SPEECH ANALYSIS WITH BESSEL FUNCTIONS

By

ROBERT GAIL HAMILTON

Bachelor of Science in

Electrical Engineering

Oklahoma State University

Stillwater, Oklahoma

1985

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 1987

SPEECH ANALYSIS WITH BESSEL FUNCTIONS

Thesis Approved:

Thesis Advisor

James E Baker

Keith A Jerome

Norman N. Durham

Dean of the Graduate College

PREFACE

A signal may be expressed as a linear combination of other functions, called the basis set. This is essentially a model of the waveform or signal. Infinitely many choices are possible for the basis set, with the most common choice being the set of trigonometric functions. This study involves the use of Bessel functions of the first kind as the basis set.

The original goal of the research was to develop an automatic speaker recognition scheme, based upon the Fourier Bessel series. But the difficulty of collecting a high quality data base and certain hardware deficiencies precluded the completion of the original goal. Also, it was found that the theoretical foundation for the use of the Fourier-Bessel series for signal analysis was practically nonexistent. For these reasons, the study was confined to general purpose speech analysis and to investigation of the computational algorithms required.

The author wishes to express his appreciation to his major advisor, Dr. Rao Yarlagadda, for his guidance and his willingness to consider nontraditional approaches to digital signal processing problems. The freedom to explore new ideas is of the utmost importance, and the research environment provided by Dr. Yarlagadda and the rest of the Electrical

Engineering faculty provided such freedom.

Special thanks are given to my wife, Janice Hamilton, and to my son, Kevin, for their forgiveness of my long hours of work and their faithful support.

The funding for this project was generously provided by the National Security Agency, under contract MDA-904-85-15-0002.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Motivation

It has long been known that arbitrary signals may be expressed as linear combinations of orthogonal functions. For example, a very common technique used in applied mathematics and electrical engineering is the expression of a signal in terms of its Fourier series coefficients. Such representations are very useful in the field of digital speech signal analysis; for example, speech signals can be compactly described by some subset of series coefficients which are, in turn, used as the features in a pattern classification scheme for speech or speaker recognition. The most commonly used series representation in the field of signal analysis is the Fourier series. But other series can also represent signals, although these alternative series are rarely as easy to derive or interpret as the Fourier series.

There has been relatively little prior research into the use of alternative basis sets for speech signal analysis. This is because the Fourier series is a very useful and convenient tool, for which there is a mathematically tractable basis set. Furthermore, the Fast Fourier Transform

algorithm (FFT) has provided impetus to the Fourier series as one of the main implements of digital signal analysis. Despite the usefulness of the Fourier series, it cannot be shown that it is optimal as a means of pattern recognition, nor is it always the most compact representation for speech coding and transmission. Therefore, investigation of alternative basis sets can be justified in the search for better basis sets for coding and classification.

In this particular study, the basis functions of interest are the Bessel functions of the first kind. This set of functions was selected for several reasons. First, they are the solutions of a family of differential equations with time-varying coefficients, thus possibly making them a better choice for the modeling of nonstationary speech signals. Second, the Bessel functions are often used in mathematical physics in situations where cylindrical boundary conditions prevail, such as vibrating drumheads and circular waveguides. The vocal tract is often modeled as a concatenation of cylindrical acoustic tubes, and thus has circular boundary conditions. Third, there are fast computer algorithms for the calculation of the Fourier-Bessel coefficients, making computation with a general purpose computer feasible. Finally, since there has been little prior investigation of the Fourier-Bessel series in speech analysis there was some value in the pure research itself.

The end application of this investigation was originally intended to be the use of Fourier-Bessel expansions in

the automatic speaker recognition problem. But during the course of the research, it was found that the theoretical basis for the application of Fourier-Bessel expansions did not seem to be well known in the literature. Therefore, the research involved theoretical and computational considerations as well as applied speech analysis. The original objective of speaker recognition was largely supplanted by the theoretical and computational investigations, due to difficulties encountered in the collection of a high quality data base for the speaker recognition problem.

Overview

Chapter II of this report gives a review of generalized basis sets and their use in speech signal analysis. The non-traditional basis sets thus exposed include the exponentially damped sinusoids and the Walsh functions. Applications include speech and speaker recognition as well as speech coding or storage. Chapter III presents the properties of Bessel functions which are useful for signal analysis, such as the asymptotic approximation. Building upon the foundation laid in Chapter III, Chapter IV describes methods used to obtain the Fourier-Bessel series coefficients, including the use of the Fast Hartley Transform. Chapter IV also includes a summary of the important properties of the Fourier Bessel series. Chapter V gives an interpretation of the Fourier-Bessel coefficients from a linear modeling point of view, as well as an approximate relationship between the

traditional frequency domain and the set of Fourier-Bessel coefficients. Thus, Chapter V serves as a bridge from the theoretical and computational details to the practical applications described in Chapter VI.

Chapter VI describes the application of the Fourier-Bessel series expansion to actual speech signals. The very practical problem of choosing the starting point of an analysis frame is addressed. Two different pitch detection methods are described, one based on the Fourier-Bessel coefficients themselves, and another based upon a nontraditional cepstrum. This nontraditional cepstrum is actually the Fourier-Bessel expansion of the log magnitude spectrum, and appears to be useful for pattern classification and coding. Finally, in Chapter VII, the conclusions reached in this research are summarized, and future areas of research are indicated.

Two appendices are included in this thesis. Appendix A includes the Fortran source code for the Fast Hartley Transform which is used in the calculation of the Fourier-Bessel series. Appendix B shows a theoretically pleasing relationship between the complex cepstrum and the Bessel functions: It is shown that representation of a sequence in terms of its complex cepstrum is equivalent to representation of the same sequence as a convolution of Bessel functions.

CHAPTER II

SERIES REPRESENTATIONS AND BASIS SETS

Introduction

The purpose of this Chapter is to review the properties of series expansions in general. This material serves as background for the later Chapters of this thesis, and is not meant to be an exhaustive survey. Unfortunately, the very important subject of fast algorithms can only be mentioned briefly; In fact, entire books have been written on the subject (Ahmed and Rao, 1975). This Chapter will deal with some of the aspects of series representations and basis sets which are often overlooked or forgotten by signal processing practitioners. For example, there seems to be a misconception that orthogonality is a requirement for a basis set. Applications of alternative basis sets to speech processing have been relatively rare in the past; So a synopsis of some of the applicable literature is provided by this Chapter. Of particular interest is the motivation behind the use of non-traditional basis sets for speech analysis. Due to the non-stationarity of speech signals, it does not seem feasible to derive an optimal transformation that is best for all speakers and all phonemes.

Overview of Series Representations: The Continuous-Time Case

The usual series representation can be considered as a model of the waveform or signal which has the assumed form of a summation of a set of linearly independent functions:

$$f(t) = \sum_i C_i g_i(t). \qquad (2.1)$$

The set of functions $\{g_i(t)\}$ is called the basis set. The model parameters are the series coefficients, or C_i's. The limits of the summation were left indeterminate, because different limits are appropriate for various basis sets. Of course, any such data model could just be assumed, but an important problem is that of completeness of the basis. It must first be determined if a linear combination of the assumed basis functions can indeed represent f(t).

Completeness of Basis Sets

Completeness of a basis set simply means that some linear combination of the basis functions can be found which converges to the desired f(t). The class of functions which can be represented by a series may be limited in some way. For example, it may be desired to represent only bandlimited functions. Another common restriction is that the data model of Equation (2.1) is valid only over a finite range of the independent variable t. But given some suitably restricted f(t), and some set $\{g_i(t)\}$, how can the completeness of the basis set be determined?

Determination of completeness can sometimes be a very difficult problem. But sometimes it is easy to verify that a basis set is <u>not</u> complete. For example, suppose that every element of the basis set is a bandlimited function. Then it would not be possible to form a linear combination of the basis set which has any energy whatsoever in the band above the highest frequency in the basis set. If modeling of a signal with higher frequencies is the goal, then another basis set must be chosen for the data model. But if only suitably bandlimited functions are to be modeled, then the basis set can actually be complete for all practical purposes. By taking the Fourier transform (or perhaps the Laplace transform) of both sides of Equation (2.1), it can be seen that the modeling of f(t) as a linear combination of the $g_i(t)$'s is equivalent to modeling the transform of f(t) as a linear combination of the transforms of the basis functions. If completeness of a basis is hard to prove or disprove in one domain, then it may be possible to easily prove or disprove the completeness hypothesis in the other domain.

An example may be in order here. Suppose that the function f(t) is to be represented on the interval (-a,a) by a linear combination of cosine functions, specified as $g_i(t) = \cos(i\pi t)$. Suppose the function f(t) is analytic in the interval (-a,a), so that the Taylor series exists. But notice that the Taylor series for the functions $\{\cos(i\pi t)\}$ only have even terms (this is essentially the "transform" of

the cosine functions to another representation). Then no linear combination of cosine functions can possibly have any odd terms in the resulting Taylor series. So if f(t) has odd components as well as even components, the assumed cosine series cannot possibly represent it on the interval (-a,a).

Orthogonality

Orthogonality of two functions can be defined in many different ways, depending on the inner product which is defined. For continuous time functions (as opposed to sampled data) it is common to use the inner product defined as

$$\langle g(t),h(t)\rangle = \int_a^b g(t)h^*(t)dt. \qquad (2.2)$$

The interval [a,b] may be finite or infinite, depending on the application. If two functions g(t) and h(t) are such that

$$\langle g(t),h(t)\rangle = 0 \qquad (2.3)$$

then the functions are said to be orthogonal. The definition of orthogonality can be generalized to include the concept of a weighting function. Let the inner product be defined as

$$\langle g(t),h(t)\rangle = \int_a^b g(t)h^*(t)w(t)dt \qquad (2.4)$$

where w(t) is called a weighting function. This concept is important because the Bessel functions are orthogonal with respect to a weighting function, as will be explained in Chapter III. Equation (2.3) is a special case of Equation

(2.4), where the weighting function is unity.

Calculation of the Series Coefficients

Even when a basis set can be shown to be complete, there remains the problem of calculation of the series coefficients (the C_i's of Equation (2.1)). If the basis set's elements are orthogonal functions (with respect to some weighting function), then the model parameters can be calculated by the following procedure. First, assume a data model of the form shown by (2.1). Then multiply both sides of the equation by the weighting function and the complex conjugate of an arbitrary member of basis set, $g_k(t)$:

$$w(t)g_k^*(t)f(t) = w(t)g_k^*(t) \sum_i C_i g_i(t). \qquad (2.5)$$

Using the fact that the summation is taken over the index i rather than k, move both $g_k^*(t)$ and $w(t)$ inside the summation:

$$w(t)g_k^*(t)f(t) = \sum_i w(t)C_i g_k^*(t)g_i(t). \qquad (2.6)$$

Now integrate both sides of Equation (2.6) over the interval of orthogonality. Depending on the basis set of interest, this interval may be infinite or finite. Therefore, the limits of the integral are stated arbitrarily:

$$\int_a^b f(t)g_k^*(t)w(t)dt = \int_a^b \left[\sum_i C_i g_k^*(t)g_i(t)w(t)\right]dt. \qquad (2.7)$$

Interchanging the order of integration and summation gives

$$\int_a^b f(t)g_k^*(t)w(t)dt = \sum_i C_i \left[\int_a^b g_k^*(t)g_i(t)w(t)dt \right] . \qquad (2.8)$$

If the set of functions $\{g_n(t)\}$ is orthogonal with respect to the weighting function w(t) then the integral inside the brackets vanishes when the subscript i (which varies over the summation) is not equal to the subscript k (which is arbitrary but fixed). Thus, only one term of the series remains. The lone term is shown in Equation (2.9):

$$\int_a^b f(t)g_k^*(t)w(t)dt = C_k \int_a^b g_k(t)g_k^*(t)dt. \qquad (2.9)$$

Isolating the C_k on one side of the equation yields the closed-form formula for any arbitrary series coefficient:

$$C_k = \frac{\int_a^b f(t)g_k^*(t)w(t)dt}{\int_a^b g_k(t)g_k^*(t)w(t)dt}. \qquad (2.10)$$

Unfortunately, this formula is sometimes very difficult to evaluate in closed form, except for a few special cases. As a theoretical tool, Equation (2.10) has much merit. But it cannot deal with the more practical case of sampled data. For example, with sampled data the integrals can only be approximated by finite summations. The sampled-data case must be treated separately, but bear in mind that it may be just as well considered as numerical quadrature. A discussion of the important sampled-data case will now follow.

Overview of Series Representations:
The Sampled-Data Case

There are basically two possibilities for sampled data sequences: (1) The sampled data consists of an infinite sequence of numbers; or (2) The sampled data consists of a finite sequence, or vector, of numbers. The latter of these possibilities is the most important in the practice of digital signal processing. After all, it is not really possible to sample data over an infinite time interval. The model of the data is as a linear combination of some vectors, as shown:

$$\underline{x} = \sum_{k=0}^{K-1} y_k \, \underline{a}_k . \qquad (2.11)$$

The y_k's are the model parameters which describe the vector (or truncated sequence) $\underline{x}$, and the set of vectors $\{\underline{a}_k\}$ is the assumed basis set for the data model. The number of vectors to be summed is K, and they have been numbered arbitrarily from 0 to K-1. The dimension of $\underline{x}$ (the vector's length) certainly does not have to be equal to K. One can simply assume, for instance, that a vector of dimension 1000 can be modeled well enough for a given purpose with a linear combination of only 100 vectors in a basis set. In that case, the original set of 1000 numbers has been described by a set of 100 numbers: Some amount of data reduction or compression has occured (but at a price: the original vector $\underline{x}$ may not be uniquely recoverable from the model parameters).

Completeness of Basis Vector Sets

A basis set of vectors can be complete only if the set has at least as many elements as the dimensionality of the vectors, N, and at least N of the vectors in the basis set are linearly independent. Testing for independence of a set of vectors is usually easier than testing the completeness of a set of continuous-time functions. Any good linear algebra text contains details of such tests, and they will not be repeated here.

Orthogonality of Basis Vectors and Computation of Model Parameters

The orthogonality of basis vectors is not a requirement for completeness. But orthogonality is a very useful computational convenience. Note that Equation (2.11) could also have been written as a matrix equation:

$$\underline{x} = A\underline{y}. \tag{2.12}$$

The columns of matrix A are the assumed basis set. If the columns of A are orthogonal to one another, then the model parameters, $\underline{y}$, are very easy to compute. First, premultiply each side of (2.12) by the complex conjugate transpose of A, yielding

$$A^*\underline{x} = A^*A\underline{y}. \tag{2.13}$$

Since the columns of A are orthogonal the matrix A^*A becomes diagonal, and hence trivially simple to invert. Now premultiply each side of (2.13) by the inverse matrix

$(A^*A)^{-1}$. The result is

$$(A^*A)^{-1}A^*\underline{x} = \underline{y}. \qquad (2.14)$$

Throughout this last discussion, it was assumed that the number of model parameters did not exceed the number of data points in $\underline{x}$. Then it can be shown that Equation (2.14) yields the $\underline{y}$ which makes $\|\underline{x}-A\underline{y}\|_2$ minimum (Menke, 1984). Thus, the property of orthogonality makes computation of the series coefficients much easier: No explicit matrix inversion needs to be performed.

Alternative Basis Expansions in Speech Processing

Several alternative basis sets have been previously used for speech processing, including Walsh functions and exponentially damped sinusoids. The purpose of this section is to catalogue some of these applications and to try to provide some insight into the reasons those alternatives were used.

Exponentially Damped Sinusoids

The set of exponentially damped sinusoids has been used several times in the past for the modeling of speech signals. Dolansky (1960) used this basis set for the continuous time modeling of speech waveforms. His choice was based on both heuristic and practical arguments. First, he reasoned that speech waveforms often look something like damped sinusoids when viewed on an oscilloscope. Second, the vacuum

tube based hardware technology of the day lent itself to the implementation of active filters with their pole locations in the left half s plane. His implementation consisted of an analog filter bank which was cleverly contrived so that the basis functions were orthogonal.

Orthogonalized damped sinusoids were also used by Manley (1963) in an early digital signal processing attempt. His goal was to build an analysis-synthesis system for the speech waveforms. The expansion into the basis functions was done by trapezoidal integration. Since this work was before the appearance of the Fast Fourier Transform on the signal processing scene, orthogonality of the basis set was almost a requirement for efficient implementation. Manley reported that the resynthesized speech was of fair quality when 16 fixed oscillation frequencies (16 pole pairs) were used. Once again, the choice of the damped sinusoids as a basis was primarily because of the supposed similarity between speech waveforms and damped sinusoids.

Another analysis-synthesis scheme was devised by L.A. O'Neill (1969). Once again, the main idea was to use a filter bank analysis-synthesis approach. His filters were based on the damped sinusoids and were also orthogonalized. This system was essentially a vocoder. O'Neill reported high-quality speech from the resynthesis scheme.

Word recognition can also be achieved using basis sets other than the traditional sine and cosine basis. An optimum basis set was derived for use in short-time analysis of a

small vocabulary (Clark, 1970). By examining the power spectra of the words he was concerned with, Clark found filter pole frequencies that corresponded to the peaks of the spectra. By expanding the speech waveform into the series, the series coefficients could then be used as the features in a pattern classification scheme. The motivation for the use of an "optimal" basis set was compactness of representation: It is usually desired to have a small set of numbers for a pattern classifier's input.

Except for Clark's approach, the choice of the basis sets was quite subjective. But at least these researchers made some attempt to justify their choice of basis, and did not slavishly choose the traditional trigonometric basis sets out of habit alone. Their motivations for the choice of damped sinusoids were reflected in the choice of Bessel functions for speech analysis in this thesis: The Bessel functions also resemble typical speech waveforms.

Walsh Functions

The Walsh-Hadamard transform has been much more widely used for image processing than for speech processing. But some researchers have seen potential advantages in the use of the Walsh-Hadamard transform for speech applications (Shum, Elliot, and Brown, 1973). The Walsh-Hadamard basis set consists of rectangular waveforms, and therefore does not resemble speech waveforms very much at all. But the advantage of this transform is the extreme simplicity of the

fast transform algorithm, which uses only additions and subtractions.

The Walsh functions were used by Shum, Elliot, and Brown in a speech coding and compression scheme. The general plan was to transform a segment of speech, and then represent it with a few dominant coefficients in the transform domain. Their segment size was only 64 samples, and they reported fair-to-good resynthesis using only the four to eight dominant coefficients.

The main reason that the Walsh-Hadamard expansion was chosen was that a very fast, simple algorithm existed for the transformation. There was no argument that the Walsh functions were optimal for speech modeling. To some extent, this reasoning was used for the selection of the Bessel functions as a basis set for the research reported in this thesis.

Chapter Summary

The definition of a series expansion has been reviewed as a preparation for the Fourier-Bessel series which will be defined in Chapter IV. An infinitude of basis functions are possible for the data model, but only a few have been used in actual speech processing due to practical difficulties and the lack of fast transform algorithms (except for a few special basis sets). Given an arbitrary transform matrix, no general method for derivation of a fast transform (which performs the same operation as the desired matrix) appears

to be known. Therefore, most practical researchers have been forced to limit their use of series expansions to those for which fast transforms have been published. In fact, the existence of a Fast Hankel Transform algorithm was one of the main factors that led to this research.

The most common basis set for speech analysis is, of course, the traditional trigonometric basis set consisting of sine and cosine functions. But some researchers have found uses for nontraditional basis sets, and their efforts inspire this work. Applications of the Bessel functions to speech processing have come about only recently. A discussion of these applications will be deferred to Chapter VI, after the introduction to the Bessel functions and Fourier-Bessel series given by Chapters III and IV, respectively.

CHAPTER III

PROPERTIES OF BESSEL FUNCTIONS

Introduction

When the series expansion of a waveform is the subject of interest, it is always desirable to have a good knowledge of the properties of the basis set. The properties of Bessel functions are well known and a brief summary will be presented here. Further information can be found in the references (Abramowitz and Stegun, 1965, Tolstov, 1962, or Watson, 1945).

Definition of Bessel Functions

The Bessel function of the first kind, $J_p(t)$, can be defined as a solution of the differential equation

$$t^2 \frac{d^2y}{dt^2} + t \frac{dy}{dt} + (t^2 - p^2)y = 0. \qquad (3.1)$$

Note that the coefficients of the differential equation are not constant: The Bessel functions are the solutions of a family of time-varying differential equations. Therefore, as a basis set, they are not shift-invariant. It was this fact that originally led to the belief that the Bessel functions may have been a better basis set, in some sense, for non-stationary speech signals. Note also that there is a param-

eter, p, in the differential equation. This parameter defines the order of the Bessel function. In this investigation, the focus was primarily upon the Bessel function of zero order, $J_o(x)$. This was done for the sake of simplicity and to limit the scope of the research, and not because the Bessel function of order zero is particularly better for signal analysis.

Plots of Bessel functions of orders zero, one, two, and three are shown in Figures 1-4. The Bessel functions have waveforms reminiscent of voiced speech, resembling the impulse response of a second-order autoregressive model.

An alternative definition of Bessel functions is based upon the generating function

$$e^{\frac{x}{2}(t - \frac{1}{t})} = \sum_{k=-\infty}^{\infty} J_k(x)\ t^k \qquad (3.2)$$

In this expression, x is a fixed parameter while the summation is taken over all integral ordered Bessel functions. In other words, if the function on the left side of Equation (3.2) is expanded into a Laurent series about the point t=0, then the coefficients of the series are the Bessel functions of order k and argument x. The generating function definition is more useful than the differential equation definition when deriving properties of Bessel functions.

A relationship between the complex cepstrum and the generating function definition of the Bessel functions has been developed. This relationship is ancillary to the main

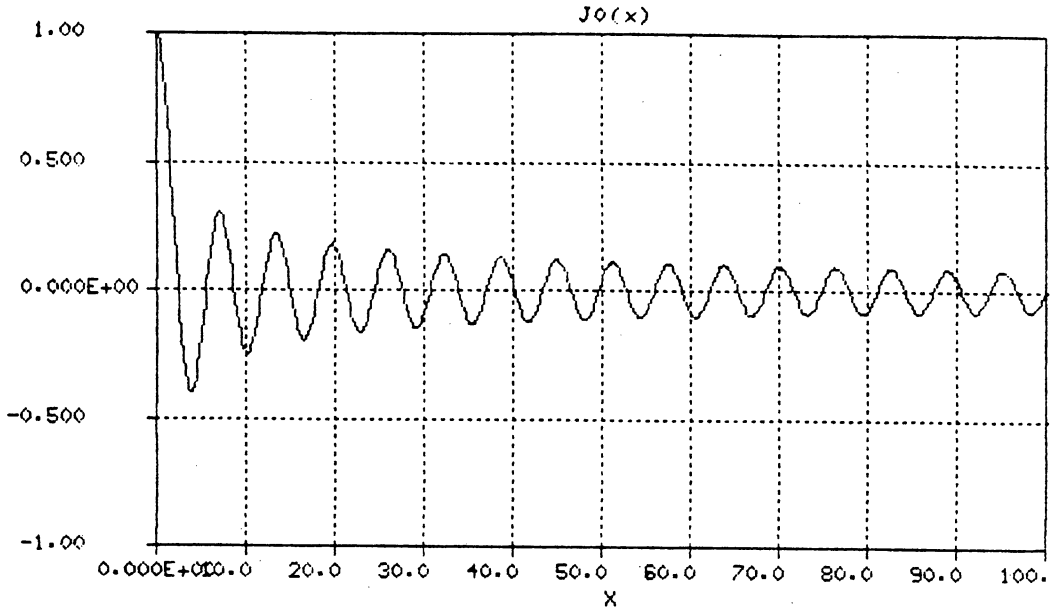


Figure 1. Bessel function $J_0(x)$.

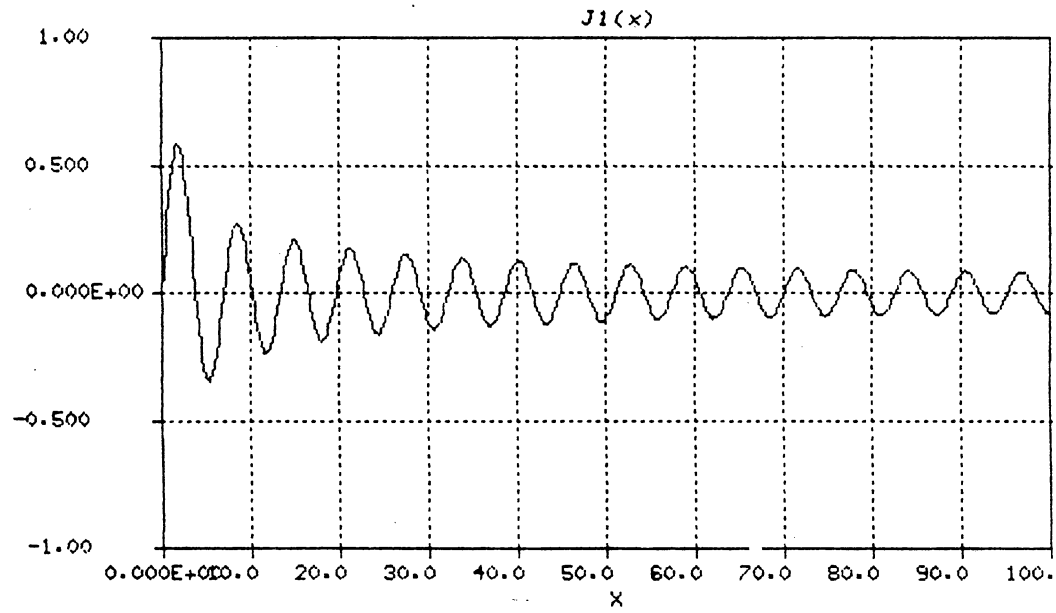


Figure 2. Bessel function $J_1(x)$.

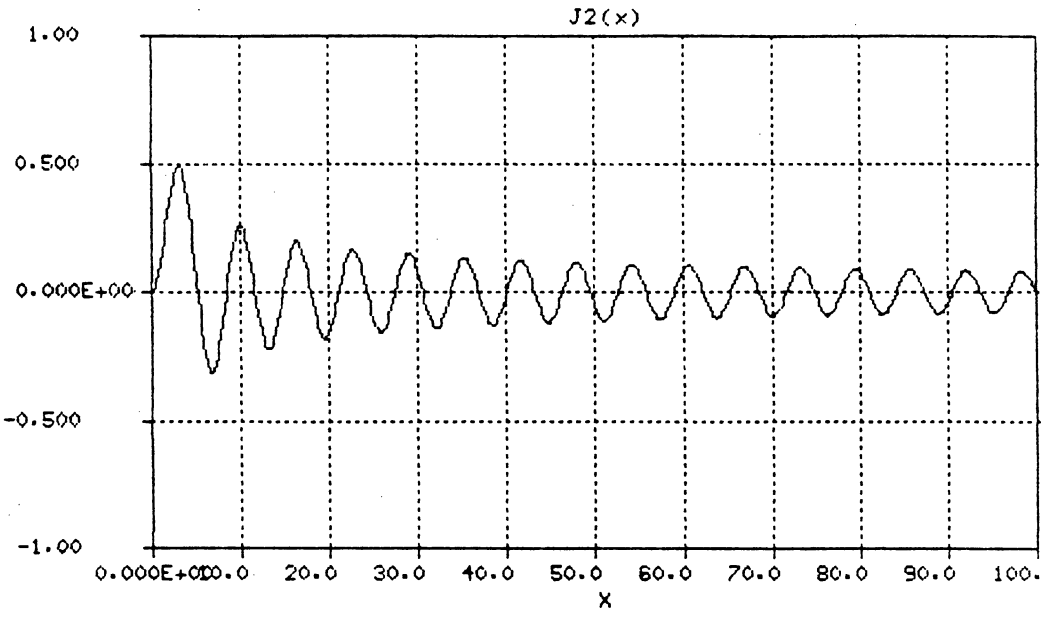


Figure 3. Bessel function $J_2(x)$.

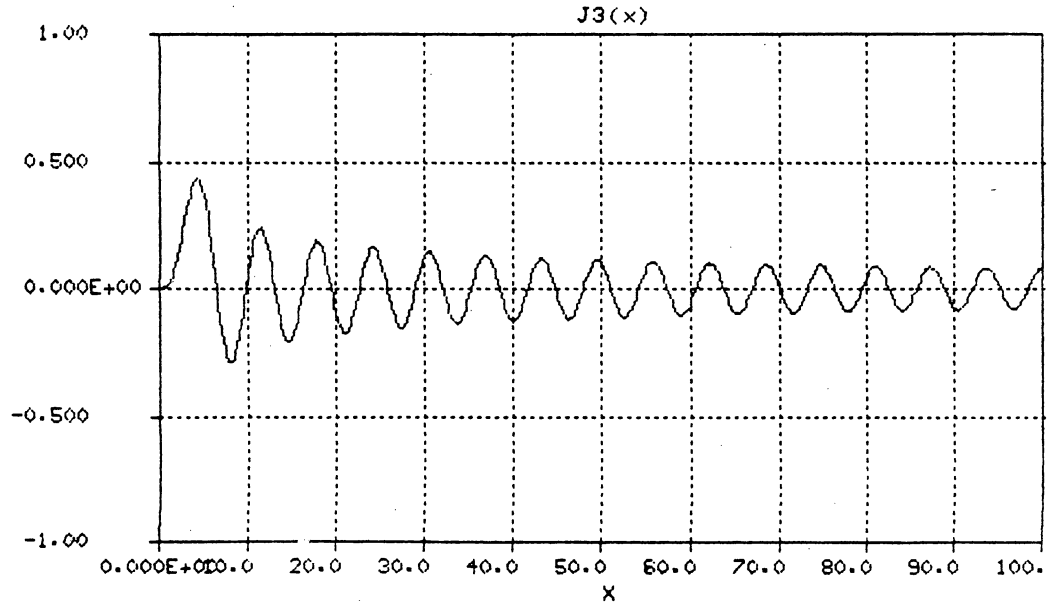


Figure 4. Bessel function $J_3(x)$.

subject of Fourier-Bessel series, and has been relegated to Appendix B of this report. However, it is believed that this relationship has not been previously published.

A family of functions known as modified Bessel functions of the first kind, $I_p(t)$, can be defined as the solutions of the differential equation

$$t^2 \frac{d^2y}{dt^2} + t \frac{dy}{dt} - (t^2 - p^2)y = 0 \qquad (3.3)$$

or may be defined by the generating function

$$e^{\frac{x}{2}(t + \frac{1}{t})} = \sum_{k=-\infty}^{\infty} I_k(x)\, t^k. \qquad (3.4)$$

Once again, the subscript indicates the order of the Bessel function, which in turn is equal to the parameter p in the differential equation definition.

Bessel functions of fractional order may also be defined, when the parameter p in Equation (3.1) is not an integer. An interesting property of the Bessel functions of half integral order is that they can always be expressed in terms of elementary functions such as the trigonometric functions and the square root (Tolstov, 1962). For example,

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\, \sin(x) \text{ and } J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\, \cos(x). \qquad (3.5)$$

This makes the Bessel functions of half integer order especially easy to manipulate.

Formulae and Properties

Many properties of Bessel functions have been discov-

ered by mathematicians over the years, but only those which pertain to the use of Bessel functions in signal analysis are presented here. Thus, the summary is by no means exhaustive. Some of the more important and useful properties are explained in the following paragraphs.

Polynomial Approximations

Polynomial approximations for the Bessel functions can be used for the generation of the Bessel functions in a digital computer. These formulae are quite lengthy and will not be repeated here, but the interested reader is referred to Abramowitz and Stegun (1965). Generally, if only integer ordered Bessel functions are required then the recursion formulae shown below can be used to generate Bessel functions of any integer order starting with Bessel functions of orders zero and one only.

Recursion Formulae

The recursion formulae for Bessel functions are

$$J_{p+1}(x) = \frac{2p}{x} J_p(x) - J_{p-1}(x) \qquad (3.6)$$

and

$$I_{p+1}(x) = \frac{-2p}{x} I_p(x) + I_{p-1}(x). \qquad (3.7)$$

These formulae can be easily obtained by differentiation (with respect to the variable t) of the appropriate generating functions. In principle, they can be used to generate Bessel functions of any order, but it has been noted that

numerical instability occurs when |p| > |x| (Olver and Sookne, 1972).

Fourier Transforms of Bessel Functions

The Fourier transform of the Bessel function $J_o(x)$ is (Abramowitz and Stegun, p. 486)

$$\int_{-\infty}^{\infty} J_o(\lambda t) e^{-j2\pi ft} dt = \frac{2}{\sqrt{\lambda^2 - (2\pi f)^2}} \quad \pi\left(\frac{f\pi}{\lambda}\right) \qquad (3.8)$$

where

$$\pi(x) = \begin{cases} 1 \text{ for } |x| < .5 \\ 0 \text{ otherwise.} \end{cases}$$

Note that the Bessel function $J_o(\lambda t)$ is strictly bandlimited and is real valued because $J_o(x)$ is an even function. Figure 5 shows the Fourier transform of $J_o(\lambda t)$. Note that the energy is concentrated near $f=\lambda/2\pi$. This is not surprising in view of the sinusoidal character of the Bessel function J_o (see Figure 1).

Laplace Transform

The Laplace transform of $J_o(\lambda t)$ is given by

$$\int_{0}^{\infty} J_o(\lambda t) e^{-st} dt = \frac{1}{\sqrt{s^2 - \lambda^2}} \qquad (3.9)$$

and the Laplace transform of the function $tJ_o(\lambda t)$ is given by

$$\int_{0}^{\infty} tJ_o(\lambda t) e^{-st} dt = \frac{s}{\sqrt{s^2 - \lambda^2}} \qquad (3.10)$$

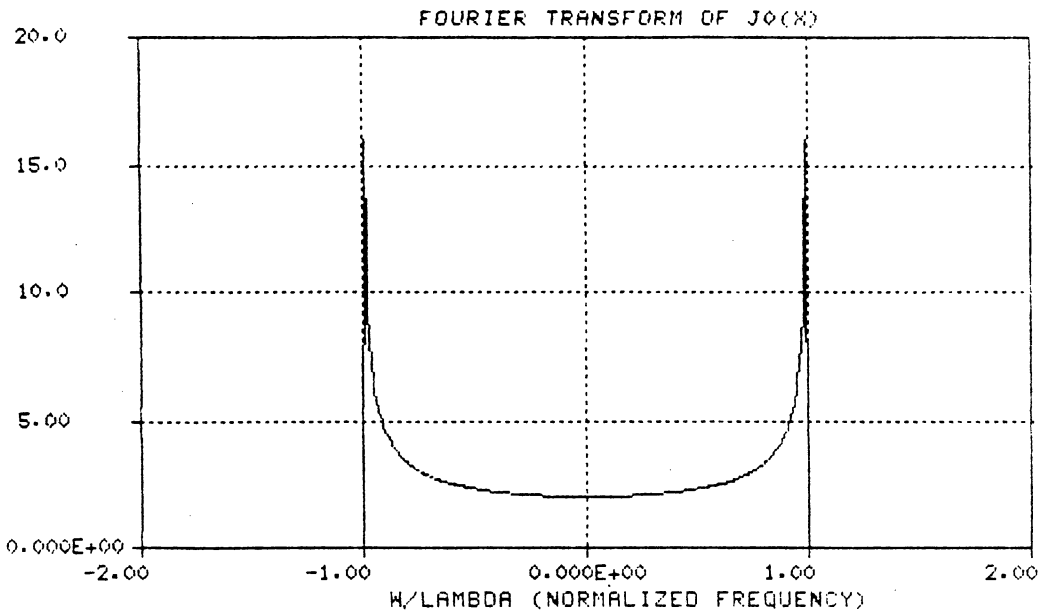


Figure 5. Fourier transform of Bessel function $J_o(\lambda t)$. Horizontal scale is normalized to $\lambda=1$.

Formulae like Equation (3.9) and Equation (3.10) can be used in conjunction with the recursion formula, Equation (3.6), to obtain Laplace transforms of higher-ordered Bessel functions. The same technique can also be applied to the Fourier transform.

Addition Formula

Unfortunately, the addition formula for Bessel functions is not as simple as the familiar trigonometric addition formulae. The formula for additive arguments is

$$J_n(x+y) = \sum_{k=-\infty}^{\infty} J_{n-k}(x)J_k(y). \qquad (3.11)$$

This formula is easily derived from the generating function definition. It is in the form of a discrete convolution. The interesting point here is that $J_o[(n+1)T]$ cannot be written as a finite linear combination of past samples of the sequence $J_o(nT)$. This means that there is no convenient shifting theorem for use with the Bessel functions.

