DESIGN OF MULTIDIMENSIONAL DIGITAL FILTERS BY SPECTRAL TRANSFORMATIONS

By

ROBERT KING

March, 1978

A Thesis submitted for the degree of
Doctor of Philosophy of the University
of London

Department of Electrical Engineering
Imperial College of Science and Technology
London S.W.7

To
Gillian

You have the right to work, but for the work's sake only. You have no right to the fruits of work. Desire for the fruits of work must never be your motive in working.

Bhagavad Gita
The Yoga of Knowledge

ABSTRACT

The thesis consists of a survey of the classical methods of two-dimensional digital filter design in the space domain and their extension to multidimensional systems.

Design techniques in the frequency domain are studied with particular reference to techniques involving spectral transformation methods between one and many dimensions. Some of the more recent methods are extended to n dimensions and the limitations of the transformations studied. Specific numerical design examples are given for three-dimensional filter specifications having approximately spherical symmetry.

New design techniques are proposed for the realization of two-dimensional fan filters of recursive form having guaranteed stability. The techniques are shown to be extendable to three-dimensional systems, in which two dimensions are linear and one is temporal.

A critical comparison is made of the several techniques proposed.

ACKNOWLEDGEMENTS

> No man is an Iland, intire of it selfe; every man is a peece of the Continent, a part of the maine; if a clod bee washed away by the Sea, Europe is the lesse, as well as if a Promontorie were ..... Any mans death diminishes me because I am involved in Mankinde. And therefore never send to know for whom the bell tolls; it tolls for thee.
>
> John Donne

I should like to thank Professor Colin Cherry for his encouragement during the conduct of this research and for creating an atmosphere in the Communication Section in which research can flourish.

Of the many who have contributed either directly or indirectly to this work, I should like to thank my colleagues on the staff of the Electrical Engineering Department, in particular Dr. A.G. Constantinides and also the many research students who have worked with me under my supervision over the past years.

Among my past research students I cannot fail to mention Dr. Majid Ahmadi. His friendship, stimulation, advice and constructive criticism contributed in no small measure to this thesis and without his encouragement I doubt if this work would have materialized.

Little of the computational work involved in the research would have been possible without the assistance of Mr. Yusif Fakhro who was always willing to give his time to advise me.

Finally my thanks are due to Mrs Shelagh Murdock for typing the manuscript and presenting the thesis in a most elegant form.

LIST OF SYMBOLS

$\sum_{i=1}^{n} a_i$	represents	$a_1 + a_2 + \dots + a_n$
$\prod_{i=1}^{n} a_i$	represents	$a_1 a_2 \dots a_n$
$\in$	represents	"is a member of"
$\cap$	represents	the intersection of two sets.
$\bigcap_{i=1}^{n} (\lvert z_i \rvert = 1)$	represents	$(\lvert z_1 \rvert = 1) \cap (\lvert z_2 \rvert = 1) \cap \dots \cap (\lvert z_n \rvert = 1)$
$\forall$	represents	"for all"

CONTENTS

		Page
Abstract		2
Acknowledgements		3
List of Symbols		4
Chapter 1	Introduction	8
1.1	Preamble	9
1.2	Representation of multidimensional signals	12
1.3	Classification and representation of multidimensional digital filters	13
1.4	Historical background	16
1.5	Outline of thesis	19
Chapter 2	Stability and Stabilization	21
2.1	Definition of stability	22
2.2	Shanks' stability theorem	24
	2.2.1 Alternative stability formulation	25
	2.2.2 Modification of Shanks' theorem	25
2.3	Huang's stability test	27
2.4	Ansell's stability test	28
2.5	Anderson and Jury stability test	29
2.6	Maria and Fahmy method	30
2.7	Comparison of stability tests	31
2.8	Stabilization techniques	32
	2.8.1 Shanks' stabilization method	32
	2.8.2 Reid and Treitel stabilization technique	36
	2.8.3 Pistor stabilization technique	40
	2.8.4 Review of stabilization techniques	45
Chapter 3	Spatial design techniques	46
3.0	Introduction	47
3.1	Kalman technique	47
3.2	Shanks' method	49
3.3	Bordner synthesis technique	52
3.4	Bertram's design technique	55
3.5	Lal's design technique	56
3.6	Criticism of design techniques in the space domain	58

Page

Chapter 4 Spectral Transformation 60

4.1 Introduction 61
4.2 One dimension to one dimension transformations 64
4.3 Two dimension to two dimension transformations 66
4.4 One dimension to two dimension transformations 68
4.4.1 Separable product 69
4.4.2 Shanks' rotated filters 70
4.4.3 Costa and Venetsanopoulos modification 73
4.4.4 McClellan transformation 74
4.4.5 Bernabo design technique 76
4.4.6 Ahmadi transformation 79
4.4.7 All-pass transformation 88
4.4.8 Comparison of two-dimensional transformations 93
4.5 One dimension to multidimension transformations 93
4.5.1 Separable product 96
4.5.2 Multidimensional McClellan transformation 98
4.5.3 Multidimensional Bernabo design technique 100
4.5.4 Multidimensional Ahmadi technique 102
4.5.5 Multidimensional all-pass transformation 106
4.5.6 Comments on multidimensional transformations 110
4.6 Application of two-dimensional circular profile filter 111

Chapter 5 Fan Filter Design 117

5.1 Introduction 118
5.2 Formulation of design problem 119
5.3 Fan filter spectral transformation 122
5.3.1 Derivation of the fan transformation 124
5.3.2 The contour approximation 129
5.3.3 Stability and stabilization 130
5.3.4 Design examples 130
5.4 High pass and multiple pass band filters 145
5.5 Multidimensional fan filters 155
5.5.1 Circular cone filters 157
5.5.2 Rectangular pyramid filters 159
5.6 Discussion of fan filters 161

Page

Chapter 6 Review, Criticism and Conclusions 162

6.0 Review 163

6.1 Stability 167

6.2 Spatial design techniques 170

6.3 Frequency transformation techniques 172

6.4 Future fields of work 175

6.5 Conclusion 176

References 177

CHAPTER ONE

INTRODUCTION

For just as a man cannot see without
eyes, so a scholar would be blind
unless he learnt from books.

"Piers the Ploughman"
William Langland

INTRODUCTION

1.1 PREAMBLE

One of the significant problems in communication is the processing of signals which have been passed through a system which has resulted in a deterioration in their quality. The need for such processing resulted, in the early days, from the distortion produced in telephone links over relatively long distances. The whole theory of analogue signal processing developed from this need and led to the classical Zobel filters and other passive network designs and later developed into the more modern aspects of active network synthesis.

More recently, digital signal processing has been introduced for the treatment either of analogue signals which have been sampled periodically to produce a set of discrete pulses or of signals which from transmission to reception are in discrete form. Such signals may be considered to be represented by a sequence of numbers representing the value of the signal at successive instants of time. Once this conceptual approach to time sequences has been accepted it may be realized that digital signal processing may be considered in the simpler manner as a means whereby an array of numbers may be modified according to some selected laws to generate an output array; these arrays may be of one or more dimensions. In one dimension the most common variable is time; in multidimensional systems the variables may be spatial, temporal or any other desired parameter. The discrete data processor usually operates on the input array sequentially; when the data is spatially or otherwise distributed, it will need to be scanned in time before processing. This may be accomplished in real time or, alternatively, the data may be stored and processed at leisure by a

relatively slower system. Occasionally a parallel processor may be used but this is generally uneconomical for large systems.

It is now shown that the data available for processing may vary in two or more dimensions. Examples of two-dimensional arrays are found in facsimile and other visual images, electron micrographs, X-ray images, isotope scanning images where the two dimensions are both spatial. Other two-dimensional arrays occur in the field of seismic exploration where the output array is displayed in the two dimensions of linear space and time.

Such arrays need to be processed in order to remove anomalous signals and also to enhance certain aspects of the data. For example in medicine, X-ray films may be processed to remove low spatial frequency variations, thereby enhancing the sudden variations of abnormal conditions. In the case of scanned isotope detection the processor may be required to remove the striations on the image produced by the finite width of the scanning lines and other distortion which occurs as a result of the finite resolution of the gamma ray scanner [1]. In electron microscopy the purpose may be to reduce the low frequency background noise which is inherent in such processes. In the field of geophysical prospecting echoes from boundaries of geological strata of a detonation are detected by a linear array of detectors placed in line with the source. Desired echoes will be received by the detectors from various changes in geological strata, whereas undesired echoes may occur from multiple reflections, and random signals may also be generated by wind noise [2,3].

It is conceptually simple to appreciate that arrays are now no longer restricted to two dimensions but may be extended to any

number of dimensions. An example of a three-dimensional array might be the output from a rectangular grid of detectors placed near to a detonation, the three coordinate axes being the two spatial dimensions of the grid and the time of reception of reflection at each detector. Again the processing of images obtained by tomographic scanning of the human body in detecting tumours will also result in a three-dimensional array. A television picture is another example of a three-dimensional signal with two spatial and one temporal dimension.

Although practical examples of higher order arrays are not available at present, it is convenient to consider multidimensional processes as opposed to the restricted two-dimensional arrays.

The processing of such multidimensional arrays poses certain problems which do not apply to one-dimensional arrays. In principle it is possible to consider a "parallel" process whereby the output for every point of an array is produced at the same instant of time by a very large number of identical processors connected to the appropriate input array data. This conceptual approach is the one most frequently adopted in the development of any theoretical work. However, the design of a system based on this concept would be absurd as each of the parallel processors would be identical. It is therefore customary to carry out the processing on each group of input data, giving one element in the output data, in a sequential manner in time using a single processor.

This idea of time scanning of space data introduces a fundamental difference between time and space series. In any two sets of time series representing cause and response, the response $r(t_o)$ at a certain time t_o can only be dependent on values of the cause $c(t)$ at

values of $t \leqslant t_0$. This is known as the principle of causality. In spatial arrays such a distinction does not exist and the response $r(x_0)$ for a one-dimensional array may depend on all values of the cause $c(x)$ for $-\infty < x < \infty$ or more rigorously for $A < x < B$ where A and B define the physical bounds of the input array. This distinction results in the inability of a time-scanned processor effectively to implement all the requirements of a digital filter operating on spatial data in one single run of a time scanner in one direction. This will be considered in greater detail in Section 1.2.

1.2 REPRESENTATION OF MULTIDIMENSIONAL SIGNALS

A multidimensional array may be represented by,

$$h = \{h(m_1,\ldots,m_n)\} \equiv \{h(\underline{m})\} \tag{1.1}$$

$$= \{h_{m_1,\ldots,m_n}\} \equiv \{h_{\underline{m}}\} \tag{1.2}$$

where $\{\underline{m}\}$ is defined as the multiple $\{m_1,\ldots,m_n\}$ and $h_{\underline{m}}$ is assumed to be zero for all subscript multiples $\{\underline{m}\}$ which do not belong to the definition set of h. The definition set of h usually corresponds to the set bounded by the physical limitations of the given array. Such an array is in general defined for positive and negative values of the elements of the subscript sets.

In the particular case in which $h_{\underline{m}} = 0$ for all subscript sets $\{\underline{m}\}$ in which any $m_i < 0$ the array is known as a first quadrant array and represented by a superscript 1 as

$$^1h = \{^1h_{\underline{m}}\} = \{^1h_{m_1,\ldots,m_n}\} \tag{1.3}$$

having non-zero values only for $\bigcap_{i=1}^{n} m_i \geqslant 0$.

Other single quadrant arrays

$$p_h = \{p_{h_{\underline{m}}}\} \tag{1.4}$$

may be defined which are non-zero only when some or all of the elements of the multiple $\{\underline{m}\}$ are negative. There are thus 2^n single quadrant arrays possible in any n-dimensional system. The definition of a first quadrant array given by equation (1.3), when considered in one dimension, is identical with the definition of a causal array in time and thus multidimensional arrays having this property are sometimes termed causal. Since causality is meaningless outside the time dimension, I shall retain the term 'first quadrant array'.

A digital array may alternatively be represented by its Z-transform, defined as

$$\begin{aligned} H(z_1,\ldots,z_n) &\equiv H(\underline{z}) \\ &= \sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_n=-\infty}^{\infty} h_{m_1,\ldots,m_n} z_1^{m_1} \cdots z_n^{m_n} \end{aligned}$$

1.3 CLASSIFICATION AND REPRESENTATION OF MULTIDIMENSIONAL DIGITAL FILTERS

Digital filters may be subdivided into two classes, linear and non-linear. A typical example of a one-dimensional non-linear filter is the adaptive or automatic equalizer in digital communication systems. In two dimensions we may quote contrast enhancement of images as a typical example.

Most digital filters at present are linear systems; for example, those used in the equalization of fixed digital communication

channels, filters used in the synthesis of speech and similar applications. This thesis will be entirely concerned with finite linear digital filters.

Linear filters may be divided into two classes. First are non-recursive filters in which the output array is a function of the input array only; such filters will have a response to an input impulse which is bounded in duration in any dimension; it is thus termed a finite impulse response (FIR) filter. Second, recursive filters are those in which the output array is dependent, not only upon the input array, but also upon neighbouring values of the output array; such filters usually have unbounded duration of their impulse response arrays in all dimensions and are thus known as infinite impulse response (IIR) filters.

The output array $\{r_{m_1,\ldots,m_n}\}$ of a non-recursive filter is related [4,5] to the input array $\{c_{m_1,\ldots,n_1}\}$ by the convolutional equation

$$r_{m_1,\ldots,m_n} = \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} a_{j_1,\ldots,j_n} \cdot c_{m_1-j_1,\ldots,m_n-j_n}$$

Similarly recursive filters may be represented by

$$r_{m_1,\ldots,m_n} = \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} a_{j_1,\ldots,j_n} \cdot c_{m_1-j_1,\ldots,m_n-j_n} - \sum_{\substack{j_1=0 \\ j_1=\cdots=j_i=\cdots=j_n\neq 0}}^{M_1} \cdots \sum_{j_n=0}^{M_n} b_{j_1,\ldots,j_n} \cdot r_{m_1-j_1,\ldots,m_n-j_n}.$$

In the one-dimensional case where the variable is time, the neighbouring values of the output must always be previous values, whereas in spatial systems they may be to the right or left, above or below, etc. Thus any scanning system in time, used to recursively process spatially distributed data will be limited to generating output functions which are constrained in only one 'quadrant' of multidimensional space.

The above digital filters may alternatively be represented by the Z-transform relationships from which the Z-transfer functions may be obtained.

For the non-recursive filter, the Z-transfer function is

$$H(z_1,\ldots,z_n) = \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} a_{j_1,\ldots,j_n} . z_1^{j_1} \ldots . z_n^{j_n}$$

For the recursive filter

$$H(z_1,\ldots,z_n) = \frac{A(z_1,\ldots,z_n)}{B(z_1,\ldots,z_n)}$$

$$= \frac{\sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} a_{j_1,\ldots,j_n} . z_1^{j_1} \ldots , z_n^{j_n}}{\sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} b_{j_1,\ldots,j_n} . z_1^{j_1} \ldots . z_n^{j_n}}$$

1.3.1 Zero Phase Filters

The spectrum of a filter $H(e^{ju_1},\ldots,e^{ju_n})$ may be determined from the Z-transform of the transfer function by setting $z_i = e^{ju_i}$. A zero phase filter is one in which

$$\text{Arg } H(e^{ju_1},\ldots,e^{ju_n}) = 0 \text{ for all values of } u_i.$$

It may be easily shown that this leads to the condition

$$\left|H(e^{ju_1},\ldots,e^{ju_n})\right| = \left|H(e^{-ju_1},e^{ju_2},\ldots,e^{ju_n})\right|$$
$$= \left|H(e^{ju_1},e^{-ju_2},\ldots,e^{ju_n})\right| = \ldots.$$
$$= \left|H(e^{-ju_1},e^{-ju_2},\ldots,e^{-ju_n})\right|$$

namely that the magnitude spectrum is symmetrical about all frequency axes.

1.4 HISTORICAL BACKGROUND

The earliest work on digital filters was performed on FIR filters using direct convolution of the filter coefficients and the input sequence. This was computationally inefficient and rendered design difficult since it needed to be carried out in the time domain.

Fourier transform methods allowed simpler design techniques to be evolved, as the only requirement was the specification of a set of weighting coefficients in the frequency domain. The advent of the Fast Fourier Transform [6] (FFT) enabled computer efficiency to be spectacularly improved. It suffered, however, from the restriction that it was limited to relatively small arrays of data if computer storage were not to become excessive; this limitation is of even greater significance when processing multidimensional arrays.

Many of the applications of two-dimensional filters require that the point spread functions in the space domain should be circularly symmetric, as any distortion is equally likely to occur in any radial direction. This results in a frequency response classificistic which is circularly symmetric about the origin. Very similar

requirements are likely to be demanded in multidimensional systems, in particular where all the 'spatial' dimensions are of the same nature.

Other situations arise in which the distortion of the desired signal has not been produced by a homogeneous medium and therefore any filter designed will similarly need to have a different response along each of the coordinate axes. An example might be the removal, or reduction, of scan lines in facsimile or television pictures. This results in frequency responses which are also not identical along the various axes. Another situation where this may occur is that in which the various axes do not represent the same type of variable, for example, in processing a television image, two axes are spatial and one temporal. In the processing of geophysical data frequently one (or two) axes are spatial and the third temporal; a classic example is the "fan filter" [2,24] .

Most of the earliest work was on the design of circularly symmetric filters (or those closely approximating that ideal). One of the earliest papers was by Darby and Davies [7] who used the two-dimensional Fourier transform to generate the impulse response of a filter specified in the frequency domain and then to process this by convolution techniques. Huang [8] extended this by investigating the use of windowing techniques of two-dimensional arrays.

McLellan [9] initiated the concept of using a one dimension to two dimensions transformation to generate a two-dimensional filter from a one-dimensional prototype.

An alternative approach was used by Merserau and Dudgeon [4] who proposed a method for the representation of two-dimensional arrays as one-dimensional sequences, in which a two-dimensional pass

or stop band filter was transformed into a multiple pass band one-dimensional design. The resulting designs have not been very encouraging.

The desire, in the one-dimensional case, to process data in real time and with small computers led to attempts to reduce the complexity of implementation by recourse to the use of recursive filters. A great deal of work has been published on the design of recursive filters in both the spatial and frequency domains. In the field of space domain synthesis we may mention the work of Kalman in 1958 [34] ; Steiglitz and McBridie, 1965 [35] ; Shanks, 1967 [38] ; Bordner, 1974 [20] ; Bertram, 1975 [39] , and Lal, 1975 [41] . The earliest of these techniques was the separable product technique suggested by Hall in 1970 [15] . This was followed by the design of Shanks et al, in 1972 [5] in which one-dimensional filters were rotated to give an approximation to a desired cut-off boundary. Costa and Venetsanopoulos, 1974, further improved this technique [16] . More recently Bernabo et al, in 1975, modified the transformation of McClellan to apply to recursive two-dimensional filters [17] .

One of the greatest problems in the design of a recursive filter is the assessment of the stability of the filter and the modifications to be applied to the transfer function to rectify any observed instability.

The determination of the stability of a two-dimensional digital filter was first studied in 1972 by Shanks [5] and Huang [10] who independently published effectively equivalent tests for the stability of two-dimensional recursive digital systems. Anderson and Jury [11] and Maria and Fahmy [12] further contributed to the work although the computational effort is still considerable.

Methods of correcting unstable transfer characteristics have been proposed for two-dimensional systems by Shanks [5] , who based his method on an unproved conjecture which has since been shown to be invalid. Reid and Treitel [53] put forward an alternative technique based on the well-known properties of the Hilbert transform. Unfortunately neither of these methods has been found to work in all cases and it is still a matter of conjecture whether this is due to a fundamental theoretical limitation or is due to approximations made in the computational implementation of the methods.

A technique proposed by Pistor [13] and expanded by Ekstrom [14] not only detects the stability or otherwise of a filter, but at the same time effects a stabilizing routine where needed. It has been suggested by Pistor that the technique, which works well with zero-phase functions, may also be applied in general. This has not been justified and examples have again provided inconclusive evidence as to the reasons for apparent failure of the technique in such cases. For zero-phase networks, the partitioning into single quadrant stable functions can be carried out in all cases if sufficient accuracy in computation is demanded.

1.5 OUTLINE OF THE THESIS

The thesis will first review the problem of stability and stabilization of multidimensional digital filters. It will then consider the problem of the design of multidimensional filters, particularly those whose responses approximate circular symmetry in the space domain. This will be followed by a review of frequency domain design techniques, introducing an extension of the Ahmadi [22] method

to n dimensions and also referring to a new method by Kap [23] which may also be extended to multidimensional systems.

Subsequent work refers to the design of two-dimensional filters with fan-shaped cut-off boundaries, giving particular emphasis to a transformation technique from one-dimensional low pass filters to two-dimensional fan filters. This method is compared with earlier techniques for fan filter design and examples designed by the technique are presented.

CHAPTER TWO

STABILITY AND STABILIZATION

To gain sagacity our mind must be trained on the very problems that other men have already solved, and it must methodically examine even the most trivial of human devices, but especially those which manifest or imply an orderly arrangement.

"Regulae ad directionem ingenii"

Descartes

STABILITY AND STABILIZATION

2.1 DEFINITIONS OF STABILITY

The most commonly used definition of stability is bounded input, bounded output (BIBO) stability; a system is defined as being BIBO stable if the output is bounded in response to a bounded input. This has been studied in considerable detail in continuous and one-dimensional systems and has more recently been extended to two- and multidimensional systems.

Consider the multidimensional input array $\{c_{m_1,\ldots m_n}\}$. This is an absolutely bounded array if

$$|c_{m_1,\ldots,m_n}| \leqslant P < \infty \tag{2.1}$$

and the array is absolutely summable if

$$\sum_{m_1=0}^{M_1} \cdots \sum_{m_n=0}^{M_n} |c_{m_1,\ldots,m_n}| \leqslant Q < \infty \tag{2.2}$$

where P and Q are positive real numbers.

We may develop the conditions for an n-dimensional system to be BIBO stable as follows. The output array, $\{r_{m_1,\ldots,m_n}\}$, is given by the convolution of the input array, $\{c_{m_1,\ldots,m_n}\}$, with the impulse response array of the filter, $\{a_{j_1,\ldots,j_n}\}$. Namely

$$r_{m_1,\ldots,m_n} = \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} a_{j_1,\ldots,j_n} \cdot c_{m_1-j_1,\ldots,m_n-j_n} \tag{2.3}$$

Application of Schwarz's inequality leads to

$$|r_{m_1,\ldots,m_n}| \leqslant \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} |a_{j_1,\ldots,j_n}| \cdot |c_{m_1-j_1,\ldots,m_n-j_n}| \quad (2.4)$$

and utilising the bounded nature of the input given by (2.1) results in

$$|r_{m_1,\ldots,m_n}| \leqslant P \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} |a_{j_1,\ldots,j_n}| \quad (2.5)$$

Comparison of (2.5) with (2.2), for a BIBO stable system, shows that $\{a_{j_1,\ldots,j_n}\}$ must be an absolutely summable array. We have, above, shown that this is a necessary condition for BIBO stability. It has also been shown by Farmer and Bedner to be a sufficient condition [25].

Thus a necessary and sufficient condition for a network to be BIBO stable is that its impulse response shall be absolutely summable.

An alternative, but less familiar, form of stability is that in which we require the output sequence to be absolutely summable if the input sequence is absolutely summable (SISO stability). By a similar application of Schwarz's inequality to the convolution equation (2.3) it may be shown that a necessary but not always sufficient condition is that the impulse response of the system shall be absolutely summable.

Although the above conditions are basic to the definition of stability, they are of little value in assessing the stability of a specified network or system. For this it is simpler to operate in the frequency or z domain. The stability of a non-recursive filter, however, can only be specified in terms of the absolute summability of the coefficients of its Z-transform, which is a direct application of the above criterion [26].

2.2 SHANKS' STABILITY THEOREM

For a recursive filter it has been shown by Shanks [5,26] that the stability of a system is controlled entirely by the properties of the denominator of the transfer function. The conditions imposed on the denominator function $B(z_1,\ldots,z_n)$ in order to be assured of stability are that

$$B(z_1,\ldots,z_n) \neq 0 \quad \forall \quad (z_i,\ i=1,\ldots,n) \in D \tag{2.6}$$

where $\quad D = \left\{(z_i,\ i=1,\ldots,n) : \bigcap_{i=1}^{n} |z_i| \leqslant 1\right\}$

Application of this theorem in one dimension is relatively straightforward since the fundamental theorem of algebra states that every real polynomial in a single variable may be factorized into real linear and quadratic factors and thus the location of the roots of the denominator may be obtained, if necessary in high-order systems, to any desired accuracy, by a computer algorithm.

In two or more dimensions the fundamental theorem does not hold. In fact it may be shown that in the general case, it is not possible to factorize a multivariable polynomial into first and second order factors. The stability problem thus devolves into a determination of the continuum of values of $z_1,z_2,\ldots,z_n$ for which $B(z_1,\ldots,z_n) = 0$ and checking whether they lie within the domain D.

Shanks' approach was to define an infinite impulse response convolution filter, $g_{m_1,\ldots,m_n}$, having a Z-transform

$$G(z_1,\ldots,z_n) = 1/B(z_1,\ldots,z_n)$$

He then showed that the stability condition

$$B(z_1,\ldots,z_n) \neq 0 \quad \forall (z_i,\ i=1,\ldots,n) \in D$$

is identical to the condition that $G(z_1,\ldots,z_n)$ shall be convolutionally stable, i.e. that there exists a stable filter, g, such that convolution of g with b shall yield the multidimensional impulse, δ ; thus

$$g * b = \delta \tag{2.7}$$

2.2.1 Alternative Stability Formulation

Anderson and Jury [27] have proposed an alternative formulation of the stability criterion of equation (2.6). This states that a system is stable if and only if (iff)

$$\begin{aligned}
&B(z_1,0,\ldots,0) \neq 0 \quad \forall\ |z_1| \leqslant 1 \\
&B(z_1,z_2,\ldots,0) \neq 0 \quad \forall\ |z_1| = 1 \cap |z_2| \leqslant 1 \\
&\cdot \\
&\cdot \\
&\cdot \\
&B(z_1,z_2,\ldots,z_n) \neq 0 \quad \forall\ (\bigcap_{i=1}^{n-1} |z_i| = 1) \cap |z_n| \leqslant 1
\end{aligned} \tag{2.8}$$

This test may be formalized and applied by using the technique of Anderson and Jury [11,27] .

2.2.2 Modification of Shanks' Theorem

It has recently been shown by Goodman [28] that the necessity of Shanks' theorem fails under certain conditions in which the transfer function numerator as well as denominator are multivariable polynomials. He has shown this for two-variable functions but the limitation is also relevant to multidimensional systems.

A two-dimensional polynomial, although not factorizable into first and second order factors, may be factorized into a set of unique irreducible real polynomials. There may be a number of points, (z_1,z_2), at which the demominator polynomial, $B(z_1,z_2)$, is zero and it is these which control the stability. In the majority of cases the numerator, $A(z_1,z_2) \neq 0$, and the point (z_1,z_2) is termed a pole. However, in some cases $A(z_1,z_2) = 0$ also; such a point is then termed a non-essential singularity of the second kind. The existence of such points modifies Shanks' theorem and may show a function to be stable despite the existence of points at which the denominator polynomial vanishes.

A modified function $F(z_1,z_2)$ will represent a stable system if $F(z_1,z_2)$ has no poles in $D_{12} = \{(z_1,z_2): |z_1| \leqslant 1 \cap |z_2| \leqslant 1\}$ and no essential singularities of the second kind, except possibly in $R_{12} = \{(z_1,z_2): |z_1| = 1 \cap |z_2| = 1\}$.

When $B(z_1,z_2) \neq 0$ in $D'_{12} = \{(z_1,z_2): |z_1| < 1 \cap |z_2| < 1\}$ but $F(z_1,z_2)$ has a non-essential singularity of the second kind in R_{12}, it appears that F may or may not be stable. Examples in which either situation may occur have been given by Goodman.

This may be summarized by saying that Shanks' stability theorem is both necessary and sufficient except when essential singularities occur in the domain R_{12}. An extension to n dimensions would suggest that the necessity of Shanks' condition might fail if essential singularities occur on the domain $R_n = \{(z_1,\ldots,z_n): \bigcap_{i=1}^{n} |z_i| = 1\}$.

2.3 HUANG'S STABILITY TEST

Shanks' theorem has proved to be difficult to apply in practice and thus alternative techniques have been proposed for its implementation. Huang [10] put forward a technique which is relatively simple for two-dimensional systems and although it might, in principle, be extended to multiple dimensions the application would be incredibly tedious.

He states that a two-dimensional function

$$H(z_1,z_2) = \frac{A(z_1,z_2)}{B(z_1,z_2)} \tag{2.9}$$

is stable iff

1) the map of $R_1 = (z_1 : |z_1| = 1)$ in the z_2 plane according to the transformation $B(z_1,z_2) = 0$ lies outside the domain $D_2 = (z_2 : |z_2| \leqslant 1)$, and

2) no point in $D_1 = (z_1 : |z_1| \leqslant 1)$ maps into the point $z_2 = 0$ by the relation $B(z_1,z_2) = 0$.

To apply this the unit circle R_1 is mapped into the z_2 plane and checked to see whether it intersects the unit circle in the z_2 plane. In addition the equation $B(z_1,z_2) = 0$ must be solved to check whether the magnitude of any root is less than unity.

Despite the simplification, the required computation is still laborious since it involves testing at an infinite number of points.

Huang showed that the test could be simplified by reduction following a technique due to Ansell [29] which would result in a finite number of steps.

2.4 ANSELL'S STABILITY TEST

Ansell's theorem [29] effectively transforms the filter function $H(z_1,z_2)$ from the z-domain to the s-domain via the two bilinear transformations

$$s_1 = \frac{1 - z_1}{1 + z_1}$$
$$s_2 = \frac{1 - z_2}{1 + z_2} \qquad (2.10)$$

to give $\hat{H}(s_1,s_2) = \dfrac{\hat{A}(s_1,s_2)}{\hat{B}(s_1,s_2)}$

The stability criteria may now be transformed into the s-domain as follows.

A filter $H(z_1,z_2)$ is stable iff

1) in all real finite u_1, the complex polynomial in s_2, $\hat{B}(ju_1, s_2)$, has no zeroes in $Re(s_2) \geqslant 0$, and

2) the real polynomial in s_1, $\hat{B}(s_1, 1)$, has no zeroes in $Re(s_1) \geqslant 0$.

Condition (2) of Ansell's test is relatively simple to apply using existing one-dimensional stability techniques. Condition (1) is, however, more difficult since it involves a study of the roots of a complex polynomial of a complex variable. It may be put into an alternative form by considering the polynomial $B(ju_1,ju_2)$. This may be written as a complex polynomial in u_2 whose coefficients are real and imaginary functions of u_1. A matrix function of u_1 may now be constructed in which the elements are functions of the above complex polynomial.

The first stability condition of Ansell is now satisfied if all the principal minors of this matrix are positive for all real u_1. The technique, albeit tedious, at least comprises a finite number of steps.

It is also apparent that the same technique may be applied to multidimensional polynomials by successive branching techniques, each set of matrices having one variable eliminated. The computational labour precludes the use of such a technique at present for anything more complex than the most trivial examples for which the stability could almost be assessed by inspection.

2.5 ANDERSON AND JURY STABILITY TEST

Anderson and Jury again tackled the significant problem of attempting to bring the stability testing procedure to a simpler form which would facilitate computation. Their technique originates from the formulation of the stability conditions of equation (2.8).

The first of these is relatively simple since it is a function of only one variable and there are a number of tests for determining whether the roots of such a polynomial will lie within unit circle. One such method is by use of the Schur-Cohn [11] matrix; The second is based on the Jury table [31] . The former involves setting up a matrix formed from the coefficients of the function to be tested. In the Schur-Cohn test the positivity of all the eigen values of the constructed matrix is assessed.

In the Jury test a sequence of polynomials is derived from the original polynomial using a simple recurrence equation. The values

of these polynomials at $z = 0$ are computed and their product obtained. The positivity of this quantity is a necessary and sufficient condition for the original polynomial to have all its zeroes outside the unit circle [55] .

The second and subsequent members of equation (2.8) are more difficult to establish.

Considering the second such polynomial which is a function of two variables, z_1, z_2, the test may be carried out in two parts.

First, a Schur-Cohn matrix is constructed considering z_2 as the variable parameter; the elements of the matrix will now be functions of z_1. For stability the Schur-Cohn matrix must be negative definite for all z_1 such that $|z_1| = 1$. This is now a problem involving self inverse polynomials. The procedure is protracted but will ultimately lead to the required assessment.

For third and higher dimensional filters, the same branching technique may be used, but for anything greater than second order filters the size of the matrices becomes so great that they may only be manipulated with great difficulty and by using inordinate computer time. The technique has been well documented for two-dimensional filters.

2.6 MARIA AND FAHMY METHOD

The second and subsequent elements of equation (2.8) have been tested by a technique developed by Maria and Fahmy [12] .

They evolved an extended form of the Jury table which is obtained from the coefficients of the original polynomial by the relatively simple technique of computing a succession of 2 x 2

determinants. The series of initial elements of the Jury table must all be non-negative (except the first, which must be non-positive).

For two-dimensional systems all these coefficients will be functions of the variable z_1; however, much simplification is achieved since the multiplying coefficients are all real and as the test is applied on the boundary of the unit circle where $|z_1| = 1$, we may write $z_1^* = z_1^{-1}$.

For multidimensional systems the branching process outlined in the case of the Anderson and Jury test may be applied, but as no determinants of higher order than 2 need to be determined, the computational effort involved is minimized.

2.7 COMPARISON OF STABILITY TESTS

The Ansell stability test suffers from the difficulty of applying the bilinear transformation before application of relatively simple tests for roots of a function in a bounded space. Certain techniques have been evolved for mechanising the application of the bilinear transformation [32,33] which simplify the computational effort of this method.

The Anderson and Jury test removes this drawback, but the formation of the Schur-Cohn matrix for higher orders is tedious since the order of the matrix is equal to the degree of the denominator of the transfer function.

The Maria and Fahmy method may perhaps be complimented on being the least tedious to implement since all the matrices involved are of second order only.

2.8 STABILIZATION TECHNIQUES

Having obtained a solution to a two-dimensional design problem and determined a multivariable transfer function which satisfies the desired specification, it is naturally a disappointment to the designer to complete a long and tedious stability assessment only to find that his carefully designed filter is unstable.

To overcome this difficulty a number of techniques have been put forward which may be used to process a transfer function in order to obtain a new and stable function which has approximately the same amplitude response. This will naturally involve a modification to the phase response.

Three techniques will be considered here. Two of them rely on varying the phase of a filter without affecting its amplitude. The third technique is mainly applicable to zero-phase functions, a class of functions which cannot be stabilized by either of the two earlier methods.

2.8 SHANKS' STABILIZATION METHOD

The basis of this method is a conjecture which was put forward by Shanks [47] which is a direct extension to two dimensions of an established property in one dimension [48]. It is regrettable that Genin and Kamp [49,50] subsequently showed that the conjecture was false in the general case by quoting a counter-example. Further counter-examples have since been studied by Tola [51].

We shall start by stating a few basic definitions.

A one-dimensional minimum phase sequence, $\{b_m\}$, is one which has no zeroes inside the z-plane unit circle, namely the z-transform $B(z)$ is such that $B(z) \neq 0 \ \forall |z| \leqslant 1$.

A two-dimensional minimum phase array, $\{b_{m_1,m_2}\}$, is defined as an array for which the Z-transform $B(z_1,z_2)$ has the following properties:

(i) $B(z_1,z_2)$ evaluated at any $\hat{z}_1$ such that $|\hat{z}_1| = 1$ has no zeroes inside the unit circle $|z_2| = 1$, namely that

$$B(\hat{z}_1,z_2) \neq 0 \quad \forall |\hat{z}_1| = 1 \cap |z_2| \leqslant 1, \text{ and}$$

(ii) $$B(z_1,\hat{z}_2) \neq 0 \quad \forall |z_1| \leqslant 1 \cap |\hat{z}_2| = 1.$$

From these definitions we may deduce that a necessary and sufficient condition for an array $\{b_{m_1,m_2}\}$ to be minimum phase is that

$$B(z_1,z_2) \neq 0 \quad \forall |z_1| \leqslant 1 \cap |z_2| \leqslant 1 \tag{2.11}$$

A multidimensional minimum phase array $\{b_{m_1,\ldots,m_n}\}$ may similarly be defined as any array whose Z-transform $B(z_1,\ldots,z_n)$ satisfies the necessary and sufficient condition that

$$B(z_1,\ldots, z_n) \neq 0 \quad \forall (z_i,\ i=1,\ldots,n) \in D \tag{2.12}$$

where $$D = \{(z_i,\ i=1,\ldots,n): \bigcap_{i=1}^{n} |z_i| \leqslant 1\}.$$

Consider now a multidimensional recursive filter having a Z-transform transfer function

$$F(z_1,\ldots,z_n) = \frac{A(z_1,\ldots,z_n)}{B(z_1,\ldots,z_n)}$$

We may restate the stability condition of equation (2.6) in the form that $F(z_1,\ldots,z_n)$ represents a stable system if $B(z_1,\ldots,z_n)$ is the Z-transform of a minimum phase array.

An unstable network with transfer function $F(z_1,\ldots,z_n)$ will have a denominator function $B(z_1,\ldots,z_n)$ which represents a non-minimum phase array $\{b_{m_1,\ldots,m_n}\}$. We may now define an array $\{p_{m_1,\ldots,m_n}\}$, whose size in any dimension m_i is arbitrary and unrelated to that of b, having the idealized property that the convolution of p with b is exactly equal to the multidimensional impulse array δ. Such an array p will be termed the planar least squares inverse (PLSI) of b.

In general it will not be possible to obtain an array of arbitrary size which will exactly satisfy this equality. It is therefore necessary to define the PLSI as that array p which satisfies the convolution equation

$$p * b = g$$

where p is chosen so that g approximates δ in the mean square sense. The mean square error $\overline{Q^2}$ between g and δ is

$$\overline{Q^2} = (1 - g_{0,0,\ldots,0})^2 + \sum_{j_1=1}^{R_1} \cdots \sum_{j_n=1}^{R_n} g^2_{j_1,\ldots j_n} \qquad (2.13)$$

where $R_i = M_i + N_i - 1$,

and M_i and N_i are the sizes of the arrays b and p respectively in the dimension m_i.

It will be seen that this error function is a quadratic function of the variables $g_{j_1,\ldots,j_n}$ and hence its minimization will result in a set of linear equations which may be solved in a relatively simple manner.

Thus for any given multidimensional array, an infinite

number of PLSI may be generated, each of a different size in one or more dimension.

The significance of the planar least squares inverse arises from the conjecture of Shanks applied to two-dimensional arrays. Shanks conjecture may be stated as follows:-

The planar least squares inverse of any two-dimensional array b_{m_1,m_2} is always minimum phase.

The conjecture is an extension of a well-established property of one-dimensional arrays which has been shown to hold for a great many two-dimensional systems. However, as noted above, the technique is not infallible and it has now been shown that the conjecture is fundamentally false. Despite this, it still does provide a useful technique for stabilizing, or at the worst reducing the instability of, an unstable filter [51] .

The stabilization technique proposed by Shanks involved determining the PLSI, p, of the denominator array, b, of an unstable filter and then obtaining a second PLSI, b, of the array, p. According to Shanks' conjecture the array b will be minimum phase and hence a filter designed using the original numerator function and a denominator function derived from b should be stable. Furthermore it may be surmised that the magnitude function of the original and the stabilized function should be approximately the same since one is a double planar least square inverse of the original. The accuracy of this approximation to the magnitude depends upon the closeness to which the array g approximates the unit impulse function δ in the process of obtaining the PLSI.

It seems likely, therefore, that the greater the size of the PLSI array chosen, the more likely is the resultant function to have a magnitude closely approximating the original magnitude function. This fairly obvious assumption has been justified by simulation studies. It has also been shown that in situations where the technique fails to stabilize an unstable two-dimensional filter function, successful stabilization may be achieved by increasing the size of the intermediate PLSI array. It is likely that these two phenomena may be related.

No work has as yet been attempted on the application of the Shanks' technique to systems having more than two dimensions. Apart from the complexity involved in any increase in the number of dimensions, the failure of the technique in certain two-dimensional systems does not encourage expansion of the idea.

At the present, after application of Shanks' stabilization technique, doubt will always remain about the success of the operation and a check on stability will need to be carried out. This again increases the computation involved in the solution of any problem.

2.8.2 Reid and Treitel Stabilization Technique

It is a well known fact that the real and imaginary parts of a minimum phase network function are related by the Hilbert transform.

In one dimension the real and imaginary parts $F_r(e^{j\omega})$, $F_i(e^{j\omega})$ of the Fourier transform of a causal sequence f(m) are related by:

$$F_i(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_r(e^{j\lambda}) \cot\left(\frac{\lambda - \omega}{2}\right) d\lambda \qquad (2.14)$$

If the sequence f(m) is, in addition, minimum phase, then the logarithm of its amplitude spectrum and its phase spectrum $\phi(e^{j\omega})$ are related by the Hilbert transform

$$\phi(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln|F(e^{j\lambda})|\cot\left(\frac{\lambda-\omega}{2}\right)d\lambda \tag{2.15}$$

This technique was used by Reid and Treitel [53] to modify the denominator function of an unstable network, which was therefore a non-minimum phase network, to become a minimum phase network. Using this in place of the original denominator function resulted in a stable network function.

In order to adapt equation (2.15) to evaluation on a digital computer it is necessary to utilize the discrete Hilbert transform instead of the continuous Hilbert transform of equation (2.15). This is represented in one dimension by

$$\phi(m) = \frac{1}{N}\sum_{i=0}^{N-1} \ln|F(i)|\left[1-(-1)^{m-i}\right]\cot\frac{\pi(m-i)}{N} \tag{2.16}$$

where F(i) is defined over the discrete range of values i = 0,1,...N-1. This procedure may be viewed alternatively as an evaluation of the integral in equation (2.15) by a trapezoidal approximation. It is thus likely that any procedure based on this technique may not be satisfactory in all cases.

The technique proposed by Reid and Treitel [53] is based on the definition of a causal array given in Chapter 1. For an n-dimensional system, a causal array $b(m_1,\ldots m_n)$ is defined as

$$b(m_1,\ldots,m_n) = 0 \quad \forall \bigcup_{i=0}^{n} m_i \geqslant M_i/2 \tag{2.17}$$

where m_i varies over the discrete set $\{0,1,\ldots, M_{i-1}\}$ for all $i = 1,\ldots,n$.

The even and odd parts of such a sequence, b_e and b_o, are defined as:

$$\begin{aligned} b_e(m_1,\ldots,m_n) &= \tfrac{1}{2}\left[b(m_1,\ldots,m_n) + b(M_1-m_1,\ldots,M_n-m_n)\right] \\ b_o(m_1,\ldots,m_n) &= \tfrac{1}{2}\left[b(m_1,\ldots,m_n) - b(M_1-m_1,\ldots,M_n-m_n)\right] \end{aligned} \tag{2.18}$$

Using these two definitions and that for causality, we may write the relationship between the even and odd parts of a minimum phase multidimensional sequence as

$$b_o(m_1,\ldots,m_n) = \left[\mathrm{sgn}(m_1,\ldots,m_n)+\mathrm{bdy}(m_1,\ldots,m_n)\right] b_e(m_1,\ldots,m_n) \tag{2.19}$$

where the multidimensional signum function is defined as:

$$\mathrm{sgn}(m_1,\ldots,m_n) = \begin{cases} 1 & \text{for } \bigcap_{i=1}^{n} (0 < m_i < M_i/2) \\ -1 & \text{for } \bigcap_{i=1}^{n} (M_i/2 < m_i < M_i) \\ 0 & \text{otherwise} \end{cases} \tag{2.20}$$

and the boundary function needed to make adjustments at the boundaries is defined as:

$$\mathrm{bdy}(m_1,\ldots,m_n) = \begin{cases} 1 & \text{for } \bigcap_{\substack{i=1 \\ i\neq j}}^{n} (m_i=0) \cap (0 < m_j < M_j/2) \\ -1 & \text{for } \bigcap_{\substack{i=1 \\ i\neq j}}^{n} (m_i=0) \cap (M_j/2 < m_j < M_j) \\ 0 & \text{otherwise} \end{cases} \tag{2.21}$$

The sequence $b(m_1,\ldots,m_n)$ may be obtained from the even and odd parts:

$$b(m_1,\ldots,m_n) = b_e(m_1,\ldots,m_n) + b_o(m_1,\ldots,m_n) \tag{2.22}$$

Now the real and imaginary parts $B_r(m_1,\ldots,m_n)$ and $B_i(m_1,\ldots,m_n)$ of the discrete Fourier transform of an n-dimensional array $b(m_1,\ldots,m_n)$ are related to the discrete Fourier transforms of the even and odd parts by

$$\begin{aligned} B_r(m_1,\ldots,m_n) &= \mathrm{DFT}\left[b_e(m_1,\ldots,m_n)\right] \\ B_i(m_1,\ldots,m_n) &= -j\ \mathrm{DFT}\left[b_o(m_1,\ldots,m_n)\right] \end{aligned} \tag{2.23}$$

Substituting equations (2.23) into (2.19) we obtain:

$$\begin{aligned} B_i(m_1,\ldots,m_n) = -j\ \mathrm{DFT}\Big[\big\{&\mathrm{sgn}(m_1,\ldots,m_n)+\mathrm{bdy}(m_1,\ldots,m_n)\big\}\cdot \\ &.\mathrm{IDFT}\big\{B_r(m_1,\ldots,m_n)\big\}\Big] \end{aligned} \tag{2.24}$$

which defines the multidimensional discrete Hilbert transform.

This may be applied to the denominator magnitude spectral array $|B(m_1,\ldots,m_n)|$ of an unstable transfer function to give the phase array, $\phi(m_1,\ldots,m_n)$, by

$$\begin{aligned} \phi(m_1,\ldots,m_n) = -j\mathrm{DFT}\Big[\big\{&\mathrm{sgn}(m_1,\ldots,m_n)+\mathrm{bdy}(m_1,\ldots,m_n)\big\}\cdot \\ &.\mathrm{IDFT}\big\{\log|B(m_1,\ldots,m_n)|\big\}\Big] \end{aligned} \tag{2.25}$$

The procedure has been applied by Reid and Treitel [53] to two-dimensional sequences and shown to give satisfactory results in many cases. However, a number of situations in which it fails have been shown to exist [54]. The cause of this may be the result of the finite truncation of the infinite array $\phi(m_1,\ldots,m_n)$ or the approximation of the integral by a finite sequence. In addition, although a uniqueness theorem has been proved in one dimension, such a theorem has not yet been discovered in two or more dimensions [56].

