THE DESIGN OF TWO-DIMENSIONAL DIGITAL FILTERS BY GENERALIZED McCLELLAN TRANSFORMATIONS

by

THOMAS FRANCIS QUATIERI, JR.
B.S., Tufts University
(1973)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June, 1975

Signature of Author: Signature redacted
Department of Electrical Engineering
May 23, 1975

Certified by: Signature redacted
Thesis Supervisor

Accepted by: Signature redacted
Chairman, Departmental Committee on Graduate Students

THE DESIGN OF TWO-DIMENSIONAL DIGITAL FILTERS BY GENERALIZED McCLELLAN TRANSFORMATIONS

by

Thomas Francis Quatieri, Jr.

Submitted to the Department of Electrical Engineering on May 23, 1975, in partial fulfillment of the requirements for the degree of Master of Science.

ABSTRACT

The design of two-dimensional linear phase FIR digital filters, optimal in the Chebyshev sense, is a very time consuming process. An alternative design technique proposed by McClellan is the transformation of a one-dimensional linear-phase FIR equiripple filter into a two-dimensional linear-phase FIR filter which is nearly optimal in the Chebyshev sense. The technique is very fast since computing time is devoted almost entirely to the one-dimensional design which is known to be efficient. Also, there exists an efficient implementation which in general requires fewer multiplies per output point than conventional techniques.

The original McClellan transformation is generalized to approximate arbitrary two-dimensional contours and compress transition width. Constraints are derived for guaranteeing a well-defined transformation and an algorithm is presented for satisfying arbitrary two-dimensional specifications in terms of one-dimensional filter parameters.

Finally, comparisons are made with the transformation technique, the optimal design, and various windowing procedures.

Thesis Supervisor: Russell M. Mersereau

Title: Research Associate

ACKNOWLEDGEMENTS

I extend my warm thanks to my advisor Russ Mersereau for his guidance and understanding during this work. I also thank Wolfgang Mecklanbräuker for his enlightening discussions. Together they have helped me to accept and appreciate the constantly changing, yet growing state of my research.

Finally, I thank Mrs. Delphine Radcliffe for the excellent typing of this manuscript.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

CHAPTER 1

INTRODUCTION

1.1 The Subject and Scope of the Thesis

In this thesis we shall be concerned with the design of two-dimensional finite duration impulse response (FIR) digital filters. A novel design technique due to McClellan [1], [2] is extended and generalized. This method transforms a one-dimensional FIR filter into a two-dimensional FIR filter by a change of variables. The transformation preserves the amplitude characteristic of the one-dimensional filter.

The remainder of this chapter presents the design problem and historical background which leads to the introduction of the McClellan transformation.

Chapter 2 is first concerned with the history and mathematics of the original McClellan transformation and its generalization to higher order transformations. The contour approximation problem for an arbitrary frequency response is then formulated. Finally, an efficient implementation scheme is presented. We shall see that with this scheme the filters can be implemented with fewer multiplies than with conventional implementations.

Chapter 3 discusses the constraints needed for a well-defined transformation and for transition width compression. Also, an algorithm is presented for meeting two-dimensional specifications.

The versatility of the technique is demonstrated in Chapter 4 by the design of arbitrary frequency responses.

Finally, Chapters 5 and 6 make comparisons with previous design methods and present some recent ideas with suggestions for further

research.

1.2 The Design Problem

Digital filters are categorized as either finite impulse response (FIR) or infinite impulse response (IIR) filters. FIR filters, unlike IIR filters, are always stable, can have exactly linear phase, and with an appropriate finite delay can be made realizable. Also, as shown by Parks and McClellan [3], the efficient 2nd Remes Exchange Algorithm is suitable for the design of equiripple one-dimensional FIR filters. This method, however, cannot be used in the design of optimal two-dimensional FIR filters. Thus, alternate techniques are being sought for both optimal and suboptimal designs.

The one-dimensional design goal is to minimize the error over the interval $[o, \pi]$ between the ideal frequency response $I(\omega)$ and the actual frequency response $H(\omega)$, under some error criterion. A two-dimensional FIR filter is characterized by the two-dimensional frequency response given by

$$H(\omega_1, \omega_2) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} h(n, m) \exp(-j\omega_1 n) \exp(-j\omega_2 m) \qquad (1.1)$$

where h(n, m) is of duration N x N. The design goal is analogous to the one-dimensional case. For example, the ideal frequency response might be specified over the two-dimensional region $[-\pi, \pi] \times [-\pi, \pi]$, in which a passband value of one and a stopband value of zero must be closely met with an allowable transition width. Figure 1.1 shows a desired set of specifications over the region $[-\pi, \pi] \times [-\pi, \pi]$ for a circularly symmetric lowpass filter.