Asymptotic Approximation

The Bessel functions can be readily approximated with a simple formula,

$$J_p(x) \approx \sqrt{\frac{2}{\pi x}} \cos(x - \frac{\pi}{2}p - \frac{\pi}{4}). \qquad (3.12)$$

Equation (3.12) is called the asymptotic approximation for the Bessel function of the first kind, of p-th order. This formula is quite good when the argument x is such that

$$x \geq \frac{3\pi}{4} + \frac{\pi}{2} n. \qquad (3.13)$$

An important special case is the asymptotic approximation to the Bessel function of order zero:

$$J_o(x) \approx \sqrt{\frac{2}{\pi x}} \cos(x - \frac{\pi}{4}). \qquad (3.14)$$

This approximation is the key to the fast algorithm for Fourier-Bessel expansion which will be discussed in a later Chapter. Figures 6 and 7 compare $J_o(x)$ to its asymptotic approximation. Figure 8 shows the absolute error between the Bessel function and its approximation. Note the rapidity with which the error dies off. As a rule of thumb, the error becomes small after the first zero-crossing of the Bessel function, and becomes practically negligible after the second zero crossing. The location of these zero crossings will now be discussed.

Zeros of the Bessel Functions

The zero crossings of $J_o(x)$ are denoted by λ_m where the subscript is the number of the zero crossing. For example, the value of x where $J_o(x)$ crosses the abscissa for the first time is denoted by λ_1. These zero crossings are well approximated by

$$\lambda_m \approx (m - \frac{1}{4})\pi \text{ for } J_o(x). \qquad (3.15)$$

For the Bessel functions of positive integer order n, this becomes:

$$\lambda_m \approx (m - \frac{1}{4} - \frac{n}{2})\pi. \qquad (3.16)$$

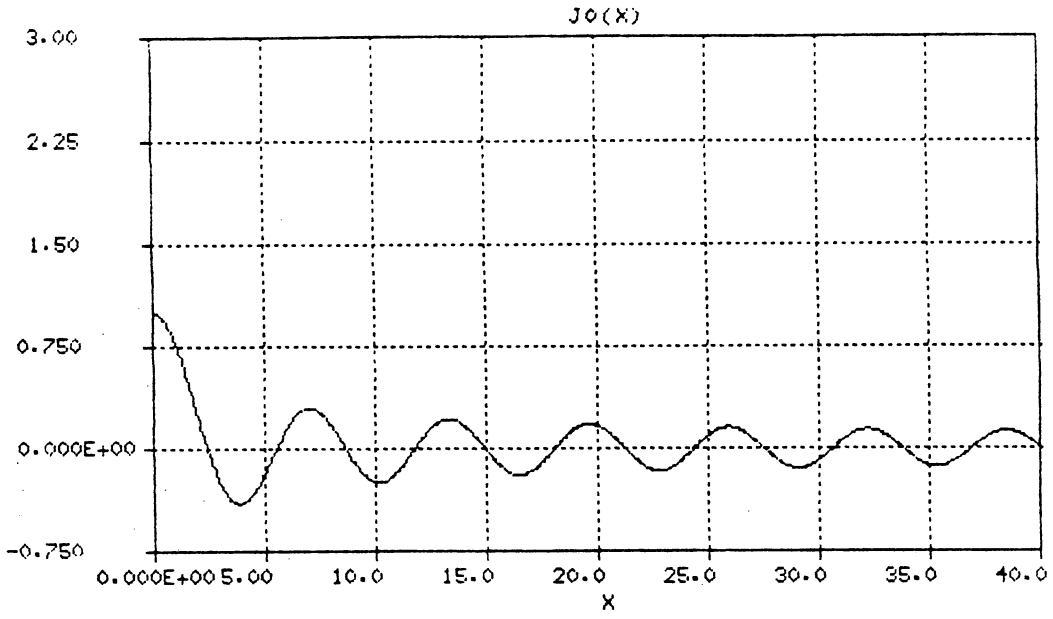


Figure 6. True Bessel function $J_o(x)$.

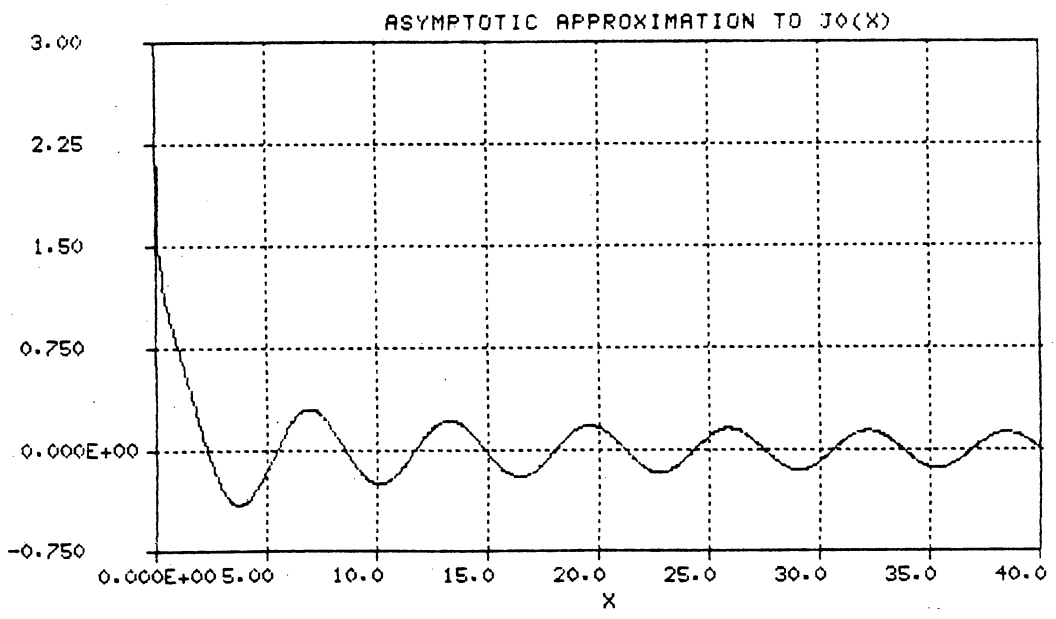


Figure 7. Asymptotic approximation to $J_o(x)$. Note that the approximation becomes large near x=0.

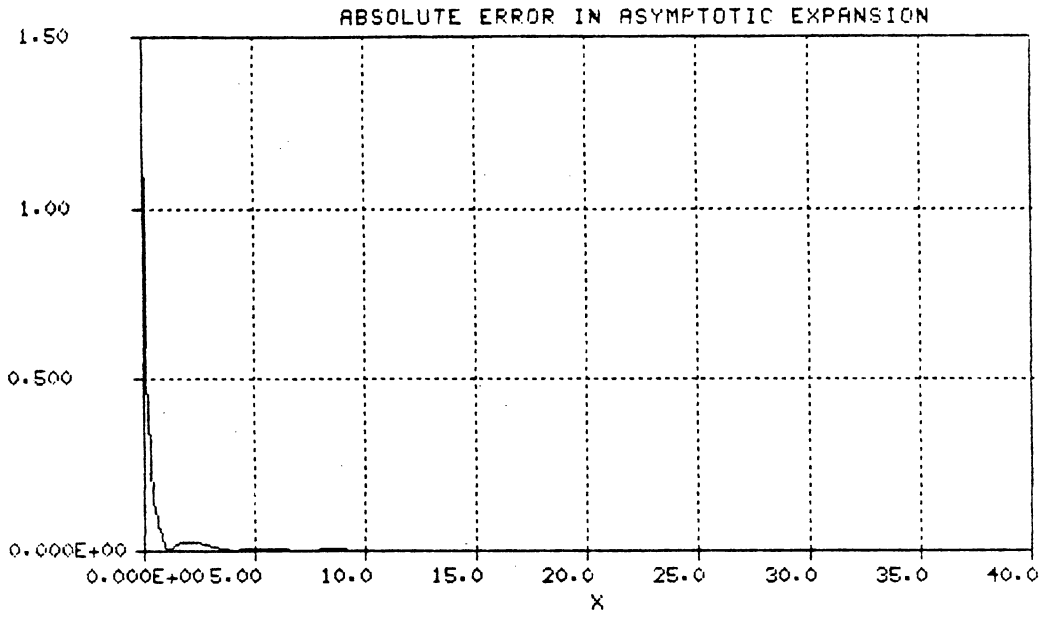


Figure 8. Absolute error in the asymptotic approximation to $J_0(x)$.

The zeros of the Bessel functions are needed in the Fourier-Bessel expansion algorithm, and can be calculated in a computer program which uses the approximation formulas given in Abramowitz and Stegun (1965). Table I compares the zeros of $J_0(x)$ computed by Equation (3.15) to the more exact values found in a table of zeros. Note that the approximation is actually quite good, improving markedly for higher ordered zeros of $J_0(x)$.

Orthogonality Properties

For the computation of a series expansion, the orthogonality properties of the basis set are always of interest. It seems to be a widely held misconception that a basis set has to be orthogonal if a series expansion is to exist. But the truth of the matter is that a set of functions can be complete on an interval, and yet need not be orthogonal. For example, the set of functions $\{x^0, x^1, x^2, x^3, \ldots\}$ can be used as the basis set in the interval $0 \leq x \leq 1$, but the set is not orthogonal. Orthogonality of the basis set is not a requirement, but is a very useful convenience.

Bessel functions are orthogonal, with respect to a weighting function, over a finite interval. Let a and b be zeros of $J_n(x)$ and let the weighting function be t. Then

$$\int_0^1 t J_n(at) J_n(bt)\,dt = \begin{cases} \dfrac{[J_n'(a)]^2}{2} & \text{if } a=b \\ 0 & \text{otherwise.} \end{cases} \qquad (3.17)$$

In this formula, $J_n'(a)$ denotes the first derivative of $J_n(x)$

TABLE I

COMPARISON OF TRUE ZEROS OF J_0 TO APPROXIMATE ZEROS.

m	True λ_m	Approximate λ_m	Relative Error, Percent
1	2.404826	2.356194	2.022228
2	5.520078	5.497787	0.403816
3	8.653728	8.639380	0.165802
4	11.79153	11.78097	0.089572
5	14.93092	14.92257	0.055941
6	18.07106	18.06416	0.038216
7	21.21164	21.20575	0.027749
8	24.35247	24.34734	0.021059
9	27.49348	27.48894	0.016525
10	30.63461	30.63053	0.013312

evaluated at x=a. For $J_0(x)$, the first derivative is $-J_1(x)$. The fact that the Bessel functions are orthogonal with respect to a weighting function other than unity is very important when interpretation of the Fourier-Bessel coefficients is to be considered.

Unfortunately, the sampled Bessel functions are not truly orthogonal. For example, consider the Bessel function of zero order, sampled at N evenly spaced points in the interval $0 \leq x < 1$. When a finite summation is performed over the interval a rather disappointing, albeit important, fact is revealed:

$$\sum_{n=0}^{N-1} \sqrt{\frac{n}{N}}\, J_p\left(a\frac{n}{N}\right) J_p\left(b\frac{n}{N}\right) \approx 0 \text{ for } a \neq b. \qquad (3.18)$$

This shows that when one is dealing with sampled Bessel functions, the mathematically elegant orthogonality property shown by Equation (3.17) degenerates into an approximation. Much of the computational work to be presented in Chapter IV is in terms of approximations, due in part to Equation (3.18). Sampled Bessel functions are mathematically difficult and certainly do not seem to be well known in the literature. In fact, only one published article was found which attempted to deal with sampled Bessel functions from a purely mathematical point of view (Jerri, 1978).

Approximation of Bessel Functions with Small Arguments

The asymptotic approximation previously discussed is

good only for large arguments. Other approximations exist for the case where the argument is small (Waldron, 1981). Starting with the generating function definitions, it can be shown that

$$J_o(x) \approx \frac{1}{6}\left[1+\cos(x)+2\cos\left(\frac{x}{2}\right)+2\cos\left(\frac{3}{2}x\right)\right]. \qquad (3.19)$$

Figures 9-11 show the Bessel function, the approximation given by Equation (3.19), and the absolute value of the error between the two. The error is very small until the argument, x, is greater than about six. At this point, however, the asymptotic approximation could be used.

An alternative definition of Bessel functions is

$$J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j[n\omega - x\sin(\omega)]}\, d\omega. \qquad (3.20)$$

For the special case of $J_o(x)$,

$$J_o(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-jx\sin(\omega)}\, d\omega. \qquad (3.21)$$

When the variable of integration is changed, and the integral is simplified, then it can be easily shown that

$$J_o(x) = \int_0^1 \cos[x\sin(\pi u)]\, du. \qquad (3.22)$$

If this integral is approximated by trapezoidal rule numerical integration then Equation (3.19) results. However, Waldron originally derived (3.19) in a very different manner. By taking more intervals in the numerical integration, better approximations to $J_o(x)$ can be found.

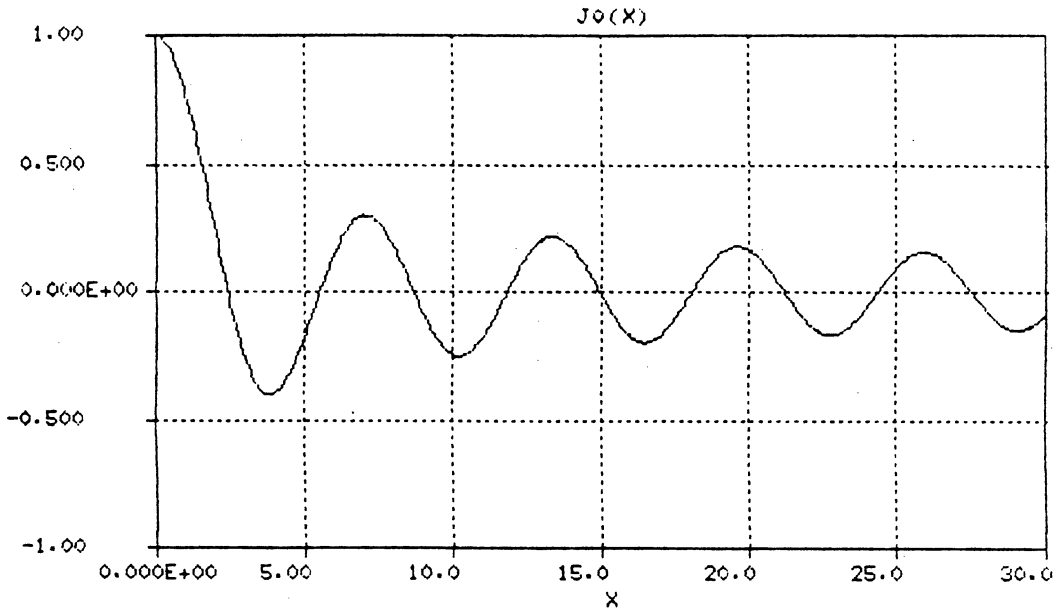


Figure 9. Bessel function $J_o(x)$.

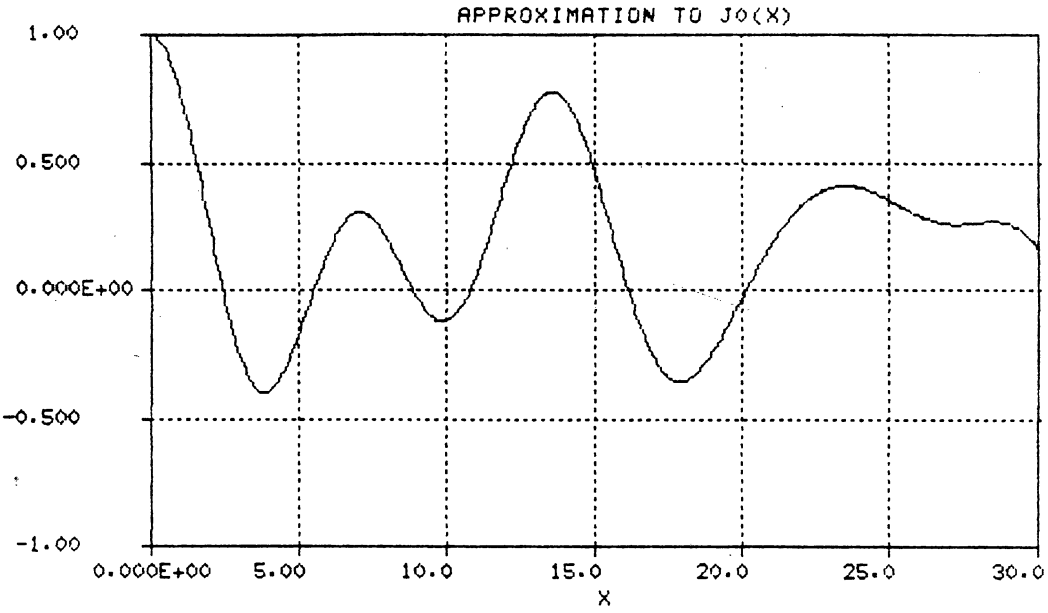


Figure 10. Waldron's approximation to $J_0(x)$ for small values of x.

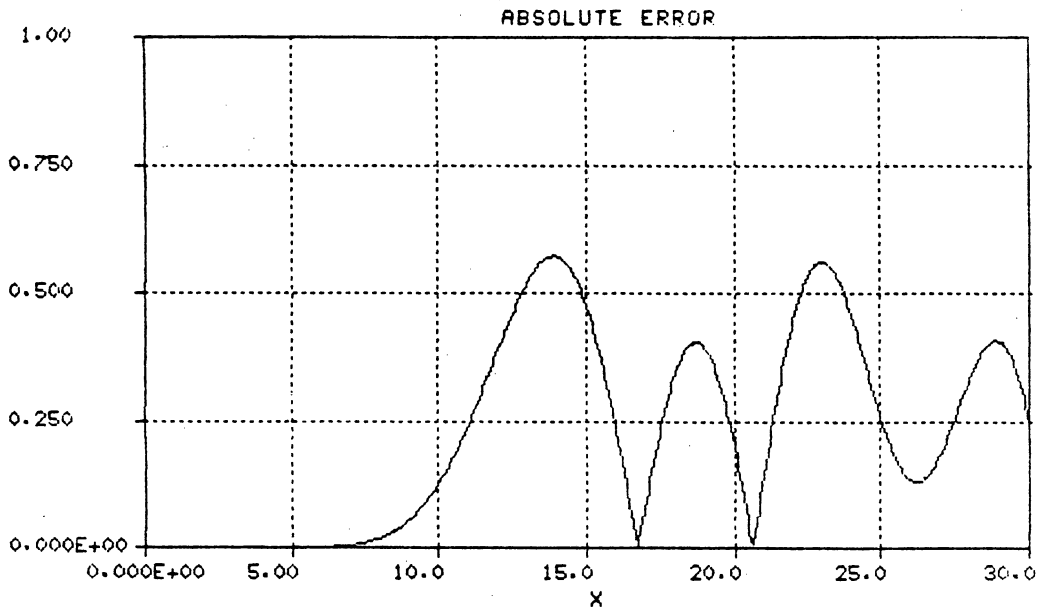


Figure 11. Absolute error of Waldron's approximation to $J_0(x)$.

Chapter Summary

An overview of some of the properties of Bessel functions has been given. Emphasis has been placed on those properties that have been found to be useful in signal analysis, including the Fourier transform, Laplace transform, and the asymptotic approximation. The asymptotic approximation is a very important element in the computational methods to be presented by Chapter IV. It is important to bear in mind that any equations which depend upon the orthogonality of uniformly sampled Bessel functions may be only approximations. However, the worth of an engineering solution should not be measured by significant digits alone; The particular application at hand may require only an approximate solution. Therefore approximate methods can bear much fruit, especially if they are fast and efficient.

CHAPTER IV

FOURIER-BESSEL ANALYSIS

Introduction

Many books and articles have appeared in the mathematical literature during the last century which expound upon the mathematical aspects of the Fourier-Bessel series (Watson, 1945, and Tolstov, 1962). But very little publication of results concerning the application of Fourier-Bessel series to signal analysis has occured. The purpose of this chapter is to define the Fourier-Bessel series and show how the coefficients of the series can be efficiently obtained using a digital computer.

Definition of the Fourier-Bessel Series

The Fourier-Bessel series is a model of a waveform or signal wherein the signal is assumed to be a linear combination of Bessel functions of the first kind:

$$f(x) = \sum_{m=1}^{\infty} C_m J_p(\lambda_m x). \qquad (4.1)$$

The C_m's are the series coefficients, or model parameters. The λ_m's are the zeros of the Bessel function. This model is assumed to be valid only over a finite interval, from x=0 to

x=1. This is not at all a restrictive requirement, because any time interval can be scaled to be in this range by dividing the actual time by the length of the analysis frame, T. Throughout this discussion, the variable x is considered as a normalized time variable in the range $0 \leq x \leq 1$.

Properties of the Fourier-Bessel Series

Convergence

The convergence of the Fourier-Bessel series is guaranteed if: (1) f(x) is a piecewise smooth continuous or discontinuous function on [0,1]; and (2) The Bessel function is of order -(1/2) or greater. The series converges to f(x) wherever f(x) is continuous, and at points of discontinuity it converges to $[f(x^+)+f(x^-)]/2$ (Tolstov, 1962, p.221). These conditions are certainly met by bandlimited signals, such as typical speech signals, because bandlimited signals are analytic (Papoulis, 1977).

Boundary Conditions

For the special case where the Bessel function is of order zero, representative elements of the basis set are shown in Figures 12 and 13. Note that all of the basis functions cross through zero at x=1. Therefore, the series can only converge to zero at x=1. This is true for all Fourier-Bessel series of the form shown by Equation (4.1), not just for those based upon $J_0(x)$. Although convergence is certain, the number of terms of the series that must be summed to

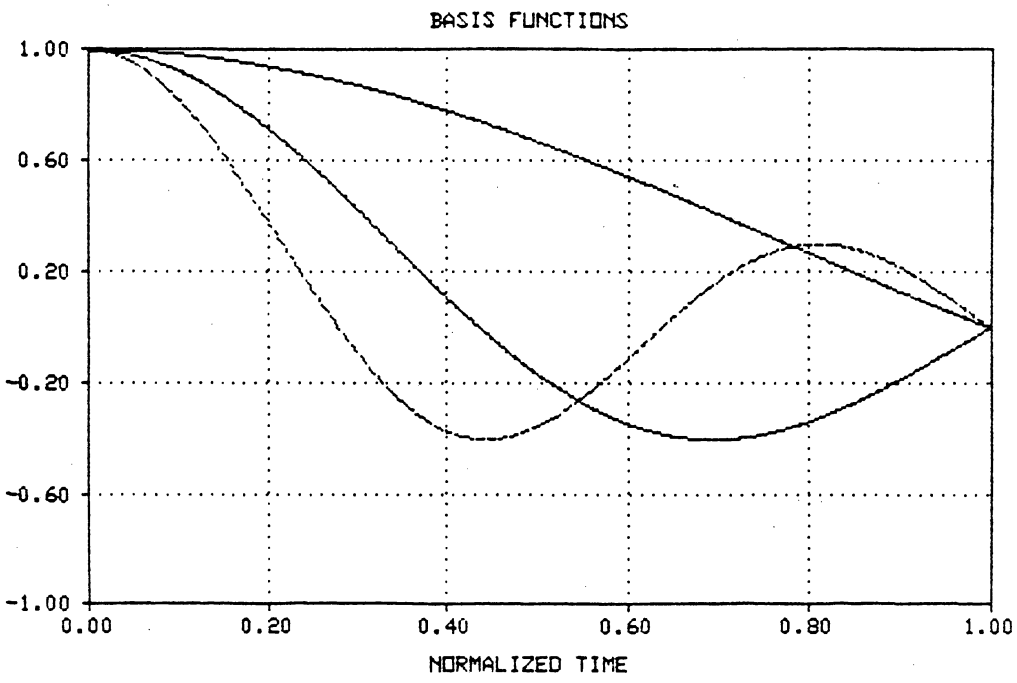


Figure 12. Basis functions $J_o(\lambda_1 x)$, $J_o(\lambda_2 x)$, and $J_o(\lambda_5 x)$. All three functions cross zero at x=1.

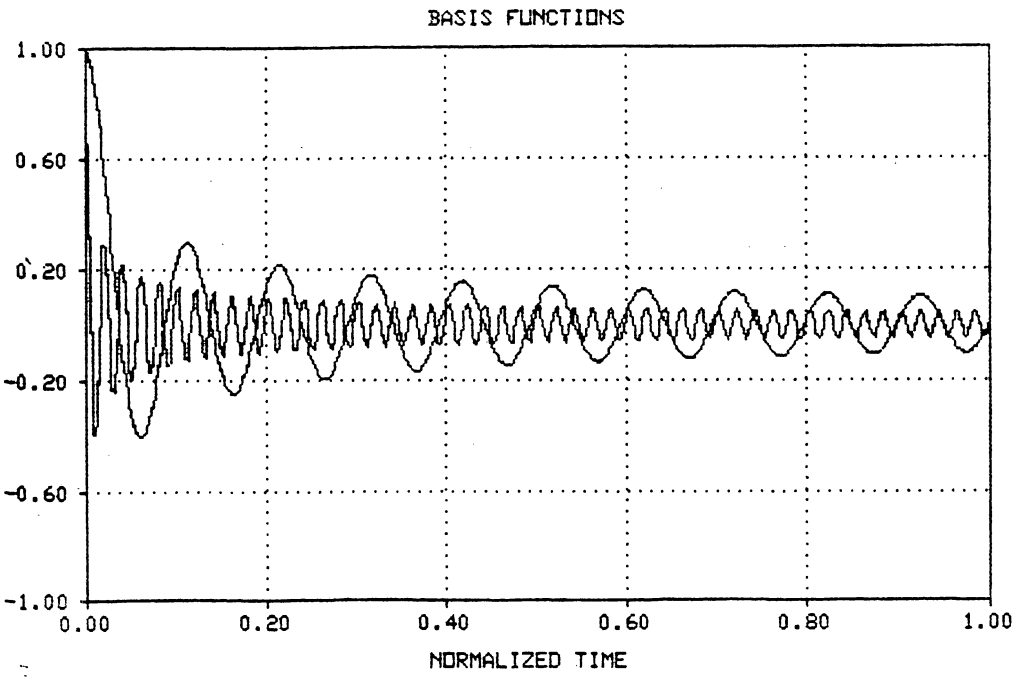


Figure 13. Basis functions $J_0(\lambda_{20}x)$ and $J_0(\lambda_{100}x)$. Note the very slow damping and sinusoidal character of the functions.

arrive at a reasonable approximation to f(x) may be large if the value of f(x) is very different from zero near x=1.

At x=0, a different set of restrictions applies. For Bessel functions of zero order, $J_o(0)=1$ and the derivative at x=0 is zero. Therefore, a linear combination of such Bessel functions must have a zero derivative near x=0. If the basis set is $\{J_p(\lambda_m x)\}$ where p is an integer greater than zero, then every element of the basis set is zero at x=0; therefore, the series converges to zero at x=0.

Exact Methods of Computing the Series Coefficients

Note that if a is not equal to b in Equation (3.17), and they are both zeros of the Bessel function, then the integral is identically zero. But if a equals b in (3.17) the result is in general nonzero. If f(x) is assumed to be of the form given in Equation (4.1) then substitution of the assumed form gives:

$$\int_0^1 x J_n(\lambda_m x) f(x)\,dx = \int_0^1 x J_n(\lambda_m x) \left[\sum_{k=1}^{\infty} C_k J_n(\lambda_k x) \right] dx$$

$$= \sum_{k=1}^{\infty} C_k \int_0^1 x J_n(\lambda_m x) J_n(\lambda_k x)\,dx \qquad (4.2)$$

$$= C_m \frac{1}{2} \left[J_n'(\lambda_m) \right]^2$$

Since the λ_m's are the zeros of the Bessel function, only one term of the entire summation can be nonzero (recall the

orthogonality formula, Equation (3.17)). Isolating the C_m on one side of the equation yields the closed-form expression for the Fourier-Bessel series coefficients:

$$C_m = \frac{2}{[J_n'(\lambda_m)]^2} \int_0^1 x J_n(\lambda_m x) f(x) dx. \qquad (4.3)$$

For the important special case where the basis set is $\{J_0(\lambda_m x)\}$, the first derivative of $J_0(x)$ is $-J_1(x)$. Calculation of the series coefficients then amounts to calculation of

$$C_m = \frac{2}{[J_1(\lambda_m)]^2} \int_0^1 x J_0(\lambda_m x) f(x) dx. \qquad (4.4)$$

Unfortunately, the integral must be evaluated using sampled data rather than a continuous f(x) in closed form, and the integral must be performed many times for the various values of the index m. Even when f(x) is known in closed form Equation (4.4) is usually very difficult to evaluate, except for a few special cases. This makes obvious the need for a fast algorithm which can quickly approximate Equation (4.4).

Approximate Methods for Calculating the Fourier-Bessel Series Coefficients

Some very useful algorithms have appeared which allow the fast computation of the Hankel transform integral,

$$F(\omega) = \int_0^\infty r J_0(\omega r) f(r) dr. \qquad (4.5)$$

There are four basic approaches for fast machine evaluation

of Equation (4.5). These four will now be described.

The Correlation Method

Siegman (1977) proposed a method based upon a change of variables. After a change of variable in Equation (4.5), the integral is cast in the form of a discrete correlation. Details can also be found in Oppenheim, Frisk, and Martinez (1980). Siegman's method allows the use of the FFT, but has a severe disadvantage for speech processing: The original speech signal has to be sampled at exponentially spaced points, and the result it produces, $F(\omega)$, is given at exponentially spaced points. Although such sampling is sometimes appropriate in optics or image processing, it is not at all suitable for one-dimensional signal processing.

Tsang's Method

Tsang et al. (1974) used a method of Hankel transformation based upon the FFT. First, the original function $f(r)$ is windowed with an exponentially tapered function. Then it is Fourier transformed via the FFT to obtain a function called $A(\lambda)$. A previously computed and stored weighting function $I(\lambda,\omega)$ is multiplied by $A(\lambda)$, and the result is numerically integrated for many values of λ to get the Hankel transform $F(\omega)$. The disadvantage of this method is that the function $F(\omega)$ is found by repeated integrations: If each numerical integration is time-consuming in a computer, then numerous integrations are dreadfully awkward for use in

any sort of real-time application such as coding or recognition.

The Projection-Slice Method

This method is based upon a theorem of two-dimensional Fourier transforms which states that the one-dimensional transform of the projection of a two-dimensional image f(x,y) onto some line (at a given angle) is equal to a radial section, or slice, of the two-dimensional Fourier transform of f(x,y) at the same angle (Mersereau and Oppenheim, 1974). The Hankel transform is related to the Fourier transform of an image, if the image is radially symmetric (Bracewell, 1965). Details are omitted here, but the most important facet of this method is that either the original data is sampled exponentially and the data is evenly spaced in the transform domain or, by duality, the data can be sampled uniformly and the result will be exponentially sampled in the transform domain (Oppenheim, Frisk, and Martinez, 1978).

Candel's Method

The most useful approach has been found to be Candel's algorithm for computation of the Fast Hankel Transform (Candel, 1981). The gist of Candel's method is that the Hankel transform kernel (a Bessel function) may be replaced by its asymptotic approximation. When such a substitution is made, the computed result is only approximate; but the great ad-

vantage is that the Hankel transform integrand is cast in terms of an ordinary trigonometric function. Also, the data is uniformly sampled in both domains. A Fast Fourier Transform may be used to approximate the integral: The algorithm becomes fast. With a fast algorithm available for the computation of the Fourier-Bessel series coefficients, and with the data sampled uniformly, the Fourier-Bessel series can be of practical use in one-dimensional signal analysis and speech processing.

In the following sections, algorithms for the process of Fourier-Bessel expansion are given where the basis set is assumed to be the set $\{J_o(\lambda_m x)\}$. Generalization to other integral-ordered Bessel functions is fairly straightforward and will not be given explicitly.

Implementation of the Approximate Fourier-Bessel Expansion Using the Fast Fourier Transform

The method described by Candel uses a technique which he called "Fourier-selection summation" to correct the estimates of the integrals which were computed by a Fast Fourier Transform, or FFT. The method requires two FFT's to be performed on each frame of data: The first FFT is used to estimate the Hankel transform of the data, and the second is used to correct the estimate of the integral. Experiments have shown that a second transform need not actually be performed because the errors are negligible for actual speech

signals. Besides, an alternative method was found for the correction of the small errors that do occur (mostly in the first few coefficients, which correspond roughly to the lowest frequencies in the speech signal). This error correction method will be discussed later.

Reconsider Equation (4.4) with the asymptotic approximation of Equation (3.14) substituted for the Bessel function. The expression thus obtained is

$$C_m \approx \frac{2}{[J_1(\lambda_m)]^2} \int_0^1 x f(x) \sqrt{\frac{2}{\pi \lambda_m x}} \cos(\lambda_m x - \tfrac{\pi}{4}) dx. \quad (4.6)$$

This expression is still not in the desired form, because in reality all that is available is sampled data and a finite summation must be performed. Suppose that rectangular-rule numerical integration is performed, as illustrated in Figure 14. The total number of integration intervals is N, so that

$$\Delta x = \frac{\text{Total Length}}{N} = \frac{1}{N} . \quad (4.7)$$

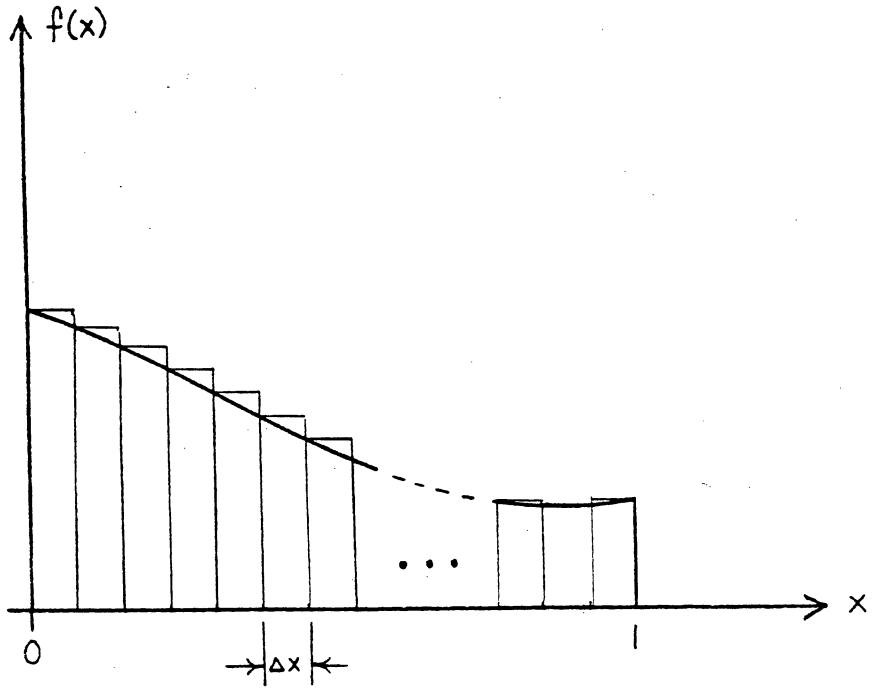
Equation (4.6) becomes

$$C_m \approx \frac{2\sqrt{2}}{[J_1(\lambda_m)]^2 \sqrt{\pi \lambda_m}} \sum_{n=0}^{N-1} \sqrt{\frac{n}{N}} f(\tfrac{n}{N}) \cos(\lambda_m \tfrac{n}{N} - \tfrac{\pi}{4}) \tfrac{1}{N} \quad (4.8)$$

Substitution of the approximation of Equation (3.15) for the zeros of the Bessel function yields, after some algebraic calisthenics,

$$C_m \approx \frac{2\sqrt{2}}{[J_1(\lambda_m)]^2 \sqrt{\pi \lambda_m}\, N^{1.5}} \sum_{n=0}^{N-1} \sqrt{n}\, f(\tfrac{n}{N}) \cos\left[\frac{nm\pi}{N} - (1+\tfrac{n}{N})\, \tfrac{\pi}{4}\right] \quad (4.9)$$

This expression is, in turn, equivalent to



$$\int_0^1 f(x)\,dx \approx \sum_{n=0}^{N-1} f(x_n)\,\Delta x = \sum_{n=0}^{N-1} f\left(\frac{n}{N}\right)\frac{1}{N}$$

Figure 14. Numerical integration. Rectangular rule used to approximate transform integrals.

$$C_m \approx K_m \, \mathrm{Re}\left\{ \sum_{n=0}^{N-1} \left[n^{.5} \, f\left(\frac{n}{N}\right) e^{j\pi(1+\frac{n}{N})/4} \right] e^{-j\frac{2mn\pi}{2N}} \right\} \qquad (4.10)$$

where the function Re(x) denotes the "real part of x" and where the set of constants, K_m, is defined by

$$K_m = \frac{\sqrt{2}\ 2}{[J_1(\lambda_m)]^2 \sqrt{\pi \lambda_m}\ N^{1.5}} \qquad (4.11)$$

The expression in Equation (4.10) is in a form suitable for summation using the Fast Fourier Transform. The algorithm will now be summarized:

STEP 1. Calculate the constants $n^{.5}\ e^{j\pi(1+n/N)/4}$ for $n = 0,1,2,\ldots,N-1$. Store in array A(n), which should be complex, dimensioned as A(0:2*N - 1).

STEP 2. Calculate constants K_m for m=1,2,... and store in array B(m). See Equation (4.11).

STEP 3. Multiply sequence f(n/N) element-by-element with array A(n) and place results in complex array C(n). Zero-pad C(n) to total length 2N.

STEP 4. Perform an FFT of length 2N on array C(n). Do not divide by 2N.

