit voice information, various experimental efforts have been made to assess the human capacity for processing information. None of the experimental results indicate the human to be capable of processing information at rates greater than the order of 50 bits/s [1]. These results indicate a redundancy in the analog speech waveform and lead to the investigation of methods which extract the information bearing parameters of the speech signal. This technique is used to attain bit rate compression for digital speech communication. Another method deals mainly with the transmission of the voice waveform in a more efficient manner than pulse code modulation (P C M). The original speech waveform is matched to the transmitted waveform in a mean square error (M S E) sense, or some other criterion, but attempts are made to preserve the structure of the original speech signal. This waveform coding method does not utilize knowledge of the physiology of the human vocal and aural apparatus and consequently is less efficient than the former techniques of parameter extraction.

1.1 WAVEFORM CODING METHODS

Two techniques are used in this category to effect waveform preservation: voice waveform digitization methods and orthogonal function expansions of the voice waveform.

1.1.1 VOICE WAVEFORM DIGITIZATION METHODS

The most elementary waveform coding method, P C M, was invented in 1937 by Reeves [2]. The speech waveform is sampled at the Nyquist rate and these samples are encoded into binary symbols which are transmitted to the receiver where the samples are reconstructed with amplitudes given by the binary numbers. The number of binary digits used to encode the samples and reconstruct the amplitude of the waveform determines the accuracy of the system.

Due to the large amplitude range of the speech waveform, various companding methods have been employed. In practical telephone P C M the logarithmic scale is utilized although other nonlinear scales are possilbe. Voice waveform distortion is decreased without increasing the number of quantization levels.

The amplitude samples of the speech waveform have varying degrees of correlation depending on the sample density and on the cutoff frequency of the filters used to band-limit the speech signal prior to digitization [3]. When adjacent samples are highly correlated, it is clearly

redundant to transmit both samples. One method of this redundancy reduction is to encode the differences between the sample amplitudes instead of the actual sample values since the amplitude variation from sample to sample requires fewer bits on the average for encoding than the total amplitude variation. Differential P C M (D P C M) employs this general approach to bit rate reduction. It actually encodes the difference between a current amplitude sample and a predicted amplitude value estimated from past samples. These past samples are weighted so as to minimize the average energy of the difference signal as calculated from long term statistics of a representative segment of speech [4].

Because speech is a non-stationary process, the long term statistics of the signal change. The adaptive prediction coding scheme operates similarly to a D P C M system except that the weight of the past samples is changed adaptively to minimize even further the M S E between the predicted and true values of the signal [5].

If the rate at which the speech signal is sampled is increased, successive samples become more and more similar and the number of bits required to encode each sample can be reduced for a given performance until only one bit is necessary. This system is called delta modulation. From an implementational standpoint this system is simple, but suffers from slope overload which occurs whenever the

incoming signal has a rapid amplitude change. This "disadvantage" of the delta modulation scheme can be used for pitch rate detection and will be mentioned later in Chapter II, Section 2.4.2

1.1.2 ORTHOGONAL FUNCTION EXPANSIONS

With the use of orthogonal transforms, the time waveform can be expressed as a series of coefficients of a complete set of orthogonal functions. Two such transforms are the Fourier transform which represents the time waveform as a summation of sinusoidal functions and the Walsh transform which decomposes the sampled time function into a summation or orthogonal square waveforms. Another possible transform is the Karhunen-Loève transform which does not have a predetermined set of orthogonal functions. Instead the functions form a basis of the transform as derived from the diagonalization of the covariance matrix of the speech segment of samples under analysis.

The basic aim of the orthogonal function expansion is to represent the voice waveform by coefficients which ideally are uncorrelated, and thereby obtain bit rate compression. The Karhunen-Loève transform coefficients are by definition uncorrelated, but the implementation of such a transform is impractical. The Fourier and Walsh transforms do not yield completely uncorrelated coefficients, but the

coefficients are less correlated than the time samples of the speech waveform {6}. Hence, bit rate reductions are possible, but are unfortunately modest. Attempts ({6},{7}) have shown that orthogonal function expansions of the speech waveforms do not yield a high bit rate reduction for the complexity involved. The bit rate reduction with the use of the Fast Fourier Transform (F F T) is only 10.5 K bit/s and the Fast Walsh Transform (F W T) yields a 7.5 K bit/s reduction when the fidelity criterion is maintained at the same level as 56 K bit/s P C M {6}.

1.2 VOCODER METHODS

Most of the techniques described under waveform coding method produce voice signals of good quality and high intelligibility. However, the bit rate requirements of these systems are of the order of 20 K bit/s. This high bit rate is the result of their failure to extract the information laden elements in the voice waveform.

Voice coding techniques which make no attempt to preserve the original speech waveform are called vocoders. The input speech signal is analyzed in terms of standardized speech features each of which can be transmitted in digital coded forms. At the receiver, these features are reassembled and the output speech signal is synthesized.

Two sources of redundancy in the voice signal are the quasi-periodicity of the short term frequency spectrum due to the presence of "pitch" and the highly structured nature of the envelope of this spectrum. The pitch information is determined by different methods. This parameter may be isolated and transmitted as an independent feature as in the channel vocoder, or it may be contained in the baseband of the voice signal as in the voice-excited vocoder. In all cases the pitch information forms an integral feature required for good quality speech synthesis.

The structured nature of the spectral envelope is utilized in almost all vocoders to different degrees. The formant vocoder utilize the regular location of humps in the spectral envelope and can thus achieve bit rates below 1 K bit/s {7}.

Vocoders built to date are generally large, bulky, power consuming machines and only the channel vocoder is presently available as an off-the-shelf device. As large scale integrated cirucit technology advances, vocoders hold the promise of low bit rate. high intelligibility, high quality digital speech transmission techniques.

1.3 INTENT OF RESEARCH

The intent of the present research is to investigate an alternative to the existing large, bulky, expensive channel vocoder with the use of the Walsh transform. The

approach will be to replace the bank of filters in the channel vocoder with sequency domain energy calculations thereby reducing the cost, size, and weight of the channel vocoder while maintaining a 2.4 K bit/s bit rate.

The reasons for the use of Walsh functions stem mainly from the implementational ease of the F W T in real time with digital hardware. Other reasons are that the sequency spectrum of a speech signal "closely" resembles the frequency spectrum of that signal, and it is known that a channel vocoder using the F F T is a feasible digital speech communication device. The proposed Walsh vocoder will be simulated on a computer and compared with a simulated version of the F F T channel vocoder.

CHAPTER II: THE SPECTRUM CHANNEL VOCODER

In 1928, H. Dudley of Bell Telephone Laboratories devised a system which reduced the information redundancy of acoustical signals derived from natural speech. This device was termed "vocoder" from VOice CODER. This method of transmitting speech signals relies on the extraction of signal parameters which influence the intelligibility and quality of the perceived speech. There is no attempt made to transmit the voice waveform in its generated form, nor is the synthesized signal an approximation to this voice waveform. A distinction is made between signals and sounds: signals are the result of sounds activating a transducer which converts air waves produced by sounds into electrical signals. It is these signals so derived that contain the information redundancy.

The vocoder is an analysis-synthesis system. The speech signal is analyzed to extract the information bearing parameters in the signal and these parameters are then used to synthesize a speech signal, the perception of which will resemble that of the original speech sounds. Also the quality of the synthesized speech determines the possibility of speaker recongition. Hence the goals of a vocoder system

are to attain highly intelligible synthetic speech of acceptable quality while maintaining a low enough bandwidth that telephone lines can be used as communication channels.

The extraction of the desired signal parameters must be achieved with minimum complexity to reduce the cost, size, and power consumption of the physical device. The parameter extraction is an engineering design problem while the choice of parameters has physiological implications.

2.1 SPEECH PRODUCTION AND PERCEPTION

Speech sounds are produced by the conversion of a steady flow of air from the lungs into a pulsating acoustic wave. This conversion, or excitation function, is accomplished in the vocal tract, which acts as an acoustic tube, by one or a combination of the following mechanisms:

1) air escaping from the lungs passes through the vocal cords which causes the air flow to vibrate

2) constrictions in the vocal tract create turbulences in the passing air stream

3) sudden releases of pressure following a closure of the vocal tract

The first mechanism creates voiced sounds which have a quasi-periodic nature. The vowel sounds are examples of

voiced sounds. Fricative sounds like"s" and"f" are generated by the second mechanism. These sounds are termed unvoiced. The third mechanism generates unvoiced plosive sounds like "p" or "t".

These mechanisms act as sources of wideband acoustic signals and the frequency spectra of these signals are modified by the vocal tract which acts as a time varying filter with transmission properties dependent on the natural frequencies of the vocal tract.

Three basic properties of the aural mechanisms are utilized in the channel vocoder design. These are:

1) the ear performs a short-term spectral energy analysis

2) for monaural perception, the ear is relatively insensitive to phase information

3) the quality of the speech perceived by the ear is closely related to the periodicity (pitch) of the speech sounds.

The speech sounds as perceived seem to be highly dependent on pitch information and the spectral envelope of the acoustic signal. The channel vocoder utilizes this dependency to reduce the information redundancy of the signals representing natural speech.

2.2 THE SPECTRUM CHANNEL VOCODER

Because the sound sources and vocal tract shape are relatively independent, they provide two parameters necessary in the modelling of the speech production and perception systems. The fundamental period of the glottal excitation of voiced sounds determines the spacing between adjacent peaks in the frequency spectrum of the acoustic signal. This fundamental period is called the pitch period and is the parameter which greatly influences speech quality. The shape of the vocal tract determines the envelope of this spectrum.

The analyzer in the channel determines the sound sources and the spectral envelope of a speech segment. If an unvoiced sound is detected then a code is transmitted to indicate that the sound source should be modelled or synthesized using a noise generator. If a voiced sound was observed then the pitch period is determined by some means (this will be discussed in section 2.4) The envelope of the spectrum is determined by calculating the energy content of adjacent bandpass regions of the frequency spectrum. This is accomplished with the implementation of a bank of contigous bandpass filters followed by squaring and averaging devices. Figure 1 depicts a typical channel vocoder analyzer.

In general, channel vocoder analyzers contains 14 to 32 channels which are identical except for the bandpass regions of the filters. The bandwidths of the filters are determined such that over a long term average each filter will contain the same average energy. The outputs from these filters are time varying, but the variation is slow enough that the outputs need only be sampled every 20 ms. This analog vocoder analyzer realization typically consists of a large number of nearly identical circuits which function in parallel to derive the spectral envelope.

The logarithmic sensitivity of the ear allows logarithmic compression of the energy quantities. These channels are also highly correlated and this correlation is exploited by a channel to channel delta modulation scheme permitting more efficient quantization and encoding.

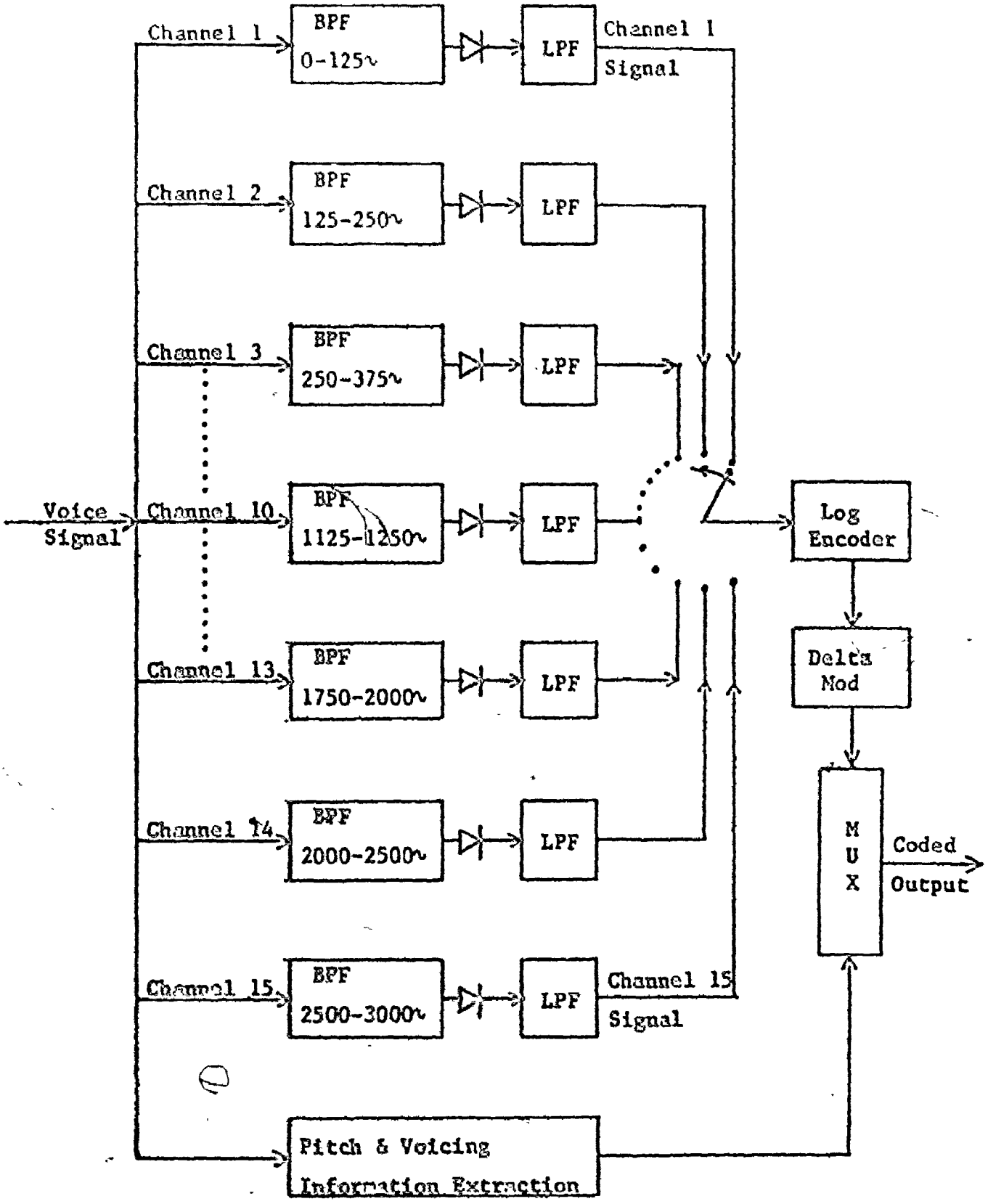
The pitch or voicing information is typically determined 100 times per second {8}, and is encoded with 8 bit words. The net result of this coding is a 32 bit representation of the frequency spectrum to which is multiplexed two 8 bit excitation parameters to form a 48 bit frame every 20 ms yielding a bit rate of 2400 bits/s.

Both the analyzer and synthesizer of a conventional channel vocoder employ identical banks of bandpass filters as indicated in Figures 1 and 2. In an analog vocoder

realization one physical bank of filters can serve both functions if the machine is of a half-duplex nature.

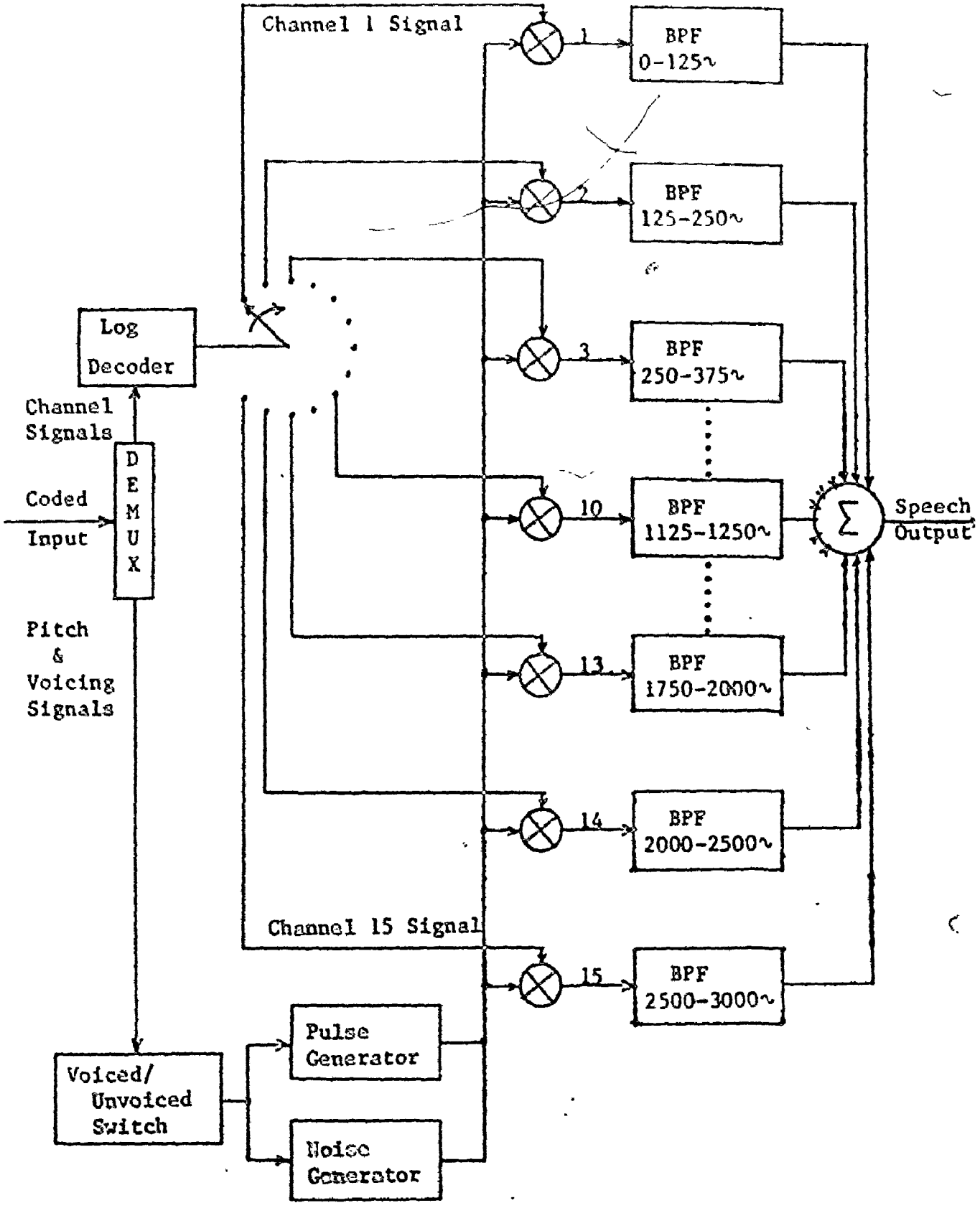
In the synthesizer the incoming 2400 bit/s data stream is decoded to yield an excitation parameter once every 10 ms. The excitation data is used to generate a train of pitch pulses at a rate corresponding to the period of the voiced sounds. If an unvoiced sound was observed then the pulse rate could be fixed at a high frequency or a noise generator could be used to excite the bandpass filters. The outputs of the bandpass filters are modulated by the channel signals determined in the analyzer, and a summation of the outputs of the modulators yields the synthesized speech signal.

The disadvantages of this analog realization of the spectrum channel vocoder stem from the implementation of the analog bandpass filters. These filters require physically large inductors and the resulting vocoder is bulky and costly (approximately 700 cu. in. at $5000 / channel [7]). A completely digital channel vocoder would be desirable to reduce the cost and size of the vocoder by replacing analog components with digital logic wich can be multiplexed to eliminate the redundancy of analog hardware. Also the entire system could then reside in digital processing machines.



BPF ≡ Bandpass Filter.

LPF ≡ Lowpass Filter 50∿



BPF ≡ Bandpass Filter

2.3 DIGITAL CHANNEL VOCODERS

There are two different methods of implementing the analog vocoder described with completely digital techniques. One method is to replace the analog bandpass filters with digital filters or one digital filter multiplexed such that each channel sequentially employs the same digital filter with a different bandpass configuration. This can be easily accomplished by altering the digital filter coefficients as required, but the size and cost of the unit are still unduly prohibitive. A more promising method is the replacement of the bank of filters with an orthogonal transform unit. One orthogonal transform employed has been the Fast Fourier Transform (F F T). This method will be briefly outlined since its structure is similar to the Walsh vocoder to be described in Chapter V.

2.3.1 THE F F T VOCODER

The standard sampling rate of speech bandlimited to 200 - 3200 Hz is 8 KHz. Let these speech samples form a sequence $x(j)$, $j \in \{0,N-1\}$ where $N=2^m$, $m \in I_+$ (m any positive integer). The discrete frequency spectrum $X(k)$ is given by

$$X(k)=1/N \sum_{j=0}^{N-1} x(j) \exp\{-i2\pi j\, k/N\},\ k=0,1,\ldots,N-1 \qquad \{1\}$$

where $i = \sqrt{-1}$. Once this transformation into the frequency

domain has been accomplished, the bandpass filters as physical devices are eliminated. To achieve the function of the bandpass filters i.e. the spectral envelope calculation, the spectral components X(k) are squared and the phase information is discarded by adding adjacent squared components. This results in a squared magnitude spectrum Y(n) given by:

$$Y(n) = X^2(2n) + X^2(2n + 1), \quad n=0,1,\ldots,N/2 - 1 \qquad \{2\}$$

Now the spectrum given by Y(n) can be divided into the bandpass regions corresponding to the regions given in the analog channel vocoder. The component Y(n) represents the squared magnitude of the spectrum at the frequency f given by:

$$f = 8000n/N \qquad \{3\}$$

for the sampling rate of 8KHz. The energy in each bandpass region can be approximated by averaging the Y(n) that fall into that bandpass region.

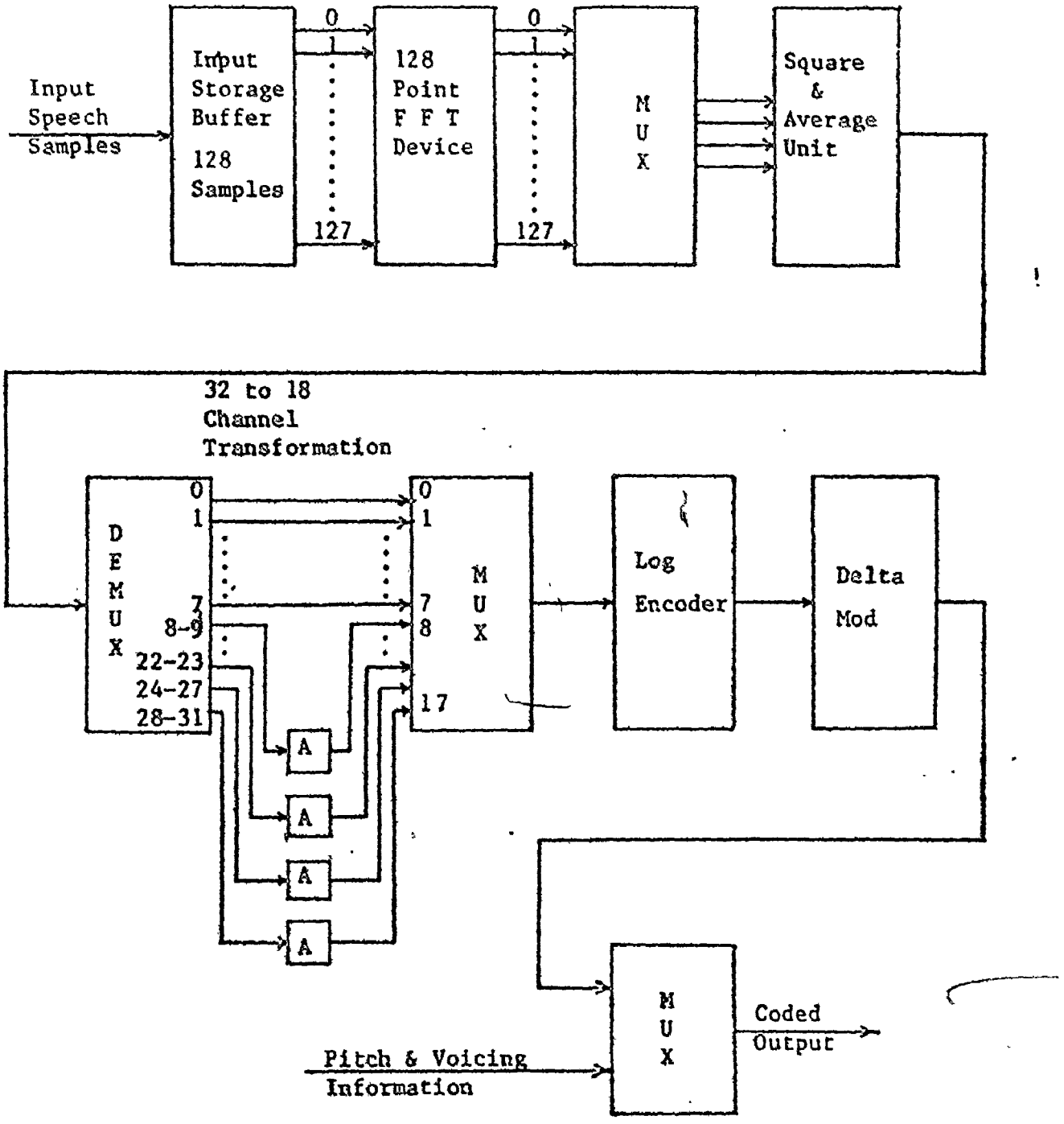
Figure 3 indicates a simple configuration for the spectral envelope detection in the analyzer section of the F F T vocoder. At first the spectrum is divided into 32 channels each of which contain a bandwidth of 125 Hz, and

the energy in these bandpass regions is calculated in a multiplexed arithmetic unit that computes the square and average operations. Since the frequency discrimination of the aural mechanism is greater at low frequencies, the higher frequency channels are averaged together by 2's and 4's to reduce the number of channels to 18. This combining of adjacent bandwidth leads to a structure equivalent to an 18 channel vocoder with non-equal bandwidth filters. It is computationally more efficient to first compute 32 equal bandwidth F F T parameters and then selectively combine them rather than to realize the nonuniform filter characteristics of the channel vocoder.

The pitch detection in the F F T vocoder can be accomplished in the same manner as in the analog channel vocoder, independently of the spectral calculations. The pitch information can again be multiplexed with the channel signals.

The synthesizer of F F T vocoder is illustrated in Figure 4. The structure of the synthesizer is similar to the analyzer in reverse. The incoming data stream is demultiplexed and decoded to yield 32 channel signals and a voicing parameter.

The channel signals are spread to form an N point frequency spectrum where the phase information is set to



A = Averaging Device

FIGURE 3 ; Typical F F T Digital Vocoder Spectral Envelope Detector

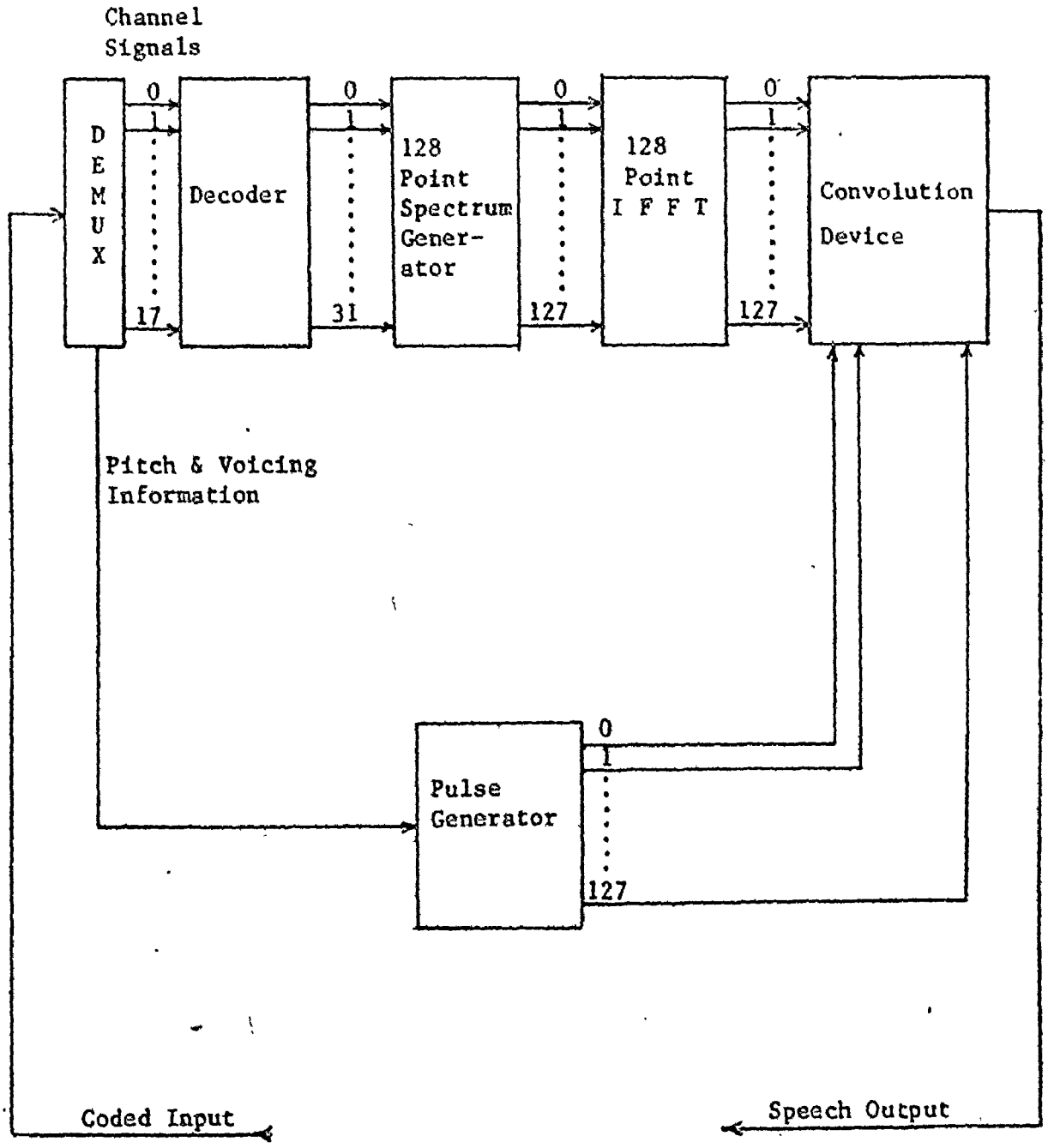


FIGURE 4 : Typical F F T Digital Vocoder Synthesizer

zero. The weightings of the spectral coefficients are the channel value in the synthesizer. The inverse F F T (I F F T) of this spectrum is then computed and the resulting time function is equivalent to the impulse response function of a bank of filters. This time function is circularly shifted by N/2 samples (or equivalently a linear phase added) to allow the major peak to appear at the center of the waveform and minimize discontinuities at the edges {8}. This is accomplished by reversing the phase of every second spectral component in the generated frequency spectrum.

The impulse response so constructed must be convolved with a train of pitch pulses generated at a rate given by the voicing parameter. If the voicing parameter indicates an unvoiced segment then a 1 ms pitch period is used and attenuated pitch pulses of random sign are produced. Otherwise a pulse train of unit magnitude is used at a pitch rate given by the voicing parameter.

The convolution of the impulse response function and pitch pulses is accomplished with the addition of output components given by this impulse response and delayed versions of the same impulse response where the delays are given by the pitch period. In this manner, the output at any particular time is given by the sum of the output components at that time. The impulse response function is updated at

one half the rate of the pitch period update since the spectral envelope of the speech segments change at approximately one half the rate of excitation changes. This type of synthesizer is called a convolution synthesizer.

The above convolution could be accomplished with the multiplication of the generated N point frequency spectrum with the F F T of the train of pitch pulses followed by an I F F T of this product. However the F F T device would have an additional computation to perform which could make the system too slow for real time operation since the F F T device is the slowest resource in the system.

Modifications are made to the procedure outlined for the analyzer and synthesizer of the spectrum channel vocoder to obtain smoother transitions between speech segments. Overlapping data windows are used so that the spectral measurements are obtained by averaging successive F F T's in order to achieve smoother spectral estimates. The spectral envelope used in the impulse function determination in the synthesizer is updated more often than 20 ms by linearly extrapolating between successive transmitted spectral envelopes. These methods are described by Bially and Anderson {8}.

The F F T vocoder contains the advantages of an all digital system. The disadvantages lie mainly in the complexity of the F F T calculations which require complex arith-

metic and multiplication operations. It is because of this complexity that a different transform, which can be more easily accomplished in real time with digital processing hardware, is investigated.

2.4 METHODS OF PITCH DETERMINATION

In order to determine the pitch period of a speech signal, the time periodicity of the source signal must be obtained from the observed speech signal. Also, voiced-unvoiced decisions require accurate determination of the presence or absence of periodicity in the speech signal. The design of an accurate pitch detector that works satisfactorily with bandlimited, noisy speech signals remains one of the unresolved areas of speech processing research.