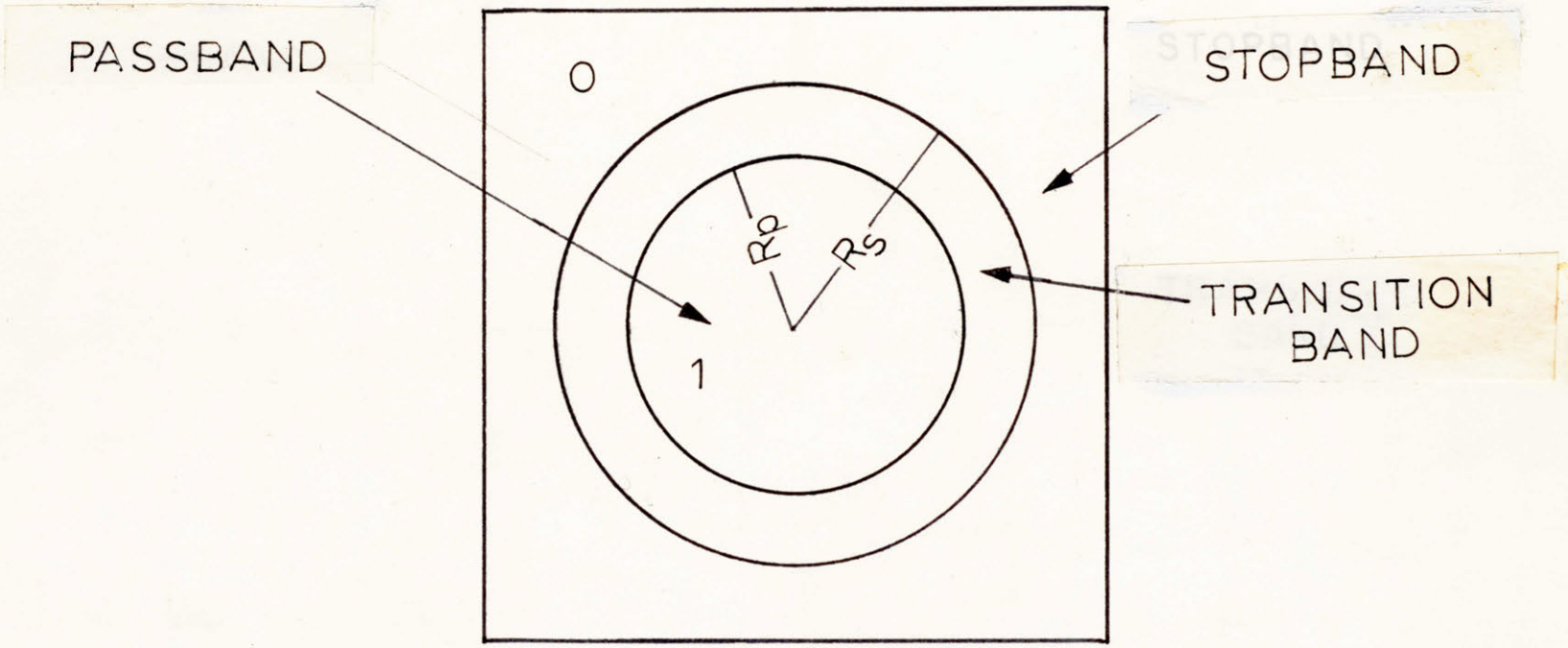

Fig. 1.1 The two-dimensional design specifications for a circularly symmetric lowpass filter.

1.3 Previous Development

A number of techniques have been applied for meeting arbitrary two-dimensional specifications. FIR filters, optimal in the least squares sense, can be designed by multiplying the unrealizable ideal frequency response by an N x N rectangular window. The rectangular window is defined by

$$w(n,m) = \begin{cases} 1 & 0 \le n \le N-1,\ 0 \le m \le N-1 \\ 0 & \text{otherwise} \end{cases} \tag{1.2}$$

Rectangular windowing, however, produces a frequency response which has large ripples or overshoots at the discontinuities of piecewise constant ideal responses. This is analogous to the Gibbs phenomenon in one dimension. See Chapter 5 for an example.

Huang [4] has shown that good two-dimensional windows can be obtained from good one-dimensional windows via the relation $w(n,m) = \hat{w}(\sqrt{n^2 + m^2})$, where $\hat{w}(t)$ is an appropriate one-dimensional continuous window sampled at the appropriate values. Another possible window is $w(n,m) = w_1(n)w_2(m)$, where $w_1(n)$ and $w_2(m)$ are good one-dimensional windows. However, since windowing is a convolution process in the frequency domain, discontinuities in the ideal response are smeared. Hence, for example, very narrow band filters with specific cutoff frequencies are hard to obtain without using very long windows. Another difficulty arises in taking the inverse transform of an ideal filter characteristic to determine the ideal impulse response. In spite of these handicaps, windowing has proven useful due to its speed and flexibility in approximating arbitrary ideal frequency responses.

Another method of design is frequency sampling [5], where points

in the transition region of an ideal DFT are left free to optimize the filter under some minimization criterion. The two-dimensional DFT is defined as

$$H(k, \ell) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} h(n, m) \exp(-j\frac{2\pi}{N}kn) \exp(-j\frac{2\pi}{N}\ell m) \qquad (1.3)$$

This design lacks flexibility in specifying cutoff frequencies; also, the approximation error tends to be highest around the transition edges.

A large concentration of error near the edges can be undesirable in some applications. This leads to the equiripple approximation design of linear phase FIR filters. By definition, a linear phase one-dimensional FIR filter is of the form

$$H(\omega) = G(\omega) \exp(j(A + B\omega))$$

where $G(\omega)$ is real. Parks and McClellan [3] have shown that one of four possible solutions is a positive symmetric, odd length impulse response, where $A = 0$, $B = -(n-1)/2$, and $h(n) = h(N-1-n)$. Since $\exp(j(A + B\omega))$ contributes only to phase, $h(n)$ is assumed centered at the origin with zero phase and length $N = 2n + 1$. Therefore, $h(n)$ is expressed as

$$H(\omega) = \sum_{m=-n}^{n} h(m) \exp(j\omega m) \qquad (1.4)$$

which can be written

$$H(\omega) = \sum_{m=0}^{n} a(m) \cos \omega m \qquad (1.5)$$

where, $a(0) = h(0)$

$a(m) = 2h(m)$, $m = 1, 2, \cdots n$.

Parks and McClellan [3] have shown that this representation of an FIR filter in terms of cosine basis functions is suitable for finding the Chebyshev approximation using the 2nd Remes Exchange Algorithm. For example, Fig. 1.2 shows an equiripple lowpass filter with passband and stopband deviations δ_p and δ_s and cutoff frequencies ω_p and ω_s.

A two-dimensional zero phase positive symmetric FIR filter of duration $(2n+1) \times (2n+1)$ is expressed as

$$H(\omega_1, \omega_2) = \sum_{\ell=-n}^{n} \sum_{k=-n}^{n} h(\ell, k) \exp(-j\omega_1 \ell) \exp(-j\omega_2 k) \tag{1.6}$$

which can be written

$$H(\omega_1, \omega_2) = \sum_{\ell=0}^{n} \sum_{k=0}^{n} a(\ell, k) \cos \ell\omega_1 \cos k\omega_2 \tag{1.7}$$

where

$$a(0,0) = h(0,0)$$

$$a(0,k) = 2h(0,k), \quad k = 1, 2 \cdots n$$

$$a(\ell,0) = 2h(\ell,0), \quad \ell = 1, 2 \cdots n$$

$$a(\ell,k) = 4h(\ell,k), \quad \ell = 1, 2 \cdots n, \quad k = 1, 2 \cdots n.$$

The above technique cannot be used in designing optimal two-dimensional filters since the cosine basis functions do not satisfy the Haar condition [6]. Thus, the 2nd Remes Exchange Algorithm does not apply. Hu and Rabiner [5] have interpreted the optimal design problem in terms of linear programming. The design of a 9 x 9 filter on an IBM 370 was extremely time consuming. Therefore alternate optimal design techniques are being sought [7], [8]. These techniques, faster than linear programming, are nonetheless slow and limit the size of the