It is again seen that after application of the technique in any situation, a stability check must be carried out to verify that the stabilization procedure has been satisfactory.

2.8.3 Pistor Stabilization Technique

The techniques studied so far may be considered as true stabilization techniques in that they cause a modification of the transfer function of the network in such a manner that the amplitude response is kept approximately unchanged, while the phase response is adjusted to ensure stability of the modified transfer function. Such procedures are thus not applicable to zero-phase functions which do not permit phase modification in any manner which would improve the stability.

In one dimension zero-phase functions can only represent, apart from a trivial case, non-causal sequences. It has been shown that such functions may be realized by processing data, first in the positive sense along the axis and cascading this with a processor working in the negative sense. As non-causal sequences do not exist in the time domain, this presents no problem; if the data is distributed in a single spatial dimension it may be stored and processed recursively (or non-recursively if required) in any manner demanded.

With this in mind, the approach used by Pistor [13] in two dimensions is to decompose the array representing the impulse response of the filter into four single quadrant arrays, ${}^{1}f$, ${}^{2}f$, ${}^{3}f$, ${}^{4}f$, each recursing in a different direction. The output array for an arbitrary

input array may be obtained by convolving the input successively with the four quadrant arrays in the appropriate directions of recursion as shown in Fig. 2.1.

The technique may be applied to any unstable filter function but in its simplest form, and probably most useful application, it is used with zero-phase functions. The essence of the problem is the determination of 4 sequences which represent stable transfer functions whose product is equal to the given unstable function and which recurse in the four cardinal directions shown in Fig. 2.1.

This problem has been solved in two dimensions by Pistor [13] who transforms the denominator of the given network function $B(e^{ju_1}, e^{ju_2})$ in the spectral domain into the function $\hat{B}(e^{ju_1}, e^{ju_2})$ in the cepstrum domain by

$$\hat{B}(e^{ju_1}, e^{ju_2}) = \ln\left\{D(e^{ju_1}, e^{ju_2})\right\} \tag{2.26}$$

The technique has been extended to n dimensions by Ahmadi and King [18,22] who show how an n-dimensional zero-phase filter may be designed as the cascade of 2^n stable recursive filters, each recursing in a different direction. The decomposition is done in the n-dimensional cepstrum domain in an identical manner to that given by Pistor.

The procedure may be outlined with application to an unstable zero-phase n-dimensional filter having a transfer function

$$F(z_1,\ldots,z_n) = \frac{A(z_1,\ldots,z_n)}{B(z_1,\ldots,z_n)} \tag{2.27}$$

in which

$$B(z_1,\ldots,z_n) = \sum_{m_1=0}^{M_1} \cdots \sum_{m_n=0}^{M_n} b_{m_1,\ldots,m_n} . z_1^{m_1} \ldots . z_n^{m_n} \tag{2.28}$$

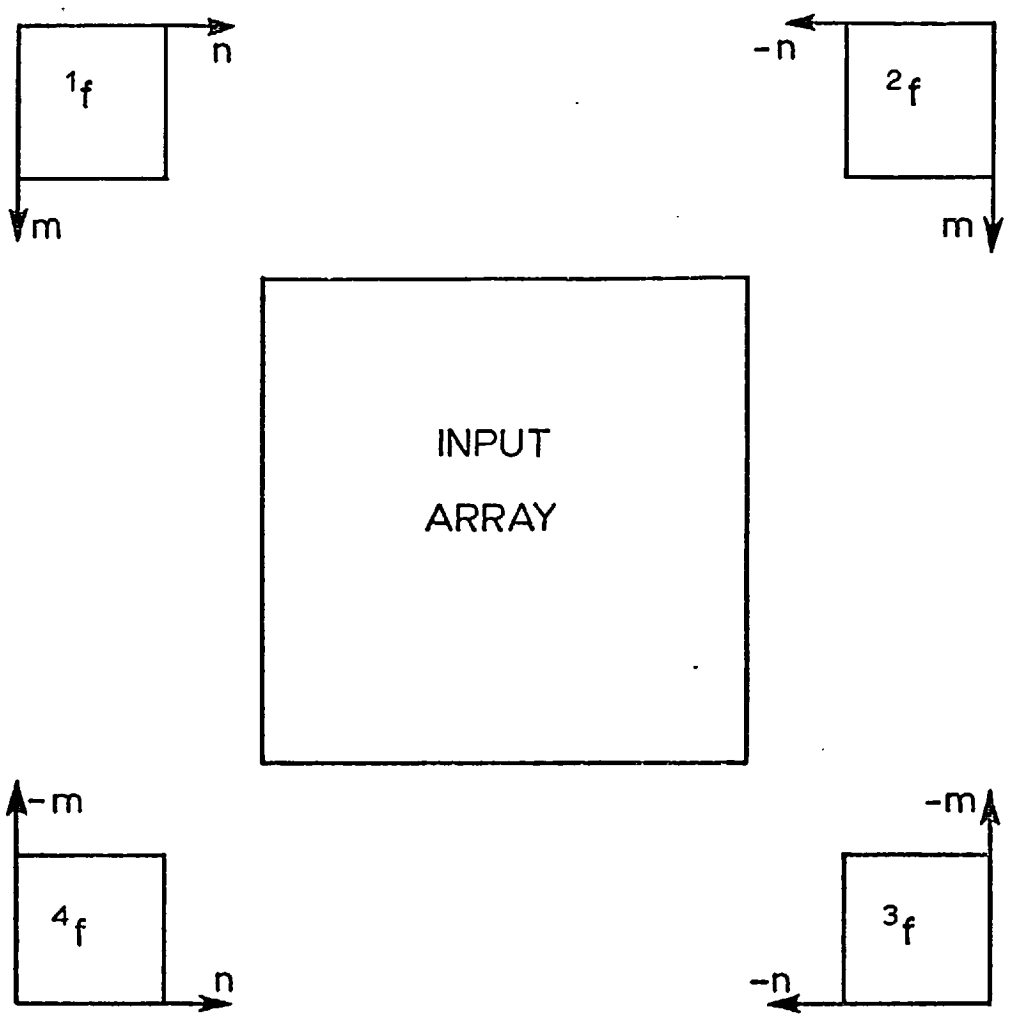


Fig. 2·1 Pistor's decomposed single-quadrant filters convolved recursively with an input array.

The spectrum of B may be evaluated on $\bigcap_{i=1}^{n} |z_i| = 1$ as

$$B(e^{-ju_1},\ldots,e^{-ju_n}) = \sum_{m_1=0}^{M_1} \cdots \sum_{m_n=0}^{M_n} b_{m_1,\ldots,m_n} \cdot e^{-j(u_1+\ldots+u_n)} \tag{2.29}$$

The cepstrum $\hat{B}$ may now be evaluated directly as

$$\hat{B}(e^{-ju_1},\ldots,e^{-ju_n}) = \sum_{m_1=0}^{M_1} \cdots \sum_{m_n=0}^{M_n} \hat{b}_{m_1,\ldots,m_n} \cdot e^{-j(u_1+\ldots+u_n)}$$

The cepstrum array $\{\hat{b}_{m_1,\ldots,m_n}\}$ is now decomposed into 2^n single quadrant arrays. No optimum procedure has been obtained for this decomposition. One satisfactory technique is to decompose $\hat{b}$ symmetrically so that for the kth quadrant filter

$${}^{k}\hat{b}_{m_1,\ldots,m_n} = g\ \hat{b}_{m_1,\ldots,m_n}$$

where $g = 1/2^n$ for $\bigcap_{i=1}^{n} (m_i = 0)$,

$g = 1/2^{n-1}$ for all values of m_i along coordinate one-dimensional axes common to quadrant k and any other quadrant,

$g = 1/2^{n-p}$ for all p-dimensional coordinate planes common to quadrant k and any other quadrant,

$g = 1$ otherwise.

The next step in the design is to transform the single quadrant array $\{{}^{k}\hat{b}_{m_1,\ldots,m_n}\}$ in the cepstrum domain to the frequency domain array $\{{}^{k}b_{m_1,\ldots,m_n}\}$. This is achieved by initially determining the value at the origin from

$${}^{k}b_{0,\ldots,0} = \exp({}^{k}\hat{b}_{0,\ldots,0})$$

and then deriving the remaining values by the recursion formula

$$^{k}b_{m_1,\ldots,m_n} = \sum_{p_1=0}^{m_1} \cdots \sum_{p_i=1}^{m_i} \cdots \sum_{p_n=0}^{m_n} \left(\frac{p_i}{m_i}\right) . {}^{k}\hat{b}_{p_1,\ldots,p_n} . {}^{k}b_{m_1-p_1,\ldots,m_n-p_n}$$

for all $i = 1,\ldots,n$ and $m_i \neq 0$

This infinite array now requires truncation to a finite length before the single quadrant denominator function ${}^{k}B(z_1,\ldots,z_n)$ may be generated.

The advantage of the Pistor technique is that, in theory, the overall transfer function of the cascade of single quadrant filters should be identical with that of the given transfer function; this is not achieved in practice since the single quadrant filters are, of necessity, truncated approximations to the infinite length filters designed by the technique. Furthermore, the technique involves the determination of multidimensional Fourier transforms, and these can only be performed to a limited accuracy by computational techniques.

The technique will always give stable single quadrant filters and it has been shown that the accuracy of the approximation to the specified impulse response is improved by using larger arrays for the intermediate Fourier transforms [13,14]. The truncation of these arrays is liable to introduce undesirable poles in the transfer function. Ekstrom and Woods [14] have proposed the introduction of weighting functions in the two-dimensional case to remove these possible poles from the unstable region. It has been shown that the same technique using multidimensional weighting sequences may improve the design procedure in these problems.

2.8.4 Review of Stabilization Techniques

Shanks' method has been found to be highly satisfactory in many cases, is simple to implement, but is based on an erroneous assumption which invalidates its application in some situations. It is not suitable to multidimensional systems or to zero-phase functions.

Read and Treitel's method is also not capable of guaranteeing the stabilization of even non-minimum phase networks. The failure appears to be due to a theoretical error rather than computational approximations although it does give satisfactory results in many cases.

Pistor's technuque may be applied to all types of zero-phase filter of any dimensionality and although stability of the resulting decomposed filters cannot be guaranteed, an increase in the size of the intermediate arrays will always lead to a stable result.

CHAPTER THREE

SPATIAL DESIGN TECHNIQUES

Metiri sua regna decet, vires - que fateri.

"Pharsalia"
Lucan

SPATIAL DESIGN TECHNIQUES

3.0 INTRODUCTION

Although the theme of this thesis is the application of transformations in the spectral domain to the problem of the design of multidimensional digital filters, it is desirable to look briefly at other design methods in order to appreciate the advantages of the spectral transformation methods.

Probably the most obvious technique which may be adopted to design a filter to process a given multidimensional input array is a direct optimization routine. This consists of designing a filter to give an output approximating the desired output and then to modify the filter coefficients in such a manner as to minimize some function of the error between the output and the specified output.

The main drawback to this approach is that there is no direct means of ensuring stability of the output array when the true error is used as the criterion and that techniques for stabilization of a filter in the space domain are not available.

However, some of the techniques available are of considerable interest and will be reviewed here.

3.1 KALMAN TECHNIQUE

In 1958 Kalman [34] proposed a technique for the design of one-dimensional recursive filters which was later extended to two dimensions.

The Kalman technique is outlined in this section in a form

applicable to multidimensional systems as derived in a direct extension from two dimensions by Nowrouzian et al [36,37].

Suppose that the desired impulse response of an n-dimensional system is $\{d_{m_1,\ldots,m_n}\}$ for all $\underline{m} \in S_K$, where

$$S_K = \left\{(m_i,\ i = 1,2,\ldots,n):\ \bigcap_{i=1}^{n} 0 \leqslant m_i \leqslant K_i\right\} \tag{3.1}$$

Then the desired filter transfer function is

$$D(z_1,\ldots,z_n) = \sum_{m_1=0}^{K_1} \cdots \sum_{m_n=0}^{K_n} d_{m_1,\ldots,m_n} . z_1^{m_1} \ldots\ldots z_n^{m_n} \tag{3.2}$$

Let the approximating n-dimensional recursive filter be

$$F(z_1,\ldots,z_n) = \frac{A(z_1,\ldots,z_n)}{B(z_1,\ldots,z_n)} \tag{3.3}$$

$$= \sum_{m_1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} f_{m_1,\ldots,m_n} . z_1^{m_1} \ldots\ldots z_n^{m_n} \tag{3.4}$$

Let $$A(z_1,\ldots,z_n) = \sum_{m_1=0}^{N_1} \cdots \sum_{m_n=0}^{N_n} a_{m_1,\ldots,m_n} . z_1^{m_1} \ldots\ldots z_n^{m_n} \tag{3.5}$$

and $$B(z_1,\ldots,z_n) = \sum_{m_1=0}^{M_1} \cdots \sum_{m_n=0}^{M_n} b_{m_1,\ldots,m_n} . z_1^{m_1} \ldots\ldots z_n^{m_n} \tag{3.6}$$

where $b_{0,0,\ldots,0} = 1$ without loss of generality.

The Z-transform of the n-dimensional unit impulse array $\{\delta_{m_1,\ldots,m_n}\}$ is defined as $X(z_1,\ldots,z_n)$.

We now wish to choose coefficients $a_{m_1,\ldots,m_n}$, $b_{m_1,\ldots,m_n}$ of $A(z_1,\ldots,z_n)$ and $B(z_1,\ldots,z_n)$ such that the coefficients of the following true, finite error $\overline{E}$ are minimized in a least mean squares sense.

$$\overline{E}(z_1,\ldots,z_n) = \frac{A(z_1,\ldots,z_n)}{B(z_1,\ldots,z_n)} X(z_1,\ldots,z_n) - D(z_1,\ldots,z_n)$$

$$= \sum_{m_1=0}^{K_1} \cdots \sum_{m_n=0}^{K_n} e_{m_1,\ldots,m_n} . z_1^{m_1} \ldots . z_n^{m_n} \quad (3.7)$$

This is a highly nonlinear optimization problem and is thus not amenable to simple computational procedures. However, an iterative solution may be formulated whereby the minimization may be carried out by a succession of linear processes.

This design technique is relatively simple and results in a good approximation to the desired impulse response specified. A solution to the problem is always guaranteed. Unfortunately, there is no assurance that the filter so designed will be stable.

3.2 SHANKS' METHOD

Shanks [38] proposed a solution to the design problem by minimizing a false error function in order to obtain a recursive filter structure. This has been extended to n-dimensions by Nowrouzian et al [36,37,18].

Consider the approximating filter function given by equation (3.3). This may be written in the form

$$A(z_1,\ldots,z_n) = B(z_1,\ldots,z_n).F(z_1,\ldots,z_n)$$

Transforming this into the space domain expresses the numerator sequence in terms of the convolution of the denominator sequence and the impulse response of the filter. Thus

$$a_{m_1,\ldots,m_n} = \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} b_{j_1,\ldots,j_n} \cdot f_{m_1-j_1,\ldots,m_n-j_n} \quad \text{for all } \underline{m} \in S_N$$

$$= 0 \quad \text{for all } \underline{m} \in \overline{S}_N \tag{3.8}$$

where $\underline{m}$ = the vector $\{m_1,\ldots,m_n\}$

and $\overline{S}_N$ = not S_N

and $S_N \triangleq \{(m_i,\ i=1,2,\ldots,n) : \bigcap_{i=1}^{n} 0 \leqslant m_i \leqslant N_i\}$

Noting that $b_{0,0,\ldots,0} = 1$ we may write equation (3.8) as

$$f_{m_1,\ldots,m_n} = -\sum_{j_1=0}^{M_1} \cdots \sum_{\substack{j_n=0 \\ \bigcap_{i=1}^{n} j_i \neq 0}}^{M_n} b_{j_1,\ldots,j_n} \cdot f_{m_1-j_1,\ldots,m_n-j_n} \quad \forall\ \underline{m} \in \overline{S}_N$$

We may now choose the coefficients $b_{j_1,\ldots,j_n}$ such that $f_{m_1,\ldots,m_n}$ closely approximates the desired impulse response $d_{m_1,\ldots,m_n}$. We may write

$$d_{m_1,\ldots,m_n} \simeq -\sum_{j_1=0}^{M_1} \cdots \sum_{\substack{j_n=0 \\ \bigcap_{i=1}^{n} j_i \neq 0}}^{M_n} b_{j_1,\ldots,j_n} \cdot d_{m_1-j_1,\ldots,m_n-j_n} \tag{3.9}$$

$$\forall\ \underline{m} \in (\overline{S}_N \cap S_K)$$

where S_K is defined in (3.1).

Now we may again define a finite error, $e_{m_1,\ldots,m_n}$, which is not a true error since it is only defined over a limited range of the given impulse response, by

$$d_{m_1,\ldots,m_n} = e_{m_1,\ldots,m_n} - \sum_{j_1=0}^{M_1} \cdots \sum_{\substack{j_n=0 \\ \bigcap_{i=1}^{n} j_i \neq 0}}^{M_n} b_{j_1,\ldots,j_n} \cdot d_{m_1-j_1,\ldots,m_n-j_n}$$

$$\forall\ \underline{m} \in \overline{S}_N \cap S_K \tag{3.10}$$

Thus

$$e_{m_1,\ldots,m_n} = \sum_{j_1=0}^{M_1} \cdots \sum_{j_n=0}^{M_n} b_{j_1,\ldots,j_n} . d_{m_1-j_1,\ldots,m_n-j_n} \qquad (3.11)$$

$$\forall \underline{m} \in (\bar{S}_N \cap S_K)$$

The error $\overline{E^2} = \sum_{\underline{m} \in (\bar{S}_N \cap S_K)} \cdots \sum e^2_{m_1,\ldots,m_n}$ is now minimized with respect to the coefficients of B. This may be carried out relatively simply since the optimization may be effected by setting the partial derivatives of $\overline{E^2}$ with respect to the $b_{j_1,\ldots,j_n}$ to zero, and these form a set of linear simultaneous equations in the denominator coefficients.

Having computed the denominator coefficients, we may obtain the numerator coefficients by minimizing the mean square difference between the coefficients of $F(z_1,\ldots,z_n)$ and the coefficients of the desired response $D(z_1,\ldots,z_n)$. Thus the numerator coefficients are chosen such that the mean square error

$$\overline{e^2} = \sum_{j_1=0}^{K_1} \cdots \sum_{j_n=0}^{K_n} \left[f_{j_1,\ldots,j_n} - d_{j_1,\ldots,j_n} \right]^2 \qquad (3.12)$$

is minimized with respect to the coefficients $a_{j_1,\ldots,j_n}$. This is, in fact, a Wiener filtering problem in n-dimensions and it again results in a set of linear simultaneous equations which are, in principle, directly soluble.

An alternative method for determination of the numerator polynomial, albeit of lower accuracy, is to compute the coefficients $a_{m_1,\ldots,m_n}$ from the convolution of $b_{m_1,\ldots,m_n}$ and $d_{m_1,\ldots,m_n}$ derived from

$$A(z_1,\ldots,z_n) = B(z_1,\ldots,z_n).D(z_1,\ldots,z_n) \qquad (3.13)$$

$$\forall \underline{m} \in S_N$$

The Shanks method appears to provide a good solution since it depends only on the solution of linear simultaneous equations. It also does guarantee a solution to the problem. However, the procedure involves minimization of a false, though finite, error and results in a filter which is not necessarily stable.

3.3 BORDNER SYNTHESIS TECHNIQUE

Bordner [20] proposed a synthesis technique which involves augmenting the given finite length impulse response array by a "tail" array to convert it into an infinite array and then minimizing the error between the filter impulse response and the augmented desired impulse response.

This procedure may be extended to n-dimensions in the following manner [36,37,18] .

Let us assume that a hypothetical sequence $\{\tilde{d}_{j_1,\ldots,j_n}\}$ is given for $\underline{j} \in S_\infty$ where S_∞ is the set of integers defined by $S_\infty = \{(j_1,\ldots,j_n) : \bigcap_{i=1}^{n} 0 \leqslant j_i \leqslant \infty \}$.

An n-dimensional recursive digital filter is to be designed with transfer function $F(z_1,\ldots,z_n)$ such that the true, infinite, mean square error

$$\overline{e^2} = \sum_{j_1=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} [\tilde{d}_{j_1,\ldots,j_n} - f_{j_1,\ldots,j_n}]^2 \qquad (3.14)$$

is minimized.

It has been shown that with appropriate choice of $\tilde{d}_{j_1,\ldots,j_n}$ minimization [20] of equation (3.14) results in a stable recursive filter.

Now consider the specified impulse response sequence $\{d_{j_1,\ldots,j_n}\}$ defined for all $\underline{j} \in S_K$ and let us define

$$\tilde{d}_{j_1,\ldots,j_n} = d_{j_1,\ldots,j_n} \qquad \forall\, \underline{j} \in S_K$$
$$= f'_{j_1,\ldots,j_n} \qquad \forall\, \underline{j} \in (\bar{S}_K \cap S_\infty) \qquad (3.15)$$

The technique originates from the simple argument that if the specified sequence were, in fact, infinite, rather than of finite length, the least mean square minimization solution would of necessity lead to a stable recursive filter. It is intuitively obvious that the augmenting sequence should be the most natural extension of the given sequence $d_{j_1,\ldots,j_n}$ and which is square summable over S_∞. This ensures stability of the designed filter.

Thus the sequence $f'_{j_1,\ldots,j_n}$ must have exactly the same functional form as $f_{j_1,\ldots,j_n}$. Its n-dimensional Z-transform is given by

$$F'(z_1,\ldots,z_n) = \frac{A'(z_1,\ldots,z_n)}{B'(Z_1,\ldots,z_n)}$$

$$= \sum_{j_1=0}^{\infty}\cdots\sum_{j_n=0}^{\infty} f'_{j_1,\ldots,j_n}.z_1^{j_1}\ldots.z_n^{j_n} \qquad (3.16)$$

where $$A'(z_1,\ldots,z_n) = \sum_{j_1=0}^{N_1}\cdots\sum_{j_n=0}^{N_n} a'_{j_1,\ldots,j_n}.z_1^{j_1}\ldots.z_n^{j_n} \qquad (3.17)$$

and $$B'(z_1,\ldots,z_n) = \sum_{j_1=0}^{M_1}\cdots\sum_{j_n=0}^{M_n} b'_{j_1,\ldots,j_n}.z_1^{j_1}\ldots.z_n^{j_n} \qquad (3.18)$$

The problem may now be formulated as the minimization of the mean square error defined by

$$\overline{e^2} = \sum_{\underline{j} \in S_K} \cdots \sum \left[d_{j_1,\ldots,j_n} - f_{j_1,\ldots,j_n} \right]^2$$

$$+ \sum_{\underline{j} \in S_\infty} \cdots \sum \left[f'_{j_1,\ldots,j_n} - f_{j_1,\ldots,j_n} \right]^2$$

$$- \sum_{\underline{j} \in S_K} \cdots \sum \left[f'_{j_1,\ldots,j_n} - f_{j_1,\ldots,j_n} \right]^2 \qquad (3.19)$$

with the only constraint that $F'(z_1,\ldots,z_n)$ must represent the transfer function of a stable recursive filter. (This constraint is identical with the requirement for square summability already mentioned.)

This minimization may be performed iteratively by first selecting a stable n-dimensional sequence which has a Z-transform, $F'(z_1,\ldots,z_n)$ and then solving the minimization of equation (3.19) to obtain $f_{j_1,\ldots,j_n}$. These may now be used as a fresh approximation to the "tail" sequence $f'_{j_1,\ldots,j_n}$. The optimization is repeated until the error $f_{j_1,\ldots,j_n} - f'_{j_1,\ldots,j_n}$ is equal to the n-dimensional zero sequence.

The minimization of equation (3.19) involves the least mean square approximation for absolutely summable infinite multidimensional sequences. This has been solved for one and two dimensions [40] and the solution is amenable to extension to n dimensions. A solution is guaranteed which minimizes the true error and which ensures that the designed filter is stable. Unfortunately the solution does not necessarily converge on a global minimum. The computational labour is very great and finally the initial choice of an appropriate n-dimensional "tail" array is arbitrary but the choice of a natural

extension array to the given impulse response array is vital to the reduction of complexity of solution.

3.4 BERTRAM'S DESIGN TECHNIQUE

A two-dimensional design technique proposed by Bertram [39] has been extended to n-dimensional systems as outlined below [37,36]. The method proposed is essentially an iterative one which starts from an arbitrary initial set of numerator and denominator coefficients and attempts to improve on these values to give a closer approximation to the desired impulse response.

At the $(p-1)^{th}$ iteration assume that we have obtained a set of coefficients $(A^{(p-1)}, B^{(p-1)}) = C^{(p-1)}$ for both numerator and denominator. Then we wish to improve this set by minimizing the mean square error

$$\overline{e^2} = \sum_{m_1=0}^{K_1} \cdots \sum_{m_n=0}^{K_n} \left[d_{m_1,\ldots,m_n} - f_{m_1,\ldots,m_n} \right]^2 \tag{3.20}$$

Now we may approximate this error by means of a truncated n-dimensional Taylor series

$$\begin{aligned} \overline{e_f^2} \simeq & \sum_{m_1=0}^{K_1} \cdots \sum_{m_n=0}^{K_n} \Bigg[d_{m_1,\ldots,m_n} - f_{m_1,\ldots,m_n}\Big|_{C=C^{(p-1)}} \\ & + \sum_{j_1=0}^{N_1} \cdots \sum_{j_n=0}^{N_n} \frac{\partial f_{m_1,\ldots,m_n}}{\partial a_{j_1,\ldots,j_n}}\Bigg|_{C=C^{(p-1)}} \Delta a_{j_1,\ldots,j_n} \\ & - \sum_{j_1=0}^{N_1} \cdots \sum_{j_n=0}^{N_n} \frac{\partial f_{m_1,\ldots,m_n}}{\partial b_{j_1,\ldots,j_n}}\Bigg|_{C=C^{(p-1)}} \Delta b_{j_1,\ldots,j_n} \Bigg]^2 \end{aligned} \tag{3.21}$$

where $$\Delta a_{j_1,\ldots,j_n} = a^{(p)}_{j_1,\ldots,j_n} - a^{(p-1)}_{j_1,\ldots,j_n} \tag{3.22}$$

and $$\Delta b_{j_1,\ldots,j_n} = b^{(p)}_{j_1,\ldots,j_n} - b^{(p-1)}_{j_1,\ldots,j_n} \tag{3.23}$$

The partial derivatives in equation (3.21) may be computed by means of a recursion formula and the problem now reduces to minimization of $\overline{e_f^2}$ with respect to $\Delta a_{j_1,\ldots,j_n}$ and $\Delta b_{j_1,\ldots,j_n}$ and since this is a quadratic function of the variables, a global minimum is assured. Values of the coefficients may now be determined from equations (3.22) and (3.23) and the iteration continued until the error is less than a prescribed minimum value.

The convergence of this algorithm is assured and a solution to the n-dimensional problem is guaranteed. The computational labour may be shown to be relatively simple. However, the main disadvantage is that the initial values need to be chosen so that they do not deviate too far from the final value. Furthermore, the mean square error given by the Taylor series is a poor approximation to the true error and once again the designed filter is not necessarily stable.

3.5 LAL'S DESIGN TECHNIQUE

Lal's [41] technique is an extension of Shanks' method. Instead of attempting to obtain a transfer function which approximates to the complete specified impulse response and hence optimized to a very high order transfer function, Lal partitioned the desired impulse response into a number of smaller arrays. Using a two-dimensional system as an example, the whole desired impulse array of magnitude $L_1 \times L_2$ may be subdivided into N_1, N_2 subgroups each of size $k_1 \times k_2$ (where $Nk_i = L_i$) as follows.

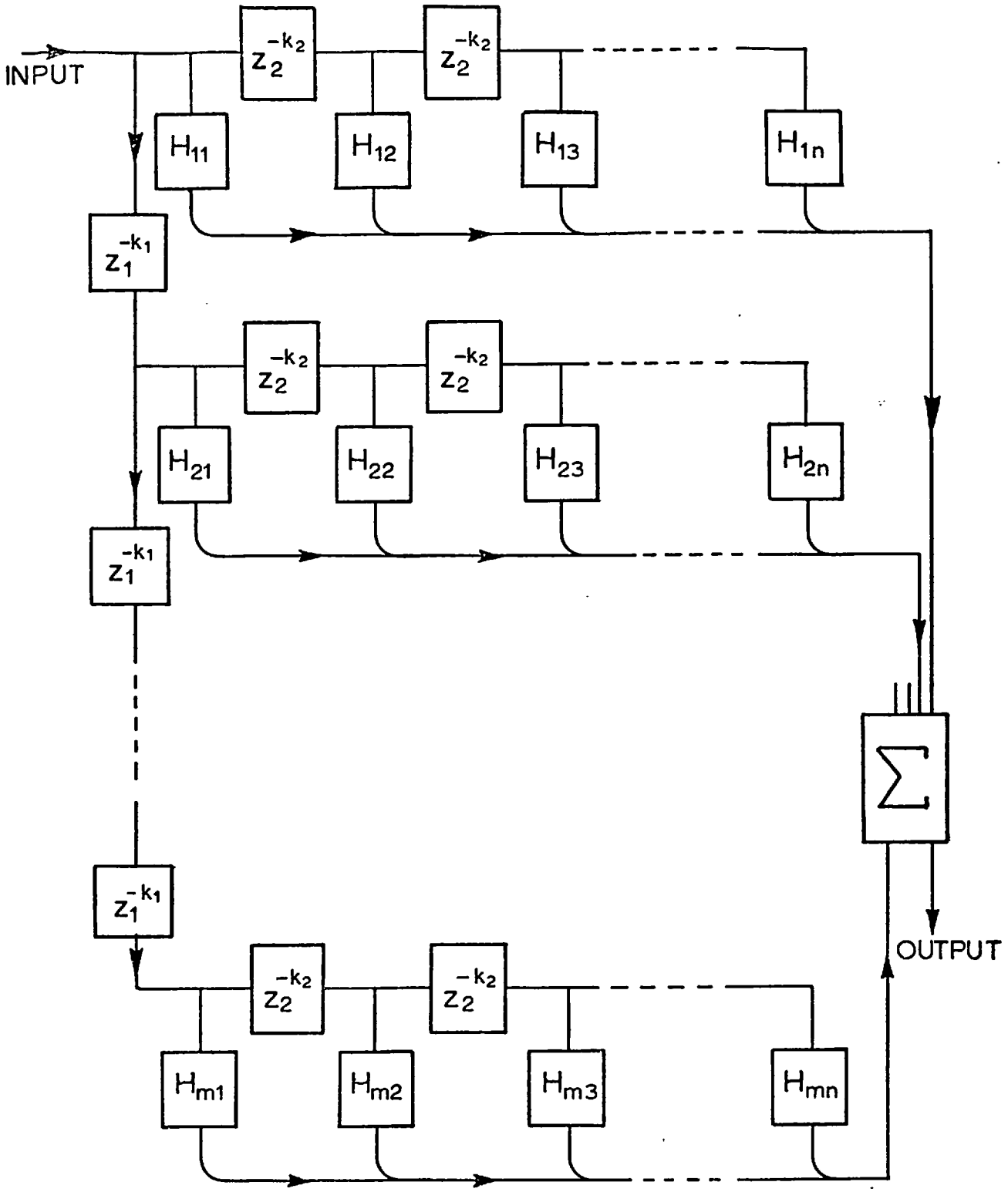


Fig. 3·1 Lal's partitioned filter.

$$\left[d_{\ell_1,\ell_2}\right]_{\substack{\ell_1=1\\ \ell_2=1}}^{L_1,L_2} = \left[d_{\ell_1,\ell_2}\right]_{\substack{\ell_1=1\\ \ell_2=1}}^{k_1,k_2} \cdot \left[d_{\ell_1,\ell_2}\right]_{\substack{\ell_1=1\\ \ell_2=k_2}}^{k_1,2k_2} \cdots\cdots \left[d_{\ell_1,\ell_2}\right]_{\substack{\ell_1=L_1-k_1+1\\ \ell_2=L_2-k_2+1}}^{L_1,L_2}$$

Each of these groups may be approximated to a relatively close degree by the basic 2-dimensional transfer function

$$H_{n_1 n_2}(z_1,z_2) = \frac{a_{11} + a_{12}z_1 + a_{21}z_2 + a_{22}z_1z_2}{b_{11} + b_{12}z_1 + b_{21}z_2 + b_{22}z_1z_2}$$

in which simple constraints on the denominator will ensure stability. The realization takes the form shown in Fig. 3.1.

3.6 CRITICISM OF DESIGN TECHNIQUES IN THE SPACE DOMAIN

The techniques have been discussed in turn at the conclusion of the relevant sections. It may be seen that the Bordner technique appears most desirable in that it is guaranteed to be stable whereas the other techniques require a concluding stability test. Such a test on a multidimensional transfer function may well involve as much computational time as is saved by one of the other techniques. It does, however, suffer the disadvantage that a global optimum cannot be guaranteed and hence one of the earlier techniques may give rise to a 'better' solution with smaller error.

One other factor in any design process of this nature is the choice of the degree of both numerator and denominator of the transfer function. In general, the higher degree transfer function chosen, the more likely one is to obtain a good approximation to the desired sequence. Inevitably such a course of action brings with it the concommitant increase in complexity of the design algorithm.

Normally the order of numerator and denominator polynomials is kept small compared with the size of the specified impulse response. If this condition is not observed it would probably be more economical to design the filter in non-recursive form.

CHAPTER FOUR

SPECTRAL TRANSFORMATIONS

It is vain to do with more
what can be done with less.

William of Occam

SPECTRAL TRANSFORMATIONS

4.1 INTRODUCTION

Having considered a number of methods for designing multidimensional digital filters in the space domain, a study of the techniques for design in the frequency domain is the obvious sequel since the performance of a system in one of these domains is directly correlated with its performance in the other domain by the multidimensional Fourier transform.

The obvious approach to frequency design would be to follow that in the space domain and attempt an approximation to a frequency response characteristic at a discrete number of frequencies and use an optimization technique to minimize the error between the desired multidimensional frequency response and the specified response. Such an approach is of no practical value since it can only minimize the error at a finite number of discrete frequencies and no control can be exercised over the behaviour of the function between these discrete frequencies. There is, in general, an infinite set of network transfer functions which can be used to approximate at a set of discrete frequencies. The space domain response obtained by taking the inverse multidimensional Fourier transform of any such transfer function may well result in an impulse response which has highly undesirable characteristics. If the realization is attempted using a recursive filter structure, a design technique based on the above may easily result in an unstable filter.

An alternative technique would involve generation of an analytic function of the multivariable frequency argument $w_1,\ldots,w_n$ which would "fit" the desired specification at an arbitrary number

of discrete points. This effectively is the approach used in one-dimensional analogue filter design when obtaining Butterworth, Chebyshev, elliptic, etc. approximations to a given specification. In this one-dimensional case it is possible to obtain the required analytic function in closed form with the assurance that the function will give rise to a stable filter.

A similar approach has been investigated by McClellan for design of a two-dimensional filter frequency response function to fit a given specification in a Chebyshev sense [24]. There are two reasons why this is not possible. First, it is impossible for any set of functions defined on a two- or multidimensional domain to satisfy the Haar condition [9]; thus the alternation theorem applies in a weaker form. Second, there is no possibility of ordering the external frequencies as in the one-dimensional case, where progress along the ordered sequence guarantees that the error changes sign from one point to the next.

It is therefore impossible to extend the design techniques used in one dimension of transforming a known stable one-dimensional analogue filter into a one-dimensional digital filter using an appropriate transformation function between the s plane variable and the z plane variable.

One further difference between one-dimensional and multi-dimensional filters lies in the response characteristics. In one dimension it is only necessary to specify the shape, either as amplitude, phase, group delay, etc. as a function of frequency, giving such parameters as pass band range, stop band range, transition band attenuation gradient, pass band ripple, minimum

pass band attenuation etc. In two- or multidimensional systems these parameters need to be defined in one or more dimensions, resulting in a specification in which the cut-off frequency is replaced by a cut-off contour in two dimensions or a cut-off hyper-surface in multiple dimensions. Most of the examples used as illustrations in this thesis will be restricted to two dimensions in order to facilitate graphical representation; however, to demonstrate the versatility of the technique some three-dimensional filters will be studied.

One other important difference between one- and multi-dimensional systems lies in the analytic properties of multivariable polynomials. In a single variable the "fundamental theorem of algebra" states that any polynomial may be factorized into the product of a number of first and second order factors with real coefficients. This permits the designer to factorize the given specified transfer function in either the continuous s-domain or the discrete z-domain into a number of first or second order functions which may be cascaded to give the required specification. No theorem corresponding to this exists for multivariable polynomials and hence such simple design techniques are not possible (this may easily be verified by a simple counter example). Thus the designer is forced to use techniques which involve the direct design of high order systems.

It is for this reason that a number of techniques have been evolved for transforming stable one-dimensional filters into two- or multidimensional filters whose cut-off boundaries have prescribed shape and whose amplitude spectrum in some given cross-sections is determined from the prototype filter.

Besides these transformations for generating higher dimensional systems from one-dimensional systems, certain other transformations have been proposed which relate two-dimensional filters having some given properties with new two-dimensional filters having different frequency characteristics; for example, transformation from two-dimensional low pass to two-dimensional band pass. These transformations are entirely analogous to the similar transformations in one-dimensional analogue and digital filters.

We may review the field of spectral transformations by considering the following categories:

1. One-dimensional to one-dimensional. This will include z to s, s to z, s to s' and z to z'.

2. One-dimensional to two-dimensional (or multidimensional) transformations to include z to $s_1, s_2, \ldots$; s to $z_1, z_2, \ldots$; s to $s_1, s_2, \ldots$; z to $z_1, z_2 \ldots$

3. Two-dimensional to two-dimensional transformations, namely, $s_1, s_2, \ldots$ to $s_1', s_2', \ldots$; $z_1, z_2, \ldots$ to $z_1', z_2', \ldots$

Further extensions to these may be envisaged but so far no work has appeared on the subject and it appears a rather sterile field.

4.2 ONE DIMENSION TO ONE DIMENSION TRANSFORMATIONS

One of the earliest uses of spectral transformations was in analogue filters to transform between low pass, high pass, band pass, band stop and other more complex multiple band filters using transformation of the form

$w \Longrightarrow 1/w$ Low pass to high pass

$w \Longrightarrow w_0/w - w/w_0$ Low pass to band pass

Their use permits a prototype low pass filter design to be transformed to a high pass, band pass or band stop filter having similar pass band, transition band and stop band characteristics.

Another transformation of this form is the bilinear transformation

$$s = \frac{1 - z}{1 + z} \tag{4.1}$$

which is the most commonly used of a whole range of transformations designed to generate the transfer function of a digital filter from that of an analogue filter having a similar nature of frequency response.

Other transformations may be derived to generate more complex filters, such as high pass, band pass, etc. from low pass prototypes. For example

$$s = \frac{1 + z}{1 - z} \tag{4.2}$$

will give a high pass from a low pass prototype.

In addition to these, there exist digital to digital transformations which change say a low pass filter $F_1(z)$ into a high pass filter $F_2(z')$ via the transformation

$$z = -z' \tag{4.3}$$

4.3 TWO DIMENSION TO TWO DIMENSION TRANSFORMATIONS

A similar concept, outlining two-dimensional to two-dimensional digital transformations, has been summarized by Prendergrass [42]. These transformations may be applied, for example, to a two-dimensional low pass digital filter having a cutoff boundary of approximately circular shape, to generate filters having a variety of combinations of low pass, high pass, band pass, band stop characteristics in the two frequency dimensions, retaining approximately the original shape of the cut-off boundary contour when transformed. This cut off boundary is rather poorly transformed in the case of low pass to high pass or transformations which covertly incorporate such a relationship.

In his consideration he has deliberately restricted himself to transformations which have the following properties:

1. First quadrant stable transfer functions generate first quadrant stable transfer functions.

2. Real functions transform to real functions.

3. Some important characteristic of the amplitude response is maintained after the transformation.

Although these are his self-dictated terms of reference it may be appreciated that the third one is the only requirement of fundamental significance and with some of the stabilizing techniques at present available neither of the first two conditions need be imposed in order to obtain usable transformation functions. The significance of the removal of the first two constraints will become apparent subsequently where a wide range of useful transformations

will be proposed which do not conform to these restrictions.

The class of transformations which he considers are all of the form of two-dimensional all pass transfer functions. Their general form is

$$z_1' = G_1(z_1,z_2)$$

$$= \pm \prod_{k=1}^{K_1} \left\{ \frac{z_1^{M_k} z_2^{N_k} \left(\sum_{i=0}^{M_k} \sum_{j=0}^{N_k} a_{ij} \, z_1^{-i} z_2^{-j} \right)}{\sum_{i=0}^{M_k} \sum_{j=0}^{N_k} a_{ij} \, z_1^{i} \, z_2^{j}} \right\} \tag{4.4}$$

and $\quad z_2' = G_2(z_1,z_2)$

which has the same functional form as equation (4.4) but with different constants a_{ij}.

It has been shown that if and only if the functions $G_1(z_1,z_2)$ and $G_2(z_1,z_2)$ represent stable two-dimensional transfer functions and that they are used as transformations on a stable two-dimensional transfer function $H(z_1,z_2)$, then the resultant transfer function $H\{G_1(z_1,z_2),\ G_2(z_1,z_2)\}$ will also be stable.

Initially he shows that transformations which involve only one variable, i.e. $z_1' = G_1(z_1)$, $z_2' = G_2(z_2)$ result in the generation of a useful class of two-dimensional filters. The simplest of these

$$z_1' = \frac{a + z_1}{1 + az_1}$$

$$z_2' = \frac{b + z_2}{1 + bz_2} \tag{4.5}$$

results in a transformation from low pass to band pass or from low pass to a low pass filter with changed cut-off characteristics along each frequency dimension. The type of transformation depends on the values chosen for the parameter a. Second degree transformations of this form introduce band pass or band stop characteristics in the appropriate dimension.

Another class of transformations studied by Prendergrass is those of the form of a second order all pass function

$$z_i' = \frac{a_i + b_i z_1 + c_i z_2 + z_1 z_2}{1 + c_i z_1 + b_i z_2 + a_i z_1 z_2} \qquad i = 1,2 \qquad (4.6)$$

The stability of this function must be considered and this imposes constraints on the coefficients a_i, b_i, c_i. This transformation is particularly useful in controlling the shape of the amplitude contours in the vicinity of the principal diagonal on the frequency plane.

4.4 ONE DIMENSION TO TWO DIMENSION TRANSFORMATION

The preceding transformations are of great value when an initial design has been achieved of, say, a low pass two-dimensional filter with appropriate cut-off boundary. However, this itself is one of the more difficult problems to solve, and the techniques of Prendergrass do not allow much modification of the boundary shape, since the limited number of parameters available in the transformation permits constraints on the transformed function at a limited number of points only.

One of the most significant problems is that of designing a two-dimensional low pass or high pass filter having a cut-off

boundary approximating to circular. It is obvious that the contours corresponding to other values of the magnitude contour must deviate from circular to a greater or lesser degree since the outermost contour must coincide with the sampling limits of the digital filter at $w_1 = \pm\pi$, $w_2 = \pm\pi$ which is essentially rectangular.

A few of the attempts at this circular boundary approximation will be summarized below.

4.4.1 Separable Product

One of the earliest attempts at a solution to this problem was put forward by Hall [15] in 1970. He suggested that a two-dimensional filter could be made from the cascade of two filters, each of which varied in one dimension only. Thus

$$H(z_1,z_2) = F_1(z_1).F_2(z_2) \tag{4.7}$$

The justification for inclusion of the separable product filter in the class of filters obtained by spectral transformation is that a two-dimensional filter $H(z_1,z_2)$ may be made by cascading the two two-dimensional filters $F_1(z_1)$ and $F_2(z_2)$; these are generated from prototype one-dimensional filters $F_1(z)$ and $F_2(z)$ by the spectral transformations $z_1 = z$ and $z_2 = z$, respectively.

The cut-off boundaries of the $F_1(z_1)$ filter are parallel to the z_2 axis, and of the $F_2(z_2)$ filter, parallel to the z_1 axis. Thus, the cut-off contour of $H(z_1,z_2)$ will be approximately rectangular with rounded corners. The only parameters available in the design are those of the prototype one-dimensional filters $F_1(z)$, $F_2(z)$ and hence no control on cut-off boundary shape is possible.

An example of a design using this technique applied to a 3rd order Butterworth prototype digital filter, having cut-off frequency at $w = \pi/2$, is shown in Fig. 4.1; the frequency response is shown by the isometric projection of Fig. 4.1(a) and the contour plot of lines of equal amplitude of response in Fig. 4.1(b). A similar design using a Chebyshev filter having the same cut-off frequency in the prototype is illustrated in Fig. 4.2. The only difference between the two is in the expected sharper cut-off of the Chebyshev response and the ripple in both frequency directions in the pass band. All subsequent designs will be carried out using a 3rd order Butterworth prototype unless any significant variations are apparent by using a Chebyshev filter.