There are many techniques of pitch determination designed for software environments. One of the more efficient and accurate methods is cepstrum pitch determination. This method is suitable for computer simulations of speech processings systems but is overly complex for hardware implementation. Hardware pitch detectors have been described in literature but are not off-the-shelf devices. Although pitch detector investigation is not the purpose of this work, a few words concerning cepstrum pitch determination and a hardware pitch detector are in order.

2.4.1 CEPSTRUM PITCH DETERMINATION

Since the logarithm of the amplitude spectrum of a periodic time signal with rich harmonic structure is periodic in frequency, cepstrum pitch determination consists of spectral analysis of this log amplitude spectrum. The term cepstrum refers to the spectrum of the log amplitude spectrum.

Basically the procedure for cepstrum pitch determination can be summarized as:

1) compute the power spectrum of a speech segment

2) obtain the logarithm of the amplitude of this power spectrum

3) compute the power spectrum of the log amplitude spectrum obtained in step 2), obtaining the cepstrum.

If the speech segment was voiced there will be a sharp peak occuring on the time scale called the quefrequency scale corresponding to the spectral periodicity. The location of the isolated peak on this axis represents the pitch period. The procedure is complicated with the problems of pitch doubling, window size choices and peak threshold level adjustments as well as the usual problems associated with accurate digital spectral estimation. Window weighting functions are commonly used to obtain more accurate spectral estimates.

Voiced speech signals can be modelled as the convolution of the vocal tract impulse response and the signal due to the vocal cord vibrations. Alternatively, the spectrum of the vocal source is multiplied with the transfer function of the vocal tract. The logarithmic operation separates the two spectra such that the spectral periodicity of the vocal source can be identified. The effect of the vocal tract is to produce low frequency ripples in the logarithm spectrum of the log power spectrum while the periodicity of the vocal source is manifested as high frequency ripples. Therefore, the spectrum of the log power spectrum has a sharp peak corresponding to the high frequency source ripples and a broader peak corresponding to the low frequency structure in the logarithm spectrum. A complete exposé on cepstrum pitch determination is given in a paper by A.M. Noll {9}.

2.4.2 A DIGITAL PITCH DETECTOR

The digital pitch detector described by Frei et al {10} basically compares bit patterns of adjacent windows of speech samples. The criteria for waveform comparison are slope, curvature and polarity of the speech signal. An adaptive delta coder determines the slope and curvature of the speech signal by exploiting the overload features of delta modulators and a hard limiter determines the polarity. Based on these criteria, logic operations determine the sim-

ilarity of the waveforms for various time delays. This algorithim produces a peak at a delay equivalent to the pitch period that is much sharper than that of the auto-correlation function [10].

CHAPTER III: WALSH FUNCTIONS

The fundamental paper on the mathematical theory of Walsh functions was published by J. L. Walsh [11]. Based on this foundation, later papers by Paley ([12]), Fine ([13], [14]), and Pichler ([15], [16]), for example have developed a mathematical basis for Walsh functions similar to the theory of trigonometric functions.

3.1 DEFINITIONS OF WALSH FUNCTIONS

There are basically four different methods of defining Walsh functions. The most common definition defines Walsh functions as products of Rademacher functions (which are a subset of Walsh functions). A different definition is the iterative difference equation defined by Harmuth [17]. Another definition derives the Walsh functions from the exponential functions [13]. Still another derivation is based on raising the base -1 to an exponent dependent on the binary representation of the Walsh functions desired. The common feature of these definitions is the requirement of dyadic representations. This representation will be outlined before the definitions are given.

A dyadic number is any number represented by a summation of the binary digits "1" or "0" modulating integer powers of 2. Let μ be expressed as a binary number given by (3.1).

$$\mu=\sum_{s=-\infty}^{\infty}\mu_s 2^s=\ldots\mu_{-2}2^{-2}+\mu_{-1}2^{-1}+\mu_0 2^0+\mu_1 2^1+\mu_2 2^2+\ldots \qquad (3.1)$$

Now μ is a dyadic number if μ_{-s} is a binary digit for all integers $s\varepsilon[-\infty,\infty]$. If $\mu_{-s}=0$ for $s>M$, $M\varepsilon I_+$ (M any positive integer), then μ is a dyadic rational. In general dyadic rationals can be expressed in an alternative form

$$\mu=(t+1)/2^M \qquad (3.2)$$

where t is an even integer and $M\varepsilon I_+$. Dyadic irrationals have the form given by (3.1) where there is an infinite number of terms to the right of the binary decimal point. With this notation all real numbers can be expressed as dyadic numbers with a one to one correspondence.

The method of defining Walsh functions as products of Rademacher functions $\mathcal{G}(k,\cdot)$ was introduced by Paley [12]. If the variable j has the dyadic representation.

$$j=\sum_{k=-M}^{0} j_k 2^{-k} \qquad (3.3)$$

ie. j is a dyadic integer, then the Walsh-Paley function $\Psi(j,\cdot)$ is defined by

$$\Psi(j,\cdot)=\prod_{k=-M}^{0}[\phi(-k,\cdot)]^{j_k} \tag{3.4}$$

The Rademacher functions $\phi(-k,\cdot)$, $k\in I$ can be defined as the functions given by

$$\phi(-k,t)=\exp[\pi i t_{1-k}] \quad , \quad i=(-1)^{1/2} \tag{3.5}$$

where t is a non negative real number denoted by

$$t=\sum_{k=-\infty}^{\infty} t_k 2^{-k} \tag{3.6}$$

The relationship between the Walsh-Paley functions $\Psi(j,\cdot)$ and the sequency[1] ordered Walsh functions $wal(n,\cdot)$ is given by

$$wal(n,\cdot)=\begin{cases}\Psi([n/2]\oplus n,\cdot), & n=0,2,4....\\ \Psi([(n-1)/2]\oplus n,\cdot), & n=1,3,5...\end{cases} \tag{3.7}$$

where $\oplus$ denotes the modulo-2 addition of dyadic numbers.

The functions $\Psi(k,\cdot)$ are different from the functions $wal(n,\cdot)$ only in the numbering of the functions. The functions $wal(n,\cdot)$ have exactly n zero crossings in the

[1]The term sequency is also referred to as generalized frequency or one half the number of zero crossings [17].

interval of orthogonality of the Walsh functions. If this interval of orthogonality is normalized to unity then n/2 becomes the generalized frequency, or sequency, of the Walsh function wal(n,·). This is analogous to exp[iπf] where f/2 is the frequency of the exponential function. The sequency ordered functions wal(n,·) are the functions most commonly used because of this correspondence with exponential functions.

Just as the exponential functions can be expressed as sines and cosines, similarly cal(j,·) and sal(j,·) are Walsh functions defined by

$$\begin{aligned} \text{cal}(j,\cdot) &= \text{wal}(2j,\cdot) \\ \text{sal}(j,\cdot) &= \text{wal}(2j-1,\cdot) \end{aligned} \tag{3.8}$$

where j is the sequency of the functions cal(j,·) and sal(j,·).

A similar definition of Walsh functions was given in a paper by Fine [13], where the Walsh-Paley functions $\Psi(y,\cdot), y \in R_+$ (R_+ denotes the non-negative real numbers) are defined by

$$\psi(y,t)=\exp[\pi i \sum_{k=-L}^{M+1} Y_k t_{1-k}] \tag{3.9}$$

where $y,\ t\in R_+$ have the dyadic representations

$$y=\sum_{k=-L}^{\infty} y_k 2^{-k} \quad \text{and} \quad t=\sum_{k=-M}^{\infty} t_k 2^{-k} \tag{3.10}$$

This definition is useful in that the Walsh functions are defined in terms of the continous parameters y and t so that Walsh integrals can be used to represent non-periodic functions as superpositions of Walsh functions.

The discrete Walsh functions wal(j,n) with $j,n\in I_+$ (I_+ denotes the non-negative integers) can be defined by

$$\text{wal}(j,n)=(-1)^{\sum j_p n_{M-1-p}} \tag{3.11}$$

where

$$j=\sum_{k=0}^{M-1} j_k 2^k \quad \text{and} \quad n=\sum_{k=0}^{M-1} n_k 2^k \tag{3.12}$$

This definition of Walsh functions is the most common for discrete implementations. The Walsh functions are defined on the interval $n\in[0,N-1]$ and the Walsh functions wal(j,n), $j\in[0,N-1]$ all start with a value of +1 such that wal(j,o)=1 for $j\in[0,N-1]$, with $N=2^m$. Changing j and n to dyadic

representations of real numbers can transform this definition of discrete Walsh functions into a continuous representation described by Blachman [18].

$$\text{wal}(j,n)=(-1)^{\sum_{k=-\infty}^{\infty}(j_k+j_{k+1})n_{-k}} \tag{3.13}$$

$$\text{where } j=\sum_{k=-\infty}^{\infty} j_k 2^{-k}, n=\sum_{k=-\infty}^{\infty} t_k 2^{-k} \tag{3.14}$$

The definition of Walsh functions given in (3.11) and (3.12) will be used because of the notational convenience of this definition. The first 16 Walsh functions derived from this definition are indicated in Figure 5 and the Rademacher functions are indicated as a subset of the Walsh functions.

3.2 WALSH FUNCTIONS AND HADAMARD FUNCTIONS

Hadamard functions had(n,·) are identical to Walsh functions wal(k,·) where the binary number n can be converted to the binary number k by following the procedure;

1) reverse the bits of n eg. for n=1011 → 1101
2) perform a Gray to binary code conversion of this bit reversed number eg. 1101 → 1001.

Therefore had(11,*) is identical to wal(9,*). Figure 5 indicates the Hadamard and Walsh function numbering. Figure 6 depicts a Fortran IV program which converts from Hadamard to Walsh function numbering. A hardware implementation of this conversion is given in [19]. The advantage of the Hadamard numbering occurs in the computation of the FHT (Fast Hadamard Transform) in that the input sequence does not have to be arranged in bit reversed order as in the computation of the FWT. This is discussed in 4.2.

3.3 WALSH MATRICES

The Walsh matrix $\underline{W}(N)$ is composed of N rows and N columns of +1's and -1's in increasing order of zero crossings. The rows are the Walsh functions in their proper order. For example wal(k-1,j-1) can be determined by locating the k^{th} row and j^{th} column of the Walsh matrix. For length 16 Walsh functions

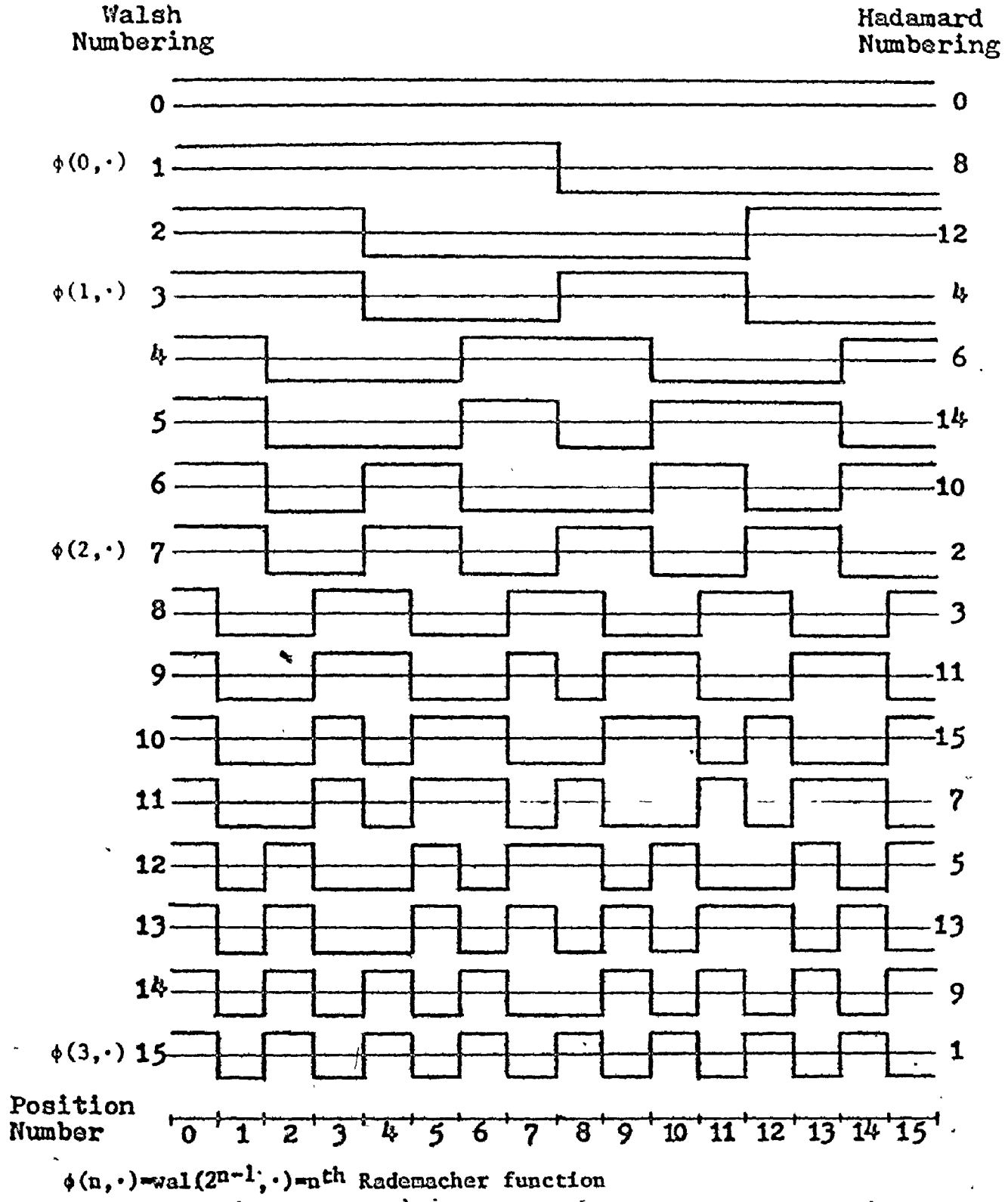


FIGURE 5 : Walsh-Hadamard-Rademacher Function Ordering for N=16

```
      SUBROUTINE CONVERT(N,M)
C  ARRAY IHAD CONTAINS THE WALSH COEFFICIENTS IN THE HADAMARD ORDER.
C  THE COEFFICIENTS WILL BE PLACED IN THE WALSH ORDER IN ARRAY IWAL.
C  N MUST BE AN INTEGER POWER OF 2 , I.E. N=2**M.
      NB(1)=0
      NB(2)=N-1
      NN=N/2
      DO 1 I=2,NN
      J=I-1
      NB(I+J)=NB(I)/2
      NB(I+J+1)=N-1-NB(I+J)
1     CONTINUE
      DO 2 I=1,N
      IWAL(NB(I)+1)=IHAD(I)
2     CONTINUE
      RETURN
      END
```

FIGURE 6 : Fortran IV Subroutine to Convert from Hadamard to Walsh Ordering of Walsh Transform Coefficients

$$\underline{W}(16) = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & - & - & - & - & - & - & - & - \\
1 & 1 & 1 & 1 & - & - & - & - & - & - & - & - & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & - & - & - & - & 1 & 1 & 1 & 1 & - & - & - & - \\
1 & 1 & - & - & - & - & 1 & 1 & 1 & 1 & - & - & - & - & 1 & 1 \\
1 & 1 & - & - & - & - & 1 & 1 & - & - & 1 & 1 & 1 & 1 & - & - \\
1 & 1 & - & - & 1 & 1 & - & - & - & - & 1 & 1 & - & - & 1 & 1 \\
1 & 1 & - & - & 1 & 1 & - & - & 1 & 1 & - & - & 1 & 1 & - & - \\
1 & - & - & 1 & 1 & - & - & 1 & 1 & - & - & 1 & 1 & - & - & 1 \\
1 & - & - & 1 & 1 & - & - & 1 & - & 1 & 1 & - & - & 1 & 1 & - \\
1 & - & - & 1 & - & 1 & 1 & - & - & 1 & 1 & - & 1 & - & - & 1 \\
1 & - & - & 1 & - & 1 & 1 & - & 1 & - & - & 1 & - & 1 & 1 & - \\
1 & - & 1 & - & - & 1 & - & 1 & 1 & - & 1 & - & - & 1 & - & 1 \\
1 & - & 1 & - & - & 1 & - & 1 & - & 1 & 1 & - & 1 & - & 1 & - \\
1 & - & 1 & - & 1 & - & - & 1 & - & 1 & - & 1 & - & 1 & - & 1 \\
1 & - & 1 & - & 1 & - & 1 & - & 1 & - & 1 & - & 1 & - & 1 & - \\
\end{bmatrix} \qquad (3.15)$$

where - denotes -1. Comparing (3.15) with Figure 5, it can be seen that the Walsh matrix is an equivalent representation of a set of Walsh functions. The Walsh matrix is symmetrical ie. $\underline{W}(N)=\underline{W}^T(N)$ where T denotes the transpose and $\underline{W}(N)\underline{W}(N)/N=\underline{I}(N)$, where $\underline{I}(N)$ is the NXN identity matrix.

If $N=2^M$, $M \varepsilon I_+$, the simplest method of generating Walsh matrices is to first obtain the Hadamard matrix and then rearrange the rows to correspond to the Walsh matrix. The Hadamard matrix is easily obtained by an iterative equation given by

$$\underline{H}(2N) = \begin{bmatrix} \underline{H}(N) & \underline{H}(N) \\ \underline{H}(N) & -\underline{H}(N) \end{bmatrix} \tag{3.16}$$

where $$\underline{H}(2) = \begin{bmatrix} 1 & 1 \\ 1 & - \end{bmatrix} \tag{3.17}$$

From (3.16) and (3.17), $\underline{H}(4)$ can be determined as

$$\underline{H}(4) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & - & 1 & - \\ 1 & 1 & - & - \\ 1 & - & - & 1 \end{bmatrix} \tag{3.18}$$

Now $\underline{H}(8)$ can be determined from (3.16) and (3.18)

$$\underline{H}(8) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & - & 1 & - & 1 & - & 1 & - \\ 1 & 1 & - & - & 1 & 1 & - & - \\ 1 & - & - & 1 & 1 & - & - & 1 \\ 1 & 1 & 1 & 1 & - & - & - & - \\ 1 & - & 1 & - & - & 1 & - & 1 \\ 1 & 1 & - & - & - & - & 1 & 1 \\ 1 & - & - & 1 & - & 1 & 1 & - \end{bmatrix} \tag{3.19}$$

If $\underline{W}(8)$ is desired then the rows of $\underline{H}(8)$ are interchanged to yield

$$\underline{W}(8) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & - & - & - & - \\ 1 & 1 & - & - & - & - & 1 & 1 \\ 1 & 1 & - & - & 1 & 1 & - & - \\ 1 & - & - & 1 & 1 & - & - & 1 \\ 1 & - & - & 1 & - & 1 & 1 & - \\ 1 & - & 1 & - & - & 1 & - & 1 \\ 1 & - & 1 & - & 1 & - & 1 & - \end{bmatrix} \qquad (3.20)$$

Walsh matrices provide a compact representation of discrete Walsh functions.

3.4 PROPERTIES OF WALSH FUNCTIONS

Walsh functions have properties similar to the properties of trigonometric functions. These properties are proven in Harmuth [17]. In all cases $N=2^m$, $m \varepsilon I_+$.

1. ORTHOGONALITY

For N length Walsh functions defined on the discrete time interval [0,N-1]

$$\sum_{j=0}^{N-1} wal(k,j)wal(n,j) = \begin{cases} N & \text{if } k=n \\ 0 & \text{if } k \neq n \end{cases} \qquad (3.21)$$

with $k,n \varepsilon [0,N-1]$. The interval of orthogonality is defined as [0,N-1] and this discrete interval corresponds to the

continous interval [0,1) subdivided into N discrete sub-intervals each of length 1/N. This property of orthogonality can also be expressed in matrix notation as

$$\underline{W}(N)\underline{W}^{T}(N)/N=\underline{I}(N) \tag{3.22}$$

2. SYMMETRY

There are two types of symmetry in Walsh functions corresponding to the symmetry in trigonometric functions. The first symmetry occurs between time and sequency with

$$wal(k,j)=wal(j,k) \quad \text{for } j,k\in[0,N-1] \tag{3.23}$$

In matrix notation this symmetry is expressed as

$$\underline{W}(N)=\underline{W}^{T}(N) \tag{3.24}$$

The second type of symmetry occurs about the midpoint of the interval of orthogonality

$$\begin{aligned} wal(k,j)&=wal(k,N-1-j) \quad \text{for } k \text{ even}, k\in[0,N-1] \\ wal(k,j)&=-wal(k,N-1-j) \quad \text{for } k \text{ odd}, k\in[0,N-1] \end{aligned} \tag{3.25}$$

3. MULTIPLICATION

The set of Walsh functions is closed under the multiplication operation ie. multiplication of Walsh functions results in another Walsh function.

$$wal(k,j)wal(n,j)=wal(n\circledast k,j) \quad k,n,j\in[0,N-1] \tag{3.26}$$

where $\circledast$ represents the bit wise modulo-2 addition of n and k expressed dyadically. This property yields an expression which defines Walsh functions by the multiplication of other Walsh functions with evenly spaced zero crossings ie. the Walsh functions which are hard limited sinusoids.

$$wal(k,j)=\prod_{n=0}^{m} wal(k_n 2^n,j) \tag{3.27}$$

where $k=\sum_{n=0}^{m} k_n 2^n$. It should be noted that for $k,n\in[0,N-1]$ then $k\circledast n\in[0,N-1]$, because N is an integer power of 2.

4. COMPLETENESS

Walsh functions form a complete orthogonal set over a discrete interval of orthogonality [17]. Hence an N length sequence can be exactly and uniquely represented by a finite series expansion of Walsh functions. For x(j) any sequence with $j\in[0,N-1]$ then

$$x(j)=\sum_{k=0}^{n-1} X(k)wal(k,j), \qquad j=0,1,\ldots,N-1 \tag{3.28}$$

$$\text{with } X(k)=1/N \sum_{j=0}^{n-1} x(j)\,wal(k,j),\ k=0,1,\ldots,N-1 \qquad (3.29)$$

In matrix notation

$$\underline{x}=\underline{W}(N)\underline{X} \qquad (3.30)$$

$$\underline{X}=\underline{W}(N)\underline{x}/N \qquad (3.31)$$

where $\underline{x}$ and $\underline{X}$ are N length column vectors.

Any N length discrete sequence can be exactly represented by N Walsh transform coefficients X(k).

3.5 MORE ON SEQUENCY

The sequency of the cal and sal functions is given by μ in the notation $cal(\mu,\theta)$ and $sal(\mu,\theta)$. Sequency is defined as one half the number of zero crossings of $cal(\mu,\theta)$ or $sal(\mu,\theta)$ in a normalized one second time interval. If μ is not a dyadic rational then $cal(\mu,\theta)$ and $sal(\mu,\theta)$ are not periodic but the interpretation of sequency as one half the number of sign changes per time interval of 1 second duration still holds true. This corresponds to the normalized frequency $\nu=fT$ of the functions $\cos(2\pi ft)$ or $\sin(2\pi ft)$ where ν is defined as the number of cycles in a time interval of duration 1.

The term sequency has also been called generalized frequency ϕ where $\phi=\mu/T$ and has the same dimension as frequency [17]. In trigonometric functions the base T drops out. If the normalized variables μ and θ in $\sin(2\pi\nu\theta)$ are replaced by the non-normalized variables $f=\nu/T$ and $t=\theta T$ then

$$\sin[2\pi\nu\theta]=\sin[2\pi(fT)(t/T)]=\sin[2\pi ft] \tag{3.32}$$

A similar form of substitution in the Walsh functions with $\phi=\mu/T$ and $t=\theta T$ yeilds

$$\mathrm{sal}(\mu,\theta)=\mathrm{sal}(\phi T,t/T)\neq\mathrm{sal}(\phi,T) \tag{3.33}$$

The only exception to (3.33) occurs when T is an integer power of 2 with μ and θ dyadic rationals. In general, since Walsh functions do not have sequency and the time base connected by multiplication, this time base is an added parameter in the general form given by

$$A\ \mathrm{sal}(\phi T,(t-\tau)/T) \tag{3.34}$$

where A is the amplitude, ϕ is the sequency, T is the time base, and τ is the time delay.

CHAPTER IV: THE WALSH TRANSFORMATION

In systems analysis the principle entities that must be dealt with are signals or time functions. Although a time function is defined by a rule assigning a value to a function for each instant of time of interest, such a description is not always optimum or even adequate for some purposes. It is convenient, therefore, to become familiar with different ways of describing time functions.

4.1 ORTHOGONAL FUNCTION EXPANSIONS

Consider a function $g(t)$ on an interval $t\varepsilon\{0,T\}$ of the time axis, with functions $\psi_k(t)$ defined on this same interval. In order to describe $g(t)$ by specifying a discrete set of coefficients, consider a series expansion of the form

$$\hat{g}(t) = \sum_{k=0}^{N-1} c_k \psi_k(t) \quad ,t\varepsilon\{0,T\} \tag{4.1}$$

in which the N coefficients c_k depend only on the functions $g(t)$ to be represented but not on time, and the N functions $\psi_k(t)$ are specified independently of $g(t)$. The notation $\hat{g}(t)$ is used to denote that $\hat{g}(t)$ is to be considered an

approximation to $g(t)$. If $g(t)$ is sufficiently well-behaved and the $\psi_k(t)$ are chosen in a suitable way, the difference between $g(t)$ and $\hat{g}(t)$ will hopefully approach zero as $N \to \infty$.

The functions $\psi_k(t)$, $k=0, 1, \ldots, N-1$ are orthogonal on the interval $(0,T)$ if

$$\int_0^T \psi_j(t)\psi_k^*(t)dt = 0 \text{ for all } j,k, \; j \neq k \tag{4.2}$$

where * denotes complex conjugate. The $\psi_k(t)$, $k=0,1,\ldots,N-1$ are orthonormal if

$$\int_0^T \psi_j(t)\psi_k^*(t)dt = \begin{cases} 1 \text{ for } j=k \\ 0 \text{ for } j \neq k \end{cases} \tag{4.3}$$

Suppose $g(t)$ has a finite square norm i.e.

$$||g||^2 = \langle g,g \rangle = \int_0^T |g(t)|^2 dt < \infty \tag{4.4}$$

and the transform coefficients g_k are defined by

$$g_k \triangleq \langle g, \psi_k \rangle = \int_0^T g(t)\psi_k^*(t)dt \tag{4.5}$$

The mean square error ξ can now be defined as

$$\xi \triangleq ||g-\hat{g}||^2 = \int_0^T |g(t)-\hat{g}(t)|^2 dt \tag{4.6}$$

According to the principle of least squares, the minimum mean square error in (4.6) will occur when the constants c_k in (4.1) are the transform coefficients defined by (4.5) and this minimum mean square error is

$$\xi_{min}(N) = ||g||^2 - \sum_{k=0}^{N-1} |g_k|^2 \tag{4.7}$$

If the $\psi_k(t)$ are a complete orthonormal sequence on the interval $\{0,T\}$ then

$$\lim_{N\to\infty} \xi_{min}(N) = 0 \tag{4.8}$$

Any function $g(t)$ of finite square norm can be exactly represented by a series of transform coefficients g_k defined by (4.5) with

$$g(t) = \sum_{k=0}^{\infty} g_k \psi_k(t) \tag{4.9}$$

If the orthogonal functions $\psi_k(t)$ are suitably chosen then a finite number of transform coefficients may sometimes exactly represent the function $g(t)$.

4.2 THE DISCRETE WALSH TRANSFORMATION

According to the sampling principle any bandlimited function $g(t)$ can be exactly represented by a sequence of

samples g(n) of that function if g(t) is sampled at or above the Nyquist rate of that function. The sequence g(n) can be exactly represented by a series of Walsh transform coefficients G(k) given by

$$G(k) = 1/N \sum_{n=0}^{N-1} g(n)wal(k,n) \qquad k=0,1,\ldots,N-1 \qquad (4.10)$$

where wal(k,n) is the k^{th} Walsh function at time interval n, and where the function g(t) is defined on the continous interval [0,1).

The function g(n) can be reconstructed exactly from the transform coefficients G(k) with

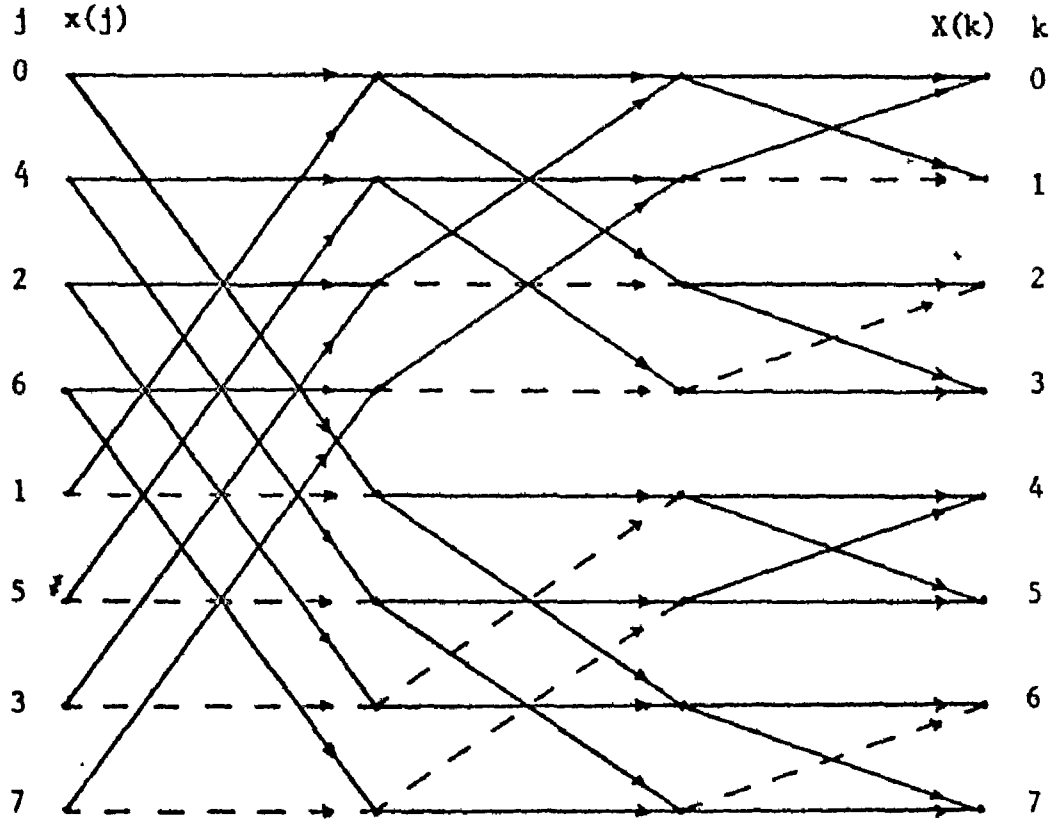
$$g(n) = \sum_{k=0}^{N-1} G(k)wal(k,n) \qquad n=0,1,\ldots,N-1 \qquad (4.11)$$

The functions G(k) and g(n) form a Walsh transform pair and this is denoted by $g(n) \overset{W}{\longleftrightarrow} G(k)$.

The transformation given in (4.10) can be computed with matrix arithmetic i.e.

$$\underline{G} = \underline{W}(N)\underline{g}/N \qquad (4.12)$$

Where $\underline{G}$ is an N length column vector whose elements are composed of G(k) in the order $\{G(0)\ G(1)\ldots G(N-1)\}^T$,



$$X(k) \underset{W}{\longleftrightarrow} x(j) \quad ; \quad \begin{matrix} j=0,1,\ldots,7 \\ k=0,1,\ldots,7 \end{matrix}$$

FIGURE 7 : Flow Graph Illustrating the F W T Operation for N=8

$\underline{W}(N)$ is the N x N Walsh matrix described in 3.3 and $\underline{g}$ is the N length column vector defined by $\{g(0)\ g(1)\ldots g(N-1)\}^T$. $(N-1)^2$ arithmetic operations are required to effect (4.12). By a special algorithm the number of operations can be minimized to $2N \log_2 N$ operations if N is constrained to be an integer power of 2 i.e. $N=2^m$, $m \varepsilon I_+$. This computationally efficient transformation is called the Fast Walsh Transformation (F W T) and the flow graph for such a transformation is illustrated in Figure 3 for N=8 This F W T is derived from the discrete Walsh transformation given in (4.12). It is based on the decimation in time algorithm (described in [19]) so that the input sequence is not in the usual order but the sequence is in the order derived from reversing the bits of the binary integer representation of the sequence numbers 0,1,...,N-1 respectively. Figure 8 indicates a Fortran program which computes the F W T of a time sequence and Figure 9 illustrates a hard ware implementation of the F W T as described in [19].