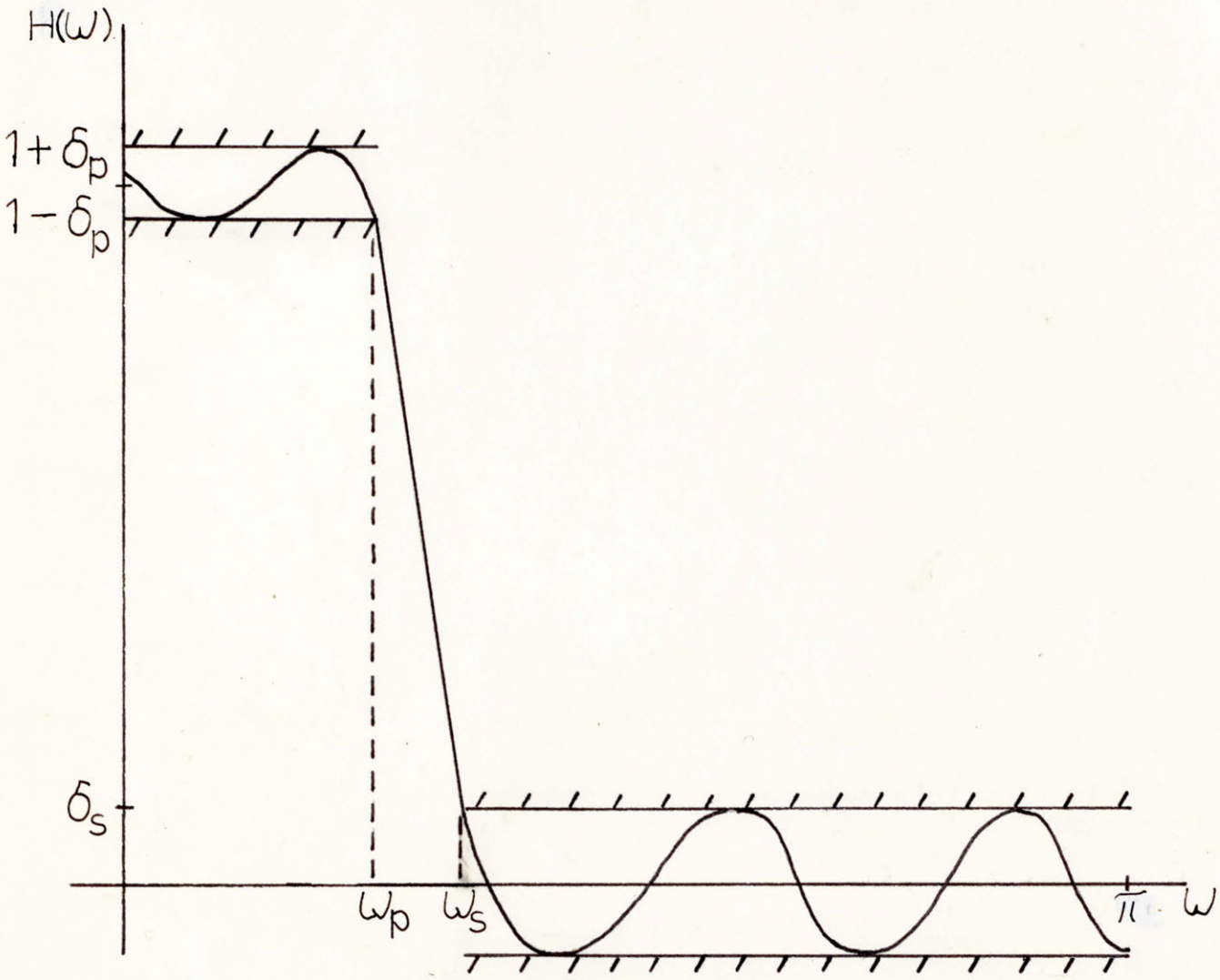

Fig. 1.2 An equiripple lowpass filter.

filter.

Hence, the McClellan transformation, a direct transformation of a one-dimensional design into a two-dimensional design, has become significant. Although sometimes suboptimum, the transformation preserves the equiripple nature of one-dimensional optimal filters. In addition, it is very fast and flexible in designing and implementing a large class of arbitrary frequency responses.

CHAPTER 2

McCLELLAN TRANSFORMATIONS

2.1 History

An alternative to the direct Chebyshev design was proposed by McClellan [1], [2]. This new approach, although generally suboptimal, is computationally and implementationally more efficient. The technique transforms a one-dimensional filter into a two-dimensional filter by a change of variables. The main advantage is its speed, since computing time is devoted almost entirely to the one-dimensional design which is known to be efficient. Hence, filters of very high order are readily obtained. Also the implementation involves on the order of $2n+1$ multiplies per output point, where $(2n+1) \times (2n+1)$ is the impulse response duration. This section presents a brief view of McClellan's original transformation.

The frequency response of a one-dimensional linear phase filter can be expressed as

$$H(\omega) = \sum_{m=0}^{n} a(m) \cos \omega m \, . \tag{2.1}$$

With $\omega = \cos^{-1} x$, $\cos(\omega m) = \cos(m \cos^{-1} x) = T_m(x)$, where $T_m(x)$ is the mth order Chebyshev polynomial. With the above change of variables, the frequency response can be expressed in x as

$$\begin{aligned} H(x) &= \sum_{m=0}^{n} a(m) \, T_m(x) \\ &= \sum_{m=0}^{n} b(m) \, x^m \, . \end{aligned} \tag{2.2}$$

Therefore,

$$H(\omega) = \sum_{m=0}^{n} b(m) \cos^m(\omega) \tag{2.3}$$

McClellan proposed the following change of variables:

$$\cos \omega = t(0,0) + t(1,0) \cos \omega_1 + t(0,1) \cos \omega_2 + t(1,1) \cos \omega_1 \cos \omega_2 \tag{2.4}$$

With the above substitution a two-dimensional zero-phase FIR filter results:

$$H(\omega_1, \omega_2) = \sum_{m=0}^{n} b(m) \left[\sum_{\ell=0}^{1} \sum_{k=0}^{1} t(\ell, k) \cos(\ell\omega_1) \cos(k\omega_2) \right]^n \tag{2.5}$$

$$= \sum_{p=0}^{n} \sum_{q=0}^{n} \hat{b}(p, q) \cos(p\omega_1) \cos(q\omega_2)$$

where

$$\hat{b}(0,0) = h(0,0) \tag{2.6a}$$

$$\hat{b}(0,q) = 2h(0,q), \quad q = 1, 2 \cdots n \tag{2.6b}$$

$$\hat{b}(p,0) = 2h(p,0), \quad p = 1, 2 \cdots n \tag{2.6c}$$

$$\hat{b}(p,q) = 4h(p,q), \quad q = 1, 2 \cdots n, \quad p = 1, 2 \cdots n \tag{2.6d}$$

An algorithm to perform this transformation is given in Appendix 1. The change of variables preserves the amplitude characteristic of the one-dimensional filter and thus results in an equiripple two-dimensional filter when the original filter is equiripple. The transformation can be viewed as a mapping where ω is mapped to a contour implicitly described by the transformation equation (2.4). This is best understood by studying two

examples due to McClellan.

The first example was designed for filtering unwanted velocity components from geophysical seismic traces. With a transformation determined by the coefficients $t(0,0) = t(1,1) = 0$ and $t(1,0) = -t(0,1) = 0.5$, the contours of Fig. 2.1 were obtained. The value associated with each contour is the one-dimensional frequency inverse image; that is, the frequency from which each contour is mapped. McClellan [2] began with an optimum Chebyshev one-dimensional lowpass filter. The passband and stopband are mapped to the two-dimensional space as shown in Fig. 2.2.