STEP 5. Multiply real part of array C element-by-element with the constants stored in array B to get the approximate coefficients:

$C_m \approx$ Re(C(m))*B(m) for m=1,2,...

Note that the preliminary Steps 1 and 2 need only to be performed once; the constants are computed once and then saved.

Recall that the asymptotic approximation is fairly

accurate only for the larger values of the argument of the Bessel function. A mitigating factor is that the Bessel function is actually multiplied by the weighting function x. The approximations to the functions $xJ_o(\lambda_m x)$ are compared to the exact functions in Figures 15 through 18 for m=1,3, 5,and 50. The errors decrease in absolute value for larger values of m. The error of the asymptotic approximation is large near x=0 (refer again to Figure 8). But when multiplied by x, the error no longer becomes unbounded near x=0. This factor tends to make the estimates of the integrals more reasonable.

Implementation of the Approximate Fourier-Bessel Expansion Using the Fast Hartley Transform

One of the modifications made to the previously used Candel algorithm was the introduction of the Hartley Transform. This transform was originally described by R.V. Hartley (1942), and has recently been revived in discrete form by Bracewell (1983). The Candel algorithm uses a Fast Fourier Transform. But it was discovered that the Fast Hartley Transform can be used for the task of Fourier-Bessel expansion because it is a real-valued transform with a built-in phase shift in its kernel which exactly matches the phase of the asymptotic approximation used in Candel's algorithm, which is equal to $\pi/4$.

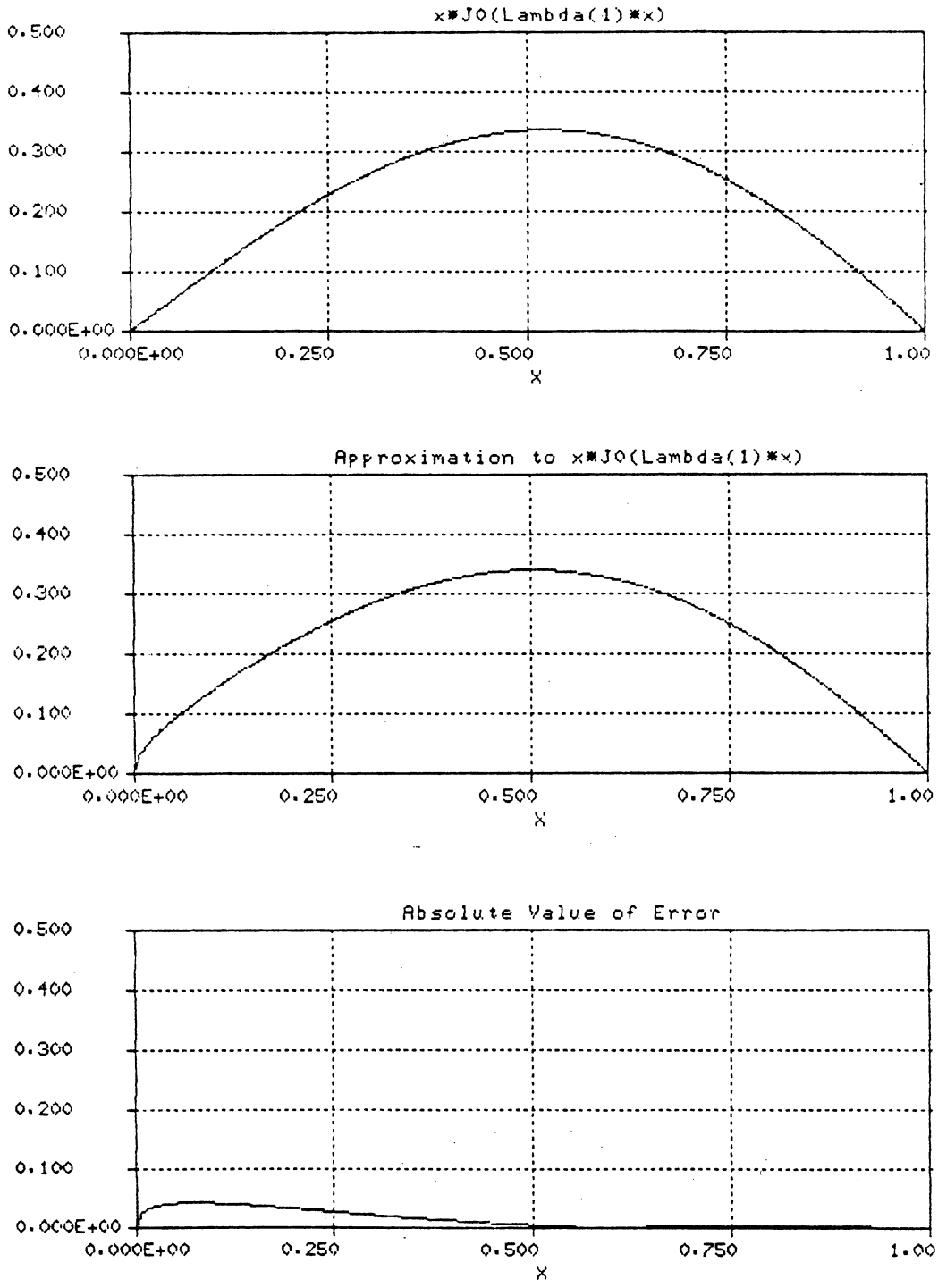


Figure 15. Comparison of true integrand $xJ_0(\lambda_1 x)$ to its approximation. Top: True integrand. Middle: Approximation. Bottom: Absolute error.

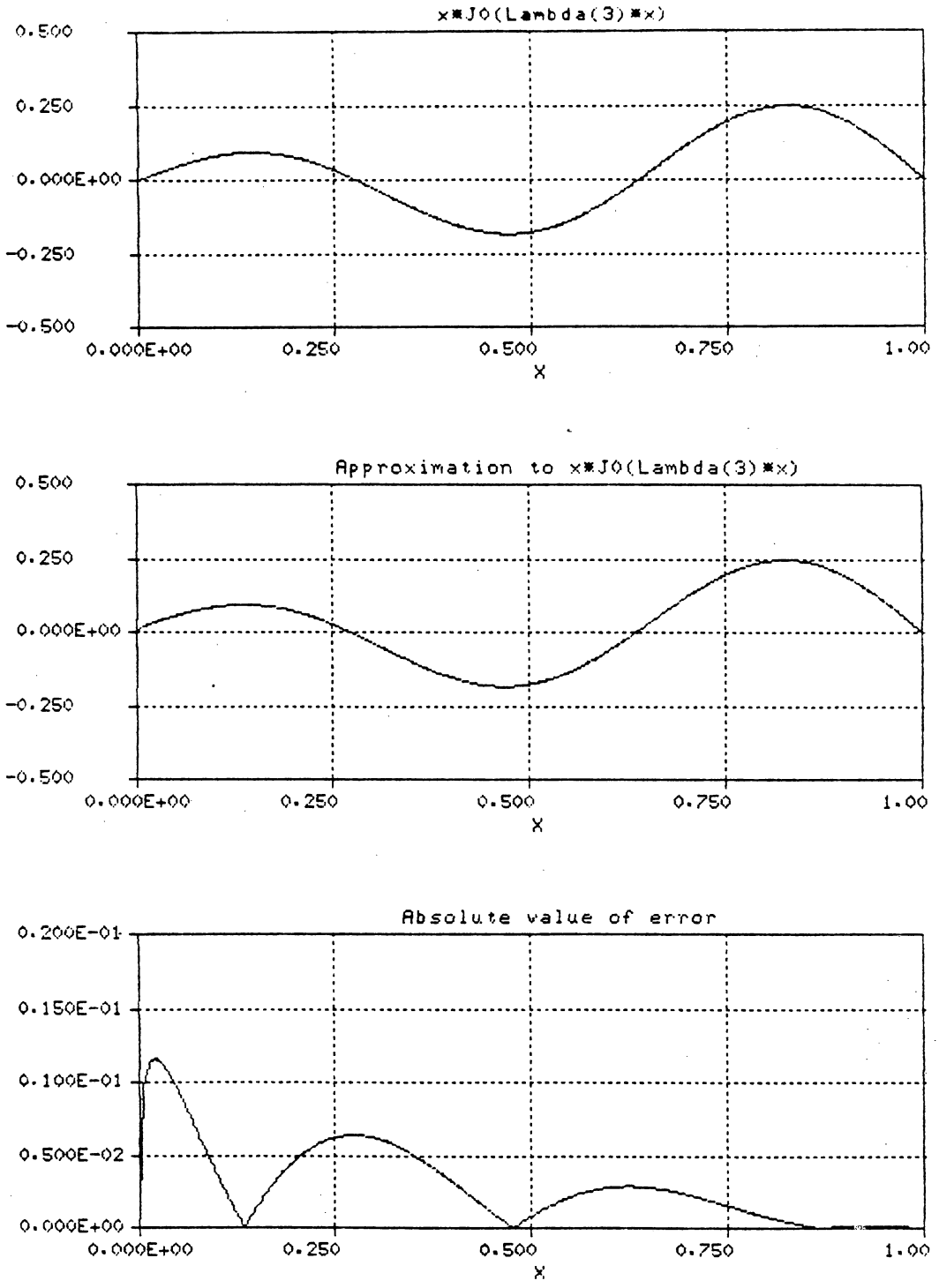


Figure 16. Comparison of true integrand $xJ_0(\lambda_3 x)$ to its approximation. Top: True integrand. Middle: Approximation. Bottom: Absolute error.

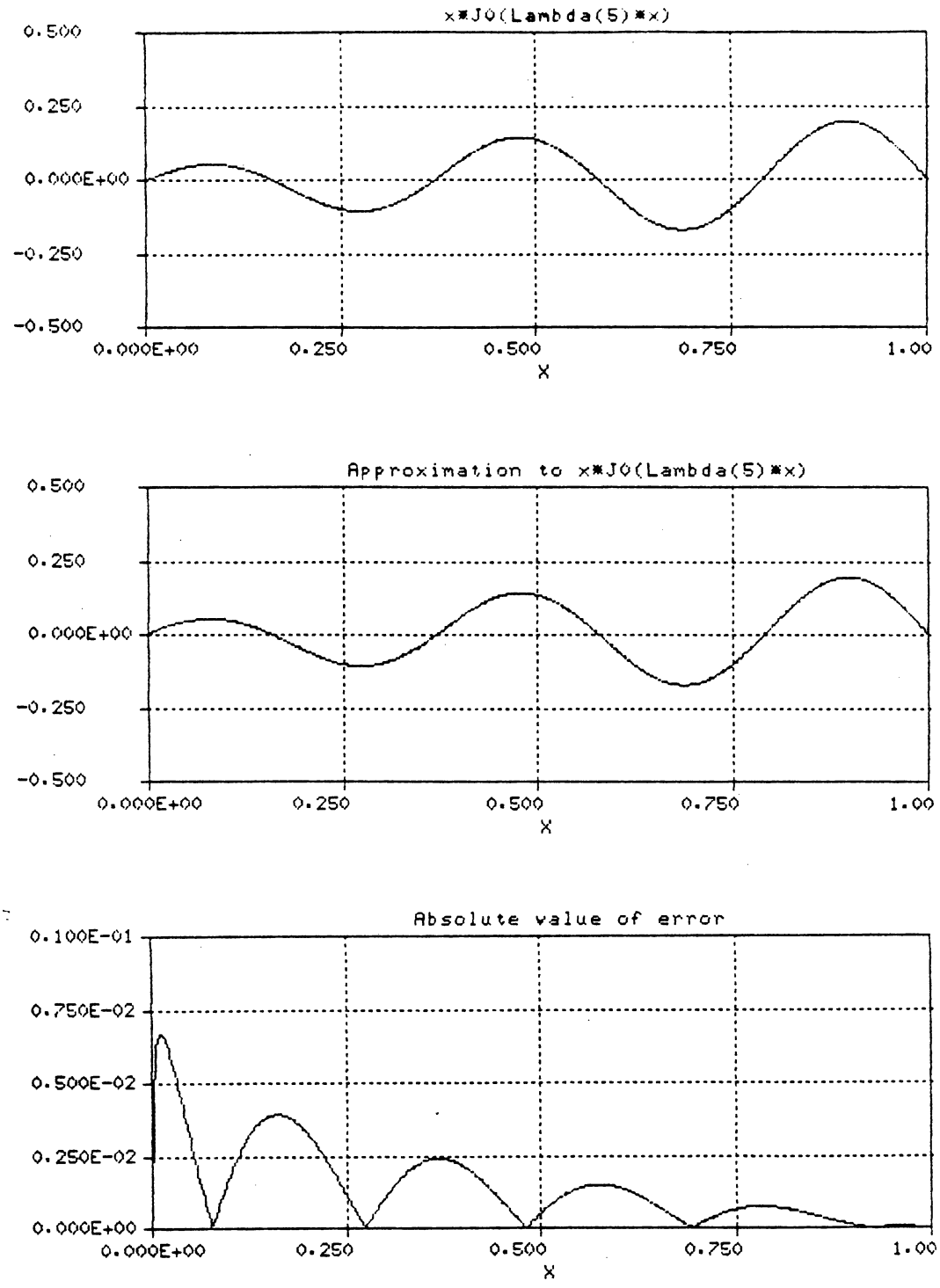


Figure 17. Comparison of true integrand $xJ_0(\lambda_5 x)$ to its approximation. Top: True integrand. Middle: Approximation. Bottom: Absolute error.

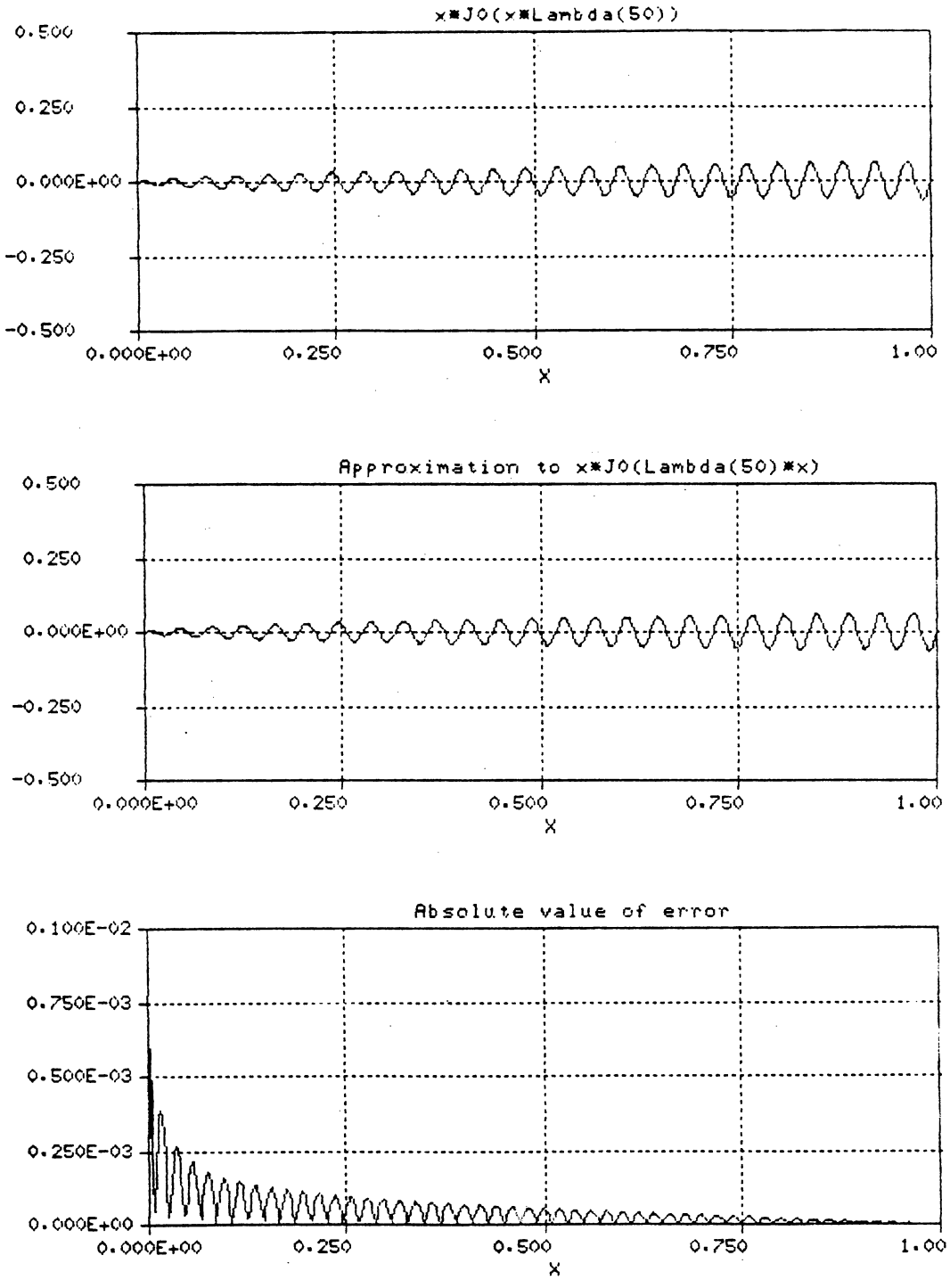


Figure 18. Comparison of true integrand $xJ_0(\lambda_{50}x)$ to its approximation. Top: True integrand. Middle: Approximation. Bottom: Absolute error.

The Hartley Transform is defined as

$$X(f) = \int_{-\infty}^{\infty} x(t)\ \mathrm{cas}(2\pi ft)\ dt \qquad (4.12)$$

and its inverse is given by

$$x(t) = \int_{-\infty}^{\infty} X(f)\ \mathrm{cas}(2\pi ft)\ df \qquad (4.13)$$

where

$$\mathrm{cas}(x) = \cos(x) + \sin(x). \qquad (4.14)$$

Note the symmetrical nature of the forward and reverse transforms.

The Discrete Hartley Transform (DHT) can be defined by

$$X(k) = \sum_{n=0}^{N-1} x(n)\ \mathrm{cas}\left(\frac{2\pi}{N} nk\right) \qquad (4.15)$$

and its inverse is then

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)\ \mathrm{cas}\left(\frac{2\pi}{N} nk\right). \qquad (4.16)$$

The Fast Hartley Transform used in this research is a variant of the one given by Bracewell (1984). The Fast Hartley Transform was modified by making it an in-place algorithm. The Fast Hartley Transform has been found to be a very useful and powerful tool, so Fortran source code for this fast transform has been included in Appendix A.

The kernel of the DHT can be rewritten using simple trigonometric identities as

$$\mathrm{cas}\left(\frac{2\pi}{N} nk\right) = 2\cos\left(\frac{2\pi}{N} nk - \frac{\pi}{4}\right). \qquad (4.17)$$

Comparison of Equation (4.17) with Equation (4.8) shows that the summation in (4.8) can be performed using a DHT. The summation can be restated as

$$C_m \approx \frac{2}{[J_1(\lambda_m)]^2\sqrt{\pi\lambda_m}} \sum_{n=0}^{N-1} \left(\frac{n}{N}\right)^{.5} f\left(\frac{n}{N}\right) \operatorname{cas}\left(\lambda_m \frac{n}{N}\right) \frac{1}{N} \qquad (4.18)$$

or

$$C_m \approx K'_m \sum_{n=0}^{N-1} n^{.5} f\left(\frac{n}{N}\right) \operatorname{cas}\left[(m-.25)\pi\frac{n}{N}\right] \qquad (4.19)$$

where

$$K'_m = \frac{2}{[J_1(\lambda_m)]^2\sqrt{\pi\lambda_m}\ N^{1.5}}. \qquad (4.20)$$

The algorithm is now summarized:

STEP 1. Compute the K'_m and store them in an array K(m). This step is preliminary and needs to be performed once.

STEP 2. Multiply the sequence $f(\frac{n}{N})$ by $n^{.5}$ for n=0,1,..., N-1 and store the results in real array A(n).

STEP 3. Zero-pad array A(n) to a total length of 8N. Array A should be real, dimensioned as A(0:8*N-1).

STEP 4. Perform a Fast Hartley Transform of length 8N on array A(n). (See Appendix A for details.)

STEP 5. For m=1,2,... let the Fourier-Bessel coefficients be approximated by $C_m \approx K'_m$*A(4m-1).

This Fast Hartley Transform method uses more memory than the Fast Fourier Transform based method. An alternative method based on the Fast Hartley Transform will now be presented which uses less memory than the FFT-based method, and

which runs faster than the DHT-based method just presented.

Consider the trigonometric identity

$$\mathrm{cas}\left(\frac{m\pi n}{N} - \frac{\pi n}{4N}\right) = \mathrm{cas}\left(\frac{m\pi n}{N}\right)\cos\left(\frac{\pi n}{4N}\right) + \mathrm{cas}\left(\frac{-m\pi n}{N}\right)\sin\left(\frac{\pi n}{4N}\right). \quad (4.21)$$

Then the summation in (4.20) can be split into two separate sums:

$$C_m \approx K_m' \left[\sum_{n=0}^{N-1} n^{.5}\cos\left(\frac{n\pi}{4N}\right) f\left(\frac{n}{N}\right) \mathrm{cas}\left(\frac{m\pi n}{N}\right) + \sum_{n=0}^{N-1} n^{.5}\sin\left(\frac{n\pi}{4N}\right) f\left(\frac{n}{N}\right) \mathrm{cas}\left(-\frac{m\pi n}{N}\right) \right] \quad (4.22)$$

Each of the sums can be computed simultaneously using a DHT of length 2N. This makes use of the Discrete Hartley Transform identity

$$\sum_{n=0}^{N-1} x(n)\mathrm{cas}\left(\frac{-2\pi}{N}kn\right) = \sum_{n=0}^{N-1} x[(N-n)\bmod(N)]\mathrm{cas}\left(\frac{2\pi}{N}kn\right). \quad (4.23)$$

The algorithm is summarized:

STEP 1. Compute the constants K_m' for m=1,2,... Save them in an array K(m). (See Equation (4.20).)

STEP 2. Dimension REAL array A(0:2*N-1).

STEP 3. For n=0 to N-1 DO:

A(n) ⟵ SQRT(n)*COS(PI/4 * n/N)*f(n).

STEP 4. For n=1 to N-1 DO:

A(2*N - n) ⟵ SQRT(n)*SIN(PI/4 * n/N)*f(n).

STEP 5. Let A(N) = 0.

STEP 6. Perform the Fast Hartley Transform of length 2N, operating on array A. (See Appendix A for details.)

STEP 7. For m=1,2,... the Fourier-Bessel coefficients are

$$C_m \approx K'_m * A(m).$$

This version avoids complex arithmetic, thus saving memory.

Matrix-Vector Interpretation of the Fourier-Bessel Expansion

An alternative method is based on the fact that the estimation of the Hankel transform (by a Fast Hartley Transform or a Fast Fourier Transform) as well as the desired true transform can each be represented by a matrix-vector multiplication. For example,

$$\underline{c} = B\underline{x} \tag{4.24}$$

where B is the desired transformation matrix, $\underline{x}$ is the data vector of sampled speech, and $\underline{c}$ is the vector of Fourier-Bessel coefficients. The approximation is expressed as

$$\underline{c} \approx M\underline{x} \tag{4.25}$$

where M is the linear transformation performed by one of the algorithms in the two previous sections. The rows of these transformation matrices contain numbers which, when plotted, look like Figures 15-18.

The difference of the desired transform matrix, B, and the approximation matrix, M, is another matrix, E, with the computationally important feature that most of its elements are very nearly equal to zero; i.e., the matrix is sparse if all of its relatively small elements are set equal to zero.

$$E = B - M \tag{4.26}$$

Typically, relatively few of the matrix E's elements have an

error greater than one percent of the largest element in absolute value on that row, and the others are set equal to zero. After all of the very small elements are set equal to zero, the resulting sparse error matrix is very easy to multiply by the input vector of speech samples to get a correction vector. The correction vector is simply added to the vector of approximate Fourier-Bessel coefficients to result in a better estimate of the true coefficient set. This method was found to be computationally efficient (the sparseness of the correction matrix is the key) and conceptually far simpler than Candel's Fourier selection-summation. In matrix notation, this becomes:

$$\underline{c} = M\underline{x} + E\underline{x}. \qquad (4.27)$$

In practice, it was found that the corrections need only to be performed on the first few Fourier-Bessel coefficients, because the errors are greatest there. The actual error correction procedure consists of multiplying the data vector $\underline{x}$ with the first few rows of the E matrix and then adding the result to the vector obtained from the fast algorithm.

Resynthesis of a Waveform From Its Fourier-Bessel Series

There are two basic methods of resynthesis: (1) exact resynthesis, involving direct summation of the series; and (2) fast approximate resynthesis, using Fast Fourier Transforms or Fast Hartley Transforms.

The direct summation method is best performed as a matrix-vector multiplication. The columns of the matrix consist of the sampled Bessel functions, and the vector is the set of coefficients. This algorithm is good for the testing of fast transformation algorithms, but is too slow for real-time applications. This method is represented by Equation (4.28):

$$f(nT) = \sum_{m=1}^{M} C_m J_o(\lambda_m nT) \qquad (4.28)$$

where M is the number of terms in the series, or the number of columns in the matrix.

The fast algorithm is based upon the same asymptotic approximation as before (Equations (3.12) or (3.14)). The summation is then

$$f(x) \approx \sum_{m=1}^{M} C_m \sqrt{\frac{2}{\pi \lambda_m x}} \cos(\lambda_m x - \frac{\pi}{4}). \qquad (4.29)$$

Manipulating the sum into a form suitable for the FFT, the result is

$$f(n) \approx \sqrt{\frac{2N}{n\pi}} \, Re\left[e^{j(\frac{\pi}{4})(1+\frac{n}{N})} \sum_{m=1}^{M} \frac{C_m}{\lambda_m} e^{-j(\frac{\pi nm}{N})} \right]. \qquad (4.30)$$

Using the Fast Hartley Transform, the result is similarly derived and will not be repeated here.

Data Windows

In traditional Fourier analysis, the use of the data windows such as Hamming, Kaiser, etc., is justified by the

argument that multiplication in the time domain is equivalent to the operation of convolution in the frequency domain. The spectral representations of signals thus obtained "look" smoother when plotted, and the results are, in general, more useful.

But there is a problem when the subject of consideration is alternative basis sets (such as Bessel functions): The convolution theorem no longer holds, so the typical arguments for using tapered windows cannot be used. There is little theoretical justification for using one window or another when Fourier-Bessel expansion is to be performed. This is a subject that is remarkably absent from the literature, and which could possibly be the subject of fruitful investigation. Further insight into the traditional sinusoidal basis sets and their usage could be a useful by-product of such research.

If the original data vector was windowed with, say, a Hamming window, then the Fourier-Bessel series accurately represented the windowed data. The Bessel functions are close relatives of the familiar trigonometric functions; so the results of windowing have been found to be similar.

In practice, it was found that data windows which forced the sequence to obey the boundary conditions set forth in this Chapter tend to give good, reliable results. For the basis set $\{J_o(\lambda_m x)\}$, good windows were the Hamming lag window (Figure 19) and the tapered cosine window (Figure 20). For Bessel functions which are zero at both endpoints

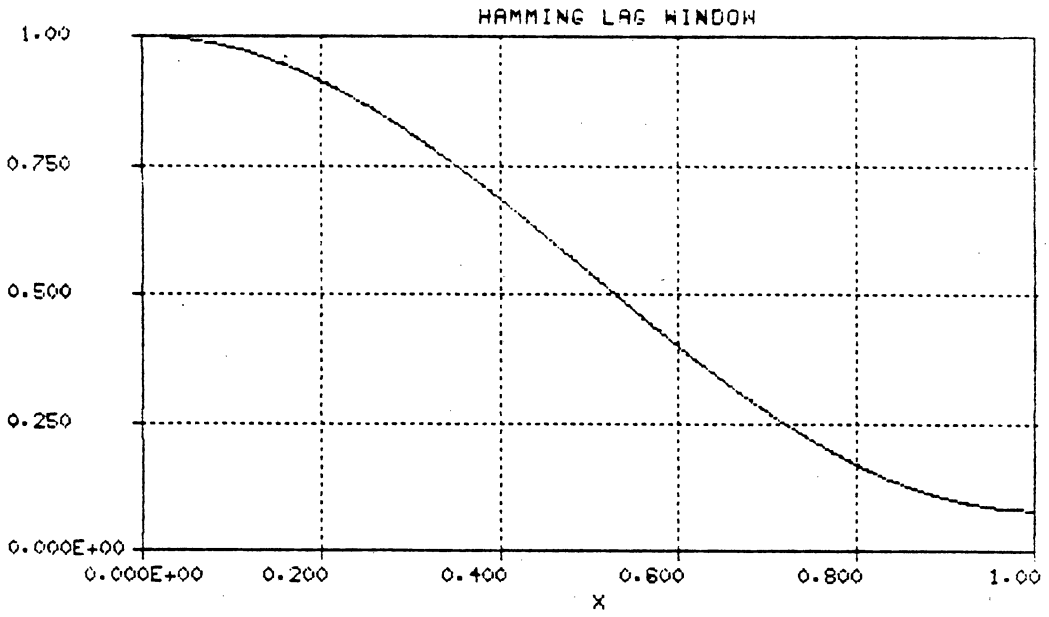


Figure 19. Hamming lag window. This is half of a traditional Hamming window.

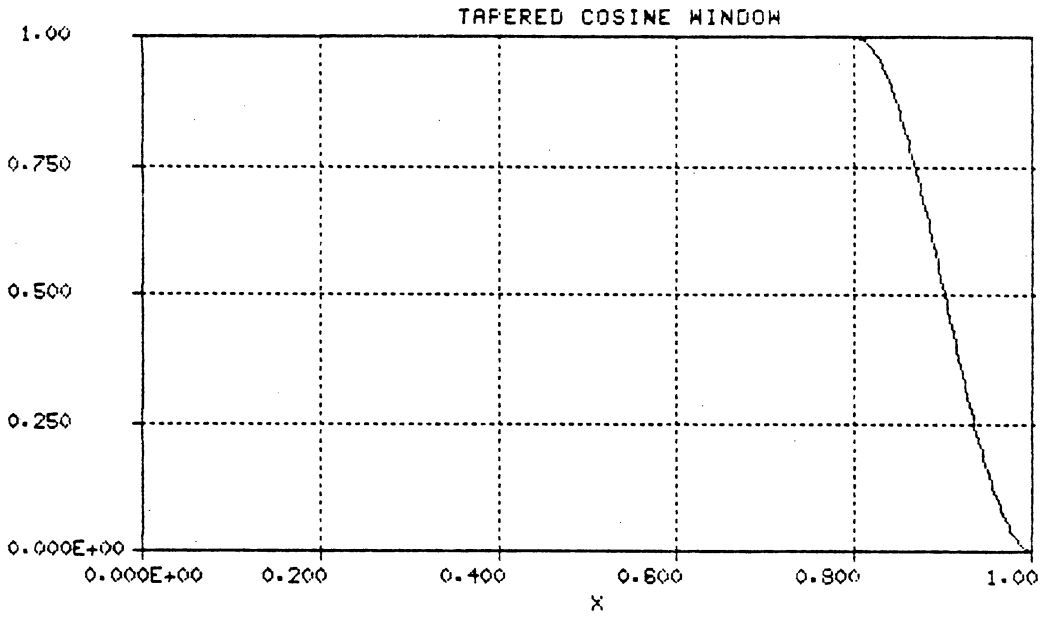


Figure 20. Tapered cosine window. Many such variations are possible.

of the interval (such as $J_1(x)$), a window which tapers to zero at both ends should be used.

Ideally, if the waveform being expanded is one of the members of the basis set then one and only one of the coefficients of the expansion should be unity, and all the others should be zero. As an example, the function $J_0(\lambda_{50}x)$ was expanded into a Fourier-Bessel series. The results obtained when using rectangular, traditional Hamming, and Hamming lag windows are shown in Figure 21. Note that the rectangular window gave the representation which was truest to the actual data: a single coefficient of unity strength at m=50. The other windows caused leakage into the adjacent coefficients.

Now as another example, let the waveform be equal to $J_0[(\lambda_{50}+\lambda_{51})x/2]$ in the interval [0,1]. This time, the "frequency" falls halfway between bins 50 and 51. The results for the rectangular, Hamming, and Hamming lag windows are shown in Figure 22. The Hamming lag window seemed to produce the better result here.

It should be noted that the very definition and interpretation of leakage is firmly rooted in the theory of spectral estimation. However, the definition of leakage is open to argument when alternative basis sets are being used.

Alternative Series Expansions

Several possibilities are suggested by the previous discussions. One is that if an asymptotic approximation

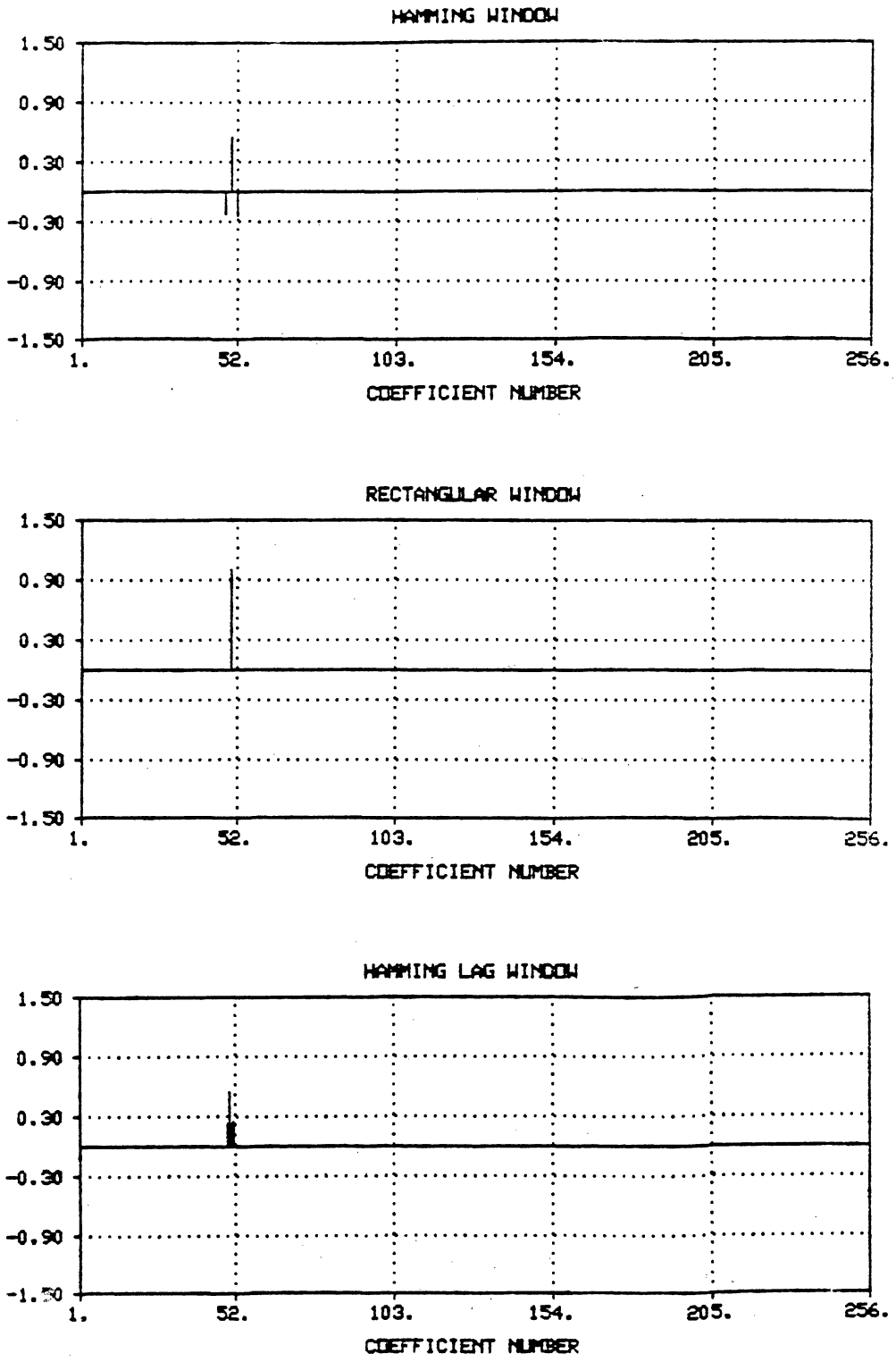


Figure 21. Comparison of windows, using function $f(x)=J_0(\lambda_{50}x)$. Top: Hamming window. Middle: Rectangular window. Bottom: Hamming lag window.