Since the Walsh functions have values of +1 and -1 only, no multiplications or divisions are required in the computation of the F W T. Only additions and subtractions are required with shifting to effect division by N where $N=2^m$, $m \varepsilon I_+$. Hence the F W T is extremely fast compared with the F F T.

4.3 PROPERTIES OF THE WALSH TRANSFORMATION

The Walsh transformation has properties analogous to the properties of the Fourier transformation. These properties will be listed here with proofs where applicable. All sequences will be of length N with $N=2^m$, $m \in I_+$.

1. Linearity

If $x(j) \overset{W}{\leftrightarrow} X(k)$ and $y(j) \overset{W}{\leftrightarrow} Y(k)$ then

$$ax(j)+by(j) \overset{W}{\leftrightarrow} aX(k)+bY(k), a,b,\text{real} \qquad (4.13)$$

Proof:
$$ax(j)+by(j) = a\sum_{k=0}^{N-1} X(k)wal(k,j)+b\sum_{k=0}^{N-1} Y(k)wal(k,j)$$

$$= \sum_{k=0}^{N-1} \{aX(k)+bY(k)\}wal(k,j)$$

Hence $ax(j)+by(j) \overset{W}{\leftrightarrow} aX(k)+bY(k)$

2. Symmetry

Due to the symmetric properties of the Walsh functions discussed in 3.2 a sequence x(j) will have a Walsh transform composed only of even number Walsh function coefficients if x(j) is symmetric about its midpoint.

Proof: Since x(j) is symmetric about its midpoint then

$$x(j) = x(N-1-j) \quad \text{for } j=0,1,\ldots, N/2-1$$

Also; $wal(k,j) = wal(k,N-1-j)$, for k even, $k\varepsilon\{0,N-1\}$

$=-wal(k,N-1-j)$, for k odd, $k\varepsilon\{0,N-1\}$

Hence $$X(k) = 1/N \sum_{j=0}^{N-1} x(j)wal(k,j)$$

$$=1/N \sum_{j=0}^{N/2-1} x(j)wal(k,j)+1/N \sum_{j=N/2}^{N-1} x(j)wal(k,j)$$

But $$1/N \sum_{j=N/2}^{N-1} x(j)wal(k,j)=1/N \sum_{j=0}^{N/2-1} x(N-1-j)wal(k,N-1-j)$$

$$=1/N \sum_{j=0}^{N/2-1} x(j)wal(k,j) \text{ for k even}$$

$$=-1/N \sum_{j=0}^{N/2-1} x(j)wal(k,j) \text{ for k odd}$$

Therefore $$X(k) = 2/N \sum_{j=0}^{N/2-1} x(j)wal(k,j) \text{ , k even, } k\varepsilon\{0,N-1\}$$

(4.14)

$=0$ k odd, $k\varepsilon\{0,N-1\}$

A similar situation exists for transforms of sequences which are asymmetric with respect to the midpoint of the sequence. Here the transform series is composed of only odd numbered coefficients.

```
      SUBROUTINE FWT(N,M,IWAL)
C  ARRAY IWAL CONTAINS THE DATA TO BE TRANSFORMED INTO THE WALSH
C  DOMAIN.  THE WALSH TRANSFORM COEFFICIENTS WILL BE PLACED IN
C  ARRAY IWAL IN THEIR NATURAL ORDER OF INCREASING SEQUENCY.
C  THE ORIGINAL DATA WILL BE LOST. IWAL HAS DIMENSION 2N.
C  N MUST BE AN INTEGER POWER OF 2 , I.E. N=2**M.
      DIMENSION IWAL(1)
      NH=N/2
      L1=0
      DO 1 L=1,M
      LP=L+1
      LM=L-1
      L1=N-L1
      L2=N-L1
      NY=0
      NZ=2**LM
      NZI=2*NZ
      NZN=N/NZI
      DO 1 I=1,NZN
      NX=NY+1
      NY=NY+NZ
      JS=(I-1)*NZI
      JD=JS+NZI+1
      DO 1 J=NX,NY
      JS=JS+1
      J2=J+NH
      IWAL(L1+JS)=IWAL(L2+J)+IWAL(L2+J2)
      JD=JD-1
      IWAL(L1+JD)=IWAL(L2+J)-IWAL(L2+J2)
1     CONTINUE
      IF(L1.EQ.0) GO TO 3
      DO 2 I=1,N
      IWAL(I)=IWAL(I+N)
2     CONTINUE
3     RETURN
      END
```

FIGURE 8 : Fortran IV Subroutine for Implementation of the Fast Walsh Transform (F W T)

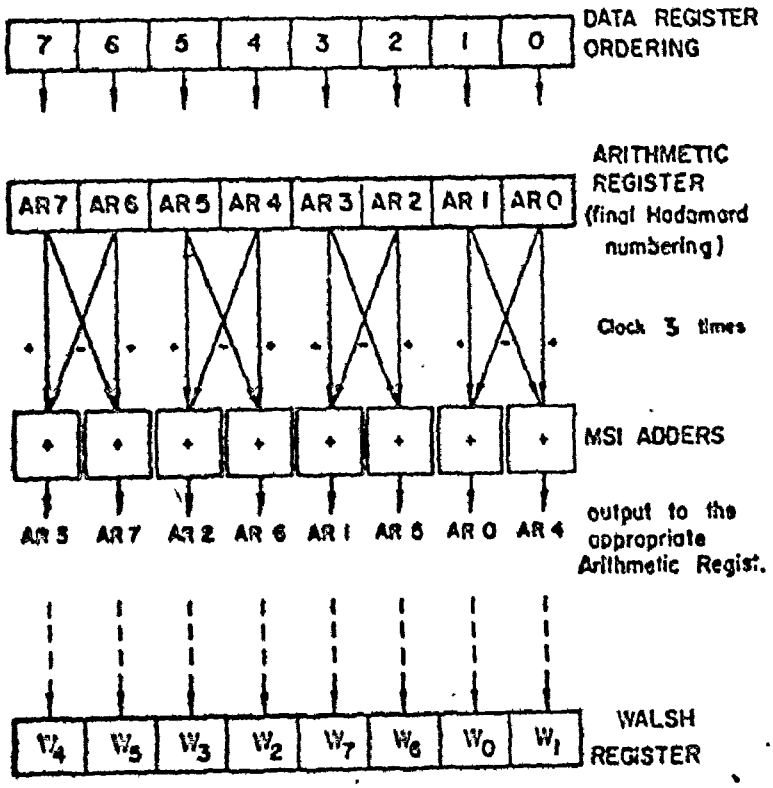


FIGURE 9 : Hardware Implementation of the F W T for N=8 .

3. Impulse Function Representation

$$\text{Define } \delta(j) = 1 \quad , \quad j=0$$
$$= 0 \quad , \quad j\neq 0 \tag{4.15}$$

Then $x(j)+a \overset{W}{\longleftrightarrow} X(k)+a\delta(k)$ (4.16)

if $x(j) \overset{W}{\longleftrightarrow} X(k)$

Proof: Define $D(k) = 1/N \sum_{j=0}^{N-1} \delta(j)\text{Wal}(k,j)$

$$= 1/N\,\text{wal}(k,0) = 1/N \tag{4.17}$$

since $\text{wal}(k,0) = 1$, $k\varepsilon\{0,N-1\}$

Similarly $\sum_{k=0}^{N-1} \delta(k)\text{wal}(k,j) = \text{wal}(0,j)=1$ (4.18)

since $\text{wal}(0,j) = 1$, $j\varepsilon\{0,N-1\}$

Hence $\delta(j) \overset{W}{\longleftrightarrow} 1/N$ and $1 \overset{W}{\longleftrightarrow} \delta(k)$ (4.19)

Therefore from the linearity property

$$x(j)+a \overset{W}{\longleftrightarrow} X(k)+a\delta(k)$$

and similarly

$$x(j)+a\delta(j) \overset{W}{\longleftrightarrow} X(k)+a/N$$

4. Signal Mean Value Calculation

For $x(j) \overset{}{\underset{w}{\longleftrightarrow}} X(k)$, $X(0)$ is the sample mean

Proof:

$$X(k)=1/N \sum_{j=0}^{N-1} x(j)wal(k,j)$$

$$X(0)=1/N \sum_{j=0}^{N-1} x(j)wal(0,j)$$

But $wal(0,j)=1$ for $j \in [0,N-1]$

Therefore

$$X(0)=1/N \sum_{j=0}^{N-1} x(j)=\text{sample mean} \qquad (4.20)$$

5. Time Shifting

For $x(j) \overset{}{\underset{w}{\longleftrightarrow}} X(k)$ then

$$x(j-\tau) \overset{}{\underset{w}{\longleftrightarrow}} \sum_{k=0}^{N-1} \alpha_{k,d}(-\tau)X(k)$$

where $\alpha_{k,d}(-\tau)$ are constants dependent only on τ.

Proof:

$$x(j)= \sum_{k=0}^{N-1} X(k)wal(k,j) \quad , \quad j \in [0,N-1] \qquad (4.21)$$

$$x(j-\tau)= \sum_{k=0}^{N-1} X(k)wal(k,j-\tau) \quad \tau \in [0,N-1] \qquad (4.22)$$

where the subtraction $j-\tau$ is accomplished with modulo-N arithmetic to effect wrap around shifting.

Now,

$$wal(k,j-\tau)=\sum_{d=0}^{N-1} \alpha_{k,d}(-\tau)wal(d,j) \tag{4.23}$$

where

$$\alpha_{k,d}(-\tau)=1/N \sum_{j=0}^{N-1} wal(k,j-\tau)wal(d,j) \tag{4.24}$$

The constants $\alpha_{k,d}(-\tau)$ are the Walsh transform coefficients that express the expansion of the time shifted Walsh functions in terms of unshifted Walsh functions and are therefore independent of the sequency x(j) and dependent only on τ.

Substituting (4.23) in (4.22) yields

$$x(j-\tau)=\sum_{k=0}^{N-1} X(k) \sum_{d=0}^{N-1} \alpha_{k,d}(-\tau)wal(d,j) \tag{4.25}$$

Rearranging the right hand side summations in (4.24)

$$x(j-\tau)=\sum_{d=0}^{N-1} [\sum_{k=0}^{N-1} X(k)\alpha_{k,d}(-\tau)]wal(d,j) \tag{4.26}$$

Comparing (4.25) with (4.21)

$$x(j-\tau) \overset{W}{\longleftrightarrow} \sum_{k=0}^{N-1} \alpha_{k,d}(-\tau)X(k) \tag{4.27}$$

For Fourier transformations, the equivalent time

shifting property is

$$x(j-\tau) \overset{F}{\longleftrightarrow} X(\omega)\exp[i\omega(-\tau)] \tag{4.28}$$

where $x(j) \overset{F}{\longleftrightarrow} X(\omega)$. The Fourier transform of the unshifted time function x(j) can be directly recoverable from the shifted time function $x(j-\tau)$, given the delay parameter τ, by multiplying the Fourier transform of $x(j-\tau)$ with $\exp[i\omega\tau]$. There is no such inverse operation for the Walsh transform ie. the phase information cannot be easily separated in the Walsh domain. The Walsh transform of x(j) can be recovered from the Walsh transform of $x(j-\tau)$ given τ by obtaining $x(j-\tau)$ using the inverse Walsh transform and then shifting $x(j-\tau)$ by $+\tau$ and obtaining the Walsh transform of this function.

The sequence x(j) can be logically or dyadically time shifted as well. Dyadic time shifts are effected by modulo-2 addition (or equivalently subtraction since modulo-2 subtraction is equivalent to modulo-2 addition) of the time parameters expressed in binary form. This dyadic time shift is denoted by $x(j \oplus \tau)$ and $\oplus$ denotes the modulo-2 addition operator. If $x(j) \overset{W}{\longleftrightarrow} X(k)$ then

$$x(j\oplus\tau)=\sum_{k=0}^{N-1} X(k)\,wal(k,j\oplus\tau) \qquad (4.29)$$

But $wal(k,j\oplus\tau)=wal(k,j)wal(k,\tau)$. Hence

$$x(j\oplus\tau)=\sum_{k=0}^{N-1} X(k)\,wal(k,j)wal(k,\tau)$$

$$=\sum_{k=0}^{N-1} [X(k)wal(k,\tau)]wal(k,j)$$

Therefore $x(j\oplus\tau)\overset{W}{\longleftrightarrow}X(k)wal(k,\tau)$ (4.30)

The Walsh transform of x(j) can be obtained from the Walsh transform of $x(j\oplus\tau)$, given τ, by multiplying the Walsh transform of $x(j\oplus\tau)$ by $wal(k,\tau)$ since $[wal(k,\tau)]^2=1$ for $k,\tau\in[0,N-1]$. The dyadic phase information can be easily separated in the Walsh transform domain in this manner.

6. Convolution

Consider the N length sequences x(j) and y(j). The linear arithmetic convolution of these two sequences is given by

$$z(j)=\sum_{\tau=0}^{N-1} x(j-\tau)y(\tau) \qquad (4.31)$$

where the subtraction $j-\tau$ is modulo-N subtraction.

Define: $x(j) \overset{W}{\longleftrightarrow} X(n)$

$y(\tau) \overset{W}{\longleftrightarrow} Y(m)$

$$x(j-\tau) \overset{W}{\longleftrightarrow} \sum_{n=0}^{N-1} \alpha_{n,k}(-\tau) X(n)$$

$z_A(j) \overset{W}{\longleftrightarrow} Z_A(k)$

Now $z_A(j)$ can be expressed as

$$z_A(j) = \sum_{k=0}^{N-1} Z_A(k)\,wal(k,j)$$

$$= \sum_{k=0}^{N-1}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \alpha_{n,k}(-\tau) X(n)\,wal(k,j)\,Y(m)\,wal(m,\tau)$$

$$= \sum_{k=0}^{N-1}\left\{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} X(n)Y(m)\sum_{\tau=0}^{N-1} \alpha_{n,k}(-\tau)\,wal(m,\tau)\right\} wal(k,j)$$

Hence $$Z_A(k) = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} X(n)Y(m)\sum_{\tau=0}^{N-1} \alpha_{n,k}(-\tau)\,wal(m,\tau) \qquad (4.32)$$

The logical convolution of $x(j)$ and $y(j)$ is defined by

$$z_L(j) = 1/N \sum_{\tau=0}^{N-1} x(j \oplus \tau)\,y(\tau) \qquad (4.33)$$

where $\oplus$ denotes modulo-2 addition.

$$z_L(j) = 1/N \sum_{\tau=0}^{N-1} y(\tau)\left\{\sum_{k=0}^{N-1} X(k)\,wal(k,\tau \oplus j)\right\} \qquad (4.34)$$

But $wal(k,\tau \oplus j)=wal(k,\tau)wal(k,j)$

Hence

$$z_L(j)=1/N \sum_{k=0}^{N-1} X(k) \sum_{\tau=0}^{N-1} y(\tau)wal(k,j)wal(k,\tau)$$

$$=1/N \sum_{k=0}^{N-1} X(k)wal(k,j)\left[\sum_{\tau=0}^{N-1} y(\tau)wal(k,\tau)\right]$$

$$= \sum_{k=0}^{N-1} X(k)Y(k)wal(k,j)$$

since $Y(k)=1/N \sum_{\tau=0}^{N-1} y(\tau)wal(k,\tau)$

Therefore

$$Z_L(k)=X(k)Y(k) \qquad (4.35)$$

where $z_L(j) \overset{W}{\longleftrightarrow} Z_L(k)$

7. Autocorrelation

The autocorrelation function $R_{xx}(\tau)$ is defined by

$$R_{xx}(\tau)=1/N \sum_{j=0}^{N-1} x(j)x(j+\tau) \qquad (4.36)$$

Substituting $x(j)= \sum_{k=0}^{N-1} X(k)wal(k,j)$

$$x(j+\tau)= \sum_{n=0}^{N-1} X(n) \sum_{d=0}^{N-1} \alpha_{n,d}(\tau)\,wal(d,j)$$

in (4.36) yields

$$R_{xx}(\tau)=1/N \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} X(k)\,wal(k,j) \sum_{n=0}^{N-1} X(n) \sum_{d=0}^{N-1} \alpha_{n,d}(\tau)\,wal(d,j)$$

$$R_{xx}(\tau)= \sum_{k=0}^{N-1} X(k) \sum_{n=0}^{N-1} X(n) \sum_{d=0}^{N-1} \alpha_{n,d}(\tau)\, 1/N \sum_{j=0}^{N-1} wal(k,j)\,wal(d,j) \qquad (4.37)$$

But $1/N \sum_{j=0}^{N-1} wal(k,j)\,wal(d,j) = 1$ for $d=k$

$=0$ otherwise

from the orthogonality of Walsh functions. Hence (4.37) reduces to

$$R_{xx}(\tau)= \sum_{k=0}^{N-1} X(k) \sum_{n=0}^{N-1} X(n)\,\alpha_{n,k}(\tau)$$

$$R_{xx}(\tau)= \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} X(k)X(n)\,\alpha_{n,k}(\tau) \qquad (4.38)$$

Similarly a cross-correlation function $R_{xy}(\tau)$ of two sequences $x(j)$ and $y(j)$ with $x(j) \overset{W}{\longleftrightarrow} X(n)$ and $y(j) \overset{W}{\longleftrightarrow} Y(m)$ is given by

$$R_{xy}(\tau)= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \alpha_{m,n}(\tau)X(n)Y(m) \qquad (4.39)$$

The logical autocorrelation and logical cross-correlation functions are the same as logical convolution since modulo-2 addition is an equivalent operation to modulo-2 subtraction.

8. Logical Covariance

The logical covariance matrix $\underline{C}_L$ is formed from the logical autocorrelation function $L(\tau)$ given by

$$L(\tau)=1/N \sum_{j=0}^{N-1} x(j\circledast\tau)x(j) \tag{4.40}$$

$$\text{Now, } \{\underline{C}_L\}_{r,c}=L(r\circledast c) \tag{4.41}$$

where r is the row number and c is the column number of the N x N covariance matrix $\underline{C}_L$. (Note: the row and column numbers start a 0 and end at N-1. The logical covariance matrix thus formed is dyadic ie. it is symmetric with respect to both diagonals. For example, for N=4

$$\underline{C}_L=\sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_3 & \rho_2 \\ \rho_2 & \rho_3 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix} \tag{4.42}$$

where $\rho_k = L(k)/\sigma^2$ and $\sigma^2 = L(0)$

The covariance matrix is diagonalized by the Walsh matrix, ie. $\underline{C}_y = \underline{W}(N)\underline{C}_L\underline{W}(N)$; where $\underline{C}_y$ is an N x N diagonal matrix with the diagonal elements given by the Walsh transform of $L(\tau), \tau \in [0, N-1]$. This arises from the properties of the $\oplus$ operation. $\underline{C}_L\underline{W}(N)$ is composed of $S_0, S_1, \ldots S_{N-1}$ where

$$S_i = \sum_{j=0}^{N-1} L(j)\,wal(i,j) \tag{4.43}$$

$$\{\underline{C}_L\underline{W}(N)\}_{r,c} = S_c\,wal(c,r) \tag{4.44}$$

For $N=4$

$$\underline{C}_L\underline{W}(N) = \frac{1}{\sqrt{4}} \begin{bmatrix} S_0 & S_1 & S_2 & S_3 \\ S_0 & S_1 & -S_2 & -S_3 \\ S_0 & -S_1 & -S_2 & S_3 \\ S_0 & -S_1 & S_2 & -S_3 \end{bmatrix} \tag{4.45}$$

where

$$S_0 = 1 + \rho_1 + \rho_2 + \rho_3$$

$$S_1 = 1 + \rho_1 - \rho_2 - \rho_3$$

$$S_2 = 1 - \rho_1 - \rho_2 + \rho_3$$

$$S_3 = 1 - \rho_1 + \rho_2 - \rho_3$$

By inspection of $\underline{C}_L\underline{W}(N)$ it can be seen that $\underline{W}(N)\underline{C}_L\underline{W}(N)$ will be diangonal and the diagonal elements will be $S_0, S_1, \ldots, S_{N-1}$, which are the Walsh transform coefficients of the logical auotcorrelation function.

9. Logical Weiner-Khintchine Theorem

Define the logical autocorrelation $L_{xx}(j)$ as

$$L_{xx}(j)=1/N \sum_{\tau=0}^{N-1} x(j\circledast\tau)x(\tau) \tag{4.46}$$

From logical convolution properties

$$L_{xx}(j)= \sum_{k=0}^{N-1} X(k)X(k)wal(k,j)$$

$$= \sum_{k=0}^{N-1} [X(k)]^2 wal(k,j)$$

Hence $L_{xx}(j) \overset{W}{\longleftrightarrow} [X(k)]^2$ (4.47)

But $[X(k)]^2=P_x(k) \equiv$ sequency power spectrum of $x(j)$.

[Note: $P_x(k)$ is variant with real arithmetic time shifts of $x(j)$ but is invariant with dyadic time shifts of $x(j)$].

Therefore

$$L_{xx}(j) \overset{W}{\longleftrightarrow} P_x(k) \tag{4.48}$$

10. Parsevals Theorem

For a finite sequency x(j) the signal energy is given by

$$E=1/N \sum_{j=0}^{N-1} [x(j)]^2$$

$$=1/N \sum_{j=0}^{N-1} \left[\sum_{k=0}^{N-1} X(k)wal(k,j) \right]^2$$

$$=1/N \sum_{k=0}^{N-1} [X(k)]^2 \sum_{j=0}^{N-1} [wal(k,j)]^2$$

$$= \sum_{k=0}^{N-1} [X(k)]^2 \qquad (4.49)$$

since $1/N \sum_{j=0}^{N-1} [wal(k,j)]^2=N$ for $k \in [0,N-1]$

Hence the energy in the time signal is preserved in the Walsh spectral energy.

11. Time Invariant Sequency Power Spectrum

The power spectrum defined by $P_x(k)=[X(k)]^2$ is time variant and hence it does not have the desired properties that a time shifted version of a periodic signal has the same power spectrum as the original signal. Hence it would be desirable to derive an expression for a time invariant sequency power spectrum.

The time invariant power spectrum can be derived from the binary orthogonal decomposition of an N length sequency $\{x(j)\}$ with $N=2^m$, $m \in I_+$, in the following way, [20]. An N- periodic sequence $\{x(j)\}$ can be decomposed into an N/2- periodic sequence $\{A_1(j)\}$ and an N/2- antiperiodic[2] sequence $\{B_1(j)\}$ with

$$\{x(j)\}=\{A_1(j)\}+\{B_1(j)\} \tag{4.50}$$

where $$\{A_1(j)\}=1/2\{x(j)+x(j+N/2)\} \tag{4.51}$$

and $$\{B_1(j)\}=1/2\{x(j)-x(j-N/2)\} \tag{4.52}$$

To continue the decomposition, $\{A_1(j)\}$ is further decomposed into an N/4- periodic sequence $\{A_2(j)\}$ and an N/4- antiperiodic sequence $\{B_2(j)\}$ with

$$\{A_1(j)\}=\{A_2(j)\}+\{B_2(j)\} \tag{4.53}$$

where $$\{A_2(j)\}=1/2\{A_1(j)+A_1(j+N/4)\} \tag{4.54}$$

and $$\{B_2(j)\}=1/2\{A_1(j)-A_1(j+N/4)\} \tag{4.55}$$

This process is continued until a 1- periodic sequence $\{A_m(j)\}$ is formed. There will then be m antiperiodic sequences and one periodic sequence with

2. A sequence $\{x(k)\}$ is said to be M- periodic if $x(k)=x(M+k)$ and M- antiperiodic if $x(k)=-x(M+k)$.

$$\{x(j)\}=\{A_m(j)\}+\{B_m(j)\}+\{B_{m-1}(j)\}+\ldots+\{B_1(j)\} \tag{4.56}$$

$$=\{A_m(j)\}+\sum_{k=1}^{m}\{B_k(j)\} \tag{4.57}$$

These sequences can be represented as N-length column vectors denoted by

$$\begin{array}{llll}
\underline{A}_m & y(0) & B_m \quad y(N-1) & B_{m-k} \quad y(N/2^k-1) \\
& y(0) & -y(N-1) & y(N/2^k) \\
& y(0) & y(N-1) & \vdots \\
& y(0) & -y(N-1) & y(2^k-1)N/2^k-1) \\
& y(0) & y(N-1) & y((2^k-1)N/2^k) \\
& y(0) & -y(N-1) & -y(N/2^k-1) \\
& y(0) & y(N-1) & -y(N/2^k) \\
& y(0) & -y(N-1) & \vdots \\
& y(0) & y(N-1) & \vdots \\
& y(0) & -y(N-1) & \vdots \\
& y(0) & y(N-1) & -y((2^k-1)N/2^k-1) \\
& y(0) & -y(N-1) & -y((2^k-1)N/2^k)
\end{array} \tag{4.58}$$

where $k=1,2,\ldots,m-1$. The components $y(n)$ will later be related to the Walsh transform coefficients $X(k)$ by comparing the average power in the time and sequency domains. Inspection of the vectors in (4.58) shows that they are all mutually orthogonal and therefore

$$||\{x(j)\}||^2=||\{A_m(j)\}||^2+\sum_{k=0}^{m-1}||\{B_{m-k}(j)\}||^2 \tag{4.59}$$

Evaluating the norm of each of the sequences from (4.58) yields

$$||\{A_m(j)\}||^2=Ny^2(0) \tag{4.60}$$

$$||\{B_m(j)\}||^2=Ny^2(N-1) \tag{4.61}$$

$$||\{B_{m-k}(j)\}||^2=2^{m-k}\sum_{n=1}^{k}\{y^2((2^n-1)N/2^k-1)+y^2((2^n-1)N/2^k\} \tag{4.62}$$

with $k=1,2,\ldots,m-1$.

The Walsh transform coefficients $X(k)$ can be derived from the Walsh transform of the sum of the sequences. In vector notation

$$\underline{X}=\underline{W}(N)[\underline{A}_m+\underline{B}_m+\underline{B}_m+\underline{B}_{m-1}+\ldots+\underline{B}_1] \tag{4.63}$$

where $\underline{W}(N)$ is the N x N Walsh matrix. From (4.63) and (4.58) it can be shown that

$$X(0)=y(0) \tag{4.64}$$

$$X(N-1)=y(N-1) \tag{4.65}$$

$$\begin{matrix} X(N/2^k-1) \\ X(N/2^k) \\ \vdots \\ X((2^k-1)(N/2^k-1) \\ X((2^k-1)N/2^k) \end{matrix} = \frac{\underline{W}(2^k)}{2^k} \begin{matrix} y(N/2^k-1) \\ y(N/2^k) \\ \vdots \\ y((2^k-1)N/2^k-1) \\ y((2^k-1N/2^k) \end{matrix} \qquad (4.66)$$

with k=1,2,...,m-1. Since the matrices $\underline{W}(2^k)$ are orthogonal it follows that

$$X^2(0)=y^2(0) \qquad (4.67)$$

$$X^2(N-1)=y^2(N-1) \qquad (4.68)$$

$$\sum_{n=1}^{k} \{X^2((2^n-1)N/2^k-1)+X^2((2^n-1)N/2^k)\}$$

$$=\frac{1}{2^{m-k}} \sum_{n=1}^{k} \{y^2((2^n-1)N/2^k-1)+y^2((2^n-1)N/2^k)\} \qquad (4.69)$$

with k=1,2,...,m-1.

The average power in the sequency {x(j)} is given by

$$Pav = 1/N \sum_{j=0}^{N-1} x^2(j) = 1/N||\{x(j)\}||^2 \qquad (4.70)$$

From (4.59) and (4.70)

$$Pav=y^2(0)+y^2(N-1)+1/2^{m-k}\sum_{n=1}^{k}\{y^2((2^n-1)N/2^{k-1})+y^2((2^n-1)N/2^k)\} \quad (4.71)$$

Hence from (4.67), (4.68) and (4.69)

$$Pav=X^2(0)+X^2(N-1)+\sum_{n=1}^{k}\{X^2((2^n-1)N/2^k-1)+X^2((2^n-1)N/2^k)\} \quad (4.72)$$

The terms $X^2((2^n-1)N/2^k-1)$ and $X^2((2^n-1)N/2^k)$ are adjacent coefficients in the Walsh coefficient spectrum.

Defining the sequency spectral density as

$$S^2(k)\triangleq X^2(2k-1)+X^2(2k) \quad (4.73)$$

and substituting (4.73) in (4.72) yields

$$Pav=S^2(0)+S^2(N/2-1)+\sum_{n=1}^{k}S^2((2^n-1)N/2^{k+1}) \quad (4.74)$$

The spectral points of the time invariant sequency power spectrum are given by

$$P_0=S^2(0) \quad (4.75)$$

$$P_1=S^2(N-1) \quad (4.76)$$

$$P_{k+1} = \sum_{n=1}^{k} S^2((2^n-1)N/2^{k+1}) \quad , \quad k=1,2,\ldots,m-1 \qquad (4.77)$$

The spectral points P_i represent the average power content of a group of sequencies. Each group contains a fundamental and the set of all odd harmonies relative to that fundamental. This power spectrum is invariant to any cyclic shift of the time sequence but the original time sequence cannot be recovered from an inverse transform of the power spectrum even with a conveniently defined phase spectrum [20].

Table 4.1 indicates the sequency components that are contained in each of the power spectral points for N=8,16 and 32. The average power in the group of frequency components which correspond to the sequency components of each respective group is exactly the same as the average power in that group of sequency components [21].

TABLE 4.1

Sequences in P_i

i	N=8	N=16	N=32
0	0	0	0
1	4	8	16
2	2	4	8
3	1,3	2,6	4,12
4		1,3,5,7	2,6,10,14
5			1,3,5,7,9,11,13,15

The time invariant sequency power spectrum can be utilized in the detection of harmonic structure in signals. The number of spectral points in the power spectrum is logarithmically dependent on the window size ie. there are m+1 spectral points for a window of length $N=2^m$. Therefore the applications of this power spectrum are limited to areas where a coarse spectral resolution is sufficient. One possible application is in the pattern recognition field where signals are classified according to their harmonic structure.

Pitch detection in speech signals was found to be

impractical using the time invariant sequency power spectrum due to the sensitivity of the human aural mechanisms to pitch inaccuracies. For practical window lengths of say N=128, there would be 8 spectral points so that the pitch could only be determined from the location of the maximum of these spectral points. For acceptible speech quality, the pitch information is usually quantized to 128 levels which would require a window length of $N=2^{127}$.

4.4 FREQUENCY & SEQUENCY

The conversion from the Fourier spectrum of frequency limited signals to the Walsh spectrum of these signals has been investigated by Blachman [18], Kitai [22], and other authors. It is useful to the understanding of the Walsh vocoder to describe these mapping from the Fourier spectrum to the Walsh spectrum for discrete signals.

Consider a frequency limited time functions x(t) which is represented on the interval $t\in[0,1)$ by the discrete sequence $x(j)$, $j=0,1,\ldots N-1$ with $N=2^m$, $m\in I_+$. The Discrete Fourier Transform (DFT) of x(j) is given by

$$F(f)=\sum_{j=0}^{N-1} x(j)\exp[-i2\pi fj/N], \quad f=0,1,\ldots N-1 \qquad (4.78)$$

The Walsh transform of x(j) is given by

$$X(k)=\sum_{j=0}^{N-1} x(j)\,wal(k,j) \quad , k=0,1,\ldots,N-1 \tag{4.79}$$

If x(t) is frequency limited such that the highest normalized frequency component is f_u then

$$x(j)=1/N \sum_{f=0}^{f_u} F(f)\exp[i2\pi fj/N],\ j=0,1,\ldots N-1 \tag{4.80}$$

Substituting (4.80) in (4.79) yields

$$X(k)=1/N \sum_{j=0}^{N-1} \sum_{f=0}^{f_u} F(f)\ \exp[i2\pi fj/N]wal(k,j)\ ,\ k=0,1,..,N-1 \tag{4.81}$$

If F(f)=c, where c is a constant, then

$$X(k)=c/N \sum_{j=0}^{N-1} \{\sum_{f=0}^{f_u} \exp[i2\pi f_j/N)\}wal(k,j),\ k=0,1,\ldots N-1 \tag{4.82}$$

Since cos(2πfj/N) has even symmetry and sin(2πfj/N) has odd symmetry over the range j=0,1,...,N-1, therefore from 4.3 prpperty 2.

$$X(2k)=c/N \sum_{j=0}^{N-1} \{\sum_{f=0}^{f_u} \cos(2\pi fj/N))wal(2k,j),\ k=0,1,\ldots N/2-1 \tag{4.83}$$

and

$$X(2k-1)=c/N \sum_{j=0}^{N-1} \{\sum_{f=0}^{f_u} \sin(2\pi f_j/N)\}wal(2k-1,j)\ k=1,2,..,N/2 \tag{4.84}$$

Equations (4.83) and (4.84) define the Walsh spectrum of a frequency limited signal the spectrum of which is flat with a constant magnitude c. Equivalently, the mapping from a frequency band into the corresponding sequency band is defined. This mapping illustrates the comparison between the FFT and Walsh vocoder envelope detectors.