There are physical situations, as in picture processing, where there is no preferred spatial frequency axis. Thus there is a need for circularly symmetric filters. McClellan [2] used the transformation given by the coefficients $-t(0,0) = t(1,0) = t(0,1) = t(1,1) = 0.5$ to design such a filter. See Fig. 2.3 for the approximately circular contours. An example is the 41 x 41 filter of Fig. 2.4. This filter was transformed from a length 41 one-dimensional equiripple filter. Note the equiripple nature of the frequency response.

The next section will present a more formal analysis of McClellan's original transformation and its generalization.

2.2 Generalization to Higher Order Transformations

We shall now generalize the original McClellan transformation to higher orders, to increase the class of possible designs.

By definition, the Mth order transformation can be expressed as

$$\cos \omega = \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \cos(\ell\omega_1) \cos(k\omega_2) . \qquad (2.7)$$

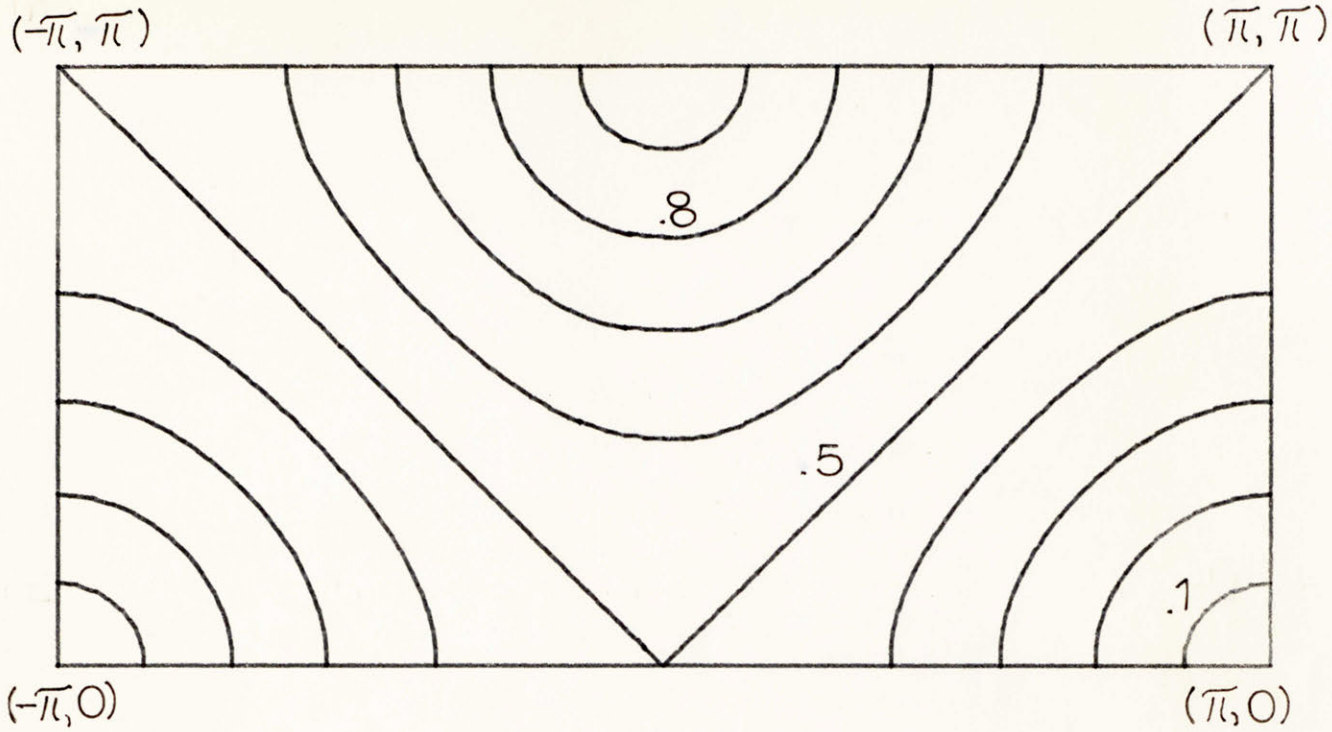

Fig. 2.1 The contours of the fan filter transformation.

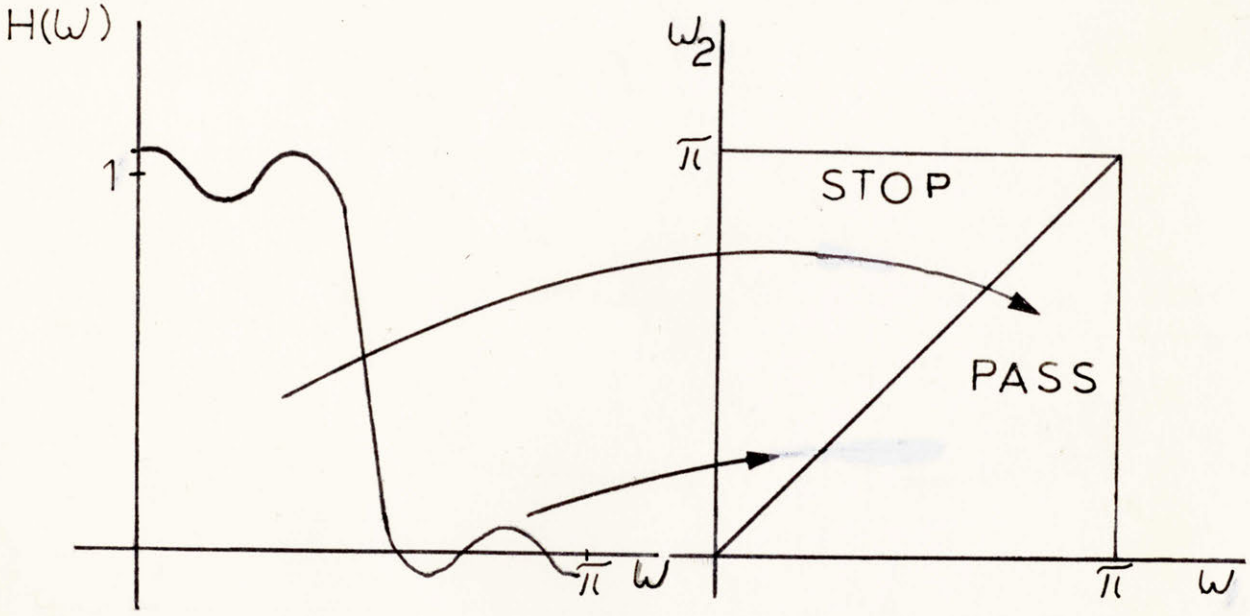

Fig. 2.2 The lowpass to fan filter mapping.