4.4.2 Shanks' Rotated Filters

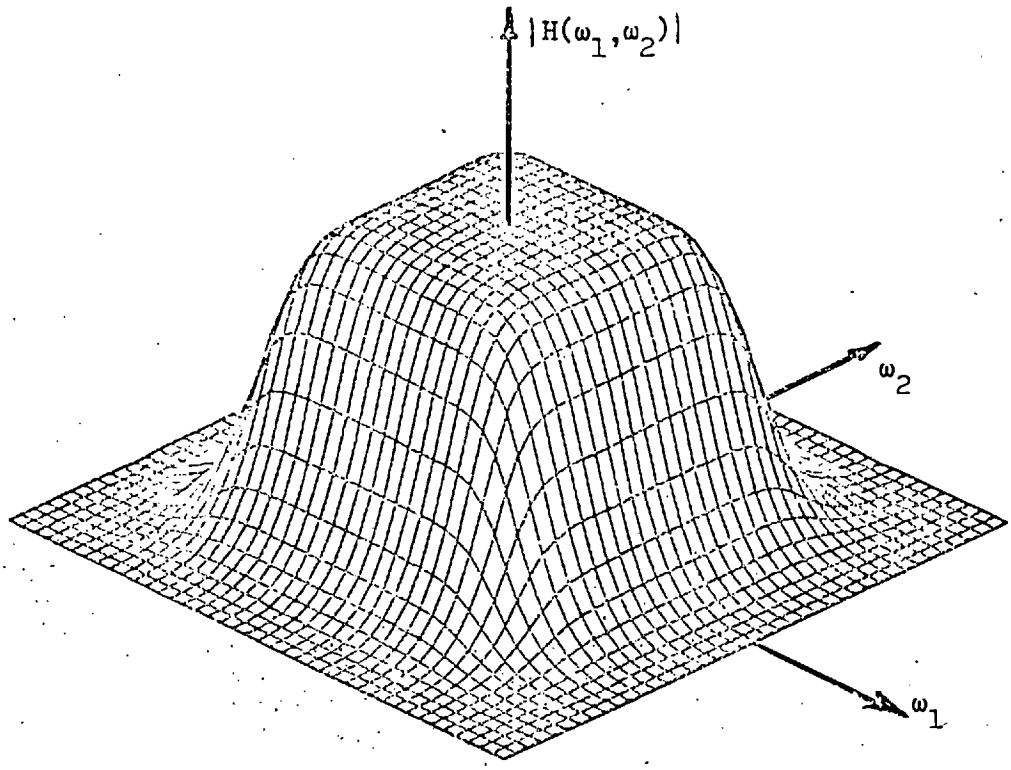
Shanks [5] proposed a technique which also originated from a one-dimensional low pass filter, in this case designed in the continuous frequency domain. Such a filter function may be represented by the product of first order numerator and denominator factors as

$$F(s) = \prod_{i=1}^{m} (s - q_i) \Big/ \prod_{i=1}^{n} (s - p_i) \qquad (4.8)$$

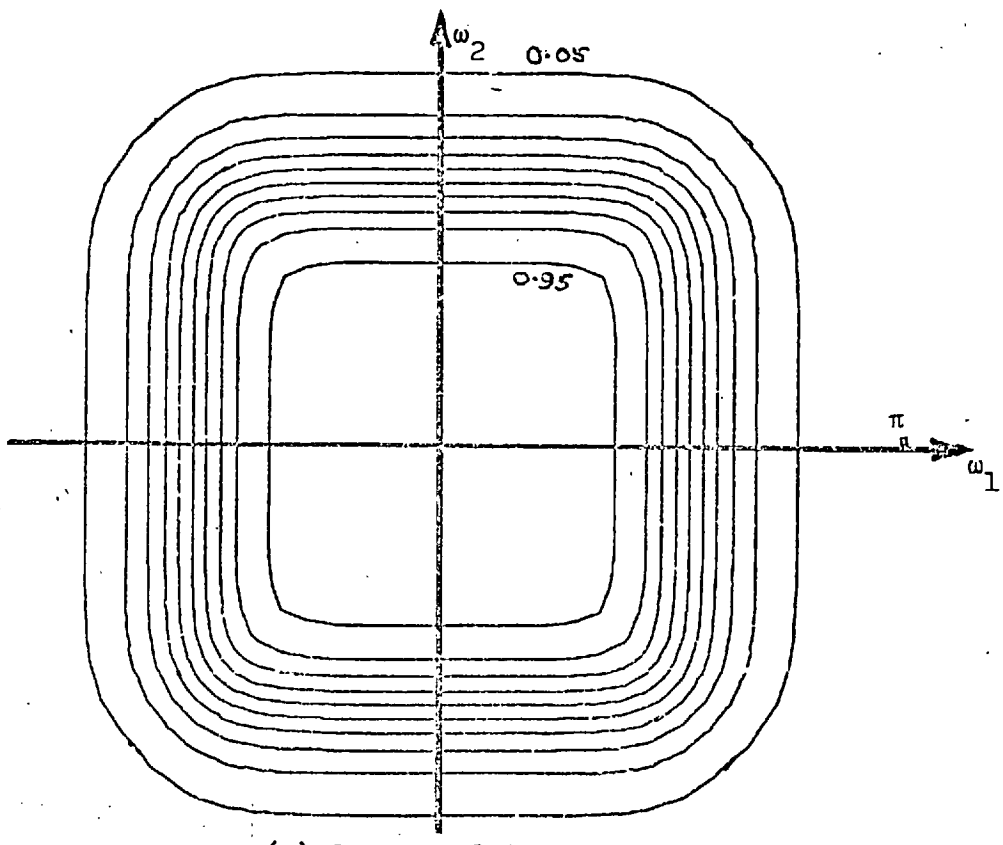
Transformation into a function of two dimensions may be achieved by setting $s_1 = s$ and leaving s_2 unspecified. The result is a two-dimensional transfer function

$$H_1(s_1, s_2) = F(s_1) \qquad (4.9)$$

having cut-off boundary parallel to the s_2 axis. This filter may now be rotated anti-clockwise through an angle θ by the transformation

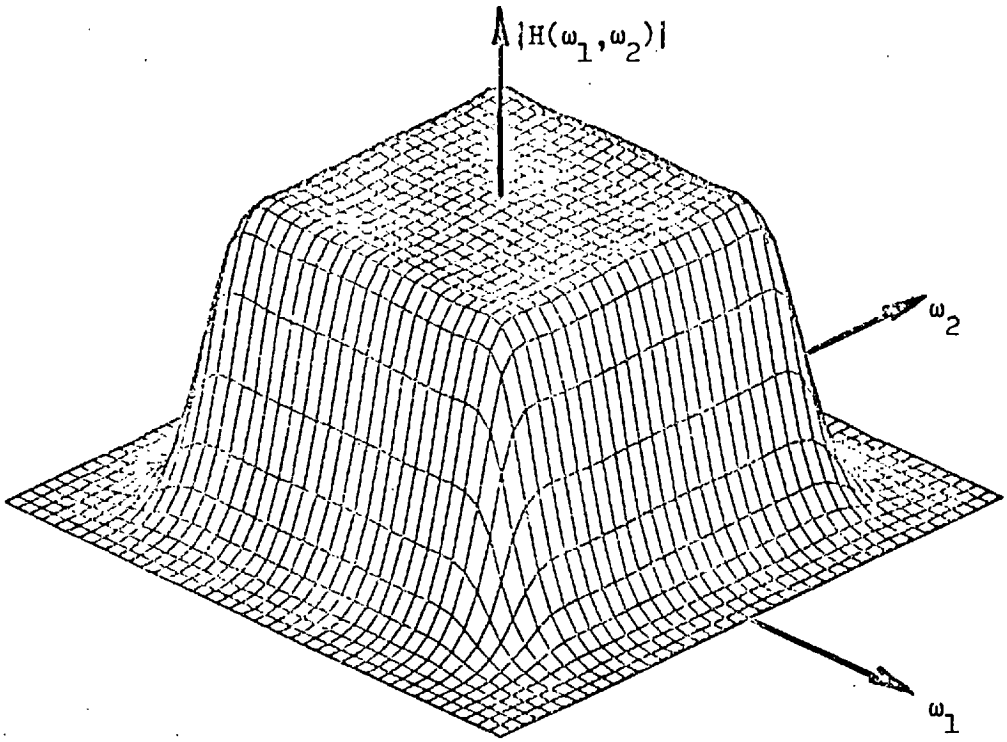


(a) Isometric plot.

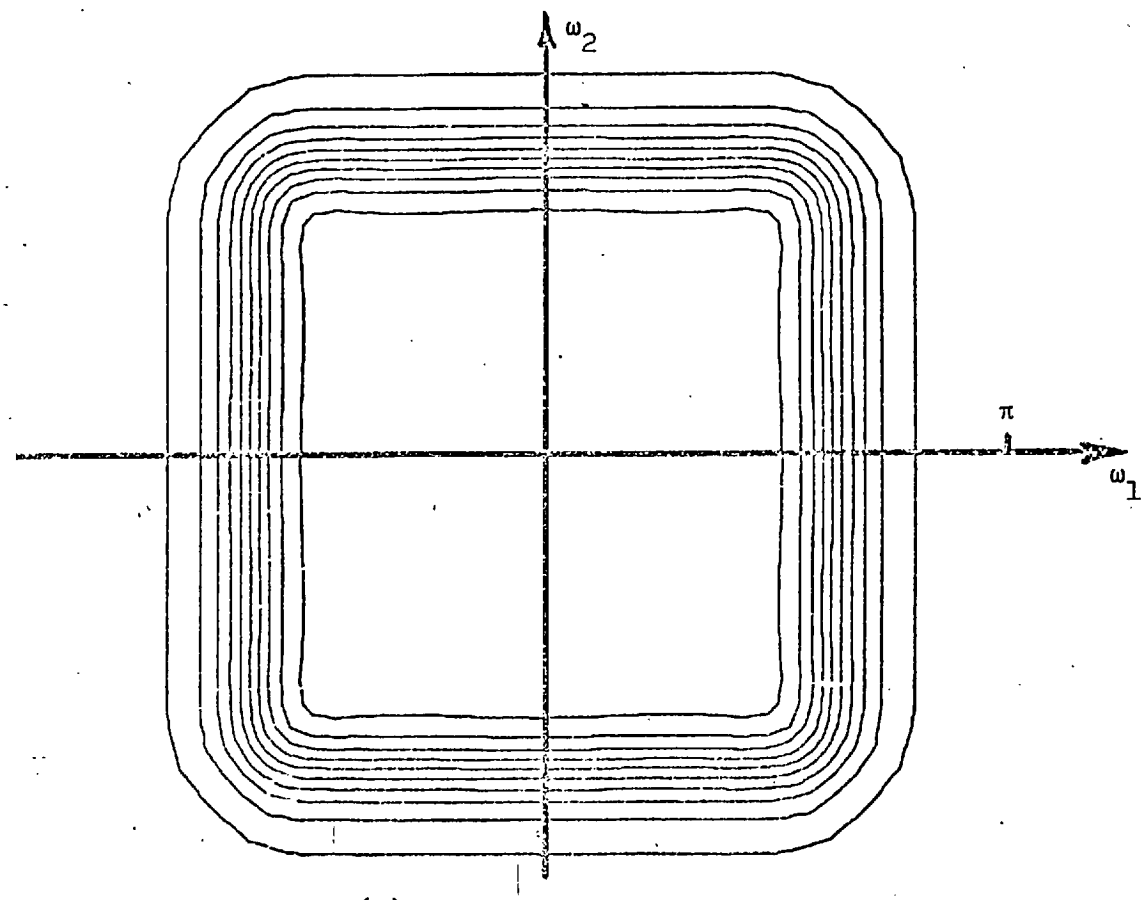


(b) Contour plot.

Figure 4.1. Separable product filter. Prototype: third order Butterworth digital filter; $\omega_o = \pi/2$.



(a) Isometric plot.



(b) Contour plot.

Figure 4.2. Separable product filter. Prototype: third order Chebyshev digital filter; $\omega_o = \pi/2$, $\delta = 1\%$.

$$s_1 = s_1'\cos\theta + s_2'\sin\theta$$
$$s_2 = -s_1'\sin\theta + s_2'\cos\theta \tag{4.10}$$

generating a two-dimensional continuous filter having a transfer function

$$H_2(s_1', s_2') = \frac{\prod_{i=1}^{m} (s_2'\cos\theta - s_1'\sin\theta - q_i)}{\prod_{i=1}^{n} (s_2'\cos\theta - s_1'\sin\theta - p_i)} \tag{4.11}$$

The analogue transfer function of equation (4.11) may be converted to a discrete form by means of the bilinear transformation (normalized to T = 1),

$$s_1' = \frac{1 - z_1}{1 + z_1}$$
$$s_2' = \frac{1 - z_2}{1 + z_2} \tag{4.12}$$

to give a discrete transfer function of the form

$$H_3(z_1, z_2) = \prod_{i=1}^{M} \frac{a_{11i} + a_{21i}z_1 + a_{12i}z_2 + a_{22i}z_1z_2}{b_{11i} + b_{21i}z_1 + b_{12i}z_2 + b_{22i}z_1z_2} \tag{4.13}$$

The cut-off boundaries of a filter designed by this technique are far from circular and there is no guarantee that the design will result in a stable filter.

4.4.3 Costa and Venetsanopoulos modification

A valuable modification to the technique of Shanks was proposed by Costa and Venatsanopoulos [43] , who attempted a design of a near-circular symmetric filter by cascading a number of 'Shanks'

filters having different angles of rotation. By this means it was possible to construct a polygonal approximation to a circular cut-off boundary so long as it was possible to rotate the original filter by a total of 180° (since each filter contributes two opposite sides to the polygon).

A stability criterion was developed which showed that angles of rotation of the designed filter from 0 to -90° resulted in stable filters. Thus the design technique could not achieve the required total angular span and the cut-off boundary was inevitably far from circular.

4.4.4 McClellan Transformation

The McClellan transformation [9,44] is a direct application of a spectral transformation to two-dimensional design techniques.

The Z-transform of a one-dimensional finite impulse response filter $H(z)$ will have a frequency response $H(e^{ju})$. For a useful class of such zero-phase filters, the frequency response may be written in the form:

$$H(e^{ju}) = h(0) + \sum_{n=1}^{N} h(n)\ (e^{jnu} + e^{-jnu}) \qquad (4.14a)$$

$$= h(0) + \sum_{n=1}^{N} 2h(n)\cos nu \qquad (4.14b)$$

In the case of two-dimensional filters, the transfer function of a class of zero-phase filters may be written

$$H(e^{ju_1}, e^{ju_2}) = \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} a(m_1,m_2)\cos m_1u_1 . \cos m_2u_2 \quad (4.15)$$

One-dimensional filters of the form (4.14b) may be converted to two-dimensional filters of the form (4.15) via the McClellan transformation

$$\cos u = \sum_{p_1=0}^{K_1} \sum_{p_2=0}^{K_2} t(p_1,p_2)\cos p_1u_1 . \cos p_2u_2 \quad (4.16)$$

The technique may be most simply illustrated by consideration of the lowest order McClellan transformation, in which $K_1 = K_2 = 1$,

$$\cos u = t_{00} + t_{10}\cos u_1 + t_{01}\cos u_2 + t_{11}\cos u_1 \cos u_2 \quad (4.17)$$

In order to apply the transformation of (4.17) to the transfer function (4.14b), this latter must be written as a power series in cos u rather than as a function of the cosines of multiple angles. We may achieve this using Chebyshev polynomials to give

$$H(e^{ju}) = \sum_{n=0}^{N} b(n)(\cos u)^n \quad (4.18)$$

Equation (4.17) is now substituted into (4.18) to give the required transfer function.

In this particular simple form of the McClellan transformation we now require to determine the four coefficients t_{ij} to fit the required contour specification.

One constraint on a transformation from one-dimensional low pass to two-dimensional low pass is that the origin in one

dimension must transform to the origin in the two-dimensional domain.

The remaining three parameters are determined by constraining the cut-off boundary to have, say, circular symmetry; this requires that, at the cut-off frequency,

$$u_1^2 + u_2^2 = R^2$$

The problem is now one of constrained optimization and usually results in a non-linear minimization of an error function expressed in closed analytic form.

For higher orders of transformation than that given by equation (4.17), it is not generally possible to specify an error function in closed form and thus any solution becomes virtually impossible without the use of unwarranted computational facilities.

A suboptimal approach has been proposed for this problem which will guarantee a solution by minimization of a false error function.

It should be noted that in the above form the McClellan transformation may only be applied to non-recursive two-dimensional filters.

4.4.5 Bernabo Design Technique

Bernabo, Emiliani and Cappelini [17] have extended the McClellan technique to the design of two-dimensional recursive digital filters.

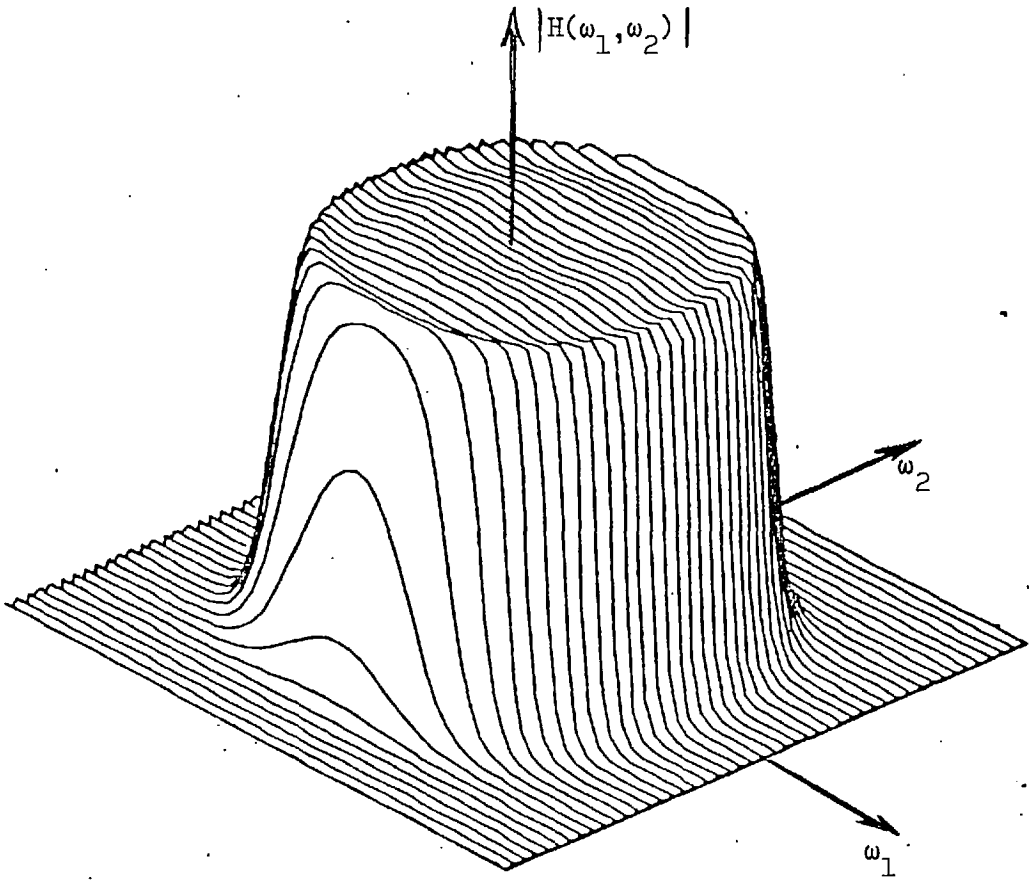


Figure 4.3. Bernabo designed filter. Prototype: fourth order Chebyshev filter; $\omega_o = 0.6\pi$, $\delta = 1\%$. [23]

The frequency transfer function of a zero-phase recursive digital filter may be written in the form

$$H(e^{ju_1}, e^{ju_2}) = \frac{\sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} p(m_1,m_2)\cos m_1u_1 . \cos m_2u_2}{\sum_{m_1=0}^{N_1} \sum_{m_2=0}^{N_2} q(m_1,m_2)\cos m_1u_1 . \cos m_2u_2} \tag{4.19}$$

The transformation equation (4.16) or, with less generality, equation (4.17), may be directly applied to the one-dimensional recursive zero-phase transfer function

$$H(e^{ju}) = \frac{\sum_{n=0}^{N} a(n) \cos nu}{\sum_{n=0}^{M} b(n) \cos nu} \tag{4.20}$$

$$= \frac{\sum_{n=0}^{N} a'(n)(\cos u)^n}{\sum_{n=0}^{M} b'(n)(\cos u)^n} \tag{4.21}$$

Now the designed filter has zero-phase property and is therefore inevitably unstable [14] . Since it is a zero-phase function, the Pistor stabilization technique may be applied [13] . In this, the denominator polynomial is factorized into four one-quadrant recursive functions which recurse in the four cardinal directions. An example of a filter designed by the Bernabo technique is shown in Fig. 4.3 [23] . There is no fundamental reason why the Bernabo technique may not be extended to more than two dimensions but the computation would be very tedious.

4.4.6 The Ahmadi Transformation

A novel transformation has been proposed by Ahmadi et al [18,19]. The relationship proposed is a simple first order two-dimensional reactance function

$$s = \frac{a_1 s_1 + a_2 s_2}{1 + b s_1 s_2} \tag{4.22}$$

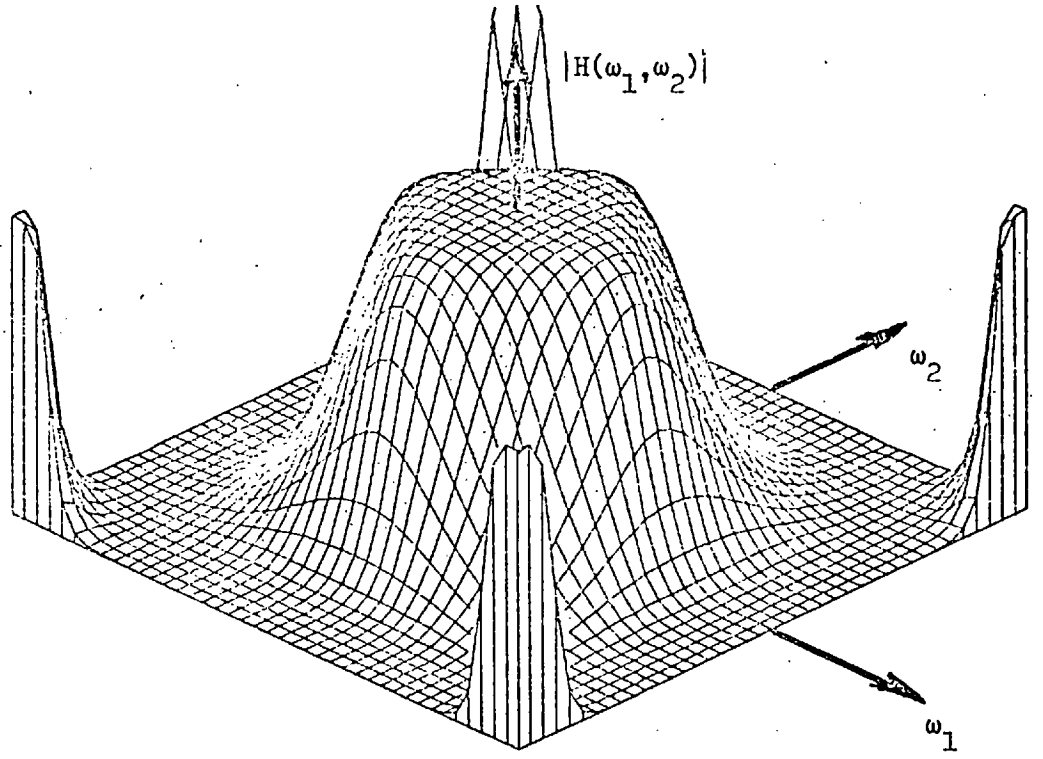
used to transform a one-dimensional low pass continuous filter function to a two-dimensional continuous low pass function. This transformation will realize a guaranteed stable first quadrant function with cut-off frequencies along the two frequency axes determined by the prototype filter characteristics and the parameters a_1 and a_2. In fact there is no loss in generality except in a frequency scaling factor along the two axes if we set $a_1 = a_2 = 1$. The parameter b controls the shape of the cut-off boundary.

In order to obtain a characteristic which has symmetry with respect to the two frequency axes and zero-phase, it is necessary to cascade four single quadrant filters to give the overall transfer function

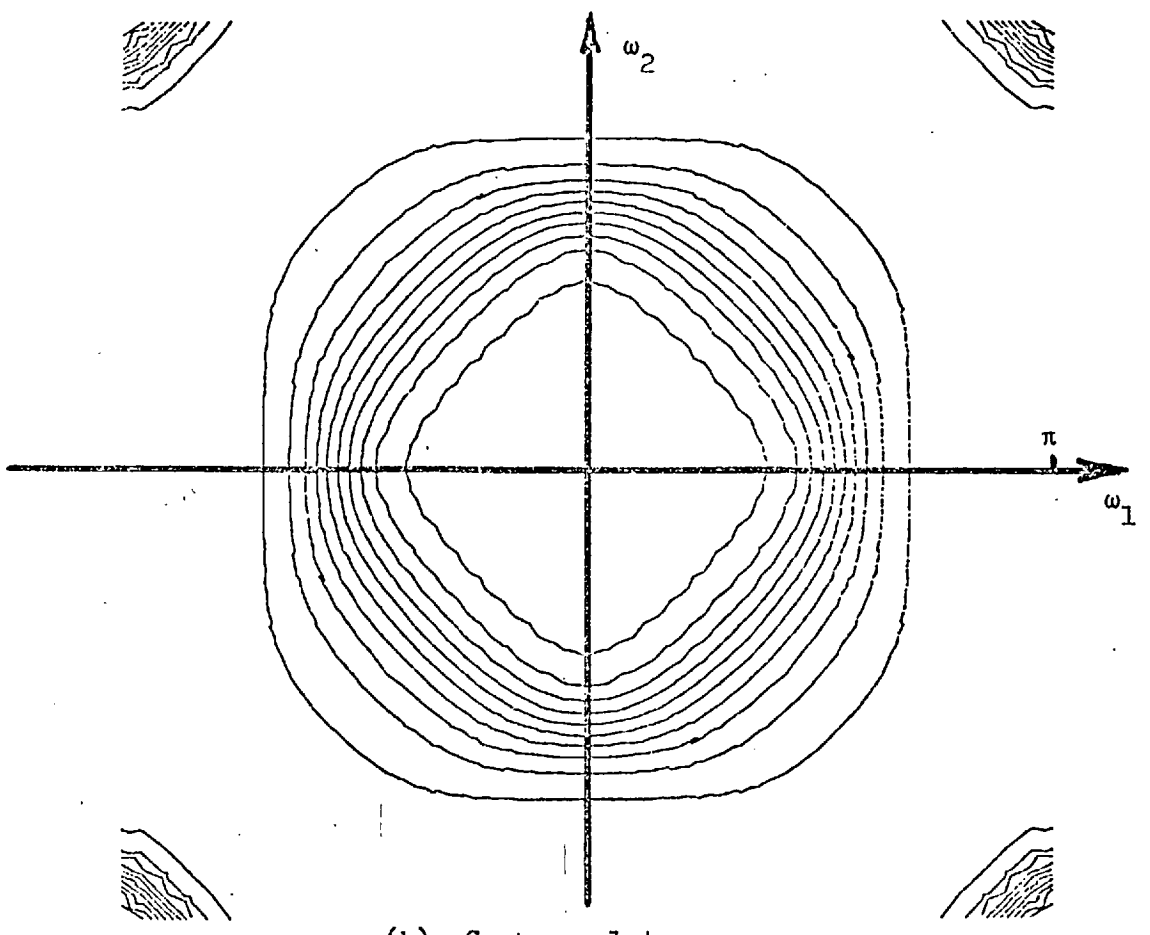
$$H(z_1,z_2) = F(z_1,z_2).F(z_1,z_2^{-1}).F(z_1^{-1},z_2).F(z_1^{-1},z_2^{-1}) \tag{4.23}$$

where $F(z_1,z_2)$ is the two-dimensional function obtained by application of equation (4.22) to a one-dimensional low pass filter. This cascade of four one-dimensional filters gives a cut-off profile which is vaguely diamond-shaped, depending on the specified cut-off frequency.

The mapping of the extreme angular frequencies, $w = 0$ and π in the one-dimensional plane shows that the one-dimensional

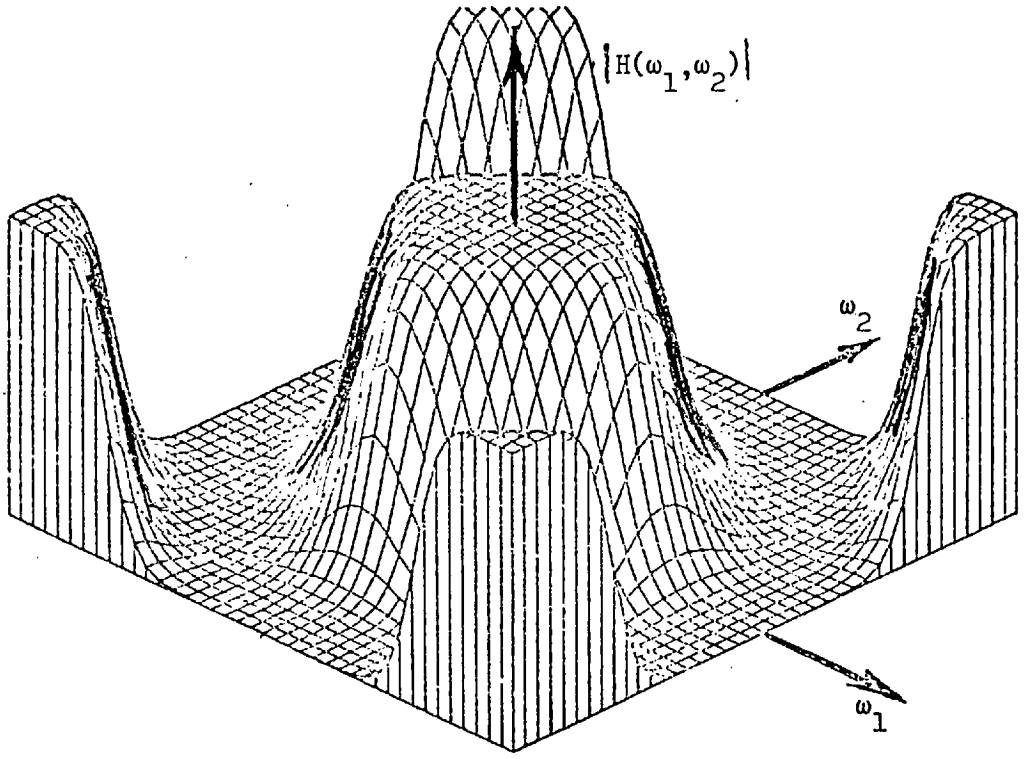


(a) Isometric plot.

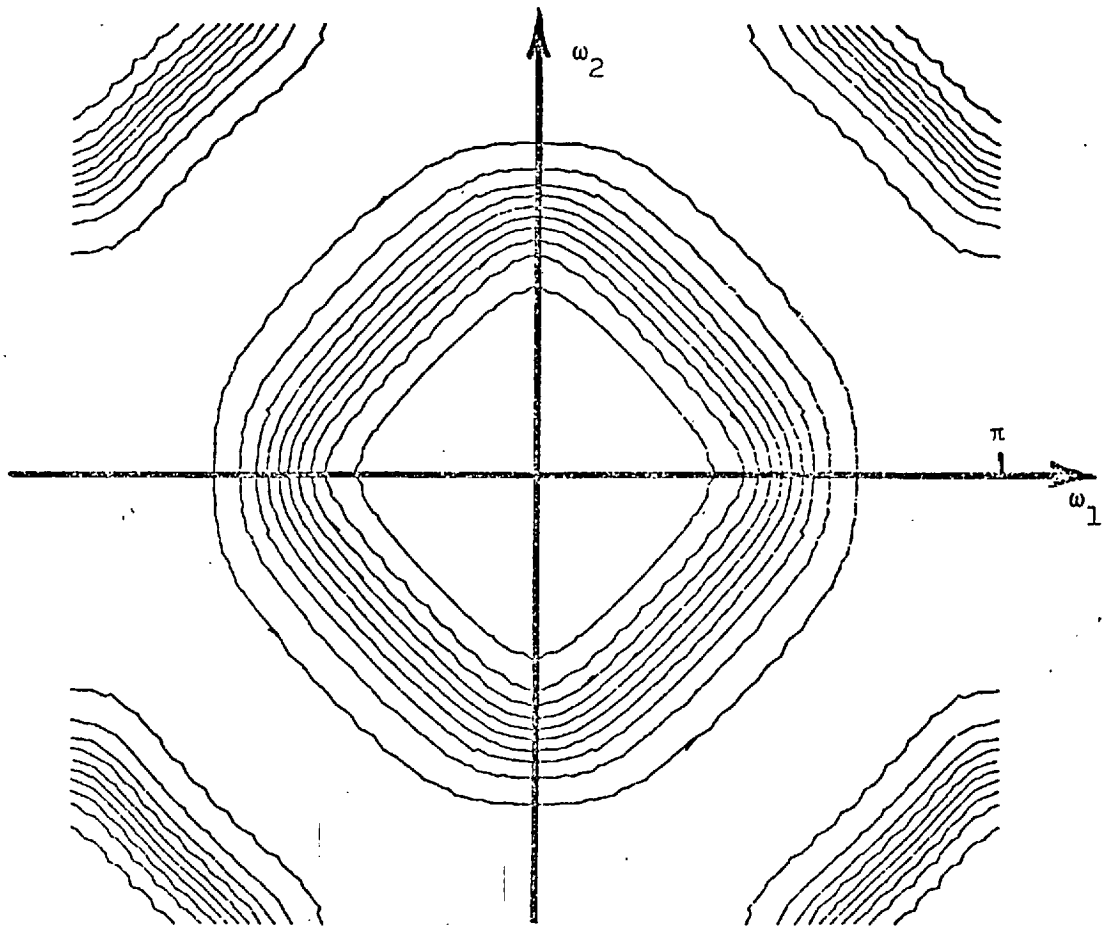


(b) Contour plot.

Figure 4.4. Ahmadi filter. Prototype: third order Butterworth analogue filter; $\omega_o = 1.1584$. Transformation; $a_1 = a_2 = 1$, $b \cong 0.2$.



(a) Isometric plot.



(b) Contour plot.

Figure 4.5. Ahmadi filter. Prototype: third order Butterworth analogue filter; $\omega_o = 1.1584$. Transformation; $a_1 = a_2 = 1$, $b = 0.6$.

origin maps into the points $(0,0)$ and (π,π) and the one-dimensional frequency π maps into the points $(0,\pi)$ and $(\pi,0)$. From this it may be seen that in addition to the low pass region around the origin there is also a second pass region located around (π,π).

The response of a two-dimensional filter designed thus from a 3rd order low-pass Butterworth prototype for two values of the parameter b are shown in Figs. 4.4 and 4.5 designed to a cut-off boundary at $w_c = \pi/2$. The additional pass band around (π,π) may be removed by the simple expedient of cascading the Ahmadi filter with a guard filter, possibly a separable low order two-dimensional filter. This modifies the characteristic to the more desirable shape shown in Fig. 4.6.

The main advantages of this technique are that stability of the filter is assured without recourse to the decomposition technique of Pistor, and also that the design procedure is extremely simple. Unfortunately as only three design parameters are available in the transformation function it is not possible to approximate very closely to the idealized circular cut-off boundary; this is seen from a comparison of Figs. 4.4 and 4.5.

It may be shown that any attempt to design a transformation function closely approximating circular is impossible in the general case using the Ahmadi transformation. One technique for approximating a contour to a circle would be a direct minimization of error with respect to b, the only free parameter available in the transformation function. This would be cumbersome, particularly in view of the fact that the designed filter is formed from the concatenation of four single-quadrant filters. A simpler technique

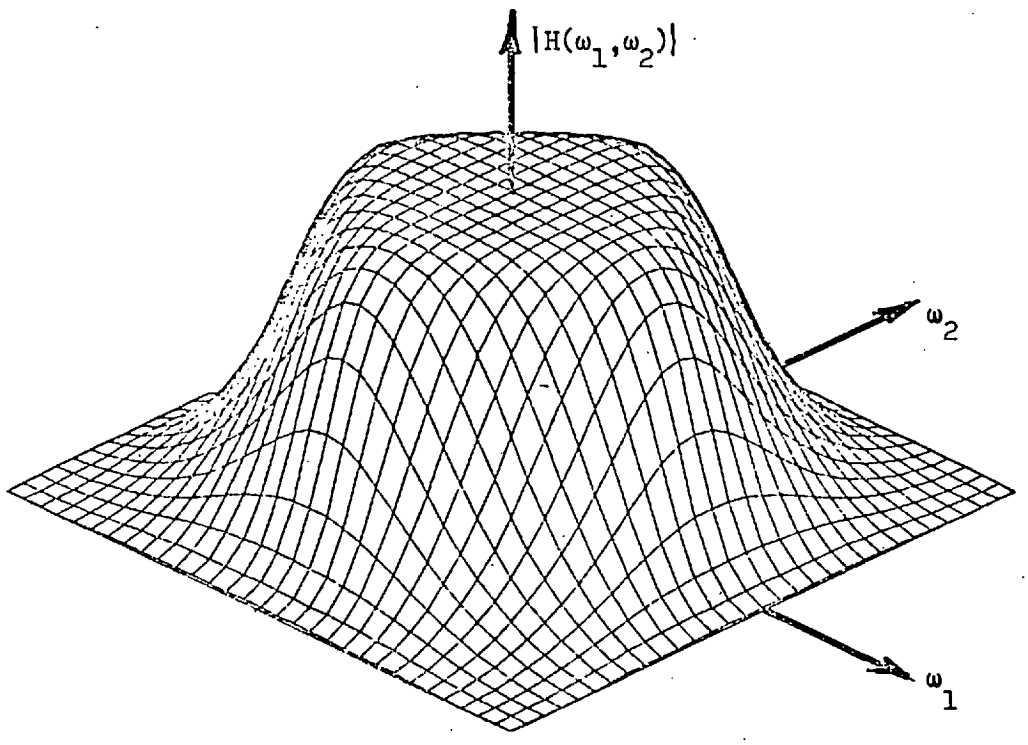


Figure 4.6. Ahmadi filter with separable low-pass guard filter. Specification as in Fig.4.4.

is to note that as the amplitude functions have polar symmetry, we need only consider the first and second quadrant functions and that the amplitude may be obtained from the square of the product of these two functions. We thus design the filter to have desired cut-off frequencies along the two axes, $w_1 = 0$, $w_2 = 0$ and along the diagonal $w_1 = w_2$; this will give a suboptimal solution to the design of a given cut-off boundary.

Consider a design in which an approximation to circular symmetry is desired and the cut-off frequency in the two-dimensional discrete frequency plane is specified as w_c. Thus at the three points $(w_c,0)$, $(0,w_c)$ and $(\frac{w_c}{\sqrt{2}}, \frac{w_c}{\sqrt{2}})$ the amplitude of the response must have the required cut-off magnitude.

At the point $(\frac{w_c}{\sqrt{2}},\frac{w_c}{\sqrt{2}})$ on the principal diagonal the amplitude is the product of that due to the first quadrant function, which is dependent upon the prototype chosen and that of the second quadrant function which is unity along the whole of the line $s_1 = s_2$ since it has the value of the prototype at $s = 0$.

At the points $(w_c,0)$ and $(0,w_c)$ the amplitude is the product of two functions, both of which are dependent on the fall-off rate of the prototype filter.

Thus we must design the transformation of equation (4.22) such that $(w_c,0)$ and $(0,w_c)$ map into the frequency Ω_c' in the analogue and $(w_c/\sqrt{2}, w_c/\sqrt{2})$ maps into Ω_c. Where the prototype filter is of Chebyshev form, the cut-off frequency of a cascade of two identical filters is the same as that of the single filter and $\Omega_c' = \Omega_c$. However, for Butterworth filters Ω_c and Ω_c' are related by

$$\Omega_c' = (2^{1/n} - 1)^{\frac{1}{2}} \Omega_c = k \Omega_c \tag{4.23}$$

where n is the order of Butterworth filter. (When n = 3, k = 0.86).

The design criteria for a two-dimensional filter having a pseudo-circular cut-off contour is that the prototype must have a cut-off frequency Ω_c given by

$$k \Omega_c = \tan w_c/2 \tag{4.24}$$

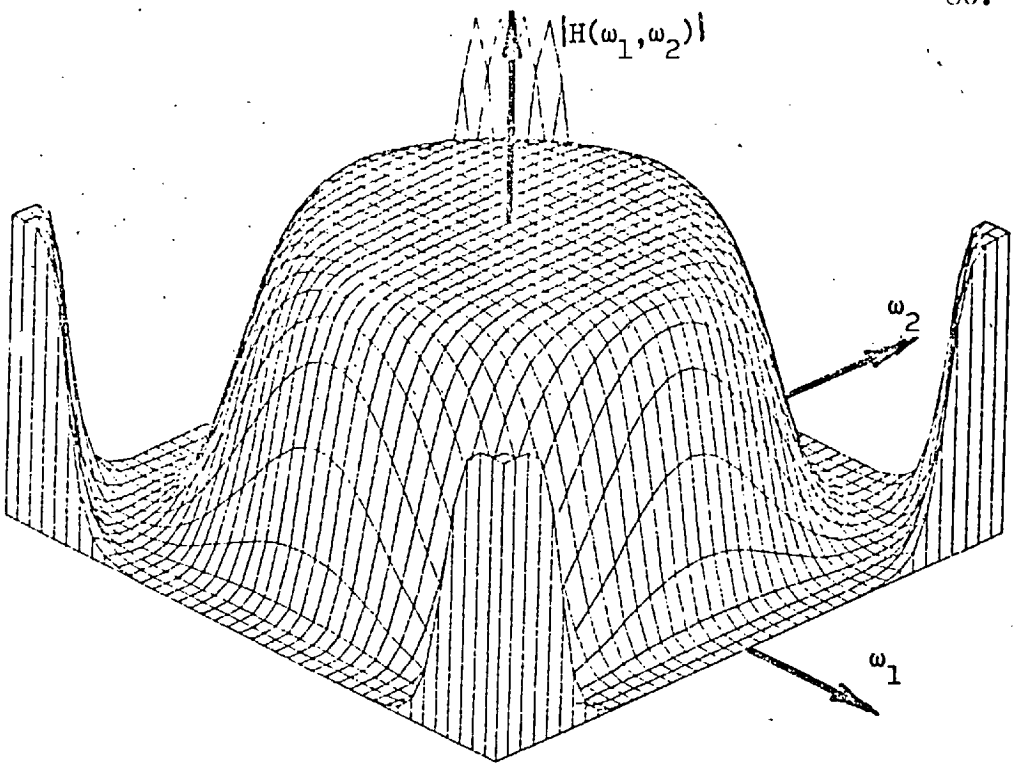
and the transformation coefficients $a_1 = a_2 = 1$ and

$$b = \frac{\tan(w_c/2) - 2k \tan(w_c/2\sqrt{2})}{\tan(w_c/2)\tan^2(w_c/2\sqrt{2})} \tag{4.25}$$

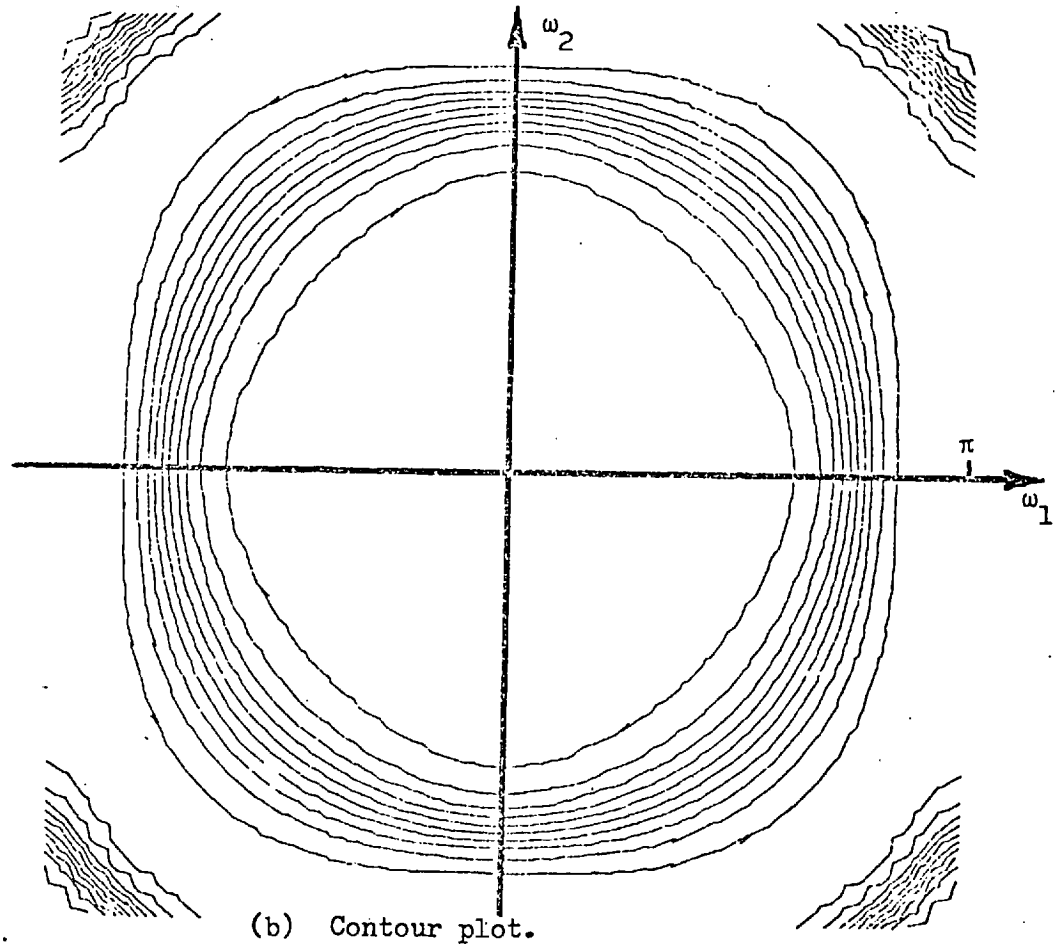
Since the transformation (4.22) must represent a stable function, $b > 0$; this imposes a constraint on the value of w_c for which a positive solution to (4.25) exists. For a third order Butterworth filter w_c must be greater than about 0.61π ; higher order Butterworth filters tend to the limits given for a Chebyshev prototype, for which $w_c > 0.70\pi$.

Fig. 4.7 shows the transfer function and contour plot of a two-dimensional filter designed to a cut-off boundary of 0.72π. It is confirmed that a circular cut-off boundary may be obtained using the Ahmadi tramsformation provided the cut-off frequency is within certain bounds. This transfer function may be compared with those in Figs. 4.4 and 4.5 which were designed for cut-off boundaries lying outside this permitted range and for which circular profiles were not obtainable.

The position is exacerbated by the fact that in the designs considered, we have only cascaded two single quadrant functions,



(a) Isometric plot.



(b) Contour plot.

Figure 4.7. Ahmadi filter. Prototype: third order Butterworth analogue filter; $\omega_o = 0.72\pi$. Transformation; $a_1 = a_2 = 1$, $b = 0.16$.

whereas to obtain a zero-phase function all four quadrants need to be cascaded.

An improvement on the above design technique has been proposed by Ali and Constantinides [57] who show that the spurious pass bands around ($\pm\pi$, $\pm\pi$) may be eliminated if b = 0. This is, of course, true and the transformation then degenerates to the simple form

$$s = a_1 s_1 + a_2 s_2 \tag{4.26}$$

This does, however, introduce the added disadvantage that a filter having pseudo-circular symmetry is now constrained by equation (4.25) to be one in which

$$\tan(w_c/2) = 2k \tan(w_c/2\sqrt{2}) \tag{4.27}$$

and thus design for a pseudo-circular cut-off boundary is only possible for one specific value of design cut-off frequency. For a Chebyshev filter equation (4.27) is satisfied only at $w_c = 0.70\,\pi$ and hence for certain specifications of cut-off frequency, the boundaries may deviate appreciably.