If the lower limit in the second summation of (4.83) and (4.84) is replaced with f_d then frequency bandpass regions could be mapped into the sequency spectrum as well as lowpass regions.

Figures 10-25 illustrate the mapping of 16 equal frequency bandpass regions into the sequency spectrum. The energy in the equivalent sequency bandpass regions are calculated and used to determine what fraction of the total energy contained in the frequency bandpass region occupies the sequency bandpass region which corresponds to the frequency bandpass region. (Note: Frequency and sequency have the same units so they are depicted on the same normalized axis in FIGURES 10-25. (Also, since a log scale is employed, -40dB is the artificial zero.)

From the inspection of Figures 10-25 it is evident that the sequency bandpass region which coincides with the frequency bandpass region contains more of the total energy

contained in the frequency bandpass region than any other sequency bandpass region of the same bandwidth.

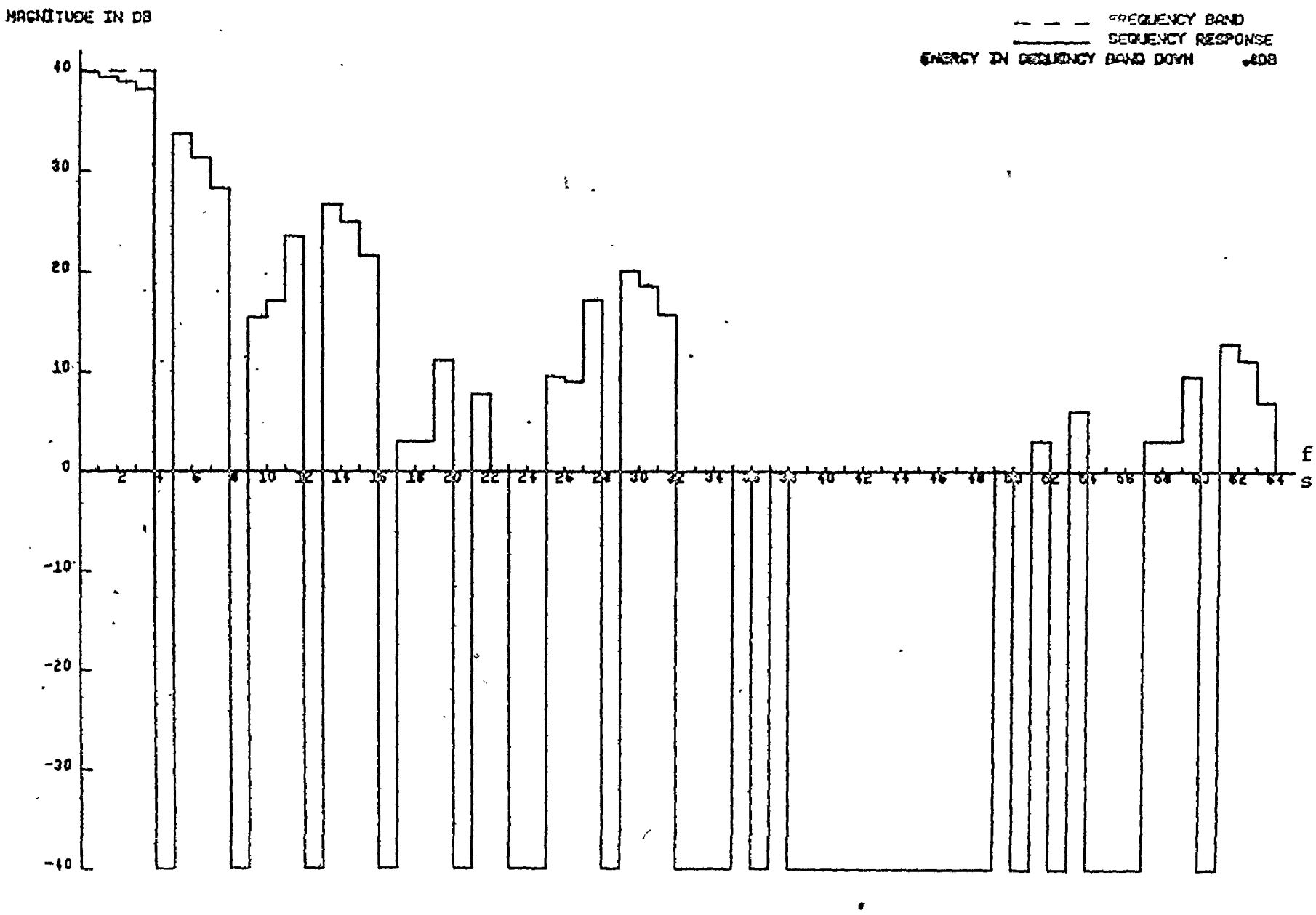


FIGURE 10 : Frequency to Sequency Conversion ; Band #1

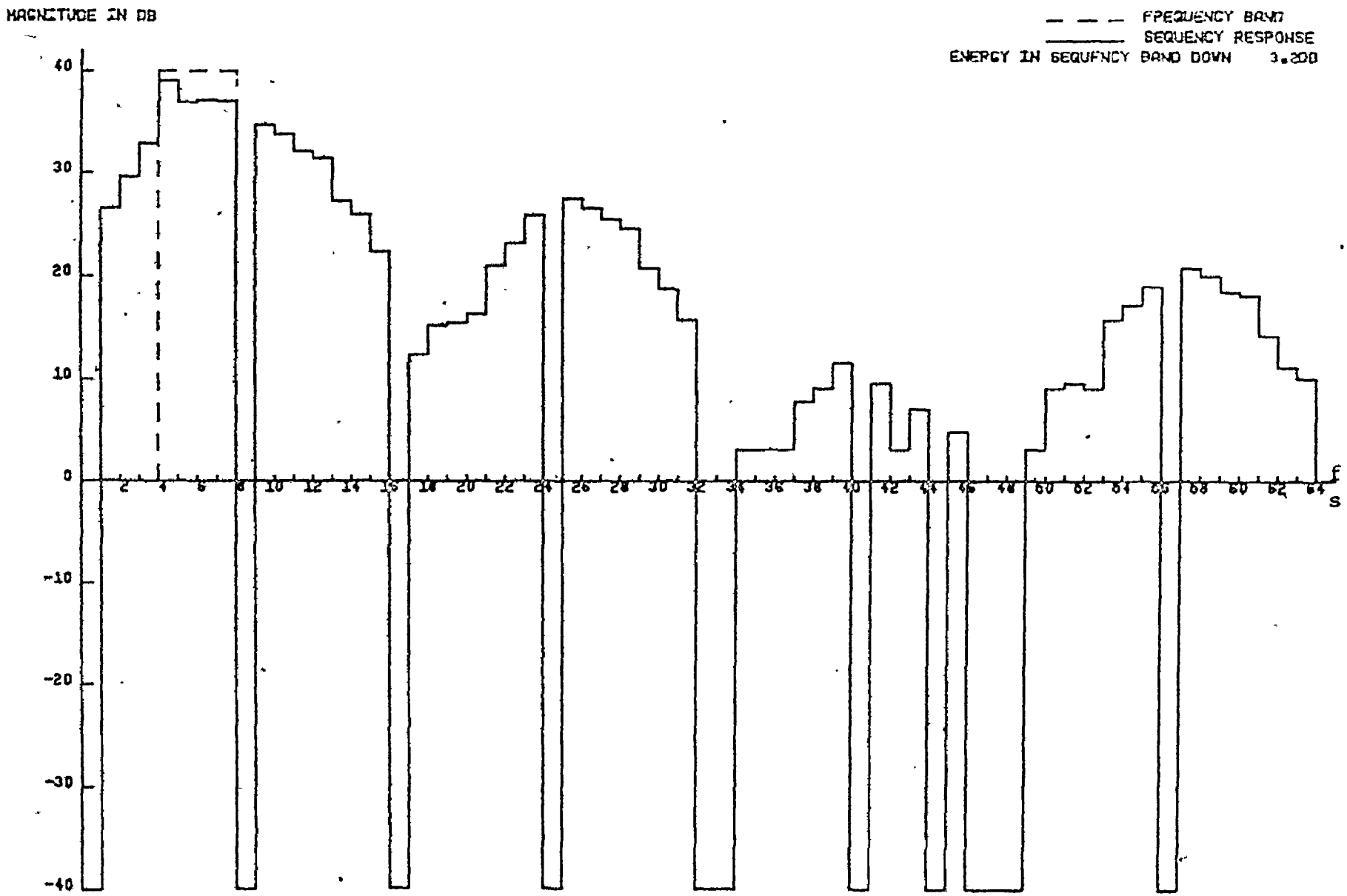


FIGURE 11 : Frequency to Sequency Conversion ; Band #2

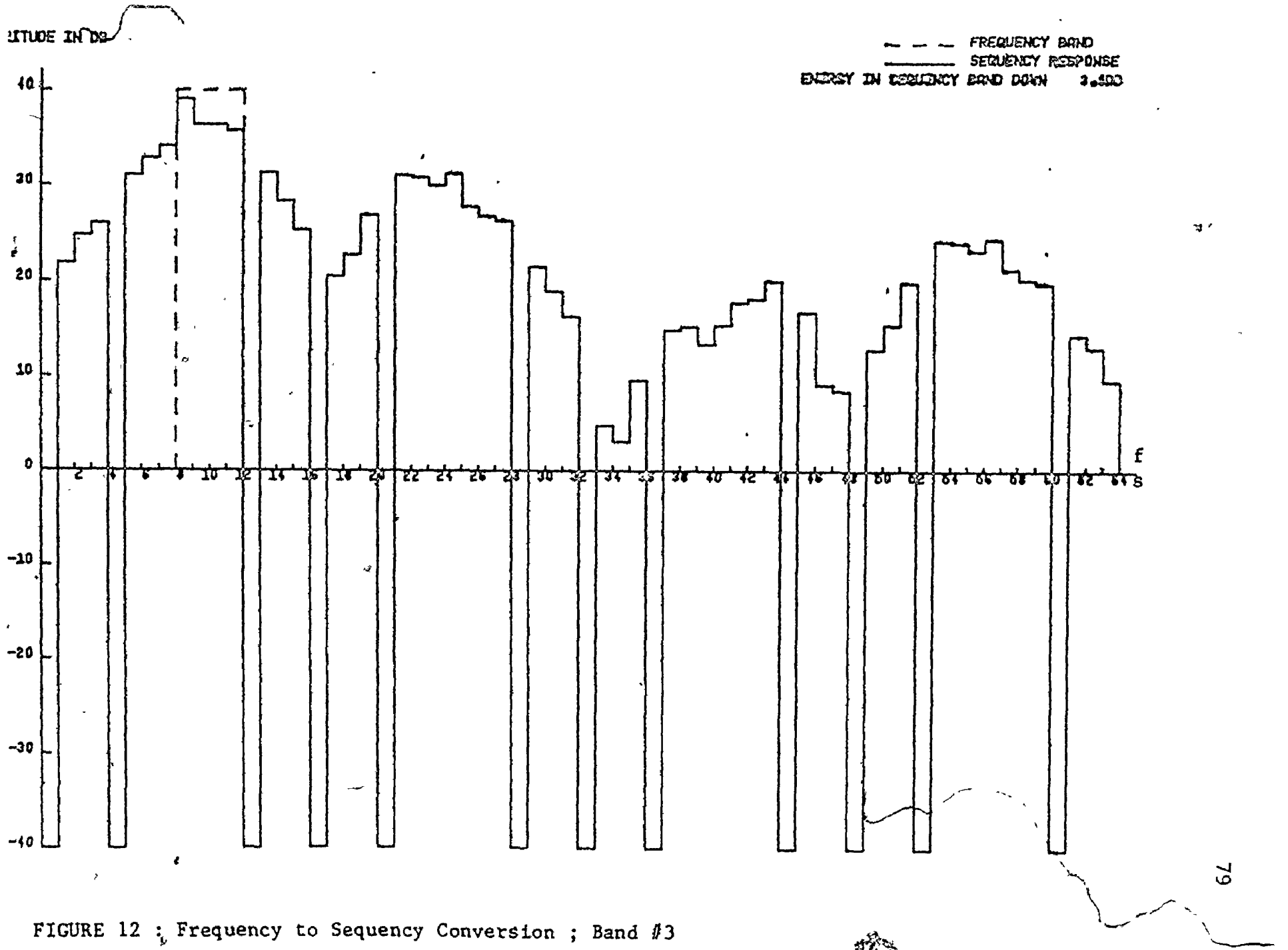


FIGURE 12 : Frequency to Sequency Conversion ; Band #3

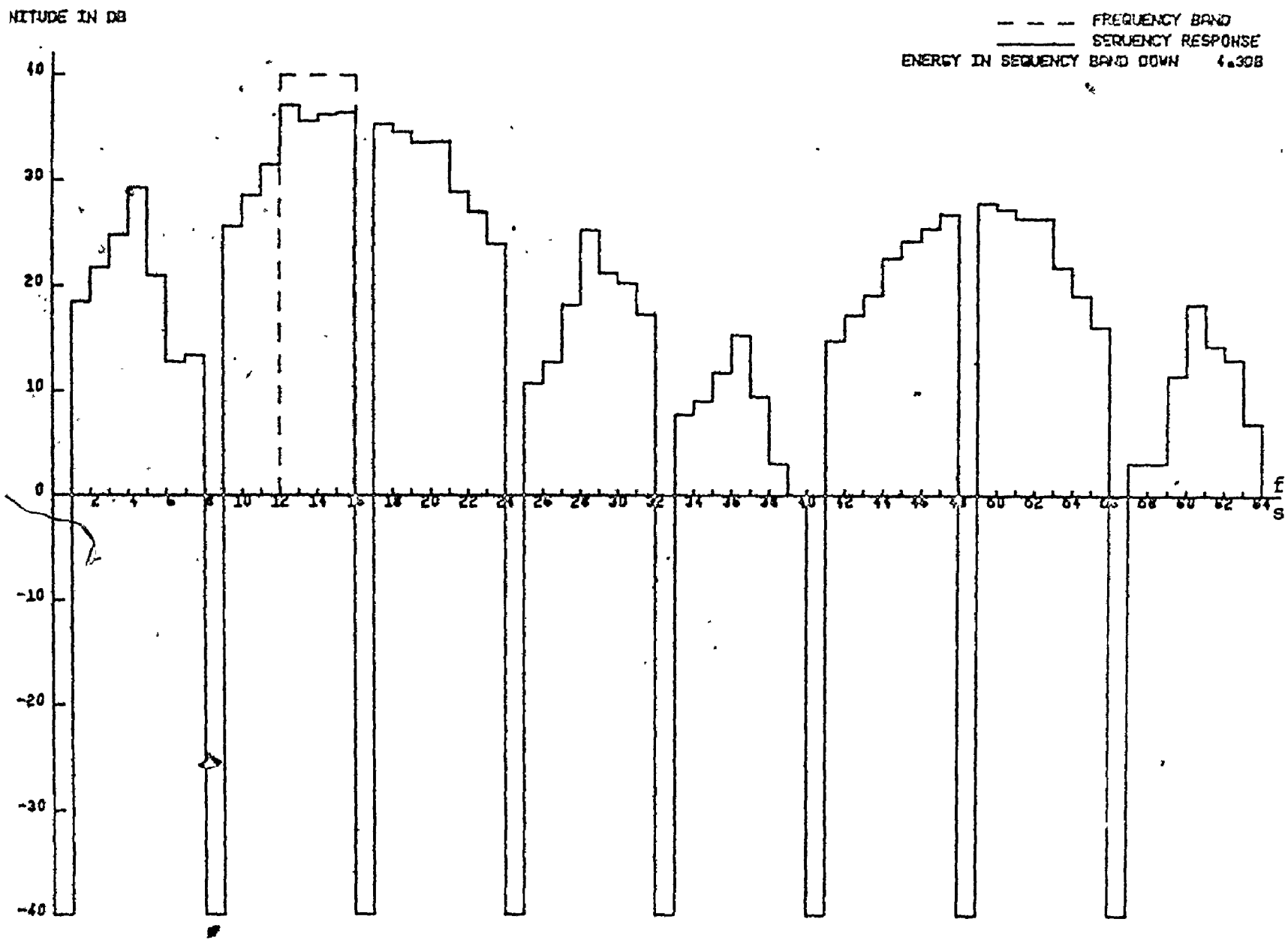


FIGURE 13 : Frequency to Sequency Conversion ; Band #4

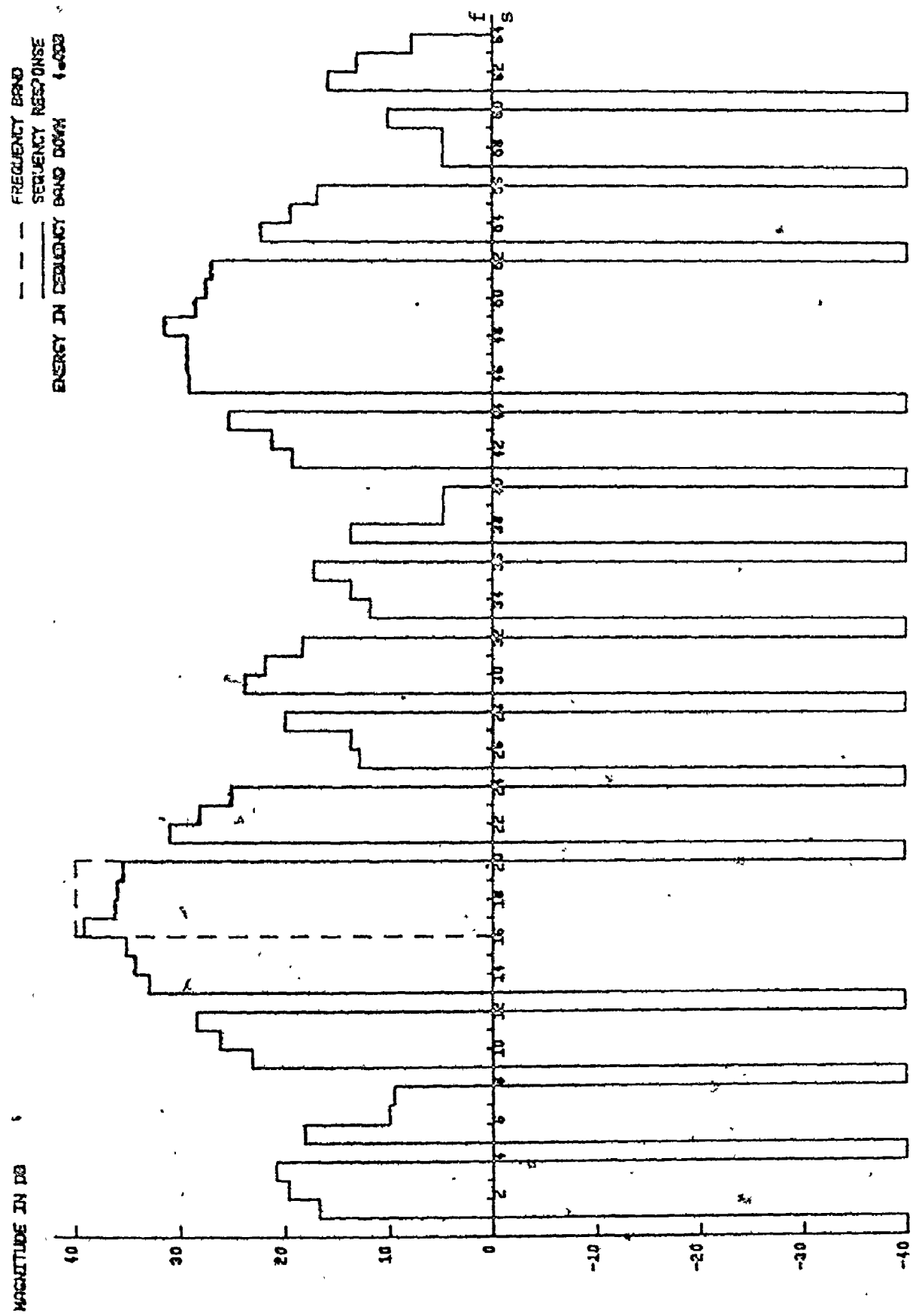


FIGURE 14 : Frequency to Sequency Conversion ; Band #5

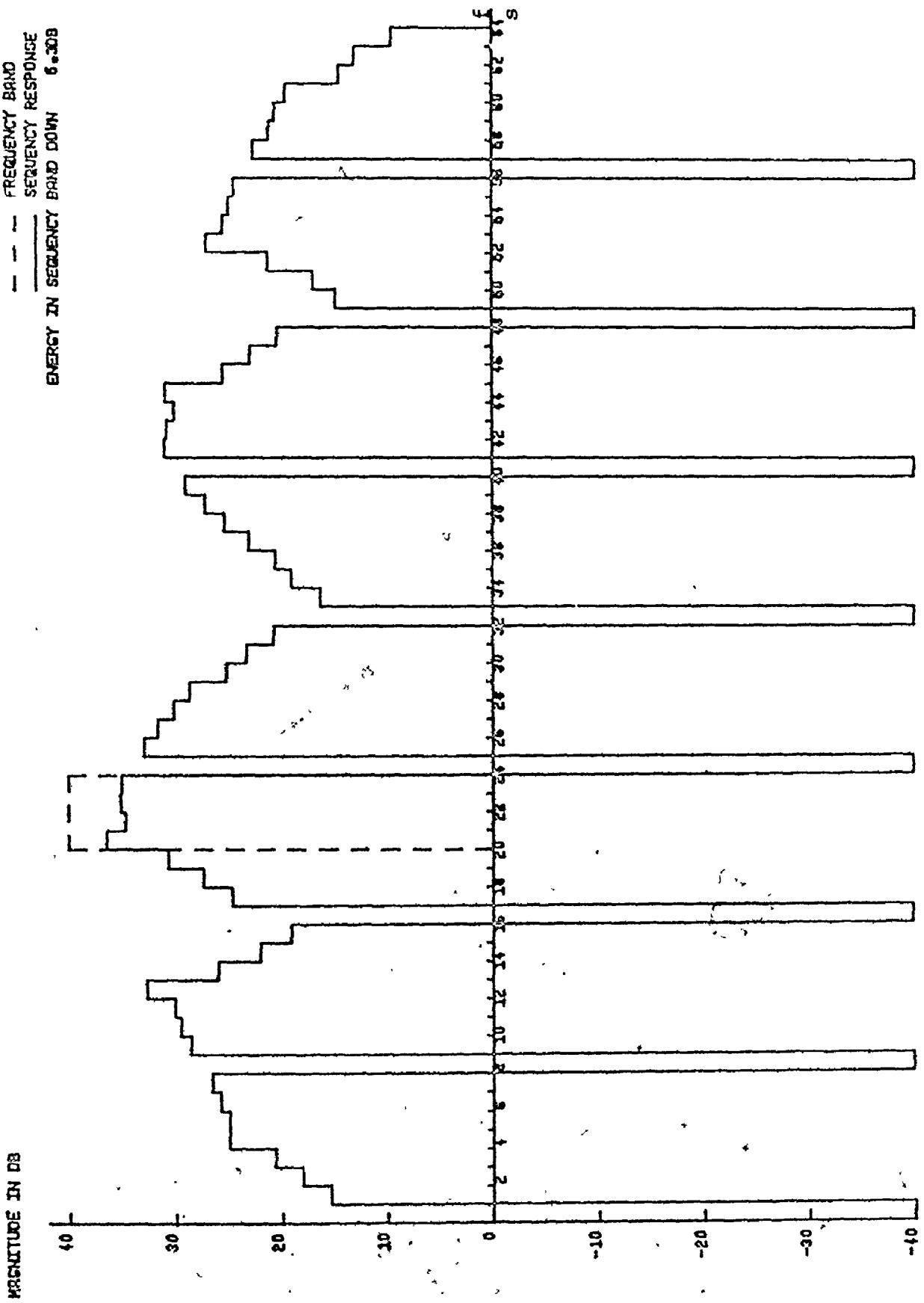


FIGURE 15 : Frequency to Sequency Conversion ; Band #6

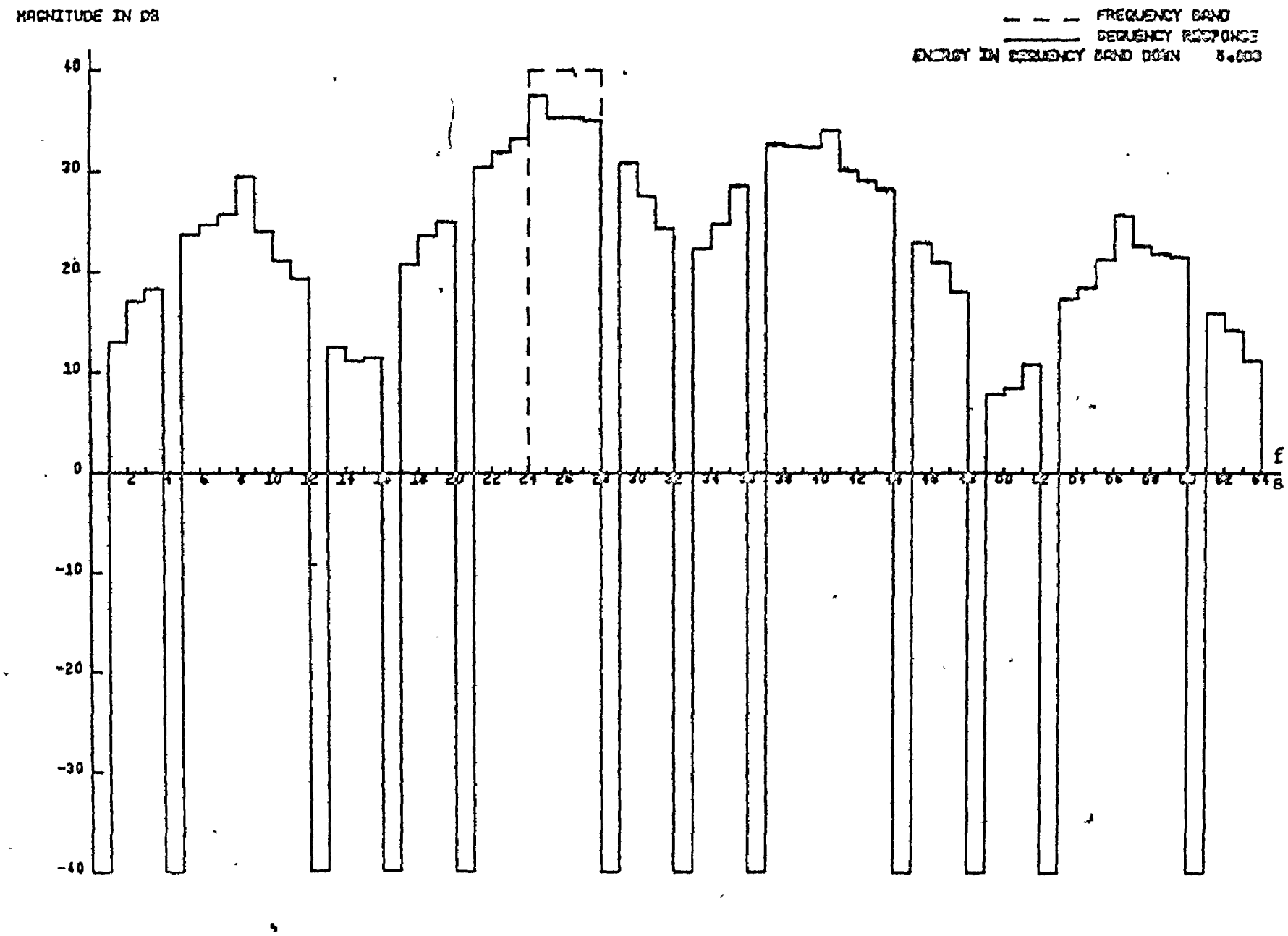


FIGURE 16 : Frequency to Sequency Conversion ; Band #7

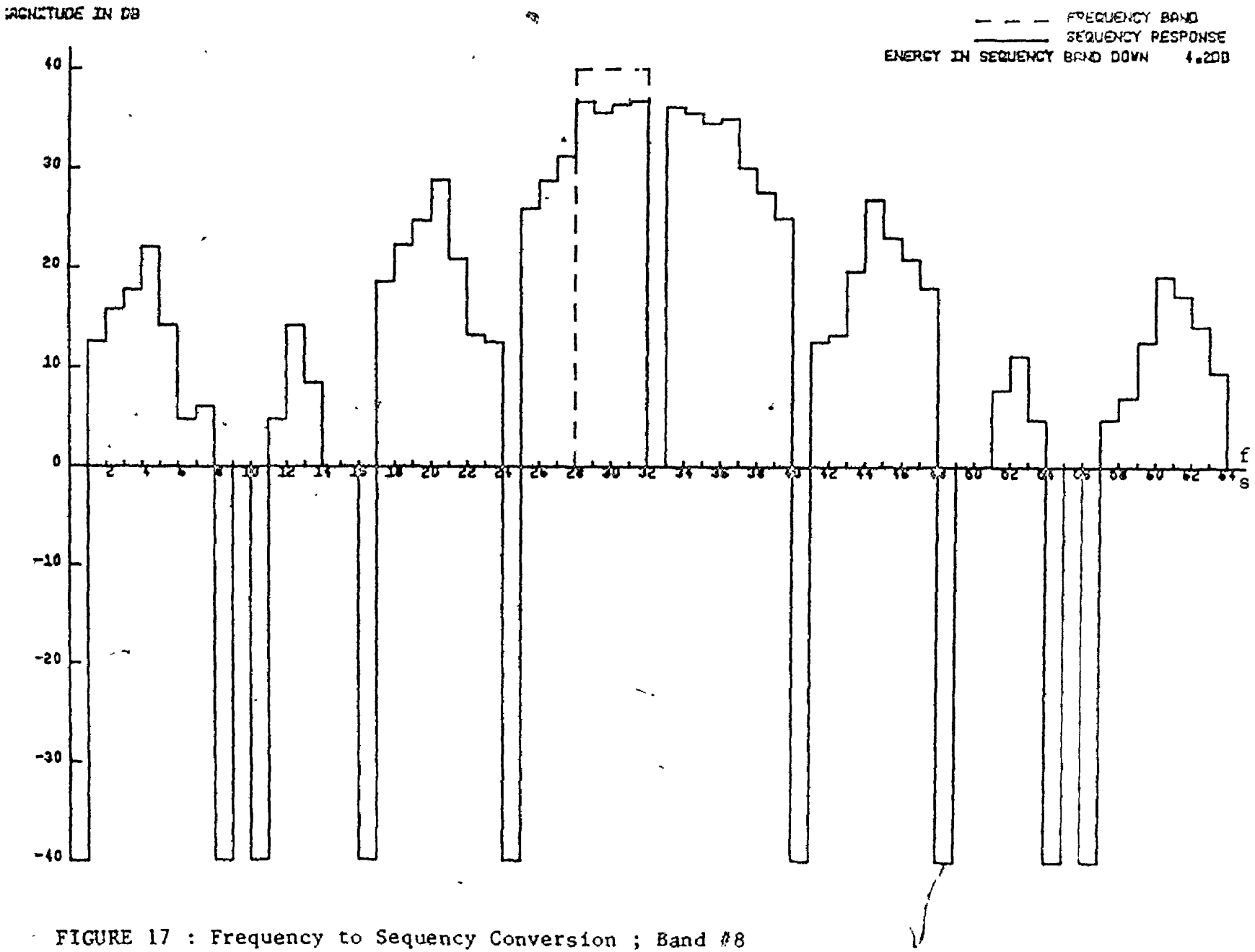


FIGURE 17 : Frequency to Sequency Conversion ; Band #8

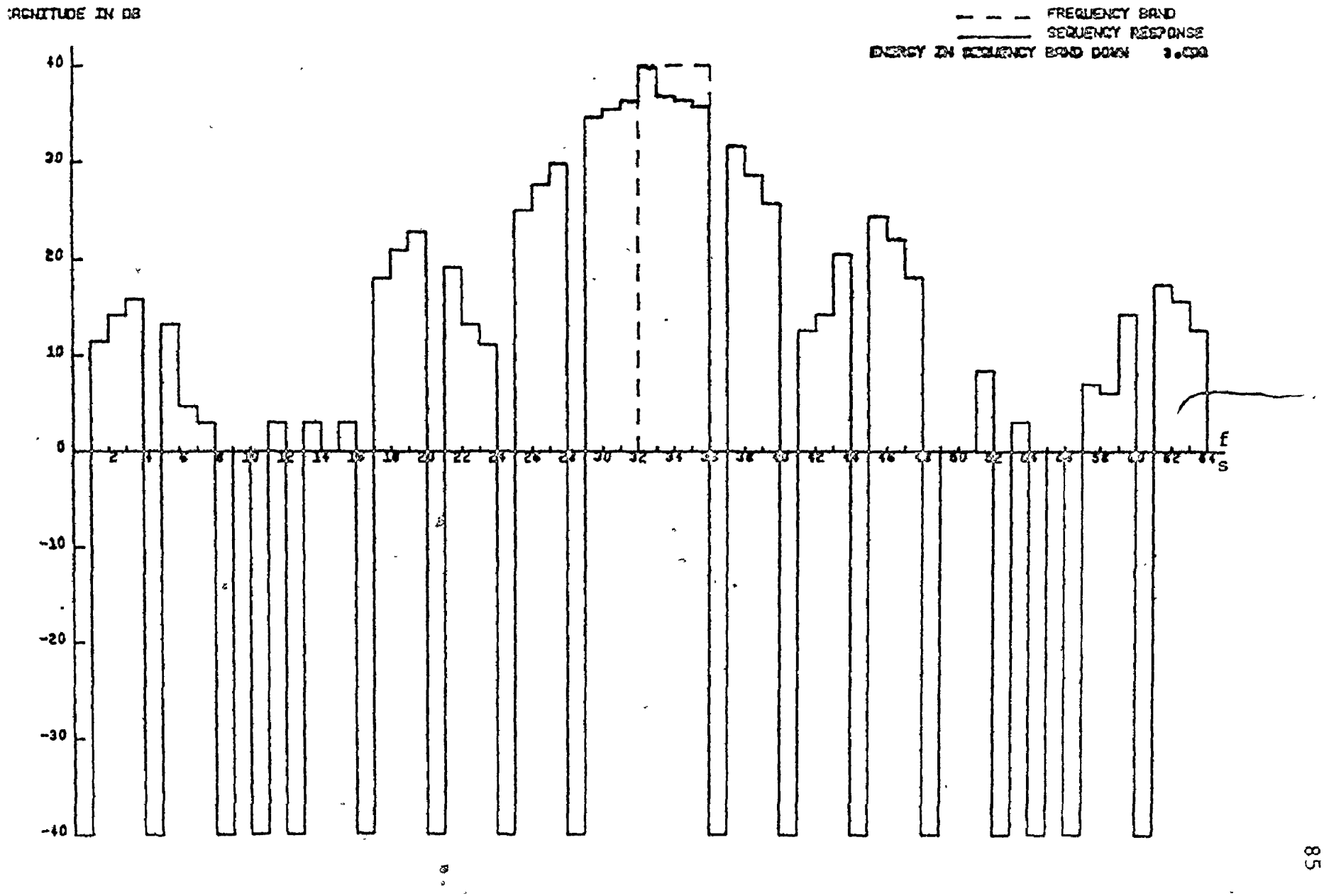


FIGURE 18 : Frequency to Sequency Conversion ; Band #9

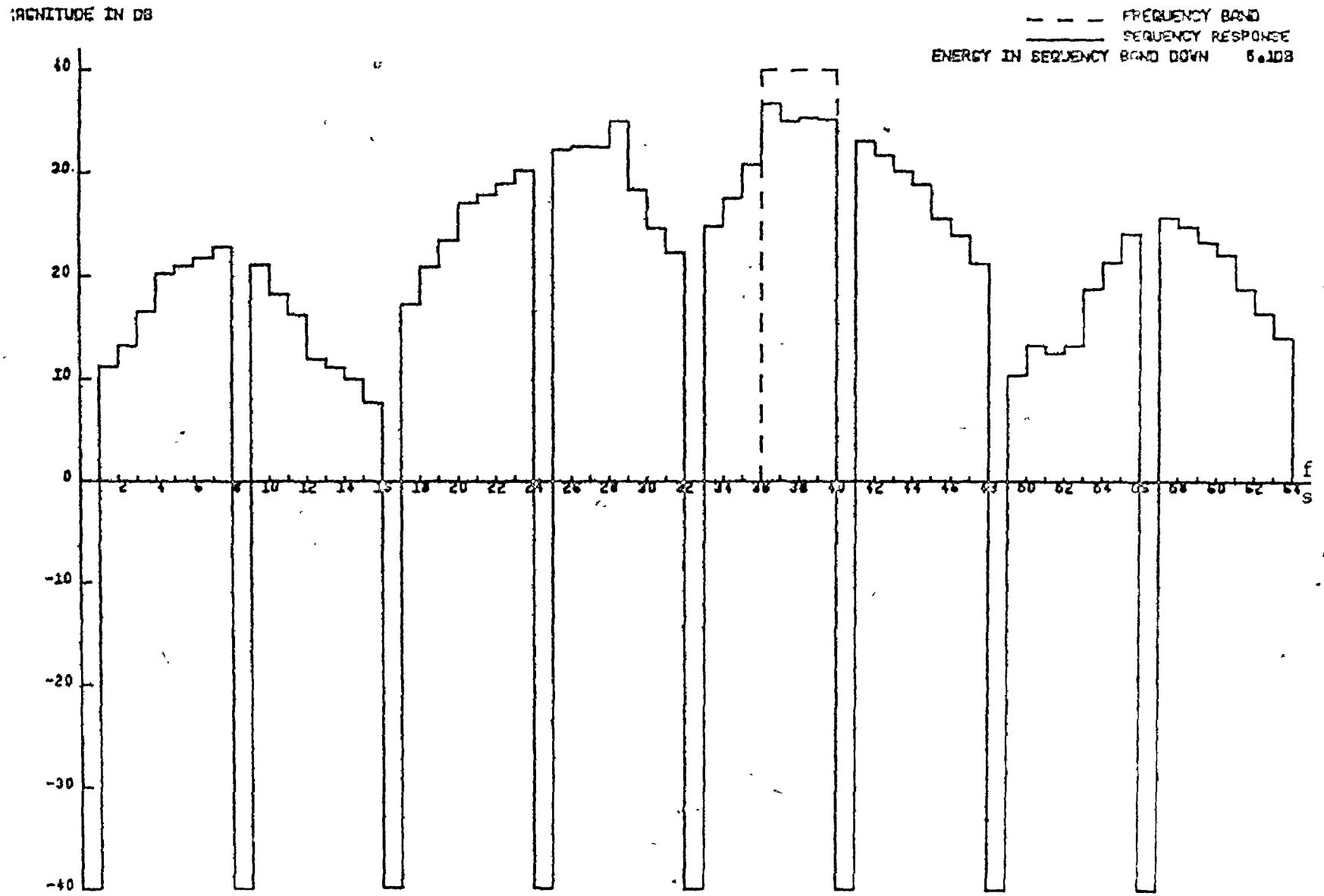


FIGURE 19 : Frequency to Sequency Conversion ; Band #10

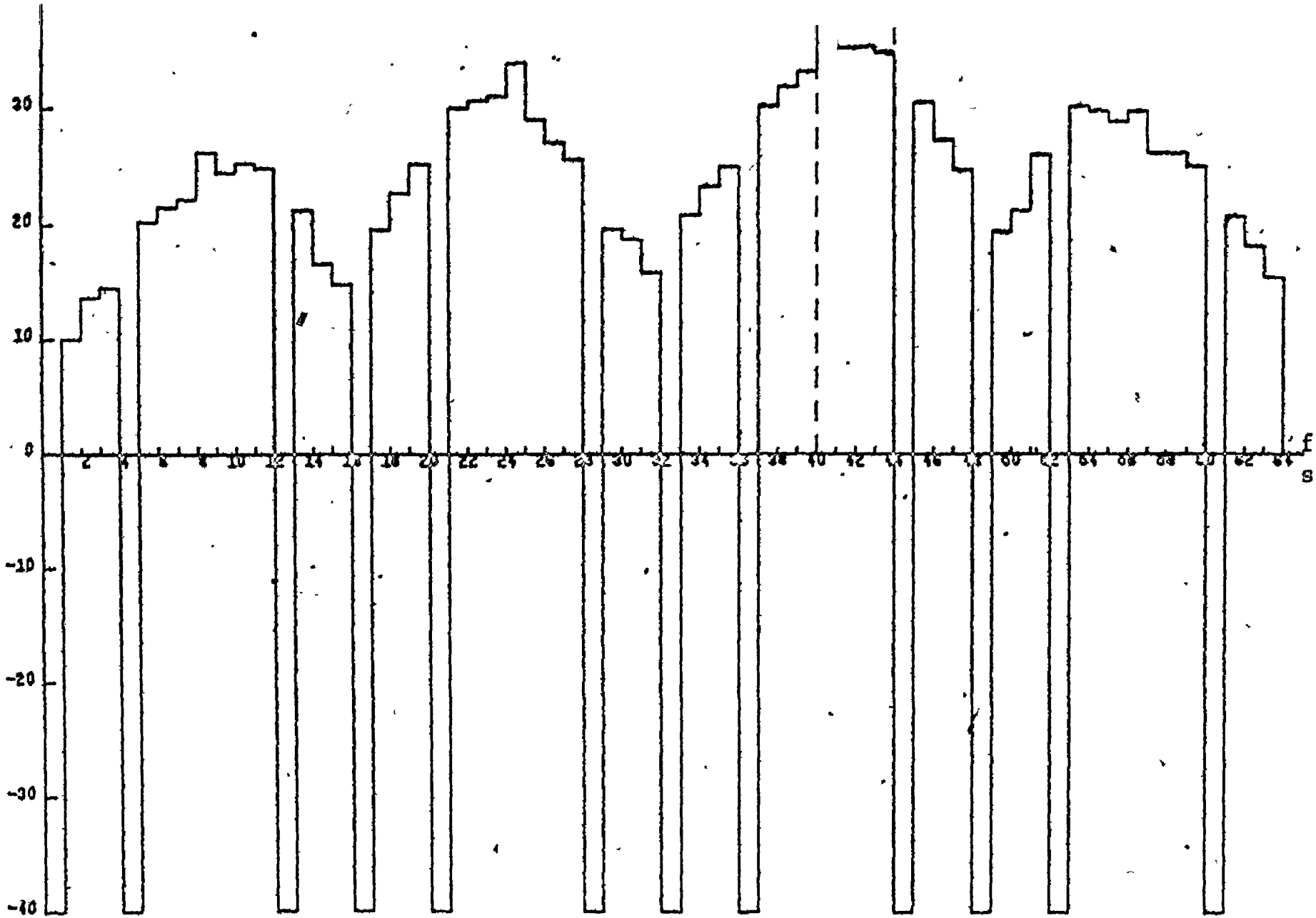


FIGURE 20 : Frequency to Sequency Conversion ; Band #11

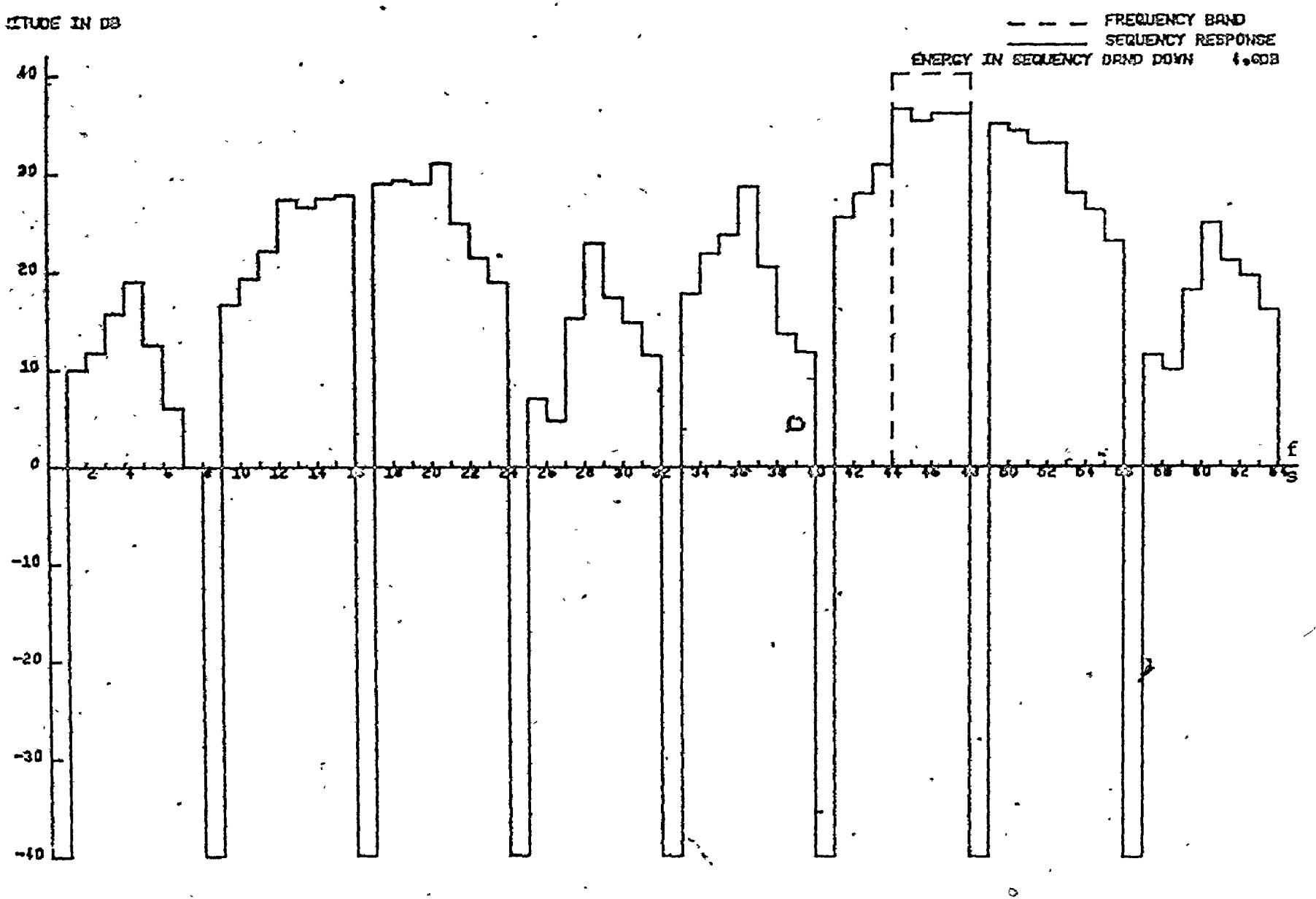


FIGURE 21 : Frequency to Sequency Conversion ; Band #12

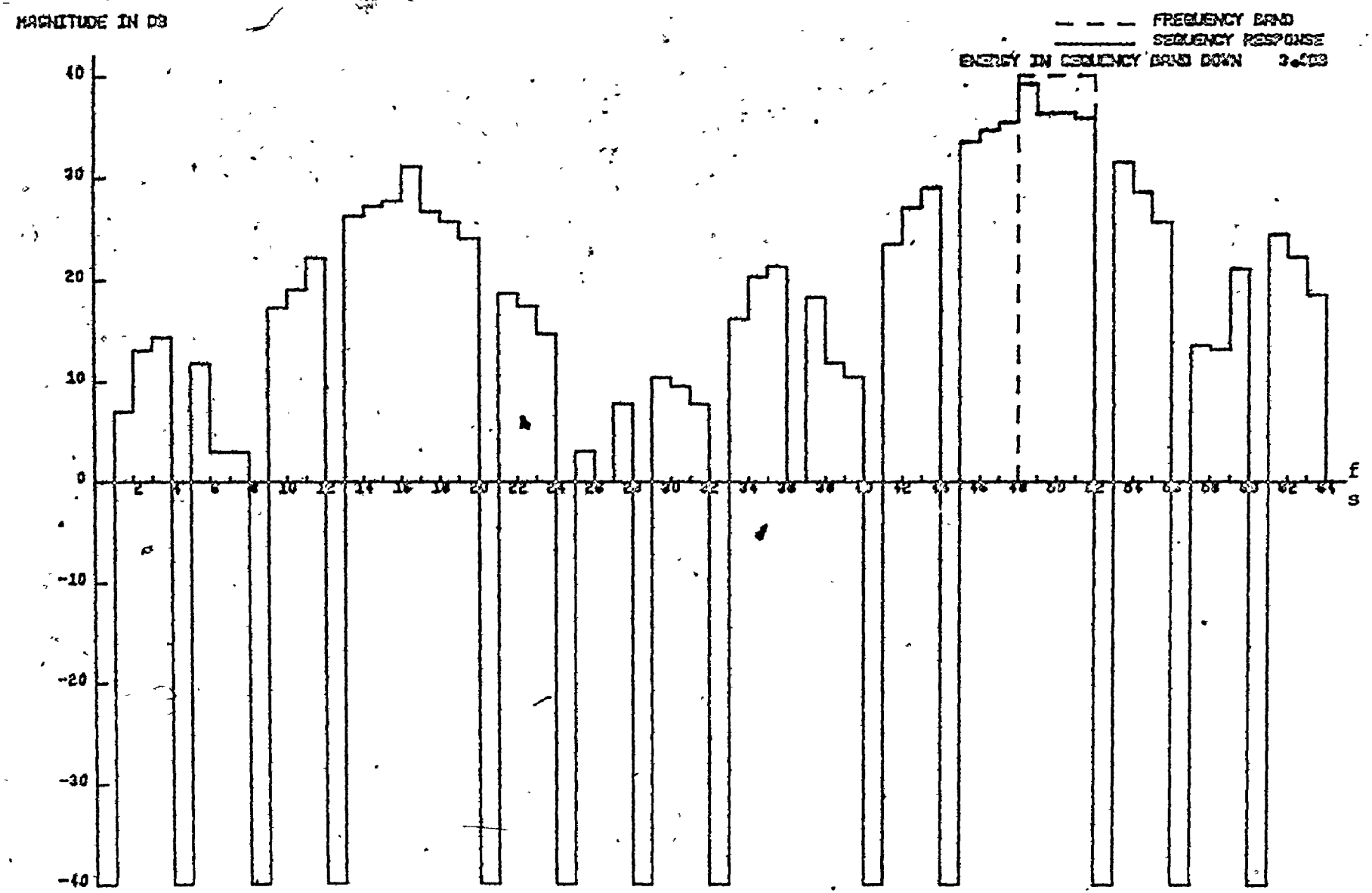


FIGURE 22 : Frequency to Sequency Conversion ; Band #13

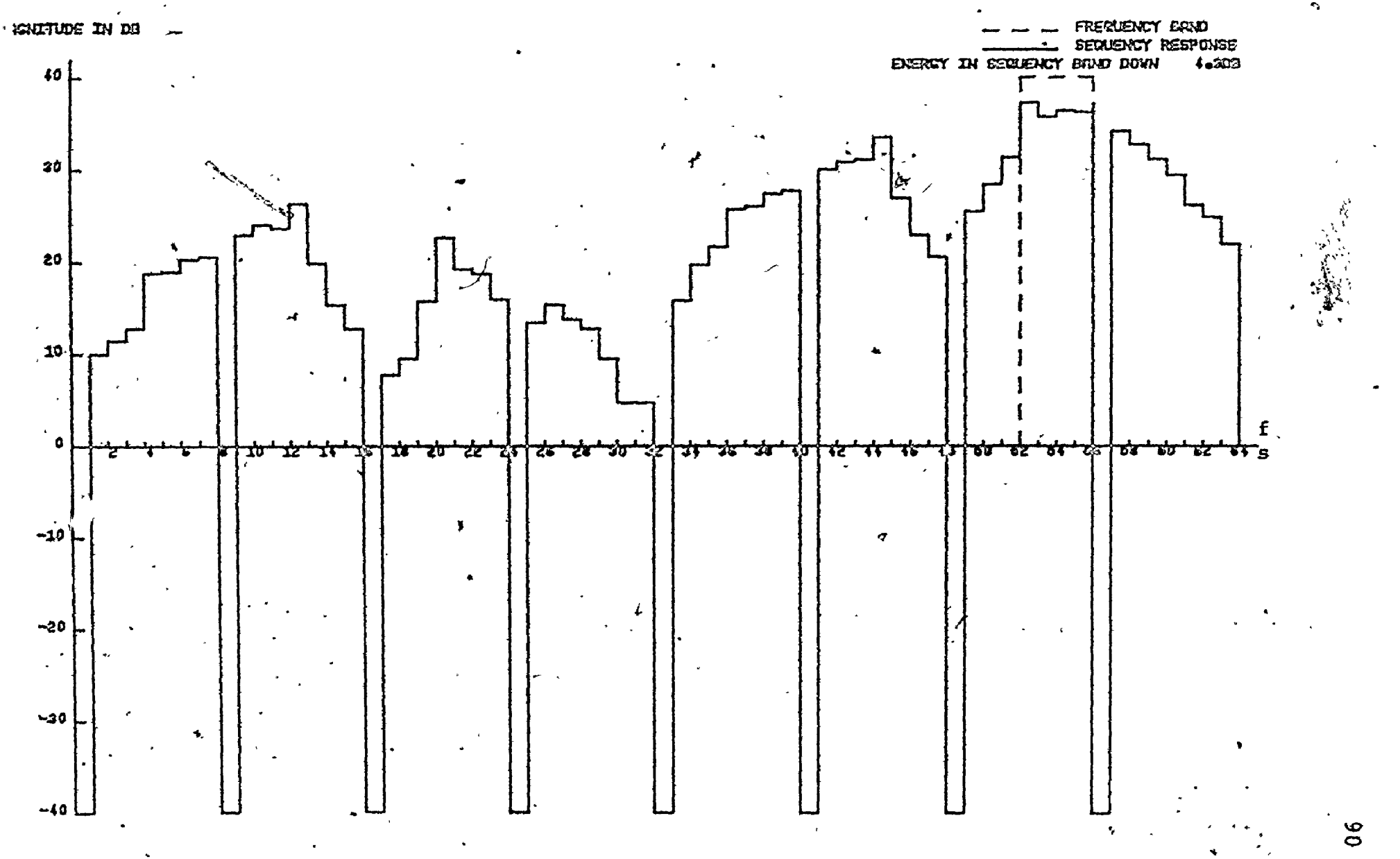


FIGURE 23 : Frequency to Sequency Conversion ; Band #14

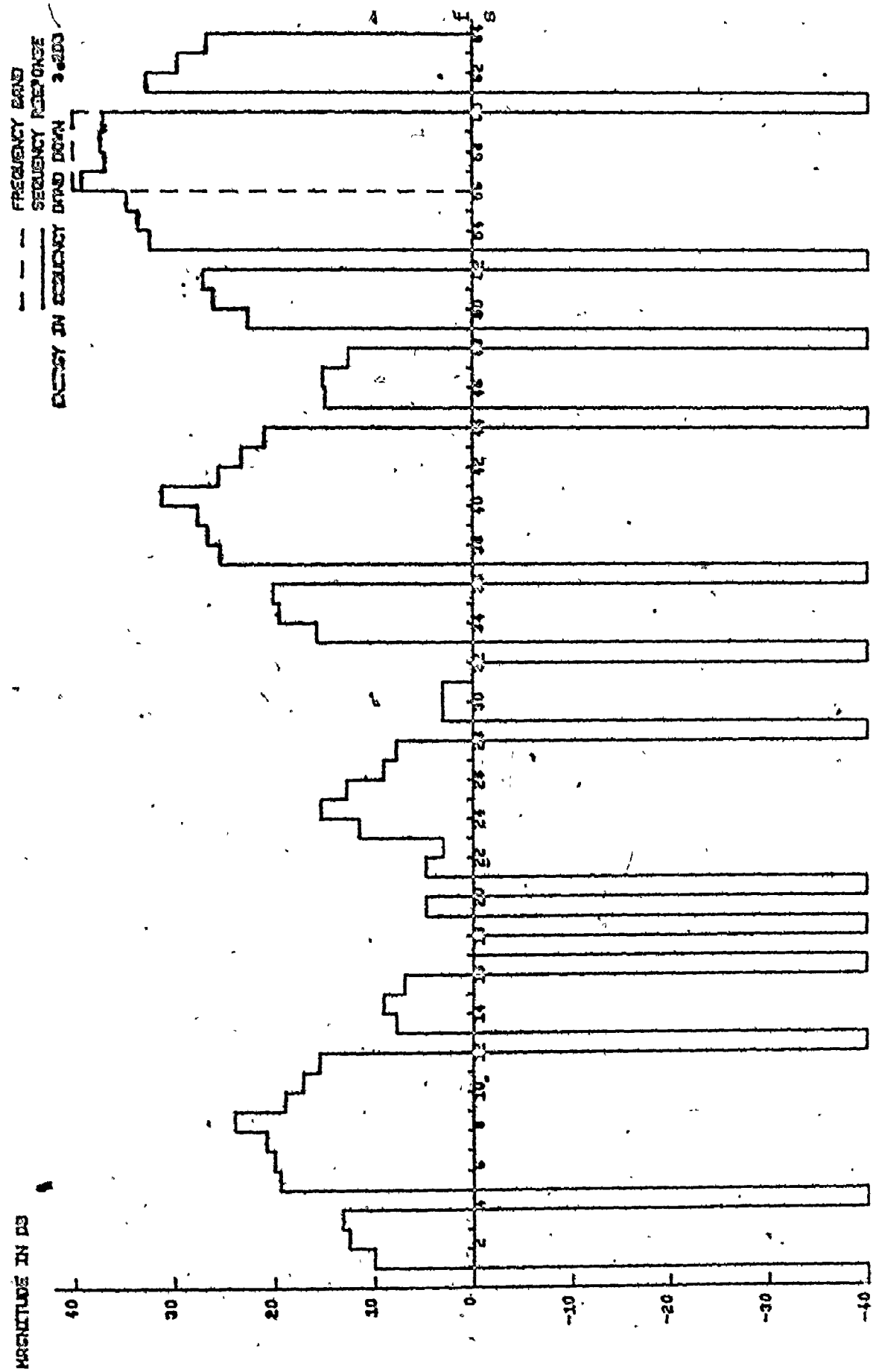


FIGURE 24 : Frequency to Sequency Conversion ; Band #15

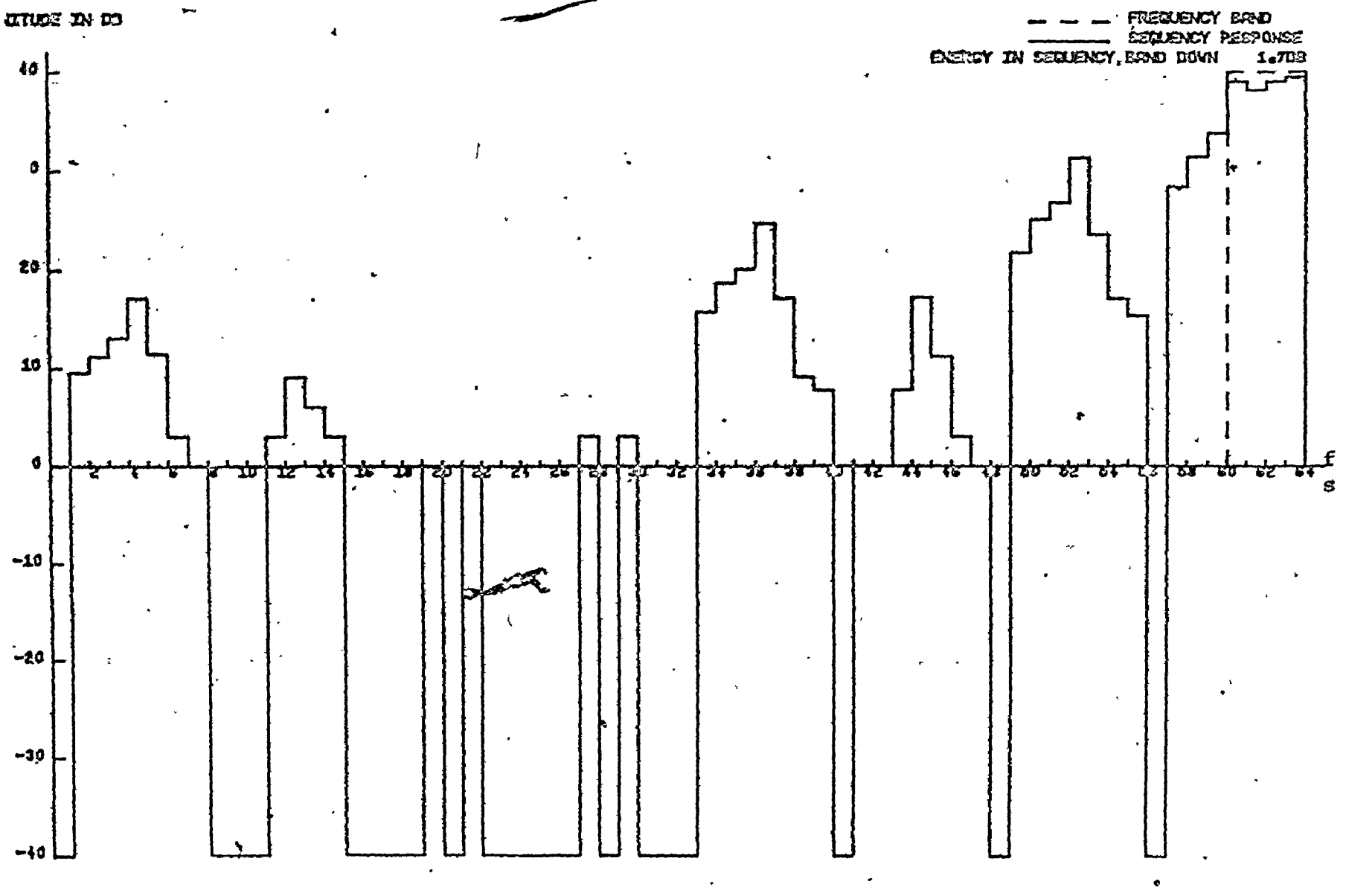


FIGURE 25 : Frequency to Sequency Conversion ; Band #16

CHAPTER V: THE WALSH VOCODER

Since a vocoder system operates in real time with speech signals, the speeds required of the FFT devices in the FFT vocoder necessitate hardware parallelism and high speed logic in the hardware implementation. This complexity and cost is reduced if FWT devices can replace the FFT devices because the hardware implementation of the FWT is less complex than the FFT implementation.

The basic configuration of the Walsh vocoder is similar to the FFT vocoder described in 2.3. The synthesized speech signals of both vocoder systems will be analyzed for comparative intelligibility and quality in order to determine whether or not FFT devices can be replaced with FWT devices in this vocoder configuration.