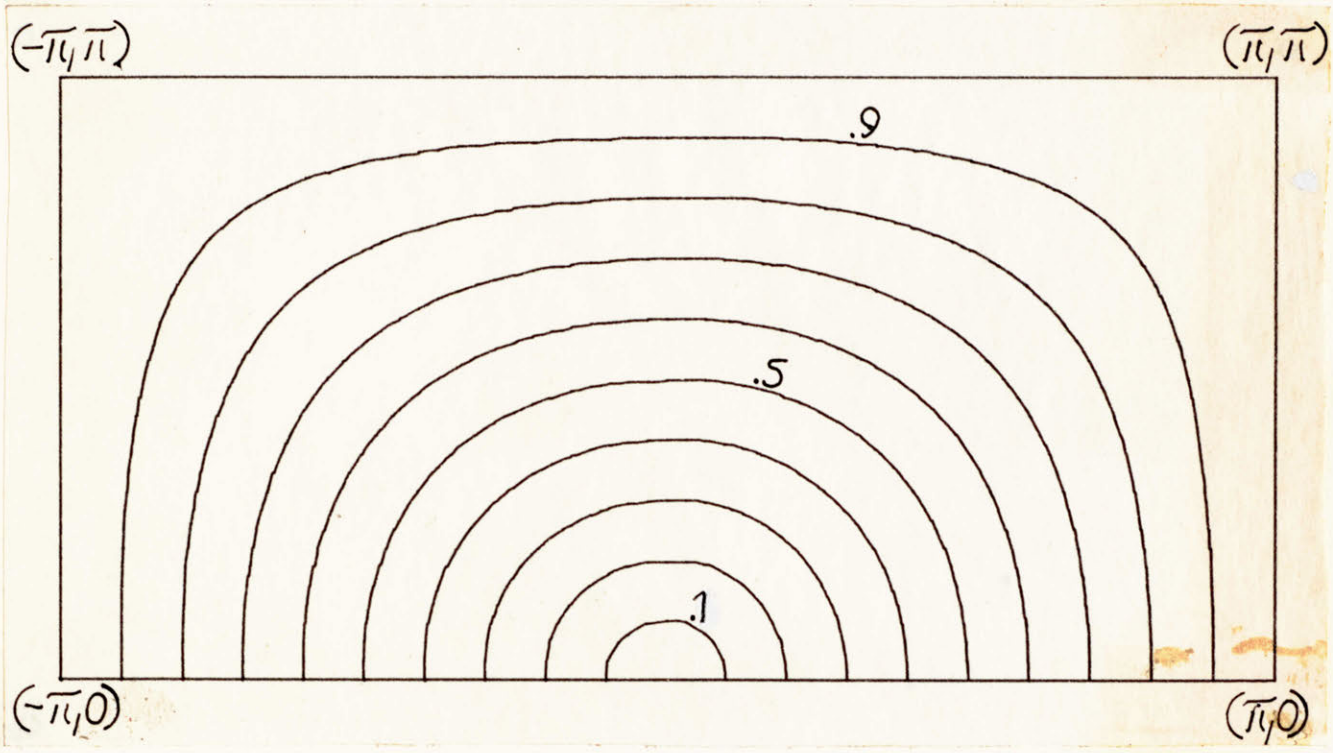

Fig. 2.3 The contours of the original McClellan transformation.

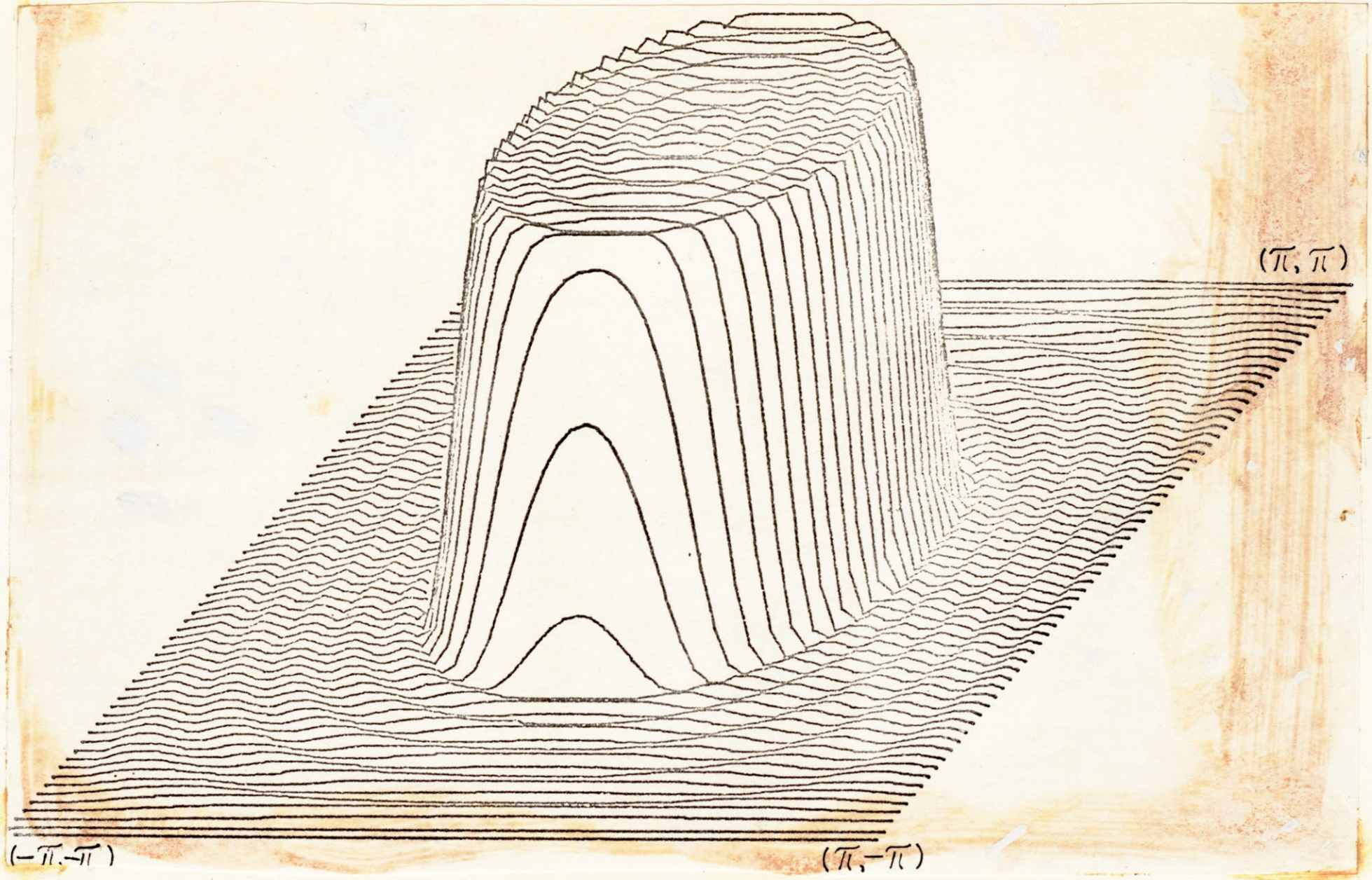

Fig. 2.4 A 41 x 41 two-dimensional filter designed from the original McClellan transformation.