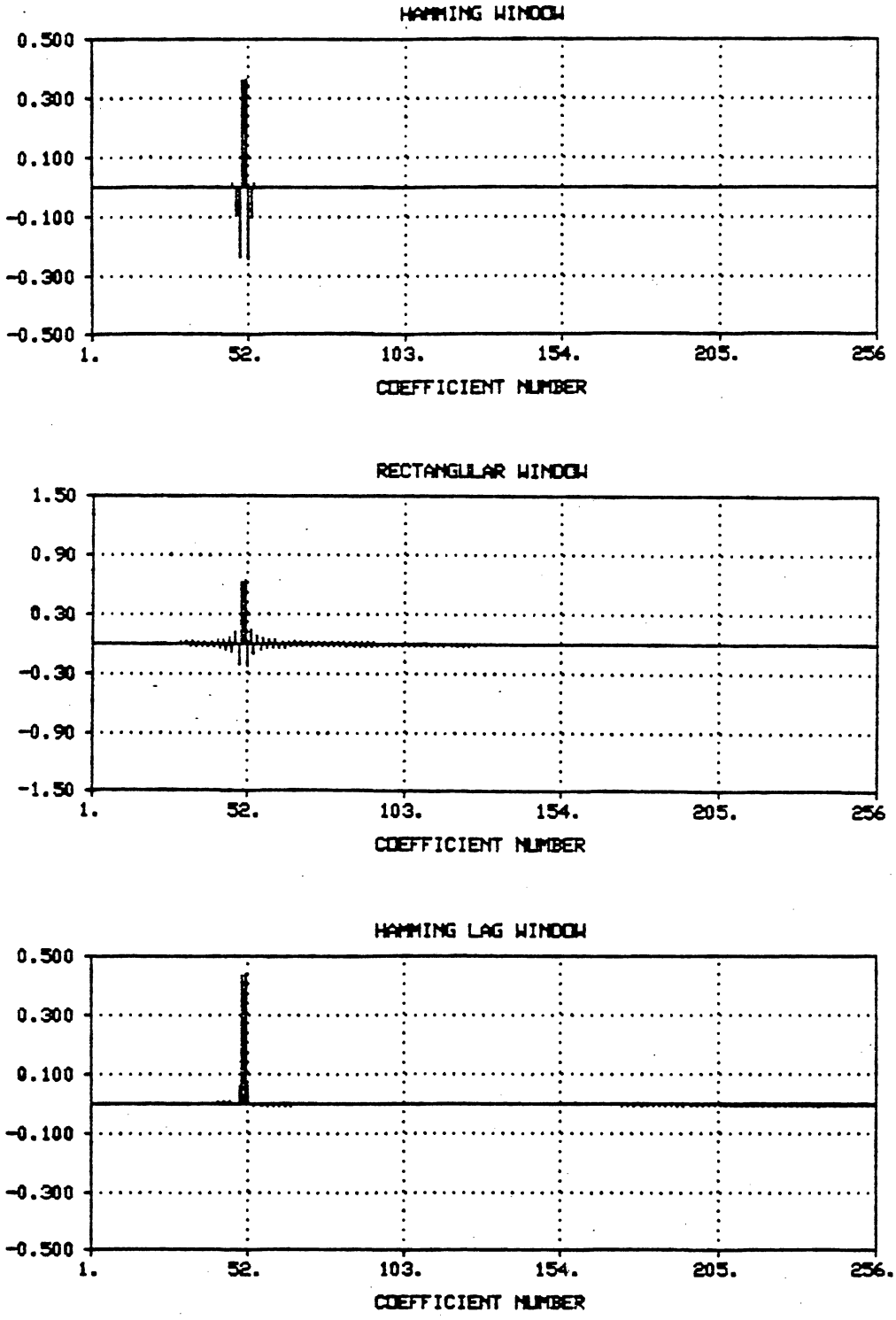


Figure 22. Comparison of windows, using function $f(x)=J_0[(\lambda_{s0}+\lambda_{s1})x/2]$. Top: Hamming window. Middle: Rectangular window. Bottom: Hamming lag window.

exists for the kernel of a transform, then an approximate method of computing that transform may be derived. This concept is quite general, and could be applied to many other types of series expansions besides the Bessel functions.

Another possibility is the representation of waveforms as linear combinations of functions which are similar to the Bessel functions. For example, the waveform could be written as a linear combination of sinc(.) functions

$$f(x) = \sum_{m=1}^{\infty} C_m \, \mathrm{sinc}(mx) = \sum_{m=1}^{\infty} C_m \frac{\sin(m\pi x)}{m\pi x} \qquad (4.31)$$

in the interval [0,1). The boundary conditions would be the same as those for the basis set $\{J_o(\lambda_m x)\}$.

From a table of integrals, it can be found that

$$\int_0^1 \sin(m\pi x)\sin(n\pi x)\, dx = \begin{cases} 0 \text{ for } m=0 \text{ or } n=0 \\ 0 \text{ for } m \neq n \\ 1/2 \text{ for } m=n \end{cases} \qquad (4.32)$$

Multiplying the Equation (4.31) on both sides by $2\pi x \sin(n\pi x)$ and integrating from 0 to 1 yields the coefficients of the sinc(.) series:

$$C_m = 2\pi m \int_0^1 x\, f(x)\, \sin(m\pi x)\, dx, \quad m=1,2,\ldots. \qquad (4.33)$$

This integral can be written in discrete approximation form as

$$C_m \approx \frac{2\pi m}{N^2} \sum_{n=0}^{N-1} n f\left(\frac{n}{N}\right) \sin\left(\frac{m\pi n}{N}\right) \qquad (4.34)$$

in the same manner as the rectangular integration in Equa-

tion (4.8). The summation could be performed by taking the imaginary part of the result of an FFT.

Another convenient expansion is

$$f(x) = \sum_{m=1}^{\infty} C_m J_{1/2}(\lambda_m x)$$

$$= \sum_{m=1}^{\infty} C_m \sqrt{\frac{2}{\pi(m\pi x)}} \sin(m\pi x) \qquad (4.35)$$

which is valid in the interval [0,1]. The integral expression for the coefficients is

$$C_m = \pi(2m)^{.5} \int_0^1 x^{.5} f(x) \sin(m\pi x)\, dx \qquad (4.36)$$

and may be suitably approximated by the same methods as before, using an FFT.

Chapter Summary

In conclusion, taking the viewpoint of rectangular-rule approximate integration and possibly using approximations to the transform kernels can lead to fast approximate algorithms for many kinds of series expansions. These approximations may be good enough for practical purposes in many cases. The general technique is to pose the desired integral as a summation whose kernel involves a function (such as sine and cosine) for which a fast algorithm is commonly known. This is a very practical technique, but seems to be little known because interest in alternative basis sets, in

general, is limited.

Three different methods for deriving the Fourier-Bessel series coefficients have been shown in this chapter. But a fast algorithm is of little value unless the results can be interpreted and used. Thus, Chapter V will attempt to interpret the transform results in terms of the traditional frequency domain, and will give a linear filtering explanation of the data obtained.

CHAPTER V

INTERPRETING THE RESULTS OF THE TRANSFORMATION

Introduction

The set of Fourier-Bessel coefficients must be given some theoretical or physical interpretation if at all possible. The result of a transformation is just a set of numbers unless logical or mathematical significance can be found for them. The purpose of this Chapter is to present some background that will enable a more intuitive insight into the result of the Fourier-Bessel transform.

FIR Filter Bank Approach

One of the reasons that the Bessel functions were chosen as the object of this investigation was that they are the solutions of a set of time-varying differential equations. At first thought, it might appear that the transformation results in a time-varying filter operation on the input data. But the actual way in which the transform is used is as a fixed FIR filter bank which acts on the input data vector (usually a windowed segment of speech).

As previously mentioned, such a linear transformation of the data can be represented as a matrix-vector multipli-

cation. Thus, each row of the transform matrix is an FIR filter. The rows do not change, but are fixed: the transformation is not time-varying. The frequency response of the FIR filter bank will now be discussed.

Frequency Response of the Filter Bank

The calculation of the frequency response of each row of the transform matrix turned out to be a very difficult problem indeed, for the following reasons. First, there is no convenient form of an addition theorem for Bessel functions. That is, there is no expression which formulates $J_o[(n+1)T]$ in terms of a finite number of past samples of the function $J_o[nT]$, where T represents a fixed time increment.

Second, the usual analysis of Fourier-Bessel series uses an integral expression for each coefficient:

$$C_m = \frac{2}{[J_1(\lambda_m)]^2} \int_0^1 x\, f(x)\, J_o(\lambda_m x)\, dx. \qquad (5.1)$$

But in a computer, only sampled signals and sampled Bessel functions can be processed. Hence, the integral can only be approximated. In fact, the fast algorithm for the approximation of the series coefficients does not represent an exact pseudoinverse of the Bessel function matrix, J, which will be defined later in this chapter.

Very little of the literature surveyed dealt with the case of sampled Bessel functions. Jerri (1978) defined a Discrete Hankel Transform, but he was not able to derive

an exact transformation: He used discrete approximations. J.P. Clero (1979) proposed a discrete Bessel transform matrix. Neither Jerri nor Clero could find mathematical proofs concerning the orthogonality of sampled Bessel functions on a finite interval.

Third, even if the fast-algorithm approximation to (5.1) is considered to be close enough (which it should be for many practical purposes) then there remains the problem of calculating

$$I(\omega,m) = \frac{2}{[J_1(\lambda_m)]^2} \int_0^1 x e^{j\omega x} w(x) J_o(\lambda_m x) dx \qquad (5.2)$$

for various values of ω, so that the frequency response can be determined. The function w(x) is a windowing function. Most of the mathematical articles surveyed containing references to integrals with Bessel functions in their integrands dealt only with the case of integration from zero to infinity, rather than with the case of finite limits.

For these reasons, the search for the closed form formula for the frequency response was (at least temporarily) abandoned. Instead, the integral in Equation 5.2 was estimated by numerical integration. The magnitude response of $I(\omega,m)$ was plotted for m=50 with w(x)=1 (rectangular window). Figure 23 shows the resultant response. The large side lobes are due to the use of the rectangular window. For the same value of m, Figure 24 shows the effect of using a Hamming window in (5.2). Note that the sidelobes are reduced,

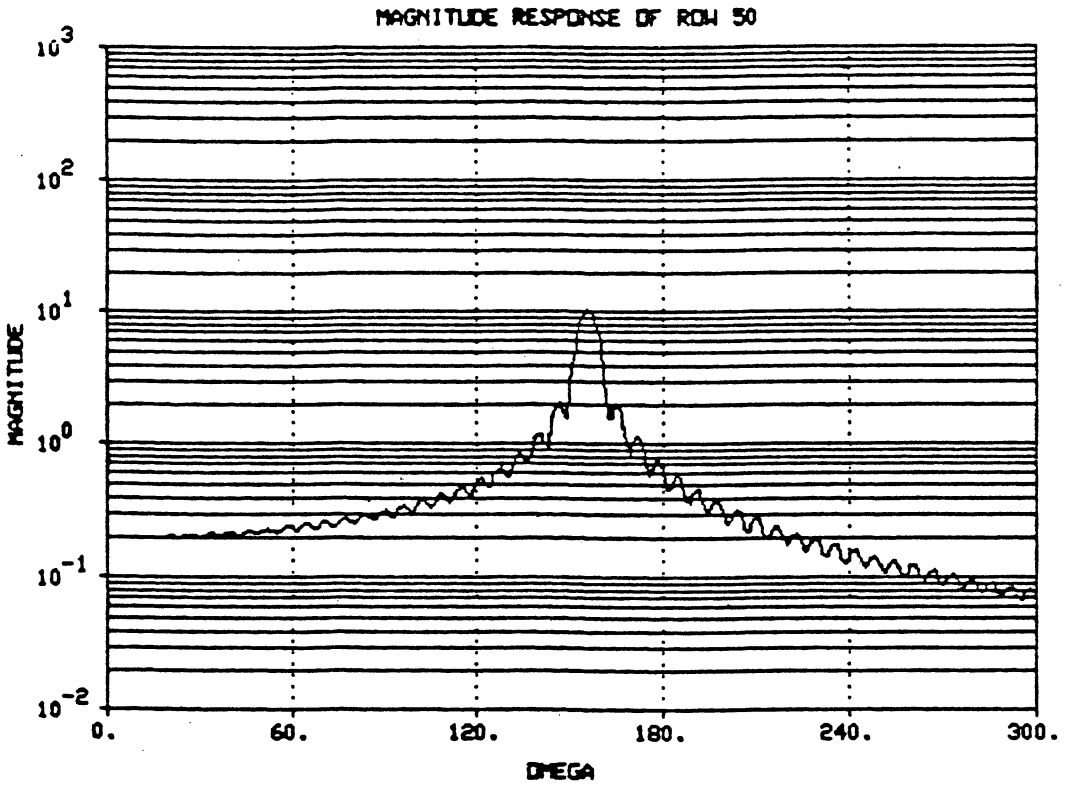


Figure 23. Magnitude response of FIR filter bank, rectangular window. Row 50 out of 256.

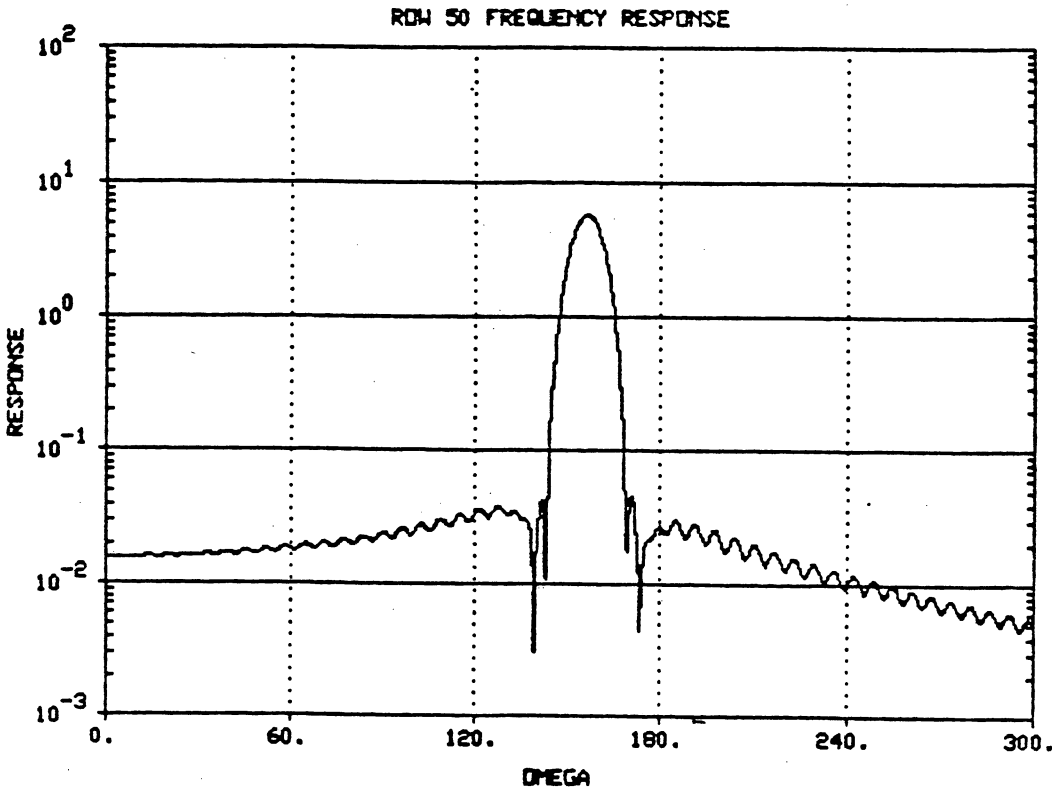


Figure 24. Magnitude response of FIR filter bank, traditional Hamming window. Row 50 out of 256.

as expected since the Bessel functions are really quite similar to the trigonometric functions. Another window function consisting of half of a traditional Hamming window was tried (Figure 19). The magnitude response of the integral transformation when this window is used is shown by Figure 25. This type of window is frequently used in the field of time series analysis for processing of autocorrelation functions. This window will be discussed again in a later section of this report.

Correspondence to Traditional Frequency Domain

A correspondence can be derived between the coefficient numbers in the Fourier-Bessel transform domain and frequency in the traditional sine-cosine expansion. The correspondence between coefficient number and traditional frequency is only approximate because each Bessel function in the series represents a whole band of frequencies, not just a single frequency. But there is a distinct concentration of energy in the frequency domain, as shown in Figures 23-25. Figures 15-18 show the sinusoidal nature of the FIR filter's impulse response, so it is really no surprise that the response of each row of the filter bank is a narrow band of frequencies.

The asymptotic approximation in Equation (3.14) is stated in terms of the normalized time, x. Making the substitution t/T for x, and substituting the expression (3.15) for λ_m, the approximation to $xJ_o(\lambda_m x)$ becomes

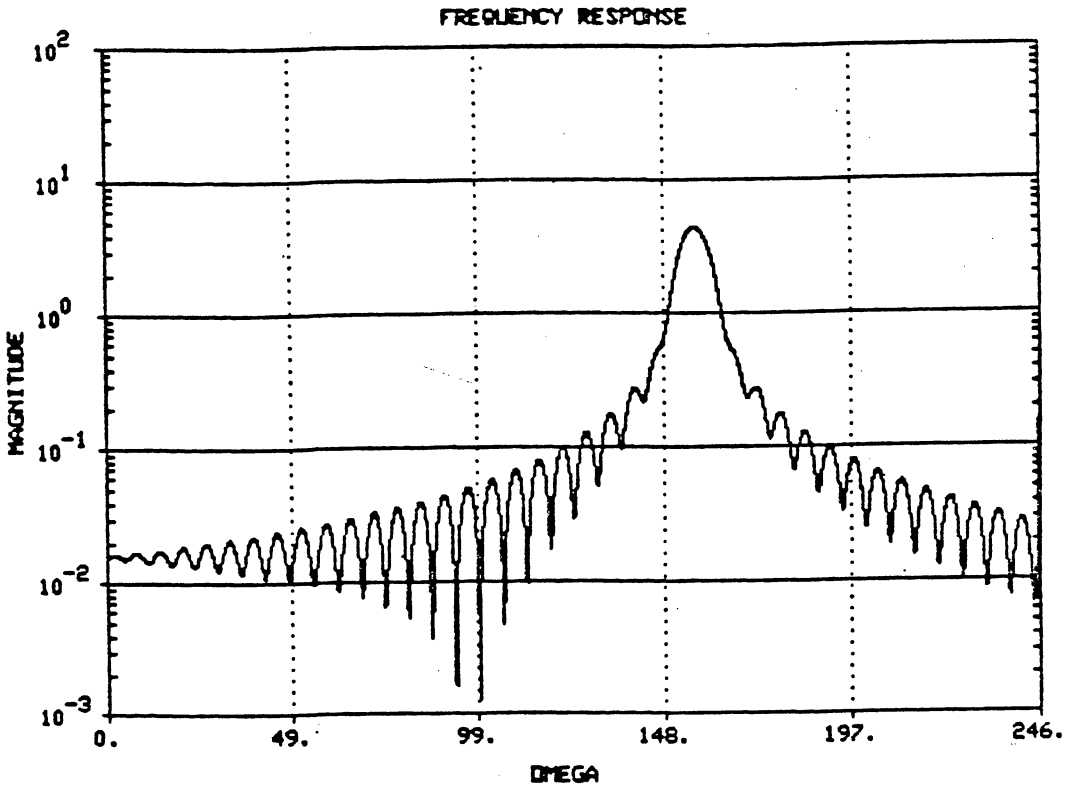


Figure 25. Magnitude response of FIR filter bank, Hamming lag window. Row 50 out of 256.

$$xJ_0(\lambda_m x) \approx \frac{1}{\pi}\left[\frac{2t}{(m-.25)T}\right]^{.5}\cos\left[\frac{(m-.25)\pi}{T}t - \frac{\pi}{4}\right]. \quad (5.3)$$

The major portion of the energy is centered at

$$\omega \approx \frac{(m-.25)\pi}{T} \quad \frac{rad}{sec} \quad (5.4)$$

where T is the length of the analysis frame, in seconds. If there are N sample points per analysis frame, and the sampling rate is f_S samples per second, then the relationship between coefficient number and frequency is

$$f \approx \frac{m - .25}{2N} f_S \quad Hz. \quad (5.5)$$

But there is a problem here: Suppose that the signal being analyzed is a sinusoid, but with a phase which makes it as orthogonal as possible to the filter's impulse response. Then the Fourier-Bessel coefficient whose approximate frequency falls nearest to the signal's true frequency may show very little response. Instead, neighboring coefficients will show a greater response. This is one of the reasons that the Fourier-Bessel series is hard to deal with: The results one obtains depend heavily upon where the analysis frame starts. The transformation is shift-variant in this sense.

Linear Modeling Interpretation

A useful viewpoint of alternative basis sets is the linear modeling interpretation (Menke, 1984). The assumed model of a waveform is as a linear combination of basis vectors, and the full power of matrix algebra can be brought to

bear on the problem.

Let J be a matrix whose columns are sampled Bessel functions. The general element of this matrix is given by

$$J_{nm} = J_0(\lambda_m \frac{n}{M}), \quad 0 \leq n \leq N-1, \quad 1 \leq m \leq M \tag{5.6}$$

where m is the column number and n is the row number. λ_m represents the m-th zero of the Bessel function used, and N is the number of rows in the matrix. The rows are numbered from 0 to N-1. The data model is as a linear combination of the columns of this matrix:

$$\begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix} = C_1 \begin{bmatrix} J_0(\lambda_1 \frac{0}{N}) \\ J_0(\lambda_1 \frac{1}{N}) \\ \vdots \\ J_0(\lambda_1 \frac{N-1}{N}) \end{bmatrix} + C_2 \begin{bmatrix} J_0(\lambda_2 \frac{0}{N}) \\ J_0(\lambda_2 \frac{1}{N}) \\ \vdots \\ J_0(\lambda_2 \frac{N-1}{N}) \end{bmatrix}$$

$$+ \ . \ . \ . + C_M \begin{bmatrix} J_0(\lambda_M \frac{0}{N}) \\ J_0(\lambda_M \frac{1}{N}) \\ \vdots \\ J_0(\lambda_M \frac{N-1}{N}) \end{bmatrix} \tag{5.7}$$

or, in matrix notation,

$$\underline{x} = J\underline{c}. \tag{5.8}$$

Note that J need not be a square matrix.

The process of obtaining the model parameters, the C_m's, is equivalent to finding the best L_2 fit to the data given the aforementioned data model. If J is a square matrix, then the model parameters may be found (in principle) by inversion of the J matrix, followed by multiplication of the data vector $\underline{x}$ by the inverse matrix, J^{-1}. If J has more rows than

columns, then the best L_2 fit can be found by calculating the pseudoinverse of J, and then multiplying the data vector by this matrix:

$$(J^T J)^{-1} J^T \underline{x} = \underline{c}. \tag{5.9}$$

In practice, this process can be well approximated by the fast algorithm discussed in the previous sections.

When viewed as a multiplication of the data vector by the pseudoinverse matrix of J, the process of finding the C_m's can be seen to be equivalent to a filter bank consisting of FIR filters: Each row of the pseudoinverse is then the impulse response of an FIR filter. The rows of the pseudoinverse are not orthogonal, because Bessel functions are not a truly orthogonal basis set.

If the pseudoinverse matrix is post-multiplied by its own transpose then the result is

$$(J^T J)^{-1} J^T [(J^T J)^{-1} J^T]^T = [(J^T J)^{-1}]^T. \tag{5.10}$$

Since $J^T J$ is symmetric, its inverse is also symmetric. Therefore,

$$[(J^T J)^{-1}]^T = (J^T J)^{-1}. \tag{5.11}$$

The matrix $J^T J$ is not diagonal, because the columns of J are not orthogonal. The inverse of a non-diagonal matrix is also nondiagonal, if it exists. Ergo, $(J^T J)^{-1}$ is not diagonal, and diagonality of this matrix would have been a necessary condition for the orthogonality of the rows of the pseudoinverse, $(J^T J)^{-1} J^T$.

A consequence of this is that correlation of the re-

sulting model parameters is to be expected even when the covariance matrix of the input data vectors is a diagonal matrix. Consider the case where the input data is a rectangular-windowed segment of sampled data from a white noise process with unity variance. Then the covariance matrix is the identity matrix. If a matrix transformation A is applied to each outcome vector $\underline{x}$, $\underline{c} = A\underline{x}$, then the covariance matrix of the resultant vector $\underline{c}$ is

$$C_{cc} = AC_{xx}A^T = AIA^T = AA^T. \qquad (5.11)$$

Substitution of the pseudoinverse of J into the place of A in Equation (5.11) yields

$$C_{cc} = [(J^TJ)^{-1}J^T][(J^TJ)^{-1}J^T]^T = (J^TJ)^{-1}. \qquad (5.12)$$

The matrix (J^TJ) is not diagonal. Therefore, its inverse is not diagonal: The model parameters in the vector $\underline{c}$ are correlated even when the input is a white noise sequence. The conclusion here is that correlation of the transform domain coefficients is to be expected even when the input sequence is white, provided that the basis set is not orthogonal.

Relationship to the Discrete Fourier Transform

The relationship between the Fourier-Bessel coefficients and the Discrete Fourier Transform (DFT) coefficients is actually quite simple to derive. It is well known that the DFT can be written as a matrix-vector multiplication

(Hershey and Yarlagadda, 1986):

$$A_{DFT}\,\underline{x} = \underline{y} \tag{5.13}$$

where $\underline{x}$ is the data vector, or sampled signal, and $\underline{y}$ is the vector of DFT coefficients. The matrix A_{DFT} consists of complex elements of the form

$$[A_{DFT}]_{kn} = e^{-j\frac{2\pi}{N}nk} \tag{5.14}$$

where the rows are numbered as k=0,1,...,N-1 and the columns are numbered as n=0,1,...,N-1. Note that the matrix $\frac{1}{\sqrt{N}} A_{DFT}$ is symmetric and also unitary.

Now let a linear model be assumed for the sampled data:

$$J\underline{c} = \underline{x}. \tag{5.15}$$

This is the same model as was presented by Equation (5.7). Premultiply both sides of Equation (5.15) by A_{DFT} to get:

$$A_{DFT}J\underline{c} = A_{DFT}\underline{x} = \underline{y} = \text{DFT coefficients.} \tag{5.16}$$

The columns of matrix $A_{DFT}J$ are the DFT's of the basis vectors of the assumed data model. Therefore, if the DFT's of basis vectors are known then these vectors can be summed with weights equal to the model coefficients to get the DFT of the original data vector, $\underline{x}$. Unfortunately, the DFT's of the model matrix's columns are difficult, or impossible, to express in a closed form (no proof has yet been found).

Chapter Summary

A linear filtering approach to interpretation of the Fourier-Bessel coefficients has been presented. The process of discrete Fourier-Bessel transformation can be modeled as

a filter bank consisting of FIR filters. Some of the difficult problems associated with discrete Fourier-Bessel analysis are: (1) Choice of windows, with the Hamming lag window seeming to be good; (2) Lack of an analytic expression for the frequency response of each row of the FIR filter model; and (3) lack of orthogonality of the basis set, making the Fourier-Bessel coefficients statistically correlated even when the input signal is uncorrelated.

The difficulty of Fourier-Bessel interpretation makes the chore of analyzing real speech signals even more problematic. For example, the point at which an analysis frame starts affects the coefficients obtained. This, and other, effects will be discussed in the context of real speech signal analysis in Chapter VI.

CHAPTER VI

APPLICATION OF FOURIER-BESSEL SERIES TO SPEECH ANALYSIS

Introduction and Survey of Applications

Classical Applications

The traditional uses of Fourier-Bessel series are related to problems in mathematical physics where circular symmetry and boundary conditions prevail. Examples are vibrating drum heads, circular waveguides, and heat conduction in cylindrical rods (Condon and Odishaw, 1967). It should also be noted that acoustic tube models of the vocal tract are cylindrical. But the literature search revealed few published reports of the use of Fourier-Bessel expansions for signal analysis or classification purposes.

Analysis-Synthesis of Speech

Fourier-Bessel series have recently been used as the basis for analysis-synthesis of speech (Chen, Gopalan, and Mitra, 1985) and a Fourier-Bessel vocoder has been built (Chang and Chen, 1986). The basic scheme used was to expand each frame of speech into a Fourier-Bessel series based upon

the Bessel function $J_1(x)$. Only dominant coefficients were retained, and the speech was resynthesized from these. Chen reported that good quality speech can be obtained using only one half to one third of the available coefficients. Understandable speech could be obtained with one tenth of the coefficients. Chen typically used 150 coefficients in the expansion because he could find 150 zeros of $J_1(x)$ in a table. But it was shown in Chapter III of this thesis that the higher-ordered zeros of $J_n(x)$ can be well approximated by Equation (3.16), so there is really no need to limit the model order to 150 or fewer.

The criterion Chen used for selection of the coefficients to be retained was that the several absolutely largest numbers in the set were kept. But note from Figure 13 that the higher λ_m becomes, the less energy is in $J_0(\lambda_m x)$. Therefore, it is proposed that the criterion for coefficient selection should not be a flat threshold, but rather a threshold function that allows the several coefficients which contribute the most energy to the signal to be kept.

Feline Cortical Potentials

P.L. Nunez published an article which described the use of the Fourier-Bessel series coefficients for characterization of the cortical evoked potential of cats due to an olfactory stimulus (1973). The rationale behind Nunez' choice of the Bessel functions as a basis set was that they "look like" the waveforms he was trying to represent. Mean-

ingful waveform representation and data compression were his goals. He conjectured that since basis functions which are the solutions of ordinary differential equations were not really suitable to describe the waveforms, then a better choice of basis could be a set of functions which are the wavelike solutions of some sort of partial differential equations. But he did not attempt to theoretically justify this proposition. In fact, his experimental results did not conclusively support the claim that the Bessel functions were fundamentally better than the traditional trigonometric functions as a basis set for the representaton of signals. But there were some promising indications that the Bessel functions could be a better basis set when an approximation to the waveform was to be formed by only a few terms of a series.

This is an important distinction: a basis set for a series representation can be judged either by its ability to (1) accurately converge to the waveform when, in the limit, many terms of the series are included, or (2) represent the waveform, well enough for a given purpose, with a only few terms of the series included in the sum. In this latter case convergence of the series is not really a requirement, for all that is desired is a description of some properties of the waveform, not an extremely accurate reproduction of the waveform. The mathematical literature seems to be preoccupied with questions of convergence and accurate representation, whereas a (possibly imperfect) description of a signal

using only a few parameters could be more desirable for pattern recognition or coding purposes.

Characteristics of Fourier-Bessel Coefficients for Typical Speech Signals

The first of several versions of the Fourier-Bessel analysis program used the method which was described by Equation (4.19). Note that the subsequent algorithm is quite inefficient: Only one out of every four frequency bins was used, so the transform had to be very long (of length 8N). The program had analysis frames which were overlapped and Hamming windowed, but which were not in any way synchronized to the speech signal's own pitch period: the frames were "free running". Some of the observations of the results are reported in the following paragraphs.

When plotted, the Fourier-Bessel coefficients tend to alternate in sign. Refer to Figure 26. Note that each large positive coefficient is usually followed or preceded by a large negative coefficient, and vice-versa. This is the most obvious characteristic of the coefficients, and is true for both voiced and unvoiced speech for all speakers.

The coefficients on either side of a given coefficient show a great deal of correlation with that coefficient. This is true especially for voiced speech. The amount of correlation is less for coefficients which are not near neighbors. Recalling Equation (5.12), this is hardly surprising since

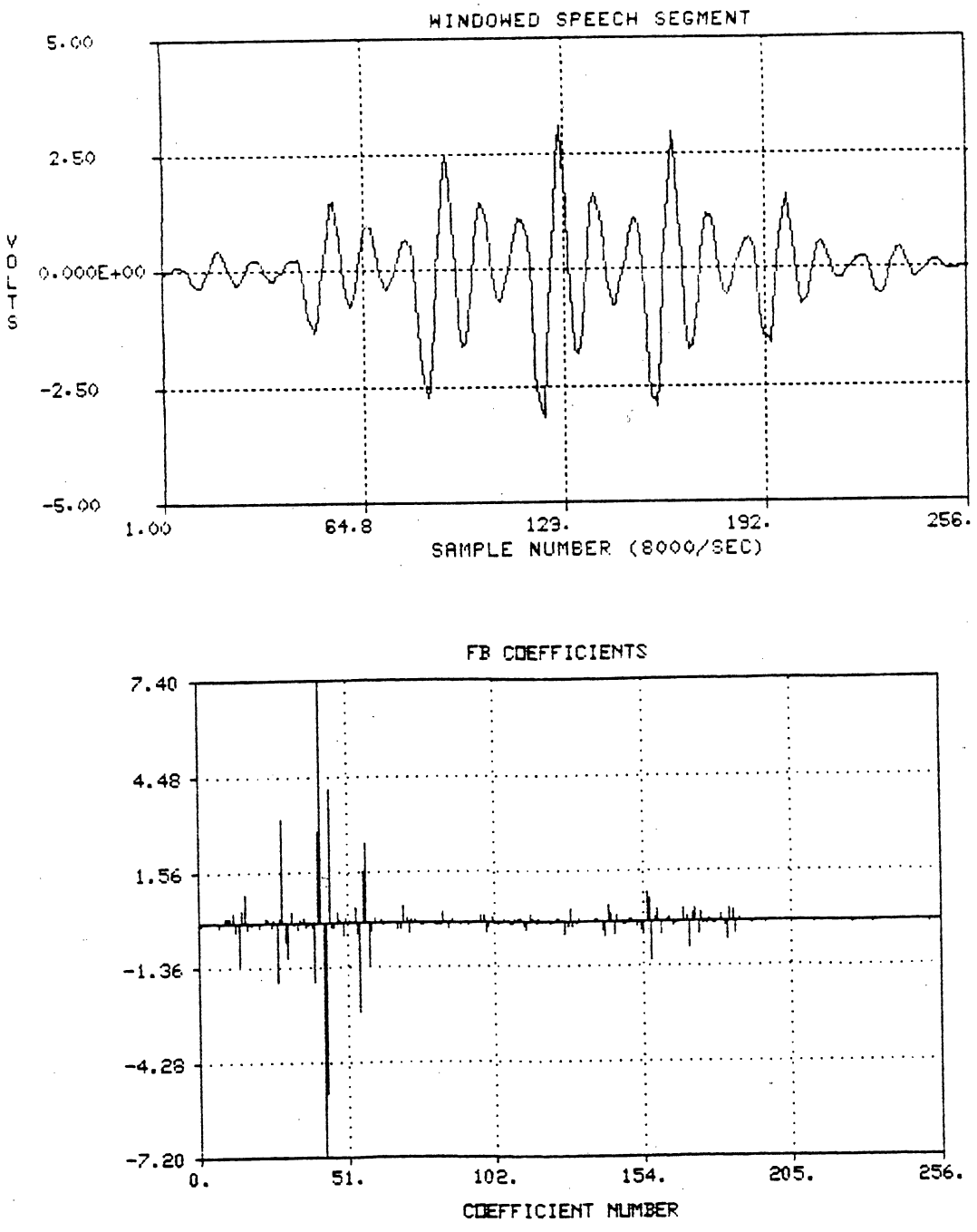


Figure 26. Windowed speech and Fourier-Bessel coefficient, Speaker 1, Trial 1. Phoneme: "a" as in "cats". Hamming window used.

some correlation of coefficients is expected.

In Figure 27, the same speech segment was again analyzed, but with a rectangular data window. Note the large alternating coefficients. This effect is very much like that which appears when dealing with the Discrete Fourier Transform. This is not surprising, since the rows of the filter bank are very nearly sinusoidal. A tapered window such as the Hamming window is definitely recommended.

The Bessel functions are shift-variant. Since the analysis frames were free-running, their starting points in a pitch period were random variables. Thus, the Fourier-Bessel representation varied quite a bit from frame to frame, even in the same phoneme. Figure 28 shows that the shift-variant property of the Bessel functions affects the coefficients obtained in neighboring analysis frames. The starting point of the 256-sample analysis frame in Figure 28 was only half a pitch period after the starting point of the analysis frame of Figure 26. Note that the general location of the large coefficients is still the same, but the exact coefficient numbers of the largest coefficients changed. Thus, the raw Fourier-Bessel coefficients are in need of further processing if reliable information is to be extracted from them.