It may thus be seen that the Ahmadi transformation may be used to design two-dimensional near-circular filters with cut-off frequencies greater than a given bound or below that bound with cut-off profiles deviating more from circular but with the disadvantage of requiring a guard filter to remove certain high frequency pass bands. The modification of Ali and Constantinides ensures that these pass bands do not exist, but also completely precludes the possibility of the design of filters having suboptimally designed circular cut-off profiles except in trivial cases.

4.4.7 All-Pass Transformation

In Section 4.3 we have stated that Prendergrass has used a two-dimensional all-pass transfer function to modify the frequency response of a designed low pass two-dimensional digital filter.

Kap [23] has used the same transformation, but applied it to a low pass one-dimensional digital filter to generate a two-dimensional low pass filter with near-circular cut-off contour. The transformation from one dimension to two is given by

$$z = G(z_1,z_2) = \frac{\alpha_o + \alpha_1 z_1^{-1} + \alpha_2 z_2^{-1} + z_1^{-1} z_2^{-1}}{1 + \alpha_2 z_1^{-1} + \alpha_1 z_2^{-1} + \alpha_o z_1^{-1} z_2^{-1}} \tag{4.28}$$

The conditions that $G(z_1,z_2)$ represents a stable transfer function are [10]

$$\begin{aligned} &|\alpha_1| < 1, \\ &|\alpha_1 + \alpha_o| < |1 + \alpha_2| \\ &|\alpha_1 - \alpha_o| < |1 - \alpha_2| \end{aligned} \tag{4.29}$$

A further consideration here is that the transformation gives low pass one-dimensional to low pass two-dimensional along both axes and also along any radius through the origin. This may be ensured by requiring that equation (4.28) maps $w = 0$ in one dimension into $(w_1,w_2) = (0,0)$ in two dimensions and also $w = \pi$ into $(w_1,w_2) = (0, \pi)$, $(\pi,0)$ and (π,π) simultaneously. The first three conditions are satisfied by any function of the form of equation (4.28) but the last demands the additional constraint that

$$\alpha_1 + \alpha_2 = 1 + \alpha_0 \tag{4.30}$$

This apparently violates the stability constraint given by the third of equations (4.29); nevertheless it may be seen that at the point $(w_1,w_2) = (\pi,\pi)$, corresponding to $(z_1,z_2) = (-1,-1)$, both the numerator and denominator of $G(z_1,z_2)$ tend to zero and thus the function does not show instability at this point as the pole and zero cancel one another [28]. This is a non-essential singularity of the second kind.

The above constraints reduce the number of design parameters to two; if, in addition, we require a symmetrical transfer function along the two axes we must set

$$\alpha_1 = \alpha_2 \tag{4.31}$$

This leaves us with a single design parameter α_o to which α_1 and α_2 are related by

$$\alpha_1 = \alpha_2 = \frac{1 + \alpha_o}{2} \tag{4.32}$$

and the stability constraints of (4.29) reduce to the single condition

$$|\alpha_o| < 1 \tag{4.33}$$

The design to an approximately circularly cut-off contour may be approached in the same manner as in the Ahmadi filter. The most obvious approach is by a direct optimization of the error between the designed filter and the specified circular profile. This results in undue computational complexity and a simpler sub-optimal solution may be obtained by constraining the profile to have a fixed radius from the origin along the coordinate axes and along $w_1 = w_2$.

The design equations for such a filter are

$$\tan(w_o/2) = \frac{1 - \alpha_o}{1 + \alpha_o} \tan(w_c/2\sqrt{2}) \tag{4.34}$$

and

$$\tan(w_o'/2) = \frac{1 - \alpha_o}{1 + \alpha_o + 2\alpha_1} \tan(w_c/2) \tag{4.35}$$

where w_o is the cut-off angular frequency of the one-dimensional prototype digital filter and w_o' is the cut-off frequency of two cascaded filters. In addition

$$\tan(w_o'/2) = k \tan(w_o/2) \tag{4.36}$$

where k is given by equation (4.23). We may see from equations (4.34) and (4.35) that a solution is possible when constrained by (4.32) for only one unique value of w_c given by

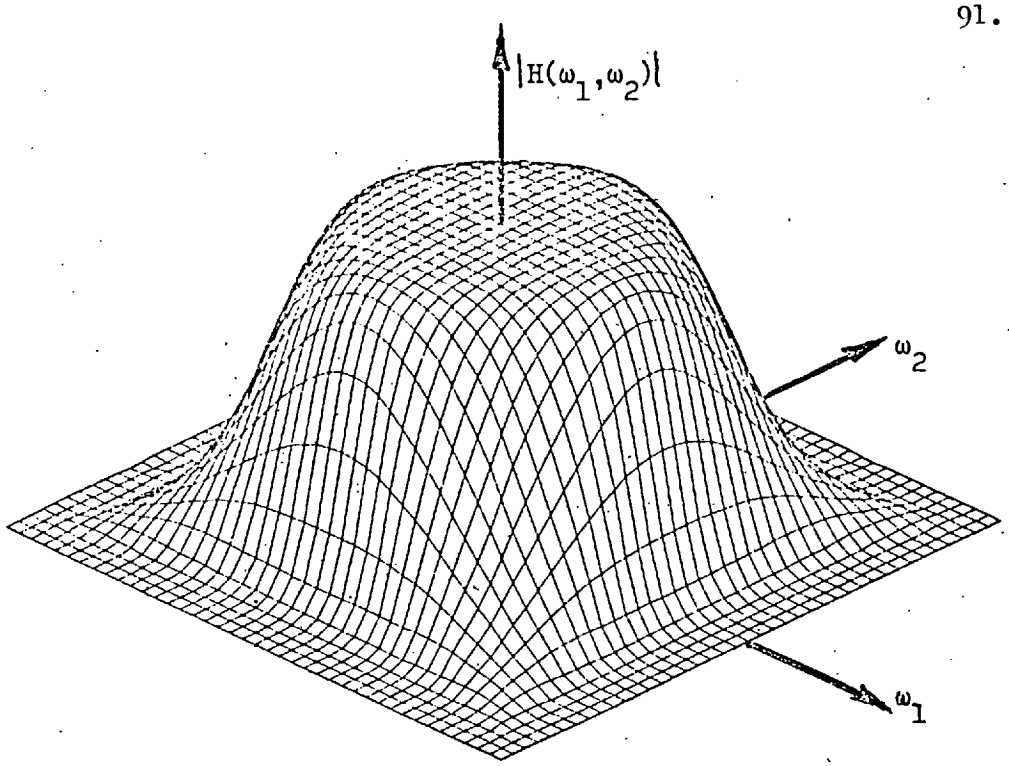
$$\tan(w_c/2) = 2\tan(w_c/2\sqrt{2}) \tag{4.37}$$

A filter designed close to this specification is shown in Fig. 4.8 with $\alpha_o = 0$ and $w_o = \pi/2$.

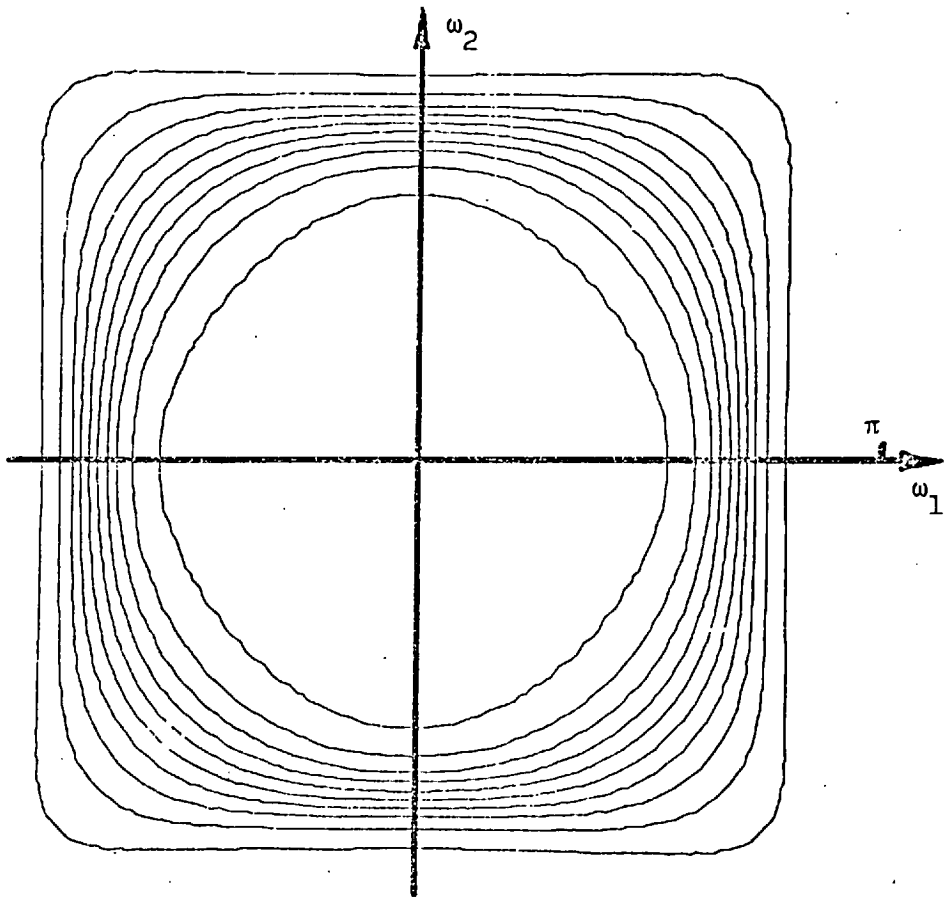
It is thus apparent that this design technique suffers from the same disadvantage as that proposed by Ali et al [57] and only approximations can be made to circularly symmetric filters except perhaps for one specific value of w_c.

This suggests that in order to obtain greater freedom in design, it might be worthwhile relaxing the constraints of equation (4.32). This will most certainly result in the reintroduction of the undesired pass bands which were also present in the Ahmadi transformation.

The stability constraints of equations (4.29) together



(a) Isometric plot.



(b) Contour plot.

Figure 4.8. All-pass transformed filter. Prototype: third order Butterworth digital filter; $\omega_c = \pi/2$. Transformation: $\alpha_0 = 0$, $\alpha_1 = \alpha_2 = 0.5$.

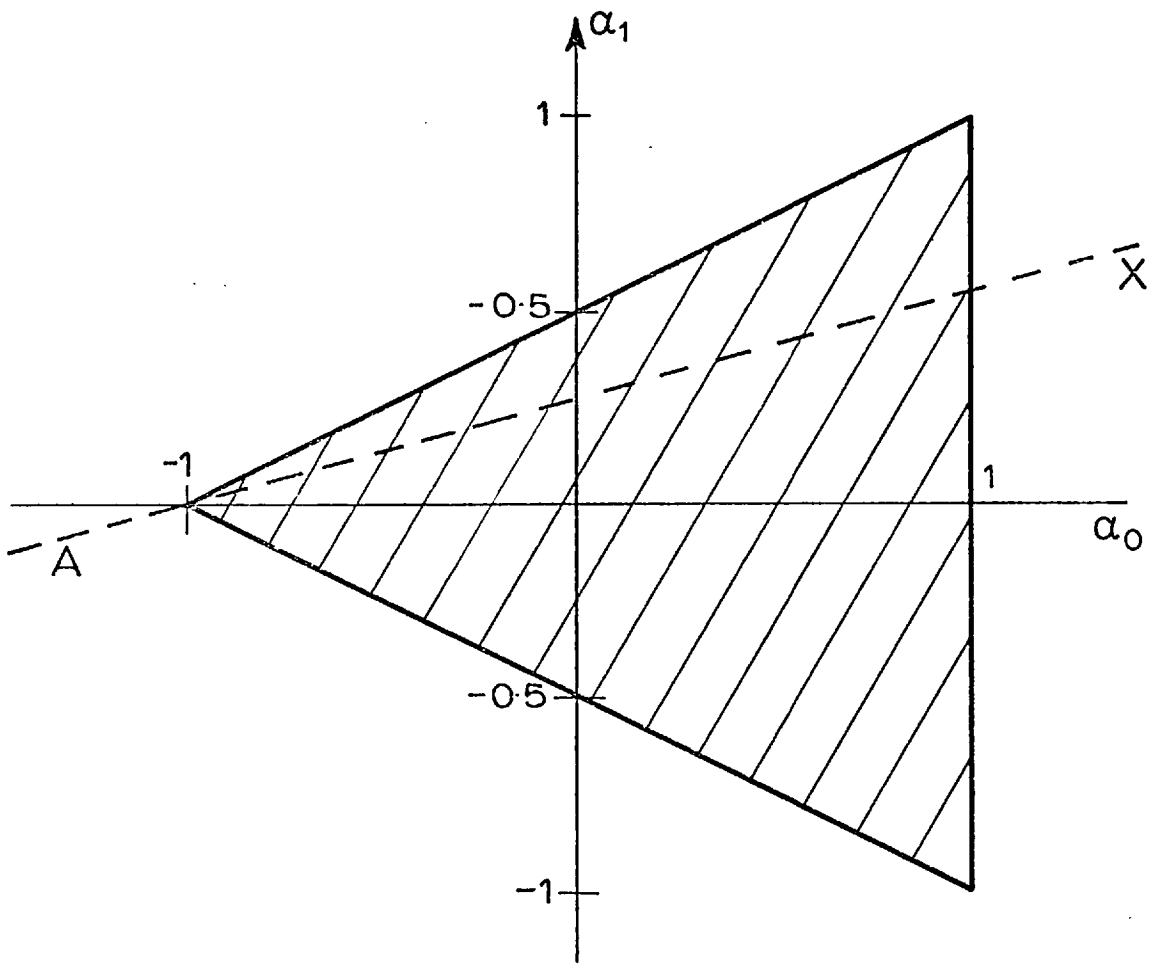


Fig.4·9 Stability region for all-pass transformation.

with the condition $\alpha_1 = \alpha_2$ may be represented diagrammatically as shown in Fig. 4.9. The all-pass transformation is stable for all (α_o, α_1) within the shaded triangle. The constraint of equation (4.30) permits values only on the oblique boundaries of this region. With this constraint removed, operation is permissible at all points within the region.

Considering design of a Chebyshev filter in which $w_o = w_o'$ equations (4.34) and (4.35) lead to

$$1 + \frac{2\alpha_1}{1 + \alpha_o} = \gamma = \frac{\tan(w_c/2\sqrt{2})}{\tan(w_c/2)} \qquad (4.39)$$

which represents the line AX having gradient $\frac{\gamma - 1}{2}$.

Thus any filter with design circular profile cut-off frequency, w_c, which satisfies

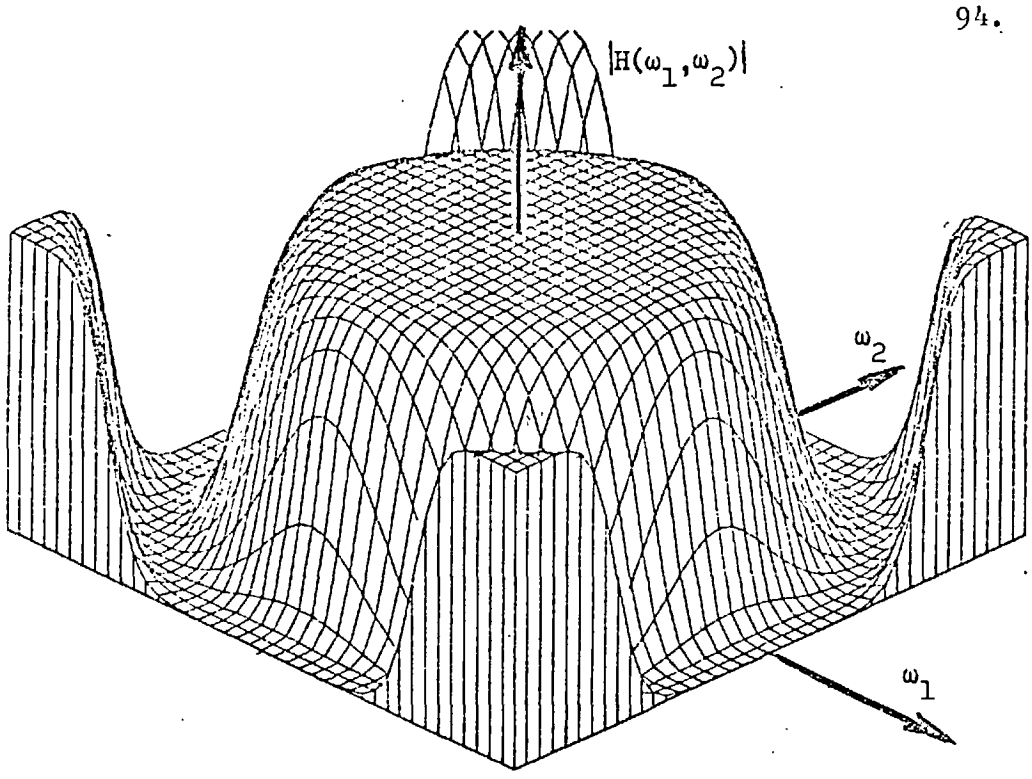
$$\left|\frac{\gamma - 1}{2}\right| < \tfrac{1}{2} \quad ; \quad \text{namely } 0 < \gamma < 2 \qquad (4.39)$$

will be capable of design by such a transformation.

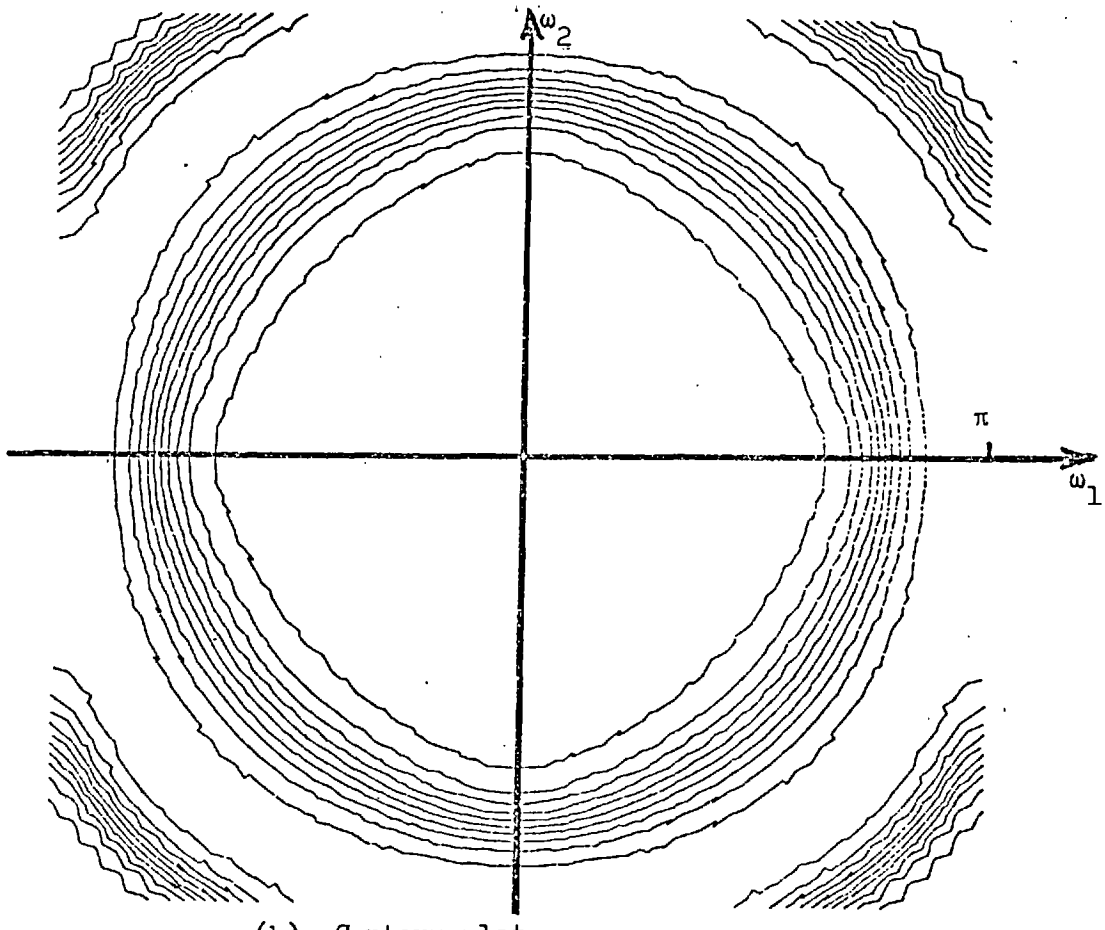
This limitation is similar to that which is imposed on the design of filters using the Ahmadi technique and restricts the transformation to design cut-off frequencies greater than approximately 2.2 radians/sec. A filter designed using this method is shown in Fig. 4.10.

4.4.8 Comparison of Two-Dimensional Transformations

It is apparent that attempts to obtain two-dimensional recursive filters having transfer functions whose cut-off boundaries closely approximate to a circle over a range of design frequencies



(a) Isometric plot.



(b) Contour plot.

Figure 4.10. All-pass tansformed filter. Prototype: third order Butterworth digital filter; $\omega_o = 0.5764$. Transformation; $\alpha_o = 0.5$, $\alpha_1 = \alpha_2 = 0.4417$.
Alternatively; $\omega_o = 1.726$, $\alpha_o = 0$, $\alpha_1 = \alpha_2 = 0.2957$.

are singularly unsuccessful. The McClellan technique is only applicable to non-recursive systems and thus results in filters of high complexity. The Bernabo technique resolves this problem but substitutes for this shortcoming a considerable increase in complexity of the design procedure; the technique is, however, capable of generating filters whose cut-off profile is closely circularly with no restrictions on the range of cut-off frequency.

The Ahmadi transformation is considerably simpler but demands the inclusion of a guard filter and cannot approximate circular profiles except over a limited cut-off frequency range. The all-pass transformation removes the spurious high-pass bands from the response and thus the need for a guard filter, but the transformation is restricted to the design of good circular profile filters at only one single cut-off frequency. The same comment applies to the modification to the Ahmadi transformation proposed by Ali et al.

The simplest technique is the separable product transformation which gives the poorest approximation to circular symmetry. However, subjective tests have shown that in processing certain images very little difference is observable, whichever of the various design techniques is used, and the insistent demand for circular cut-off profiles is probably misplaced.

4.5 ONE DIMENSION TO MULTIDIMENSION TRANSFORMATION

Although the application of digital filters of higher order than three appears remote, the design of the general case may easily be included in any extension of the one to two

dimensional transformations. One of the most likely cut-off profiles which might be specified would be that having multidimensional spherical boundaries. We shall therefore consider three techniques whereby this may be achieved.

To illustrate the amplitude responses of the designed multidimensional filters graphically would be difficult and so all examples are restricted to three dimensions. In three dimensions it is possible to present the response of the filter visually as a plot of the cut-off isometric boundary surface. All the filters will be designed with low pass characteristics, namely having full transmission at the origin and the cut-off boundary plotted will correspond to the "3 dB" surface, namely a value of 0.7. As all the filters will be symmetrical in all eight primary sectors, the profile will be plotted for positive values of w_3 and both polarities of w_1 and w_2.

4.5.1 Separable Product Technique

The two-dimensional separable product technique discussed in Section 4.4.1 may be extended to any number of dimensions in a relatively trivial manner. The cut-off boundary surface will approximate to a multidimensional rectangular parallelipiped.

The profile of the cut-off isometric surface of a three-dimensional separable product filter is shown in Fig. 4.11. The design is extremely simple; the resulting filter may be made symmetric about all axes by simple transformation and stability is ensured. The designed filter is, of course, not zero-phase. To achieve this, one would need to cascade two identical sets of filters recursing in opposite directions.

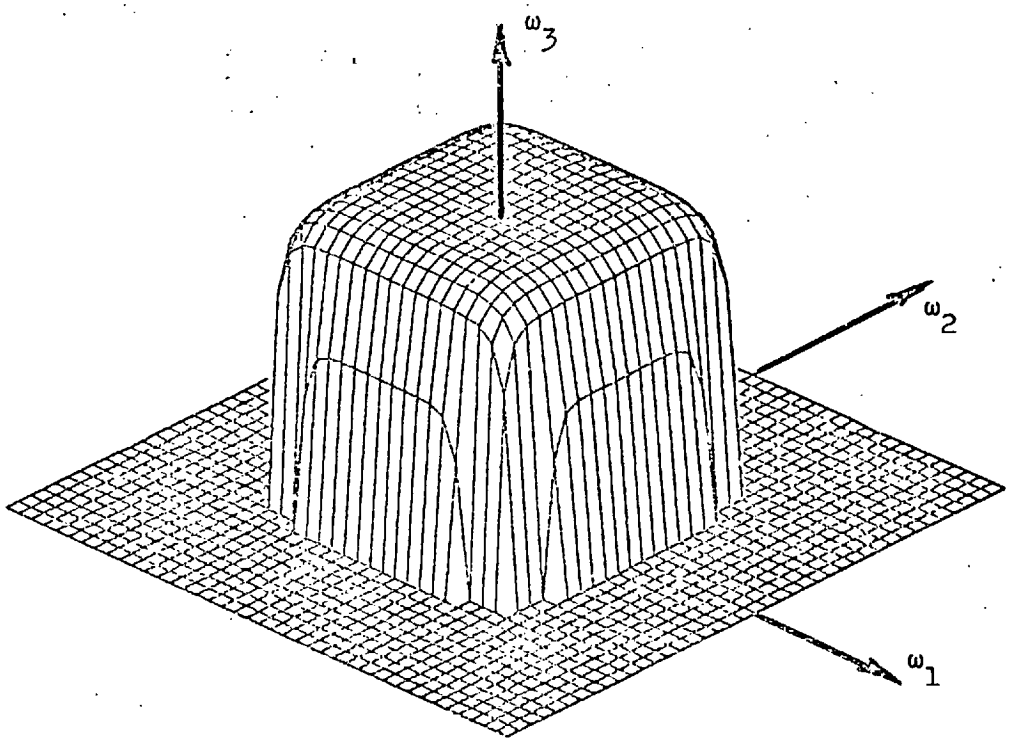


Figure 4.11. Cut-off surface of three-dimensional separable product filter. Prototype: third order Butterworth digital filter; $\omega_o = \pi/2$.

4.5.2 Multidimensional McClellan Transformation

The McClellan transformation [9,44] discussed in Section 4.4.4 may be easily extended to many dimensions [58].

A class of N-dimensional zero-phase FIR filters may be shown to have a frequency response which is of the form

$$H(e^{ju_1},\ldots,e^{ju_n}) = \sum_{m_1=0}^{M_1} \cdots \sum_{m_n=0}^{M_n} a(m_1,\ldots,m_n)\cos m_1u_1 \ldots \cos m_nu_n \tag{4.40}$$

The transformation, by analogy with the two-dimensional case, takes the form

$$\cos u = \sum_{p_1=0}^{P_1} \cdots \sum_{p_n=0}^{P_n} t(p_1,\ldots,p_n)\cos p_1u_1 \ldots \cos p_nu_n \tag{4.41}$$

This transformation may be applied to the one-dimensional transfer function

$$H(e^{ju}) = \sum_{n=0}^{N} h(n) \cos nu \tag{4.42}$$

which may be alternatively written, using the Chebyshev polynomial functions, as

$$H(e^{ju}) = \sum_{n=0}^{N} b(n) (\cos u)^n$$

The resultant multidimensional transfer function is

$$H(e^{ju_1},\ldots,e^{ju_n}) = \sum_{n=0}^{N} b(n) \left[\sum_{p_1=0}^{P_1} \cdots \sum_{p_n=0}^{P_n} t(p_1,\ldots,p_n)\cos p_1u_1 \ldots \cos p_nu_n \right]^n \tag{4.43}$$

Again by application of Chebyshev polynomials this may be rewritten in the form of equation (4.40).

The original prototype transfer function is specified by the parameters h(n), which determine b(n), and are responsible for the shape of the amplitude response in the pass, stop and transition regions. The parameters $t(p_1,\ldots,p_n)$ control the mapping of a single frequency in one dimension into a region in the multidimensional frequency domain and thus are responsible for determining the shape of the cut-off profile.

The most pertinent profile in our case is that in which the cut-off profile is of the form of a hypersphere having a pass band edge determined by

$$\sum_{i=1}^{n} u_i^2 = R^2$$

The simplest transformation to consider is that in which $P_1 = P_2 = \ldots = P_n = 1$ in equation (4.41). This provides 2^n parameters $t(p_1,\ldots,p_n)$ which may be used to control the shape of the cut-off profile. One constraint imposed in the usual case of a low pass to low pass transformation is that the one-dimensional origin is mapped into the n-dimensional origin, requiring

$$\sum_{p_1=0}^{1} \cdots \sum_{p_n=0}^{1} t(p_1,\ldots,p_n) = 1$$

The remaining 2^n-1 parameters may be determined by solving equation (4.41) for one of the frequency variables, say u_n, in terms of the other parameters and the design cut-off frequency; the error between this value and the desired value is then minimized on the

assumption that the required profile is a perfect hypersphere. In the multidimensional case this task is usually too formidable. A suboptimal approach is frequently used which assumes that if the mapping were exact, the value of the function would be constant on the desired hypersphere; the error between the calculated value of the function on the hypersphere and its ideal value may be minimized to give a solution.

4.5.3 Multidimensional Bernabo Design Technique

The extension of the application of the McClellan transformation to recursive filters given by Bernabo et al [17] to n-dimensions has been given by Ahmadi [18]. It follows the general procedure for two dimensions, of transforming the magnitude square of a one-dimensional amplitude response to n dimensions; this is followed by the n-dimensional decomposition technique [22] to obtain a set of stable one-quadrant filters.

A zero-phase recursive filter transfer function may be written

$$H(e^{ju_1},\ldots,e^{ju_n}) = \frac{\sum_{m_1=0}^{M_1}\cdots\sum_{m_n=0}^{M_n} p(m_1,\ldots,m_n)\cos m_1u_1\ldots\cos m_nu_n}{\sum_{\ell_1=0}^{L_1}\cdots\sum_{\ell_n=0}^{L_n} q(\ell_1,\ldots,\ell_n)\cos \ell_1u_1\ldots\cos \ell_nu_n} \qquad (4.44)$$

Consider a one-dimensional recursive filter with response function

$$H(e^{ju}) = \frac{\sum_{n=0}^{N} a(n)\cos nu}{\sum_{m=0}^{M} b(m)\cos mu} \qquad (4.45)$$

which may alternatively be written as

$$H(e^{ju}) = \frac{\sum_{n=0}^{N} a'(n)(\cos u)^n}{\sum_{m=0}^{M} b'(m)(\cos u)^m} \tag{4.46}$$

Let this be transformed to an n-dimensional filter via

$$\cos u = \sum_{p_1=0}^{P_1} \cdots \sum_{p_n=0}^{P_n} t(p_1,\ldots,p_n)\cos p_1u_1 \ldots \cos p_nu_n \tag{4.47}$$

resulting in

$$H(e^{ju_1},\ldots e^{ju_n}) = \frac{\sum_{n=0}^{N} a'(n)\left[\sum_{p_1=0}^{P_1}\cdots\sum_{p_n=0}^{P_n} t(p_1,\ldots,p_n)\cos p_1u_1\ldots\cos p_nu_n\right]^n}{\sum_{m=0}^{M} b'(m)\left[\sum_{p_1=0}^{P_1}\cdots\sum_{p_n=0}^{P_n} t(p_1,\ldots,p_n)\cos p_1u_1\ldots\cos p_nu_n\right]^n} \tag{4.48}$$

This may be rearranged using the recurrence formulae of Chebyshev polynomials to give an expression of the form of equation (4.44).

The approximation to a desired boundary condition may be obtained by an optimization technique similar to that used in the two-dimensional case, after first having designed the one-dimensional prototype filter to give the desired frequency response in the pass, stop and transition bands. The final step in the design technique is to stabilize the resultant filter. As the procedure inevitably results in a zero-phase filter, the decomposition technique of Ahmadi and King [22] must be used to generate a set of single quadrant stable recursive filters.

4.5.4 Multidimensional Ahmadi Technique

The obvious extension of the Ahmadi technique to n-dimensions would be to use the general n-dimensional first-degree reactance function

$$F(s_1,\ldots,s_n) = \frac{\sum_{i=1}^{n} a_i s_i + \sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=j}^{n} a_{ijk} s_i s_j s_k + \cdots}{1 + \sum_{i=1}^{n}\sum_{j=i}^{n} b_{ij} s_i s_j + \sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=j}^{n}\sum_{\ell=k}^{n} b_{ijk\ell} s_i s_j s_k s_\ell + \cdots} \tag{4.49}$$

However, this general form of function is not stable. This may be seen from a consideration of the stability conditions for a multi-dimensional system, which may be formulated

$$\mathrm{Re}\left[F(s_1,\ldots,s_2)\right] > 0 \qquad \forall\left\{s_i,\ i=1,\ldots,n : \bigcap_{i=1}^{n} (\mathrm{Re}(s_i) > 0)\right\}$$

Superficial inspection of expression (4.49) shows that this cannot hold for unconstrained values of the coefficients in any but the most trivial cases. We shall therefore restrict our studies to the three-dimensional case and show how we may obtain necessary and sufficie nt conditions for the stability of a multidimensional function generated from lower dimensional functions. It will not, of course, be possible to obtain closed form expressions for the stability conditions as functions of the coefficient parameters.

Consider the first degree three-dimensional reactance function

$$F(s_1,s_2,s_3) = \frac{a_1 s_1 + a_2 s_2 + a_3 s_3 + a_4 s_1 s_2 s_3}{1 + b_1 s_2 s_3 + b_2 s_3 s_1 + b_3 s_1 s_2} \tag{4.49a}$$

It may be noted that this function in general does not satisfy the stability conditions. However, we may derive a set of necessary and

sufficient conditions for (4.49a) to represent a three-dimensional reactance function.

Let us make the one- to two-dimensional transformation of equation (4.22) as

$$s = G_1(s_1', s_2') = \frac{\alpha_{11}s_1' + \alpha_{21}s_2'}{1 + \beta_1 s_1' s_2'} \tag{4.50}$$

One or both of these variables, s_1', s_2', may be transformed by similar expressions to give a three- or four-dimensional function. Considering the three-dimensional situation we make the substitution

$$\begin{aligned} s_1' &= G_2'(s_1) = s_1 \\ s_2' &= G_2''(s_2, s_3) = \frac{\alpha_{12}s_2 + \alpha_{22}s_3}{1 + \beta_2 s_2 s_3} \end{aligned} \tag{4.51}$$

Concatenation of these two substitutions results in an expression representing a third order reactance function of the form of equation (4.49a) but with the added constraints that

$$\begin{aligned} a_3 b_2 &= a_2 b_3 \\ \text{and} \qquad a_1 b_1 &= a_4 \end{aligned} \tag{4.52}$$

Now the necessary and sufficient conditions for (4.22) to represent a reactance function are

$$a_1 > 0, \quad a_2 > 0 \text{ and } \quad b > 0.$$

These may be applied directly to (4.50) and (4.51). So long as these conditions are satisfied, we are able, by choice of the α_{ij}, β_j to generate the complete set of all three-dimensional reactance functions, since

$$F(s_1, s_2, s_3) = G_1\left[G_s'(s_1),\ G_2''(s_2, s_3)\right]$$

where G_1, G_2' and G_2'' are positive real reactance functions.

This approach may be used to generate a transformation from one dimension to many dimensions in a series of steps each increasing the dimensionality of the function by at least one. (For higher order filters two or more transformations may be simultaneously applied.)

Considering the three-dimensional case, we note that in practice only five design parameters are available since no loss of generality ensues by setting $\alpha_{12} = 1$. We then have three parameters defining the cut-off frequencies along the three coordinate axes and two other parameters which may be used to control the shape of the boundary surface at two intermediate points.

The shape of the profile is probably most simply determined by constraining it to pass through five points mapped by (4.50) and (4.51) in the three-dimensional space. For convenience these may be chosen as $(w_c,0,0)$, $(0,w_c,0)$, $(0,0,w_c)$, $(\frac{w_c}{\sqrt{2}}, \frac{w_c}{\sqrt{2}}, 0)$ and $(0, \frac{w_c}{\sqrt{2}}, \frac{w_c}{\sqrt{2}})$ for a pseudo-spherical cut-off boundary. The response at $(\frac{w_c}{\sqrt{2}}, 0, \frac{w_c}{\sqrt{2}})$ will depend upon the order in which the transformations (4.50) and (4.51) are executed; in the example considered it will be identical with that at $(\frac{w_c}{\sqrt{2}}, \frac{w_c}{\sqrt{2}}, 0)$. An alternative would be to force the cut-off profile to pass through $(\frac{w_c}{\sqrt{3}}, \frac{w_c}{\sqrt{3}}, \frac{w_c}{\sqrt{3}})$.

An example of a filter having spherical profile designed from a Butterworth third order filter with cut-off frequency of 0.5π is shown in Fig. 4.12. The cut-off surface is a close approximation to spherical shape although the parameters chosen for the design are outside the range where it is possible to satisfy the constraints for pseudo-spherical boundary accurately.

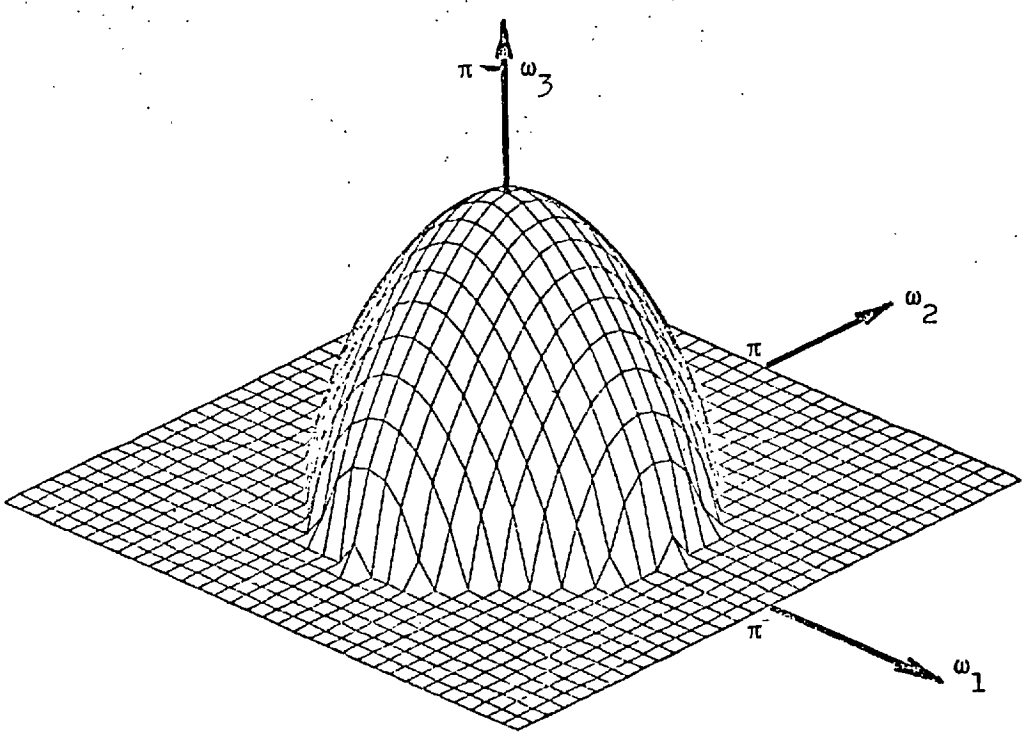


Figure 4.12. Cut-off surface of three-dimensional Ahmadi filter. Prototype: third order Butterworth analogue filter; $\omega_o = \pi/2$. Transformation in both planes; $a_1 = a_2 = 1$, $b^o = 0.2$.

It should be observed that four three-dimensional filters need to be cascaded (eight for zero-phase functions) and hence the cut-off frequencies along the axes $(w_1,0,0)$, $(0,w_2,0)$, $(0,0,w_3)$ will be the result of the superposition of four filters; at points such as $(w_1,w_2,0)$ only two low pass filters will be cascaded; and at points such as $(\pm w_1, \pm w_2, \pm w_3)$ only one filter. When using a Butterworth prototype filter this will modify the design equations at three sets of points.

It may be seen that an approximately spherical filter is obtained but spurious high frequency pass bands will be present unless an appropriate guard filter is provided to eliminate them as was done on the example shown in Fig. 4.12.

4.5.5 Multidimensional All-Pass Transformation

The all pass transformation function may be used to transform to a multidimensional filter. The general n-dimensional all pass function of first degree is

$$G(z_1,\ldots,z_n) = \frac{z_1\cdot\cdot\cdot\cdot\cdot z_n\left(1 + \sum_{i=1}^{n} \alpha_i z_i^{-1} + \sum_{i=1}^{n}\sum_{j=i}^{n} \alpha_{ij} z_i^{-1} z_j^{-1} + \ldots\right)}{1 + \sum_{i=1}^{n} \alpha_i z_i + \sum_{i=1}^{n}\sum_{j=i}^{n} \alpha_{ij} z_i^{-1} z_j^{-1} + \ldots} \qquad (4.53)$$

Before such a transformation function can be applied to a one-dimensional digital filter, it is necessary to establish conditions for its stability. This may be done numerically for any given function by one of the methods discussed in Chapter 2. However, no closed form conditions have been established in the general case.

It is, however, possible to generate a limited class of multidimensional first order all pass functions by successive application of the one- to two-dimensional transformation. Thus the transformation

$$z = \frac{1 + \alpha_{11}z_1' + \alpha_{21}z_2' + \alpha_{01}z_1'z_2'}{\alpha_{01} + \alpha_{21}z_1' + \alpha_{11}z_2' + z_1'z_2'} \tag{4.54}$$

followed by

$$\begin{aligned} z_1' &= z_1 \\ z_2' &= \frac{1 + \alpha_{12}z_2 + \alpha_{22}z_3 + \alpha_{02}z_2z_3}{\alpha_{02} + \alpha_{22}z_2 + \alpha_{12}z_3 + z_2z_3} \end{aligned} \tag{4.55}$$

will realize the overall transformation

$$z = \frac{1 + a_1z_1 + a_2z_2 + a_3z_3 + b_1z_2z_3 + b_2z_3z_1 + b_3z_1z_2 + cz_1z_2z_3}{c + b_1z_1 + b_2z_2 + b_3z_3 + a_1z_2z_3 + a_2z_3z_1 + a_3z_1z_2 + z_1z_2z_3} \tag{4.56}$$

It will be noticed that there are seven parameters in equation (4.56) which may be equated with the six parameters of equation (4.54) and (4.55). Thus the most general form of all pass third order network cannot be generated in this manner. However, it is not difficult to show that the third (and higher) order functions so constructed will be stable if and only if the individual two-dimensional transformations are stable. The necessary and sufficient conditions for these transformations to represent stable network functions are

$$\begin{aligned} &|\alpha_{1i}| < 1 \\ &|\alpha_{1i} + \alpha_{0i}| < |1 + \alpha_{2i}| \\ &|\alpha_{1i}\ \alpha_{0i}| < |1 - \alpha_{2i}| \end{aligned} \tag{4.57}$$

for all $i = 1,\ldots,n$; where n is the dimensionality of the filter.

A relatively direct procedure is now indicated for the design of a multidimensional filter having approximately hyperspherical band edge boundary. It is, of course, true in the multidimensional design, as in the two-dimensional design, that an accurate pseudo-hyperspherical pass band boundary can only be achieved at one frequency of approximately 2.2 rads/sec. However, filters closely approximating hyperspheres may be generated at a wide range of frequencies around this value.

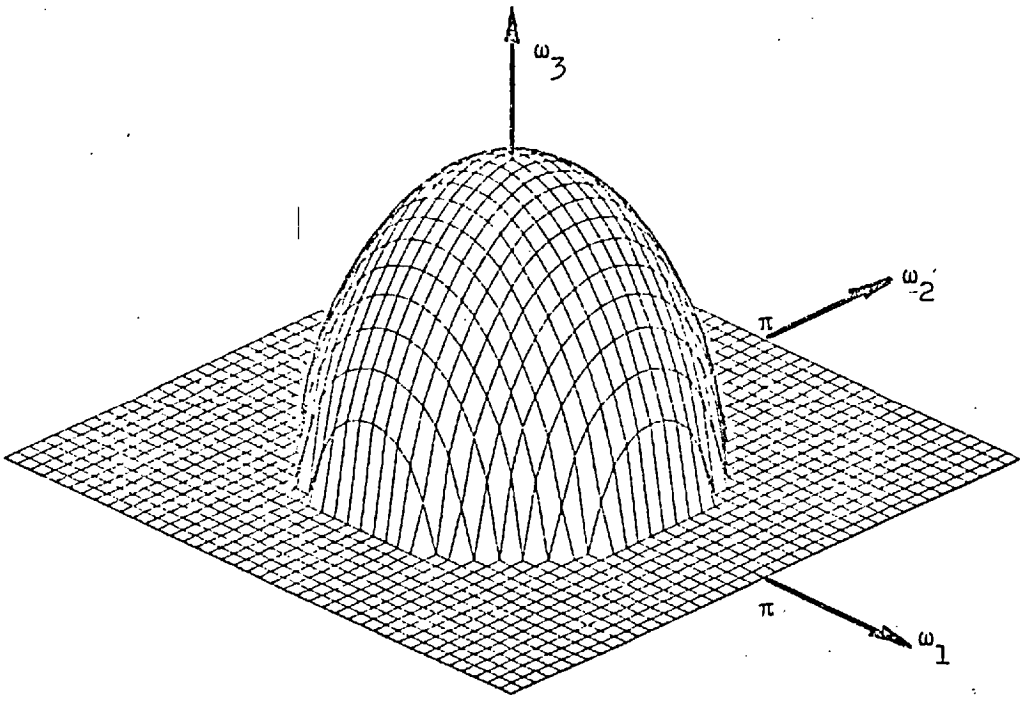
The design is achieved in a number of steps. Starting with the specified design cut-off boundary, w_{cn}, a filter of one dimension lower is specified having cut-off boundary, $w_{c(n-1)}$, given by

$$\tan \frac{w_{c(n-1)}}{2} = \frac{1 - \alpha_{0n}}{1 + \alpha_{0n} + 2\alpha_{1n}} \tan \frac{w_{cn}}{2} \qquad (4.58)$$

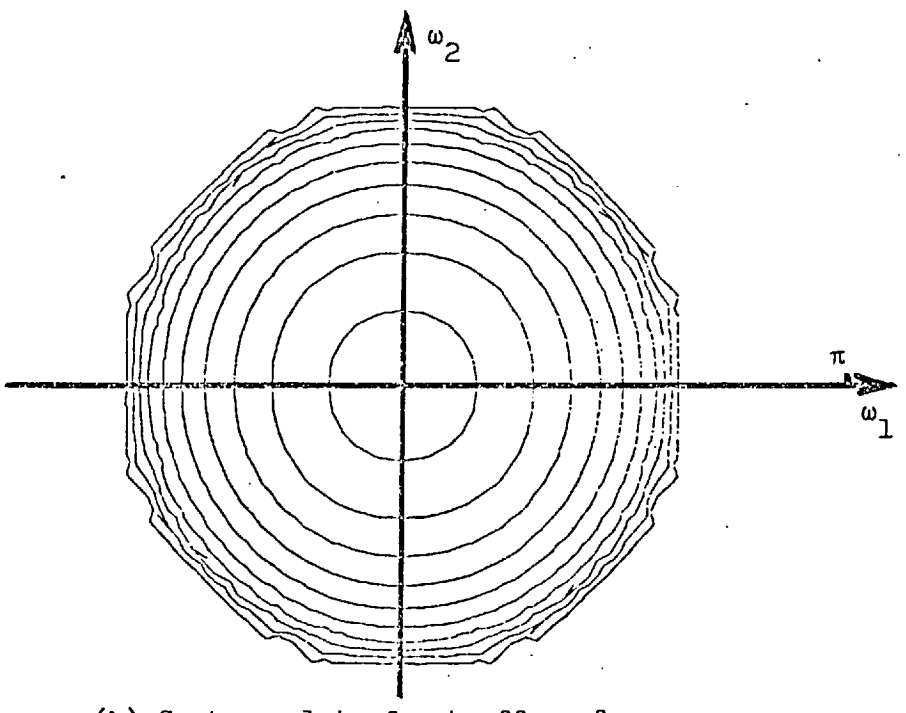
where n is the order of dimensionality of the filter and assuming $\alpha_{1n} = \alpha_{2n}$. This procedure will also permit design of the parameters α_{0n} and α_{1n} ensuring that they are constrained to lie within the stability triangle given in Fig. 4.9.