For M = 1 this general expression simplifies to that in equation (2.4). Substituting for cos ω in the one-dimensional FIR filter of equation (2.3) results in

$$H(\omega_1, \omega_2) = \sum_{m=0}^{n} b(m) \left[\sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \cos(\ell \omega_1) \cos(k\omega_2) \right]^m . \qquad (2.8)$$

As with the first order case, this expression simplifies to that of a zero-phase two-dimensional FIR filter. Appendix 1 presents a generalization of a recursive algorithm developed by McClellan [2] for the first order case of section 2.1. The resulting filter is expressed as

$$H(\omega_1, \omega_2) = \sum_{p=0}^{nM} \sum_{q=0}^{nM} \hat{b}(p, q) \cos(p\omega_1) \cos(q\omega_2) \qquad (2.9)$$

where the order of the filter is $(2nM+1) \times (2nM+1)$ and where the relation between the b(p, q) and the impulse response coefficients h(p, q) is described by equations (2.6a, b, c, d). Note that the two-dimensional order depends on both the one-dimensional impulse response length and the order of the transformation.

We can view the Mth order transformation as a three-step mapping, demonstrated in Fig. 2.5. Let $\omega = \cos^{-1} x$, $\omega_1 = \cos^{-1} u$, and $\omega_2 = \cos^{-1} v$. Then, mapping T_1 of Fig. 2.4 maps the interval $[0, \pi]$ to $[-1, 1]$, where segment A maps to segment D, and segment B maps to segment C. Substituting for ω, ω_1, and ω_2 leads to the expression

$$\begin{aligned} x &= \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) R_\ell(u) S_k(v) \\ &= \sum_{\ell=0}^{M} \sum_{k=0}^{M} q(\ell, k) u^\ell v^k \end{aligned} \qquad (2.10)$$

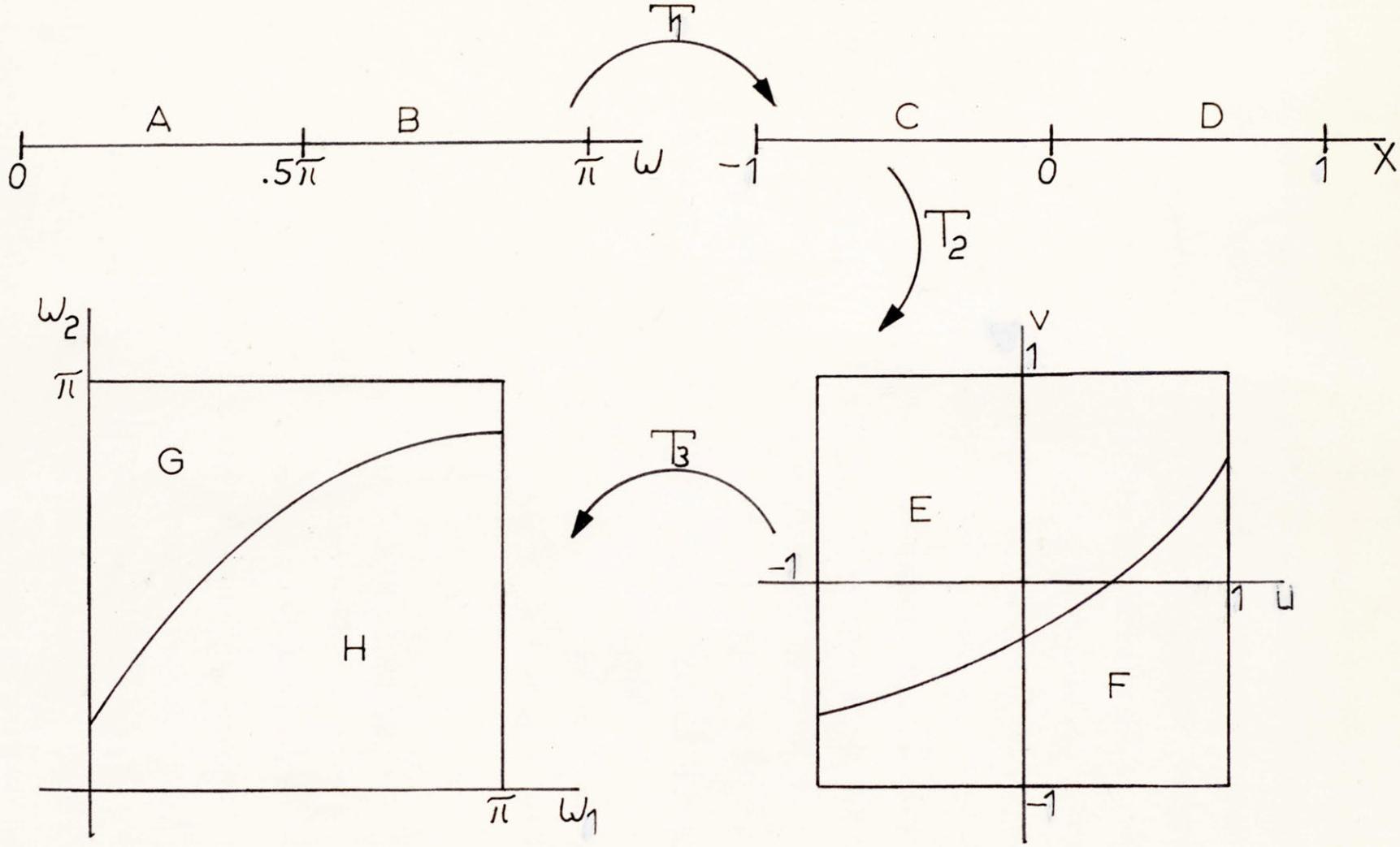

Fig. 2.5 The three-step mapping of the transformation.

where $R_\ell(u)$ and $S_k(v)$ are the Chebyshev polynomials for the variables u and v. Thus a particular value of x maps to an Mth order polynomial which is implicitly described by the transformation equation in the u, v space. This is indicated in Fig. 2.5 by T_2 which maps segment C to region E and segment D to region F in the u, v space. Regions E and F depend on the choice of transformation coefficients $t(\ell, k)$. Finally, since $H(\omega_1, \omega_2)$ is positive symmetric in both ω_1 and ω_2, it is necessary to consider only the mapping between the u, v space and the first quadrant of the ω_1, ω_2 space bordered by π. With $u = \cos \omega_1$ and $v = \cos \omega_2$, T_3 maps region E into region H and maps region F into region G. We assume the mapping is well-defined, a concept to be discussed in Chapter 3. Essentially, a well-defined mapping is one which guarantees that only points from the interval $[0, \pi]$ are mapped to the two-dimensional region of interest. Note that the general Mth order mapping preserves the amplitude characteristic of the one-dimensional filter.