The shift-variant effect just noted was partially remedied by synchronizing the beginning of each analysis frame with the same point in each pitch period. This resulted in more consistent Fourier-Bessel coefficients from frame to

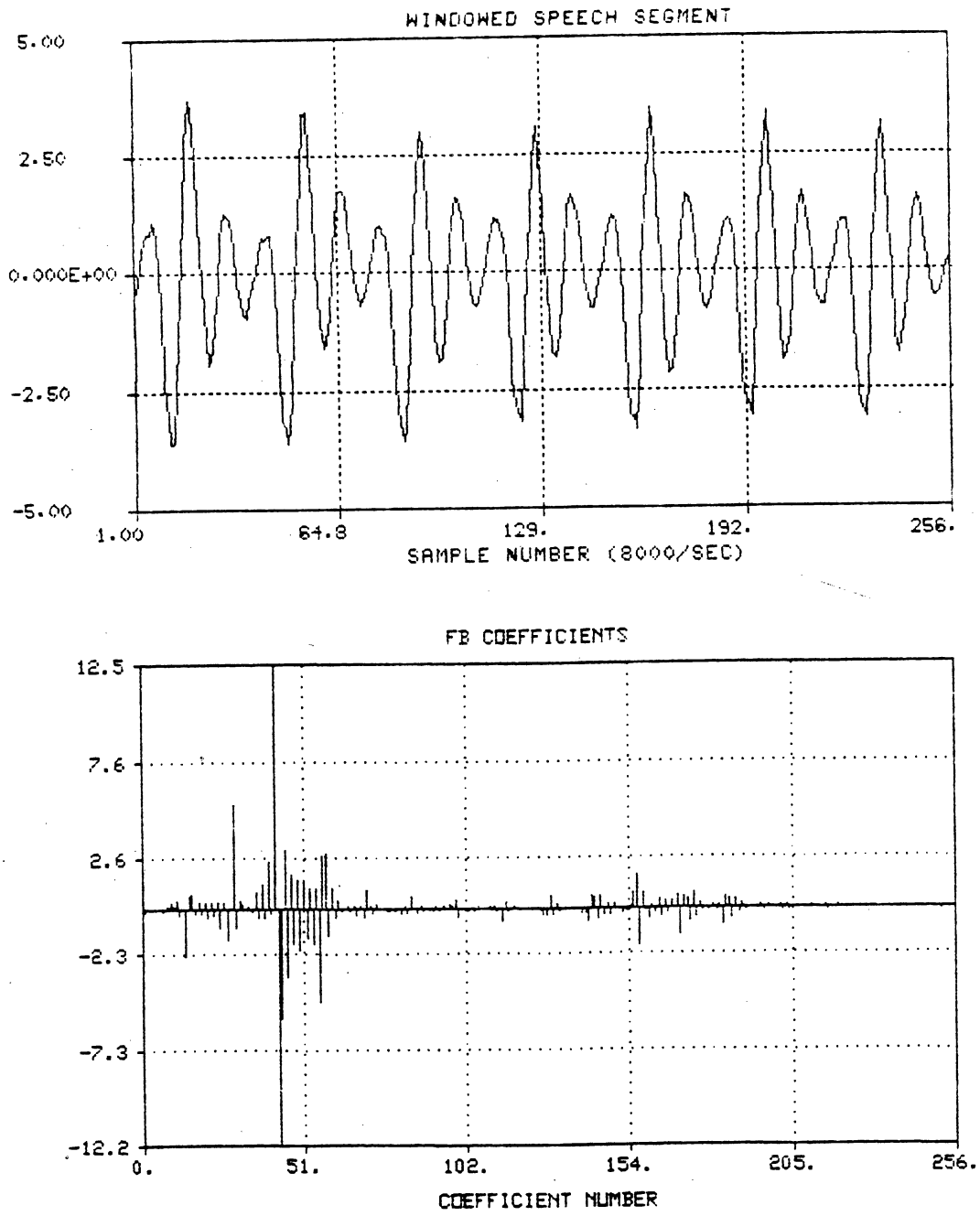
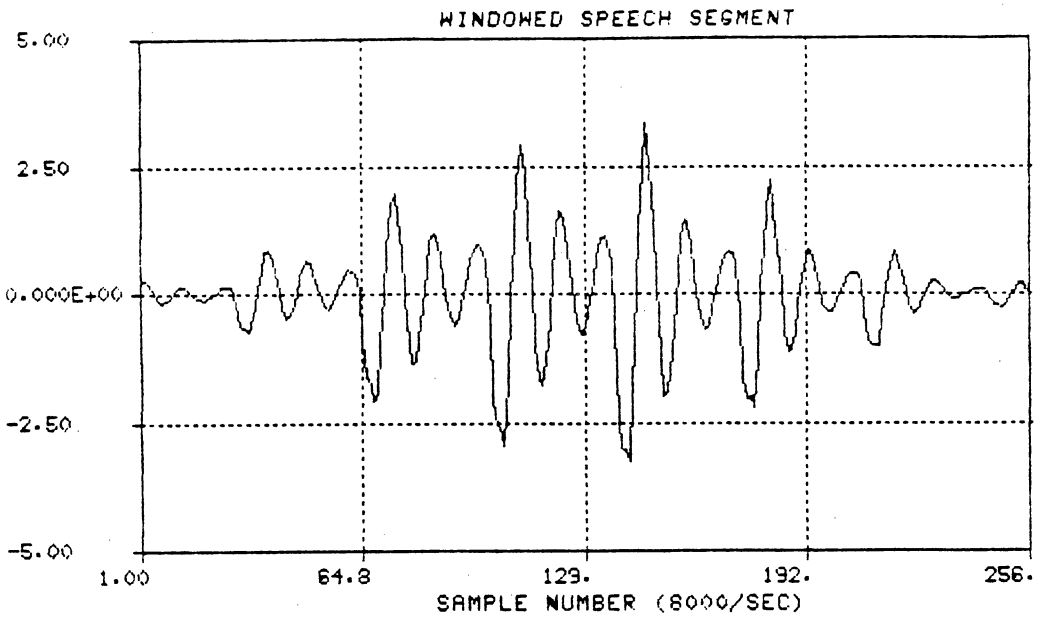


Figure 27. Windowed speech and Fourier-Bessel coefficients, rectangluar window.Speaker 1, Trial 1. Phoneme: "a" as in "cats".



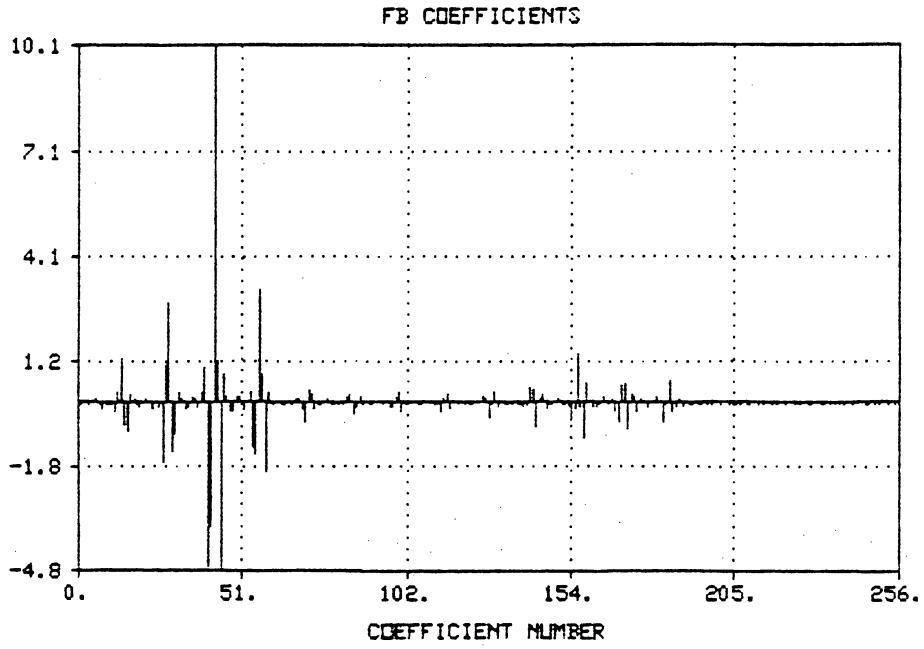


Figure 28. Windowed and time shifted speech and Fourier-Bessel coefficients. Speaker 1, Trial 1. Phoneme: "a" as in "cats". Hamming window used. Time shift of about half a pitch period.

frame within the same phoneme. Figure 29 shows the windowed speech segment and FB coefficients obtained when the beginning of the analysis frame was exactly one pitch period after that in Figure 26. Note that the Fourier-Bessel coefficients in Figures 26 and 29 are more consistent with each other than those shown in Figure 28.

Two other analysis frames, each with a different voiced phoneme, are shown in Figures 30 and 31. The general location of the largest coefficients changes from phoneme to phoneme. Since there is a correspondence between coefficient number and frequency, it is seen that these regions are the formants of the speech spectrum.

Comparison of the Fourier-Bessel coefficient set for the same phoneme ("a" as in "cats") is made in Figures 26 and 32-35. Figure 32 shows the result for a second trial of the same word for speaker number 1. Comparison to Figures 33 and 34 for speaker number 2 does not reveal any immediately obvious features which could be used to differentiate between speakers. Figure 35 shows the same phoneme for speaker four, a male speaker (speakers 1 and 2 were female). The major difference is that the spacing between the large coefficients is less for the lower-pitched male speaker than for the female speakers.

The Fourier-Bessel coefficients have less regularity and more high-frequency energy for unvoiced speech than for voiced speech. Figure 36 shows the result of expanding an "h" sound into a Fourier-Bessel series. Once again, this re-

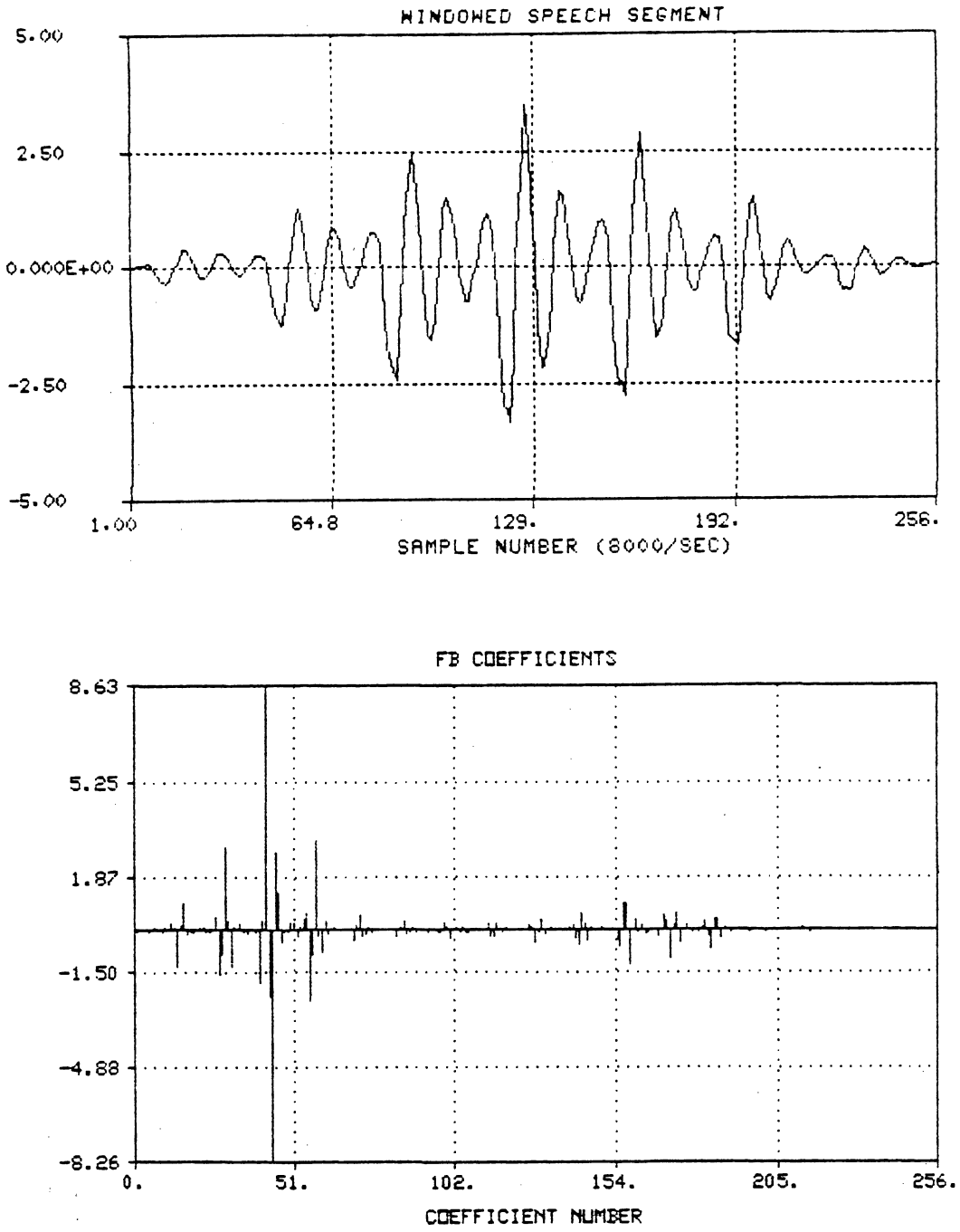


Figure 29. Shifted speech segment, with Fourier-Bessel coefficients. Speaker 1, Trial 1. Shifted one pitch period with respect to Figure 24.

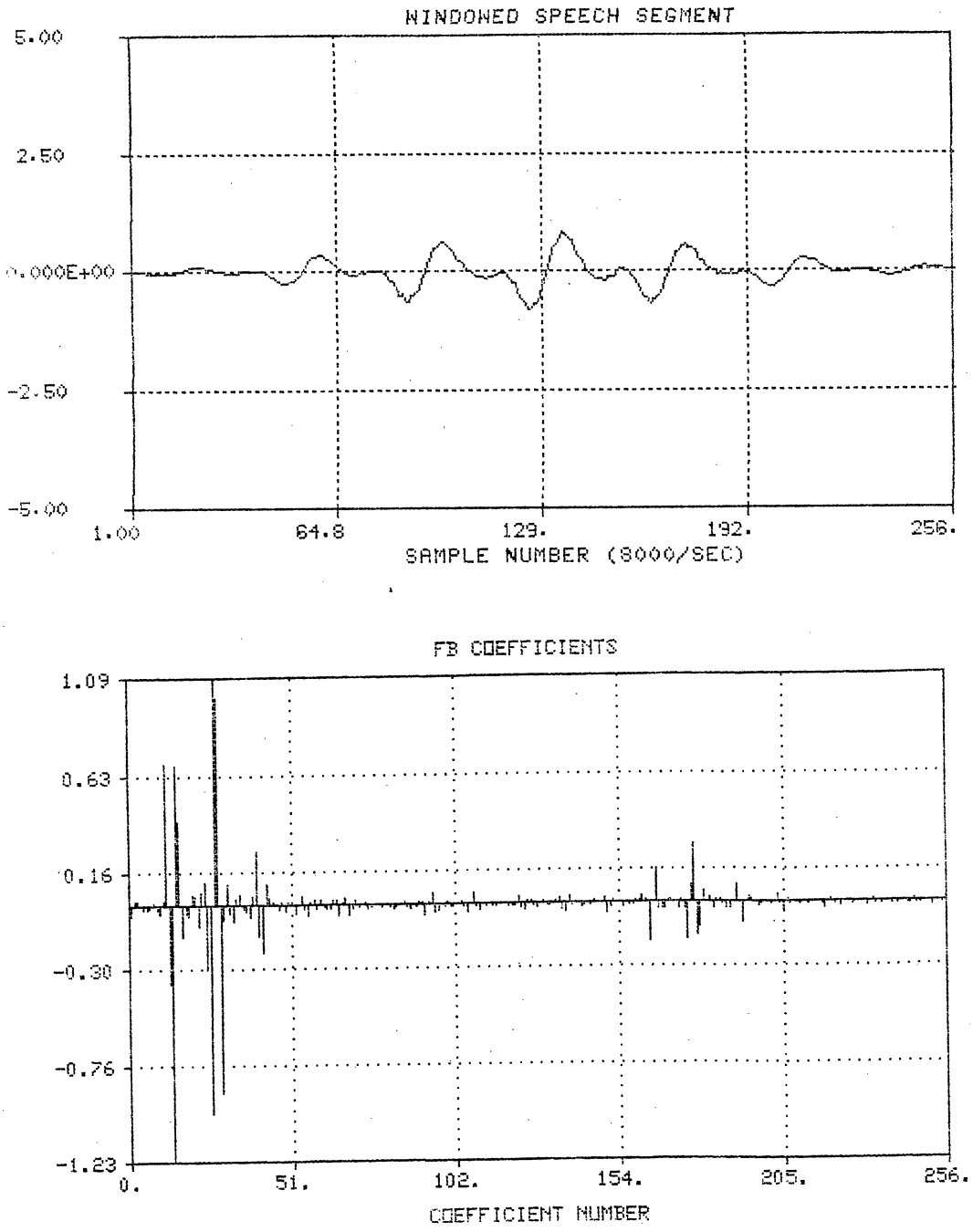


Figure 30. Speech segment,"e" as in "each", with Fourier-Bessel coefficients. Hamming window used.

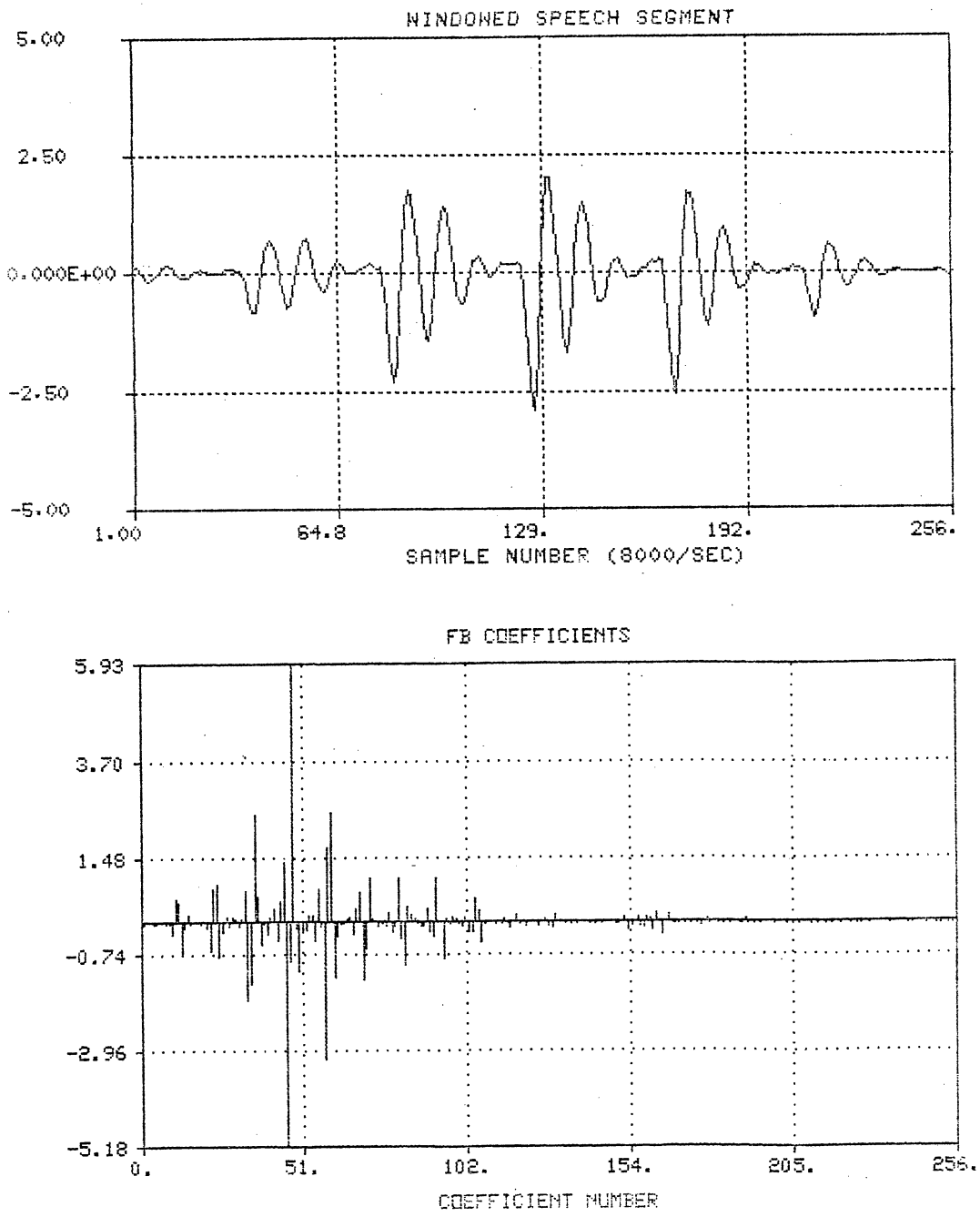
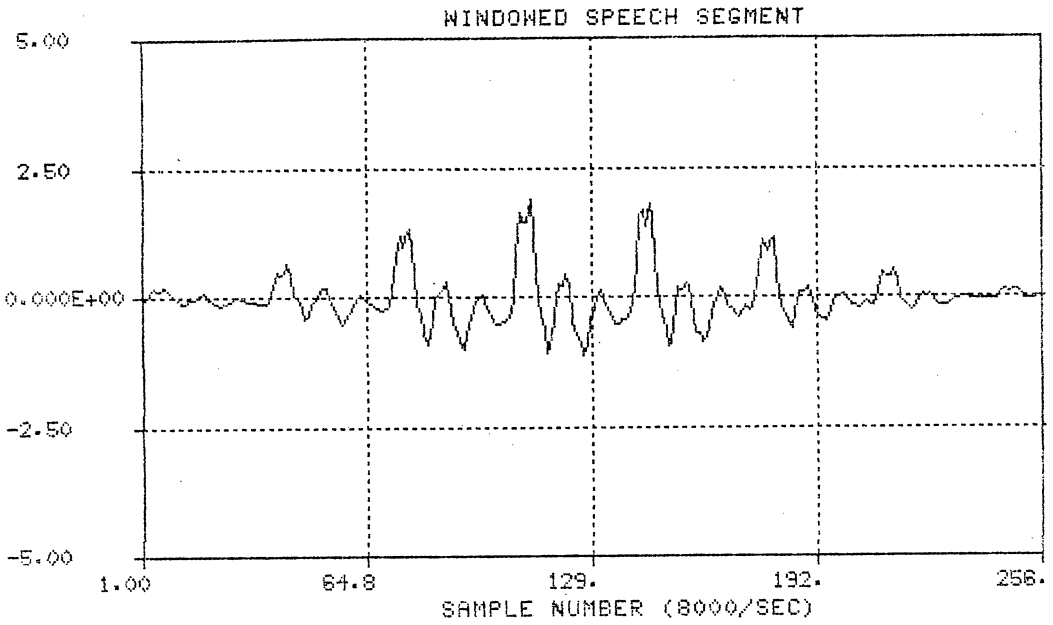


Figure 31. Speech segment,"o" as in "dogs", with Fourier-Bessel coefficients. Speaker 1. Hamming window used.



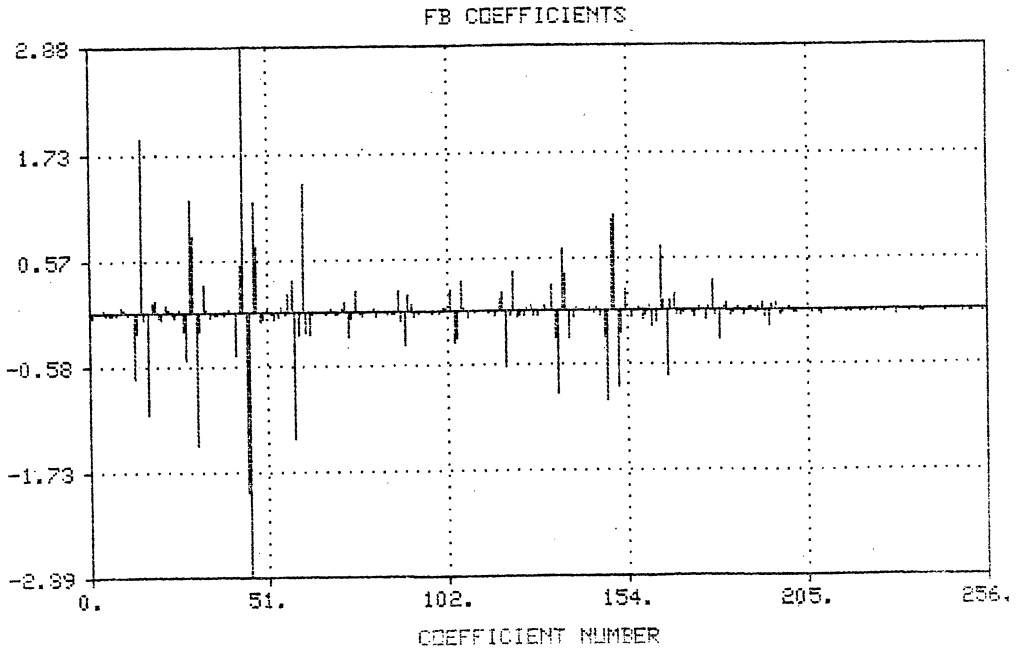


Figure 32. Speech segment,"a" as in "cats", with Fourier-Bessel coefficients. Speaker 1, Trial 2. Hamming window used.

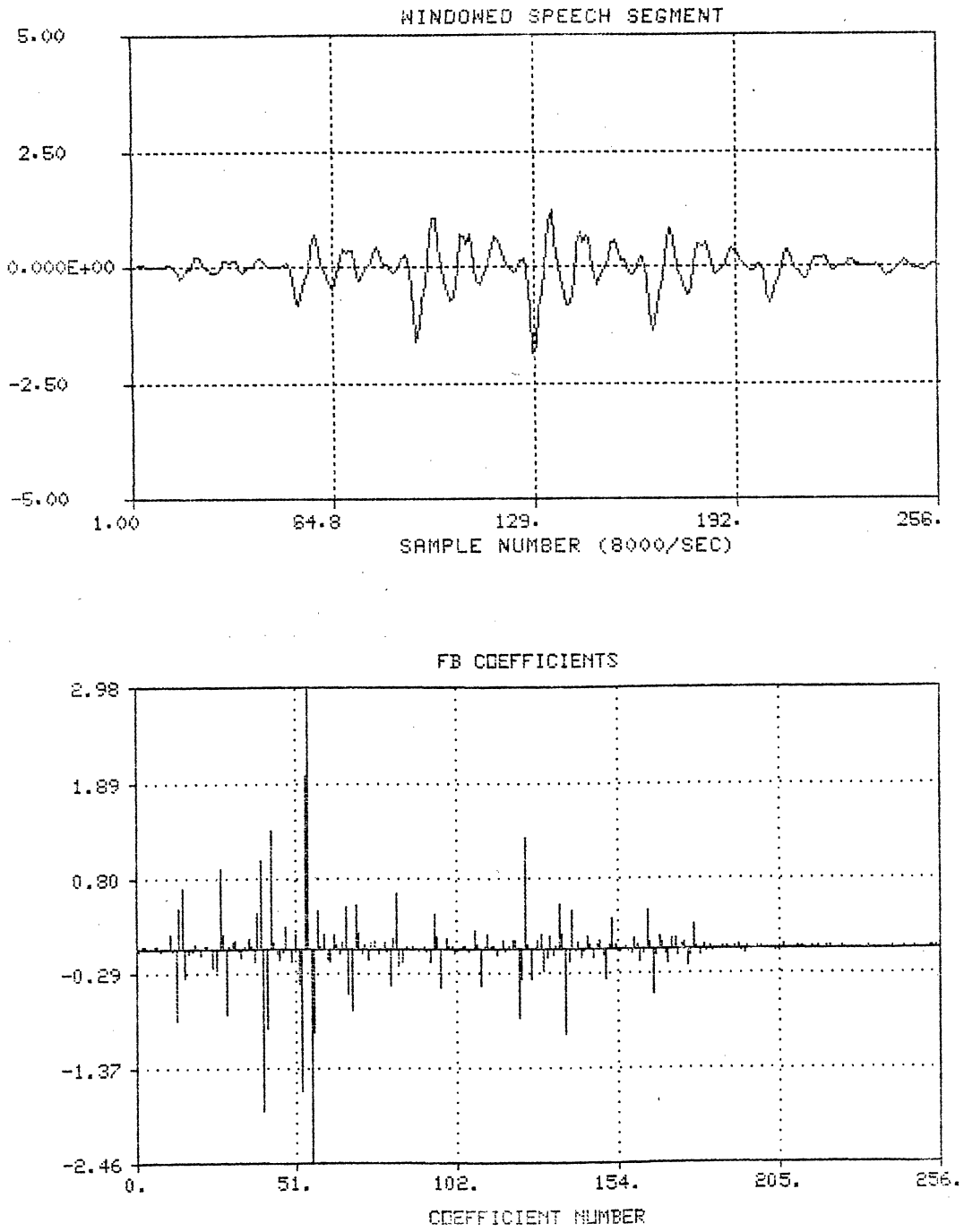
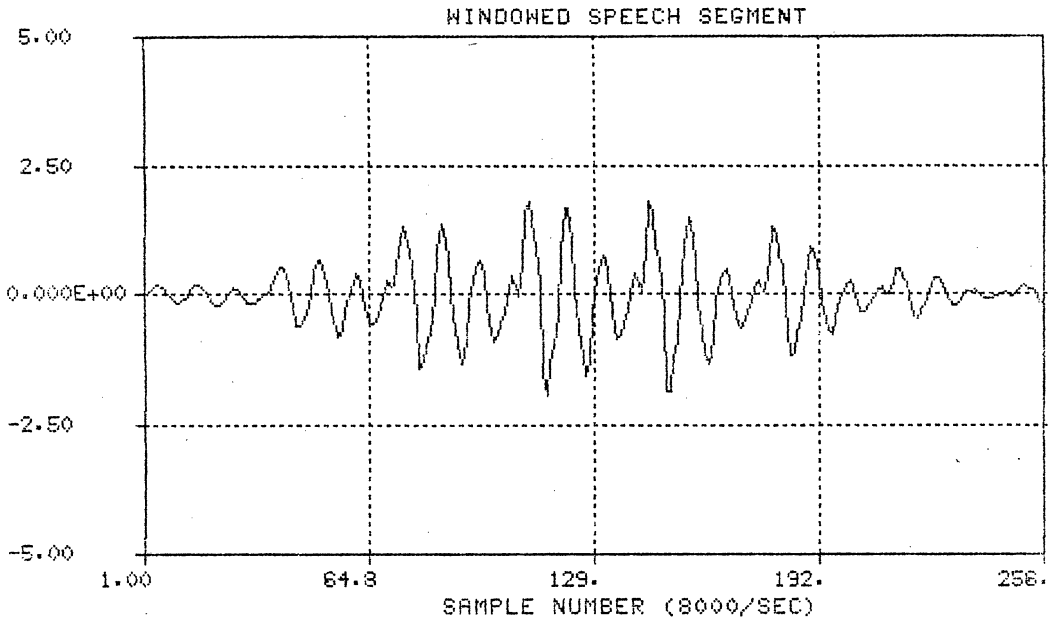


Figure 33. Speech segment,"a" as in "cats", with Fourier-Bessel coefficients. Speaker 2, Trial 1. Hamming window used.



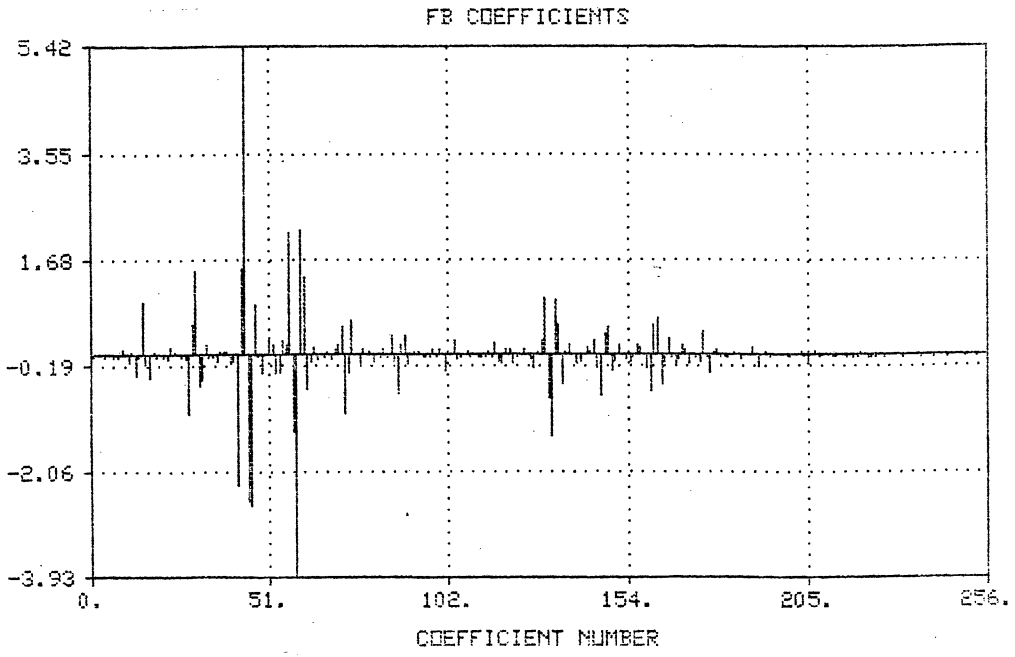
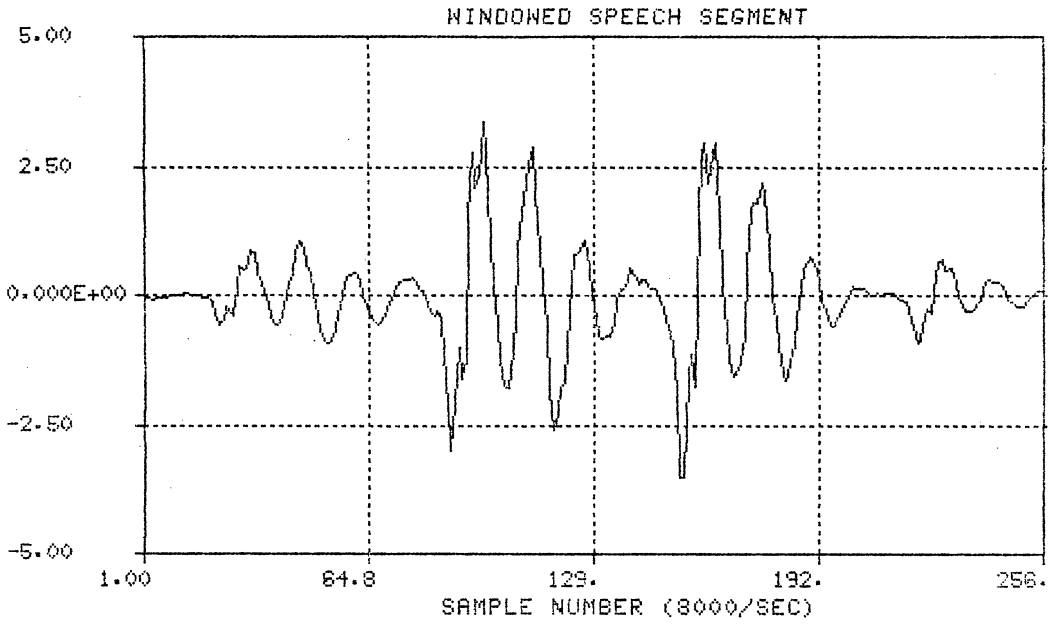


Figure 34. Speech segment,"a" as in "cats", with Fourier-Bessel coefficients. Speaker 2, Trial 2. Hamming window used.



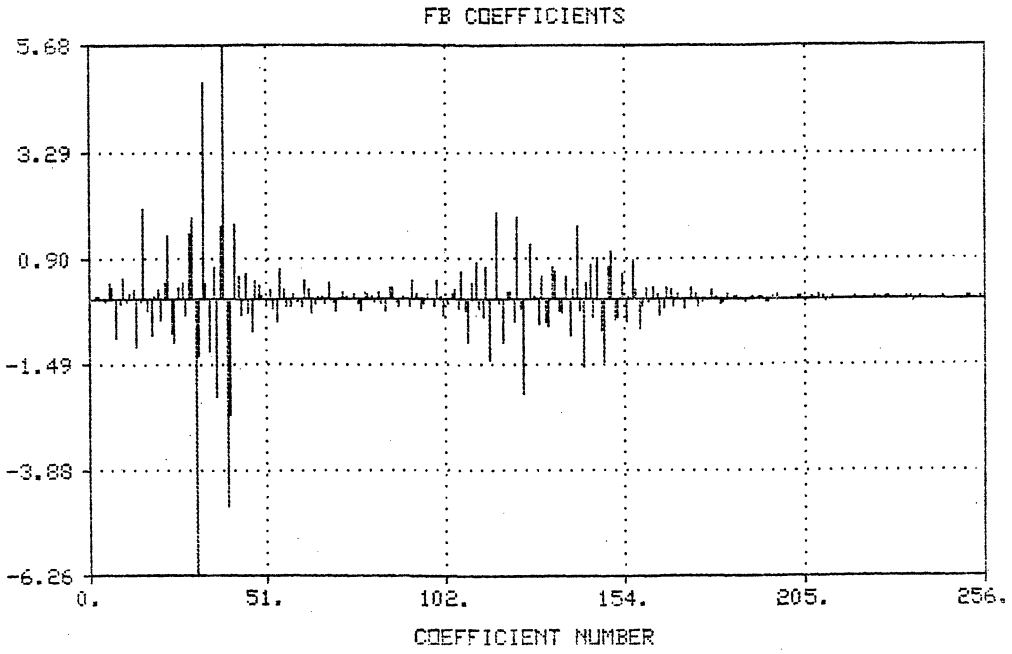


Figure 35. Speech segment,"a" as in "cats", with Fourier-Bessel coefficients. Speaker 4, Trial 1. Hamming window used.