5.1 THE WALSH VOCODER SYSTEM

The Walsh vocoder system to be described can be separated, like all vocoder systems, into analysis and synthesis sections. Some operations performed in the synthesis of the voice signals are the reverse of the same operations performed in the analysis of voice signals.

Hence the hardware employed in the vocoder system can in some cases be multiplexed for use by the analyzer and synthesizer of the vocoder system. This would be the case where the analyzer and synthesizer are physically approximate to each other in a system that both transmits and receives vocoer signals in a two way communication system.

In the computer simulation of the Walsh and FFT vocoder systems, this hardware sharing finds expression in software subroutines which are called by both the analyzer and synthesizer sections of the simulation. Parameters are used to indicate whether forward or reverse operations are required for the analysis or synthesis operations respectively.

The input speech signal is sampled at 8KHz after lowpass filtering with a third order filter with a cutoff frequency of 3KHz. These samples are stored on a mass storage file since the simulation is not in real time. In the simulation magnetic tape was utilized as the mass storage medium to more closely approximate the sequential data access present in an actual real time vocoder system.

5.1.1 THE WALSH VOCODER ANALYZER

Since the N-length FWT is only applicable for $N=2^m$

mel$_+$, the Walsh spectra of the speech signals were calculated using a 128 length FWT. This allowed spectral envelope calculations every 16ms (for the 8KHz sampling rate) which is less than the spectrum sampling interval of 20ms employed in the analog channel vocoder described in 2.2. Increasing the window length would increase the spectral resolution but since the speech signal is not stationary ie. the power spectrum changes with time, the spectral changes would be missed since in effect longer window lengths result in a longer spectral sampling period, unless unnecessary window overlapping is employed.

The pitch parameters were determined with cepstrum pitch detection techniques in the simulation using 32ms windows to obtain adequate pitch detection resolution. The 256 length sample windows allow 64 different pitch values from 62.5Hz to 4000Hz on a nonlinear frequency scale. The frequency scale is the inverse of the linear quefrency (or time) scale. In effect there are 55 usable pitch parameters ranging from 62.5 Hz to 400 Hz (the pitch frequency rarely exceeds 400 Hz). In addition, a parameter is needed to represent unvoiced segments and this is set to 1000 Hz. These 56 pitch parameters are encoded with a 6 bit representation every 8ms. The 256 length sample windows are overlapped

with the previous sample window employed in the pitch detection to obtain pitch parameters every 8 ms. Longer sample windows would increase pitch parameter resolution but the possibility then increases that the longer window will contain multiple pitch parameters ie. the window could contain speech segments of different pitch values and the cepstrum pitch detector is not able to distinguish single and multiple pitch values.