The idea of viewing the mapping as a three-step process, the second of which maps x to a polynomial in the u, v space, will be of great interest in later chapters.

2.3 The Approximation Problem in Contour Design

We have seen that the transformation maps a particular ω into a curve implicitly described by the transformation equation. The goal of this section is to formulate the general approximation problem in terms of mapping an ω to some desired continuous curve $\omega_2 = g(\omega_1)$ in the ω_1, ω_2 space. The approximation problem is then made linear, avoiding the great difficulty in confronting the nonlinear version. That is, ω_2 can be written in terms of ω_1 by manipulating the transformation equation. However, as we see from equation (2.7), this involves writing ω_2 in terms

of nonlinear functions such as cosines and arcosines. Hence a very complicated nonlinear problem develops.

Zero error is desired between the ideal curve $\omega_2 = g(\omega_1)$ and the contour from which ω is mapped. Since this contour will not necessarily follow the ideal curve exactly, an error can be found in traversing $g(\omega_1)$ expressed as

$$e(\omega_1) = \cos \omega - \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \cos (\ell\omega_1) \cos (kg(\omega_1)) \tag{2.11}$$

$$= \cos \omega - \sum_{\ell=0}^{M^2} \hat{t}(\ell)\, \varphi_\ell(\omega_1)$$

where $\varphi_\ell(\omega_1) = \cos (\ell\omega_1) \cos (kg(\omega_1))$. The objective is to find a best approximation to $\cos \omega$ over a given interval in ω_1 by choosing the appropriate $\hat{t}(\ell)$, the free coefficients associated with the basis functions $\varphi_\ell(\omega_1)$. The basis functions and the ideal function, $\cos \omega$, will change when constraints are placed on the coefficients, as demonstrated later by examples.

A slight modification of the problem can be made by allowing $x = \cos \omega$ itself to be a free variable. The minimization procedure will find not only the optimum coefficients $\hat{t}(\ell)$, but also the best frequency mapping to approximate $g(\omega_1)$.

Nevertheless, the problem can always be viewed as a linear approximation problem where an ideal function $H_d(\omega_1)$ is to be approximated by a set of linearly independent basis functions. Two possible error criteria are the Chebyshev (minimax) or least squares. The Chebyshev is desirable when approximating discontinuous ideal curves.

The approximation problem in this case can be formulated as

$$\text{minimize} \qquad \max |e(\omega_1)| = \max \left| H_d(\omega_1) - \sum_{\ell=0}^{L} \hat{t}(\ell)\varphi_\ell(\omega_1) \right|. \quad (2.12)$$

A very fast and efficient algorithm developed by Hersey and Mersereau [9] is used in the minimax approximation procedure. The set of basis functions, derived from constraint considerations, in general does not satisfy the Haar condition and so the 2nd Remes Exchange Algorithm does not apply. Hersey and Mersereau have modified an ascent algorithm by Cheney [10]to perform a multiple rather than a single exchange of error deviation, thus quickening the algorithm. This works in the absence of the Haar condition and so is applicable to the given set of basis functions.

The least squares approximation, being faster than the above algorithm (only one iteration is required), is desirable for smooth ideal curves. The approximation problem can be formulated as

$$\text{minimize} \qquad \sum_{\Omega} e^2(\omega_1) = \sum_{\Omega} \left[H_d(\omega_1) - \sum_{\ell=0}^{L} \hat{t}(\ell)\, \varphi_\ell(\omega_1) \right]^2 \qquad (2.13)$$

where the summation is over discrete values of ω_1 in the interval (Ω) of interest. The least squares solution procedure is well known.

Some examples with the first order transformation follow. They demonstrate the above techniques and the effect of constraints on the approximation problem. The ideal function is $g(\omega_1) = (R_p^2 - \omega_1^2)^{\frac{1}{2}}$ over the interval $[0, R_p]$ where R_p is the desired passband edge radius (see Fig. 1.1) for a circularly symmetric lowpass filter.

Suppose no constraints are applied to the coefficients. Then the

minimax approximation problem can be expressed as

$$\text{minimize} \qquad \max|e(\omega_1)| = \max\left|\cos\omega - \sum_{\ell=0}^{1}\sum_{k=0}^{1} t(\ell,k)\cos(\ell\omega_1)\cos(k\omega_2)\right| \tag{2.14}$$

where $\omega = \omega_p$, the passband edge of the transformed one-dimensional lowpass filter. Since $\cos\omega = \cos\omega_p$ is a constant, the obvious trivial solution with zero error is

$$t(0,0) = \cos\omega_p \tag{2.15a}$$

$$t(1,0) = t(0,1) = t(1,1) = 0\,. \tag{2.15b}$$

Hence, the passband edge in one dimension is mapped to the entire two-dimensional region $(-\pi,\pi) \times (-\pi,\pi)$, as well as the ideal curve $g(\omega_1)$. Some set of constraints, therefore, must be applied. A reasonable set of constraints for a circularly symmetric lowpass filter is the following

$$0 \rightarrow (0,0) \tag{2.16a}$$

$$\pi \rightarrow (\omega_1,\pi) \tag{2.16b}$$

$$H(\omega_1,\omega_2) = H(\omega_2,\omega_1)\,. \tag{2.16c}$$

These constraints guarantee that the origin in one dimension maps to the origin in two dimensions and that π maps to the outer horizontal boundary $\omega_2 = \pi$. Constraint 2.16c implies octagonal symmetry and so π also maps to the outer vertical boundary $\omega_1 = \pi$. Also, constraining octagonal symmetry results in $t(\ell,k) = t(k,\ell)$, reducing the number of degrees of freedom in the approximation. This constraint saves

computational time.

The first constraint (2.16a) generates the constraint equation

$$1 = t(0,0) + t(1,0) + t(0,1) + t(1,1) . \tag{2.17}$$

(2.16b) implies that

$$-1 = t(0,0) - t(1,0)\cos\omega_1 - t(0,1) - t(1,1)\cos\omega_1 \tag{2.18}$$

or

$$-1 = t(0,0) - t(0,1) \tag{2.19a}$$

$$0 = t(1,0) + t(1,1) . \tag{2.19b}$$

Finally, (2.16c) yields the equation

$$t(1,0) = t(0,1). \tag{2.20}$$

This results in four equations with four unknowns. Solving,

$$-t(0,0) = t(1,0) = t(0,1) = t(1,1) = 0.5$$

Hence, this set of constraints gives McClellan's original transformation.