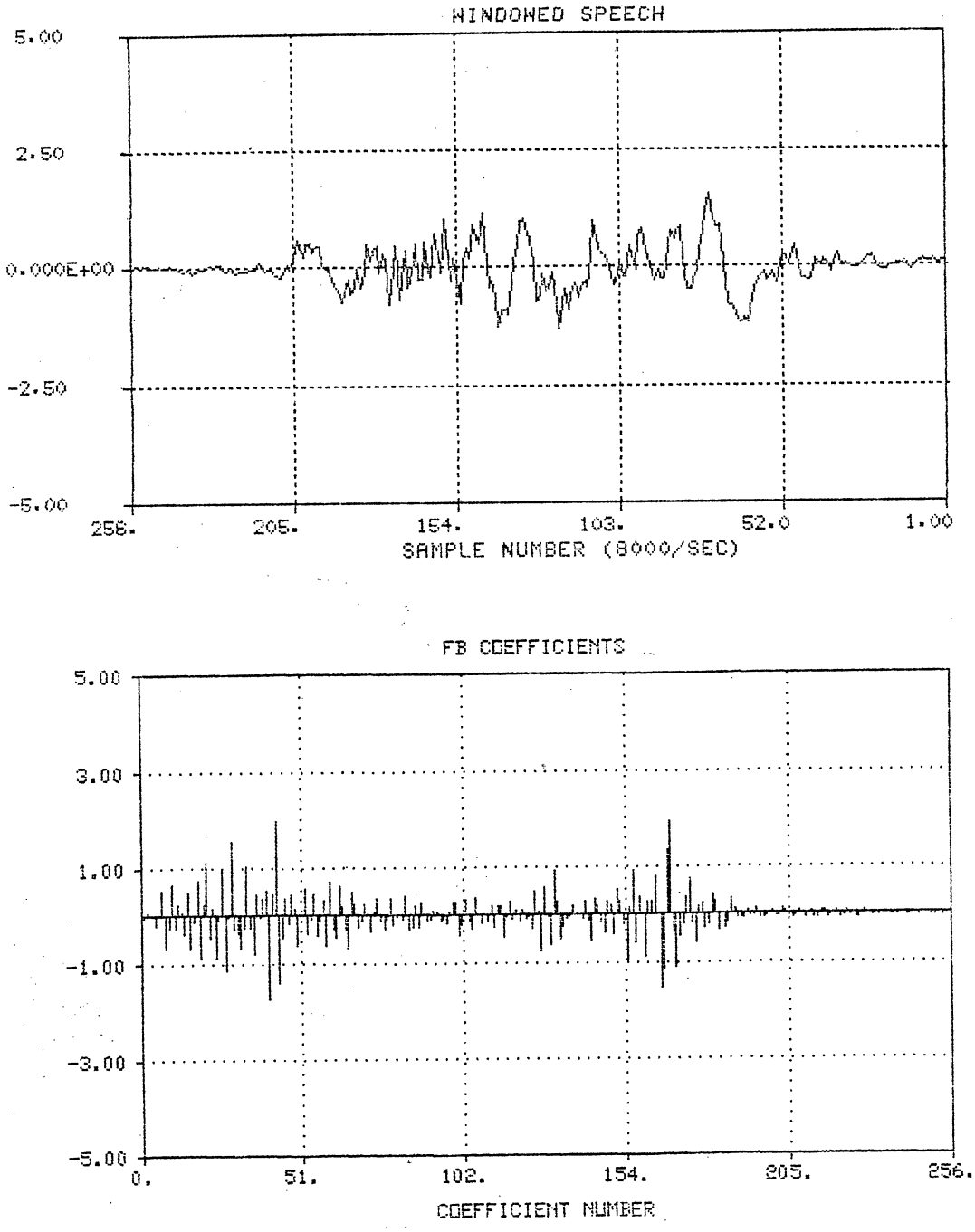


Figure 36. Unvoiced speech segment, with Fourier-Bessel coefficients. Phoneme: "h" fricative. Speaker 4. Note lack of regularity of coefficients.

sult is hardly surprising, in view of the fact that the process of Fourier-Bessel expansion is equivalent to a bank of FIR digital filters, each of which has a narrow frequency response.

Extraction of Pitch Information From the Fourier-Bessel Coefficients

It was noticed during many trials of Fourier-Bessel expansion that there was some regularity to the occurrance of large peaks: The spacing of the largest peaks seemed to correspond to the pitch frequency. The pitch information for voiced speech is fairly easily extracted from the coefficients if a simple algorithm is performed:

STEP 1. Perform Fourier-Bessel expansion of a windowed segment of speech.

STEP 2. Perform a three-point maximum-filter operation on the coefficient set. This is similar to a median filter, except that the maximum absolute value of every three samples is used as the output.

STEP 3. Perform a three-point median filter operation on the result of Step 2.

STEP 4. Measure the average distance between peaks of the resulting waveform. The pitch frequency is then well approximated by

$$f_{pitch} \approx \frac{\Delta m - .25}{2N} f_S \qquad (6.1)$$

where Δm is the average number of samples between

peaks.

The result of performing this process on the set of coefficients from Figure 26 is shown in Figure 37. This is essentially the same as measuring pitch from the fine structure of the magnitude spectrum.

Formants can be distinguished in the Fourier-Bessel coefficients, but there does not seem to be any great advantage in using Fourier-Bessel coefficients instead of Fourier coefficients for determination of formant frequencies and relative energies. It was found that the formant frequencies could be distinguished, but none too clearly, from the raw Fourier-Bessel coefficient set. When the simple nonlinear filter mentioned in the previous pitch detection scheme was applied to the coefficients, the formants were easier to discern.

Figure 38 shows the result of applying the maximum-median filter to the coefficient set shown in Figure 36. Note that the peaks are irregular as compared to the case of voiced speech.

Refinement of Fourier-Bessel Coefficients and Feature Extraction

All of the same information is contained in the Fourier-Bessel representation as in the common Fourier representation of the signal. However, the information is in a different form. It should be noted that most automatic speaker recognition schemes use as a feature set the magni-

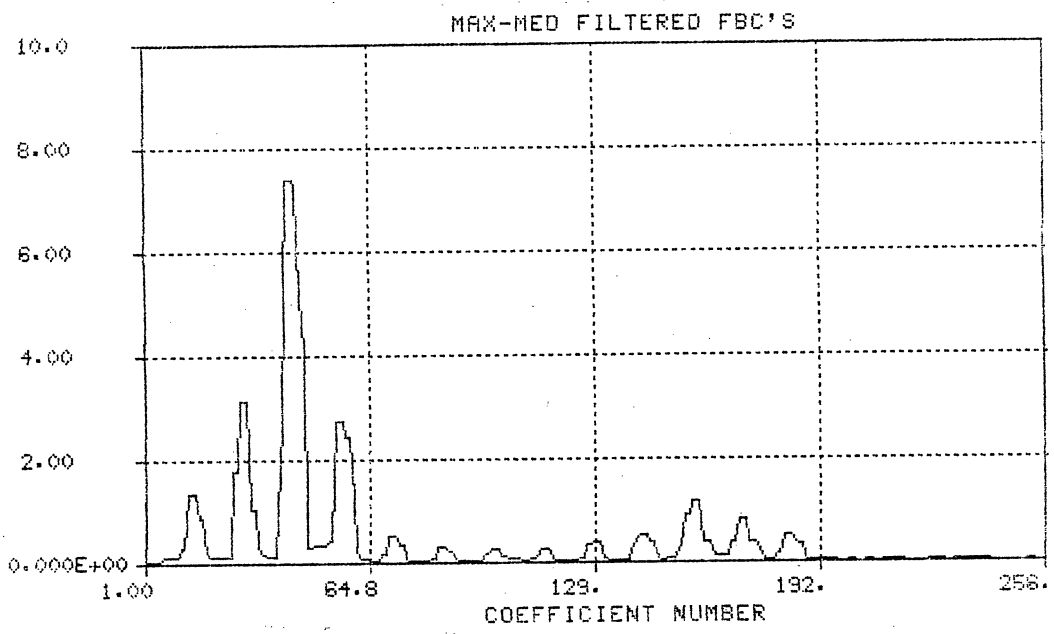


Figure 37. Result of maximum-median filtering operation on the Fourier-Bessel coefficients of Figure 26. Voiced speech.

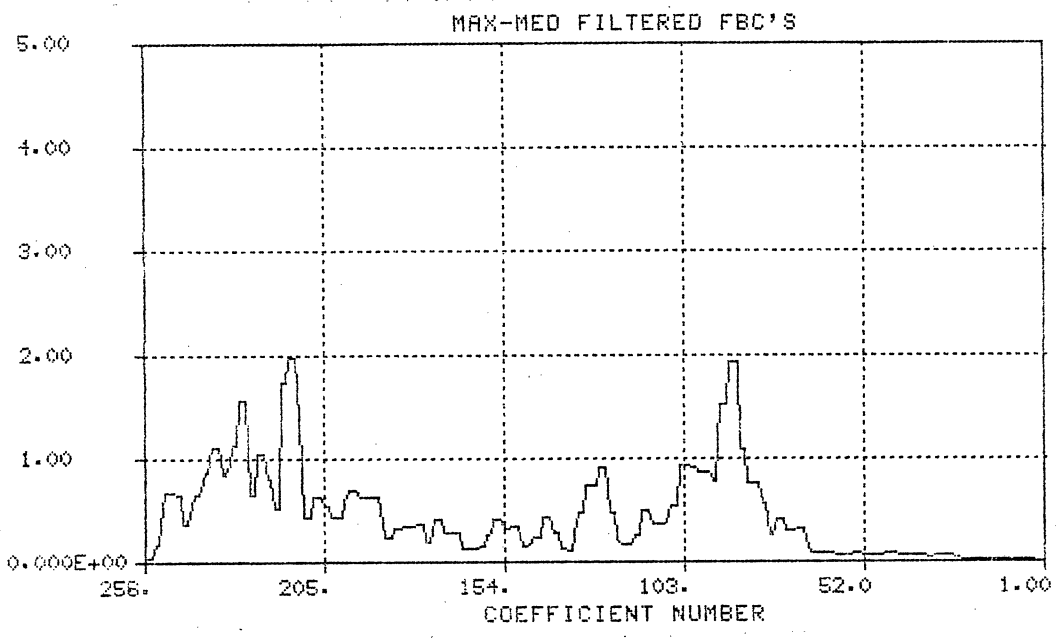


Figure 38. Result of maximum-median filtering operation on the Fourier-Bessel coefficients of Figure 36. Unvoiced speech.

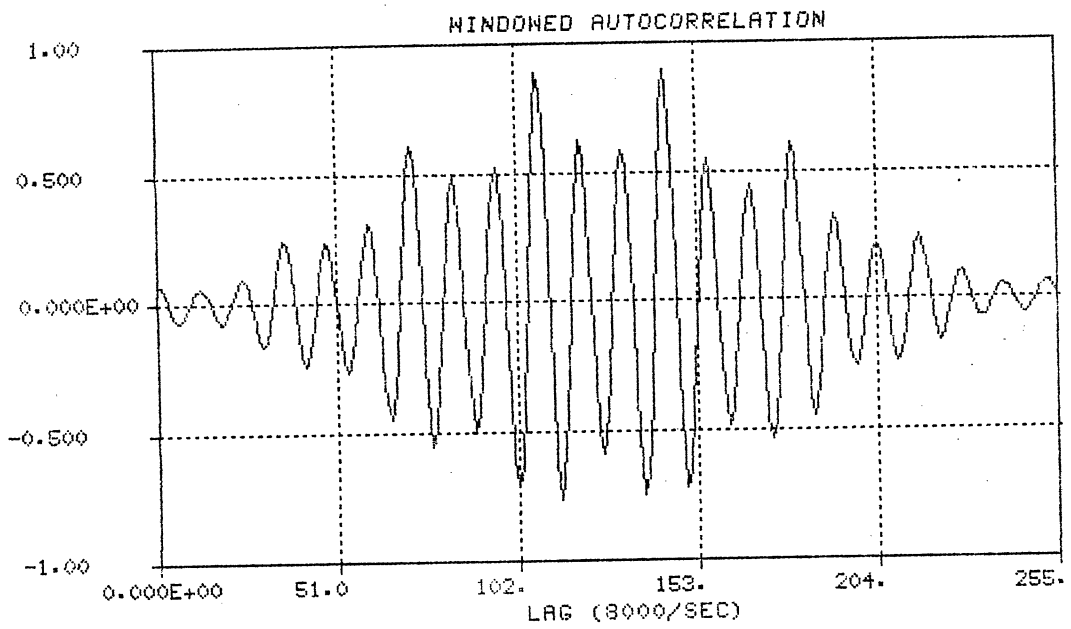
tude spectrum of the speaker, either in direct or indirect form. In the following discussions, it should be borne in mind that the main purpose was to try to find a feature set which was suitable for pattern recognition.

Fourier-Bessel Expansion of Autocorrelation Functions

In the previous portion of this Chapter, it was found that Fourier-Bessel expansion of windowed speech segments did not always result in a reliable and easily reproduced set of coefficients. Depending on the point where the analysis window started, a different set of coefficients resulted. Therefore, there seemed to be two options: (1) Make the analysis windows pitch-synchronous, or (2) perform Fourier-Bessel expansion of some function of the data which is robust in the face of different analysis-frame starting points. The latter method was chosen; Autocorrelation functions can be easily expanded into Fourier-Bessel series.

An example of this technique is shown in Figure 39. In this example, a Hamming window was used. The result of using the Hamming lag window is shown in Figure 40. The transformation of the autocorrelation rather than the original time domain waveform not only provides a means of normalizing the beginning of the analysis frame, but also provides a convenient opportunity to normalize the magnitudes by dividing by the signal's variance.

In all cases note how few coefficients are really re-



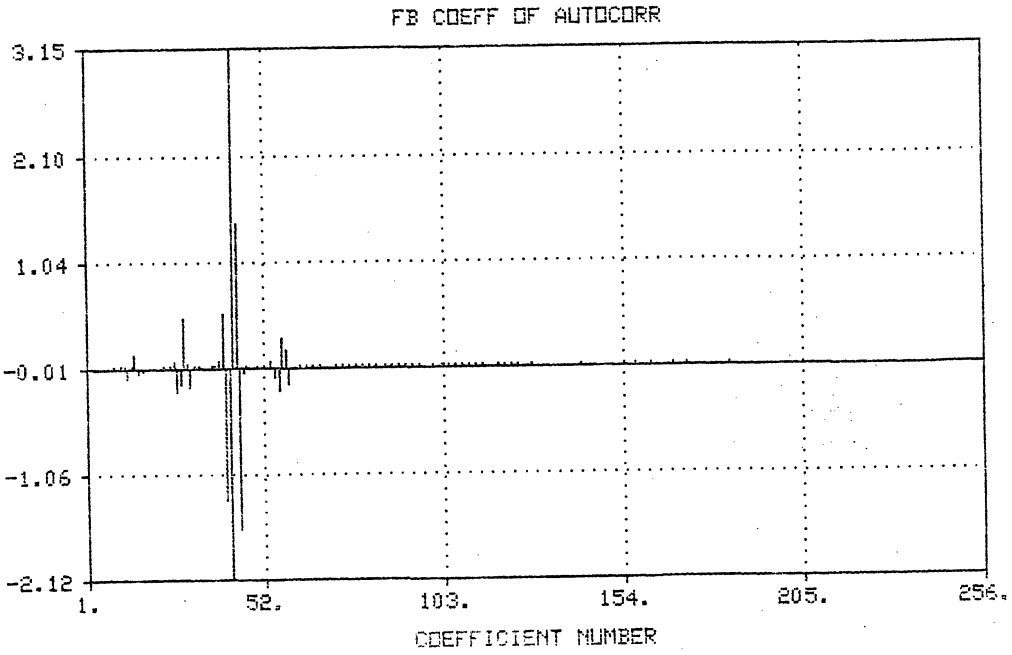
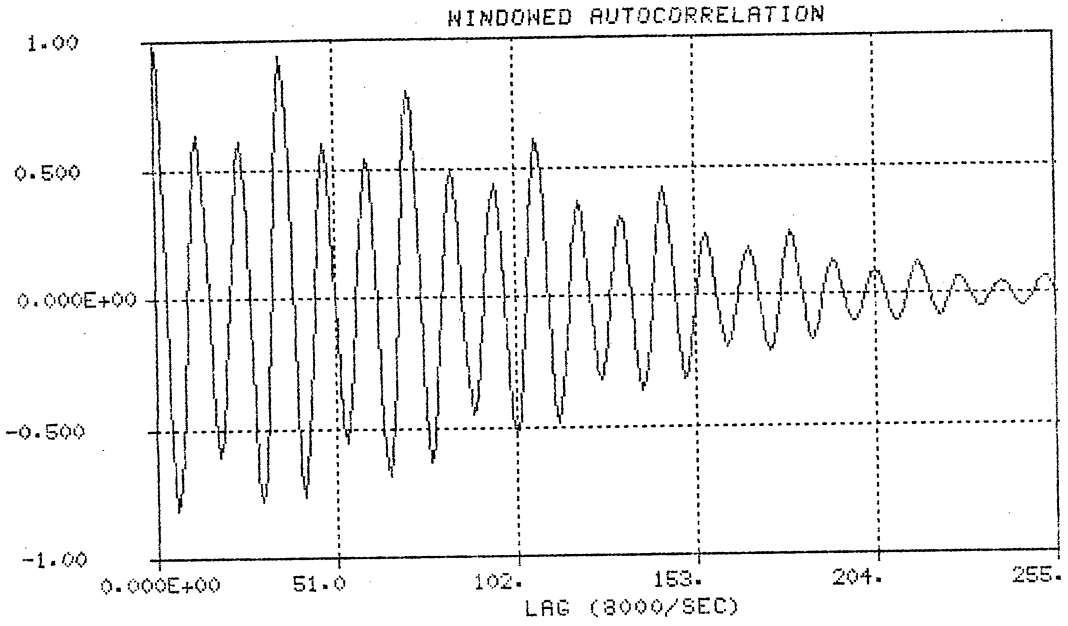


Figure 39. Windowed autocorrelation, with Fourier-Bessel coefficients. Same speech segment as in Figure 26. Hamming window.



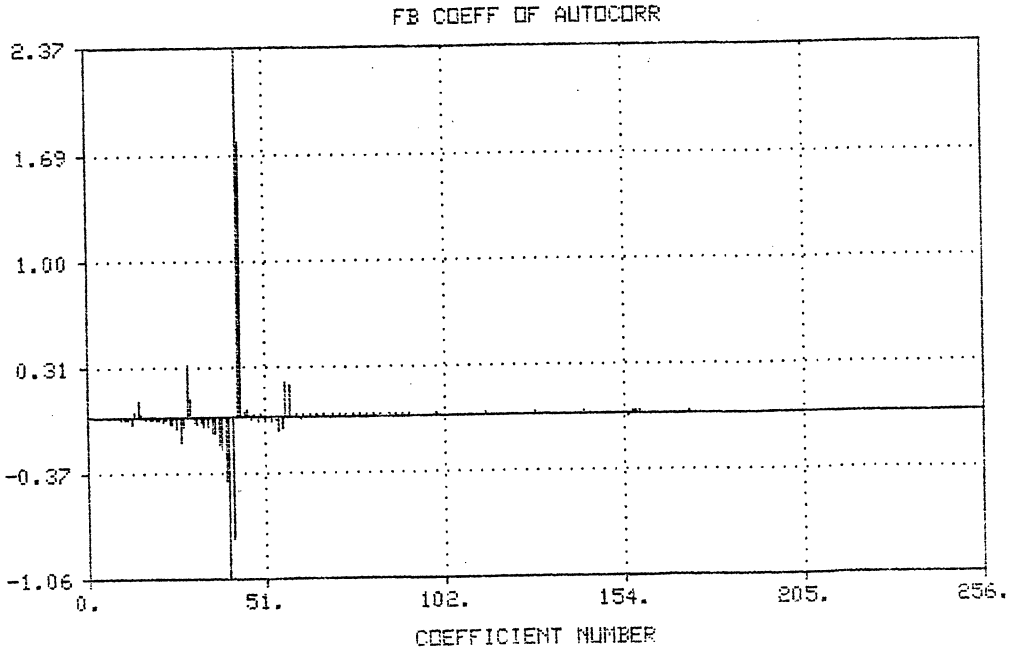


Figure 40. Windowed autocorrelation, with Fourier-Bessel coefficients. Same speech segment as in Figure 26. Hamming lag window was used.

quired for representation of the windowed autocorrelation. Definite differences between speakers are also noticable. The distance between positive peaks can be directly related to the pitch, by the same method as was explained in the previous section. Also, the distance between adjacent positive and negative peak coefficients is a possible feature for pattern recognition.

But there is still the problem of extraction of a suitably compact feature set. From the methods already presented here, only the pitch can be extracted. The set of Fourier-Bessel coefficients obtained by transformation of the autocorrelation is too large a set to be used in a typical pattern classifier. Therefore, a different description of the magnitude spectrum was needed. Such a description, a form of generalized cepstrum, will now be discussed.

Fourier-Bessel Expansion of the Magnitude Spectrum

Suppose that the log magnitude spectrum is obtained by windowing, transforming, and then taking the logarithm by traditional methods. The magnitude spectrum is a waveform which can be expressed as a linear combination of Bessel functions on the interval $(0, f_S/2)$. Note that such a representation is homomorphic with repect to convolution, and thus could be considered as a sort of generalized cepstrum.

For example, consider two signals x(n) and y(n). If x(n) and y(n) are convolved, then their log magnitude spec-

tra add. Let $X(\omega)$ and $Y(\omega)$ represent the log magnitude spectra of $x(n)$ and $y(n)$. If $X(\omega)$ and $Y(\omega)$ are each modeled as series with some basis set $g_i(x)$,

$$X(\omega) = \sum_i C_i g_i(\omega) \quad , \quad Y(\omega) = \sum_i B_i g_i(\omega), \qquad (6.2)$$

then the log magnitude spectrum of $x(n)*y(n)$ can be written as the sum of the series for $X(\omega)$ and $Y(\omega)$:

$$X(\omega) + Y(\omega) = \sum_i (C_i + B_i) g_i(\omega). \qquad (6.3)$$

In Equations (6.2) and (6.3), the limits of the summation were left indeterminate, because different limits would be appropriate for various choices of the basis set $g_i(x)$. Equation (6.3) states that a series representation of the log magnitude spectrum is a homomorphic system for convolution. The basis set could be any of a wide variety of functions, including orthogonal polynomials, Bessel functions, Walsh functions, or the traditional trigonometric series. It is shown in Appendix B that the trigonometric series representation of the log spectrum is equivalent to the cepstrum. But there is no reason why alternative basis sets should not be used for homomorphic system characterization, deconvolution, or filtering.

Now consider an example. Figure 41 shows a windowed speech segment which will be analyzed. It is the same as the segment used in Figure 26. An FFT was used to estimate the power spectrum. The log power spectrum is shown in Figure 42. A tapered cosine window was applied to the magnitude

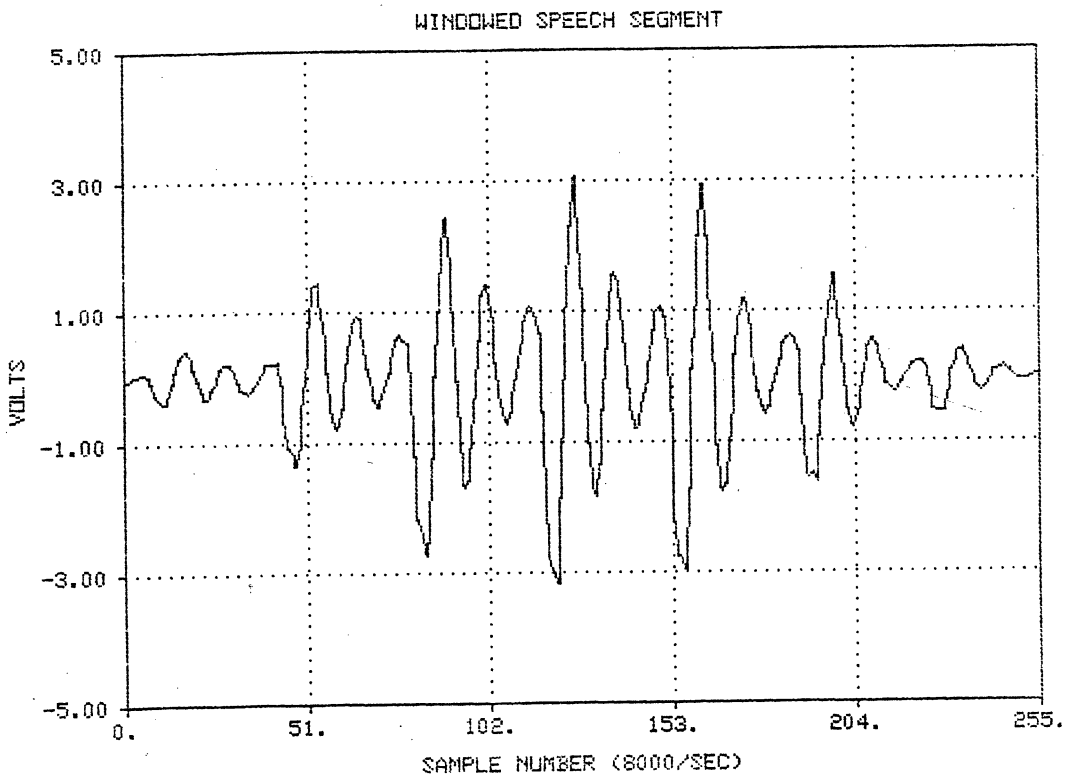


Figure 41. Windowed speech segment for cepstrum analysis. Speaker 1, phoneme "a" as in "cats". Hamming windowed.

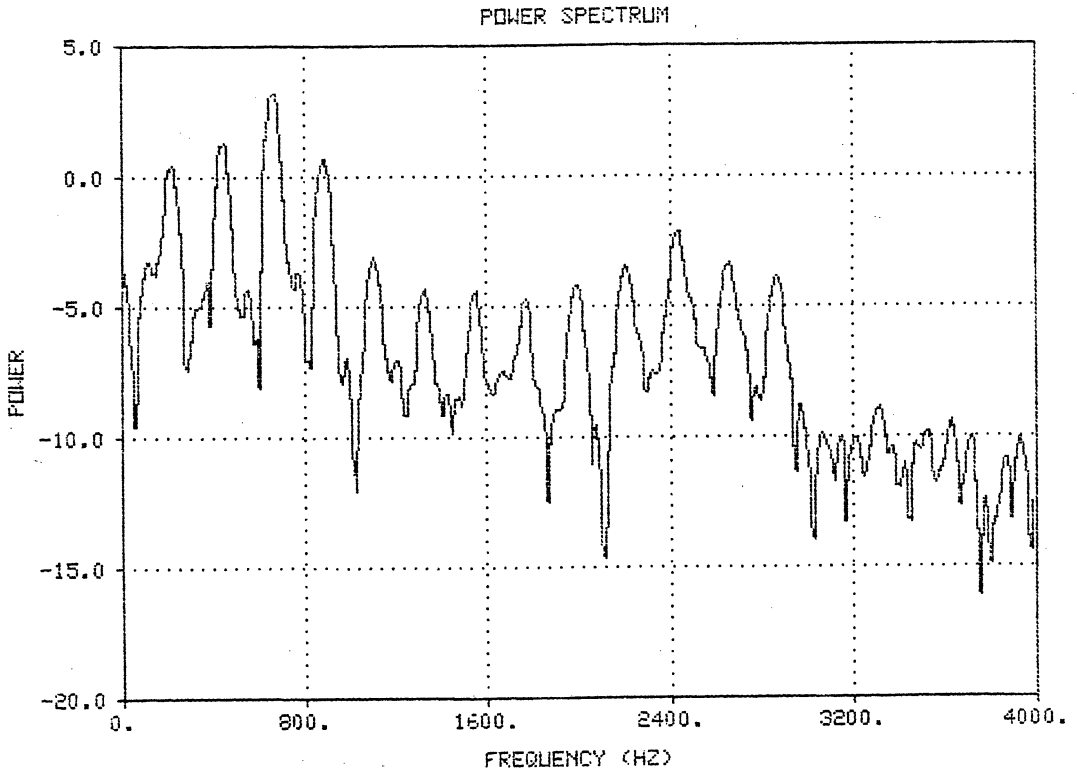


Figure 42. Log magnitude spectrum of speech segment from Figure 41.

spectrum, and the result is shown in Figure 43. Also, a constant was added to the log power spectrum so that the boundary condition for the Fourier-Bessel series can be met at the right end of the interval. The log power spectrum was then expanded as a series of Bessel functions. The first 100 coefficients of the expansion are shown in Figure 44.

As with the traditional cepstrum, the "low-time" part of the cepstrum represents the gross shape of the magnitude spectrum, while the pitch period is clearly shown as a very large coefficient. The interpretation of the horizontal axis in Figure 44 as "time" is a subject open for discussion.

The degree to which a limited number of these coefficients can represent the magnitude spectrum is illustrated in Figures 45-50. The general shape of the spectrum can be represented by the first nine to twelve coefficients.

The pitch estimate is formed as follows. First, it should be recalled that the approximate frequency associated with the m-th Fourier-Bessel coefficient is

$$f \approx \frac{m - .25}{2} \quad \frac{\text{cycles}}{\text{analysis frame}}. \tag{6.4}$$

When the pitch frequency is f_p, and the analysis frame goes from f=0 to $f=f_s/2$, then there should be $f_s/(2f_p)$ cycles of the fine structure of the spectrum in the analysis interval. Then, setting these two results equal to each other yields an expression that can be solved for f_p in terms of m (the coefficient number) and f_s (the sampling rate):

$$\frac{m - .25}{2} \approx \frac{f_s}{2 f_p}. \tag{6.5}$$

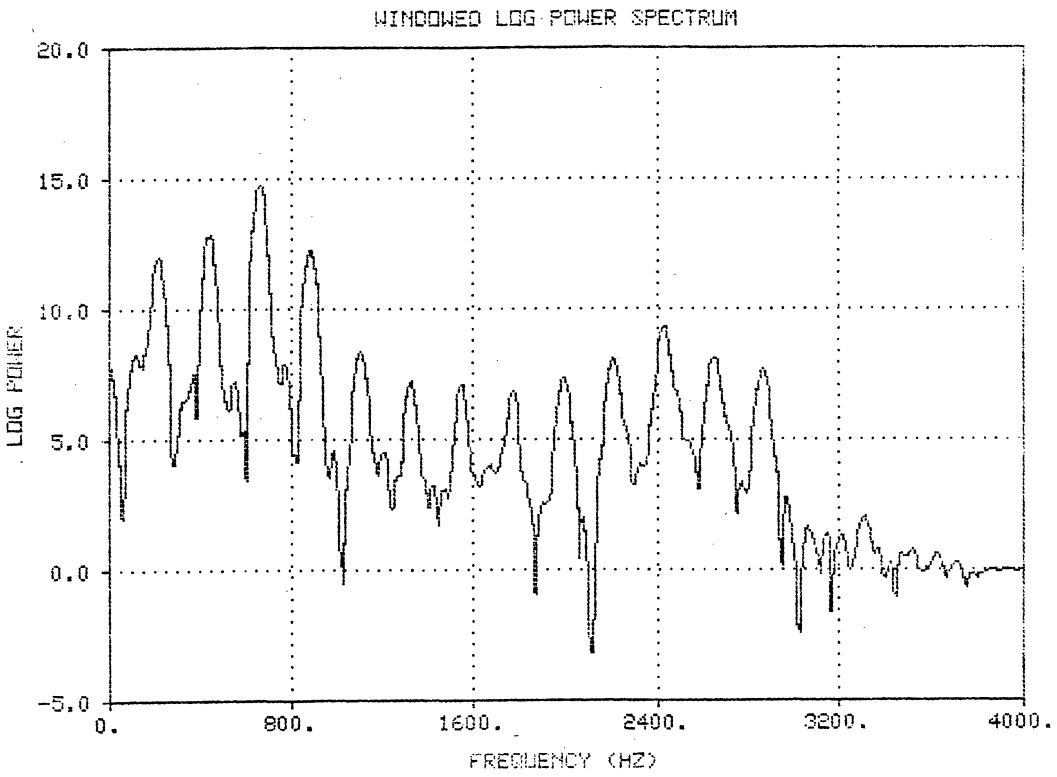


Figure 43. Windowed and normalized power spectrum from Figure 42. Tapered cosine window of Figure 20 used. Note that the log power has been forced to equal zero at f=4000 Hz.

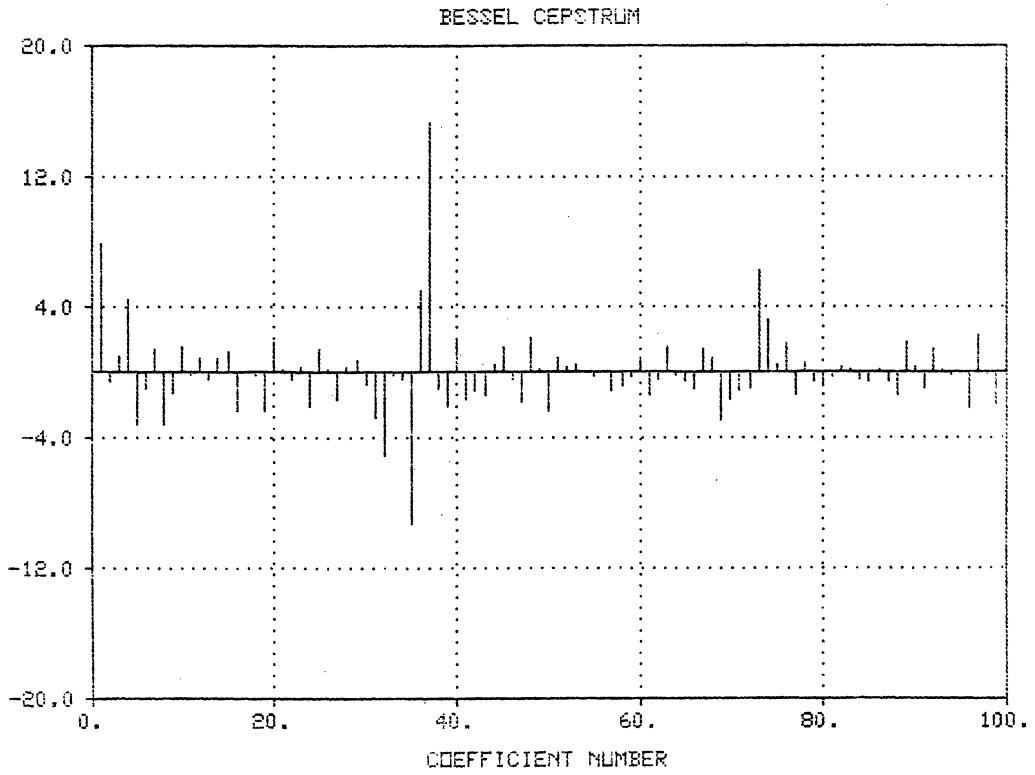


Figure 44. Fourier-Bessel expansion of the log power spectrum of Figure 43. Coefficient number 37 can be used for pitch estimation.

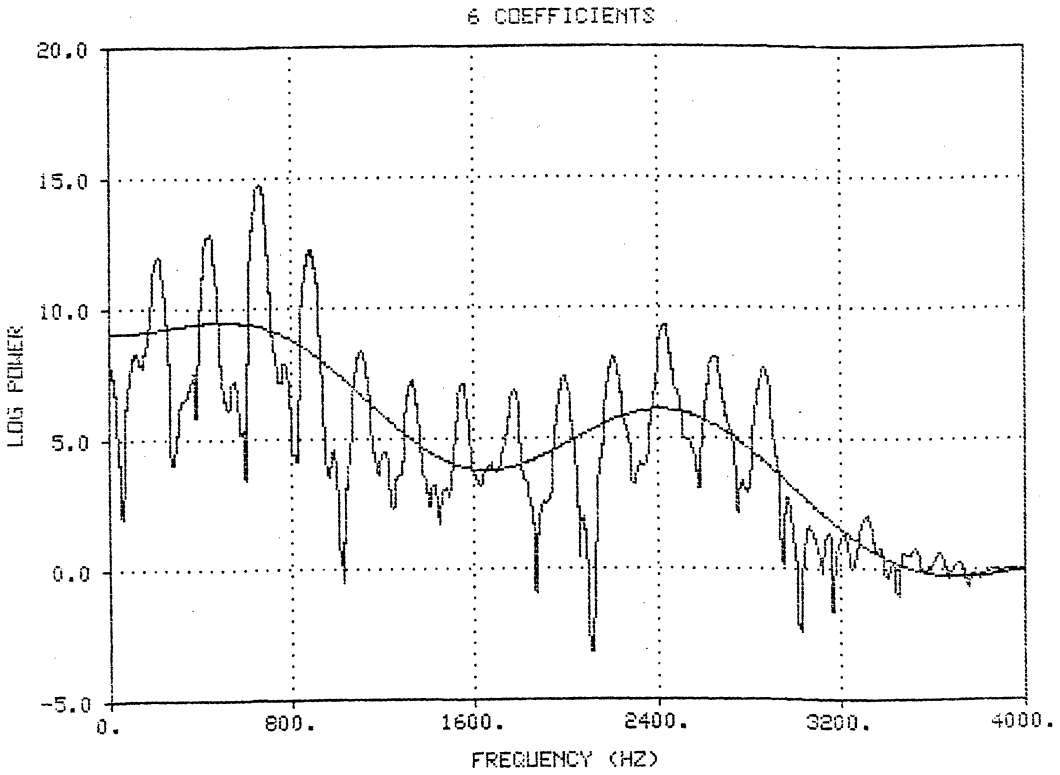


Figure 45. Comparison of actual magnitude spectrum to approximation formed by six coefficients. The smooth curve is the approximation.