The timing sequence for the pitch detection and Walsh spectral envelope determination is illustrated in FIGURE 26. Initially the 256 length array ITST is filled with speech samples in locations 64 to 255 with locations 0 to 63 filled with zeros. The following algorithm defines the sample transition through the analyzer:

1) samples 191 to 64 inclusive in ITST are shifted into the 128 length array IHAD in preparation for the FWT operation.
2) the first pitch parameter PF1 is determined from the 256 samples in ITST by subroutine CPITCH.
3) the samples in ITST are shifted down 64 places with samples 0 to 63 inclusive discarded ie. ITST(I)= ITST(I+64), I=0,191.

4) the next 64 speech samples are placed in ITST in locations 192 to 255 inclusive in that order as indicated in step 2 of FIGURE 26.

5) the second pitch parameter PF2 is now calculated from the 256 samples in ITST by subroutine CPITCH.

6) repeat step (3).

7) repeat step (4).

8) if transmit switch is still on, ie. if there are any more speech samples, repeat steps (1) through (7); if not repeat steps (1) and (2) and then go to (9).

9) PF2=PF1

10) STOP

The above algorithm defines two pitch parameters for every spectral envelope calculation without any interference of one operation with the other. This is depicted in FIGURE 26 where the timing sequence indicates the simultaneous determination of PF1 from the samples in ITST and the computation of the FWT from the samples in IHAD. This was necessary to simulate the independence of a hardware pitch detector with spectral envelope calculations in a hardware vocoder system.

From FIGURE 26 it can be seen that there are three

possible 128 length sample sets in ITST which could be used in the spectral envelope calculations. These sets are samples 0 to 127 inclusive, 64 to 191 inclusive or 128 to 255 inclusive. The only differences in choosing one set over the others are in the intialization of ITST and the validity of the pitch parameters associated with the spectral envelopes. Clearly the samples used to compute the FWT should also be somewhere near those used in the pitch detection. Choosing the set 64 to 191 inclusive for the spectral envelope calculation, we calculate PF1 with all of the samples used in the FWT operation as well as 64 samples of the previous and future sets used in the spectral envelope calculations. PF2 is determined from the 128 samples used in the present spectral calculations as well as the 128 samples from which the next spectral envelope will be determined. This sample allocation promotes smoother pitch transition from one window to the next in place of isolating two pitch parameters with one spectral envelope.

Once the FWT operation is completed, the 128 Walsh transform coefficients are placed in 128 length array IWAL and IHAD is free to accept the next 128 samples to be transformed into the Walsh domain. The Walsh sequency power spectrum is then calculated by squaring and adding Walsh coefficients of the same sequency. There are thus 64 sequency power

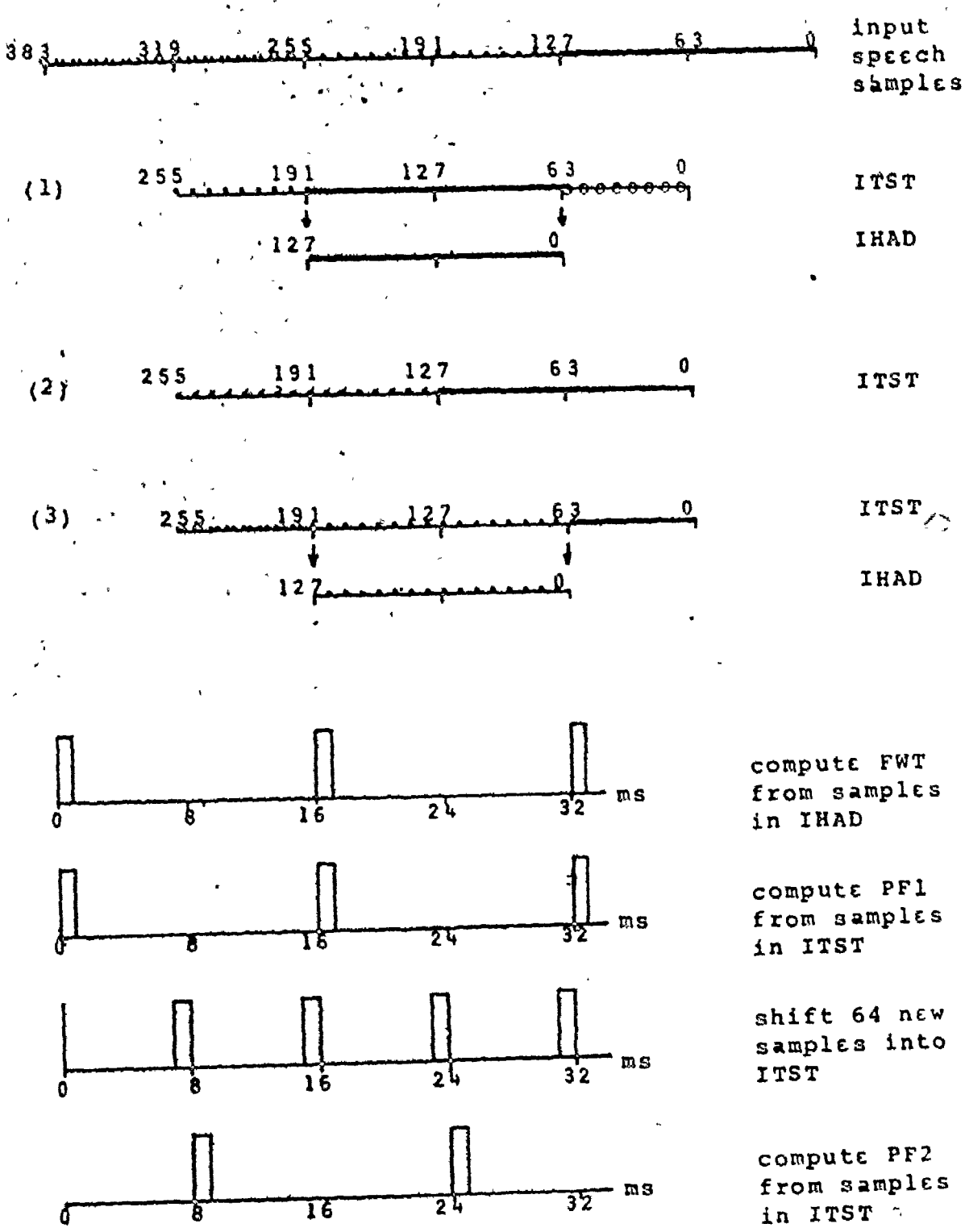


FIGURE 26 :Sample allocation and timing sequence for pitch detection and Walsh transform calculation.

spectral points which are used in the spectral envelope calculation.

The spectral envelope is calculated by the subroutine GSPS operating in the forward mode using the 64 sequency power spectral points which were placed in array SS. Subroutine GSPS averages the spectral points in a given sequency band yielding the average energy in that sequency band. These average energies in 14 such bands form a set of 14 channel values representing an approximation to the spectral envelope of that 16ms speech segment. These 14 channel values are placed in array BAND for input to a log encoder.

The output of the log encoder is not sent to the usual delta modulator in the simulation of the vocoder system since it has been shown that the 14 channel signals in a log encoded format can be further encoded such that the output of a channel to channel delta modulation yields a 26 bit representation of the set of channel values [8]. The two pitch parameters associated with this 16ms speech segment are encoded with a 6 bit representation for each. The 6 bit codes are then multiplexed with the delta modulator output. The result of this encoding would yield a 38 bit data stream

for each 16ms speech segment or a bit rate of 2375 bits/s.

The analyzer section of the Walsh vocoder simulation is depicted in FIGURE 27 in flowchart form. FIGURE 29 lists subroutine GSPS which calculates average sequency spectral energies in the 14 predefined spectral channels from the Walsh sequency power spectrum in the forward mode of operation ie. the analyzer has control of GSPS. The channel bandwidths and locations were determined from the long term average spectral envelope illustrated in FIGURE 40a. In FIGURE 40a the sequency axis represents a scale from dc to 4000 Hz for the 8KHz sampling rate. There are 64 sequency components for 16ms windows. The spectral envelope indicates that most of the spectral energy is located in the region dc to 2000Hz (0 to 32 sequency) and then a lower energy zone between sequencies 32 and 44 occurs followed by a higher energy zone. The 14 channel signals, which represent an approximation to the spectral envelope, are calculated in the following 14 spectral regions:

6 regions of bandwidth 125Hz (2 sequency components)

4 regions of bandwidth 250Hz (4 sequency components)

1 region of bandwidth 750Hz (12 sequency components)

3 regions of bandwidth 500Hz (8 sequency components)

The channel signals are calculated by averaging the sequency components whch lie in each of the respective spectral regions as indicated in FIGURE 29.

The bandpass regions of the sequency spectrum which are used in the channel signal calculations are contiguous but not equal. The widths of the bandpass regions were determined from the inspection of FIGURE 40a. The regions of the spectrum which contained more of the average signal energy were given harrower bandwidths in the channel signal calculations. This sequency spectral subdivision is different from the subdivision of the frequency spectrum as indicated in FIGURE 28 where subroutine GFPS (which is used in the FFT vocoder simulation) is listed. Here, the frequency spectrum is subdivided into 6 regions of bandwidth 125Hz, 4 regions of bandwidth 250Hz, 3 regions of bandwidth 500Hz and 1 region of bandwidth 750Hz. This subdivision was based on a long term average frequency spectral envelope illustrated in FIGURE 41a where the frequency components 0 to 63 represent the frequency scale dc to 4000Hz for the 8KHz sampling rate and 16ms windows. Comparing FIGURES 41a and 40a it can be seen why the sequency spectrum is subdivided differently from the frequency spectrum.

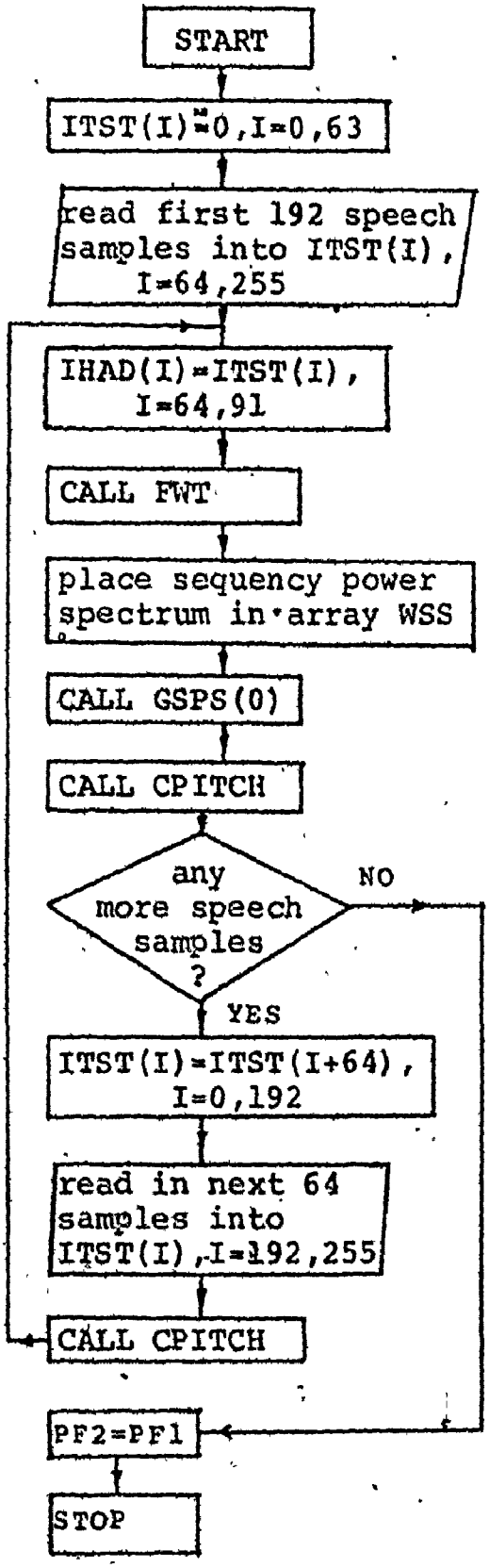


FIGURE 27:Flowchart of Walsh vocoder analyzer

From FIGURES 40a and 41a it is also apparent that the signal energy is spread more evenly in the sequency spectrum. This implies that a greater number of channel signals are required to represent the sequency spectrum than are required to represent the frequency spectrum for a given level of accuracy in the spectral envelope approximations. This can be explained by the greater correlation between sequency components than frequency components [6] and hence the channel signals of the Walsh vocoder will be more correlated and a channel to channel delta modulation scheme will be more effective in reducing the bit rate required to transmit the channel signals. However, in order to attain a comparison of the Walsh and FFT vocoders, both vocoder simulations utilized 14 channel signals as indicated in FIGURES 29 and 28 for the Walsh and FFT vocoder simulations respectively.

In the reverse mode of operation, GSPS (or GFPS) generates a 64 point sequency (or frequency) power spectrum from the channel values under the control of the synthesizer of the vocoder. Although the operations performed by GSPS (or GFPS) in the forward and reverse modes are not identical (in the reverse mode GSPS (or GFPS) does not perform

arithmetical operations) it is advantageous to use the same program to compress and generate the sequency (or frequency) power spectrum due to the exactly reverse memory operations. In a hardware configuration the sequency power spectrum would be stored in shift registers forming array WSS and array BAND would contain the channel signals. In the forward (or analyzer) mode of operation an operative averaging device would interconnect the two memory arrays. In the reverse (or synthesizer mode) the averaging device could operate as a demultiplexer. The sequency (or frequency) power spectral points in WSS that are averaged to yield a channel signal in BAND, in the analyzer, will be given that channel value in the synthesis of a speech signal. This spectrum generation and compression is illustrated in FIGURES 28 and 29. FIGURE 30 illustrates the sequency power spectrum of a typical 16ms speech signal and the sequency power spectrum generated by subroutine GSPS(1) from the channel signals computed by GSPS(0) for this 16ms speech segment. FIGURE 31 illustrates the frequency spectrum of this same 16ms speech segment and the corresponding frequency spectrum generated by GFPS(1) from the channel signals computed by GFPS(0). The frequency and sequency axis of FIGURES 31 and 30 respectively are indicated as components 0 to 63 which represents a scale of dc to 4000Hz (in both

```
                SUBROUTINE GFPS(IAS)
                COMMON WSS(64),BAND(14)
C IF IAS=1 THEN THE 64 POINT FREQUENCY POWER SPECTRUM IS TO
C BE GENERATED FROM THE 14 CHANNEL VALUES IN ARRAY BAND AND
C PLACED IN ARRAY WSS.
C IF IAS≠1 THEN THE CHANNEL VALUES ARE CALCULATED FROM THE
C FREQUENCY POWER SPECTRUM IN WSS AND PLACED IN BAND.
                IF(IAS.EQ.1) GO TO 10
                DO 1 I=1,14
1               BAND(I)=0
                DO 3 I=1,6
                DO 2 J=1,2
2               BAND(I)=BAND(I)+WSS(J+(I-1)*2)
3               BAND (I)=BAND(I)/2.
                DO 5 I=7,10
                DO 4 J=1,4
4               BAND (I)=BAND(I)+WSS(J+(I-1)*4-12)
5               BAND (I)=BAND(I)/4
                DO 7 I=11,13
                DO 6 J=1,8
6               BAND (I)=BAND(I)+WSS(J+(I-1)*8-52)
7               BAND (I)=BAND(I)/8.
                DO 9 I=14,14
                DO 8 J=1,12
8               BAND (I)=BAND(I)/12
                GO TO 100
10              DO 11 I=1.6
                DO 11 J=1,2
11              WSS(J+(I-1)*2)=BAND(I)
                DO 12 I=7,10
                DO 12 J=1,4
12              WSS(J+(I-1)*4-12)=BAND (I)
                DO 13 I=11,13
                DO 13 J=1,8
13              WSS(J+(I-1)*8-52)=BAND(I)
                DO 14 I=14,14
                DO 14 J=1,12
14              WSS(J+52)=BAND(I)
100             RETURN
                END
```

FIGURE 28: Subroutine GFPS: Simulation of FFT vocoder channel signal calculations or frequency spectrum generation for analyzer and synthesizer modes respectively.

```
      SUBROUTINE GSPS(IAS)
      COMMON WSS(64), BAND(14)
C IF IAS=1 THEN THE 64 POINT SEQUENCY POWER SPECTRUM IS
C TO BE GENERATED FROM THE 14 CHANNEL VALUES IN ARRAY
C BAND AND PLACED IN ARRAY WSS.
C IF IAS≠1 THEN THE CHANNEL VALUES ARE CALCULATED
C FROM THE SEQUENCY POWER SPECTRUM IN WSS AND
C PLACED IN BAND
      IF(IAS.EQ.1) GO TO 10
      DO 1 I=1,14
1     BAND(I)=0
      DO 3 I=1,6
      DO 2 J=1,2
2     BAND(I)=BAND(I)+WSS(J+(I-1)*2
3     BAND(I)=BAND(I)/2.
      DO 5 I=7,10
      DO 4 J=1.4
4     BAND (I) = BAND(I)+WSS(J+(I-1)*4-12)
5     BAND(I)=BAND(I)/4.
      DO 7 I=11,11
      DO 6 J=1,12
6     BAND(I)=BAND(I)+WSS(J+28)
7     BAND(I)=BAND(I)=BAND(I)/12.
      DO 9 I=12,14
      DO 8 J=1,8
8     BAND(I)=BAND(I)+WSS(J+(I-1)*8-48)
9     BAND(I)=BAND(I)/8.
      GO TO 100
10    DO 11 I=1,6
      DO 11 J=1,2
11    WSS(J+(I-1)*2)=BAND(I)
      DO 12 I=7,10
      DO 12 J=1,4
12    WSS(J+(I-1)*2) = BAND (I)
      DO 13 I=11,11
      DO 13 J=1,12
13    WSS(J+28) = BAND (I)
      DO 14 I=12,14
      DO 14 J=1,8
14    WSS(J+(I-1)*8-48)=BAND(I)
      RETURN
      END.
```

FIGURE 29: Subroutine GSPS: Simulation of Walsh vocoder channel signal calculations or sequency spectrum generation for analyzer and pectively.

figures) subdivided into intervals of 62.5Hz The different spectral allocations for the channel signals of the Walsh and FFT vocoder simulations can be seen from the inspection of FIGURES 30 and 31 respectively. Also, a resemblance between the frequency and sequency spectra for this speech segment can also be noticed. This resemblance is not the general case as can be seen from inspection of FIGURES 40a and 41a.

5.1.2 THE WALSH VOCODER SYNTHESIZER

The structure of the synthesizer is similar to the analyzer in reverse. The incoming data stream is demultiplexed and decoded to yield 14 channel signals and two pitch parameters every 16ms. In the simulation the decoded channel signals are placed in array BAND and the pitch values in Hz are placed in locations PF1 and PF2.

The goal now is to convolve a pitch pulse train derived from the pitch values and the impulse response function obtained by the inverse FWT of the Walsh spectrum represented by the 14 channel values.

Since the pitch pulse train is a series of impulses, this convolution is achieved by additions of delayed impulse

response functions where the delay is determined by the spacings of the pitch pulses.

In the simulation of the vocoder the pitch pulse generator was simulated by means of a 128 bit register IPPS which is called the pitch pulse scale. This pitch pulse scale represents a 16ms time axis with 125 μs subdivisions. The presence of the value 1 in a location in IPPS represents an impulse in the pitch pulse scale at that point in time. The separation between impulses is determined by the pitch values located in PF1 and PF2. The pitch pulse scale IPPS is set up according to the following algorithm:

1) initialize IPPS(1)=1; IPPS(I)=0,I=2,128;LI=1
2) SD1=4000/PF1; SD2=4000/PF2
3) NI=LI+SD1
4) IF(NI.GT.65) GO TO (9)
5) IF(NI.LE.0) GO TO (3)
6) IPPS(NI)=1
7) LI=NI
8) GO TO (3)
9) NI=LI+SD2
10) IF(NI.GE.129) GO TO (14)
11) IPPS(NI)=1

12) LI=NI

13) GO TO (9)

14) LI=LI-128

15) READ IN NEW VALUES OF PF1 AND PF2

16) WAIT UNTIL READY FOR NEW PITCH PULSE SCALE THEN GO TO (2)

The last impulse due to PF1 cannot go beyond IPPS(65) because PF1 is the pitch parameter associated with the 8ms time frame represented by IPPS(I),I=1,64 and PF2 is associated with IPPS(I), I=65,128. The only overlap allowed is at the end of one and beginning of the other time frame.

For an example of how the pitch pulse scale is set up consider PF1=125Hz and PF2=250Hz. At the end of step (14) in the above algorithm IPPS(I)=1 for I=1,33,65,81,97,113 and LI=1.

In order to attain a smoother transition between adjacent spectral envelopes, three more sets of 14 channel signals are determined from a linear interpolation between two adjacent sets of channel signals. In this manner the spectral envelope approximations are updated every 4ms.

Subroutine GSPS, operating in the synthesizer mode creates a 64 point sequency power spectrum from the channel signals stored in BAND and places the power spectrum in array WSS. This operation is listed in FIGURE 29.

A Walsh coefficient spectrum is now created from the Walsh sequency power spectrum contained in WSS. Array IHAD is set up with every cal term in the Walsh coefficient spectrum given a value determined from the square root of the power spectral point of the same sequency in WSS. Every second cal term was premultiplied by -1 to minimize the discontinuities at the edges of the windows since all cal terms have values +1 at t=0. The sal terms in the Walsh coefficient spectrum are preset to zero for simplicity.

Subroutine FWT converts the Walsh transform coefficients in IHAD into the time domain yielding a I28 length impulse response function. (The inverse FWT is identical to the FWT except for a multiplying factor as shown in Chapter IV). The four resulting impulse response functions are stored in 128 length arrays IRFI, IRF2, IRF3, and IRF4 in chronological order.

The pitch pulse scale corresponding to the 16ms time frame associated with the four impulse response functions is subdivided into 4 equal 32 length sections. Each section will be associated with the respective impulse response function. If a 1 appears in IPPS(I), I=1,32 then IRF1 is added to the current output component, which is stored in 256 length array

IOUT, starting at IOUT(I). Similarly if IPPS(J)=1,J=33,64, then IRF2 is added to IOUT starting at IOUT(J), and so on, IRF1, IRF2 and IRF3 are determined from the inverse transform of sequency power spectra determined from linear interpolations between adjacent transmitted channel signals to attain smoother transitions between spectral windows. IRF1 is associated with the first 4ms segment of IPPS ie. IPPS(I),I=1,32 and IRF2 is associated with the next 4ms segment ie. IPPS(I),I=33,64. These impulse response functions are then associated with PF1 which is used to set up IPPS(I),I=1,64. Similarly IRF3 and IRF4 are associated with the next two 4ms segments of IPPS respectively and hence with PF2. Recall that PF1 was determined from the 128 samples used to compute the current set of channel signals as well as 64 samples from the previous and future spectral calculations. IRF1 and IRF2 are also linked with the past spectral calculations by virtue of a linear interpolation between the past and current channel values. PF2 was determined from the samples used to compute the current channel signals and the 128 samples which will be used to compute the next set of channel signals. IRF3 is still associated with the past set of channel signals by the interpolation mentioned and IRF4 is determined solely from the current channel values. In this manner, there is a

current and future speech samples

The summation of impulse response functions as dictated by the pitch pulse scale to form the output synthesized speech samples is illustrated in FIGURE 33. The formation of a 16ms output component is illustrated as a summation of delayed impulse response functions. The impulse response functions and the output component are truncated at the 16ms time interval. Actually, the impulse response functions are all of length 16ms (or 128 samples). The truncated portions would be added to the output component beyond 16ms.

This process is continually repeated with the creation of a new pitch pulse scale and four impulse response functions. The output component array IOUT is backspaced 128 places every 16ms. These backspaced 128 samples of IOUT form the output synthesized speech samples which are to be desampled for an analog speech waveform output.

A flowchart of the Walsh vocoder synthesizer appears in FIGURE 34.

5.2 THE FFT VOCODER SIMULATION

The FFT vocoder is simulated in the exact same manner as the Walsh vocoder with FFT operations replacing FWT operations. The terms sequency power spectrum and Walsh coefficient spectrum in the Walsh vocoder simulation are replaced by the

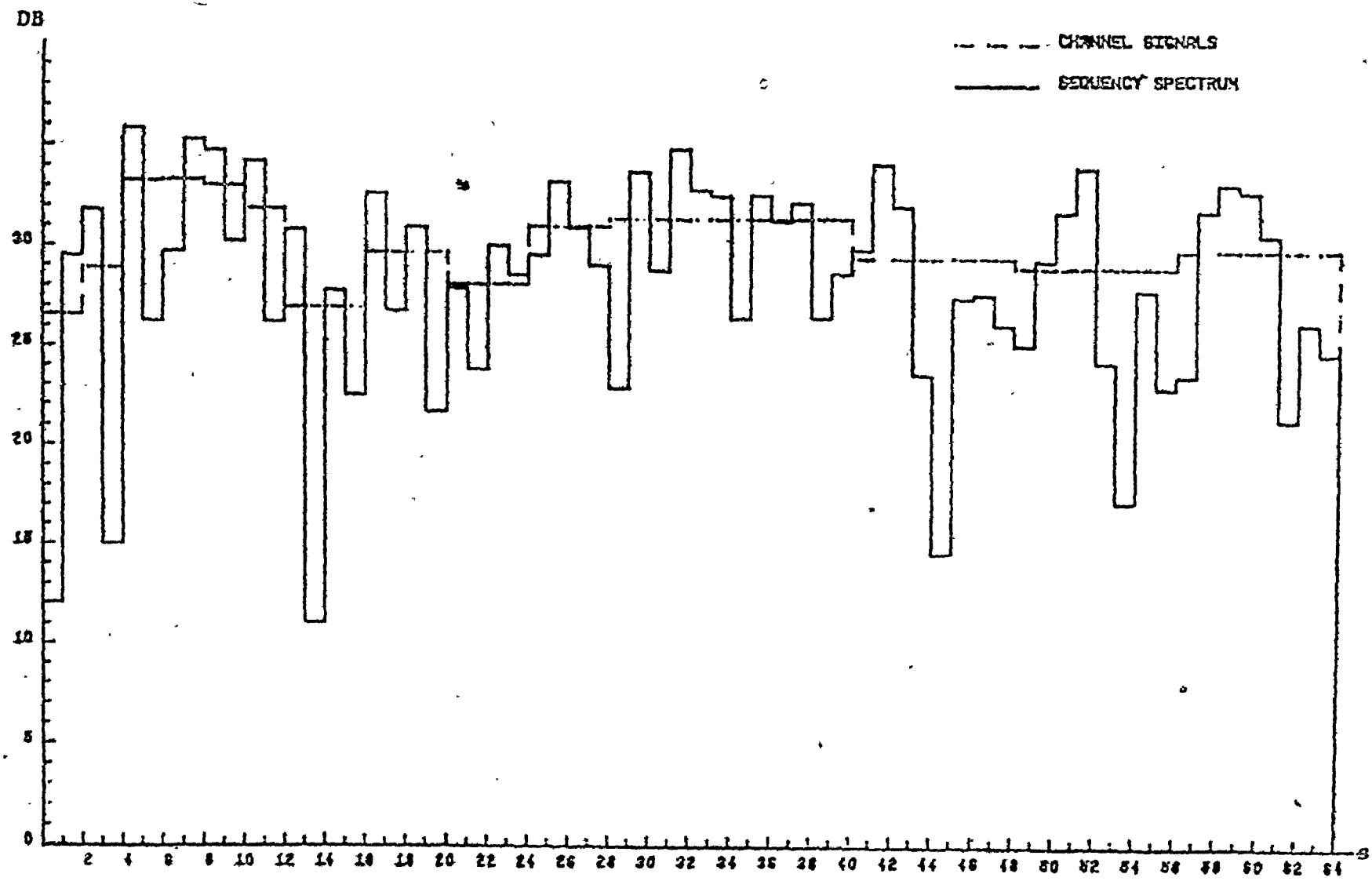


FIGURE 30 :Sequency spectrum & the corresponding channel signals derived by the simulated Walsh vocoder analyzer for a 16ms speech segment.

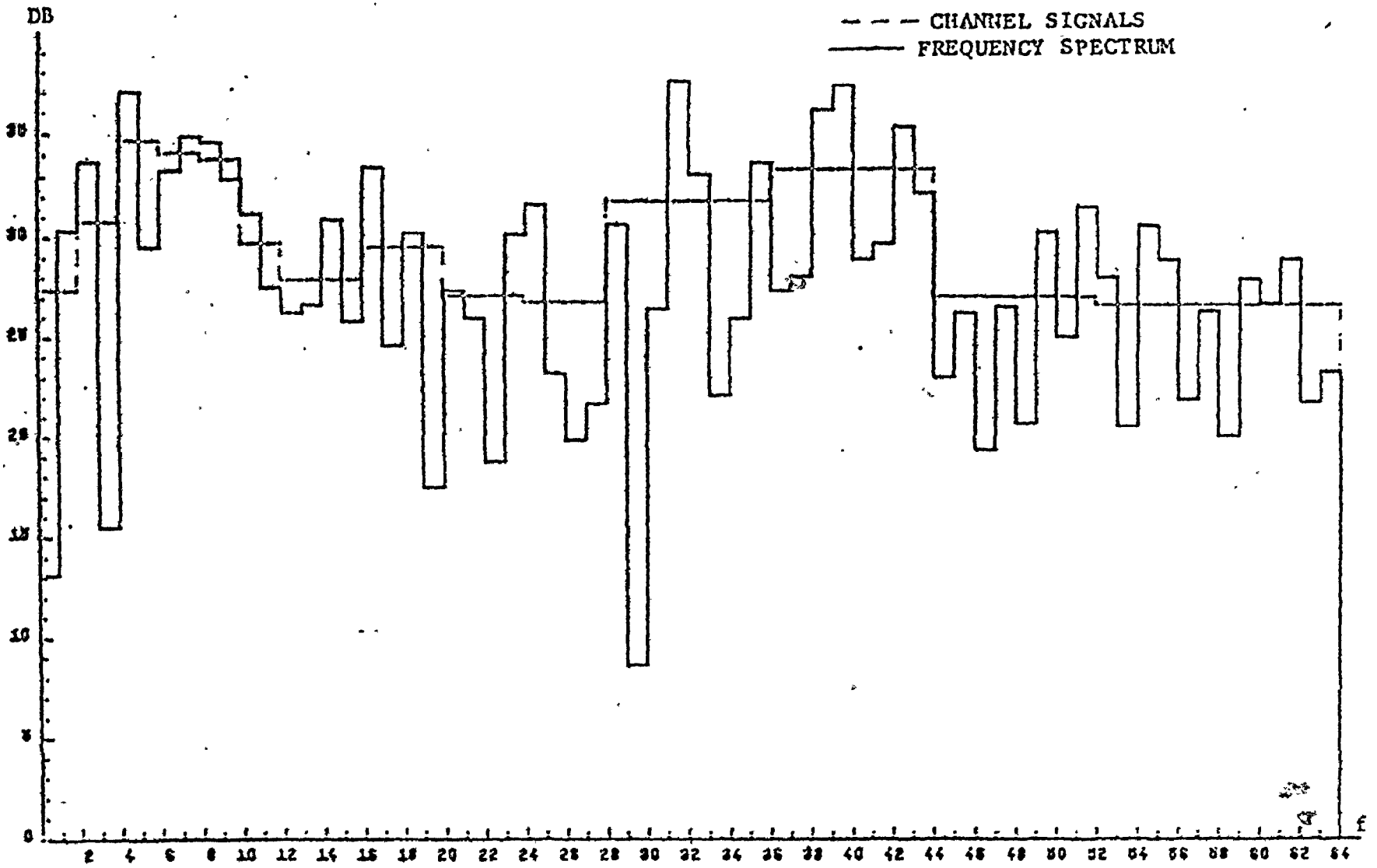


FIGURE 31 :Frequency spectrum & the corresponding channel signals derived by the simulated FFT vocoder analyzer for the same 16ms speech segment used in FIGURE 30.