This successive reduction of the dimensionality of the filter is continued until a one-dimensional filter is obtained which may then be designed according to the specified requirements in the pass, stop and transition bands of the filter.

The choice of parameters α_{0n} and α_{1n} in equation (4.58) may be made either using the Kap restricted form of transformation or the more general form. In the Kap form α_{0n} and α_{1n} are related by



(a) Isometric plot of cut-off surface.



(b) Contour plot of cut-off surface.

Figure 4.13. Cut-off surface for three-dimensional filter via all-pass transformation. Prototype: third order Butterworth digital filter: $\omega_o = \pi/2$. Transformations in both planes; $\alpha_o = 0$, $\alpha_1 = \alpha_2 = 0.5$.

$$1 + \alpha_{0n} = 2\alpha_{1n} \tag{4.59}$$

and a low pass $(n-1)^{th}$ dimension filter is transformed into a low pass n^{th} dimension filter. Using the more general form, added freedom is available in design but one suffers from the disadvantage of having the extra undesirable pass bands at high frequencies.

A three-dimensional filter was designed using this procedure based on a Butterworth low pass prototype digital filter. The profile of the 3 dB isometric surface is shown in Fig. 4.13(a), from which one may subjectively observe satisfactory spherical symmetry. Fig. 4.13(b) shows contours of Fig. 4.13(a) from which it may be seen that good spherical symmetry of this particular pass band boundary is maintained throughout the three-dimensional space.

4.5.6 Comments on Multidimensional Transformation

Of the several transformation techniques studied, the Bernabo design gives good results but requires stabilization techniques for satisfactory implementation. Of the other circular symmetric designs, the all pass transformation gives more closely spherical boundaries, although similar results may be obtained by use of the Ahmadi technique followed by a guard filter.

In all probability, however, it will be shown that the separable product technique, although giving cut off profiles which deviate greatly from spherical, performs in a great many practical situations as satisfactorily as the techniques giving better spherical profiles. It also has the outstanding merit that it is very much simpler to design and to implement.

4.6 APPLICATION OF TWO-DIMENSIONAL CIRCULAR PROFILE FILTER

In the preceding sections we have outlined a number of procedures for designing two-dimensional filters having cut-off profiles which approximate to a circle as it is intuitively felt that such a filter will be most likely to counteract the effects of isotropic distortion introduced between the object and the image.

To add conviction to this argument and to assess the value of the various design techniques, it would be valuable to make a subjective assessment of the improvement in image quality after processing by filters designed by different techniques. In addition the effect of the order of the prototype filter may also be considered. The results given below are mainly due to Kap [23] . The original recorded image is shown in Fig. 4.14 and represents an X-ray picture taken of a diseased human liver; the vertical axis represents the intensity of the picture elements and the horizontal axes the linear dimensions in a lateral plane through the body. It may be observed that the image is corrupted by considerable high frequency noise and an improvement in picture quality for diagnostic purposes may be obtained by the use of a two-dimensional low pass filter.

A number of filter structures were designed and used to process this image.

Fig. 4.15 shows the effect of using a Bernabo designed filter based on a fourth order Chebyshev prototype with $w_c = 0.6\pi$ and passband ripple of 1%; the filter has been decomposed using the Pistor technique into four stable single quadrant filters, each having a denominator function truncated to an 8 x 8 array.

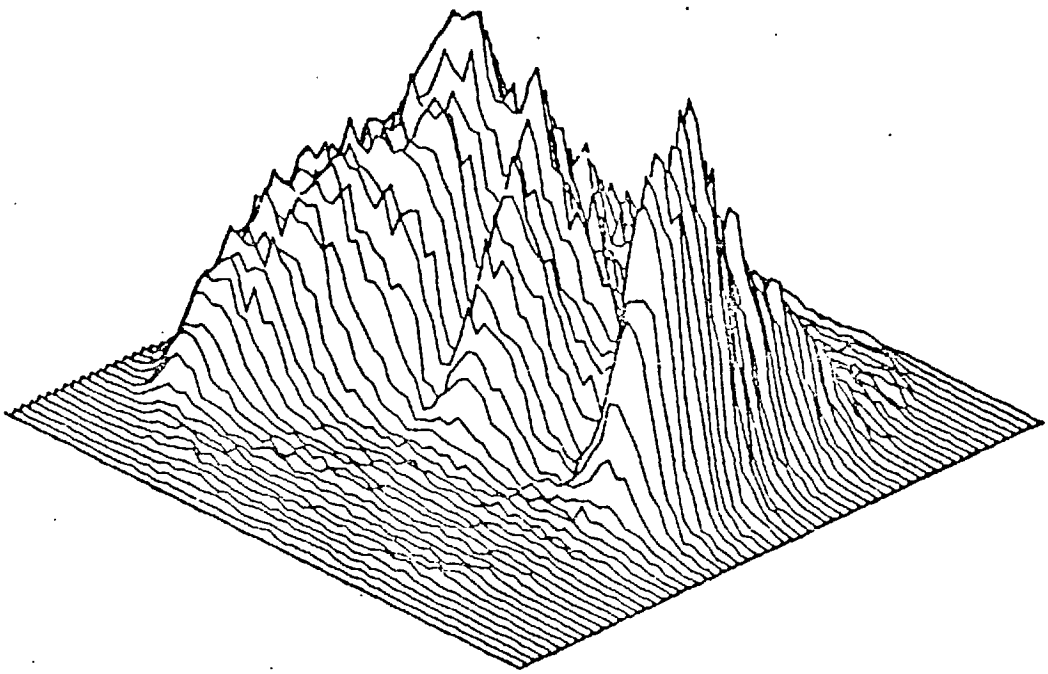


Figure 4.14. Original X-ray image of diseased liver.

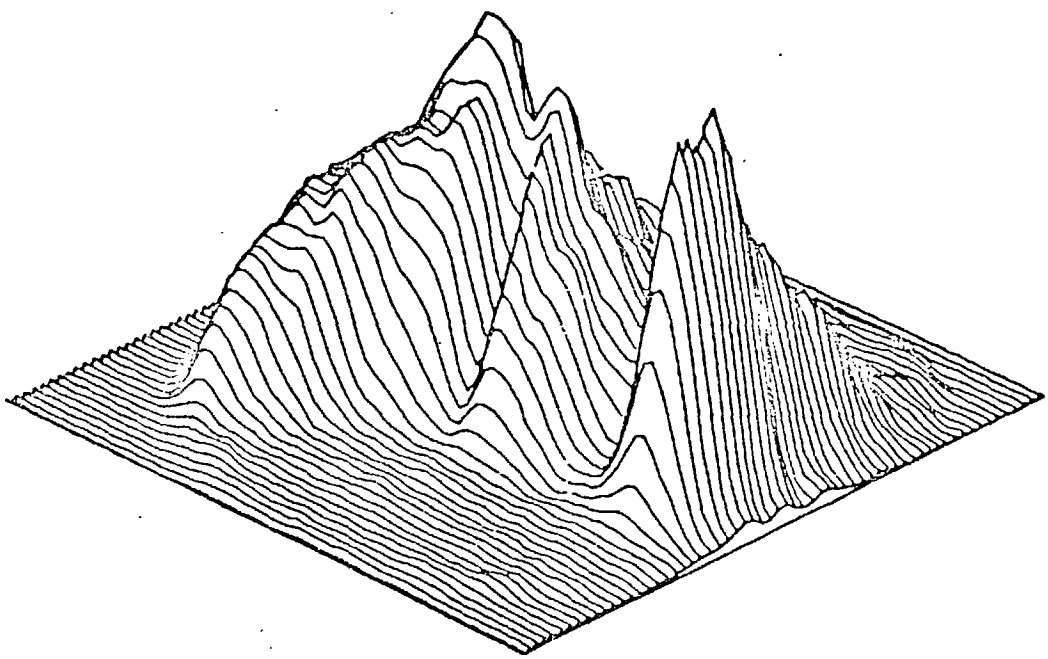


Figure 4.15. Image of Fig. 4.14 processed by a fourth order Bernabo designed filter.

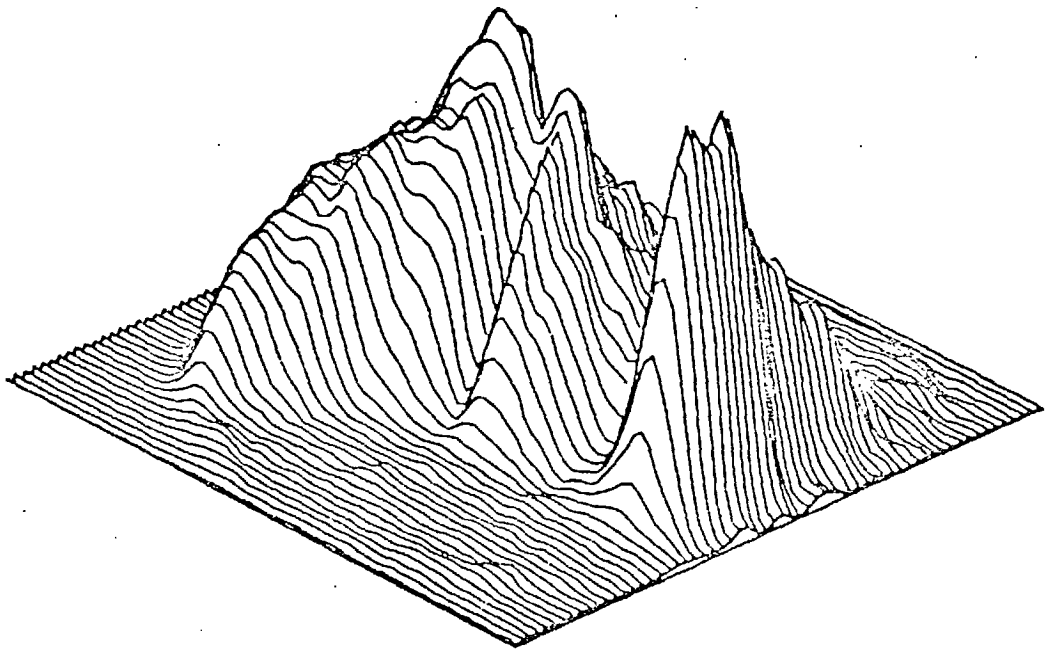


Figure 4.16. Image of Fig. 4.14 processed using a fourth order Ahmadi designed filter.

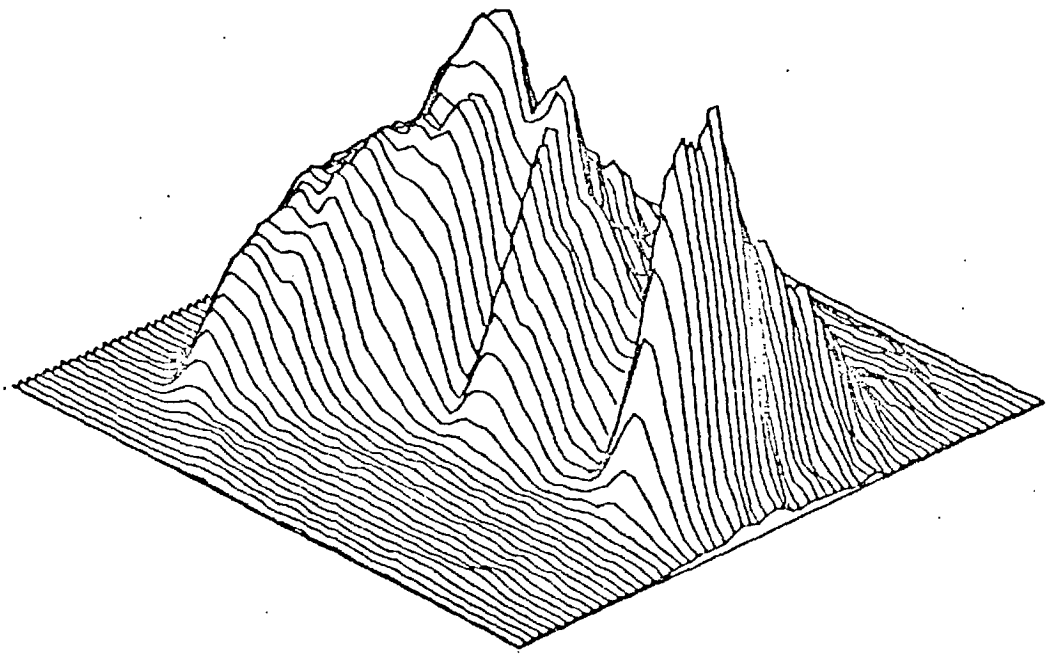


Figure 4.17. Image of Fig.4.14 processed using a fourth order all-pass transformation designed filter.

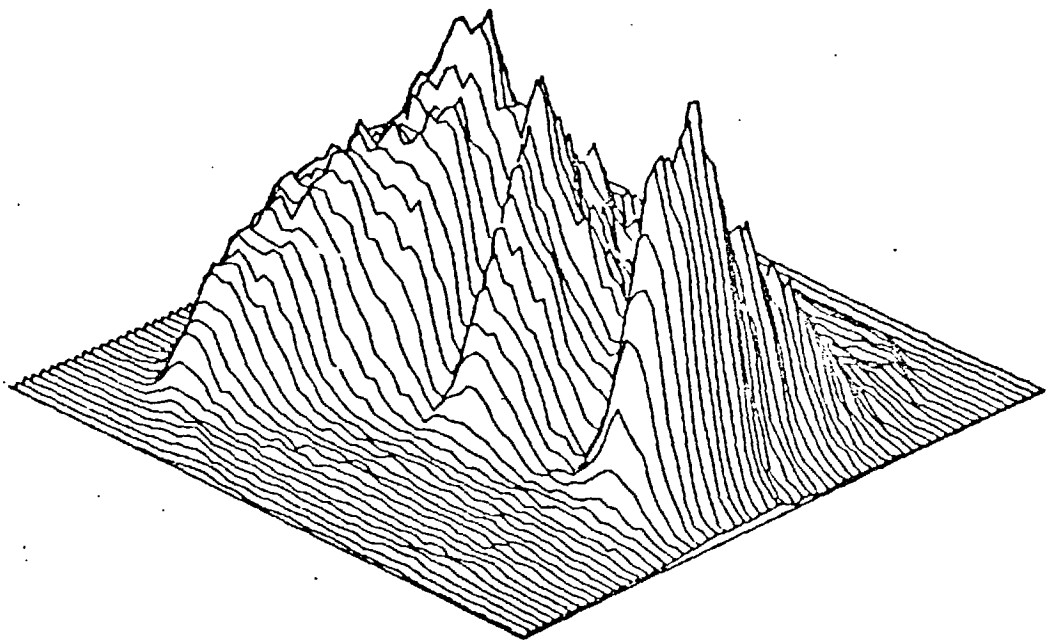


Figure 4.18(a) Image of Fig 4.14 processed by a second order separable product filter.

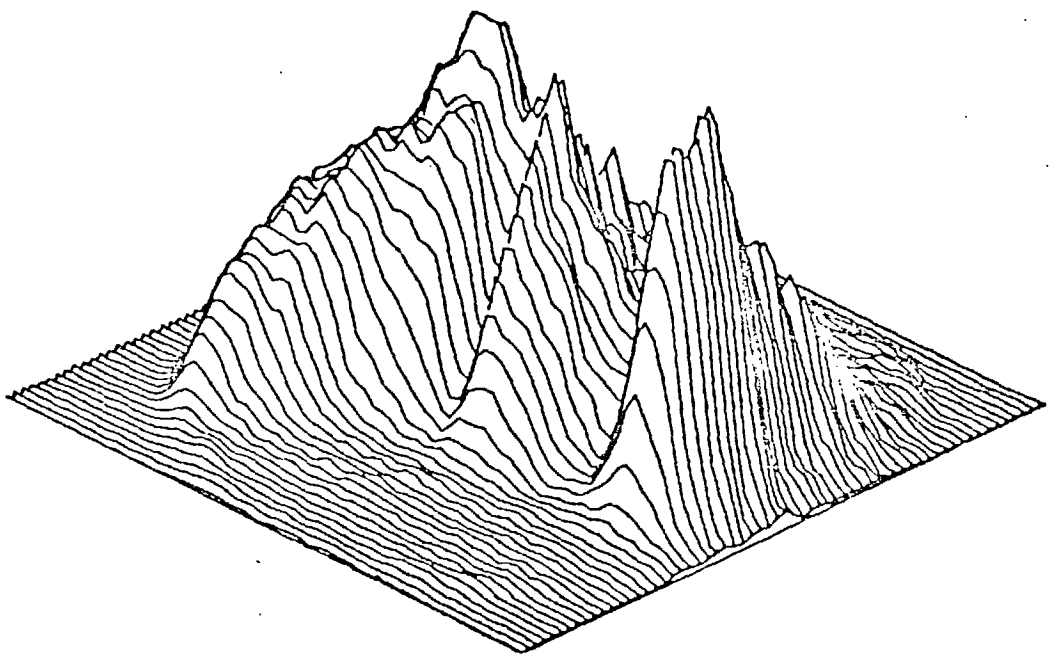


Figure 4.18(b). Image of Fig 4.14 processed by a fourth order separable product filter.

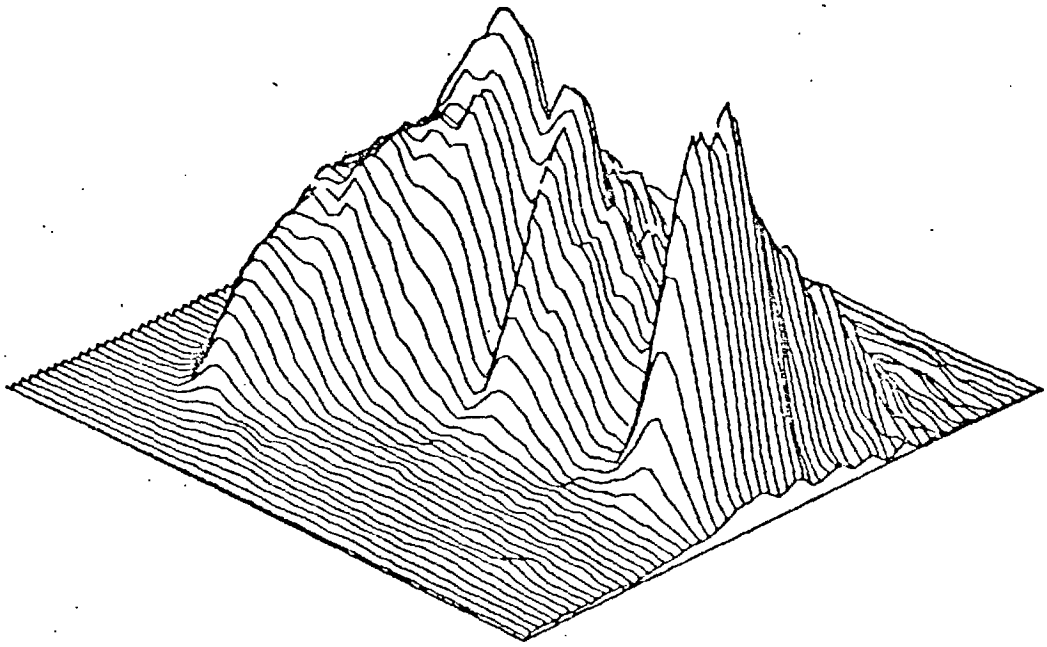


Figure 4.18(c). Image of Fig. 4.14 processed by an eighth order separable product filter.

In Fig. 4.16 the effect of using an Ahmadi filter is seen. This filter has again been designed from the same fourth order Chebyshev prototype using transformation coefficients $a_1 = a_2 = 1$, $b = 0.5$. It is cascaded with a third order Butterworth separable guard filter.

In Fig. 4.17 the response using a filter designed using the modified all-pass transformation of Kap; the prototype is the same Chebyshev filter as above.

Finally in Fig. 4.18 the use of a single separable product filter is investigated. The three plots show the effect of varying the order of the prototype Chebyshev filter. The ripple is maintained at 1% in the passband and the order of the filter progressively increased. In Fig. 4.18(a) a second order prototype is used; in 4.18(b) a fourth order, and in 4.18(c) an eighth order.

Comparison of these responses shows very little difference between the filtered images resulting from the use of any of the filters designed using a fourth order prototype. Further, the increase in complexity from fourth order to eighth order seems hardly justified by subjective comparison of Figs. 4.18(b) and (c), although the increase from second to fourth order is clearly significant.

From the above very limited subjective assessment it would appear that a rectangular cut-off profile is as satisfactory as a circular one and that there appears no justification for resort to the more complicated design techniques for circular profile filters. Such deductions must, of course, be treated with reserve as this may only be fortuitous and a result of the properties of the distorting noise in the original signal.

CHAPTER FIVE

FAN FILTERS

The wise see knowledge and action as one:
They see truly.
Take either path
And tread it to the end:
The end is the same.
There the followers of action
Meet the seekers after knowledge
In equal freedom.

"Bhagavad Gita"
The Yoga of Renunciation

FAN FILTER DESIGN

5.1 INTRODUCTION

Most of the design techniques considered so far have been for filters having cut-off boundaries which are approximately circular or hyperspherical. In this chapter we shall consider a different type of profile, one having a fan or wedge shape. The practical significance of such filters originates in the field of geological survey.

One technique used in geophysical prospecting is to detonate an explosive charge near the surface of the ground and to detect the impulses which have been reflected by interfaces between geological strata and other discontinuities by a set of seismographs situated some distance away. The sequence of echoes received on the array of detectors constitutes a two-dimensional received array in which one dimension is time and the other dimension is linear displacement between the elementary seismographs forming the detection array. It is customary for the detectors to form a linear spatial set, frequently placed vertically in a borehole some distance from the primary detonation. The filtering problem is one in which it is desirable to segregate the echoes into two groups, one which is travelling upwards from low strata and the others which are travelling downwards, probably as the result of echoes from the surface or other higher discontinuities; this would minimize the spurious responses obtained by multiple echoes from several strata discontinuities.

In other situations the detectors may be placed along a horizontal line on the surface of the ground. A similar filtering

problem applies in this case. In more complex situations the array of detectors may be arranged as a two-dimensional array and hence the output from such a set of detectors will form a three-dimensional output array, one dimension of which is time and the other two dimensions being the spatial distribution of the detector elements.

The foregoing indicates that two- and three-dimensional digital filters would be appropriate devices for processing the output data from a set of seismic detectors. The two-dimensional problem has been studied for a considerable time; the earliest solution was given by a simple convolution filter [45,46]. Since then a number of improvements have been made, including the use of recursive filters. Some of these solutions will be discussed later.

5.2 FORMULATION OF DESIGN PROBLEM

A filter is required which will select those signals which are travelling with an apparent velocity whose magnitude lies within a certain bound. This demands a transfer function $Y(w_1,w_2)$, given by [2]

$$Y(w_1,w_2) = \begin{cases} 1, & -\frac{|w_1|}{V} \leqslant w_2 \leqslant \frac{|w_1|}{V} \\ 0, & \text{otherwise} \end{cases} \tag{5.1}$$

where w_1 is the angular frequency of the time varying signal at each detector and w_2 is the spatial angular frequency along the array of detectors.

Such a filter frequency contour is shown in Fig. 5.1.

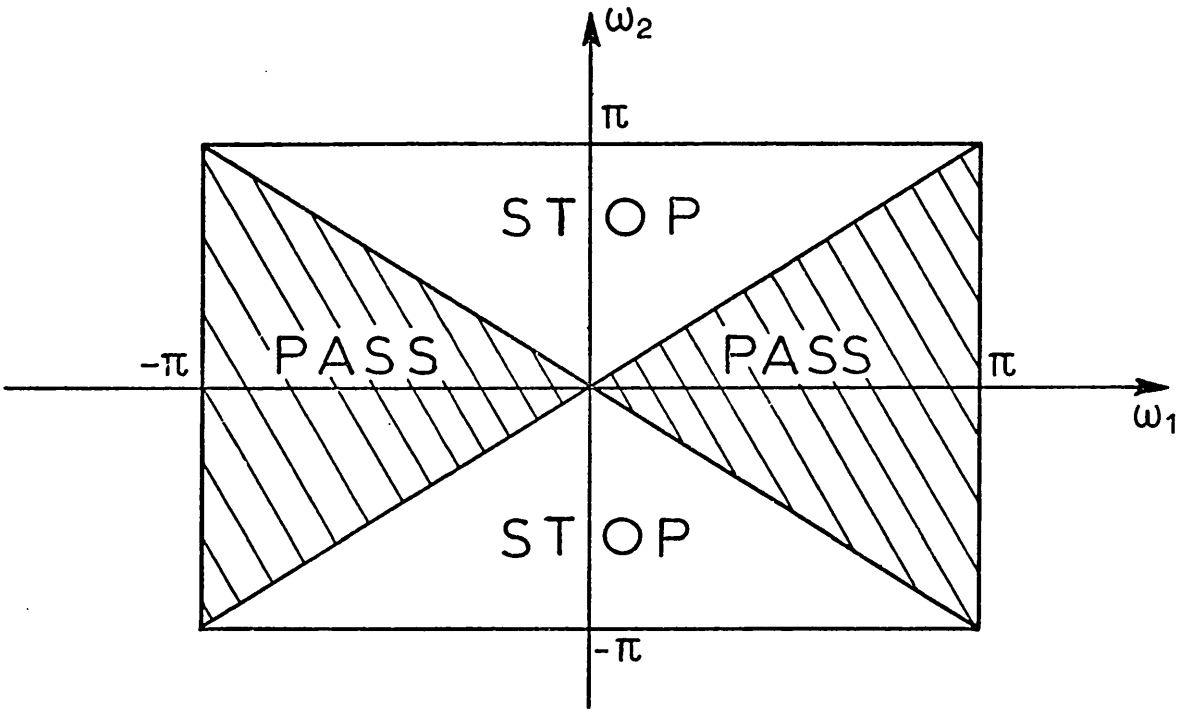


Fig.5·1 Ideal fan filter characteristic.

In practice the temporal frequency response is band limited by the need to eliminate high frequency noise introduced by wind and other extraneous sources. The spatial frequency response is limited by the finite interval between adjacent detectors, This consideration leads to an appreciation of the suitability of digital filters to this signal processing problem.

The earliest attempt at solving this problem was carried out by Embree et al [46] who obtained the inverse Fourier transform of $Y(w_1,w_2)$ defined in equation (5.1) as

$$y(t,x) = \int_{-w_{1N}}^{w_{1N}} \int_{-w_{2N}}^{w_{2N}} Y(w_1,w_2)e^{j(w_1 t - w_2 x)} \, dw_1 \, dw_2 \qquad (5.2)$$

where $\quad w_{2N}V = |w_{1N}|$.

This integral was evaluated directly by Embree who thereby obtained a time-space array which could be convolved with the two-dimensional input sequence to give an output sequence which enhanced the echoes travelling upwards (or, alternatively, downwards depending on the location of pass and stop zones) and eliminated directly-transmitted waves.

Treitel et al [2] improved on the algorithm given by Embree, considerably reducing the computational complexity involved, by taking advantage of certain symmetries of the space-time impulse response array. They also introduced a technique whereby a convolution filter could be similarly designed to give a fan band rejection filter.

Subsequent work by McClellan et al [59] offered a solution to the problem working directly in the frequency domain, transforming

a one-dimensional filter to a two-dimensional one by the well-known McClellan transformation discussed in Section 4.4.4. This technique is shown to permit an approximation to Chebyshev characteristics to be maintained in two dimensions as an attempt at a realization of a minimum ripple approximation in two dimensions.

Subsequent work by Sengbush et al concentrated on the design of optimum velocity filters based on a Wiener optimization process. They designed both band pass and band reject filters by the process by which they were able to reduce the noise in the output data [60,61] .

5.3 FAN FILTER SPECTRAL TRANSFORMATION

The transformation techniques studied in Chapter 4 have, in the main, been constrained by the conditions suggested by Prendergrass [42] in which he considers only those transformations which result in stable filters and also that generate real two-dimensional network functions from real one-dimensional functions.

We have seen that some transformations, for example that of Bernabo et al, do not, per se, result in stable two-dimensional systems but that application of one of the well-documented stabilization techniques may be used to render the system stable without materially affecting its performance.

We will now consider a transformation which represents a complex function of the two-dimensional variables and show that the limitation imposed by Prendergrass on useful transformation functions is by no means mandatory. We shall now apply such a complex

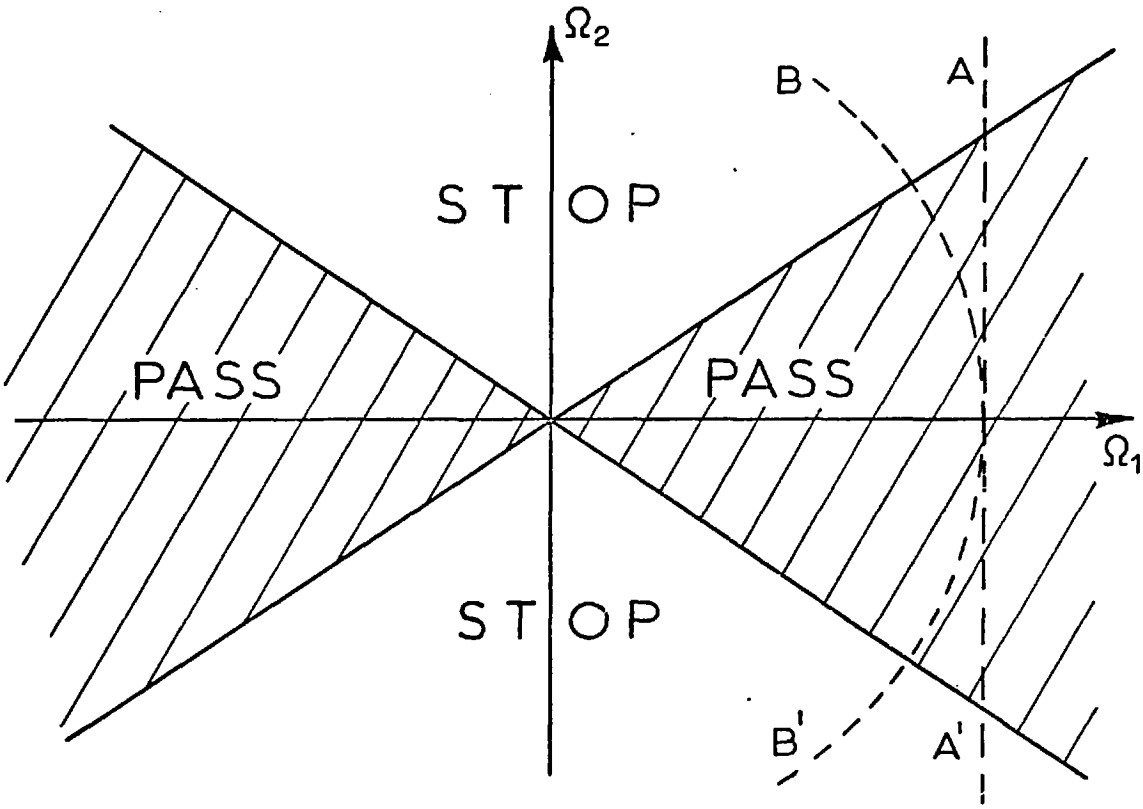


Fig. 5·2 Analogue low-pass filter transformed to analogue fan filter.

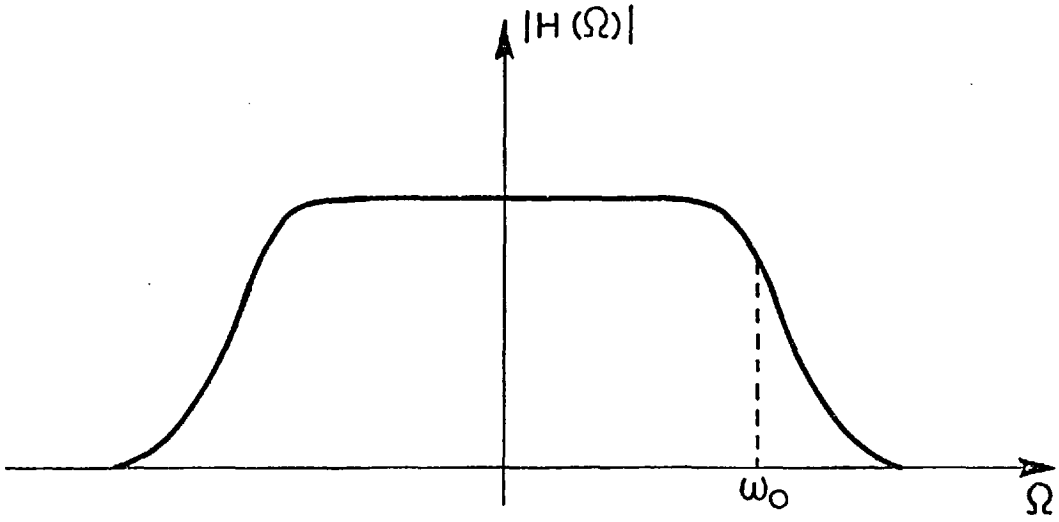


Fig. 5·3 Prototype low-pass analogue filter.

transformation to a one-dimensional low pass filter to generate a low pass fan filter.

5.3.1 Derivation of the fan transformation

An ideal fan filter, ignoring for the moment the band limitations necessary in a practical situation, may be illustrated on the two-dimensional continuous frequency domain by Fig. 5.2, which is a graphical representation of equation (5.1).

This may be achieved by a digital filter which will have the form of Fig. 5.1 and may be obtained from the transfer function represented by Fig. 5.2 by use of the bilinear transformations

$$s_1 = \frac{z_1 - 1}{z_1 + 1}$$

$$s_2 = \frac{z_2 - 1}{z_2 + 1}$$

where s_1, s_2 are the continuous variables and for real frequencies $s_1 = j\Omega_1$, $s_2 = j\Omega_2$ and z_1, z_2 are the discrete variables which for real frequencies are given by $z_1 = e^{jw_1 t}$, $z_2 = e^{jw_2 t}$.

A cross section across the filter parallel to the Ω_2 axis may have a transfer function magnitude which is of any of the classical forms, for example, Butterworth, Chebyshev, elliptic, etc. of the form of Fig. 5.3.

Reference to Fig. 5.4 suggests that we may transform the prototype filter $H(w)$ to $H(\Omega_1, \Omega_2)$ by the relationship

$$\Omega = \Omega_N \text{ arc tan } \Omega_2/\Omega_1 \qquad (5.3)$$

where Ω_N is a normalizing constant.

This will not cause the cross-section AA' (Fig. 5.2) to have the given one-dimensional characteristic but rather the arc of a circle

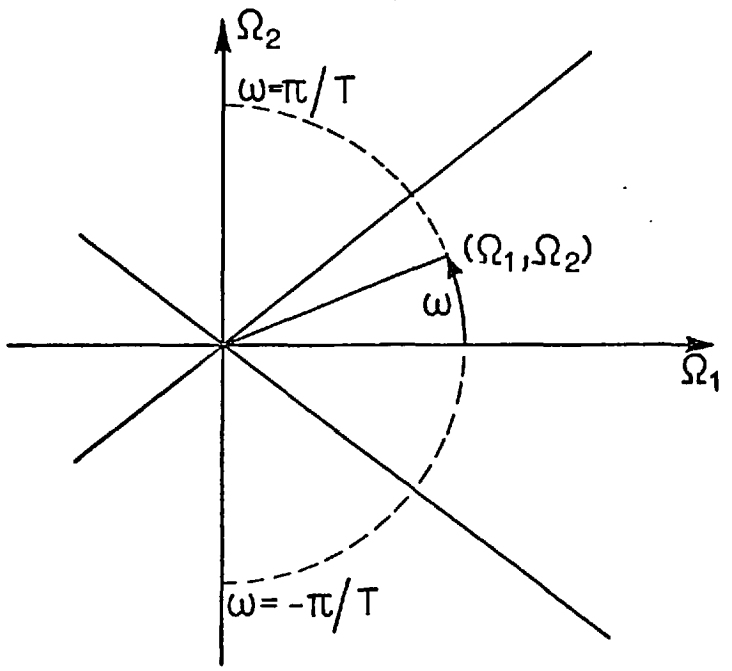


Fig. 5·4 Digital low-pass filter transformation to analogue fan filter.

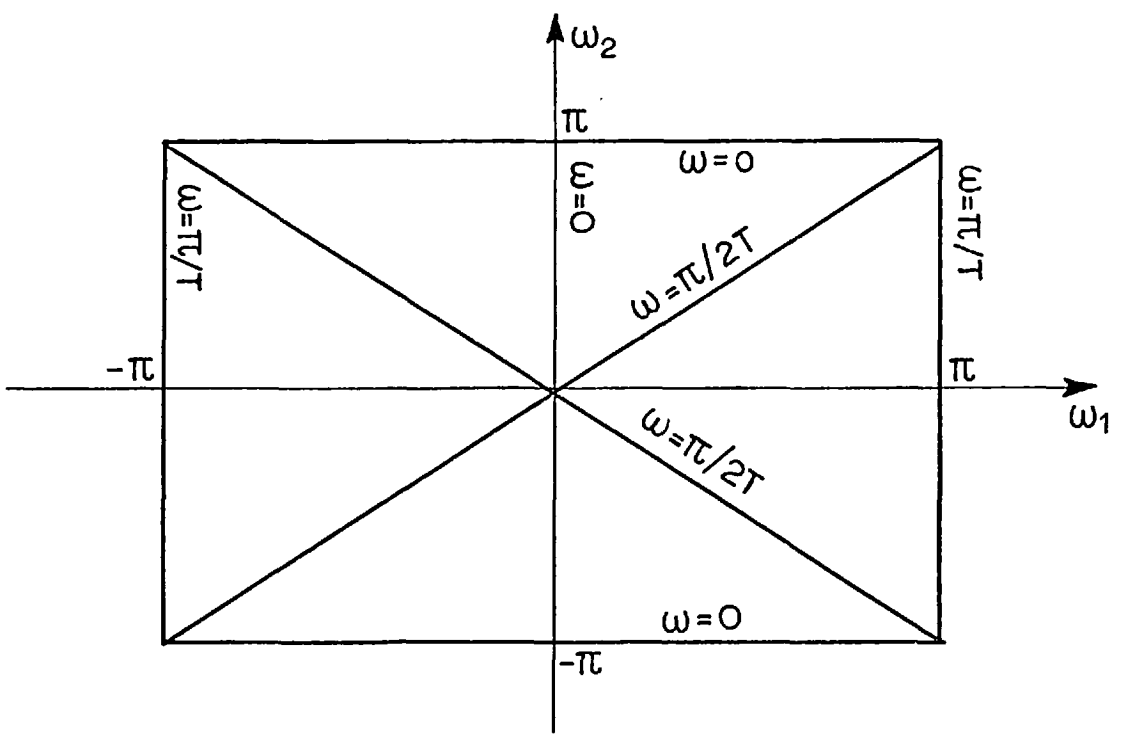


Fig. 5·5 Boundary values of low-pass digital fan filter.

such as BB'. It is further seen that the range of Ω is bounded between $-\pi/2$ and $\pi/2$ for the complete range of transfer function covering both pass and stop bands, which suggests that a transformation from the discrete one-dimensional frequency variable w to the two-dimensional continuous variables Ω_1, Ω_2 would be the more obvious choice.

We shall thus consider the transformation

$$w = w_N \text{ arc tan } \Omega_2/\Omega_1 \tag{5.4}$$

Making the substitution $z = \exp jwT$, and $s_1 = jw_1$, $s_2 = jw_2$, leads to

$$z = \exp\left[j\, w_N\, T \text{ arc tan } s_2/s_1\right] \tag{5.5}$$

Now we may note that the extreme values of w/w_N from (5.4) are $\pm\pi/2$ and thus our prototype digital filter for which wT ranges between $-\pi$ and π imposes the constraint that $w_N = 2/T$.

Further manipulation of (5.5) using the identity

$$\exp(2jn \text{ arc tan } x_1/x_2) = \left(\frac{x_2 + jx_1}{x_2 - jx_1}\right)^n \tag{5.6}$$

leads to the simplified transformation function

$$z = \frac{s_1 + js_2}{s_1 - js_2} \tag{5.7}$$

and in this form of a rational function it is more amenable to algebraic manipulation. It should be noted that this is a complex transformation and will therefore generate a two-dimensional function $F(s_1, s_2)$ which has complex coefficients. We may note that the fan filter which we are attempting to generate is symmetrical about the w_1 axis and hence a transformation

$$z = \frac{s_1 - js_2}{s_1 + js_2} \tag{5.8}$$

applied to the same prototype will produce another complex function $F(s_1, -s_2)$ which will have an identical amplitude characteristic.

Moreover it is obvious that

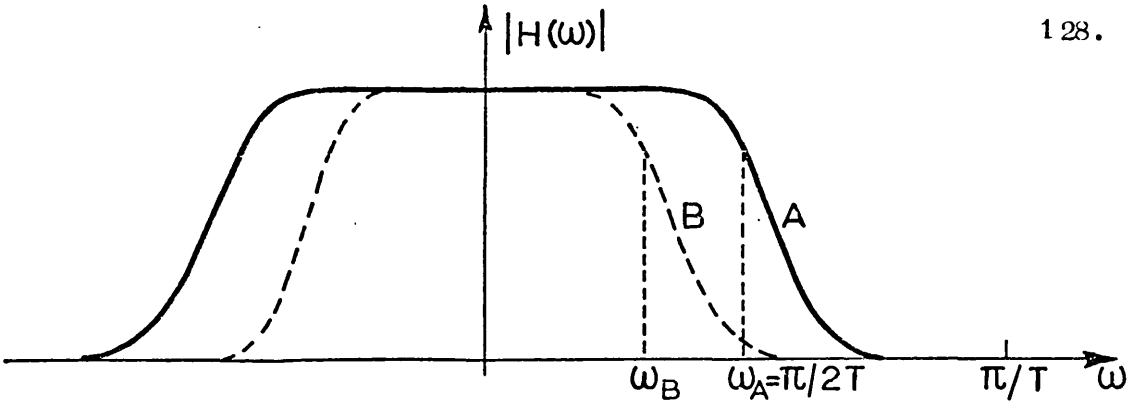
$$F(s_1, -s_2) = F^*(s_1, s_2) \tag{5.9}$$

and hence we may cascade two filters, one using the transformation (5.7) and the other the transformation (5.8). An alternative method would be to use the single transformation (5.7) applied to the product of two prototype functions $\hat{F}(z).\hat{F}(z^{-1})$, resulting, of course, in an identical two-dimensional filter function

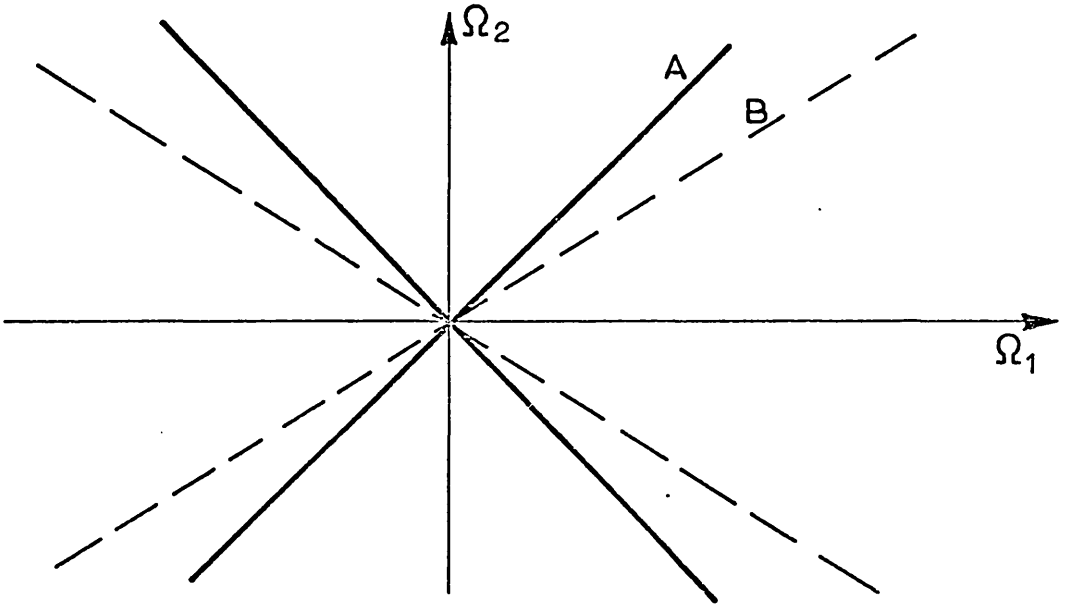
$$H(s_1, s_2) = \hat{F}(z_1).\hat{F}(z_1^{-1})\Big|_{z = \frac{s_1 - js_2}{s_1 + js_2}}$$

The ultimate two-dimensional digital filter may be derived from this analogue filter function by means of the bilinear transformation applied to the two variables s_1 and s_2.