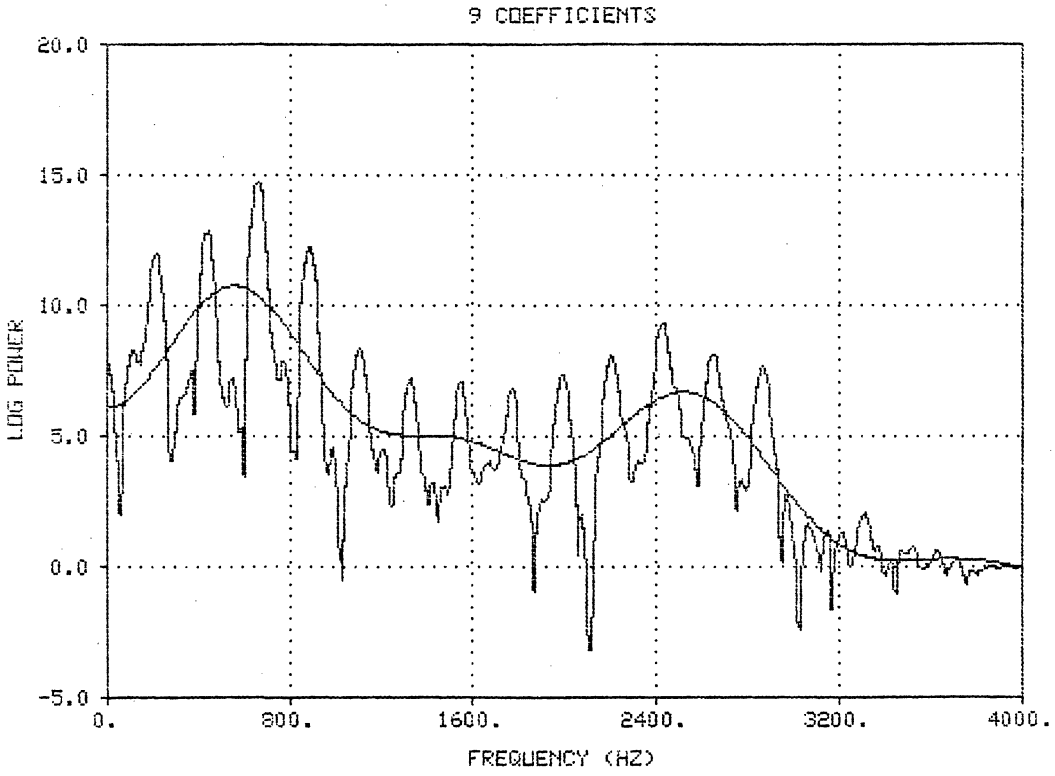


Figure 46. Comparison of actual magnitude spectrum to approximation formed by nine coefficients. The smooth curve is the approximation.

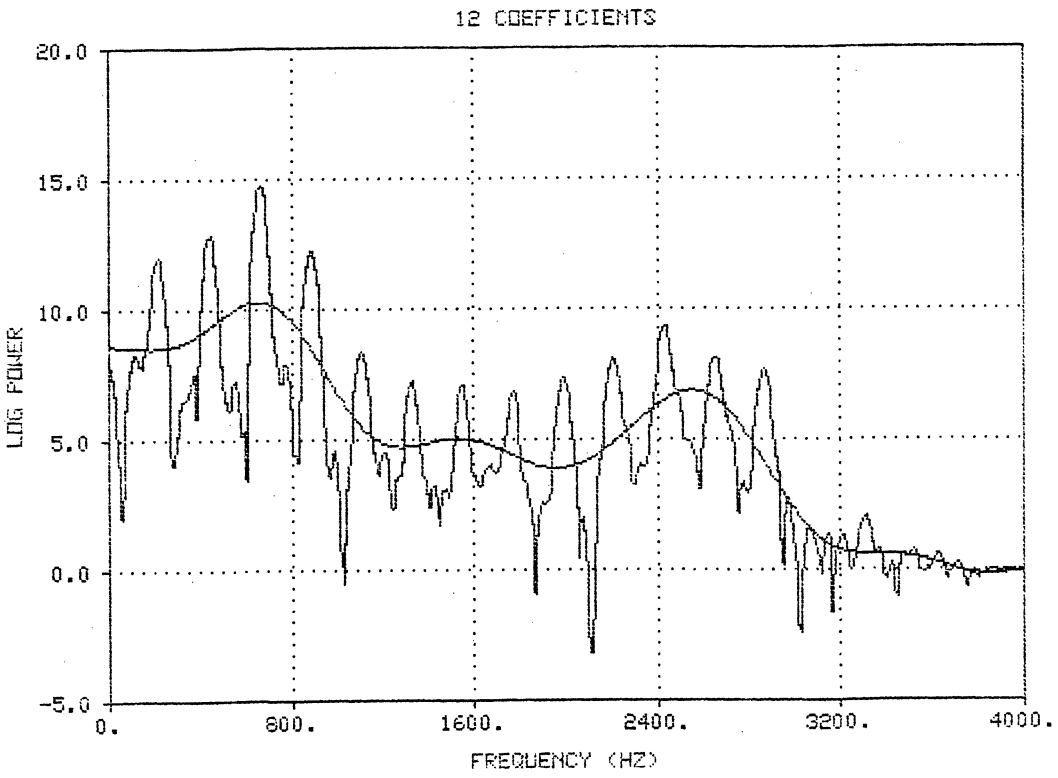


Figure 47. Comparison of actual magnitude spectrum to approximation formed by 12 coefficients. The smooth curve is the approximation.

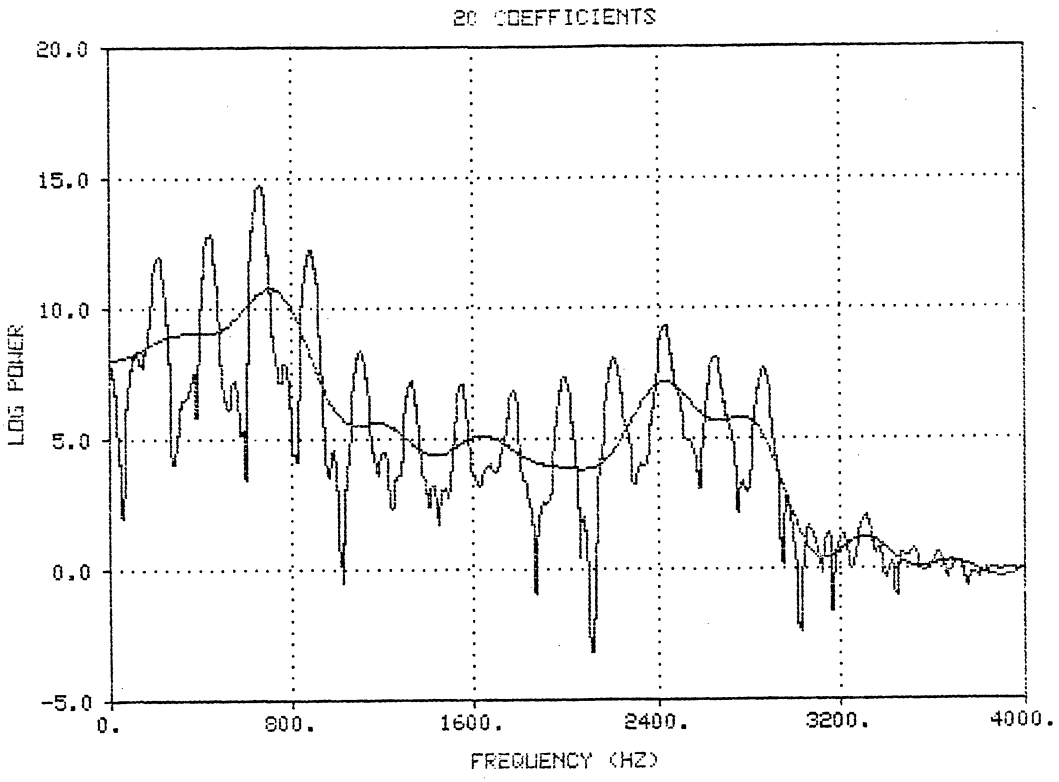


Figure 48. Comparison of actual magnitude spectrum to approximation formed by 20 coefficients. The smooth curve is the approximation.

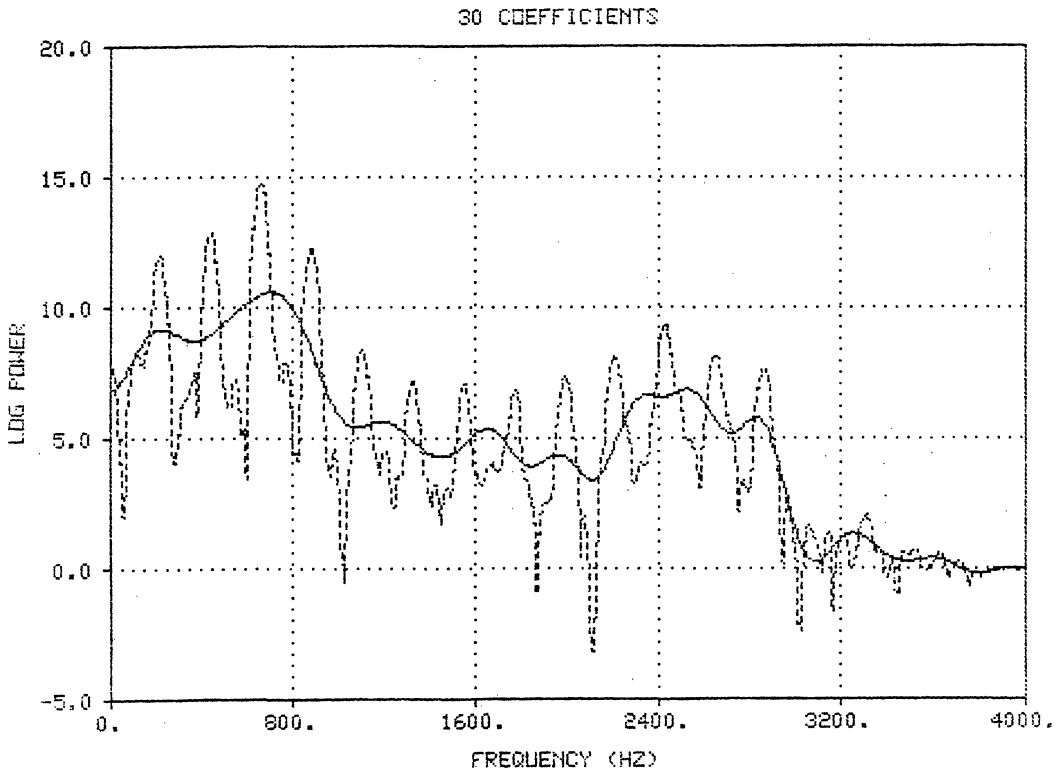


Figure 49. Comparison of actual magnitude spectrum to approximation formed by 30 coefficients. Dashed line: actual spectrum. Solid line: Approximation.

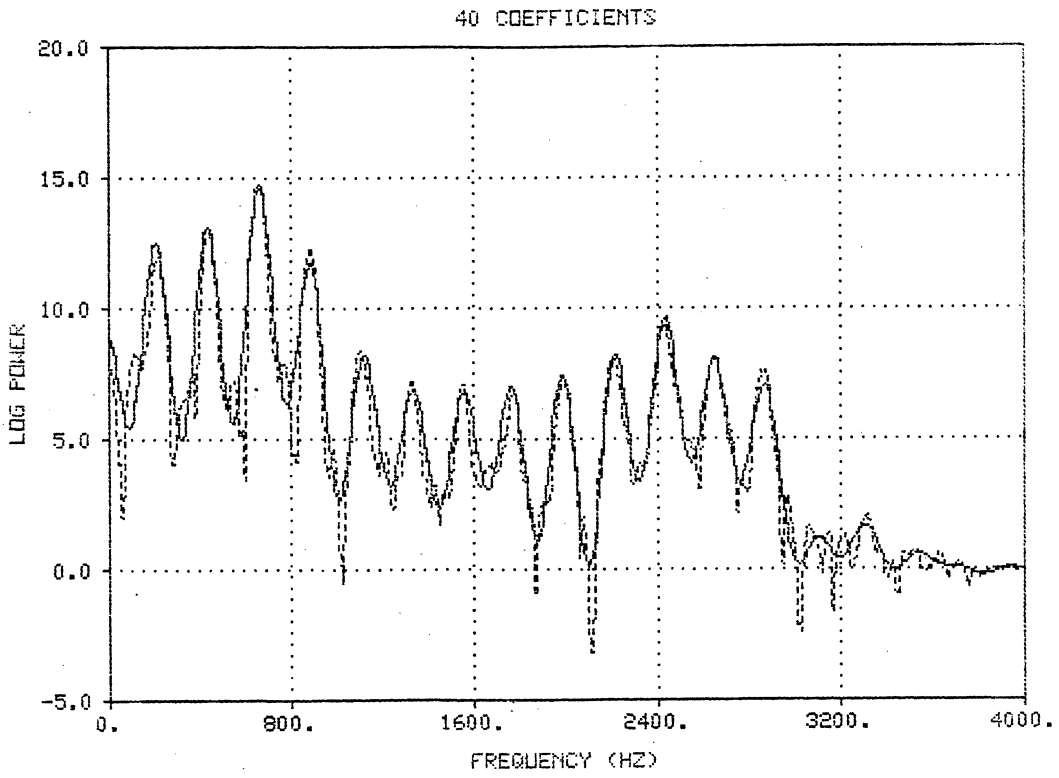


Figure 50. Comparison of actual magnitude spectrum to approximation formed by 40 coefficients. Dashed line: actual spectrum. Solid line: Approximation.

Solving the equation for f_p gives

$$f_p \approx \frac{f_s}{m - .25} . \qquad (6.6)$$

If m is the coefficient number of the largest Fourier-Bessel coefficient, then Equation (6.6) yields the pitch estimate. Additional work is required to test this method when the speech data is noisy.

Another speech example is shown in Figures 51-54. Again, note the prominence of the coefficient corresponding to the pitch period.

The most common method of deconvolution or filtering with a homomorphic system consists of gating, or windowing, the cepstrum coefficients. Then the signal is reconstructed from the modified cepstrum. But when the traditional cepstrum is used, deconvolution by this method can only be effective if the cepstra of the two signals to be separated do not overlap in the quefrency domain (Tribolet, 1979).

Now suppose that a cleverly chosen basis set and its companion series expansion are applied to the log magnitude spectrum of the original signal. If the basis set is chosen properly, then the quefrency domain coefficients may be separated so that they do not significantly overlap. (The quefrency domain can be generally defined as the domain that is the result of a transform of the log power spectrum.) Then gating or windowing can be used to eliminate the undesired signal, and the deconvolved signal can be reproduced.

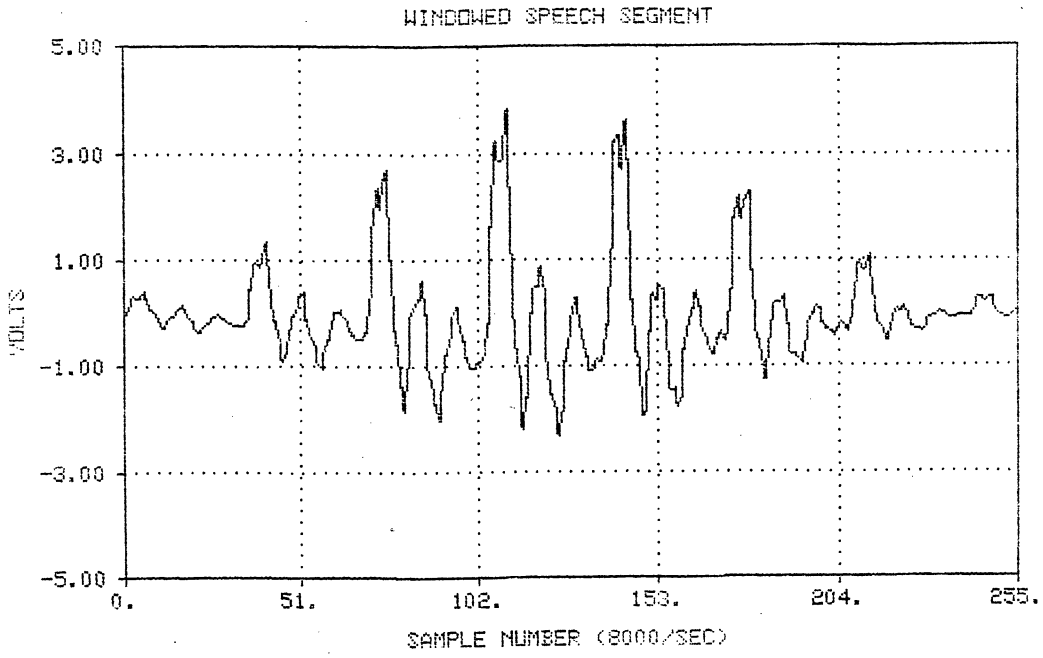


Figure 51. Windowed speech segment, "a" as in "cats". Hamming windowed. Speaker 1, Trial 2.

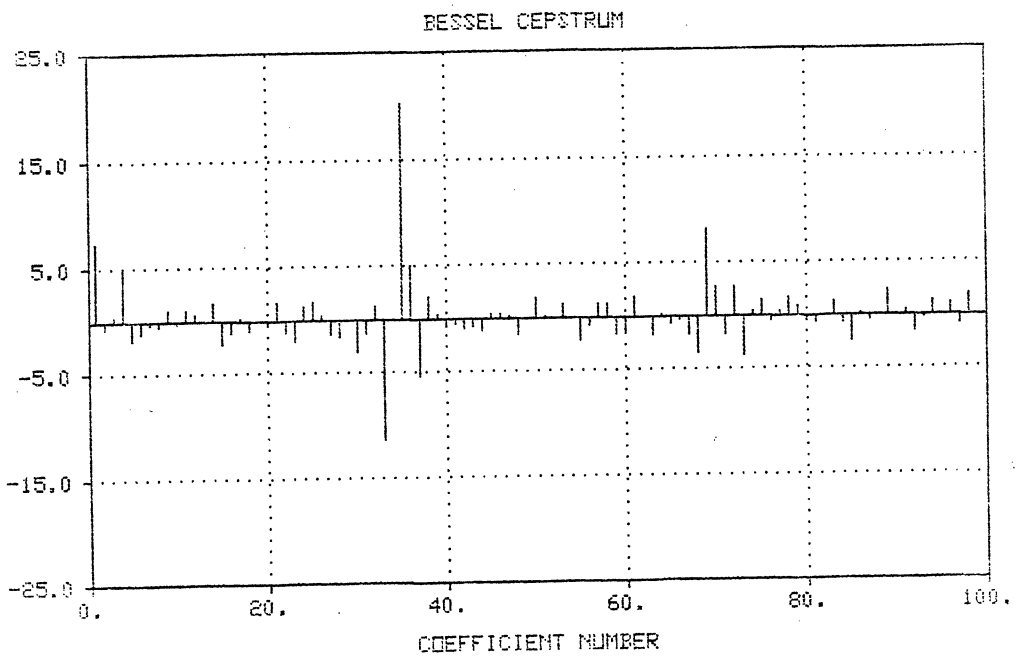


Figure 52. Fourier-Bessel cepstrum of speech segment of Figure 51. Large coefficient at m=35 corresponds to pitch.

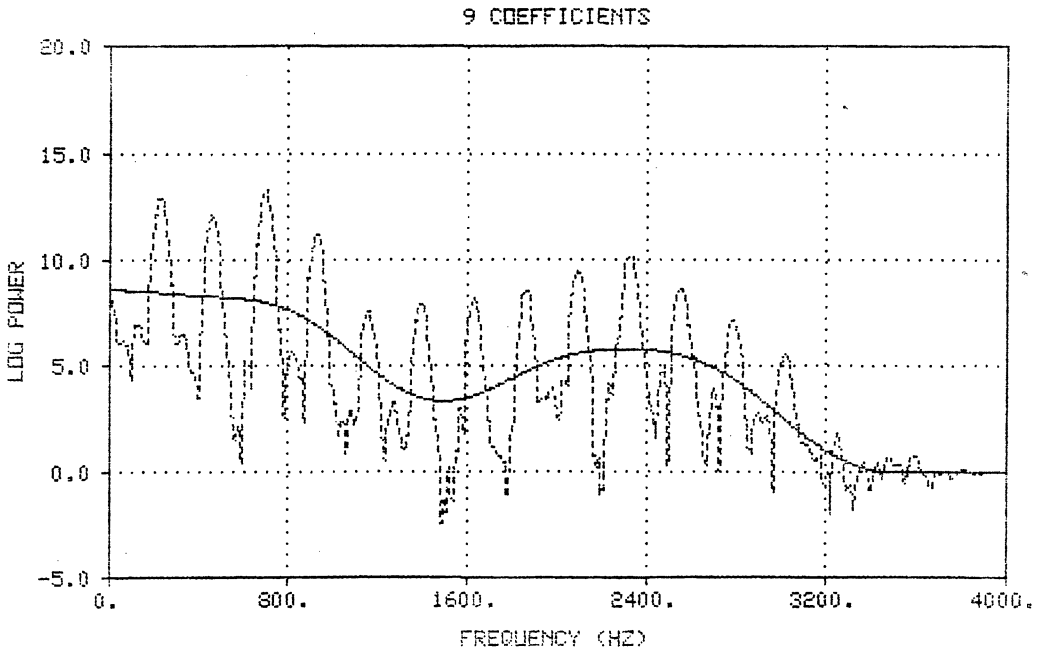


Figure 53. Approximation to power spectrum of the speech segment of Figure 51, using nine coefficients. Dashed line: Actual spectrum. Solid line: Approximation.

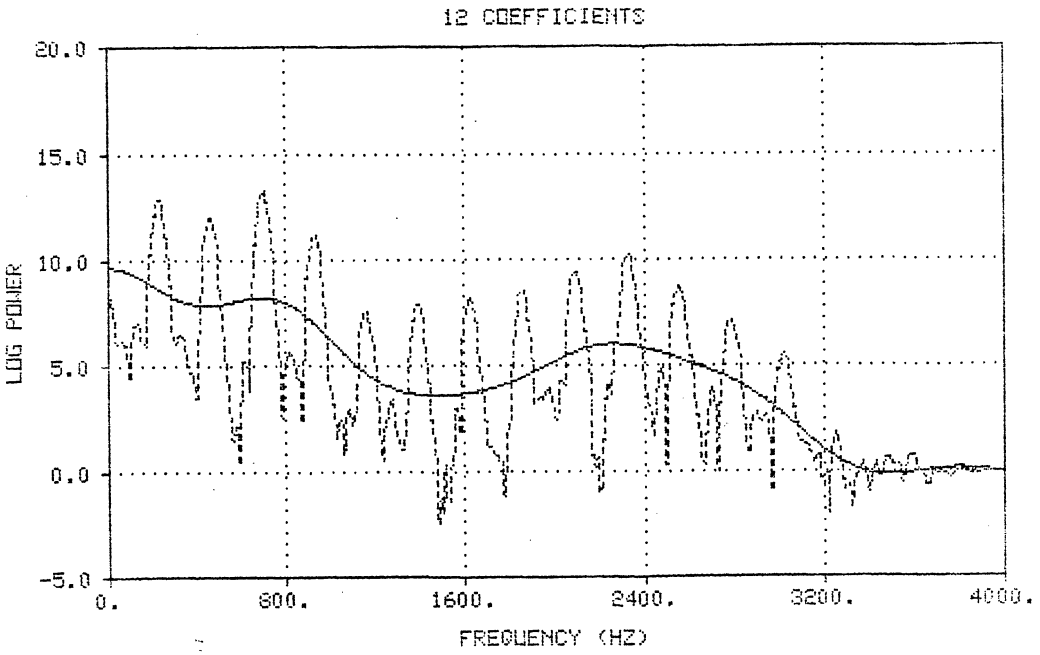


Figure 54. Approximation to power spectrum of the speech segment of Figure 51, using 12 coefficients. Dashed line: Actual spectrum. Solid line: Approximation.

Feature Sets for Pattern Recognition

In theory, any of the Fourier-Bessel expansions presented in this report could be used to generate feature vectors. But the practical problems with large feature sets preclude the use of the raw Fourier-Bessel coefficients of windowed speech segments. The problems are: (1) dimensionality of the feature space; (2) efficient extraction of the features; and (3) reproducibility and robustness of the features.

The problem of efficient calculation of the Fourier-Bessel series coefficients has been practically alleviated by use of the fast transforms of Chapter IV. The problem of reproducibility of the features can be mitigated by expanding the autocorrelation or the magnitude spectrum, because each of these waveforms has a naturally defined starting point. But the problem of dimensionality of the feature space is reasonably solved only by the Fourier-Bessel expansion of the log magnitude spectrum.

After much experimentation with various aspects of the Fourier-Bessel expansion, it appears that about the best feature set would consist of the pitch plus the first few Bessel cepstrum coefficients. Unfortunately, comprehensive testing of this proposed feature set was precluded by the lack of a high quality database for the speaker recognition problem. Such a data base should consist of several different sentences uttered by at least 20 speakers, with at least

25 trials of each. Also, the data must be collected on different days and under good recording conditions. Obviously, the collection of such a data base can be a formidable task.

The results are expected to be much the same as with other cepstrum-based methods. Enough information is available in the literature to make comparisions possible with other feature sets.

Chapter Summary

Some of the practical aspects of Fourier-Bessel analysis have been presented. One of the chief difficulties is that the starting point of the analysis frame affects the coefficients obtained. Since there is no shifting theorem for the Bessel function basis set, this makes subsequent analysis of the coefficients very difficult. One solution to this problem is to transform some waveform which has a well-defined, natural starting point such as the autocorrelation or the magnitude spectrum.

Expression of the log magnitude spectrum as a linear combination of some basis functions amounts to a generalized cepstrum. Almost any convenient basis could be used for this purpose, and it is traditional to use the trigonometric basis sets for this purpose. It has been proposed that the use of these alternative expansions can enable the use of homomorphic filtering or deconvolution even when the traditional cepstrum coefficients overlap in the quefrency domain.

CHAPTER VII

CONCLUSION

The original purpose of this investigation was to find an alternative method of automatic speaker recognition. At first, it was thought that the Bessel functions would be a good basis set for speech signal modeling because: (1) The Bessel functions are the solutions to a class of differential equations with time-varying coefficients; (2) The Bessel functions resemble voiced speech waveforms. Other potential advantages included the existence of a fast transform algorithm for the Hankel transform integral, and also the fact that the result was real-valued rather than complex.

It has been shown that the result of a Fourier-Bessel transformation of a speech waveform is a rather complicated set of numbers, which are hard to interpret. This is because there is a lack of theoretical knowledge concerning sampled Bessel functions and the fact that alternative transform domains are usually unfamiliar to digital signal processing practitioners. It was found that the set of numbers produced by Fourier-Bessel expansion of a segment of speech was too large and variable to be used in a pattern classification scheme. It was also found that there is an approximate correspondence between the concept of traditional frequency

and the Fourier-Bessel coefficient numbers. Also, a method of pitch computation was devised.

The fast transformation for the Hankel transform integral was streamlined for use with the Fourier-Bessel expansion. It was found that the Fast Hartley Transform was very useful and efficient for this purpose. The Fast Hartley Transform shows much promise as a general-purpose signal processing tool. The errors involved in the approximate computation of the Fourier-Bessel expansion were small enough to ignore in most cases. This allowed further streamlining of the fast transformation algorithm by the elimination of the computationally costly "Fourier selection summation" method used by Candel (1981). Also, this reinforces the idea that an engineering solution does not have to be precise to a great many decimal places to be a useful tool.

One of the reasons that the Fourier-Bessel series was selected as a topic of research was that a fast algorithm, Candel's method, had previously been devised. This points to a very real problem of digital signal processing. The general transformation is equivalent to a matrix-vector multiplication. But given an arbitrary transform matrix, can a fast transform algorithm be derived which performs the same linear transformation as the desired matrix transformation? Up to the present time, most fast transformation algorithms have depended upon a simple, redundant structure of the transform matrix to allow a fast transformation to be derived. A general method of deriving fast transformation

algorithms equivalent (or approximately equivalent) to a given linear transformation would be very useful indeed.

This thesis also explored an alternative cepstrum, consisting of the Fourier-Bessel expansion of the log power spectrum of a segment of speech. This representation was found to be compact enough for possible application to pattern recognition. But more importantly, it points out that many alternative cepstra are possible, each corresponding to a different series expansion of the log power spectrum of a signal. Using such alternatives, homomorphic filtering may be applied to some classes of problems for which it has not been previously considered.

An area of future research using the Fourier-Bessel series is the estimation of pitch in the presence of noise. The generalized cepstrum method of pitch estimation could be applied to this problem. It is also possible that other transforms, such as the Walsh-Hadamard transform, could operate on the log magnitude spectrum to get other more easily computed pitch estimation schemes.

The area of speaker authentication could still be profitably explored. Building on the results presented in this thesis, the Fourier-Bessel representation of the autocorrelation could be further processed to extract features for pattern recognition. For example, pitch and formant information can be obtained. The "Bessel cepstrum" described in Chapter VI could also be used as a description of the magnitude spectrum.

Another area for future work is speech coding with the Fourier-Bessel series. It may be possible to use the Bessel cepstrum coefficients for speech coding. For example, the first few coefficients may be used to describe the magnitude spectrum's general shape, and the largest coefficients can be used for pitch detection. Once again, a carefully chosen basis set for the alternative cepstrum could yield a fairly compact and cheaply computed description of the log power spectrum.

REFERENCES CITED

Abramowitz, Milton and Irene Stegun, Eds. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. New York: Dover Publications, 1965.

Ahmed, N. and K.R. Rao. Orthogonal Transforms for Digital Signal Processing. New York: Springer-Verlag, 1975.

Bracewell, R.N. "Discrete Hartley Transform." J. Opt. Soc. Am., Vol. 73, No. 12, Dec. 1983, pp. 1832-1835.

_____. "The Fast Hartley Transform." Proc. IEEE, Vol. 72, NO. 8, Aug. 1984, pp. 1010-1018.

_____. The Fourier Transform and Its Applications. New York: McGraw-Hill, 1965.

Candel, S.M. "Dual Algorithms for Fast Calculation of the Fourier-Bessel Transform." IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-29, No. 5, Oct. 1981, pp. 963-972.

Chang, Y. and C.S. Chen. "The Hardware Implementation of Fourier-Bessel Vocoder." IEEE-Academia Sinica Workshop on ASSP Proc., Beijing China, April 1986.

Chen, C.S., K. Gopalan, and P. Mitra. "Speech Analysis and Synthesis Via Fourier-Bessel Representation." Proc. ICASSP, Tampa, Fla., March 1985, pp. 497-500.

Clark, Moyett T. "Word Recognition by Means of Orthogonal Functions." IEEE Transactions on Audio and Electroacoustics, Vol. AU-18, No. 3, September 1970, pp. 304-312.

Clero, J.P. "A New Discrete Transform Matrix." Proc. IEEE, Vol. 67, No. 10, Oct. 1979, p. 1463.

Condon, E.U. and H. Odishaw. Handbook of Physics. 2nd Ed. New York: McGraw-Hill, 1967.

Dolansky, L. "Choice of Base Signals in Speech Signal Analysis." IRE Transactions on Audio, Vol. AU-8, No. 6, Nov.-Dec. 1960, pp. 221-229.

Hartley, R.V.L. "A More Symmetrical Fourier Analysis Applied to Transmission Problems." Proc. IRE, Vol. 30, No. 3, March 1942, pp. 144-150.

Hershey, John E., and R.K. Rao Yarlagadda. Data Transportation and Protection. New York: Plenum Press, 1986.

Jerri, A.J. "Towards a Discrete Hankel Transform and its Applications." Applicable Analysis, Vol. 7, 1978, pp. 97-109.

Manley, Harold J. "Analysis-Synthesis of Connected Speech in Terms of Orthogonalized Exponentially Damped Sinusoids. The Journal of the Acoustical Society of America, Vol. 35, No. 4, April 1963, pp. 464-474.

Menke, William. Geophysical Data Analysis: Discrete Inverse Theory. New York: Academic Press, Inc., 1984.

Mersereau, R. and A. Oppenheim. "Digital Reconstruction of Multi-dimensional Signals from their Projections." Proc. IEEE, Vol. 62, pp. 1319-1338, 1974.

Nunez, Paul L. "Representation of Evoked Potentials by Fourier-Bessel Expansions." IEEE Trans. on Biomedical Engineering, Sept. 1973, pp. 372-374.

Olver, F.W.J. and D.J. Sookne. "Note on Backward Recurrence Algorithms." Mathematics of Computation, Vol. 26, No. 120, Oct. 1972, pp. 941-947.

O'Neill, L.A. "The Representation of Continuous Speech with a Periodically Sampled Orthogonal Basis." IEEE Transactions on Audio and Electroacoustics, Vol. AU-17, No. 1, March 1969, pp. 14-21.

Oppenheim, Alan V., George Frisk, and David Martinez. "An gorithm for the Numerical Evaluation of the Hankel Transform." Proc. IEEE, Vol. 66, No. 2, Feb. 1978, pp. 264-265.

Oppenheim, Alan V., George Frisk, and David Martinez. "Computation of the Hankel Transform Using Projections." J. Acoust. Soc. Am., Vol. 68, No. 2, Aug. 1980, pp. 523-529.

Papoulis, A. Signal Analysis. New York: McGraw-Hill, 1977.

Rabiner, L.R. and R.W. Schafer. Digital Processing of Speech Signals. Englewood Cliffs, N.J.: Prentice-Hall, 1978.

Shum, F., A. Ronald Elliot, and W. Owen Brown. "Speech Processing With Walsh-Hadamard Transforms." IEEE Transactions on Audio and Electroacoustics, Vol. AU-21, No. 3, June 1973, pp. 174-179.

Siegman, A.E. "Quasi-Fast Hankel Transform." Optics Letters, Vol.1, No. 1, July 1977, pp.13-15.

Tolstov, Georgi P. Fourier Series. New York: Dover, 1962.

Tribolet, J.M. Seismic Applications of Homomorphic Signal Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

Tsang, L., R. Brown, J.A. Kong, and G. Simmons. "Numerical Evaluation of Electromagnetic Fields Due to Dipole Antennas in the Presence of Stratified Media." J. Geophysical Res., Vol. 79, 1974, pp. 2077-2080.

Waldron, R.A. "Formulas for Computation of Approximate Values of Some Bessel Functions." Proc. IEEE, Vol. 69, No. 12, Dec. 1981, pp. 1586-1588.

Watson, G.N. Theory of Bessel Functions. 2nd Ed. New York: The Macmillan Company, 1945.

APPENDIXES

APPENDIX A

FAST HARTLEY TRANSFORM ALGORITHM

Introduction

Bracewell (1984) introduced a BASIC language version of the Fast Hartley Transform. It was not an in-place algorithm so it required a lot of memory. Some improvements were made in the fast transform, so that memory was saved and the lookup table for sines and cosines which Bracewell provided was eliminated. To avoid the (time consuming) sine and cosine evaluations, recursive calculation was installed. But due to numerical instability, some users may wish to disable this feature.