TIME SEQUENCE OF CEPSTRUM PITCH DETECTION FOR <SPEED>

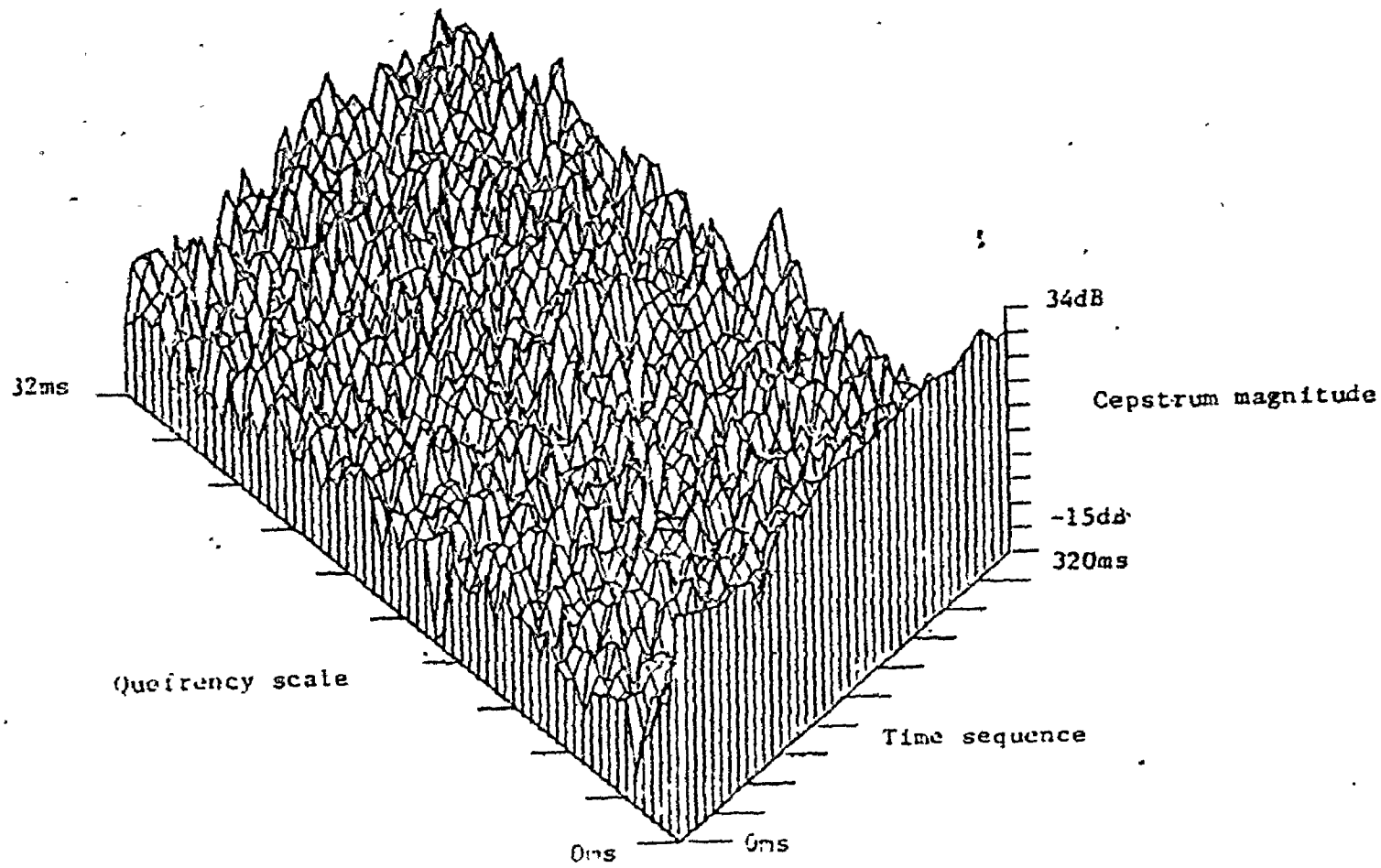


Fig 11 32 :Cepstrum for speech signal "speech" (duration 320ms) calculated every 8ms

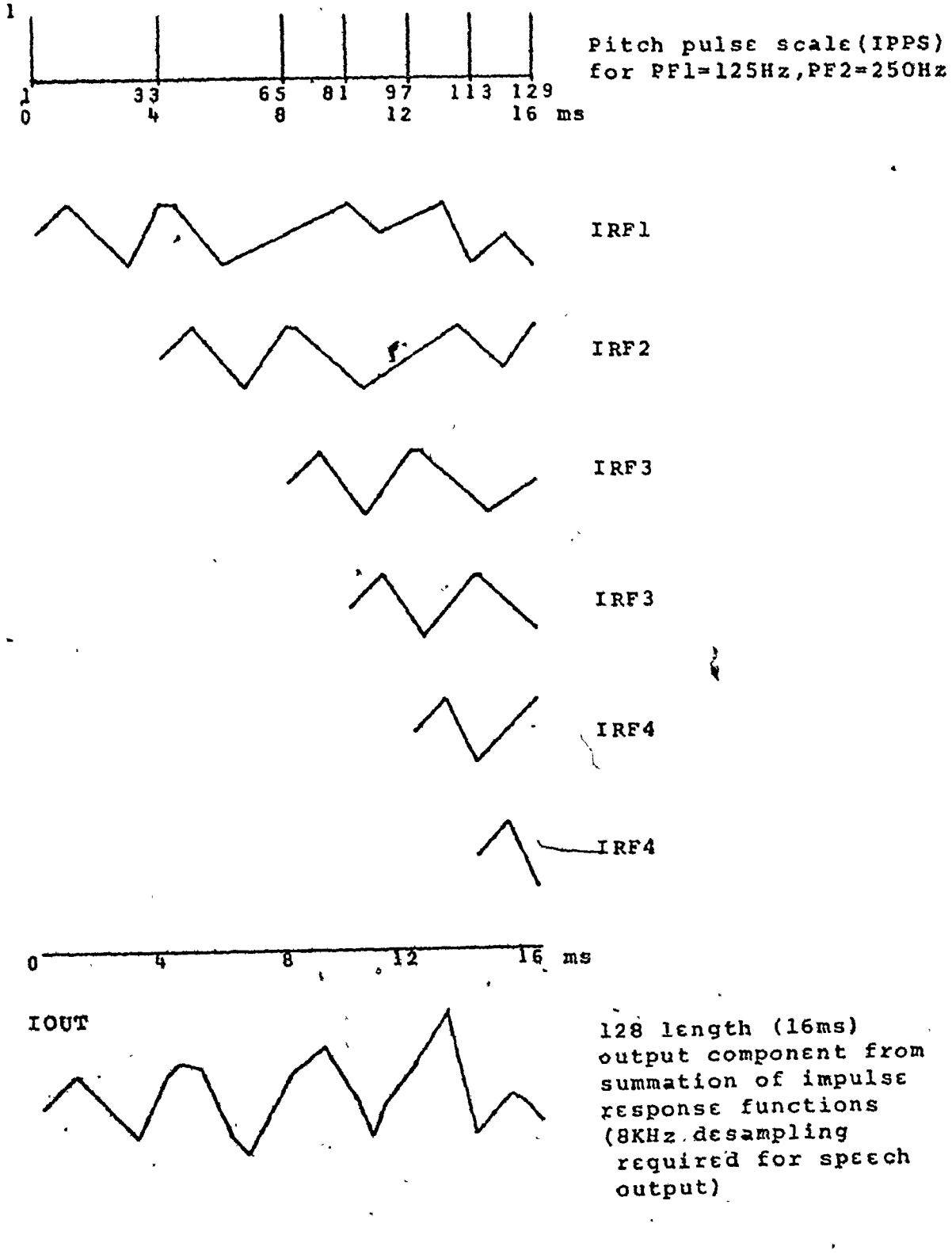
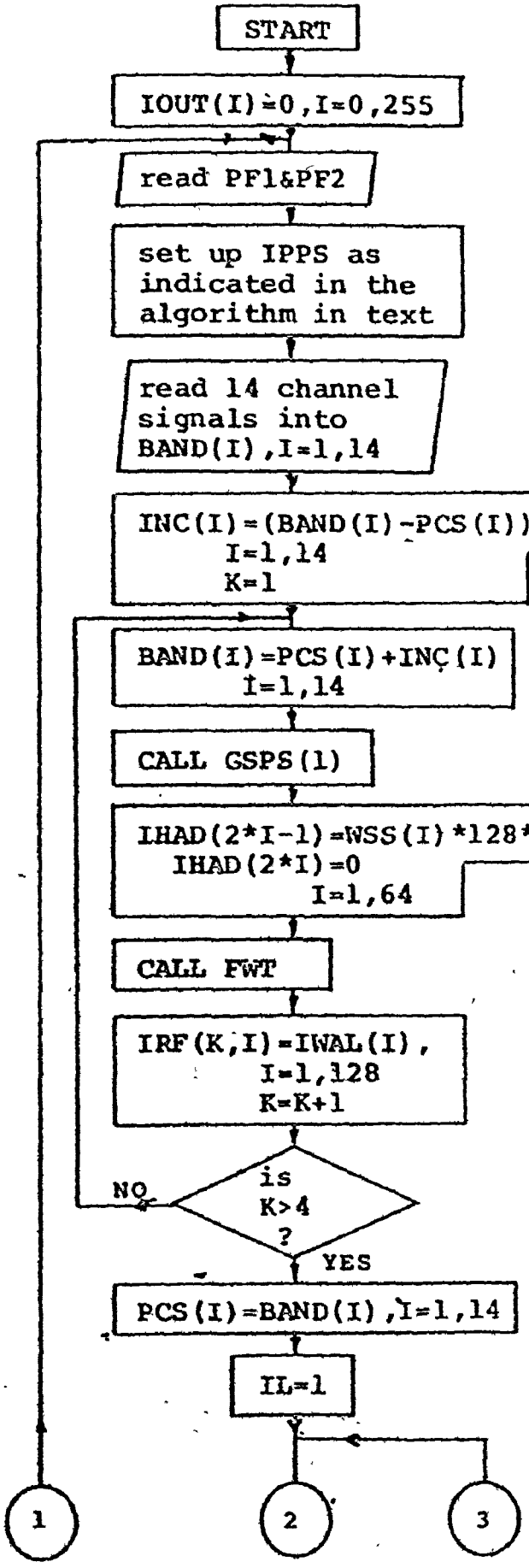


FIGURE 33 :Synthesized speech output from pitch pulse scale & delayed impulse response functions.



/IOUT contains the
/current output samples

/PF1&PF2 are the pitch
/parameters associated
/with this 16ms speech
/segment to be synthe-
/sized(they are used
/to set up IPPS,the
/pitch pulse scale)

/PCS contains the chan-
/nel signals of the
/previous 16ms segment.

/GSPS(1) generates a
/sequency power spect-
/rum in WSS.

/FWT converts the
/Walsh coefficients in
/IHAD into the time
/domain samples placed
/in IWAL.
/IRF(K,I) is the 128
/length impulse response
/function IRFK,K=1,4.

/continued

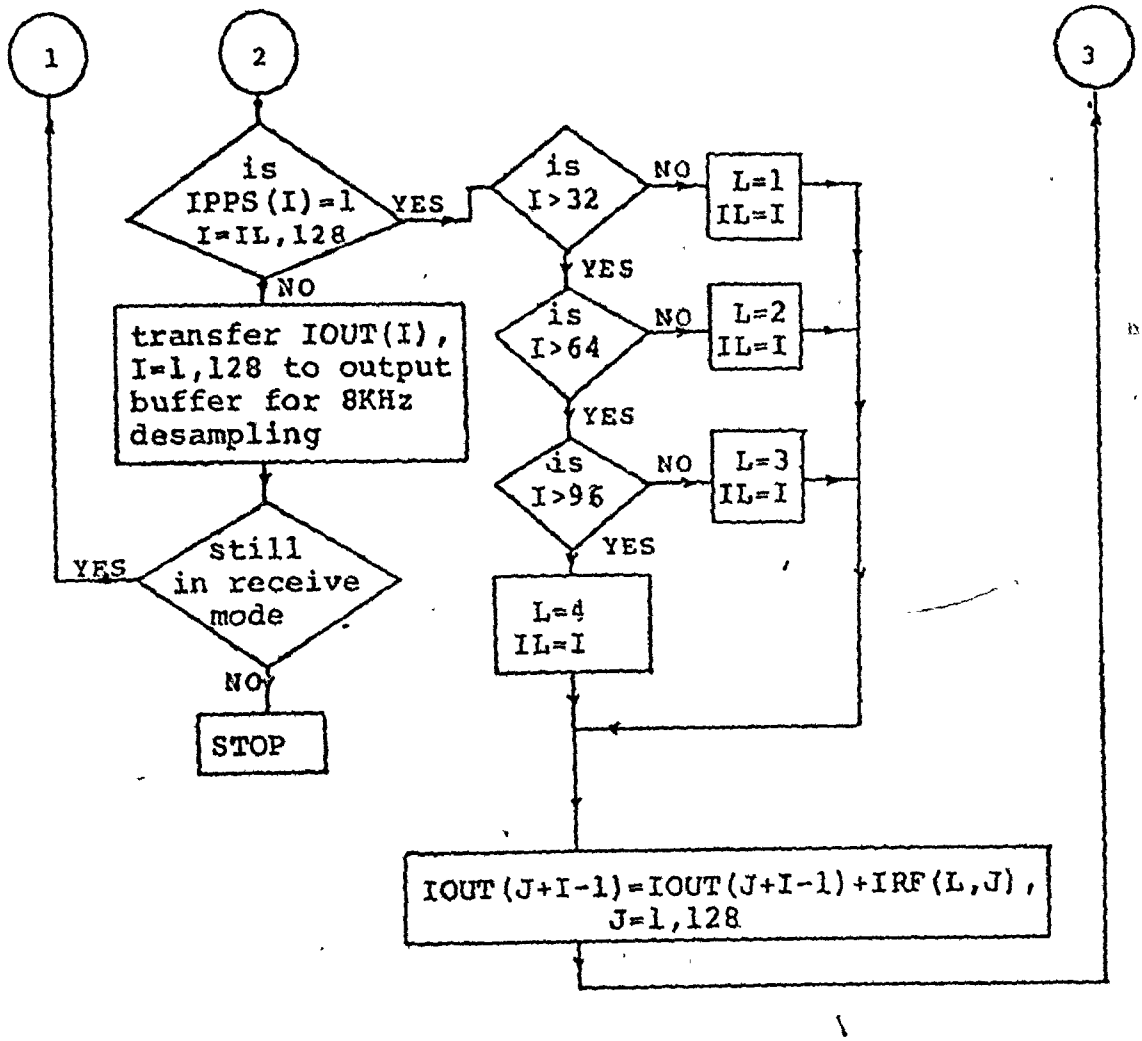


FIGURE 34 :Flowchart of Walsh vocoder synthesizer simulation

terms frequency power spectrum and Fourier coefficient spectrum respectively in the FFT vocoder simulation. Subroutine GFPS (FIGURE 28) is used in place of subroutine GSPS (FIGURE 29) because of different spectral subdivisions as previously mentioned in 5.1.1.

5.3 REMARKS CONCERNING THE SIMULATIONS

A hardware vocoder system would naturally contain a hardware pitch detector as would all real time practical vocoder systems. Pitch detection is a difficult task that has as yet not been resolved. An indication of the pitch detection problem, even in software detection techniques, can be seen in FIGURE 32 where the cepstrum of the word "speed" is shown in a 3 dimensional view. The cepstrum is calculated every 8ms for the 320ms speech segment. The cepstrum pitch detector must locate peaks in the cepstrum and the locations of these peaks yield the reciprocal of the pitch. If no peaks above a set threshold value are found, then the segment is unvoiced. An example of the cepstrum of unvoiced speech segments can be seen in the first segments of the word "speed" in FIGURE 32. There are no distinguishable peaks. Near the end of the time sequence the vowel sounds, which are

voiced, are distinguished by the cepstra which have prominent peaks. However, even these peaks might be located at multiple values of the pitch period. The cluttered image of the cepstra serves to indicate the difficulty of pitch detection even in sophisticated software techniques like cepstrum pitch detection.

FIGURE 35 illustrates a time sequence of the Walsh spectra of the word "speed" before synthesis and FIGURE 37 illustrates the Fourier spectra of the same 320ms speech segment. FIGURES 36 and 38 illustrate the time sequence of the Walsh and Fourier spectra after synthesis in the Walsh and FFT vocoder simulations respectively. In each of the above four cases a 16ms window (or 128 sample window) was utilized in the spectral calculations. The sequency and frequency axis represent a scale from dc to 4000Hz with 64 spectral coefficients.

Comparing FIGURES 35 and 37, the similarities between the sequency and frequency spectra can be noticed. The first spectra in the sequence are flat and contain low energy. These spectra are determined from the unvoiced fricative sound "s" in "speed". The "ee" sound produces jagged spectra and the repetitive peaks in these spectra indicate the presence

of pitch. Another plosive "d" follows and the flatness of this spectrum especially in the low frequency portion of the Fourier spectrum, can also be seen.

FIGURES 36 and 38 indicate the somewhat preservative nature of the spectra of the speech segments of the Walsh and FFT vocoder simulations respectively. Without the pitch information, these spectra would contain perfectly flat regions in the areas of the spectrum where the channel signals were calculated. (see FIGURES 30 and 31 respectively). The fact that the spectra illustrated in FIGURES 36 and 38 do not exactly match those of FIGURES 35 and 37 respectively indicates that the synthesized speech signals are not identical to the original and hence the synthesized speech output will be "different" from the original speech input. This difference is usually manifested in a degredation in quality and intelligibility of the output speech sounds with respect to the input speech sounds. Some speech sounds will be distorted more than others as can be seen from the segments in FIGURES 36 and 38 which are poorer approximations to the corresponding segments of FIGURES 35 and 37 respectively.

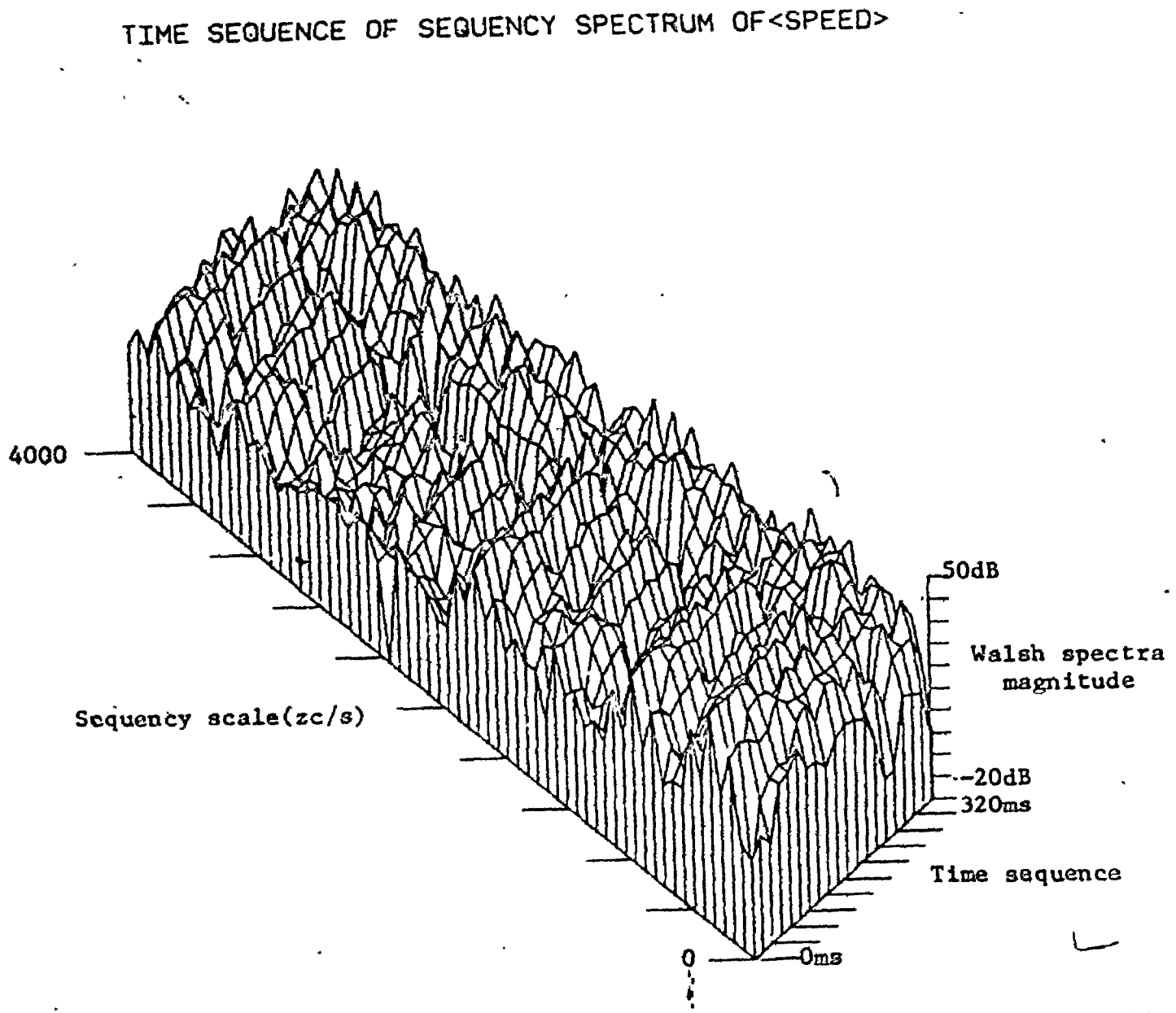


FIGURE 35: Walsh spectra of speech signal "speed" (duration 320ms) calculated every 16ms.

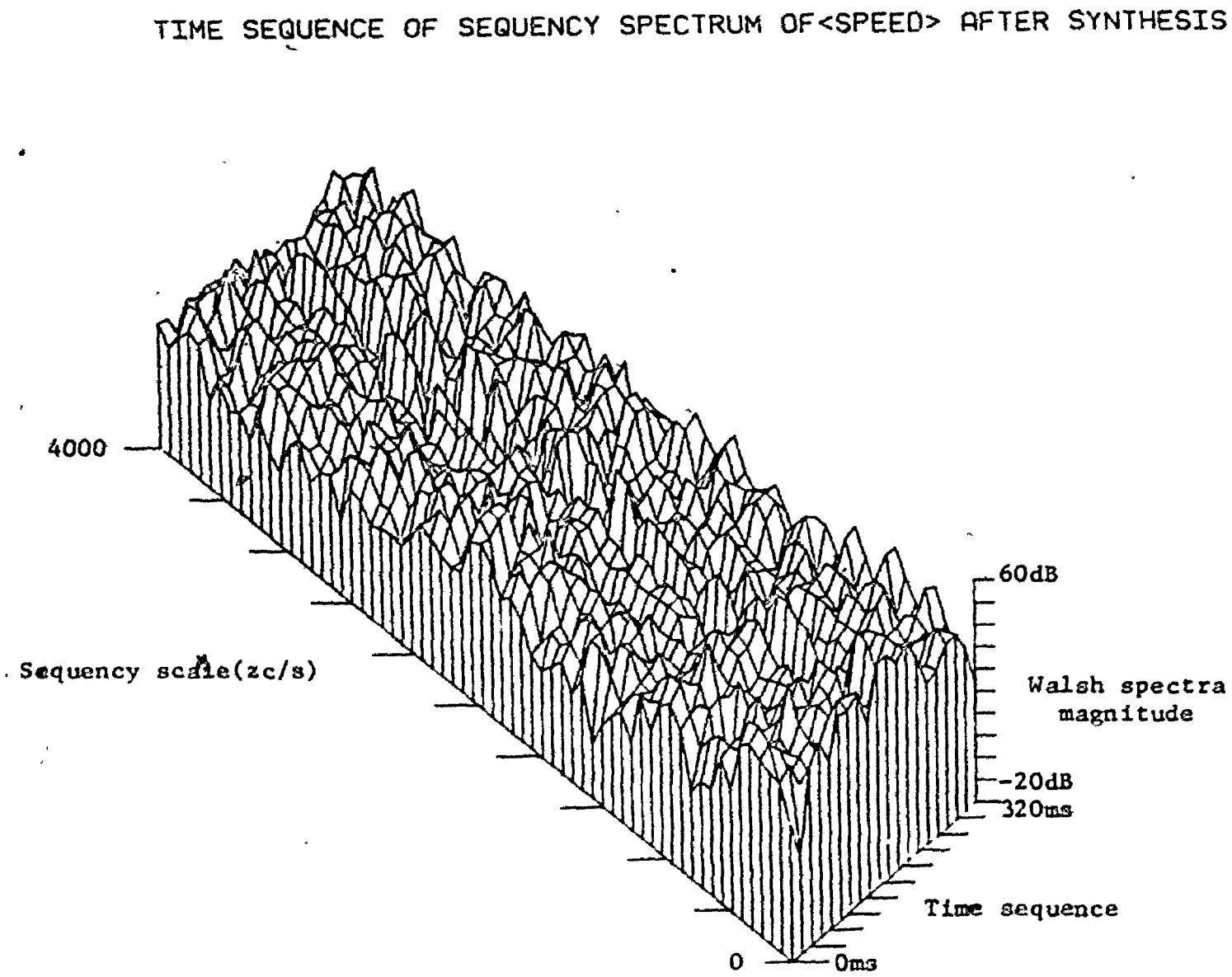


FIGURE 36: Walsh spectra of synthesized speech signal "speed" (duration 320ms) calculated every 16ms.

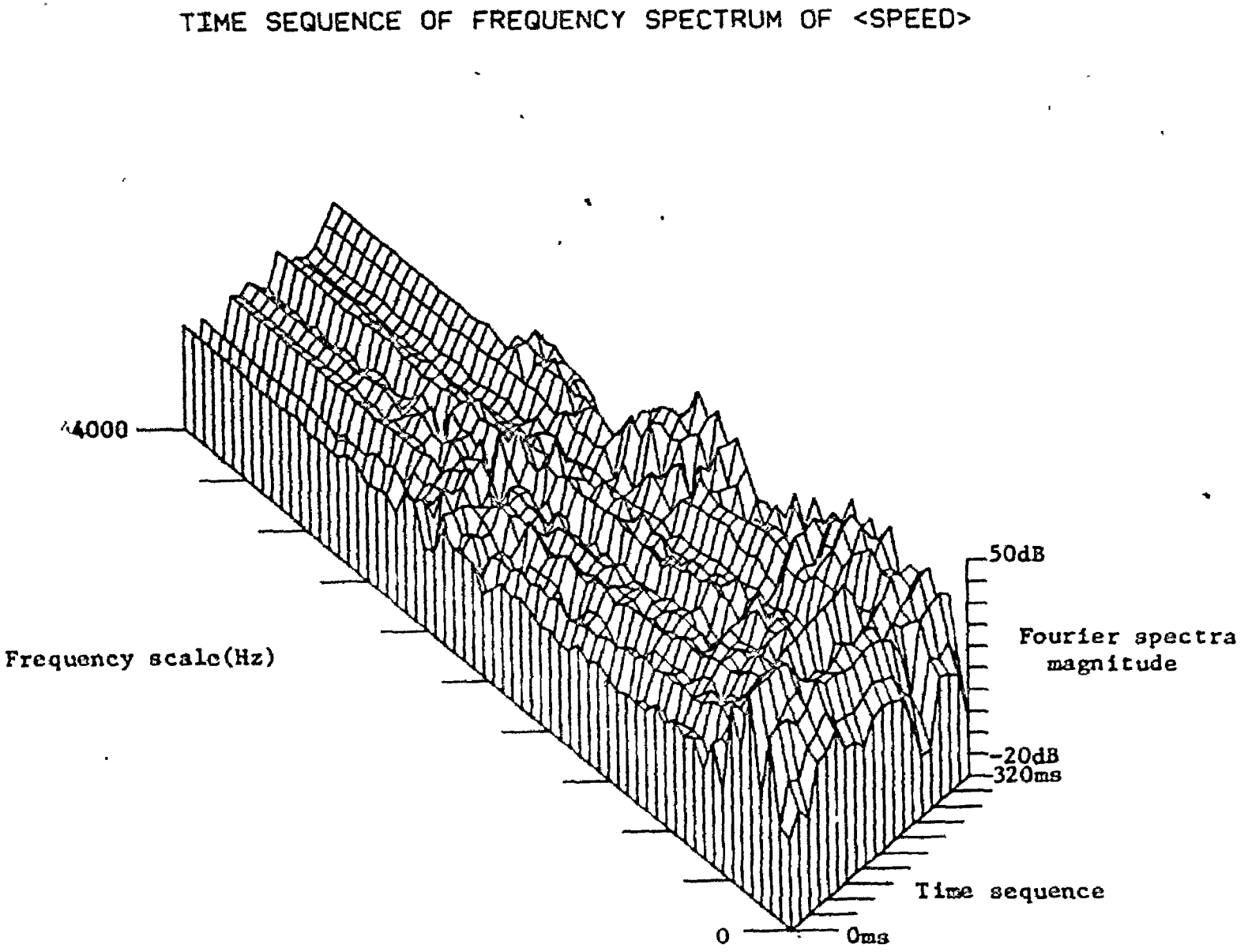


FIGURE 37: Fourier spectra of speech signal "speed" (duration 320ms) calculated every 16ms.

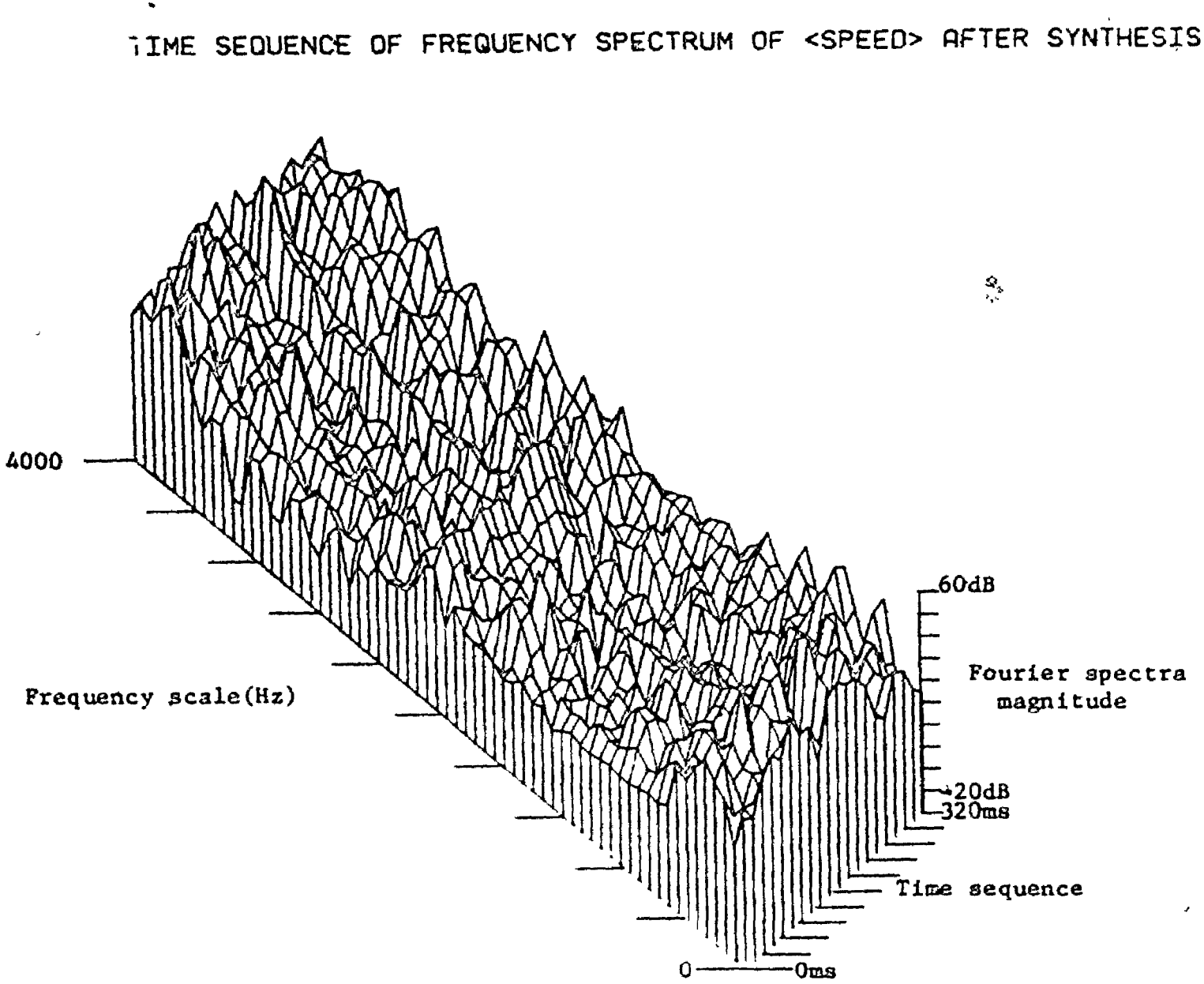


FIGURE 38: Fourier spectra of synthesized speech signal "speed" (duration 320ms) calculated every 16ms.

5.4 TIME INVARIANT SEQUENCY POWER SPECTRUM VS SPECTRAL ENVELOPE DETECTION

The information extracted from input speech signals in the analyzers of the Walsh, FFT, and analog spectrum channel vocoders is the spectral envelope and the pitch parameters associated with segments of the speech signals. An approximation to the spectral envelope is represented by a set of channel signals which are derived by calculating the energy in arbitrary adjacent bands of the spectrum. The number of channel signals is chosen such that a sufficient approximation to the spectral envelope is attained with a minimum number of channels. The portion of the spectrum from which a channel signal is derived utilizes knowledge of the human aural mechanisms and long term spectral envelope inspections. For instance, inaccuracies in the lower end of the spectrum are more noticable than at higher spectral points (and more energy is concentrated at the lower spectral end) and therefore, the bands at the lower end of the spectrum are narrower ie. they contain fewer spectral points.