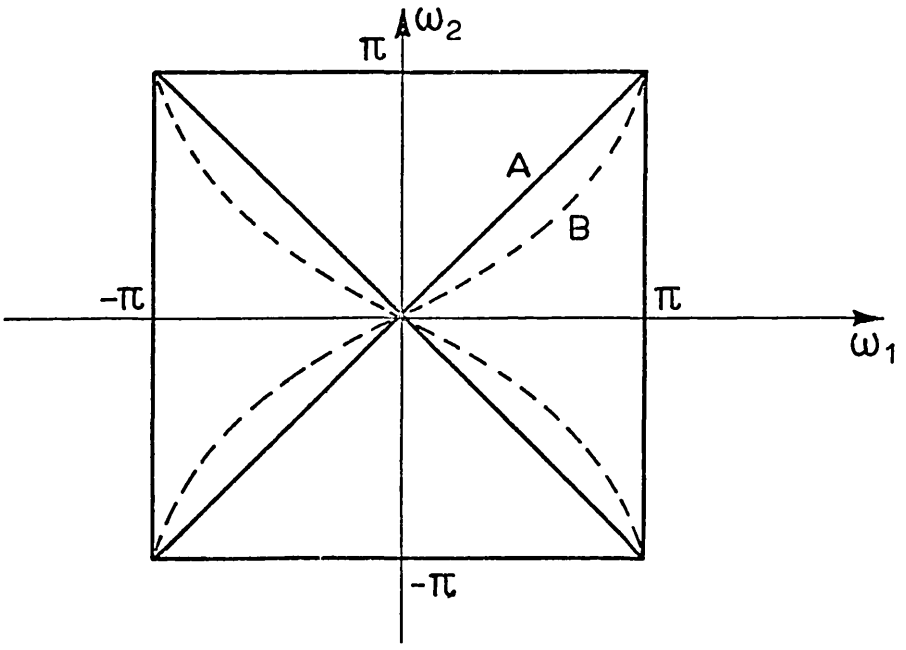
We may study the mapping from the one-dimensional discrete frequency domain w, to the two-dimensional discrete frequency plane w_1, w_2 with reference to Fig. 5.5. The lines $w_1 = \pm\pi$ and $w_2 = 0$ map into the point $w = 0$; the lines $w_1 = 0$ and $w_2 = \pm\pi$ map into $w = \pi/T$; the principal diagonals $w_1 = \pm w_2$ map into $w = \pi/2T$. It may be noted that these boundary transformations create discontinuities at the points $(0,0)$, $(\pm\pi, \pm\pi)$ at which the transfer function will be unspecified. These singular points may be defined as desired, either in the pass or stop band. The transformation is thus symmetrical about both axes, as could also be observed from



(a) Prototype digital low-pass filter.



(b) Cut-off boundaries of continuous fan filter.



(c) Cut-off boundaries of digital fan filter.

Fig.5·6 Cut-off boundaries of fan filter.

equations (5.7) and (5.8); this shows that the resulting function $H(s_1, s_2)$ will be even in both s_1 and s_2.

5.3.2 The Contour Approximation

The basic transformation given by equation (5.7) maps the diagonal $w_1 = w_2$ into the prototype digital filter frequency $\pi/2T$. Let us consider a prototype low pass digital filter having cut-off frequency $k\pi/T$; then the cut-off boundary in the Ω_1, Ω_2 continuous plane will lie along the lines $\Omega_2 = \pm \Omega_1 \tan k\pi/2$ and we apparently have design freedom in choice of the angle of the cut-off boundary (corresponding to velocity in the geophysical problem). However, in transforming to the digital frequency plane we observe that the continuous frequencies $\Omega_1 = \Omega_2 = \infty$ will always transform into $w_1 = w_2 = \pi$, the discrete frequencies. This will result in a distortion of the filter cut-off contour for any prototype filter for which the cut-off boundary does not coincide with half the sampling frequency. This is illustrated in Fig. 5.6. It may be noted that in the vicinity of the origin in the w_1, w_2 plane, the cut-off profile is tangential to the lines $w_2 = \pm w_1 \tan k\pi/2$ and is thus similar to that in the continuous domain.

Since, in the application in which this design technique is most important, namely geophysical prospecting, the variables in the two directions are of differing nature, it is always possible to scale one of the variables, or its sampling rate, so that the prototype may be designed to a cut-off frequency of $\pi/2$. In some other applications it may not be possible to do this and so the need for filters whose cut-off profiles do not lie along the 45°

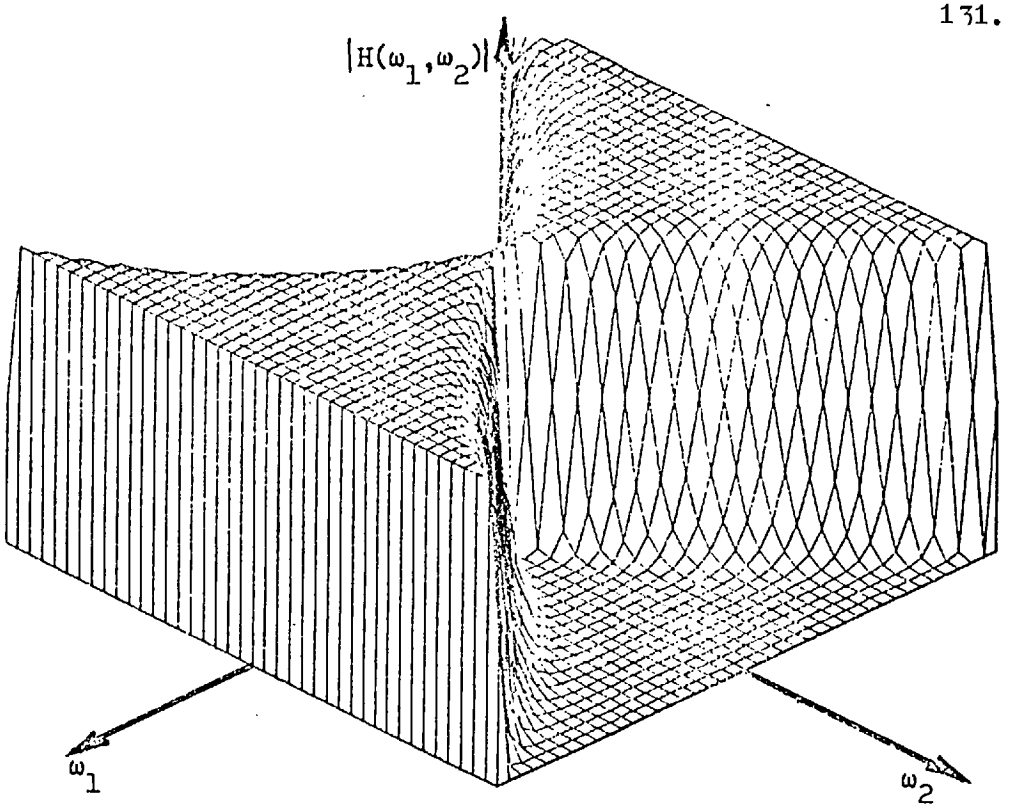
line may arise. In such cases it will only be possible to approximate the required cut-off contour. However, since in all digital filter designs of this nature a guard filter is required to remove high frequency inversions of the basic filter profile, it may be possible to design the guard filter to remove as much of the outer part of the curved profile B (Fig. 5.6c) as is necessary to approximate the desired filter characteristic. Such a guard filter need only have a stop band cut-off boundary along the w_1 axis where it affects the pass band. Along the w_2 axis it is only necessary for the removal of the higher frequency bands and so a lower order guard filter may be adequate.

5.3.3 Stability and Stabilization

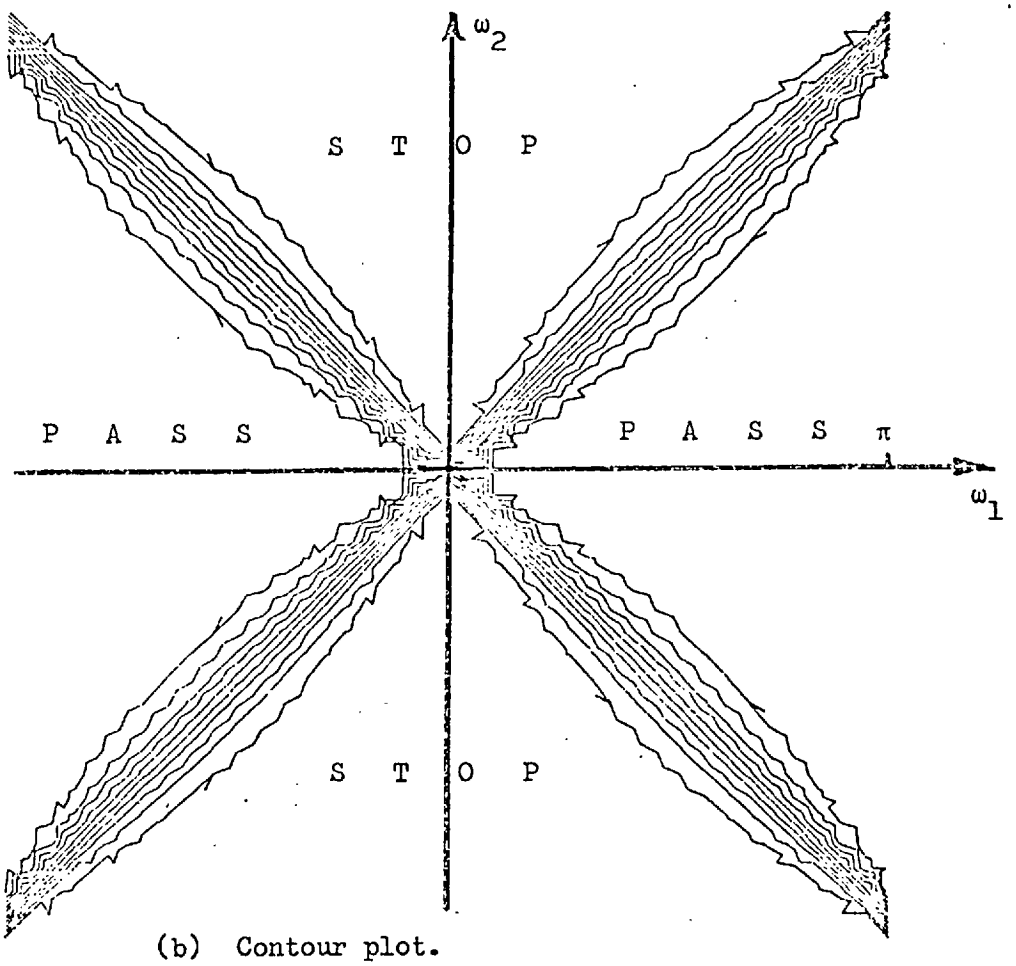
Although the transformation function (5.7) is complex, this does not, of itself, give any indication of the stability of the resulting transformed network. However, it was noted in Section 5.3.1 that the network function $H(z_1,z_2)$ is even in both z_1 and z_2. Moreover, since it is obtained by transforming a function derived from the cascade of $F(z)$ with $F(z^{-1})$ it must be a zero phase function and by definition unstable. As it is a zero phase function we may stabilize it using the technique of Pistor [13] ; thus the final procedure before implementation must be the decomposition of the filter function into four single quadrant functions, each recursing in a different direction.

5.3.4 Design Examples

A number of fan filter designs were implemented to show the significance of the preceding transformation. Where quoted,



(a) Isometric plot.



(b) Contour plot.

Figure 5.7. Low-pass fan filter. Prototype: third order Butterworth low-pass digital filter; $\omega_o = \pi/2$.

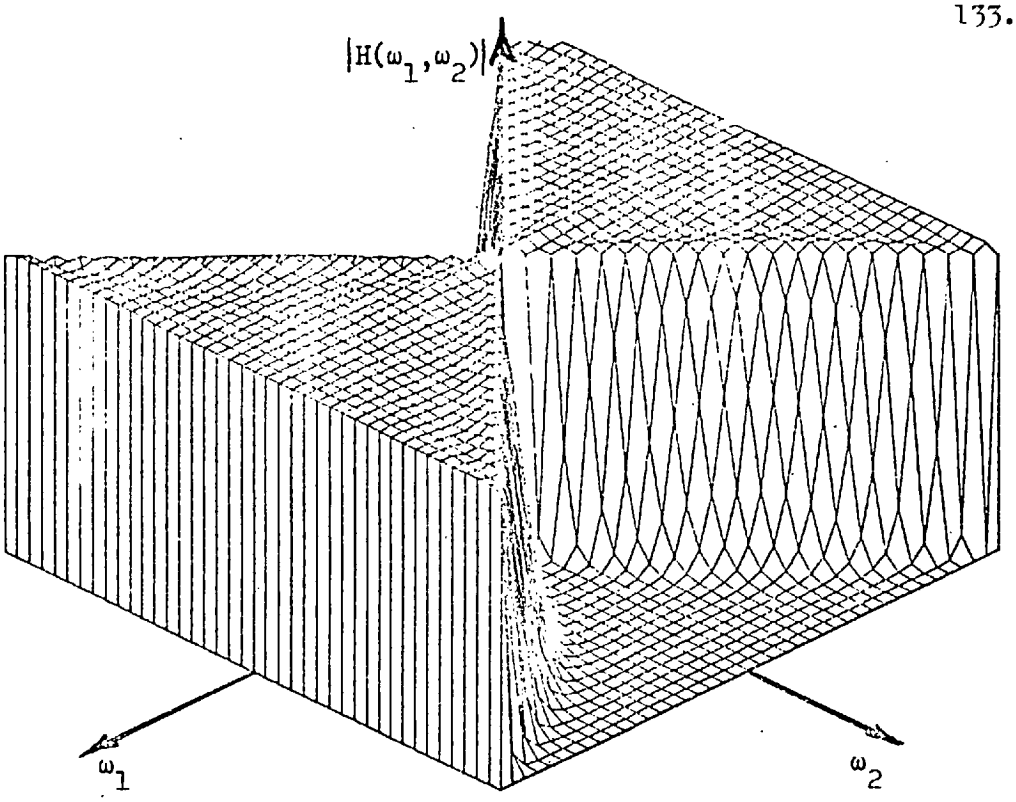
NUMERATOR

1.7776	10.6667	26.6667	35.5556	26.6667	10.6667	1.7776
-10.6667	-64.0000	-160.0000	-213.3333	-160.0000	-64.0000	-10.6667
26.6667	160.0000	400.0000	533.3333	400.0000	160.0000	26.6667
-35.5556	-213.3333	-533.3333	-711.1111	-533.3333	-213.3333	-35.5556
26.6667	160.0000	400.0000	533.3333	400.0000	160.0000	26.6667
-10.6667	-64.0000	-160.0000	-213.3333	-160.0000	-64.0000	-10.6667
1.7776	10.6667	26.6667	35.5556	26.6667	10.6667	1.7776

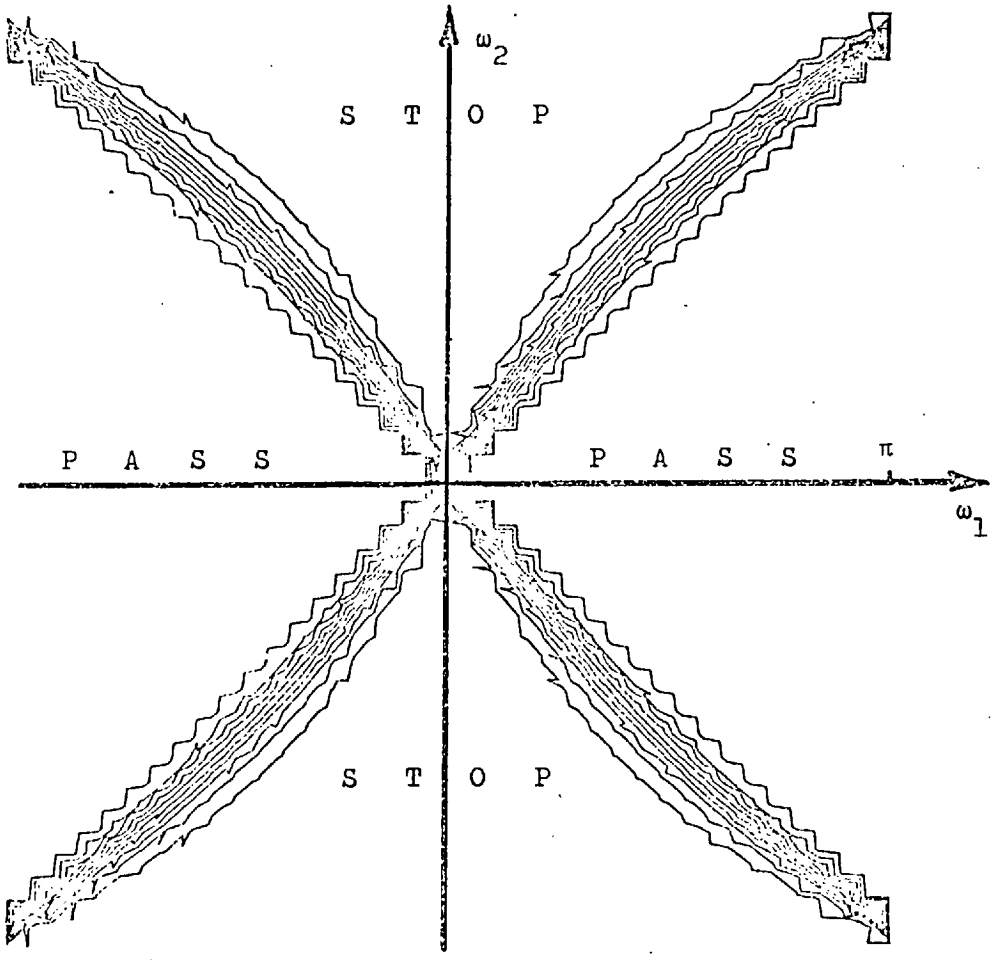
DENOMINATOR

3.5556	0.	53.3333	0.	53.3333	0.	3.5556
0.	-128.0000	0.	-426.6667	0.	-128.0000	0.
53.3333	0.	800.0000	0.	800.0000	0.	53.3333
0.	-426.6667	0.	-1422.2222	0.	-426.6667	0.
53.3333	0.	800.0000	0.	800.0000	0.	53.3333
0.	-128.0000	0.	-426.6667	0.	-128.0000	0.
3.5556	0.	53.3333	0.	53.3333	0.	3.5556

Table 5.1. Coefficients of transfer function of fan filter of Fig. 5.7; prototype, Butterworth, $\omega_o = \pi/2$.



(a) Isometric plot.



(b) Contour plot.

Figure 5.8. Low-pass fan filter. Prototype: third order Chebyshev low-pass digital filter; $\omega_o = \pi/2$, $\delta = 1\%$.

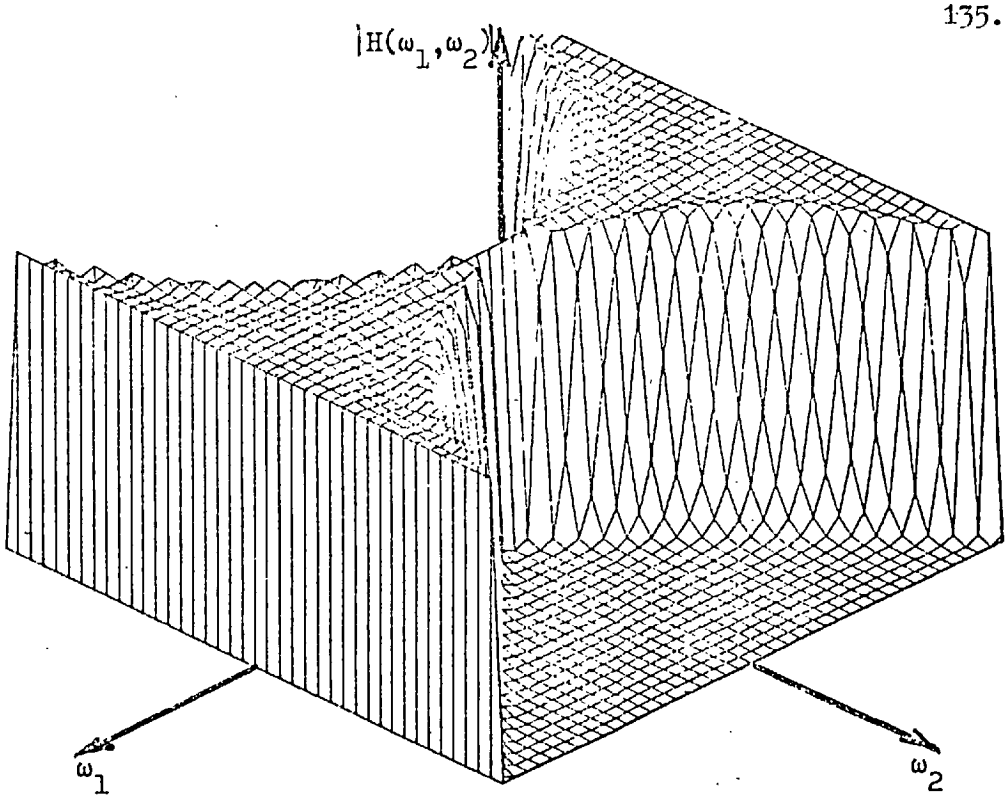
NUMERATOR

3.4917	20.9504	52.3760	69.8347	52.3760	20.5904	3.4917
-20.9504	-125.7025	-314.2562	-419.0083	-314.2562	-125.7025	-20.9504
52.3760	314.2562	785.6405	1047.5207	785.6405	314.2562	52.3760
-69.8347	-419.0083	-1047.5207	-1396.6943	-1047.5207	-419.0083	-69.8347
52.3760	314.2562	785.6405	1047.5207	785.6405	314.2562	52.3760
-20.9504	-125.7025	-314.2562	-419.0083	-314.2562	-125.7025	-20.9504
3.4917	20.9504	52.3760	69.8347	52.3760	20.5904	3.4917

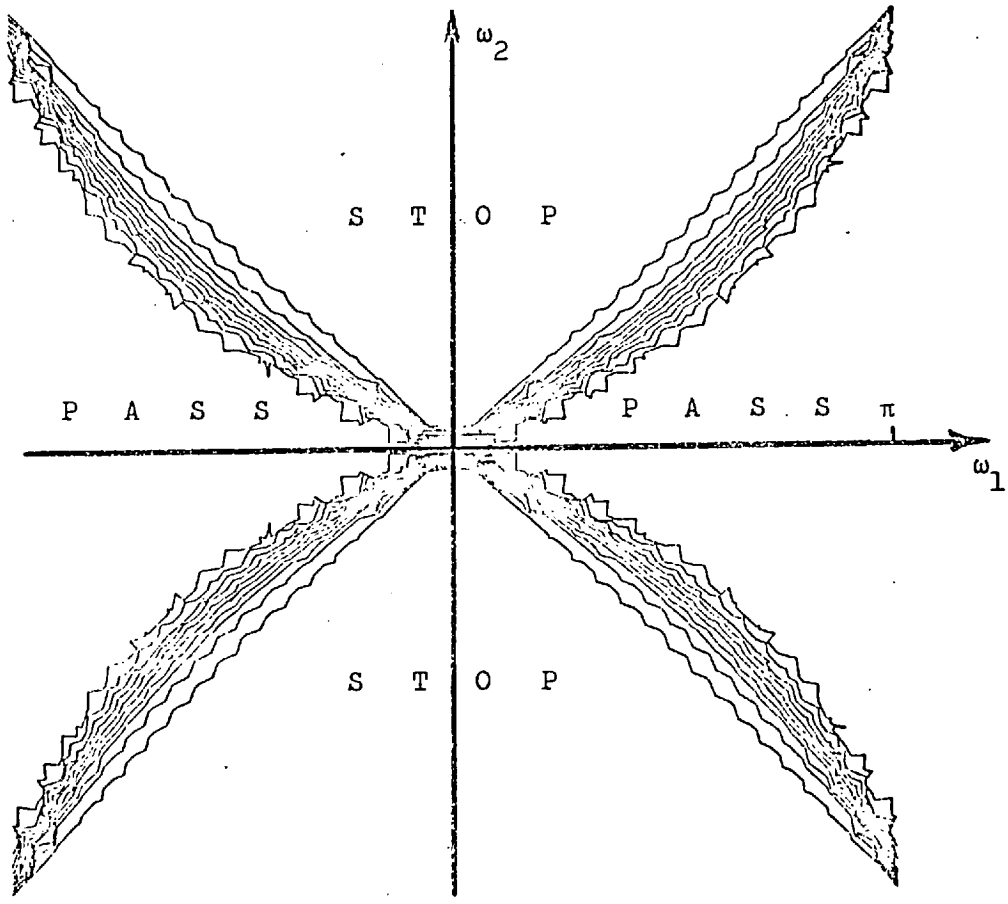
DENOMINATOR

3.5614	18.8185	70.4558	37.7997	70.4558	18.8185	3.5614
-18.8185	-162.2803	-207.4913	-563.6463	-207.4913	-162.2803	-18.8185
70.4558	207.4913	1039.8029	716.5663	1039.8029	207.4913	70.4558
-37.7997	-563.6463	-716.5663	-1833.3970	-716.5663	-563.6463	-37.7997
70.4558	2o7.4913	1039.8029	716.5663	1039.8029	207.4913	70.4558
-18.8185	-162.2803	-207.4913	-563.6463	-207.4913	-162.2803	-18.8185
3.5614	18.8185	70.4558	37.7997	70.4558	18.8185	3.5614

Table 5.2. Coefficients of transfer function of fan filter of Fig. 5.8; prototype: Chebyshev, $\omega_o = \pi/2$.



(a) Isometric plot



(b) Contour plot.

Figure 5.9. Low-pass fan filter. Prototype: third order Chebyshev low-pass digital filter; $\omega_o = 0.207\pi$, $\delta = 1\%$.

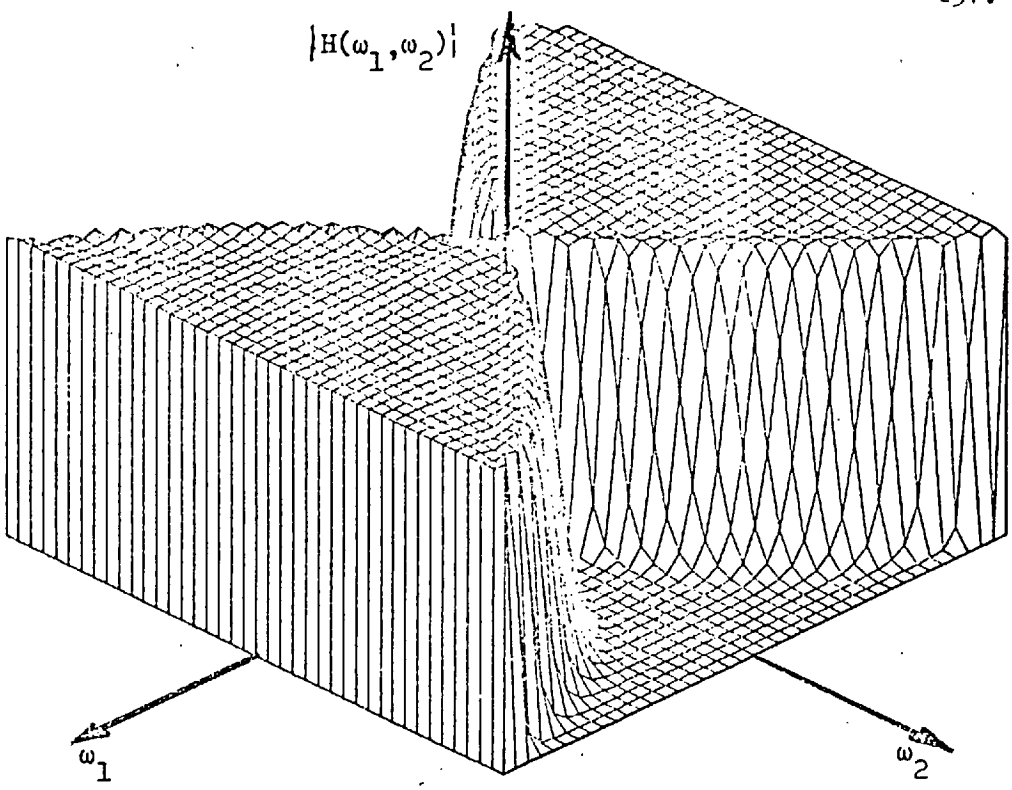
NUMERATOR

0.1690	1.0138	2.5344	3.3793	2.5344	1.0138	0.1690
-1.0138	-6.0827	-15.2067	-20.2756	-15.2067	-6.0827	-1.0138
2.5344	15.2067	38.0167	50.6889	38.0167	15.2067	2.5344
-3.3793	-20.2756	-50.6889	-67.5852	-50.6889	-20.2756	-3.3793
2.5344	15.2067	38.0167	50.6889	38.0167	15.2067	2.5344
-1.0138	-6.0827	-15.2067	-20.2756	-15.2067	-6.0827	-1.0138
0.1690	1.0138	2.5344	3.3793	2.5344	1.0138	0.1690

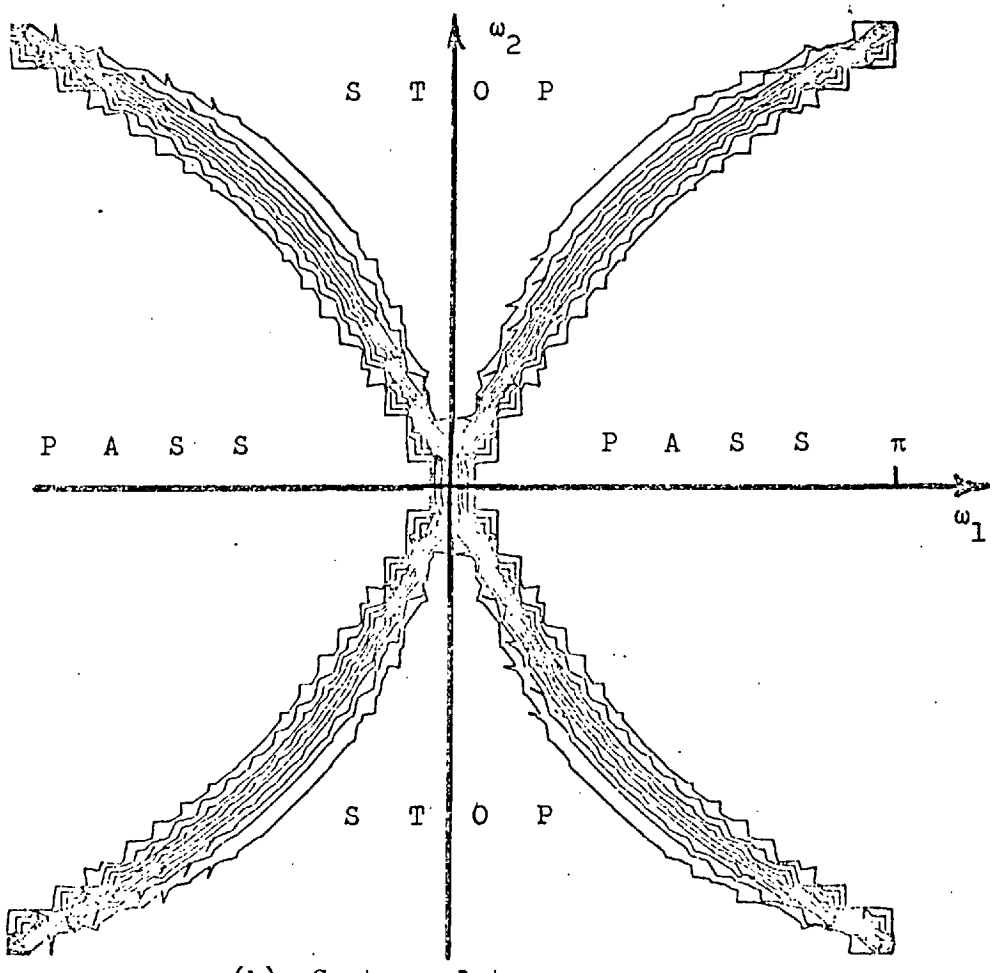
DENOMINATOR

8.4205	-58.2435	168.1705	-225.8815	168.1705	-58.2435	8.4205
58.2435	-386.8641	968.8619	-1345.3643	968.8619	-386.8641	58.2435
168.1705	-968.8619	2480.6951	-3197.8030	2480.6951	-968.8619	168.1705
225.8815	-1345.3643	3197.8030	-4372.9131	3197.8030	-1345.3643	225.8815
168.1705	-968.8619	2480.6951	-3197.8030	2480.6951	-968.8619	168.1705
58.2435	-386.8641	968.8619	-1345.3643	968.8619	-386.8641	58.2435
8.4205	-58.2435	168.1705	-225.8815	168.1705	-58.2435	8.4205

Table 5.3. Coefficients of transfer function of fan filter of Fig. 5.9; prototype; Chebyshev, $\omega_o = 0.207\pi$.



(a) Isometric plot.



(b) Contour plot.

Figure 5.10., Low-pass fan filter. Prototype: third order Chebyshev low-pass digital filter; $\omega_o = 0.75\pi$, $\delta = 1\%$.

NUMERATOR

8.6329	51.7976	129.4939	172.6586	129.4939	51.7976	8.6329
-51.7976	-310.7854	-776.9635	-1035.9513	-776.9635	-310.7854	-51.7976
129.4939	776.9635	1942.4087	2589.8783	1942.4087	776.9635	129.4939
-172.6586	-1035.9513	-2589.8783	-3453.1710	-2589.8783	-1035.9513	-172.6586
129.4939	776.9635	1942.4087	2589.8783	1942.4087	776.9635	129.4939
-51.7976	-310.7854	-776.9635	-1035.9513	-776.9635	-310.7854	-51.7976
8.6329	51.7976	129.4939	172.6586	129.4939	51.7976	8.6329

DENOMINATOR

8.7476	53.3798	133.3187	161.6153	133.3187	53.3798	8.7476
-53.3798	-319.1229	-751.7447	-1066.5494	-751.7447	-319.1229	-53.3798
133.3187	751.7447	1997.6754	2522.1330	1997.6754	751.7447	133.3187
-161.6153	-1066.5494	-2522.1330	-3549.5519	-2522.1330	-1066.5494	-161.6153
133.3187	751.7447	1997.6754	2522.1330	1997.6754	751.7447	133.3187
-53.3798	-319.1229	-751.7447	-1066.5494	-751.7447	-319.1339	-53.3798
8.7476	53.3798	133.3187	161.6153	133.3187	53.3798	8.7476

Table 5.4. Coefficients of transfer function of fan filter of Fig. 5.10; prototype: Chebyshev, $\omega_o = 0.75\pi$.

the transfer functions, $H(z_1,z_2)$, will be represented by numerator and denominator matrices A and B, thus

$$H(z_1, z_2) = \frac{Z_1 \, A \, Z_2^T}{Z_1 \, B \, Z_2^T}$$

where Z_1 is the row vector $[1 \quad z_1^{-1} \quad z_1^{-2} \quad z_1^{-3} \quad z_1^{-4} \quad z_1^{-5} \quad z_1^{-6}]$

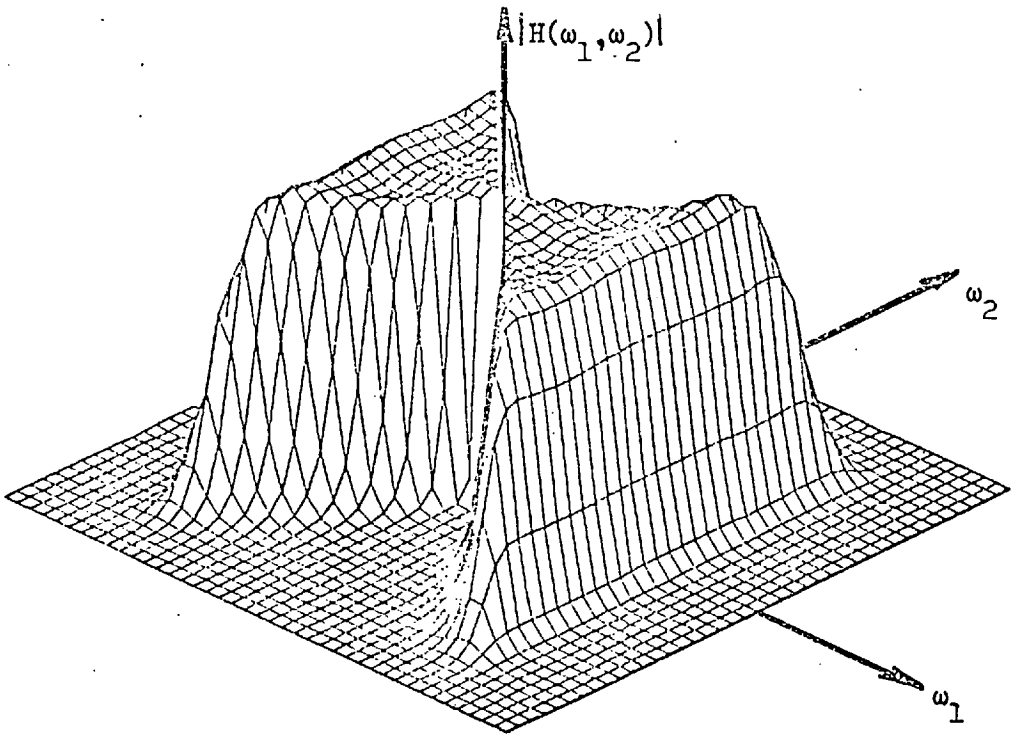
and Z_2 is the row vector $[1 \quad z_2^{-1} \quad z_2^{-2} \quad z_2^{-3} \quad z_2^{-4} \quad z_2^{-5} \quad z_2^{-6}]$

Fig. 5.7 shows the amplitude response and isometric amplitude contours of a fan filter derived from a third order Butterworth filter prototype with discrete cut-off frequency at $w_0 = \pi/2$. The transfer function matrices of this filter are given in Table 5.1.

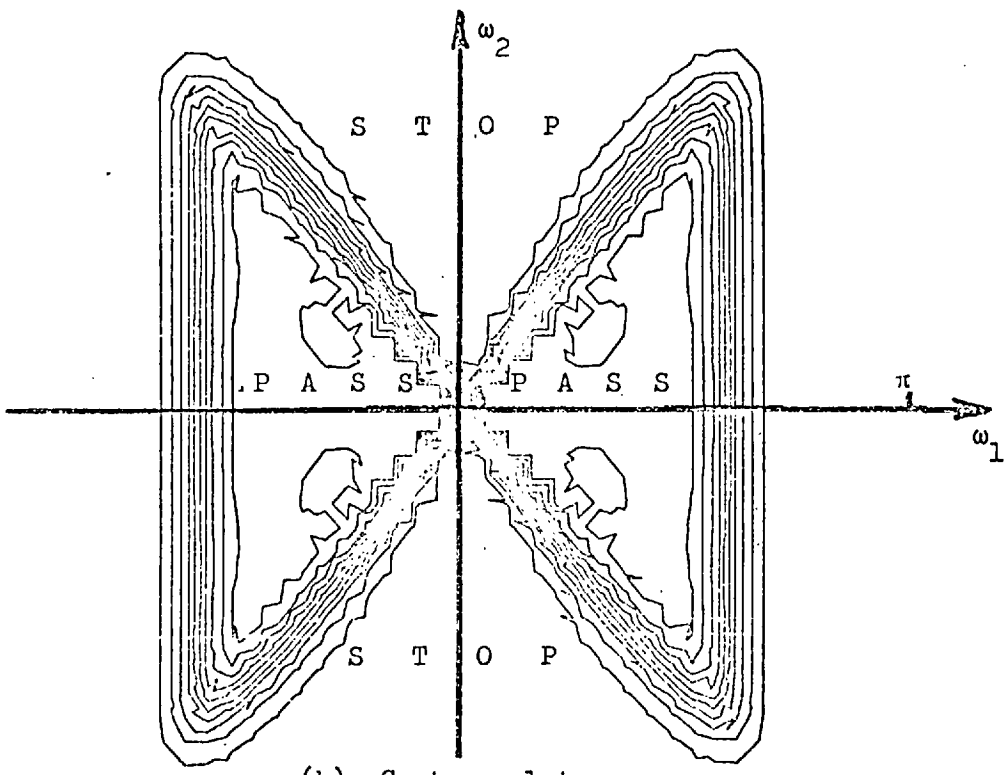
A similar response of a fan filter derived from a third order Chebyshev filter prototype with 1 dB pass band ripple and discrete cut-off frequency at $w_0 = \pi/2$ is shown in Fig. 5.8. The transfer function matrices are given in Table 5.2.

Figs. 5.9 and 5.10 show the predicted effect of variation of the cut-off frequency from the ideal value of $w = \pi/2$. Both are designed to the Chebyshev specification with 1 dB pass band ripple. Fig. 5.9 has a prototype with $w_0 = 0.207\pi$. Fig. 5.10 is based on a prototype with $w_0 = 0.75\pi$. The transfer function matrices are given in Tables 5.3 and 5.4.

One may observe from these that, as expected, the pass band edge lies almost exactly along the diagonals for $w_0 = 0.5\pi$; the Chebyshev filter shows about 2% ripple in the pass band as a result of cascading two similar filters and a considerably steeper transition band.

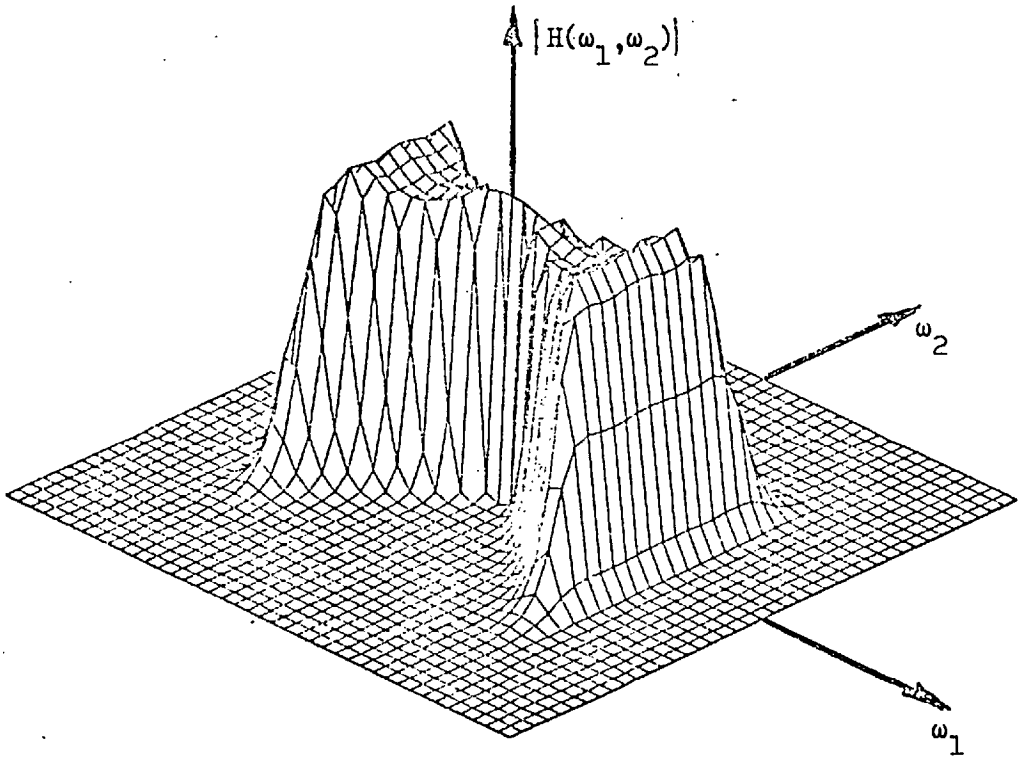


(a) Isometric plot.

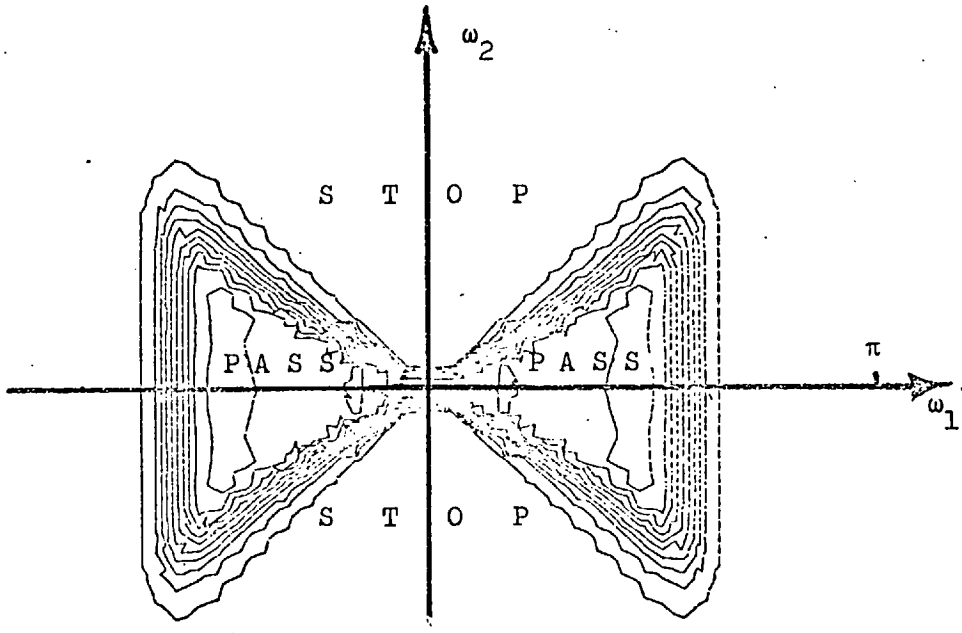


(b) Contour plot.

Figure 5.11. Low-pass fan filter. As in Fig. 5.8 cascaded with a sixth order separable guard filter.



(a) Isometric plot.