The Program

```
      SUBROUTINE FHT(F,N,FWD)
C---------FAST IN-PLACE HARTLEY TRANSFORM
C---------ROBERT HAMILTON
C          APRIL 6, 1986
C          This algorithm is an improvement of the previous
C          one given by Bracewell. This subroutine requires
C          no workspace or extra arrays internally. Also,
C          the sine/cosine lookup-table has been eliminated
C          and recursive computation of the trig functions
C          has been installed.

C     INPUTS:
C       (1) F = REAL array ; put the input data here
C       (2) N = INTEGER; length of the transform; should be
C           a power of 2.
C       (3) FWD = INTEGER, which tells the FHT subprogram
```

```
C          whether to divide the result by N. If FWD=1,
C          the division will be performed; otherwise, it
C          will not be performed.

C     OUTPUTS:
C       (1) F = array of output values.
C       Arguments N and FWD are not affected by the sub-
C       routine.

       INTEGER FWD,N
       REAL F(0:N-1)
       INTEGER P,N2,N4,S0,S2,S4,D,E,Q
       DATA PI/3.14159265/

       IF(N.LT.2) STOP 'N.LT.2 in FHT.....'

       IF(N.NE.LASTN)THEN    ! find P, where 2**P = N
         P=1
         I=2
5        IF(N.EQ.I) GO TO 10
           I=I*2
           P=P+1
           IF(I.GT.N) STOP 'N not a power of 2 in FHT call.'
           GO TO 5
10       CONTINUE
         N2=N/2
         N4=N/4
         W=2.*PI/FLOAT(N)
       ENDIF

C------PERMUTE, PLACE IN BIT-REVERSED ORDER

       J=-1
       I=-1
1100   I=I+1
       E=N     ! E=2**P
1110   E=E/2
       J=J-E
       IF(J.GE.-1)GO TO 1110
       J=J+2*E
       IF(I.LE.J)GO TO 1100
       TEMP=F(I+1)
       F(I+1)=F(J+1)
       F(J+1)=TEMP
       IF(I.LT.(N-3))GO TO 1100

C-----ENDPERMUTE

       DO I=0,N-2,2
         TEMP=F(I)
         F(I)=TEMP+F(I+1)
         F(I+1)=TEMP-F(I+1)
       ENDDO
```

```
      IF(P.EQ.1)GO TO 500 ! DONE

      I=0
      J=2
130   TEMP=F(I)
      F(I)=TEMP+F(J)
      F(J)=TEMP-F(J)
      I=I+1
      J=J+1
      TEMP=F(I)
      F(I)=TEMP+F(J)
      F(J)=TEMP-F(J)
      I=I+3
      J=J+3
      IF(I.NE.N) GO TO 130

      IF(P.EQ.2) GO TO 500 ! DONE

      S0=N4                    ! Now start the main iteration.
      S4=4                     ! Uses decimation-in-time.
      DO L=2,P-1
        S2=S2*2
        S0=S0/2
        TEMP=W*FLOAT(S0)
        CX=COS(TEMP)
        SX=SIN(TEMP)
        DO Q=0,N-1,S2
          I=Q
          D=I+S4
          TEMP=F(I)
          F(I)=TEMP+F(D)
          F(D)=TEMP-F(D)
          K=D-1
          SNX=0.
          CNX=1.
          DO J=S0,N4,S0
            I=I+1
            D=I+S4
            E=K+S4
1697        TEMP=CNX*CX-SNX*SX    ! recursive computation of
1698        SNX=SNX*CX+CNX*SX     ! sines and cosines. For
1699        CNX=TEMP              ! slightly better numerical
C1700       CNX=COS(FLOAT(J)*W)     ! accuracy, comment out
C1701       SNX=SIN(FLOAT(J)*W)     ! lines 1697,1698, 1699
            TEMP=F(D)*CNX+F(E)*SNX ! and use lines 1700 and
            Z=F(D)*SNX-F(E)*CNX    ! 1701.
            F(D)=F(I)-TEMP
            F(E)=F(K)-Z
            TEMP=F(I)+TEMP
            F(K)=F(K)+Z
            F(I)=TEMP
            K=K-1
```

```
          ENDDO ! Close loop J
          E=K+S4
        ENDDO       ! Close loop Q
        S4=S2
      ENDDO         ! Close loop L

500   IF(FWD.EQ.1)THEN    ! Divide by N
        TEMP=1./FLOAT(N)
        DO I=0,N-1
          F(I)=F(I)*TEMP
        ENDDO
      ENDIF
      LASTN=N
      RETURN
      END
```

APPENDIX B

RELATIONSHIP OF BESSEL FUNCTIONS TO THE COMPLEX CEPSTRUM

Introductory Remarks

During the course of researching the properties of the Fourier-Bessel series, an interesting interpretation of the complex cepstrum in terms of the Bessel functions was discovered. This relationship is not believed to have been previously published in the electrical engineering or mathematical literature, although it is actually quite simple to derive.

Before commencing the main discussion, the definition of the cepstrum will be reviewed. Briefly, the cepstrum of a sequence is defined as the inverse Z transform of the complex natural logarithm of the sequence's ordinary Z transform. Figure 55 illustrates the definition of the cepstrum with a block diagram.

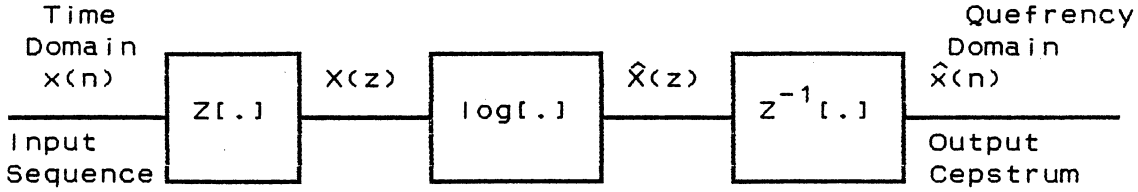


Figure 55. Illustration of the complex cepstrum.

The process of recovering a sequence from its cepstrum is illustrated by Figure 56.

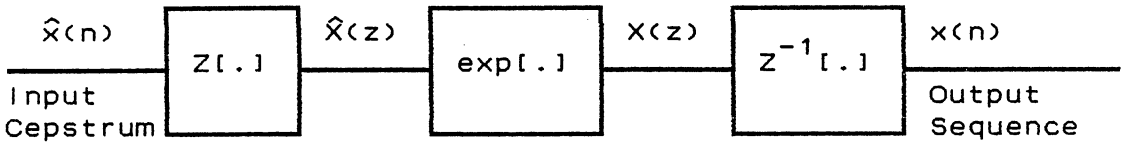


Figure 56. Illustration of recovery of a sequence from its cepstrum sequence.

The outstanding feature of the cepstral representation is that convolution in the time domain is equivalent to addition in the cepstral domain. Therefore deconvolution or filtering can, in principle, be performed in the cepstral domain by simply adding the cepstra of the sequences to be convolved. It is not the purpose of this section to give an exhaustive tabulation of all of the properties of the cepstrum. The interested reader is referred to the excellent book by Tribolet (1979), or for speech applications see Rabiner and Schafer (1978).

Relationship of Bessel Functions to the Complex Cepstrum

It will now be shown that the representation of a sequence by its complex cepstrum is equivalent to representing

the sequence's log magnitude spectrum with a cosine series and the sequence's phase with a sine series. Furthermore, it will be shown that representation of a signal by its cepstrum is equivalent to representing the signal as a convolution of Bessel function sequences. Rather than state these properties as theorems at this point, they will be delayed until special notation has been introduced. The key formula that relates a sequence $x(n)$ to its cepstrum $\hat{x}(n)$ is shown in Equation (B.17). Equations (B.24) and (B.25) relate the cosine series for the log magnitude spectrum, and the sine series for the phase spectrum, to the time domain sequence $x(n)$. The derivation of these formulae will now be given.

The generating function of the Bessel functions, $J_n(a)$, has been stated as

$$e^{\frac{a}{2}(t - \frac{1}{t})} = \sum_{n=-\infty}^{\infty} J_n(a)t^n. \qquad (B.1)$$

Substituting $1/z$ for the dummy variable t, an expression for the Z transform of the sequence of Bessel functions results:

$$Z\{J_n(a)\} = e^{\frac{a}{2}(\frac{1}{z} - z)}. \qquad (B.2)$$

Note that the argument of the Bessel functions, a, is now considered to be a constant parameter while the order of the Bessel function is varied over the entire range of integers.

The properties of this sequence are interesting. For example, the sequence may be generated using the recursion formula

$$J_{n+1}(a) = \frac{2n}{a} J_n(a) - J_{n-1}(a). \qquad (B.3)$$

If $J_0(a)$ and $J_1(a)$ are known, then $J_n(a)$ can, in principle, be calculated using (B.3). This formula is linear, but it does not have constant coefficients. The formula exhibits a numerical instability when |n| exceeds |a|. For this reason, it is sometimes recommended to use a backwards recursion formula to calculate the sequence (Olver and Sookne, 1972). An easy, but still very effective, way to approximate the sequence is to simply use the recursion formula with the initial conditions, $J_0(a)$ and $J_1(a)$, given above and then truncate the sequence when |n|>|a|. The elements of the sequence become very small indeed after that point and can be safely ignored.

The Fourier transform of the sequence can be obtained from the Z transform of the sequence by substitution of $\exp(j\omega)$ for z in Equation (B.2), yielding

$$e^{-ja\sin(\omega)} = \sum_{n=-\infty}^{\infty} J_n(a) e^{-jn\omega}. \qquad (B.4)$$

Note that the Fourier transform has a flat magnitude response and has a phase response which is sinusoidal in shape with a variation that is governed by the parameter a.

Now suppose that the original sequence of Bessel functions has k zeros inserted between each element:

$$\{J_n^k(a)\} = \{\ldots,0,J_{-1}(a),\underbrace{0,\ldots,0}_{k \text{ zeros}},J_0(a),0,\ldots,0,J_1(a),0,\ldots\} \qquad (B.5)$$

The notation for such a sequence shall be $\{J_n^k(a)\}$ where the superscript denotes the number of zeros inserted between each sample of the Bessel function sequence, the subscript is the order of the Bessel function, and a is the argument of the Bessel function, considered as a constant parameter. The Z transform of this sequence is derived simply from the Z transform of the original sequence by replacing z with z^{k+1} in (B.2). The resulting Fourier transform is

$$e^{-ja\ \sin[(k+1)\omega]} = \sum_{n=-\infty}^{\infty} J_n(a)\ e^{-j[k+1]\omega n}. \qquad (B.6)$$

Note that the phase response is no longer periodic with period 2π, but now has period $(2\pi)/(k+1)$. The shape of the phase response curve is still sinusoidal.

A similar situation exists for the case of the Bessel functions of the second kind, $I_n(a)$. The generating function for this set of functions is

$$e^{\frac{a}{2}(t + \frac{1}{t})} = \sum_{n=-\infty}^{\infty} I_n(a)t^n. \qquad (B.7)$$

Again substituting 1/z for t (where the parameter a is held constant), the Z transform of the sequence of functions is obtained:

$$e^{\frac{a}{2}(z + \frac{1}{z})} = \sum_{n=-\infty}^{\infty} I_n(a)z^{-n}. \qquad (B.8)$$

Letting $z = \exp(j\omega)$, the Fourier transform of the sequence is

$$F(\omega) = e^{a\cos(\omega)} = \sum_{n=-\infty}^{\infty} I_n(a)\, e^{-j\omega n}. \qquad (B.9)$$

Note that this Fourier transform has zero phase and has a magnitude response which is periodic with period 2π. The natural logarithm of the magnitude response is

$$\ln|F(\omega)| = a\cos(\omega). \qquad (B.10)$$

If k zeros are inserted in between each of the samples of the sequence of Bessel functions of the second kind then by an argument similar to the one used for the derivation of (B.6) the Fourier transform of the sequence becomes

$$e^{a\cos[(k+1)\omega]} = \sum_{n=-\infty}^{\infty} I_n(a)\, e^{-j(k+1)\omega n}. \qquad (B.11)$$

Consider the inverse system for homomorphic deconvolution shown in Figure 56. If the sequence $\hat{x}(n)$ in the quefrency domain (the domain in which $\hat{x}(n)$ exists) is equal to

$$\hat{x}(n) = \{\ldots,0,0,a/2,\underbrace{0,\ldots,0,\ldots,0}_{(2k+1)\text{ zeros centered around } n=0},a/2,0,\ldots\} \qquad (B.12)$$

then $\hat{X}(z) = \frac{a}{2}(z^{k+1} + 1/z^{k+1})$ so that $X(z) = \exp[\hat{X}(z)]$. Finally, when the inverse Z transform of $\exp[\hat{X}(z)]$ is computed as in Figure 56, the sequence of Bessel functions of the second kind is obtained (with k zeros inserted between samples):

$$x(n) = \{I_n^k(a)\} = \{\ldots 0, I_{-1}(a), 0, \ldots, 0, I_0(a), 0, \ldots, 0, I_1(a), 0, \ldots\}. \qquad (B.13)$$

Now consider the case where $\hat{x}(n)$ is an odd function of n in

the quefrency domain:

$$\hat{x}(n) = \{\ldots 0,0,\underbrace{-a/2,0,\ldots,0,\ldots,0,a/2}_{(2k+1)\text{ zeros centered around } n=0},0,0\ldots\}. \qquad (B.14)$$

Going through the inverse homomorphic system again, a sequence of Bessel functions of the first kind is obtained (with k zeros inserted between samples):

$$x(n) = \{J_n^k(a)\} = \{\ldots,0,J_{-1}(a),0,\ldots,0,J_0(a),0,\ldots,0,J_1(a),0,\ldots\}. \qquad (B.15)$$

Now suppose that some cepstrum, $\hat{x}(n)$, has been given. Then the cepstrum can be decomposed into the sum of even and odd parts

$$\hat{x}(n) = \hat{x}_{even}(n) + \hat{x}_{odd}(n) \qquad (B.16)$$

where $\hat{x}_{even}(n)$ is an even function $\hat{x}_{odd}(n)$ is an odd function of n. Each of these, in turn, can be decomposed into a sum of elementary cepstra as in Equations (B.12) and (B.14). A sum of cepstra represents convolution in the time domain. Therefore, the time domain sequence x(n) can be represented as a convolution of sequences of Bessel functions, where the order of each Bessel function is considered as a time variable and the argument of each Bessel function sequence is calculated from the cepstral coefficients:

$$x(n) = e^{\hat{x}(0)}\delta(n) * \overset{\infty}{\underset{k=0}{\ast}} \left[I_n^k[\hat{x}(k+1)+\hat{x}(-k-1)]\right] * \overset{\infty}{\underset{k=0}{\ast}} \left[J_n^k[\hat{x}(k+1)-\hat{x}(-k-1)]\right] \qquad (B.17)$$

The large asterisk is defined as the notation for the convolution of many sequences, in analogy with capital sigma for a summation or capital pi for a product. Equation (B.17) establishes a direct link between time and quefrency.

Another interpretation of the cepstrum is as a parametric representation of the magnitude and phase spectra of a time-domain sequence. Consider the log-magnitude spectrum of a real sequence, x(n):

$$\ln|X(\omega)| = \ln\left|\sum_{n=-\infty}^{\infty} x(n)\, e^{-j\omega n}\right| \qquad (B.18)$$

For simplicity, suppose that the sequence has zero phase so that it is completely described by (B.18). Suppose that the magnitude of the Fourier transform is nonzero in the interval $-\pi<\omega<\pi$. Then the log magnitude exists, and is a periodic function of the variable ω with a period of 2π. Also, the log magnitude is an even function. The log magnitude can therefore be expressed as a cosine series. The constant term will be ignored here, for it merely represents the contribution of a flat spectrum, which could be caused by a large sample at n=0 in the time domain. The cosine series is

$$\ln|X(\omega)| = \sum_{n=1}^{\infty} a_n \cos(n\omega) \;;\; -\pi < \omega \leq \pi \qquad (B.19)$$

where the coefficients of the series are related to the cepstrum by

$$a_n = \hat{x}(n) + \hat{x}(-n). \qquad (B.20)$$

The coefficients of the Fourier cosine series are directly

related to the representation of the signal as a convolution of Bessel function sequences,

$$x(n) = \bigast_{k=0}^{\infty} \left[I_n^k(a_{k+1}) \right] \qquad (B.21)$$

where the a_i's are the coefficients of the cosine series in Equation (B.19).

Now suppose that the signal has a flat magnitude spectrum. The phase spectrum can be treated in a similar manner. Let the phase spectrum be expressed as a sine series:

$$\arg(X(\omega)) = \sum_{n=1}^{\infty} b_n \sin(n\omega). \qquad (B.22)$$

Note that this is not always possible for any arbitrary sequence x(n). However, if the phase cannot be expressed as a Fourier sine series, then the signal can be negated (phase shifted by 180 degrees) so that the phase is an odd function of ω with a period of 2π. A signal with a flat magnitude spectrum, but with a phase which is expressible as a Fourier sine series can then be written as the convolution of Bessel function sequences of the first kind:

$$x(n) = \bigast_{k=0}^{\infty} \left[J_n^k(-b_{k+1}) \right] \qquad (B.23)$$

where the b_i's are the coefficients of the series (B.22).

Summarizing these results, let $\phi(w)$ be the phase spectrum of x(n), and let $\ln|X(\omega)|$ be the magnitude spectrum. Suppose that these functions are expressible as series:

$$\ln|X(\omega)| = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega)$$

and (B.24)

$$\emptyset(X(\omega)) = \sum_{n=1}^{\infty} b_n \sin(n\omega).$$

Then x(n) can be written as an infinite convolution of Bessel function sequences, with the arguments of the Bessel functions expressed in terms of the cosine and sine series coefficients:

$$x(n) = e^{a_0}\delta(n) * \overset{\infty}{\underset{k=0}{\ast}} \left[I_n^k(a_{k+1})\right] * \overset{\infty}{\underset{k=0}{\ast}} \left[J_n^k(-b_{k+1})\right]. \quad (B.25)$$

Experimental Verification

As an experimental verification of some of these formulae, a sequence was synthesized by the direct convolution of Bessel function sequences. The log magnitude of the sequence was chosen as sketched in Figure 57. The phase was set to zero, so that the corresponding time domain sequence should be an even function. Expanding the magnitude spectrum into a cosine series, the coefficients are as shown in Equation (B.26):

$$\ln|F(\omega)| = a_0 + \sum_{n=1}^{\infty} a_n \cos(\omega n) \quad (B.26)$$

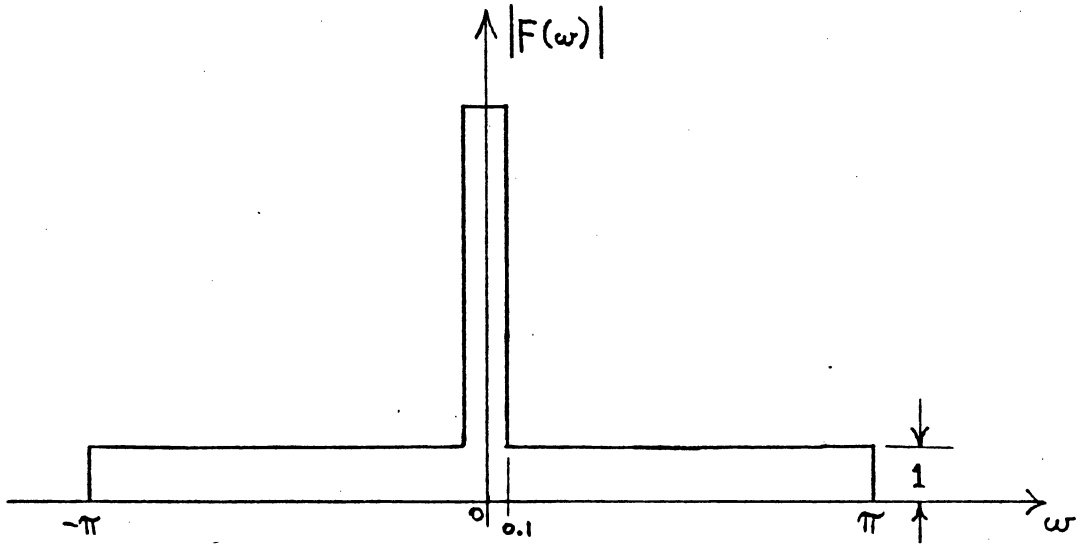


Figure 57. Magnitude spectrum of test sequence.

where $a_0 = \frac{A\omega_0}{\pi}$ and $a_n = \frac{A}{\pi n} \sin(\omega_0 n)$ for n=1,2,.... The value of ω_0 was arbitrarily chosen as 0.1. The time domain sequence was synthesized by a direct convolution of Bessel function sequences, as shown in Equation (B.21). Of course, only a finite number of such functions can actually be convolved. Figures 58-63 show the effect of including more sequences in the convolution. The function converged to x(n) slowly, for this is a sort of "worst case" for a Fourier series: There are step discontinuities at ω=-0.1 and ω=0.1, so that there are many significant coefficients in (B.26). On the practical side, it should be noted that the sequences were actually truncated when $|n|>|a|$ to avoid numerical instability.

Summary

An unusual relationship between the complex cepstrum of a sequence and the Bessel functions has been presented. The cepstrum representation is equivalent to the representation of a log power spectrum by a cosine series with the phase spectrum represented as a sine series. Unfortunately, this simple explanation of the cepstrum is usually obnubilated when the cepstrum is defined as $\hat{x}(n) = Z^{-1}\{\ln[Z\{x(n)\}]\}$. As of this writing it is not believed that the representation as a convolution of Bessel function sequences is very practical if used directly as a synthesis method. But as a tool for theoretical work or derivations it may be useful.

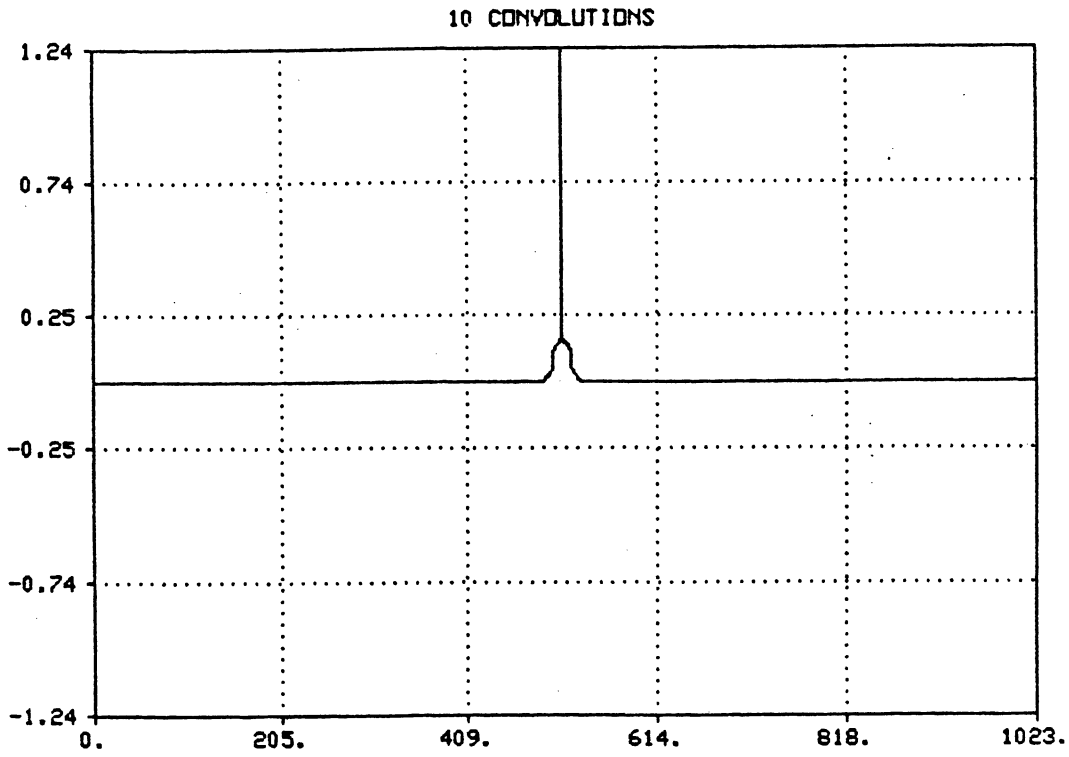


Figure 58. Resynthesized sequence, ten convolutions.

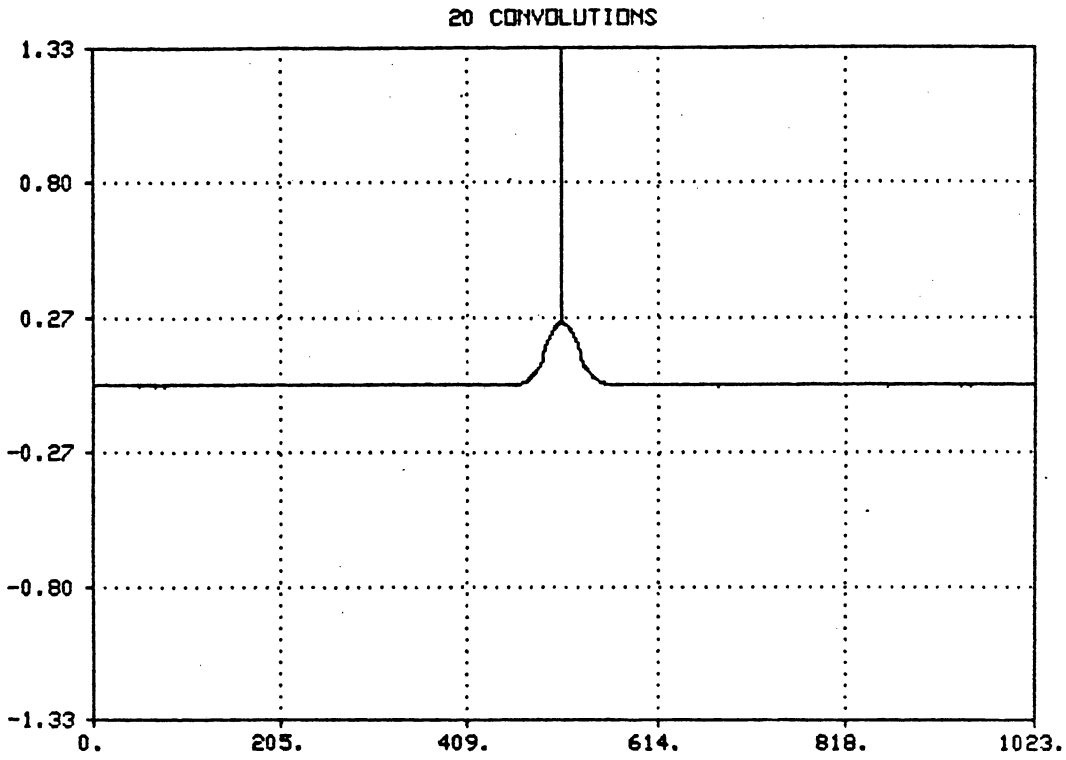


Figure 59. Resynthesized sequence, 20 convolutions.

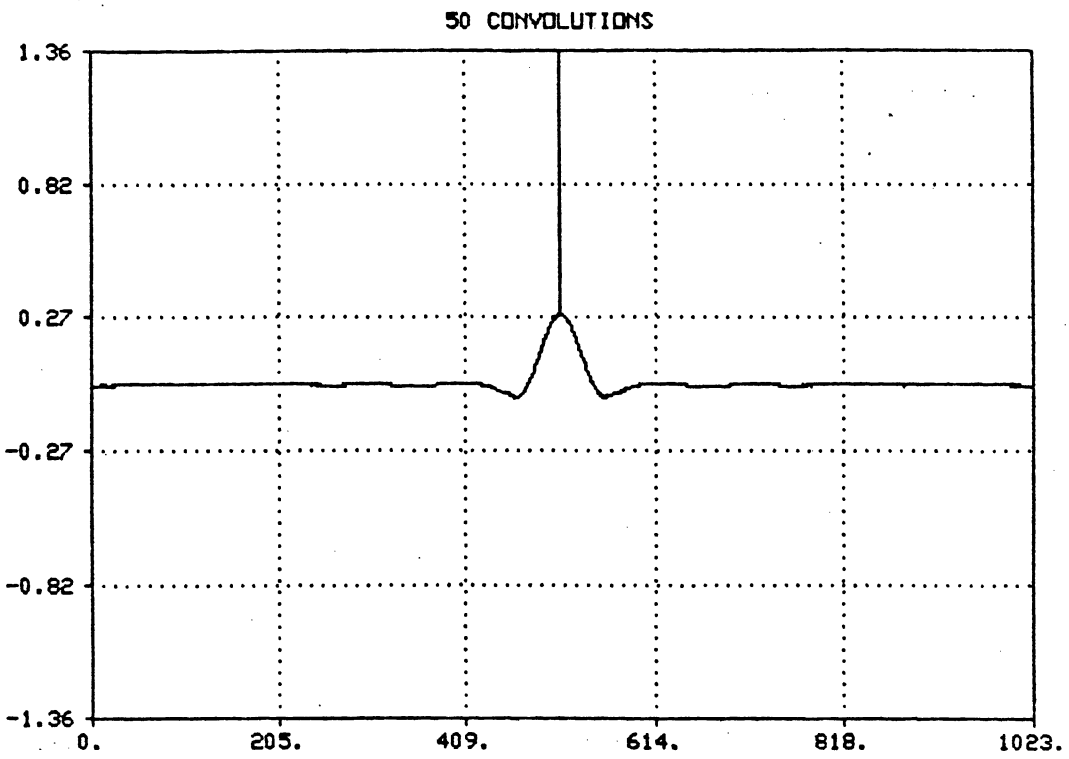


Figure 60. Resynthesized sequence, 50 convolutions.

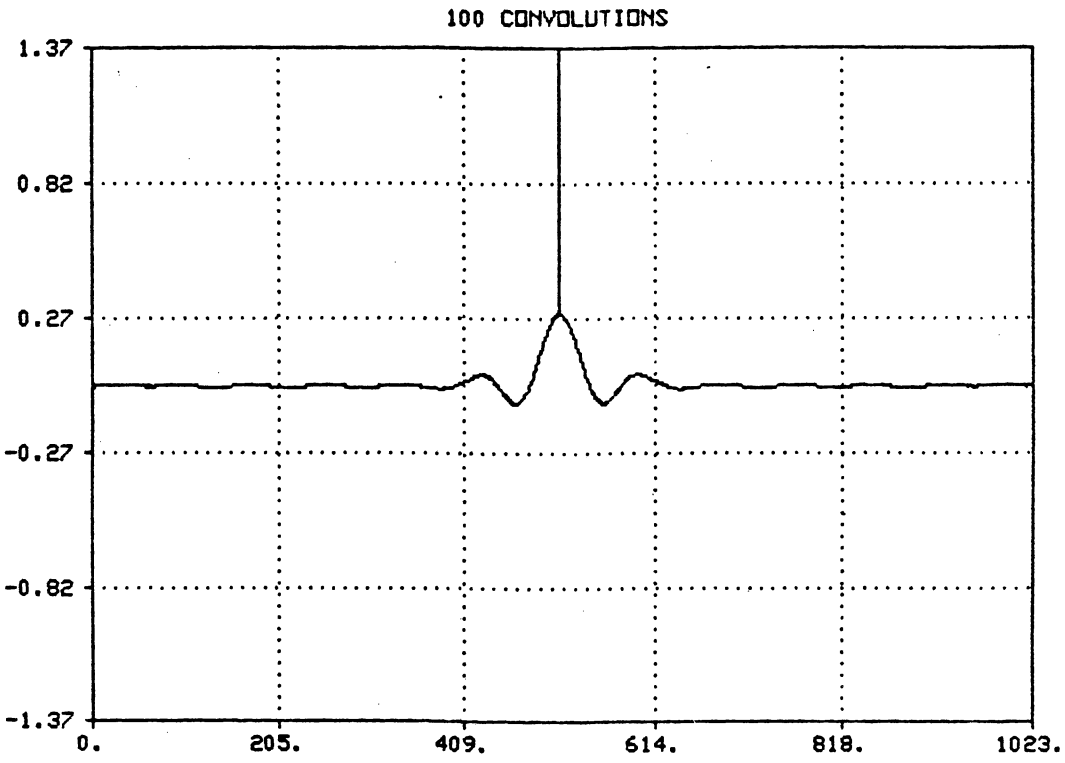


Figure 61. Resynthesized sequence, 100 convolutions.

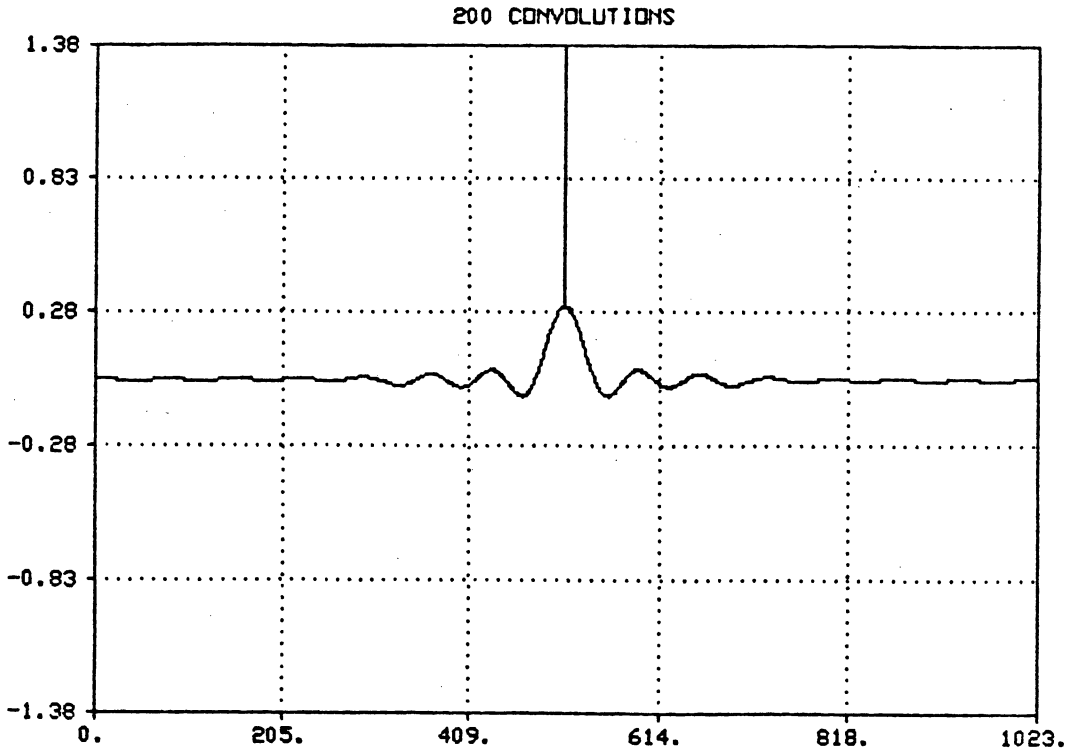


Figure 62. Resynthesized sequence, 200 convolutions.

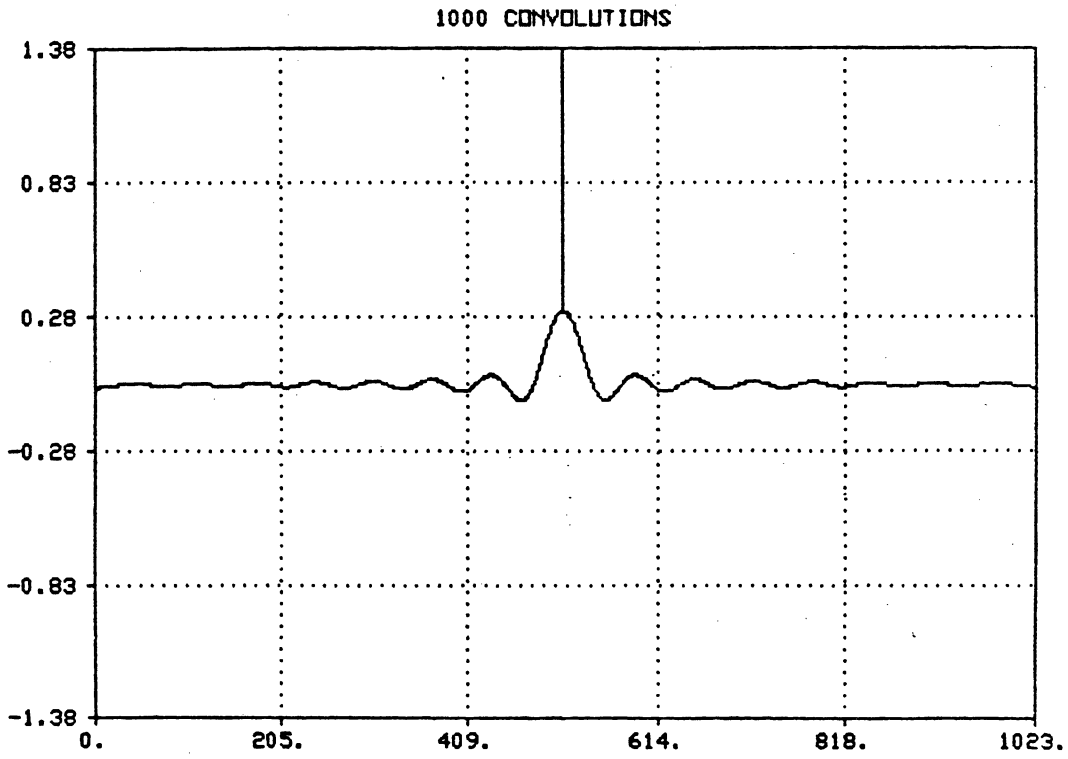


Figure 63. Resynthesized sequence, 1000 convolutions. The result is a sinc(.) function, with an added impulse at the origin (n=512) as there should be.

VITA

Robert Gail Hamilton

Candidate for the Degree of

Master of Science

Thesis: SPEECH ANALYSIS WITH BESSEL FUNCTIONS

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Lawton, Oklahoma, September 19, 1957, the son of William E. and Dorthy E. Hamilton. Married to Janice K. Dague on October 4, 1980.

Education: Graduated from Edmond Memorial High School, Edmond, Oklahoma, in May 1975; Received Bachelor of Science Degree in Electrical Engineering from Oklahoma State University in December, 1985; Completed the requirements for the Degree of Master of Science in Electrical Engineering from Oklahoma State University in May, 1987.

Professional Experience: Electronics Technician for U.S. Navy from August 1975 to March 1982. Holder of commercial radiotelephone license from the Federal Communications Commission. Research Assistant, School of Electrical Engineering, Oklahoma State University from December 1985 to present.