In Chapter IV the term time invariant sequency power spectrum was introduced. This spectrum is a measure of the harmonic content of a signal. Instead of approximating a

spectral envelope with channel signals derived from the energy content of adjacent spectral bands, the time invariant sequency power spectrum represents the average power content of a group of sequencies (or frequencies). Each group contains a fundamental sequency and all of the odd harmonies of that fundamental sequency.

A vocoder which utilizes the time invariant sequency power spectrum (TlSPS) was simulated in order to evaluate the possibility of the utilization of this spectrum in place of the spectral envelope approximations. It was not clear how the pitch parameters could be used to shape the TlSPS.

According to the model of the vocal apparatus, the spectral envelope was determined by the shape of the vocal tract and the pitch of voiced sounds was due to glottal excitation which consisted of a series of pulses produced at the pitch rate. The relationship between the spectral envelope and the pitch was the convolution of a pitch pulse train with an impulse response function determined by the inverse transform of the spectral envelope. There is no such known model which relates the TlSPS to the pitch.

Since the TlSPS is easily determined, a Walsh vocoder utilizing the TlSPS was simulated. The following hypothesis

was used as a basis for the T1SPS Walsh vocoder:
If speech can be represented by a combination of tones of various frequencies (or sequencies) and their harmonies (some musicians can make their instruments "talk"), then the T1SPS would approximate "tones" with spectral components whose frequencies are a function of the sampling rate. The T1SPS is an average measure of the power in these "tones" such that the spectrum derived from the T1SPS would have unlikely symmetries ie. every spectral point in each group would have the same magnitude. Hence the spectrum derived from the T1SPS was "shaped" by a long term average sequency spectral envelope.

The T1SPS was calculated for 32ms speech segments. This yielded 9 power spectral points. Subroutine T1PS(0) was used to calculate the T1SPS in the analyzer mode of the T1SPS Walsh vocoder from the Walsh sequency power spectrum determined from the Walsh coefficients produced by subroutine FWT. Subroutine TIPS(1), operating in the synthesizer mode, generates a shaped sequency power spectrum which is then used to create a Walsh coefficient spectrum in the same manner as indicated in the synthesizer of the Walsh vocoder. The Walsh coefficient spectrum is then transformed into the time domain

by FWT yielding 256 synthesized speech samples.

Since the average power in the group of frequency components which correspond to the sequency components of each respective group is exactly the same as the average power in that group of sequency components, TIPS(1) could also be used to generate a frequency power spectrum. This spectrum could then be shaped by a long term average frequency spectral envelope. The resulting spectrum could then be used to create a Fourier coefficient spectrum which can be transformed into the time domain by an inverse FFT operation yielding synthesized speech samples.

Subroutine T1PS is listed in FIGURE 39 and the long term average sequency and frequency spectral envelopes used in the T1SPS Walsh vocoder and T1SPS Fourier vocoder respectively are shown in FIGURES 40b and 41b. The envelopes used in the T1SPS vocoder simulations were normalized such that the largest component has a value of 64. This was done to compensate for the averaging effect of the T1SPS calculation. The largest sequency or frequency group of the T1SPS has 64 components and hence the average power is obtained by dividing the sum of the spectral components in that group by 64. The windows were of length 256 and hence the 128 frequency of sequency components represent a scale of dc to 4000Hz in divisions of 31.25Hz.

```
      SUBROUTINE TIPS(IAS)
      COMMON WSS(128),BAND(9)
C IF IAS=1 THEN THE 128 POINT SEQUENCY (OR FREQUENCY) SPECTRUM
C IS TO BE GENERATED FROM THE TISPS POINTS IN BAND AND PLACED
C IN WSS.
C IF IAS≠1 THEN THE TISPS POINTS ARE TO BE CALCULATED FROM
C THE SEQUENCY (OR FREQUENCY) SPECTRUM IN WSS AND PLACED
C IN BAND.
      N=256
      NN=N/2
      M=8
      IF(IAS.EQ.1) GO TO 10
      BAND(1)=WSS(1)
      BAND(2)=WSS(NN)
      DO 1 I=2,M
      II=I-1
      SUM=0
      FIX=FLOAT(2**(M-I))
      JU=2**(M-I)
      DO 2 J=1,JU
      INX=2**(I-2)+(J-L)*(2**II)+1
2     SUM=SUM+WSS(INX)
1     BAND(I+1)=SUM/FIX
      GO TO 9
10    WSS(1)=BAND(1)
      WSS(NN)=BAND(2)
      DO 7 I=2,M
      II=I-1
      JU=2**(M-I)
      DO 7 J=1,JU
      INX=2**(II-1)+(J-1)*(2**II)+1
7     WSS(INX)=BAND(I+1)
C SUBROUTINE MODIFY SHAPES WSS BY MULTIPLYING THIS SEQUENCY (OR
C FREQUENCY) SPECTRUM WITH A LONG TERM AVERAGE SEQUENCY (OR
C FREQUENCY) SPECTRAL ENVELOPE SHOWN IN FIGURE 40b (AND
C FIGURE 41b)
      CALL MODIFY
9     RETURN
      END
```

FIGURE 39: Subroutine TIPS: Simulation of the TISPS Walsh and Fourier vocoder spectral calculations or spectrum generation for analyzer and synthesizer modes respectively.

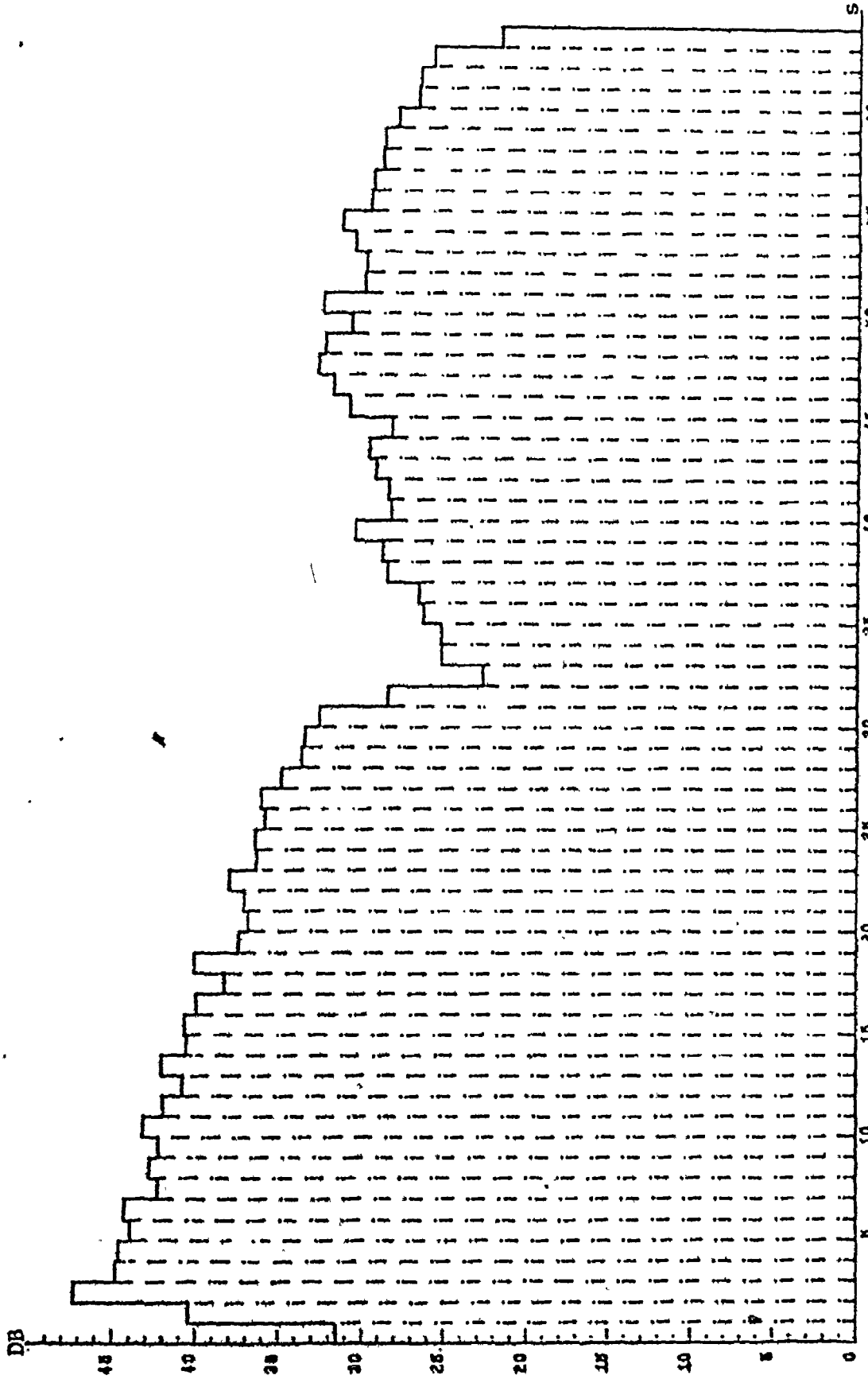


FIGURE 40a :Long term average sequency spectral envelope for 16ms speech segments sampled at 8Kiz.

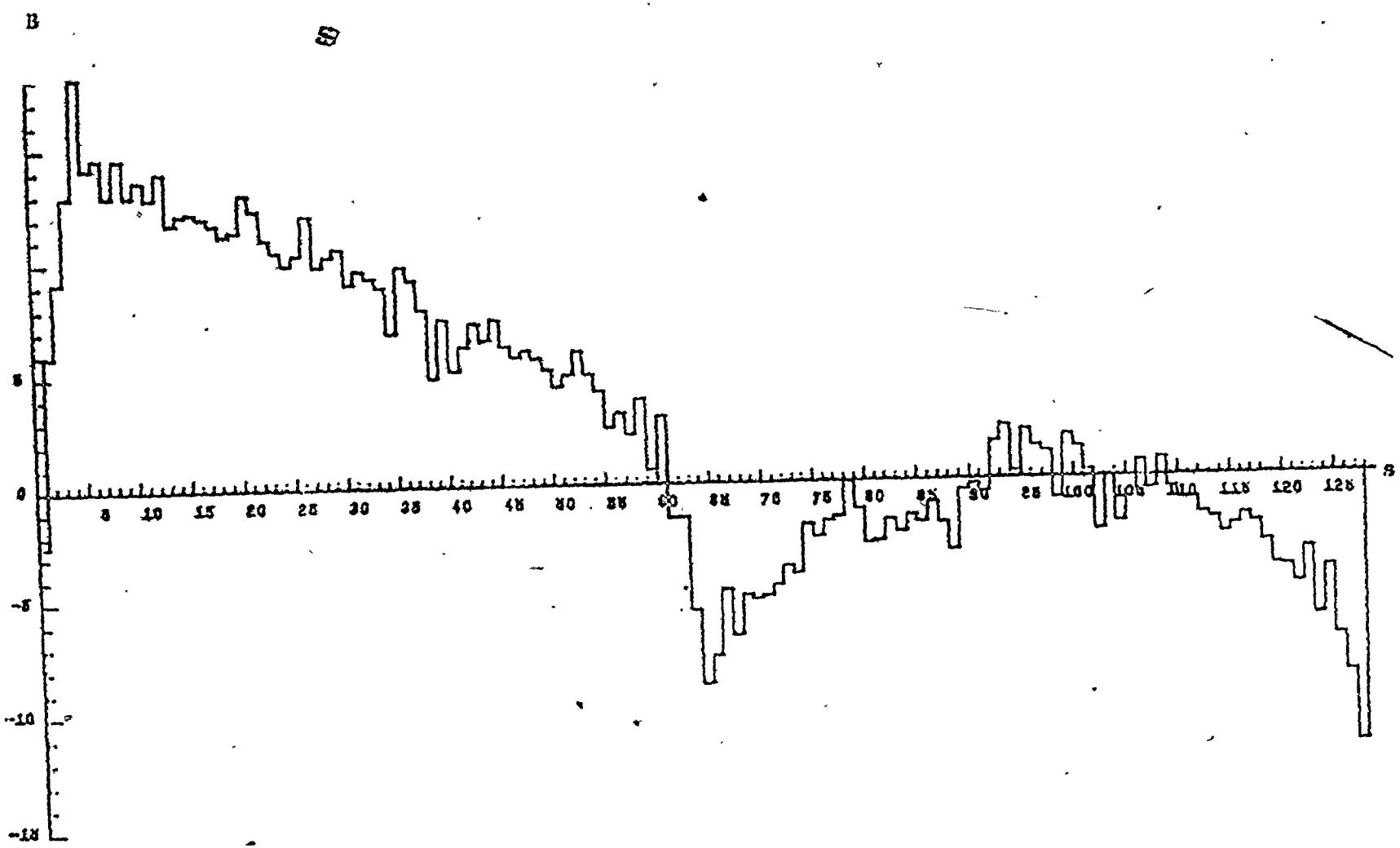


FIGURE 40b :Long term average sequency spectral envelope employed in the TISPS Walsh vocoder.

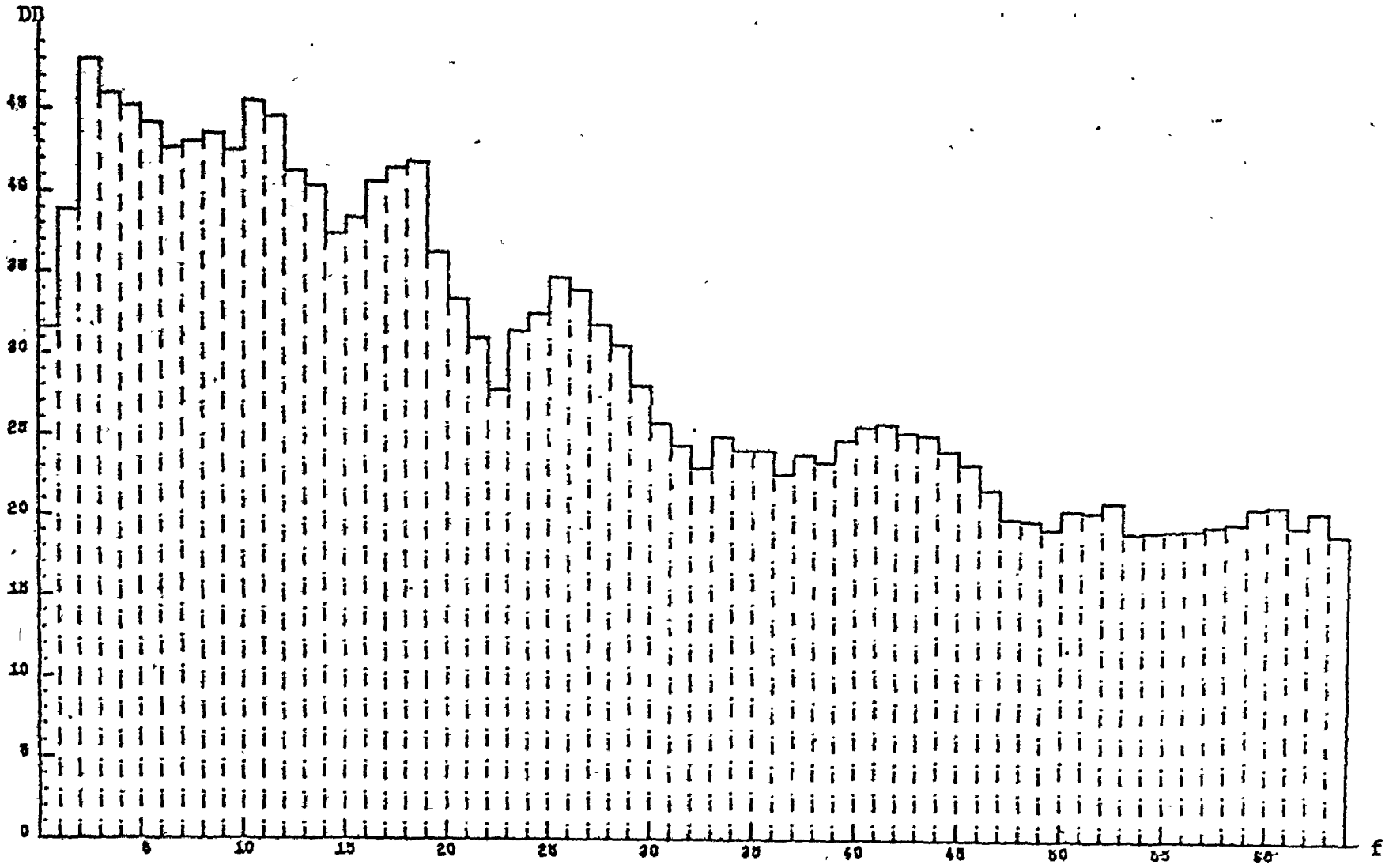


FIGURE 41a :Long term average frequency spectral envelope for 16ms speech segments sampled at 8KHz.

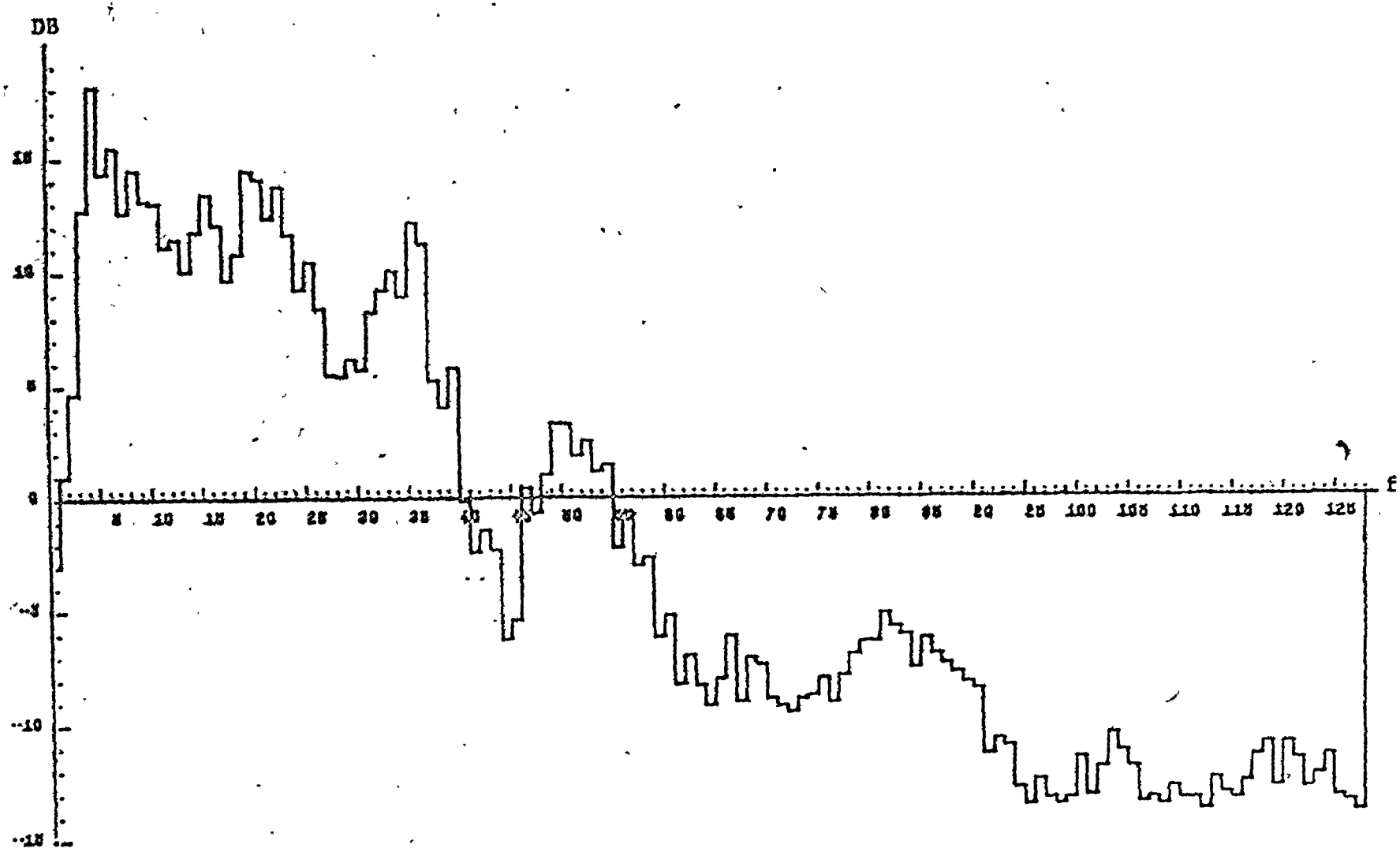


FIGURE 41b :Long term average frequency spectral envelope employed in the TISPS Fourier vocoder.

5.5 RESULTS OF THE SIMULATIONS

The vocoders simulated required that the phase information contained in the speech signal be redundant or unnecessary because only magnitude spectra could be considered in the synthesis of the speech signal. All phase information is discarded during the calculation of the magnitude spectrum in the vocoder synthesizers and no attempt is made in the analyzers to replace this phase information. To reduce the discontinuities at the window edges, every second cosine term in the Fourier spectrum and every second cal term in the Walsh spectrum were given a negative value and all of the sine terms in the Fourier spectrum and all of the sal terms in the Walsh spectrum were preset to 0 magnitude in order to reduce the complexity of the vocoder systems. If the phase information is truly unnecessary for speech intelligibility ([23]), then the spectra so constructed would be the simplest to implement.

In order to establish whether or not the phase information was necessary to retain the intelligibility of the original speech signals, the phase information was removed from the original speech signals and a new speech signal was derived from the Walsh and Fourier spectra.

In the first case, for a discrete Fourier magnitude spectrum $F_s^2(n)$, $n=0,1...,N/2-1$, the Fourier coefficient spectrum $F_c(j)$, $j=0,1,...,N-1$ is reconstructed as follows:

$$F_c(2n) = F_s(n)\cos(n\Pi) \qquad \text{(cosine terms of frequency } n/NT\text{)}$$
$$n=0,1...,N/2-1 \qquad (5.1)$$
$$F_c(2n+1)=0 \qquad \text{(sine terms of frequency } n/NT\text{)}$$

where N=128 for the 16ms windows utilized with $1/T=8KHz$ sampling rates. This is the same method of coefficient spectrum reconstruction utilized in the FFT vocoder simulation. (The sentences used appear in Appendix A).

The speech signals derived from the inverse transform of $F_c(j)$ had lost a great deal of the quality contained in the original speech signals and the intelligibility was somewhat reduced. The pitch variations in the reconstructed signal were suppressed. This was evident from the resultant monotone quality of the speech signals. Voiced sounds seemed more distorted than unvoiced sounds. Intelligibility was reduced by an echoing effect during vowel sounds and a slight suppression of the fricative sounds. Plosive sounds seemed unaffected.

In order to establish whether this particular method of Fourier coefficient spectrum reconstruction was an isolated case, another method was employed. The coefficient spectrum

$F_c(j), j=0,1...,N-1$, was reconstructed from the magnitude spectrum $F_s^2(n), n=0,1...,N/2-1$, as follows;

$$F_c(2n)=\frac{F_s(n)}{\sqrt{2}}\cos(n\Pi) \qquad \text{(cosine terms of frequency } n/NT)$$

$$n=0,1,...N/2-1 \qquad (5.2)$$

$$F_c(2n+1)=\frac{F_s(n)}{\sqrt{2}} \qquad \text{(sine terms of frequency } n/NT)$$

The speech signal derived from the inverse transform of F_c was indistinguishable from the speech signal derived from the coefficient spectrum given in (5.1).

The effect of phase removal in the Walsh domain (I will define phase in the Walsh domain to have the same definition as phase in the Fourier domain ie. $\phi(n)=\arctan[W_c(2n)/W_c(2n+1)]$, $n=1,2,...N/2-1$ where W_c is the Walsh coefficient spectrum and ϕ is the phase spectrum to compare with $\theta(n)=\arctan[F_c(2n+1)/F_c(2n)]$ where F_c is the Fourier coefficient spectrum and θ is the phase spectrum) was accomplished in a similar manner to the phase removal procedure followed in the Fourier cases. The Walsh coefficient spectrum W_c was derived from the Walsh magnitude spectrum W_s^2 (which was obtained by adding the squared Walsh coefficients of the same sequency) as follows:

$$W_c(2n)=W_s(n)\,sal(N/2,n) \qquad \text{(cal terms of sequency } n/NT\text{)}$$
$$n=0,1,\ldots N/2-1 \qquad (5.3)$$
$$W_c(2n+1)=0 \qquad \text{(sal terms of sequency } n/NT\text{)}$$

The speech signals derived from the Walsh transform of W_c had similar qualities to the speech signals derived from the reconstructed Fourier coefficient spectra. The intelligibility seemed to be affected in an almost identical manner, but the quality did not seem to be degraded as much as in the Fourier case. The quality seemed to improve when the lowpass filter used in desampling was set to 2HKz from 3HKz in the Walsh case but no such improvement was apparent in the Fourier case.

The above results seemed to indicate that some quality and intelligibility of the speech waveforms is lost when the phase information is removed from those waveforms. However the method of reconstruction seemed to have no effect. Hence the FFT and Walsh vocoder simulations utilized the reconstructions given in (5.1) and (5.3) respectively.

5.5.1 RESULTS OF THE FFT VOCODER SIMULATION

Four sentences were processed by the FFT vocoder simulation (see Appendix A). The synthesized speech signals were intelligible and appeared to retain some of the quality

of the original speech signals. The pitch variations that were lost in the speech signals which were reconstructed after the phase information was discarded, seemed to be intact even though the pitch seemed higher than that of the original speech signals. There was a slight background noise that seemed to be the result of a reverberation effect during voiced sounds. Fricative and plosive sounds did not appear degraded. Because of the higher pitch present, speaker recognition was not possible but speaker differentiation was possible. The synthesized speech did not have a monotone machine-like quality.

5.5.2 RESULTS OF THE WALSH VOCODER SIMULATION

The sentences processed by the FFT vocoder simulation were also processed by the Walsh vocoder simulation. The resultant synthesized speech was very similar to the FFT vocoder simulation output. The pitch did not seem to be as artificially high in the Walsh vocoder simulation and the background noise during voiced sounds seemed stronger. However, the same natural quality of the speech signals was retained and the intelligibility was only slightly degraded by the background noise. It would be difficult to distinguish the synthesized speech signal from the FFT and Walsh vocoder simulations.

Again, in the Walsh case, the speech quality seemed to improve when the desampling lowpass filter was set to 2KHz from 3KHz

5.5.3 RESULTS OF THE T1SPS VOCODER SIMULATIONS

The four sentences processed by the FFT and Walsh vocoder simulations were processed by the Fourier and Walsh T1SPS vocoder simulations discussed in 5.4. In both T1SPS cases the resultant speech was unintelligible and of very poor quality. All that remained of the original speech structure was the rythm of the sentence. It appeared to resemble the output of a kazoo or some other one note low frequency wind instrument if the "musician" was trying to make the instrument appear to speak. There was no pitch variation present and the output was very "bassy" ie. no high frequency tones were perceptible.

5.6 CONCLUSIONS

The FWT can replace the FFT in vocoder applications investigated without any perceptible degradation of the synthesized speech output. The replacement of the FFT with the FWT can realize simpler hardware logic which can operate in real time for speech signals. The bank of filters used in the analogue channel vocoder is thereby replaced with

digital hardware which is compact, less expensive, and more reliable. All of the disadvantages of the analogue channel vocoder, pitch detection excepted, are solved with the use of the FWT in the Walsh vocoder configuration discussed.

The reverberant quality of the synthesized speech signals of the Walsh and FFT vocoders during voiced sounds seems to be due to phase distortions that correspond to delay distortions [24] and are perceived as an "electrical accent" in the vocoder output. The phase spectrum of a voiced excitation signal is pulselike [24] so perhaps in the reconstruction of these signals a less arbitrary phase spectrum can be utilized to lessen the effect of phase distortions.

The cepstrum pitch detector used in the simulations appears to yield some pitch values that are higher than the true pitch. This pitch error is slightly more noticeable in the FFT vocoder simulation than in the Walsh vocoder simulation.

The quality of the vocoder outputs could also be upgraded with the use of longer time windows to attain a spectral resolution that is less coarse. Overlapping windows could be utilized to attain longer sample windows without missing spectral changes. For example, 32ms windows overlapping the previous window by 16ms could be used to double the spectral resolution.

Although it has been stated that the Walsh vocoder can operate at 2400 bps, the total encoding procedure was not simulated. It was proven that 14 channel signals can represent the spectral envelope and the pitch parameters can be encoded with 6 bit words every 8ms. Further studies should be made to investigate word length requirements for arithmetical operations given a 64Kbps PCM speech signal.

The T1SPS vocoders as simulated are not practical devices. Perhaps their performance can be improved with a spectral envelope that is updated at a frequent rate. The spectral envelope might be represented by filter coefficients or by adaptive predictive coding techniques. The long term spectral envelopes utilized do not contain enough information to yield acceptible speech output or the T1SPS itself does not contain the required information. Since the basic rythm of the sentence is retained, it would seem that some information is contained in the T1SPS but it is not known at this time what additional information is required.

Generally, more precise and definitive speech quality and intelligibility measurements should be made on the vocoder speech outputs to more accurately evaluate the vocoder systems. It would seem from the subjective tests performed that the

Walsh vocoder is a promising vocoder system in that the FFT vocoder has been previously effected with acceptible results and the Walsh vocoder output was indistinguishible from the FFT vocoder output when both vocoder configurations were identical except for the FWT replacing the FFT.

APPENDIX A

Four sentences used to evaluate the simulation performance:

1) "Speed and efficiency were stressed" (2 seconds)

2) "Black pup was a throwback" (2 seconds)

3) "Will the rest follow soon?" (2 seconds)

4) "Nature, as we often say, makes nothing in vain, and man is the only animal whom she has endowed with the gift of speech; and it is a characteristic of man that he alone has any sense of good and evil." (15 seconds)

[7] Birch, J. N. and Getzin, N., "Multipath /RFI/ Modulation Study for DRSS-RFI Problem: Voice Coding and Intelligibility Testing for a Satellite Based Air Traffic Control System, Magnavox Company Technical Report No. ASAO-PR20020-2, April 1971.

[8] Bially, T., and Anderson W. M., "A Digital Channel Vocoder". IEEE Trans. Communication Technology Vol. Com-18, #4, August 1970 pp 435-442.

[9] Noll, A. M. "Cepstrum Pitch Determination", Journal Acoustical Society of America, Vol. 41, #2, 1967, pp 293-309.

[10] Frei, A. H. et al, "Adaptive Predictive Speech Coding Based on Pitch-Controlled Interruption/ Reiteration Techniques", IEEE International Conference on Communications Vol 2 June 1973.

[11] Walsh, J. L., "A closed set of normal orthogonal functions". American Journal of Mathematics, Vol. 45 1923 pp 5-24.

[12] Paley, R. E., "A remarkable series of orthogonal functions", Proc. London Math. Society, Vol. 34 1932, pp 241-279.

[13] Fine, N. J., "On the Walsh functions", Trans. American Math. Society, Vol. 64, 1949, pp 372-414

[14] Fine, N. J., "The generalized Walsh functions". Trans. American Math. Society. Vol. 69, 1950, pp. 66-77.

[15] Pichler, F., "Walsh functions and linear system theroy". Proc. Applications of Walsh Functions April, 1970, pp 175-182.

[16] Pichler, F., "On state-space description of linear dyadic invariant systems", Proc. Applications of Walsh Functions, April 1971, pp 166-170.

[17] Harmuth, H. F., Transmission of Information by Orthogonal Functions, Springer-Verlay, Berlin/ New York, 1969.

[18] Blachman, N. M., "Spectral Analysis with Sinusoids and Walsh Functions". IEEE Trans. Aerospace and Electronic Systems, Vol. AES-1, #5, September 1971, pp 900-905.

[19] Shum, F., Fast Transformations with Walsh - Hadamard Functions, Master of Engineering Thesis, McMaster University, Hamilton, Ont. Canada, September 1972.

[20] Ahmed, N., et al, "Bifore or Hadamard Transform", IEEE Trans. Audio and Electroacoustics, Vol. AU-19 #3, September 1971, pp 225-234.

[21] Campanella, S. J., and Robinson, G. J., "Digital Sequency Decomposition of Voice Signals", Proc. 1970 Applications of Walsh Functions, April, 1970 pp 230-237.

[22] Kitai, R., and Siemens, K. H., "A Non-Recursive Equation for the Fourier Transform of a Walsh Function". IEEE Trans. Vol. EMC-15, 1973, pp 81-83.

[23] Shroeder, M. R. "Vocoders: Analysis and Synthesis of Speech", Proceedings IEEE, Vol. 54, No. 5, May 1966, pp 720-734.

[24] Gold, B., and Rader, C.M., "The Channel Vocoder", IEEE Transactions on Audio and Electroacoustics, Volume AU-15, No. 4, December 1967.