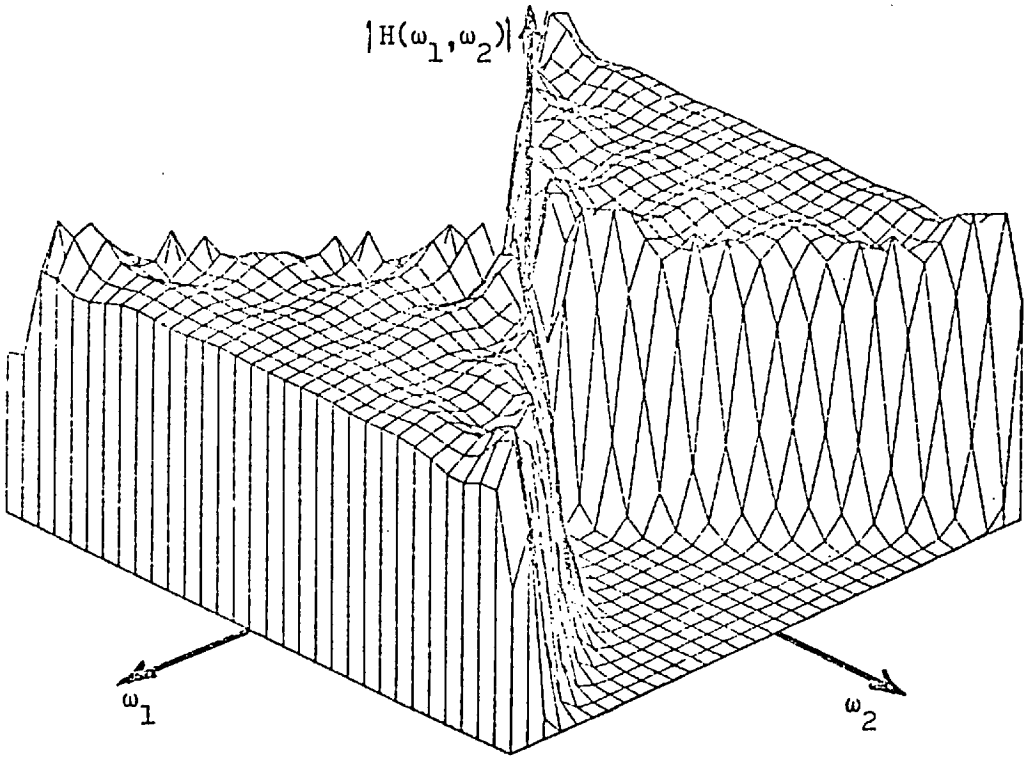


(b) Contour plot.

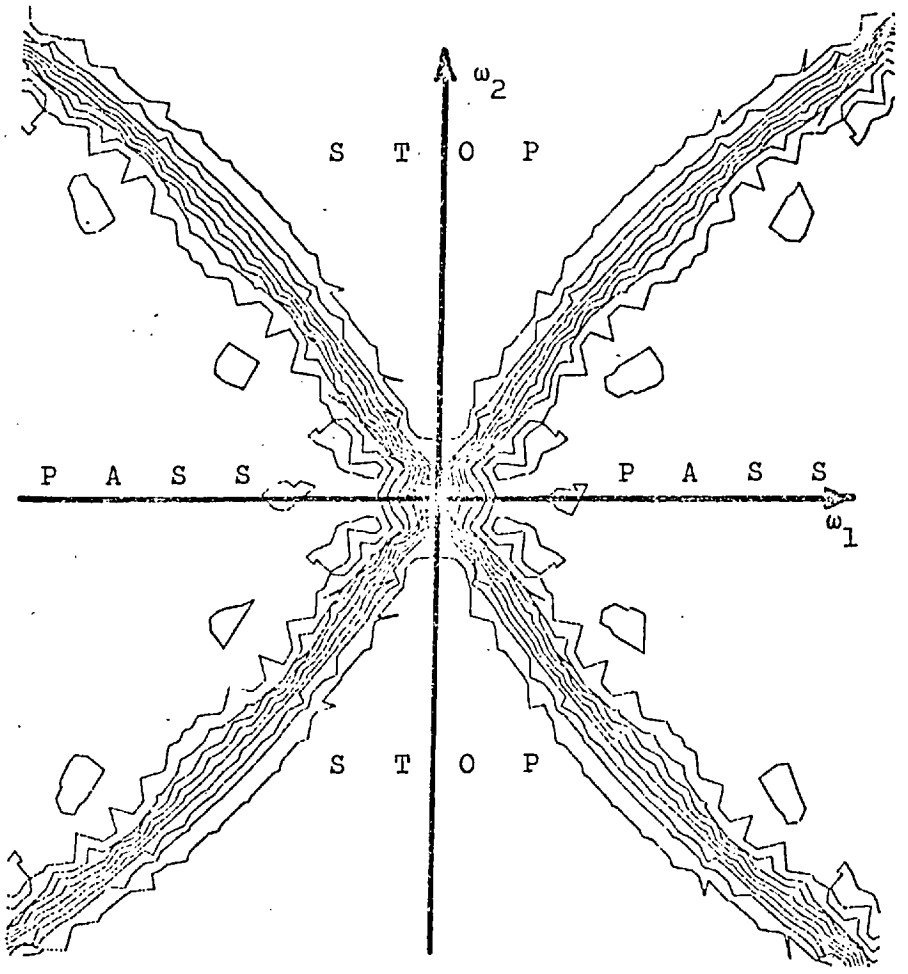
Figure 5.12. Low-pass fan filter. As in Fig. 5.9 with guard filter having $\omega_o = 0.207\pi$.

4.250	0.5157	-0.5435	-0.02542	-0.08840	-0.003895	-0.02907
-0.5157	-5.195	-0.1713	0.1602	-0.01168	0.02557	-0.003535
-0.5435	0.1713	1.495	-0.001834	0.1447	-0.0006941	0.04754
0.02542	0.1602	-0.001834	0.01514	0.002249	0.008339	0.0006964
-0.08840	0.01168	0.1447	-0.002249	0.01006	0.00004112	0.003474
0.003895	0.02557	0.0006941	0.008339	-0.00004112	0.002285	0.00006407
-0.02907	0.003535	0.04754	-0.0006964	0.003474	-0.00006407	0.001121

Table 5.5. Coefficients of denominator of the first quadrant function after decomposition of the denominator of Table 5.2 by the Pistor technique.



(a) Isometric plot.



(b) Contour plot.

Figure 5.13. Low-pass filter. As in Fig. 5.8, after stabilization using the Pistor technique.

The effect of deviations of prototype cut-off frequency from $w_0 = \pi/2$ are more pronounced with Chebyshev prototypes than with Butterworth type filters. It may be observed that a fan filter having other than a $\pm \pi/4$ angle would be realized by cascading one such filter with a high order guard filter. This is shown in Fig. 5.11 for a fan filter derived from a Chebyshev third order prototype cascaded with a sixth order Chebyshev guard filter.

The effect of a guard filter on a fan filter designed from a prototype with cut-off frequency less than $\pi/2$ is shown in Fig. 5.12, from which it may be seen that a close approximation to linear fan filter specification is obtained over a limited band of frequencies.

As noted in Section 5.3.3 any designed filter requires stabilization. This may be achieved using the Pistor technique. This procedure has been applied to the transfer function given in Table 5.2 using a 32 point FFT algorithm; the final decomposed denominator transfer function was truncated to a 7 x 7 array.

The first quadrant decomposed denominator array of the transfer function is given in Table 5.5. The frequency response of the transfer function obtained by cascading four single quadrant filters and the original numerator function is shown in Fig. 5.13. It is seen from this that considerable ripples are introduced in the reconstituted response; these may be caused partly by the truncation of the infinite array which forms the denominator polynomial, but also by the limitations of a fast Fourier transform based on a finite number of discrete frequencies. An increase in the number of elements used in the FFT algorithm has been shown

to reduce the ripple. An extension of the size of the denominator array will also usually improve the approximation, but this involves an increase in complexity of the transfer function realization.

5.4 HIGH PASS AND MULTIPLE PASS BAND FILTERS

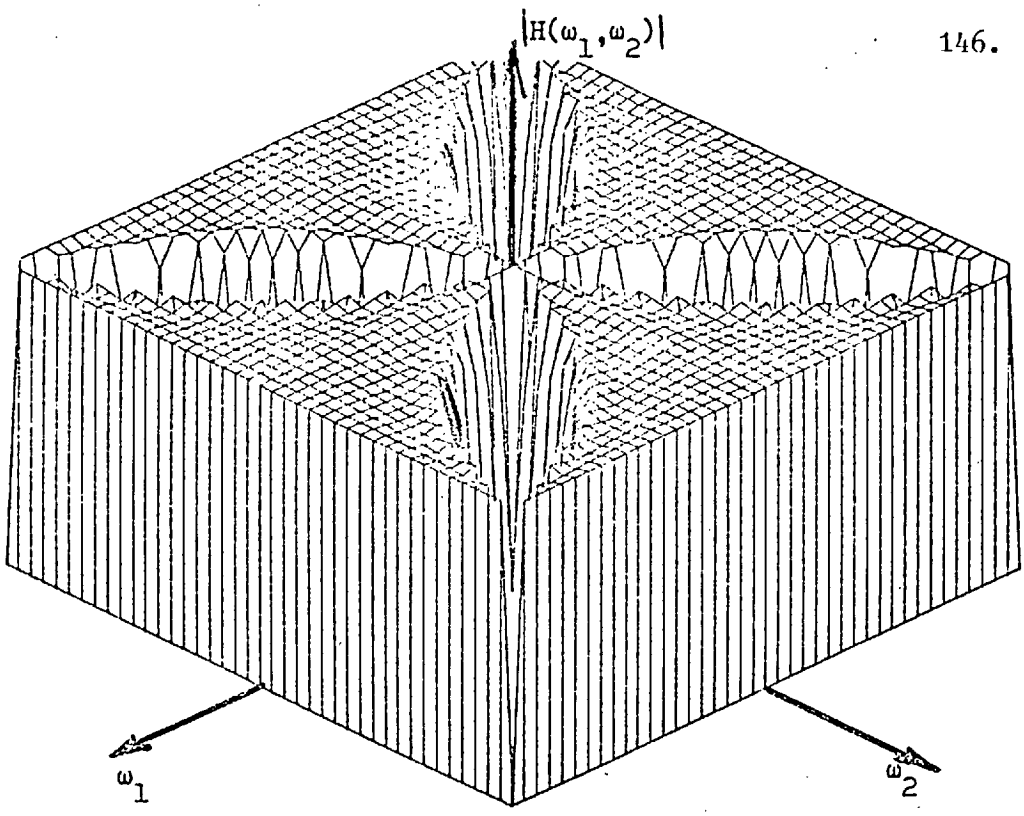
The design of high pass fan filters by the proposed technique is almost trivial as it may be effected by direct interchange of the two variables z_1 and z_2. This is, of course, identical to designing a high pass prototype from the low pass prototype by the classical substitution $z = -z'$.

The design of band pass and band stop filters may be performed in a similar manner by conversion of the prototype filter into the desired form by one of the standard one dimension substitutions.

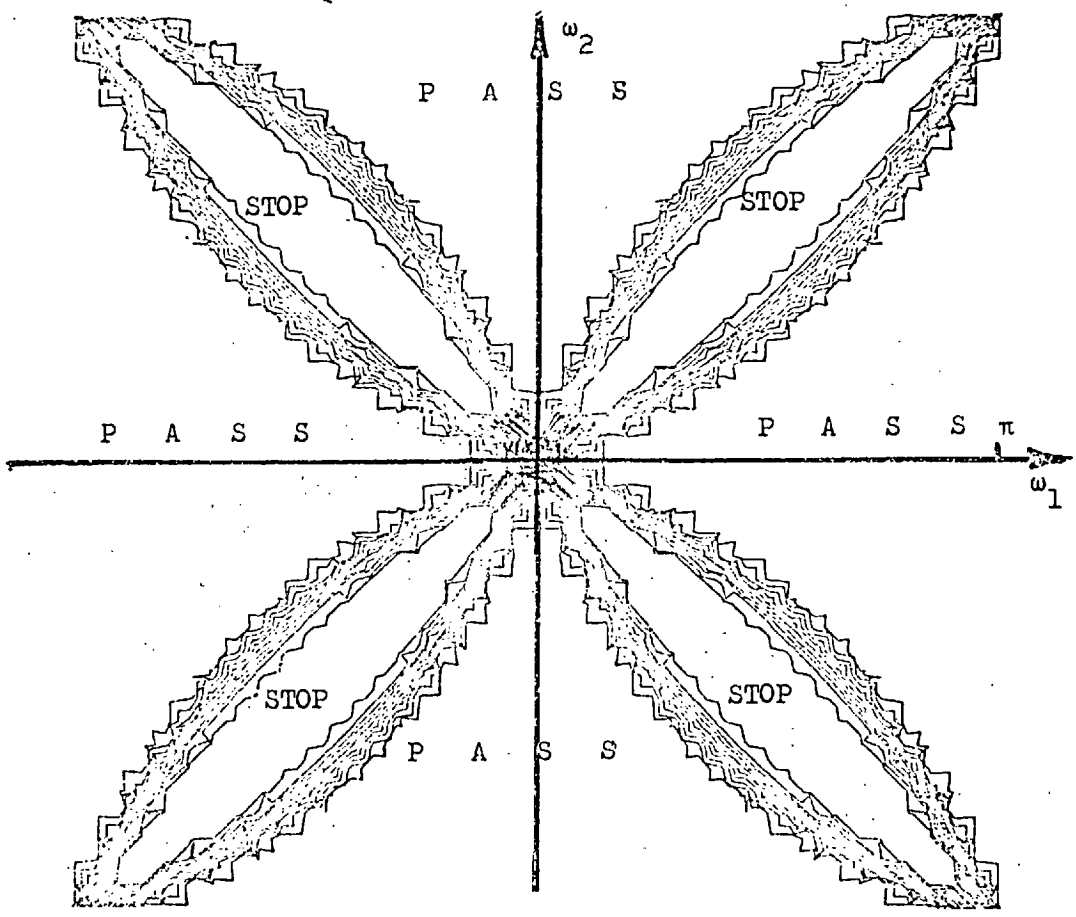
An alternative approach to this design problem may be obtained by a direct transformation from a one-dimensional low pass filter. If we wish to obtain a band stop filter we must map $w = 0$ of the one-dimensional low pass prototype filter response into the $w_2 = 0$ the $w_1 = 0$ axes of the two-dimensional filter response and $w = \pi$ into the diagonal line $w_1 = w_2$, for example. The cut off frequency of the one-dimensional filter at, say, $k\pi/2$, will no longer map into the principal diagonal in the two-dimensional filter.

The form of transformation required to achieve this is obtained by reference to equation (5.4) which may be modified to satisfy the above boundary constraints to give

$$w = 2w_N \arctan \Omega_2/\Omega_1 \qquad (5.10)$$

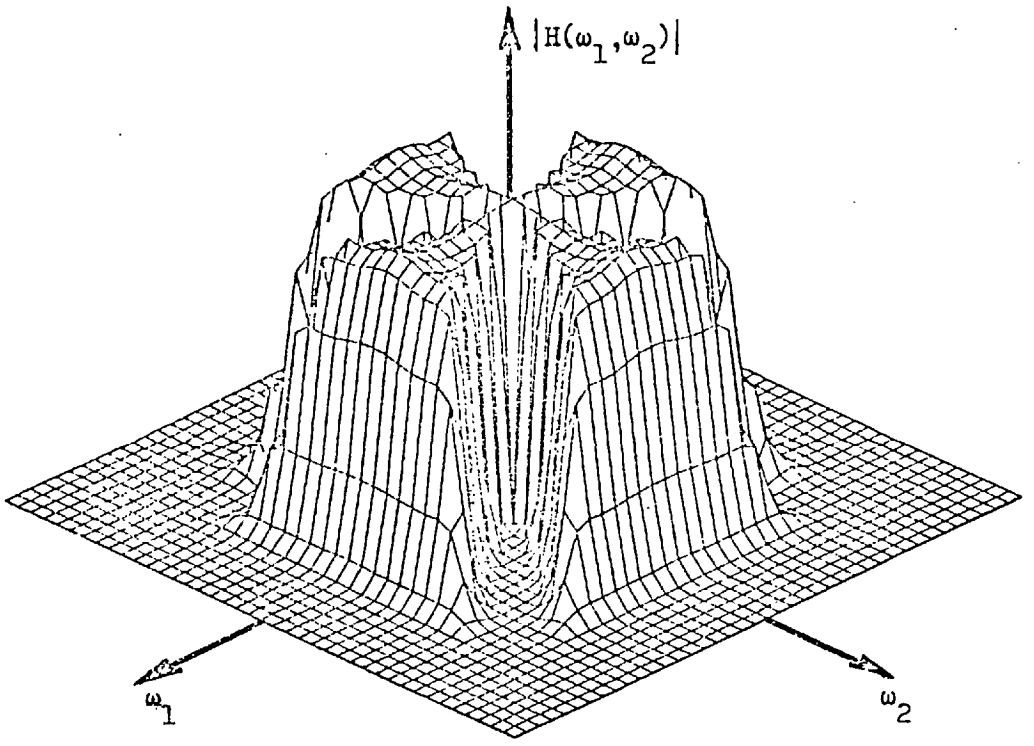


(a) Isometric plot.

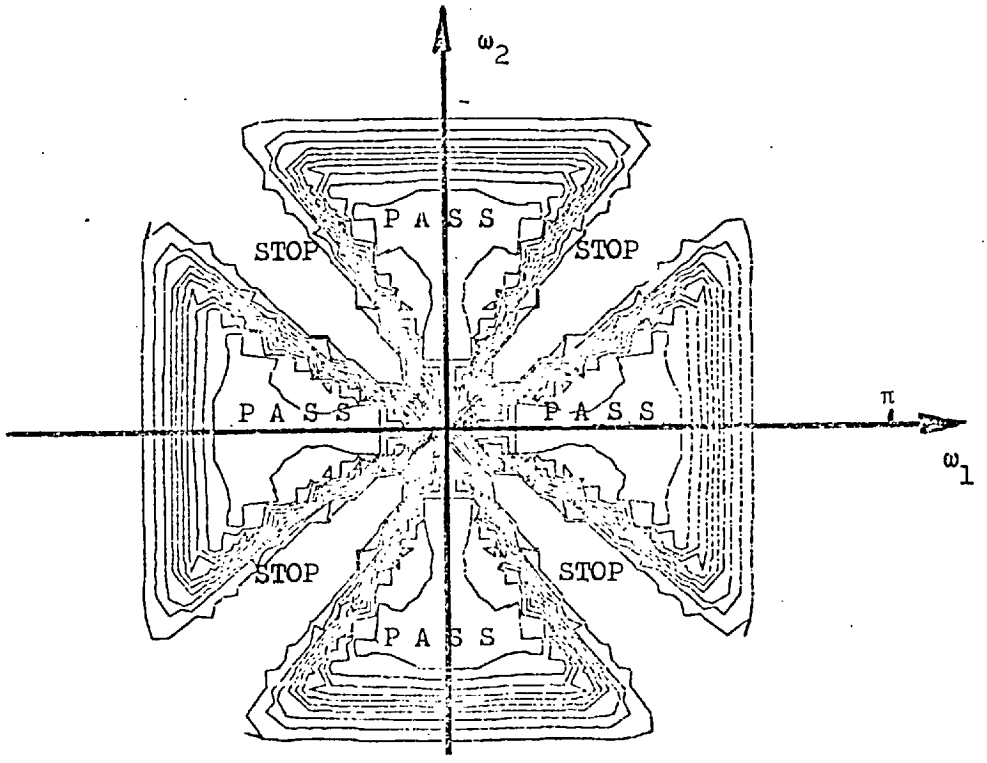


(b) Contour plot.

Figure 5.14. Band-stop filter. Prototype: third order Chebyshev low-pass digital filter; $\omega_o = \pi/2$, $\delta = 1\%$.

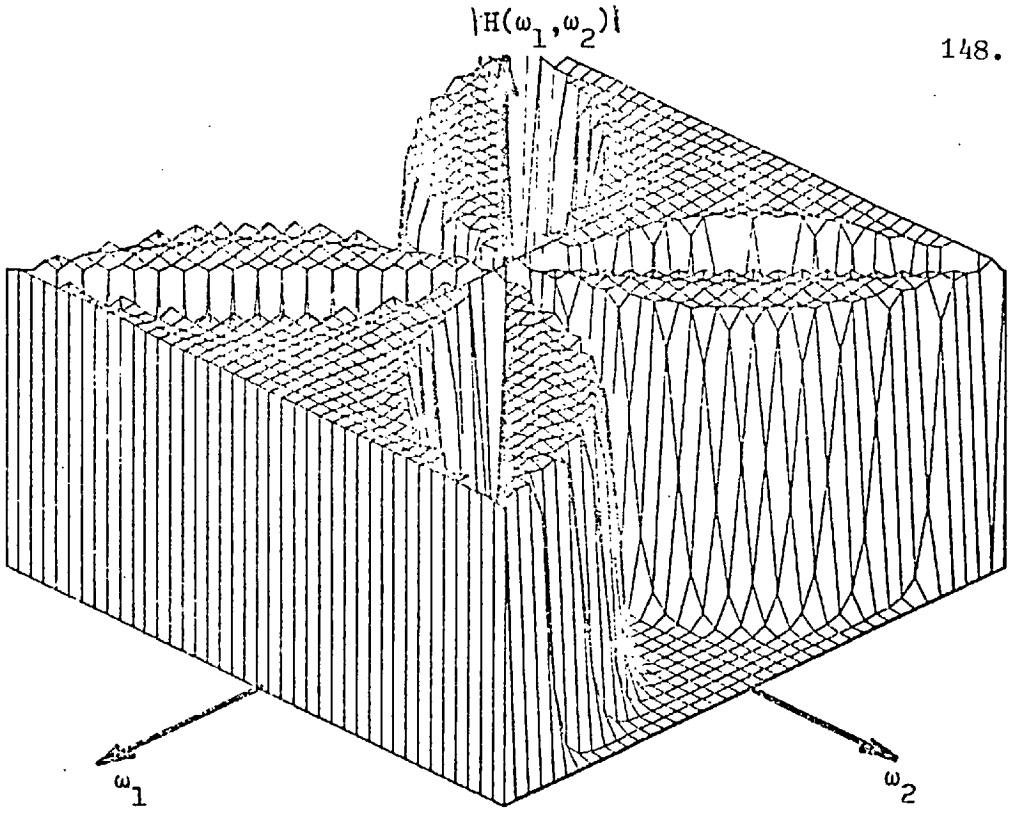


(a) Isometric plot.

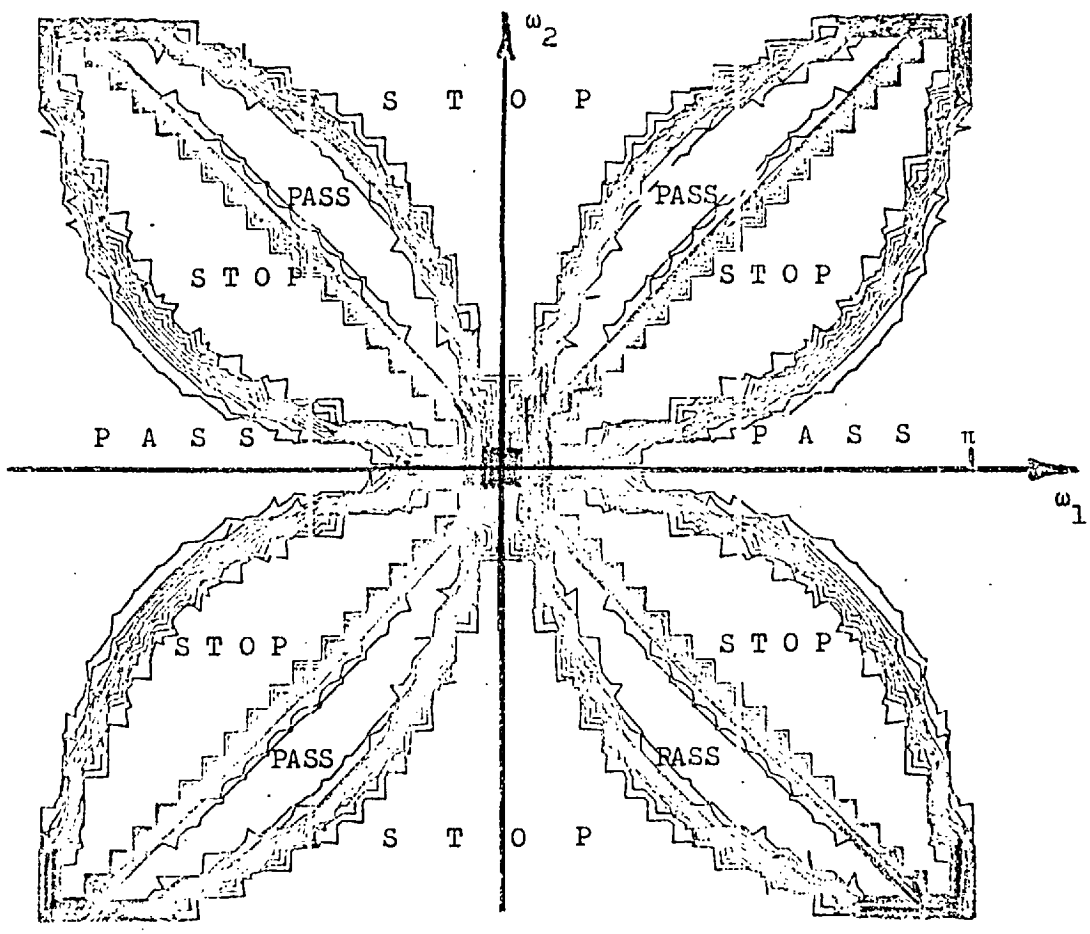


(b) Contour plot.

Figure 5.15. Band-stop filter. As in Fig. 5.13 cascaded with a sixth order guard filter.



(a) Isometric plot.



(b) Contour plot.

Figure 5.16. Multiple-band fan filter. Prototype: third order Chebyshev low-pass digital filter; $\omega_o = \pi/2$, $\delta = 1\%$.

Using identity (5.6) this leads to

$$z = \left(\frac{s_1 + js_2}{s_1 - js_2}\right)^2 \tag{5.11}$$

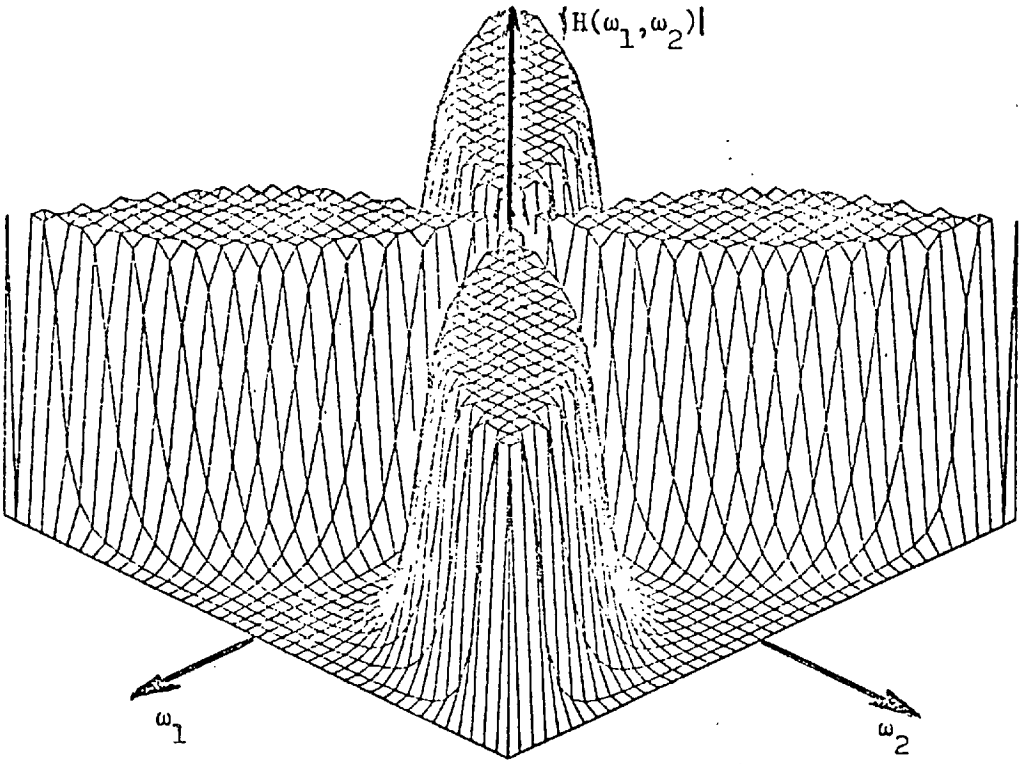
If this is used as a transformation applied to a one-dimensional low pass prototype it will give a double fan filter with pass bands along both axes and stop band along the region of the principal diagonals. This is illustrated by the frequency response plot shown in Fig. 5.14. A realization approximating more closely to a band stop filter may be obtained by cascading the previous filter with a guard filter as shown in Fig. 5.15; the ripples in the pass band are, of course, considerably accentuated as a result of the superposition of a number of Chebyshev-type responses in this region.

Multiple pass band fan filters may be obtained using higher degree transformation functions of the form

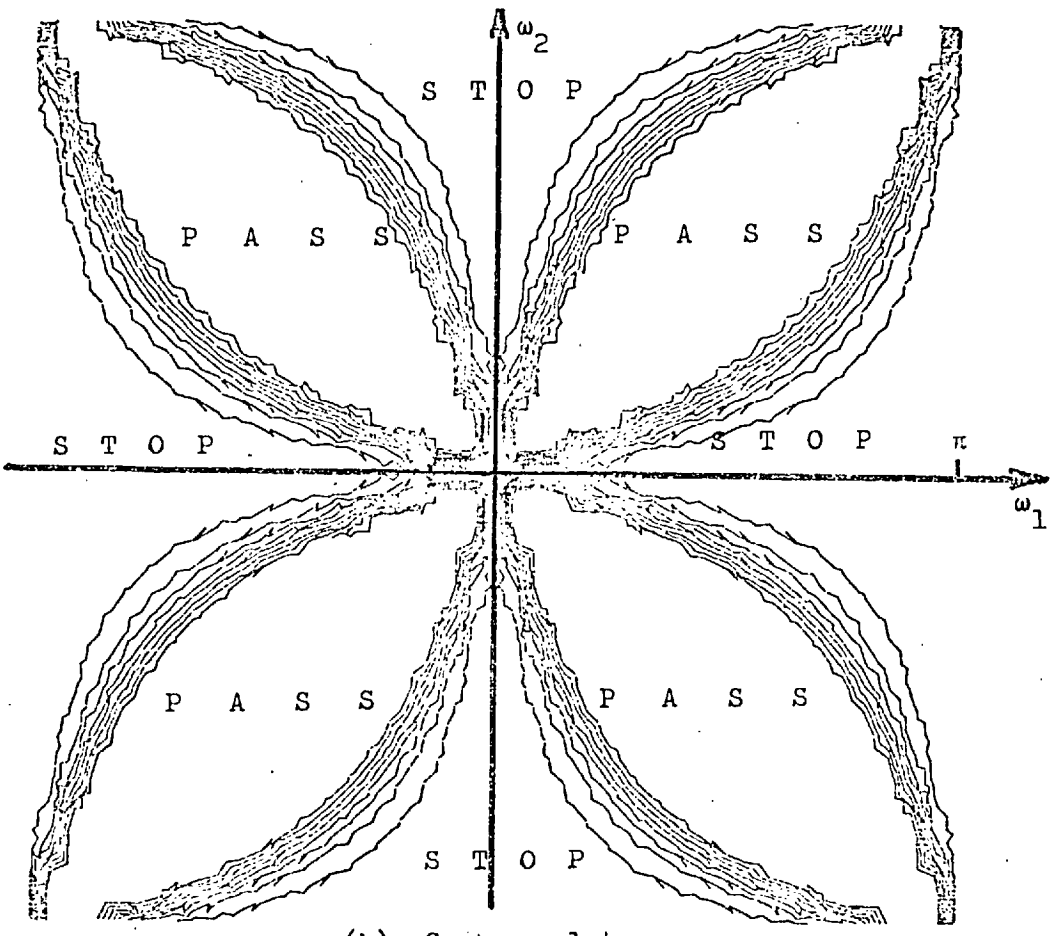
$$z = \left(\frac{s_1 + js_2}{s_1 - js_2}\right)^n \tag{5.12}$$

but they would be of limited utility; furthermore, it is obvious that both the pass and stop band regions are uniformly distributed angularly around the origin since only a single low pass prototype is used. If such filters are required they would be better designed by the use of the simple transformation of equation (5.7) applied to a more gener ally designed one-dimensional multiband filter. A filter using a third degree transformation is shown in Fig. 5.16.

One application of the transformation which does, however, appear useful, is the generation of a band pass filter by using the second order transformation (5.11) on a high pass prototype filter.

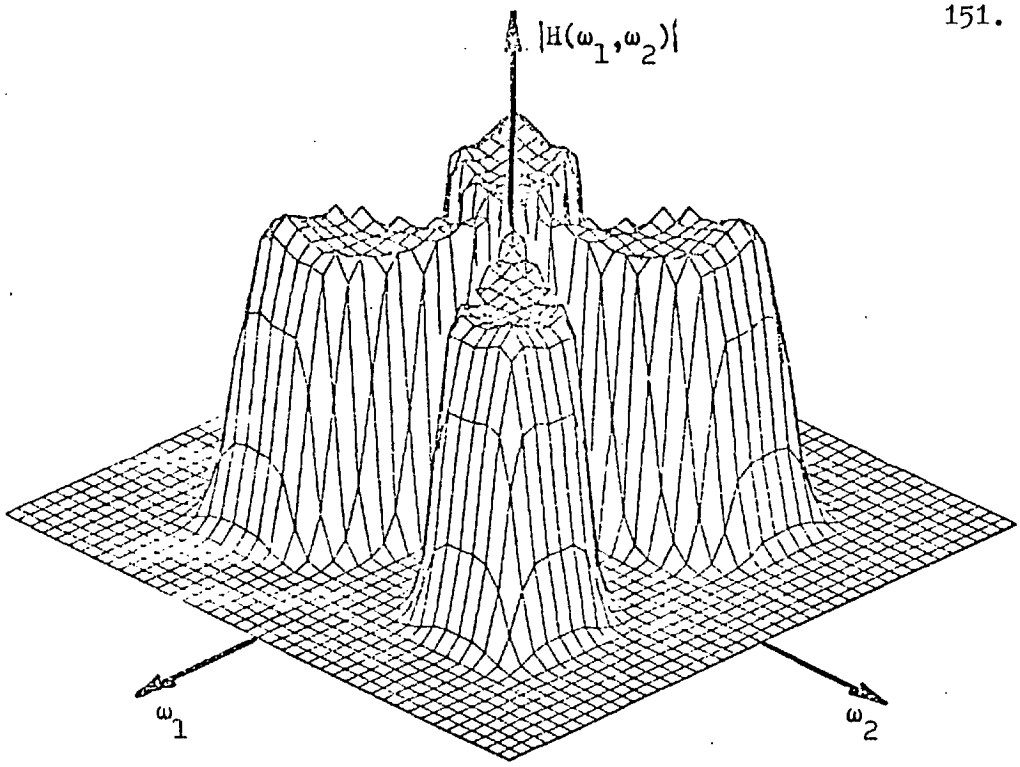


(a) Isometric plot.

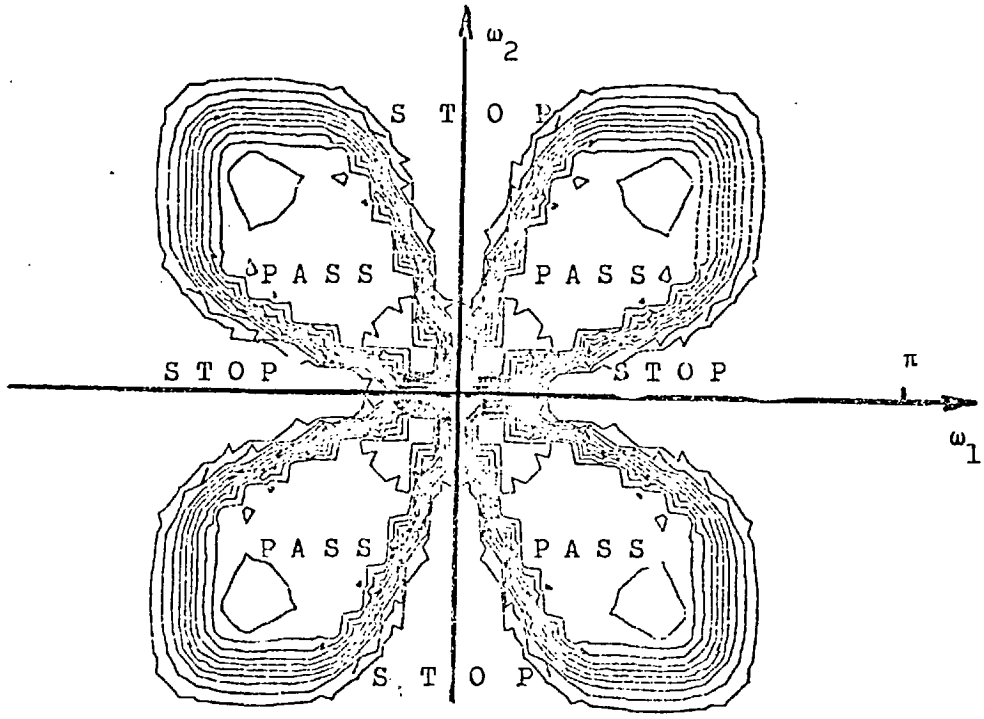


(b) Contour plot.

Figure 5.17. Band-pass fan filter. Prototype: third order Chebyshev high-pass digital filter; $\omega_o = \pi/2$; $\delta = 1\%$.

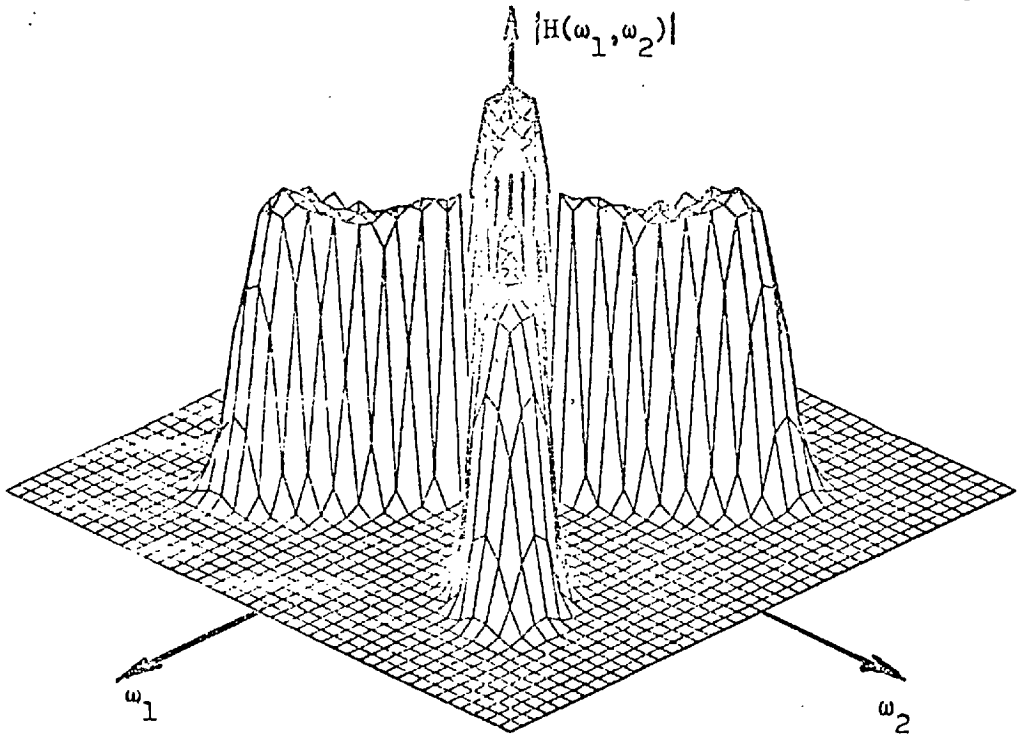


(a) Isometric plot.

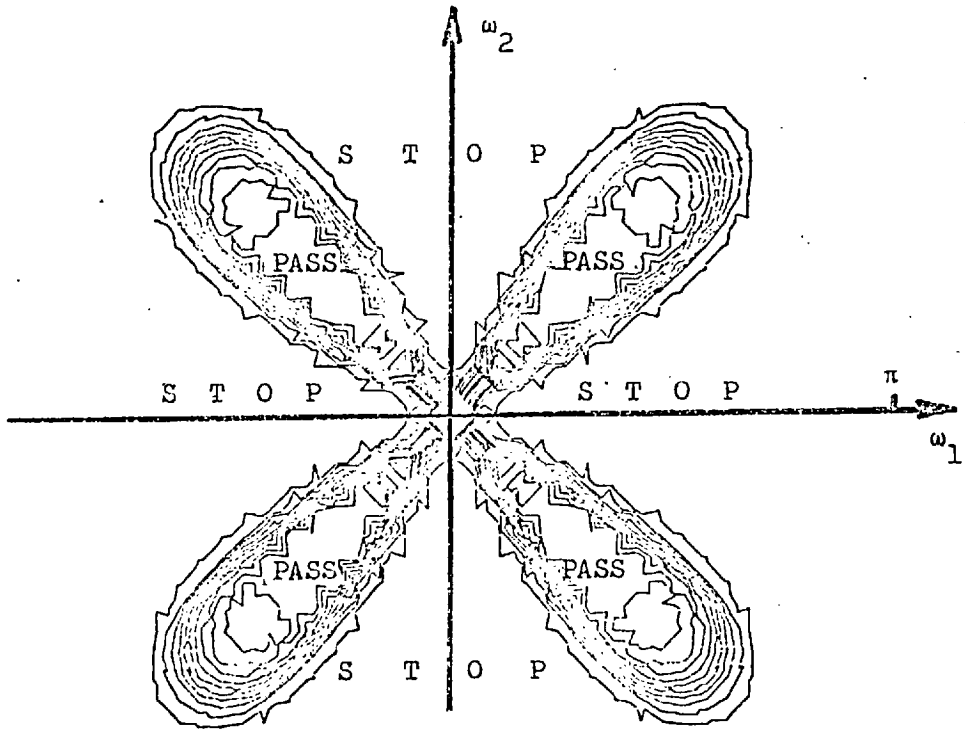


(b) Contour plot.

Figure 5.18. Band-pass fan filter. As in Fig. 5.16, cascaded with a sixth order low-pass guard filter.



(a) Isometric plot.



(b) Contour plot.

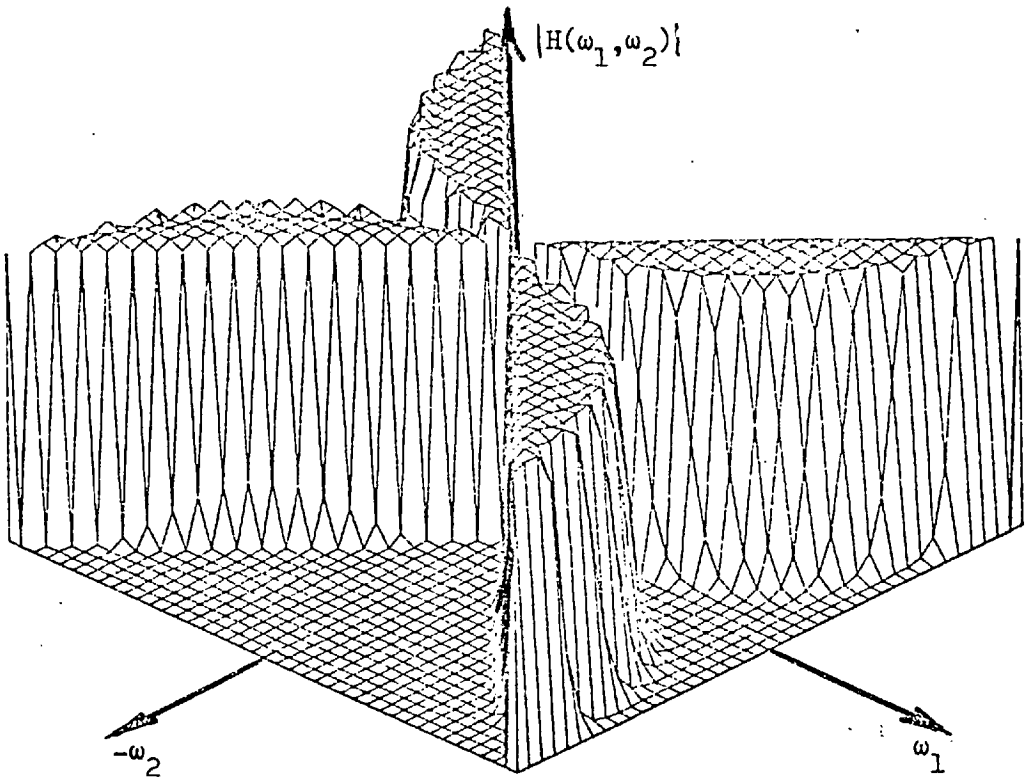
Figure 5.19. Band-pass fan filter. Prototype: third order Chebyshev high-pass filter; $\omega_o = 0.207\pi$, $\delta = 1\%$; cascaded with a low-pass guard filter; $\omega_o = \pi/2$, $\delta = 1\%$.

This is illustrated in Fig. 5.17 using as prototype a high pass filter with cut-off frequency $w_0 = 0.5\pi$; it may be noted that the pass band lies along the diagonals $w_1 = \pm w_2$. This would permit the design of a filter for geophysical prospecting which selected waves which were travelling with apparent velocities lying within a band $V_1 < |\sigma| < V_2$.

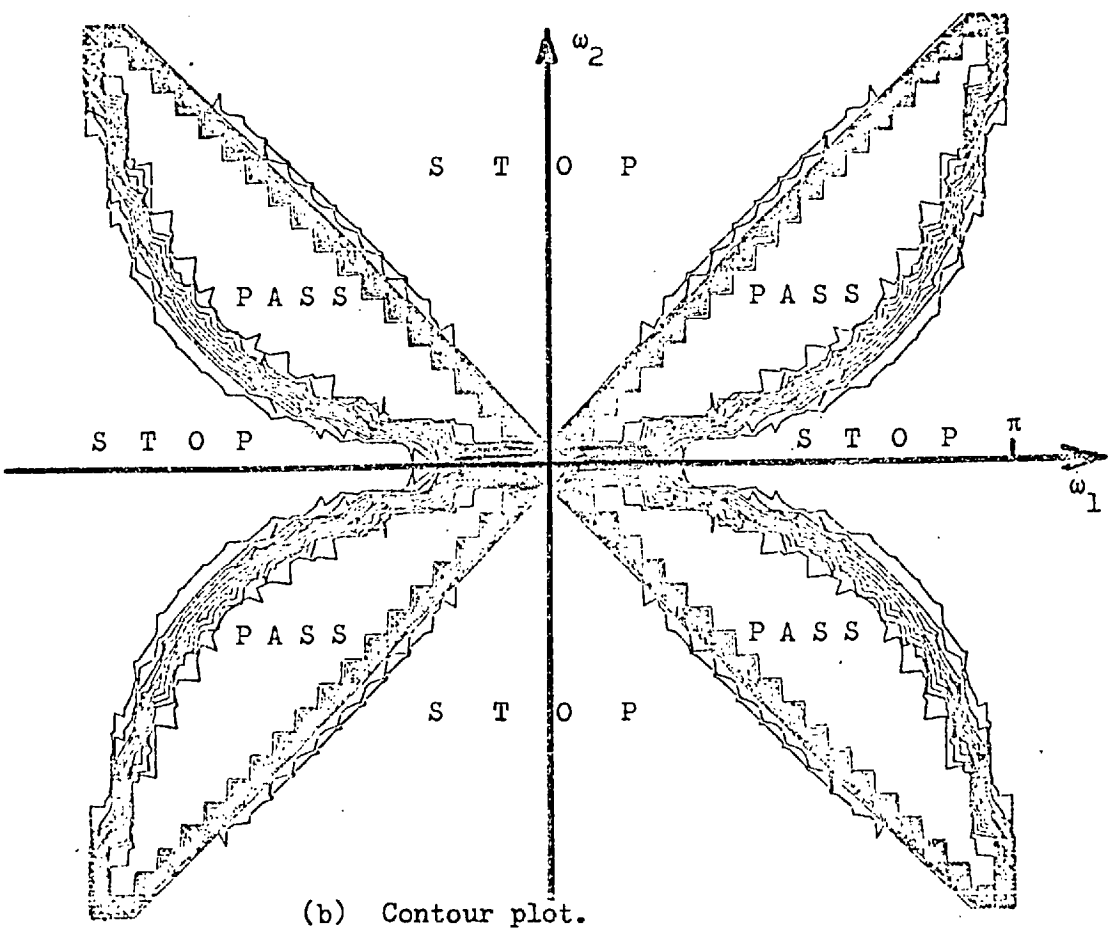
In order to achieve an approximation to fan design with linear cut-off boundaries, we need to cascade this filter with a low pass guard filter. The effect of this is shown in Fig. 5.18. A better response would be obtained using a higher order guard filter, although if this is based on a Chebyshev prototype it would introduce additional ripple in the pass band.

A narrower band pass fan filter is shown in Fig. 5.19 based on a Chebyshev high pass prototype with cut-off frequency at 0.8π ; the guard filter is the same as in the previous example with $w_0 = 0.5\pi$.

Although the particular filters designed in this manner are symmetrical about the principal diagonal, it may be desirable to design band pass fan filters having arbitrary angular orientation of the band edges. These may be closely approximated in the following manner. We may cascade two filters; the first is obtained by a third degree transformation applied to a low pass filter and has the frequency response shown in Fig. 5.16; the second is a simple high pass filter such as that shown in Fig. 5.8 with the w_1 and w_2 axes interchanged. This cascade of filters will give rise to a frequency response as shown in Fig. 5.20; this particular characteristic is designed throughout using a one-dimensional Chebyshev



(a) Isometric plot.



(b) Contour plot.

Figure 5.20. Band-pass fan filter. Cascade of filter of Fig. 5.15 with high-pass filter obtained by transposition of axes of filter of Fig 5.8.

prototype with cut-off frequency at $\pi/2$. A guard filter will then permit isolation of those parts where the cut-off boundary is adequately linear.

It may be seen that a very wide range of filters may be obtained by this technique, giving great freedom to the designer to obtain special profile filters of basically fan form.

5.5 MULTIDIMENSIONAL FAN FILTERS

The use of the one-dimensional to fan transformation (5.7) applied twice in succession along different dimensional axes will generate a class of filters having cut-off profiles which, at present, appear to have little practical utility. They will therefore not be considered.

However, we have noted in Section 5.1 that it is possible for the array of seismographic detectors used in geophysical prospecting to be two- or even three-dimensional. We may therefore need to generate a fan filter which filters out all signals, the magnitude of whose apparent velocity is less than some given value V in any angular direction. This suggests that we start with a two-dimensional filter with a circular cut-off boundary and transform this into a three-dimensional filter by the transformation (5.7). Ideally this should produce a filter with a conical cut-off boundary as shown in Fig. 5.21(a) derived from the two-dimensional filter shown in Fig. 5.21(b).

The prototype in Fig. 5.21(b) has been drawn with a circular cut-off boundary corresponding to a radius $\pi/2$ as we have observed in using the one- to two-dimensional fan transformation

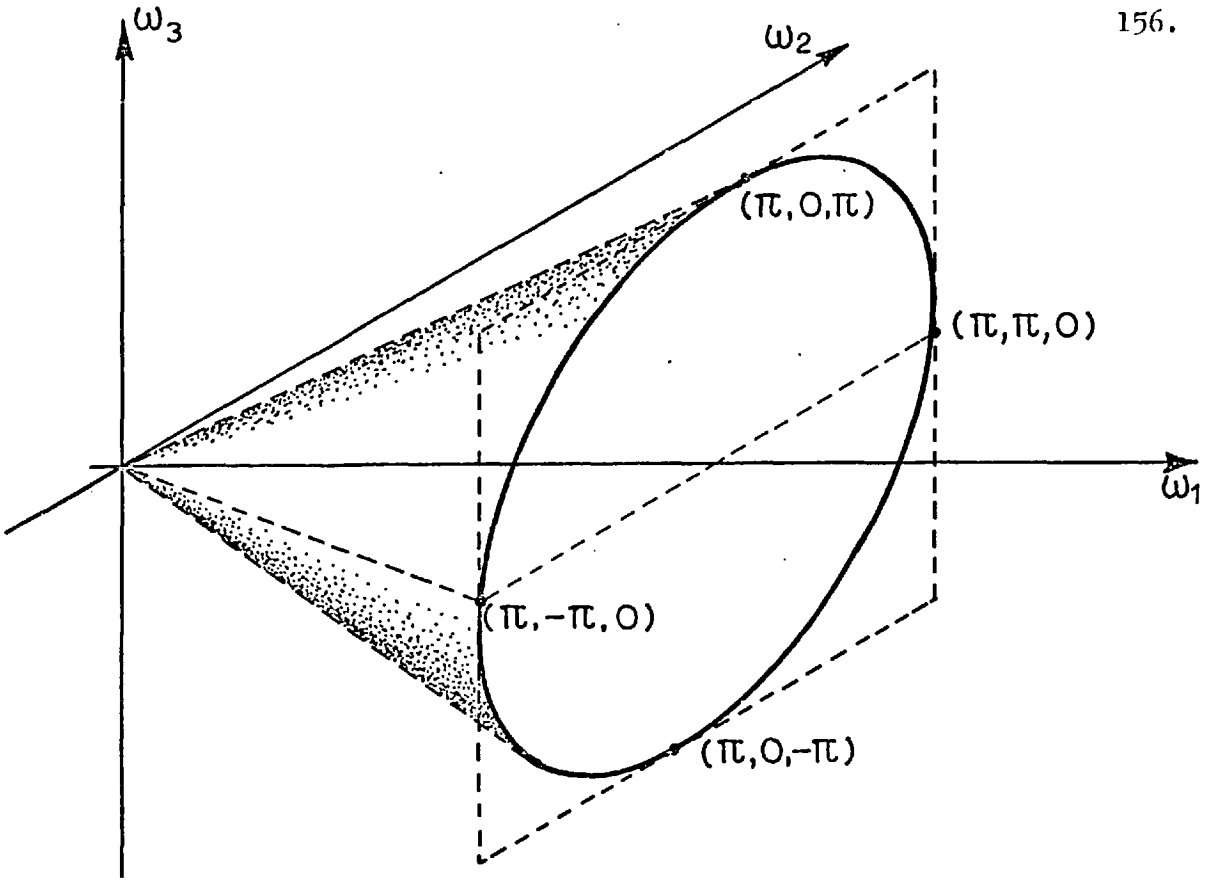


Fig. 5·21(a) Ideal boundary of three-dimensional fan filter.

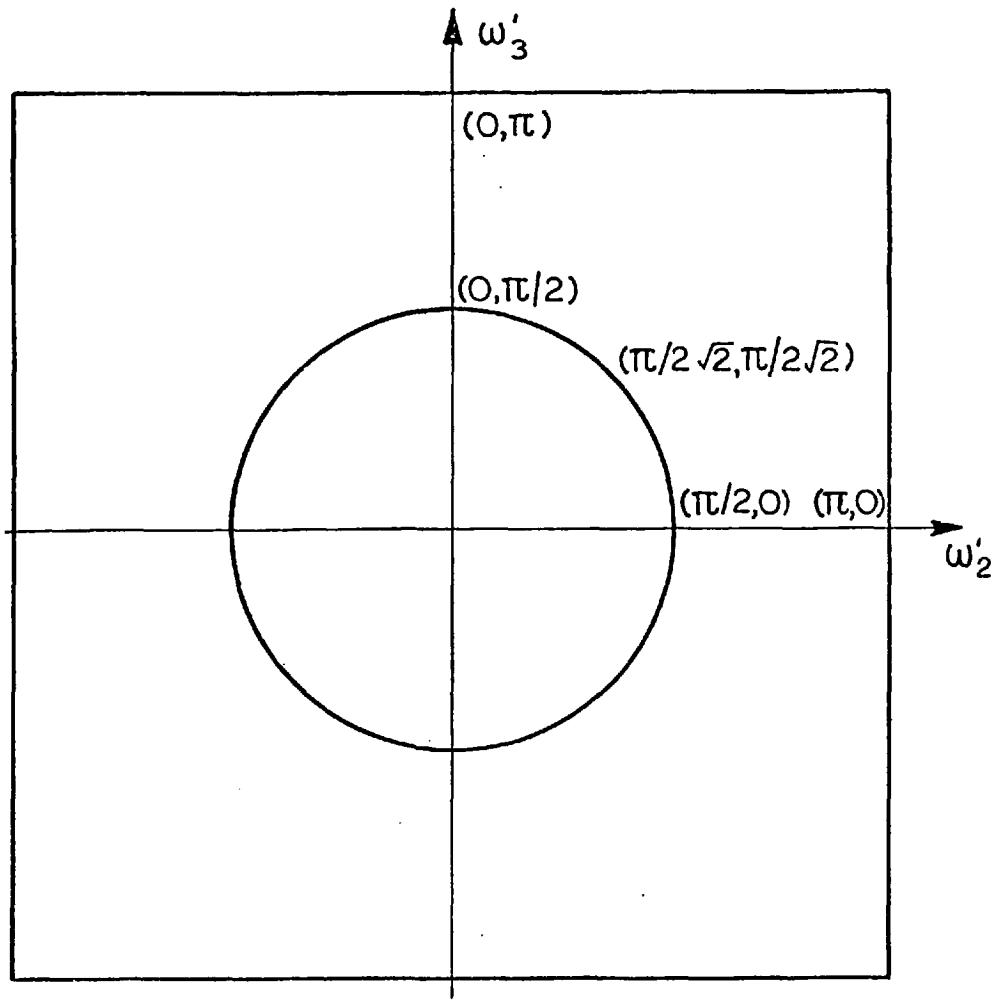


Fig.5·21(b) Cut-off profile for prototype 2-dimensional low-pass filter to design ideal 3-dimensional fan filter.

that this will ensure the cut-off boundary of the filter is linear along the diagonal. Such a prototype in this case will ensure that the point $(\pm\pi/2, 0)$ on the (w_2', w_3') plane maps into all points along the line $w_1 = \pm w_2$, $w_3 = 0$ and $(0, \pm\pi/2)$ maps into the straight line $w_1 = \pm w_3$, $w_2 = 0$.

The nature of the transformation from the w_2', w_3' plane to the w_1, w_2, w_3 space is

$$z_2' = \frac{s_1 - js_2}{s_1 + js_2}$$
$$z_3' = \frac{s_1 - js_3}{s_1 + js_3} \qquad (5.13)$$

followed by the bilinear transformation applied to s_1, s_2, s_3.

It is unfortunate that such a transformation will also map the point $(\pi/2\sqrt{2}, \pi/2\sqrt{2})$ into the corner point (π, π, π). This implies that any shape of cut-off profile of two-dimensional prototype filter will map into the boundary square in three-dimensional space on the planes $w_1 = \pm\pi$.

5.5.1 Circular Cone Filters

The ideal 3-dimensional fan filters based on a two-dimensional prototype discussed in the last section may be termed cone filters from the shape of their cut-off boundaries. We may attempt a design of such a filter using the all-pass transformation to give a two-dimensional prototype filter having cut-off boundary along the two axes at $(w_1', w_2') = (0, \pi/2)$ and $(\pi/2, 0)$. This will ensure that these points map into the straight lines $w_2 = 0$, $w_1 = \pm w_3$ and $w_3 = 0$, $w_1 = \pm w_2$. The boundary along the major diagonal passing

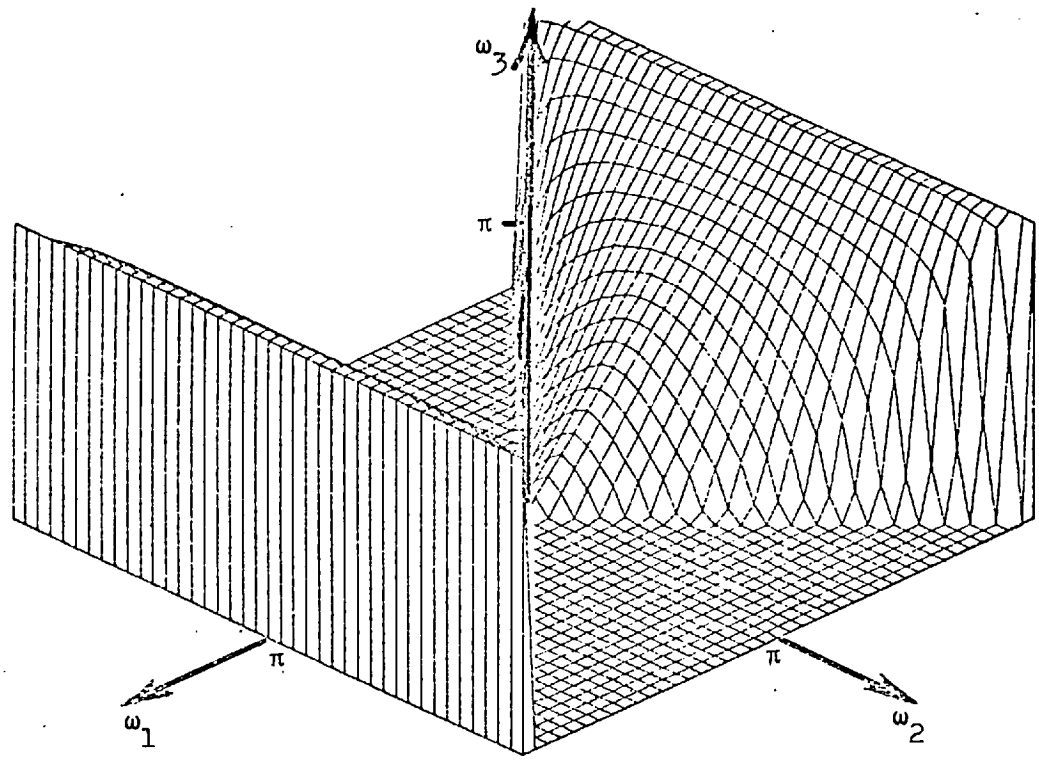


Figure 5.22. Three-dimensional fan filter derived from an all-pass transformed two dimensional digital filter. Prototype: third order Butterworth low-pass digital filter; $\omega_o = \pi/2$.

through (π,π,π) is, however, concave and only in the vicinity of the origin will this boundary lie along a generator of the desired conical surface. A design using this specification is shown in Fig. 5.22. Observations of the sections perpendicular to the w_1-axis show that the boundaries are near-circular close to the origin becoming almost square towards $w_1 = \pi$.

An alternative specification would be to demand that the profile have linear generators along $w_1 = w_2 = w_3$ by choosing a prototype based on a cut-off frequency 0.58π. This will force the boundaries along the coordinate planes, $w_2 = 0$ and $w_3 = 0$, to be convex.

It appears that the nearest approximation to a truly conical filter would be one designed using a prototype filter with quasi-circular cut-off boundary midway between these at 0.54π approximately.

5.5.2 Rectangular Pyramid Filters

As an alternative solution to this problem, we may abandon any attempt at obtaining a quasi-conical filter and settle for a much simpler square pyramidal filter based on the use of transformations (5.13) on a prototype two-dimensional separable filter, having cut-off boundary at $(\pi/2, 0)$ and $(0,\pi/2)$. Such a filter will have an approximately square cross-section at all values of w_1. A "pyramid" filter is shown in Fig. 5.23.

It would be interesting to make a subjective comparison of the results of filtering data from a two-dimensional set of geophysical data by "cone" and "pyramid" filters. Judging by

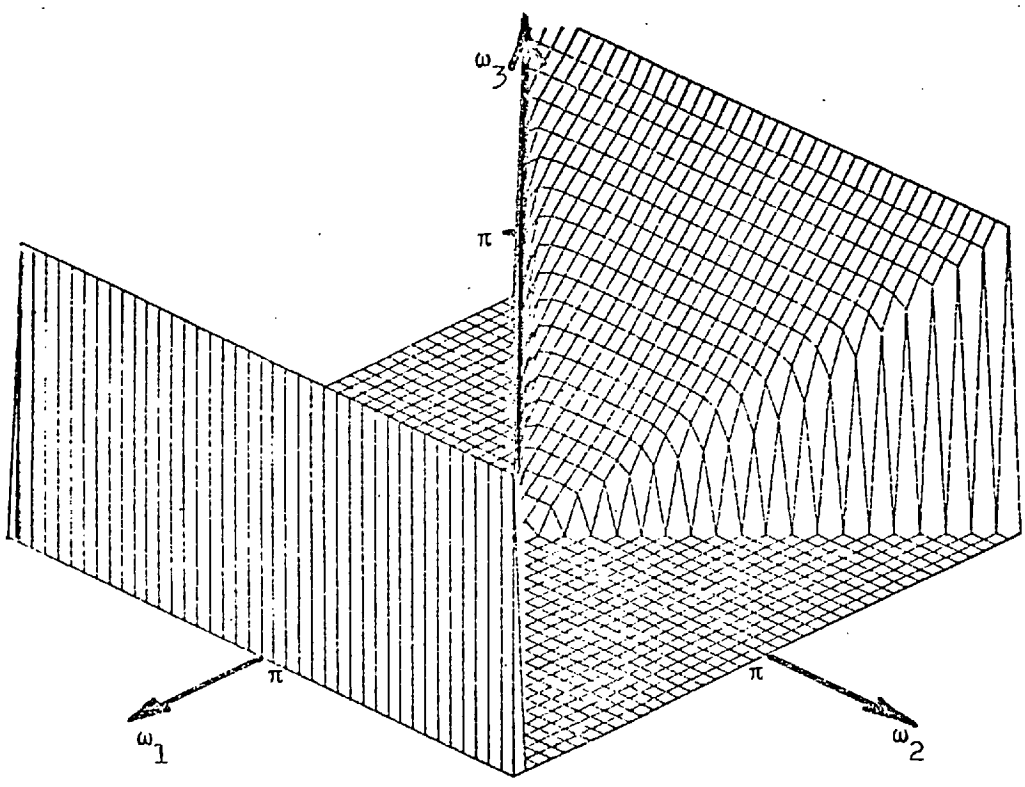


Figure 5.23. Three-dimensional fan filter derived from a separable product two-dimensional filter. Prototype: third order Butterworth low-pass digital filter; $\omega_o = \pi/2$.

comparative assessments on the two-dimensional prototypes, it is likely that the simpler pyramid will be as satisfactory as the quasi one.

5.6 DISCUSSION OF FAN FILTERS

The transformation proposed leads to a family of fan filters. Only in very special cases do these have truly linear cut-off boundaries.

In the majority of cases considered, the cut-off boundary may be made adequately linear by limiting the upper bound on the frequency response of the filter.

Techniques for the design of filters having narrow bands with specified upper and lower cut-off frequencies are also illustrated.

The versatility of the transformation technique is seen in the ability to design a range of filters which otherwise have not been capable of realization.

The extension to systems in which two variables are spatial and one temporal has been demonstrated leading to the design of conical and pyramidal filters.

CHAPTER SIX

REVIEW, CRITICISM AND CONCLUSIONS

Myself when young did eagerly frequent
Doctor and Saint, and heard great argument
About it and about: but evermore
Came out by the same door as in I went.

"Rubaiyat"

Omar Khayyam.

REVIEW, CRITICISM AND CONCLUSIONS

6.0 REVIEW

Although this thesis is, according to its title, concerned with the design of multidimensional digital filters via spectral transformations, it is desirable to consider the purpose of designing multidimensional filters at all, and further the desirability of designing them by the proposed technique.

As outlined in the introduction, the purpose of filters is to process images to modify their properties or to remove undesirable distortion or noise which may have contaminated them so that they only bear a vague resemblance to the object from which they originated. It is therefore apparent that the use of the filter must define the specification of the filter. The specification may be made, in the simplest way, by specifying the point spread function which gives the response of the filter to a unit impulse at the spatial origin. In an alternative representation, the filter specification may be defined in the multidimensional frequency domain. The two specifications are directly linked by the Fourier transform.

However, in any physical system the data is always presented as a function of multidimensional space; a comparison may be made with a one-dimensional temporal system where the data is a function of time. Thus in using a multidimensional signal processing device it is always necessary to consider the input and output data as spatial functions and this affects the manner in which the input data is processed and also, as a consequence, the form of design of the filter.

The earliest signal processors in one dimension consisted of a filter sequence which was convolved in time with the input sequence to give an output sequence; the so-called transversal filters. Such filters in general were termed convolution filters and were obviously of finite length and had a finite impulse response.

A second approach to the problem was via the discrete Fourier transform (DFT) whereby a DFT of the input sequence was determined; this was multiplied at each frequency by the response function of the filter to give the frequency spectrum of the output sequence and finally an inverse DFT gave the desired output. Depending on the nature of the filter function, the output sequence might be of finite length or of infinite duration in response to a finite input sequence.

Either of these techniques can easily be extended to multiple dimension for processing image signals. One of the shortcomings of these processes lay in the considerable volume of computation required. In using convolution filters it was found that filter sequences of thirty or forty elements might be needed to produce adequate filtering in a given situation or to approximate a required transfer characteristic to a desired accuracy. Increase of the dimensions of the system from one to many increases the complexity exponentially. Attempts were made by a number of techniques, including truncation, windowing and, more recently, phase correction to reduce the length of the filter without seriously impairing the response characteristic. Considerable progress has been achieved in this respect although large point spread function arrays are still common.

The second approach to this problem was again hindered by the difficulty of obtaining Fourier transforms of functions by computational techniques. The advent of the Fast Fourier Transform algorithm [6,62] with its many modifications and sophistications went far to overcoming this difficulty. However, in order to obtain accurate results it is necessary to use a large number of points in obtaining the Fourier transform and, although not so seriously, the problem of computational complexity returns.

It is apparent that a convolution filter may process a signal "in real time"; namely, the output value of each element is obtained sequentially and is available a finite time after the corresponding element is read. When the signal is processed via the Fourier transform, it is necessary first to read and store all the input data, then process it as outlined above, whereupon the whole of the output data becomes simultaneously available; thus real time processing is not possible. This also shows that the computer requires sufficient storage for all the input data in addition to that required during the processing algorithm. In one dimension this presents little problem; however, in two and more particularly in three and more dimensions computer storage limits the size of the array which may be handled and thus forces quantization of the image into undesirably large picture elements (pixels). It is, of course, possible to handle large numbers of elements but this involves technical problems in the organization of the computer to store such large arrays and a concomitant increase in the time required for each complete image to be processed. If one of the dimensions of the image is time, the duration of the computational process may completely preclude the use of this technique for filtering anything other than slowly changing images.

For these reasons a third technique of image processing was developed based on the finite difference equations of a recursive filter. By taking the Z-transform of these equations one may derive a recursive transfer function. However, the processing is implemented by a direct application of the finite difference equations as outlined for two dimensions by Shanks [63]. It immediately became apparent that the length of the filter, as defined by the number of coefficients in numerator and denominator to approximate to a given filter response, could be drastically reduced below that required for a comparable convolution filter. Furthermore, it may also be appreciated that each output pixel may be generated sequentially from the input pixels after a finite delay. There is no need for the large storage requirements demanded by the Fourier transform processing technique. This reduction in complexity is extended to multidimensional systems and has stimulated considerable research to be carried out in the design of recursive filters, both one- and multidimensional.

This thesis discusses two of the principal methods by which recursive multidimensional filters may be designed. The first basic class is those which work directly in the space domain; the second class carries out the design exclusively in the frequency domain. Since all recursive filters are essentially feedback systems, the possibility of an unstable design being obtained is always present. It is thus essential for completeness that a survey of the stability tests on systems is undertaken.

6.1 STABILITY

The basic definition of the stability of a system in terms of the absolute summability of the multivariable spatial array is shown to lead to the classical conditions derived by Shanks [5,26] in the Z-domain. Goodman [28] has shown that under certain circumstances these conditions are no longer necessary although always sufficient; an example of such a function appears in Chapter 5.

Many attempts have been made to obtain closed form expressions for the stability of multidimensional recursive filters. For a number of simple cases this has been achieved [10,63] but no general formulation has been obtained. As a consequence of this, a number of computational algorithms have been proposed. These are mainly based on the extension of Shanks theorem to an assessment of the stability of a number of polynomials (equal to the dimensionality of the system) at all points within unit circle considered as a function of one variable. This has been discussed in considerable detail by Jury [64] and doubt has been cast on its validity.

Of the techniques for carrying out the assessment of stability, that proposed by Maria and Fahmy is probably the simplest to use and the most economical in computer time. No comprehensive assessment of the relative efficiencies of the various techniques for high dimension filters has been carried out.

The immediate sequel to a check on the stability of a system is the derivation of techniques for stabilization. Two techniques are available for stabilization of non-zero phase

functions; Shanks [47] approach is based on a conjecture that the planar least squares inverse of a filter is stable and that the double planar least squares inverse will have a magnitude spectral function that closely approximates that of the original function; Read and Treitel [53] base their procedure on an extension of the well known one-dimensional property of the Hilbert transform linking the logarithm of the magnitude and the phase of a minimum phase function.

Unfortunately neither of these techniques is infallible. It has been proven by a number of counterexamples [47,51] that the Shanks conjecture is false even in two dimensions and therefore is unreliable as a stabilization technique in any multi-dimensional system. The Read and Treitel approach appears more hopeful in that the failure to achieve stabilization has been attributed to the necessity to truncate the generated minimum phase array to a finite size before use as the denominator coefficients of a recursive filter transfer function. Bose [56] has also shown that the Hilbert transform is not, in general, applicable to multidimensional systems because of the inability to obtain an appropriate boundary corresponding to the one-dimensional boundary.

Despite these shortcomings, a number of test cases have shown that both these techniques under certain conditions can yield a stable transfer function having a magnitude response closely approximating that of a given unstable filter. It appears that both techniques are fairly satisfactory when the system is grossly unstable but tend to fail in situations where the system is only marginally unstable.

For zero phase systems neither of these techniques is applicable. Pistor [13] proposed a method applicable to two dimensions and Ahmadi and King [18,22] extended it to multiple dimensions which partitions a zero phase filter into a cascade of stable single quadrant recursive filters. The technique is in principle exact and should result in a set of perfectly stable filters, each of which may be implemented by recursion in the appropriate direction; if infinite precision computational facilities were available this would be true. Unfortunately the technique relies on transformation from the spatial to the frequency domain and this may be achieved only by a computational algorithm which evaluates the Fourier transform or its inverse at a finite number of frequencies and to a finite accuracy. Examples have shown that considerable improvement in the accuracy of the procedure may be achieved by increasing the size of the array used for the Fourier transform, which justifies the above argument.

One other limitation of the procedure is that the algorithm decomposes the denominator of the transfer function into a set of single quadrant functions (for a two-dimensional system, four such functions). These single quadrant functions are infinite multidimensional polynomials which for practical purposes require truncation to a finite length; if the decomposed polynomials are fast converging functions this truncation is unlikely to introduce serious error, but in cases in which they are only slowly converging considerable errors may be introduced.

It is likely that the truncation of the decomposed polynomials is responsible for the failure of the Pistor technique to realize a set of stable filters in certain cases, whereas the

finite size of the Fourier transform array may cause the considerable errors which are noticed between the magnitude response of some stabilized filter functions and that of the original functions; considerable errors may be noticed in the fan filter in Chapter 5 which was stabilized using this technique, implemented using an intermediate transform array of dimensions 32 x 32, which is too small to take account of the sharp cut-off profile effectively.

An extension of the technique proposed by Ekstrom and Woods [14] removes some of the errors produced in the amplitude response function by using certain weighting functions on the decomposed sequences. Another improvement has been proposed [65] which uses an optimization procedure to modify the numerator of the transfer function to compensate for the errors introduced by the denominator. Yet another proposal has been made by Rousogiannakis [66] in which the phase correction algorithm [67] may be used to modify the truncated demoninator arrays to give better approximations to the desired magnitude response without affecting the frequency response; this has shown promise of only minor improvement. However, modification of the numerator polynomial by the phase correction technique to compensate for the errors due to truncation of the partitioned denominator shows greater potential.

6.2 SPATIAL DESIGN TECHNIQUES

Initially a survey was made of the presently available techniques for design in the space domain of a recursive filter having a given multidimensional spread function. No constraints were placed on the functional form of the specification;

however, designs could only be achieved if the point spread function was causal and absolutely summable. The principal disadvantage of most of these techniques is that the designed filter may not necessarily be stable, since the filter transfer function is always derived as an approximation to the specified stable point spread function. It is therefore imperative to apply one of the stability tests discussed in Chapter 2; in those cases in which the filter is shown to be unstable, one of the stabilization techniques (either Shanks or Read and Treitel) must be used. It has been remarked that the original design is, at best, an approximation to the desired point spread function. Moreover it is well known that both these stabilization techniques inevitably distort the amplitude response function as discussed in Section 6.1, and that they do not guarantee that the resulting filter will be stable; thus a final stability check needs to be carried out on the designed transfer function.

As an alternative stabilization method, we could consider that of Pistor. However, the specifications in all the spatial design techniques are of single quadrant filters and although there is no evidence that the Pistor stabilization method is unsatisfactory in such cases, no evidence is available showing the conditions, if any, under which the technique may be used for non-zero phase systems.

The only spatial design technique which guarantees a stable filter is that proposed by Bordner. Unfortunately, it suffers from the other drawbacks, the principal of which is that the solution does not converge on a global minimum. It also relies

on the choice of an "extension array" to the given finite size point spread array which is a "natural extension". The choice of this "tail array" is critical in reducing the complexity of computation.

It is thus seen that of the spatial design techniques considered, only that of Bordner may really be considered as a viable technique and even here the computational problems may well be very great.

The above deductions are based on designs in two dimensions. Attempts at three- and higher dimensional systems do not appear to have been attempted although the theory has been established [36,37]. However, for such higher-dimensional systems the preceding criticisms will carry even greater weight.

6.3 FREQUENCY TRANSFORMATION TECHNIQUES

A number of frequency transformation techniques have been considered with particular reference to transfer functions having cut-off boundaries which approximate to a hypersphere (or, in the two-dimensional case, to a circle).

The earliest of these was proposed by McClellan, who used a one- to two-dimensional transformation applied to a non-recursive filter to obtain a circularly symmetric two-dimensional filter. The technique has been extended by Bernabo et al to recursive filters and results in a filter which can be made to have a very good circular cut-off boundary but without any guarantee of stability of the designed filter; in fact, since the

designed filter is zero phase it will of necessity be unstable and in order to obtain a practical implementation it is necessary to decompose it by the Pistor technique. The criticism of the method lies in the large amount of computation time required, although the designed filter has the best circular symmetry of any of the considered techniques.

The Ahmadi procedure provides a transformation from one to two dimensions via a simple two-dimensional reactance function. It ensures a stable filter with cut-off profile approximating circular, in particular for higher values of cut-off frequency. It is unfortunate that spurious pass bands exist when this transformation is used, but these may be removed by a low order low pass separable guard filter. The method has the advantage that it is relatively simple to extend the technique to systems of any number of dimensions. A simplified form of the Ahmadi transformation postulated by Ali eliminates the spurious pass bands but is even more restrictive in the generation of cut-off boundaries of approximate hyperspherical form.

The all-pass transformation postulated has the advantage over the Ahmadi transformation that better cut-off boundary shape may be achieved, but similar constraints hold on the frequency ranges over which this is satisfactorily obtained. Spurious pass bands exist when using the general form of the transformation, which may be removed by a guard filter; the special form of the transformation put forward by Kap eliminates these but restricts the cut-off frequency range for circular symmetry. The all-pass transformation may also be very simply extended to multidimensional systems, either in its general or more restrictive form.

In contrast to these transformations we may consider the simplest and earliest transformation, namely that obtained from multiplying together two filter functions, each of which is dependent upon only one frequency variable and which may therefore be trivially generated from a one-dimensional filter. Although in two dimensions the cut-off profile is almost rectangular, this defect appears to be of little significance in many low pass filtering problems. The technique has the overriding advantage that the design procedure is simpler than any of the other transformations and very little increase in complexity accrues when it is extended to any number of dimensions.

The design of fan filters has also been attacked using a new spectral transformation. In the simple situation where the two axes are time and distance, fan filters with accurately linear cut-off profiles may be designed having the steepness of the cut-off boundary defined in a simple manner by the order of the prototype low pass filter from which it was generated. Such filters have superior characteristics to any of those designed by presently available techniques.

More complex fan filters may also be designed by the same technique but, except in trivial cases, the cut-off boundaries are curved; approximation to linearity may be achieved by limiting the frequency range of the filter in both directions by means of a low pass guard filter. Superposition of two types of fan filter has been shown to produce the equivalent of a fan band pass filter; this technique opens up many multiple band possibilities.

A final application of the fan transformation is shown in a three-dimensional application which could have value in geophysical prospecting where two-dimensional spatial arrays of geophones are used.

6.4 FUTURE FIELDS OF WORK

Circular cut-off filters have been the subject of much of this thesis and the methods of designing to a pseudo-circular cut-off boundary have in general consisted of ensuring a fit at three points. Such a technique will not necessarily ensure an approximation in a mean square error sense and it is open to investigation whether a better design method could be derived based on a minimization of some error function. The subjective assessments obtained using separable filters suggest that it is unlikely that this will yield filters which have superior properties in any practical applications.

In the use of both the Ahmadi and the all-pass transformations, frequency bands have been given, outside which quasi-circular filters may not be designed. It would be of interest to determine what deviation from circular symmetry results when the design of filters having cut-off frequencies below this limit is attempted. It is conjectured that these frequency limits are not rigid but that the parameters of the transformation will enable a choice of filter to be made to optimize the cut-off boundary. In the restricted forms of either of these transformations in which spurious pass bands are eliminated this becomes of considerable significance.

Similar investigations into the linearity of boundaries of fan filters may also show preference for certain forms of transformation function and prototype filter.

With the wide range of spectral transformations now available it would be interesting to compare the various filters for factors such as ease of implementation, quantization errors, etc.

6.5 CONCLUSION

A very brief survey of the application of some circular symmetry filters in the processing of one particular image has shown that by a crude subjective assessment of the results, the more complex filters have little advantage over the simpler ones, once again justifying the words of William of Occam.

REFERENCES

1. R.W. Selzer, "Improving Biomedical Image quality with Computers", NASA Technical Report 32-1336, 1968.

2. S. Treitel, J.L. Shanks and C.W. Frasier, "Some Aspects of Fan Filtering", Geophysics, Vol. 32, p. 789, Oct. 1967.

3. M.B. Dobin, Introduction to Geophysical Prospecting, McGraw-Hill, New York, 1960.

4. R.M. Mersereau and D.E. Dugeon, "The Representation of Two-dimensional Sequences as One-Dimensional Sequences", IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-22, No. 5, p. 320, Oct. 1974.

5. J.L. Shanks, S. Treitel and J.H. Justice, "Stability and Synthesis of Two-Dimensional Recursive Filters", IEEE Trans. on Audio and Electroacoustics, AU-20, p. 115, June 1972.

6. J.W. Cooley and J.W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series", Mathematics of Computation, Vol. 19, p. 297, 1965.

7. D.K. Darby and E.B. Davies, "The Analysis and Design of Two-dimensional Filters for Two-Dimensional Data", Geophysical Prospecting, Vol. 15, p. 383, 1967.

8. T.S. Huang, "Two-dimensional Windows", IEEE Trans. on Audio and Electroacoustics, AU-20, p. 88, March 1972.

9. J.H. McClellan, "On the Design of One-Dimensional and Two-Dimensional FIR Digital Filters", Ph.D. Dissertation, Dept. of Electrical Engineering, Rice University, 1973.

10. T.S. Huang, "Stability of Two-dimensional Recursive Filters", IEEE Trans. on Audio and Electroacoustics, AU-20, p. 158, June 1972.

11. B.D.O. Anderson and E.I. Jury, "Stability Test for Two-dimensional Recursive Filters", IEEE Trans. on Audio and Electroacoustics, AU-21, p. 366, Aug. 1973.

12. G.A. Maria and M.M. Fahmy, "On the stability of Two-dimensional Digital Filters", IEEE Trans. on Audio and Electro-acoustics, AU-21, p. 470, Oct. 1973.

13. P. Pistor, "Stability Criterion for Recursive Filters", IBM Journal of Research and Development, 18, p. 59, Jan. 1974.

14. M.P. Ekstrom and J.W. Woods, "Two-dimensional Spectral Factorization with Applications in Recursive Digital Filtering", IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-24, No. 2, April 1976.

15. E.L. Hall and S.J. Dwyer, "Frequency Domain Design of Spatial Digital Filters", Proc. 1970 International Symposium on Information Theory, June 1970.

16. J.M. Costa and A.N. Venetsanopoulos, "Design of Circularly Symmetric Two-dimensional Recursive Filters", IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-22, p.432, Dec. 1974.

17. M. Bernabo, P.L. Emiliani and V. Cappellini, "A method for designing 2-D Recursive Digital Filters having a Circular Symmetry", Proc. of Florence Conference on Digital Signal Processing, Florence, September 1975.

18. M. Ahmadi, "Design, Stability and Applications of Two-dimensional Recursive Digital Filters", Ph.D. Thesis, Imperial College of Science and Technology, University of London, 1977.

19. M. Ahmadi, A.G. Constantinides and R.A. King, "Design Technique for a Class of Stable Two-dimensional Recursive Digital Filters", Proc. 1976 IEEE International Conference on Acoustics, Speech and Signal Processing, Philadelphia, April 1976.

20. G.W. Bordner, "Time Domain Design of Stable Recursive Digital Filters", Ph.D. Dissertation, State University of New York at Buffalo, 1974.

21. M. Lal, "Space Domain Design of Two-dimensional Recursive Digital Filters", Proc. 1975 IEEE International Symposium on Circuits and Systems, Boston.

22. M. Ahmadi and R.A. King, "A Stability Criterion for N-dimensional Zero-phase Recursive Digital Filters", Proceedings of IEEE, 65, No. 6, p. 893, June 1977.

23. M. Kap, "A Comparison of Different Two-Dimensional Recursive Digital Filter Design Techniques", M.Sc. Thesis, Imperial College of Science and Technology, University of London, 1977.

24. J.H. McClellan and T.W. Parks, "Equiripple Approximation of Fan Filters", Geophysics, 37, p. 573, August 1972.

25. C. Farmer and J.B. Bednar, "Stability of Spatial Digital Filters", Mathematical Biosciences, 14, p. 113, 1972.

26. J.H. Justice and J.L. Shanks, "Stability Criterion for N-dimensional Digital Filters", IEEE Trans. on Automation and Control, AC-18, p. 284, June 1973.

27. B.D.O. Anderson and E.I. Jury, "Stability of Multi-dimensional Recursive Filters", IEEE Trans. on Circuits and Systems, CAS-21, p. 300, March 1974.

28. D. Goodman, "Some Stability Properties of Two-dimensional Recursive Filters", IEEE Trans. on Circuits and Systems, CAS-24, p. 201, April 1977.

29. H.G. Ansell, "On Certain Two-variable Generalizations of Circuit Theory, with Applications to Networks of Transmission Lines and Lumped Reactances", IEEE Trans. on Circuit Theory, p. 214, June 1964.

30. A. Cohn, "Ueber die Anzahl der Wurzeln einer algebraischen Gleichung in einer Kreise", Math. J., p. 110, 1922.

31. E.I. Jury, "A note on the Reciprocal Zeros of a Real Polynomial with respect to the Unit Circle", IEEE Trans. on Circuit Theory, p. 292, June 1964.

32. E.I. Jury, "Remarks on the Mechanics of Bilinear Transformation", IEEE Trans. on Audio and Electroacoustics, AU-21, p. 380, August 1973.

33. E.I. Jury and O.W.C. Chan, "Combinatorial Rules for some Useful Transformations", IEEE Trans. on Circuit Theory, CT-20, p. 476, Sept. 1973.

34. R.E. Kalman, "Design of a Self-Optimizing Control System", Trans. ASME, 80, p. 468, 1958.

35. K. Steiglitz and L.E. McBride, "A Technique for the Identification of Linear Systems", IEEE Trans on Automatic Control, AC-10. p. 461, 1965.

36. B. Nowrouzian, "Space Domain Design and Stability of N-dimensional Recursive Digital Filters", M.Sc. Thesis, Imperial College of Science and Technology, University of London, 1976.

37. B. Nowrouzian, M. Ahmadi and R.A. King, "On the Space Domain Design Techniques of N-dimensional Recursive Digital Filters", Proc. IEEE International Conference on Acoustics, Speech and Signalling, May 1977.

38. J.L. Shanks, "Recursion Filters for Digital Processing", Geophysics, 32, p. 33, 1967.

39. M.S. Bertran, "Approximation of Digital Filters in One and Two Dimensions", IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-23, p. 438, 1975.

40. W.H. Huggins and R.N. McDonough, "Best Least Squares Representation of Signals by Exponentials", IEEE Trans. on Automatic Control, AC-13, p. 408, 1968.

41. M. Lal, "Space Domain Design of Two-dimensional Recursive Digital Filters", Proc. 1975 IEEE Trans. on Automatic Control, AC-13, p. 408, 1968.

42. N.A. Prendergrass, S.K. Mitra and E.I. Jury, "Spectral Transformations for Two-dimensional Digital Filters", IEEE Trans. on Circuits and Systems, CAS-23, No. 1, p. 26, January 1976.

43. J.M. Costa and A. Venetsanopoulos, "Design of Circularly Symmetric Two-dimensional Recursive Filters", IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-22, p. 432, December 1974.

44. R.M. Mersereau et al., "McClellan Transformation for Two-dimensional Digital Filtering", IEEE Trans. on Circuits and Systems, CAS-23, No. 7, July, 1976.

45. J.P. Fail and G. Grau, "Les Filtres en Eventail", Geophysical Prospecting, 11, p. 131, 1963.

46. P. Embree, J. Burg and M. Backus, "Wide Band Velocity Filtering - the Pie-Slice Process", Geophysics, 28, p. 948, 1963.

47. J.L. Shanks, S. Treitel and J.H. Justice, "Stability and Synthesis of Two-dimensional Recursive Filters", IEEE Trans. on Audio and Electroacoustics, AU-20, p. 115, June 1972.

48. E.A. Robinson, "Statistical Communication and Detection", Hafner, New York, 1967.

49. Y. Genin and Y. Kamp, "Counterexample in the Least Squares Inverse Stabilization of 2-D Recursive Filters", Electronics Letters, 15, p. 330, July 1975.

50. Y. Genin and Y. Kamp, "Comments on the Stability of the Least Mean-Square Inverse Process in Two-dimensional Digital Filters", IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-25, p. 92, February 1977.

51. M. Tola, "Stabilization Techniques for Two-dimensional Recursive Digital Filters", M.Sc. Thesis, Imperial College of Science and Technology, University of London, 1977.

52. V. Cizek, "Discrete Hilbert Transform", IEEE Trans. on Audio and Electroacoustics", AU-18, No. 4, p. 340, December 1970.

53. R.R. Read and S. Treitel, "The Stabilization of Two-dimensional Recursive Filters via the Discrete Hilbert Transform", IEEE Trans. on Geoscience Electronics, GE-11, No. 3, p. 153, July 1973.

54. J.L. Shanks and R.R. Read, "Techniques for Stabilizing Two-dimensional Finite Difference Operator", Proceedings of Colloque International on Point Transformations and their Applications, Toulouse, September 1973.

55. M.G. Strintzis, "Tests of Stability of Multidimensional Filters", IEEE Trans. on Circuits and Systems, CAS-24, No. 8, August 1977.

56. N.K. Bose, "Problems in Stabilization of Multidimensional Filters via Hilbert Transform", IEEE Trans. on Geoscience Electronics, GE-12, No. 1, p. 146, Oct. 1974.

57. A.M. Ali and A.G. Constantinides, "On Two-Variable Reactance Functions for Two-dimensional Recursive Filter Design", Electronics Letters, Vol. 14, No. 1, Jan. 1978.

58. M. Ahmadi and R.A. King, "Transformation Technique for N-dimensional FIR Filter", Proc. 20th Midwest Symposium on Circuits and Systems, August 1977.

59. J.H. McClellan and T.W. Parks, "Equiripple Approximation of Fan Filters", Geophysics, 37, No. 4, p. 573, 1972.

60. R.L. Senghush and M.R. Foster, "Design and Application of Optimal Velocity Filters in Seismic Exploration", IEEE Trans. on Computers, C-21, No. 7, p. 648, 1972.

61. R.L. Senghush and M.R. Foster, "Optimum Multichannel Velocity Filters", Geophysics, Vol. 33, No. 1, 1968.

62. B. Gold and C. Rader, "Discrete Fourier Transforms", Digital Processing of Signals, New York, McGraw-Hill, 1969.

63. D. Davis and L. Souchon, "General Stability Conditions for N-dimensional Digital Filters", Record of the Nineth Annual Asilomar Conference on Circuits, Systems and Computers, November 1975.

64. E.I. Jury, "Inners and Stability of Dynamic Systems", John Wiley and Sons, 1974.

65. F. Bernabo, P.L. Emiliani and V. Cappellini, "Design of Two-dimensional Recursive Digital Filters", Electronics Letters, Vol. 12, No. 11, May 1976.

66. P. Rousogrannakis, Private communication.

67. M.T. Manty and J.K. Aggarwal, "The Design of Multidimensional FIR Digital Filters by Phase Correction", Proc. 1975 International Symposium on Circuits and Systems, Boston, U.S.A., April 1975.