Another reasonable set of constraints is

$$0 \rightarrow (0,0) \tag{2.21a}$$

$$\pi \rightarrow (\pi,\pi) \tag{2.21b}$$

$$H(\omega_1,\omega_2) = H(\omega_2,\omega_1). \tag{2.21c}$$

These constraints imply the equations

$$1 = t(0,0) + t(1,0) + t(0,1) + t(1,1) \tag{2.22a}$$

$$-1 = t(0,0) - t(1,0) - t(0,1) + t(1,1) \tag{2.22b}$$

$$t(1,0) = t(0,1) . \tag{2.22c}$$

We see that one free variable remains.

The approximation problem is formulated in the following way. Solving in terms of t(0, 0),

$$t(1,0) = t(0,1) = 0.5 \tag{2.23a}$$

$$t(1,1) = -t(0,0) . \tag{2.23b}$$

Therefore,

$$\begin{aligned} \cos\omega &= t(0,0) + 0.5(\cos\omega_1 + \cos\omega_2) - t(0,0)\cos\omega_1\cos\omega_2 \\ &= 0.5(\cos\omega_1 + \cos\omega_2) + t(0,0)(1 - \cos\omega_1\cos\omega_2) . \end{aligned} \tag{2.24}$$

With $\cos\omega = x$ also a free variable, we can express the approximation error as

$$e(\omega_1) = H_d(\omega_1) + t(0,0)\,\varphi_1(\omega_1) + x\,\varphi_2(\omega_1) \tag{2.25}$$

where

$$H_d(\omega_1) = 0.5(\cos\omega_1 + \cos\omega_2) \tag{2.26a}$$

$$\varphi_1(\omega_1) = 1 - \cos\omega_1\cos\omega_2 \tag{2.26b}$$

$$\varphi_2(\omega_1) = 1 \tag{2.26c}$$

and where

$$\omega_2 = (R_p^2 - \omega_1^2)^{\frac{1}{2}} \tag{2.26d}$$

The minimax error criterion with $R_p = 0.8\pi$ resulted in $t(0,0) = 0.3529$, $\omega = \cos^{-1} x = 0.685\pi$ and nearly perfect circularity. The contours are shown in Fig. 2.6.

The least squares error criterion with $R_p = 0.8\pi$ resulted in $t(0,0) = -0.3531$, $\omega = \cos^{-1} x = 0.683\pi$, and again almost perfect circularity as shown in Fig. 2.7. We see that the contours of Fig. 2.6 and Fig. 2.7 are nearly identical.

These examples have involved equality constraints. If inequality constraints are imposed, then linear programming with the minimax error criterion can be used in the approximation problem. Linear programming, a single exchange algorithm, is much slower than that of Hersey and Mersereau, but has the advantage of handling inequality constraints. This technique will be demonstrated in Chapter 4 for the design of fan filters.

Lastly, we note that the minimax error $e(\omega_1)$ does not necessarily lead to an equiripple contour about the ideal $g(\omega_1)$. The error $e(\omega_1)$ is not the actual error deviation of the contour from the specified curve, but rather the error between $H_d(\omega_1)$ and the set of basis functions. The actual contour deviation can be measured in the following way. Letting ω, the frequency associated with each contour, vary so that the error is zero, we write

$$\cos\omega - \sum_{\ell=0}^{M^2} \hat{t}(\ell)\, \varphi_\ell(\omega_1) = 0 . \tag{2.27}$$

Therefore,

$$\omega = \arccos\left[\sum_{\ell=0}^{M^2} \hat{t}(\ell)\, \varphi_\ell(\omega_1)\right] . \tag{2.28}$$

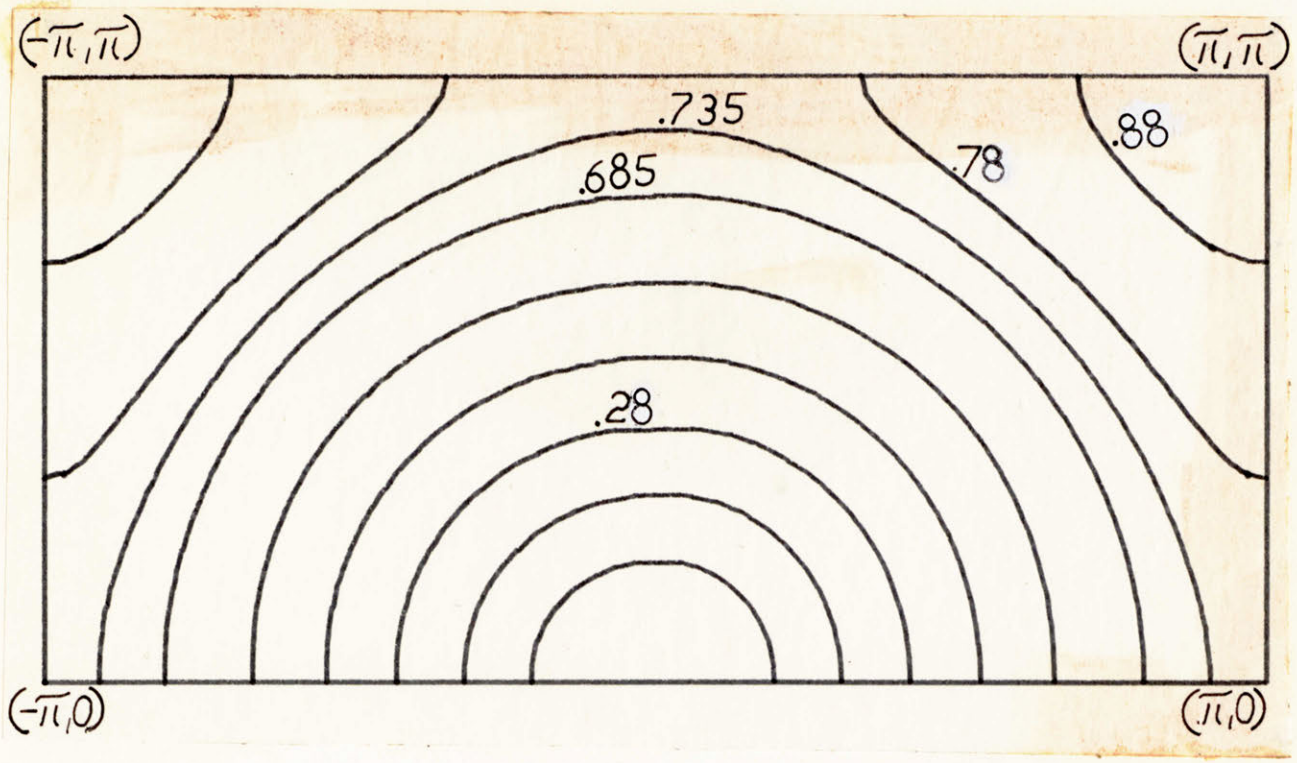

Fig. 2.6 The contours of the first order transformation with cos ω free using the minimax criterion for the ideal radius 0.8 π.

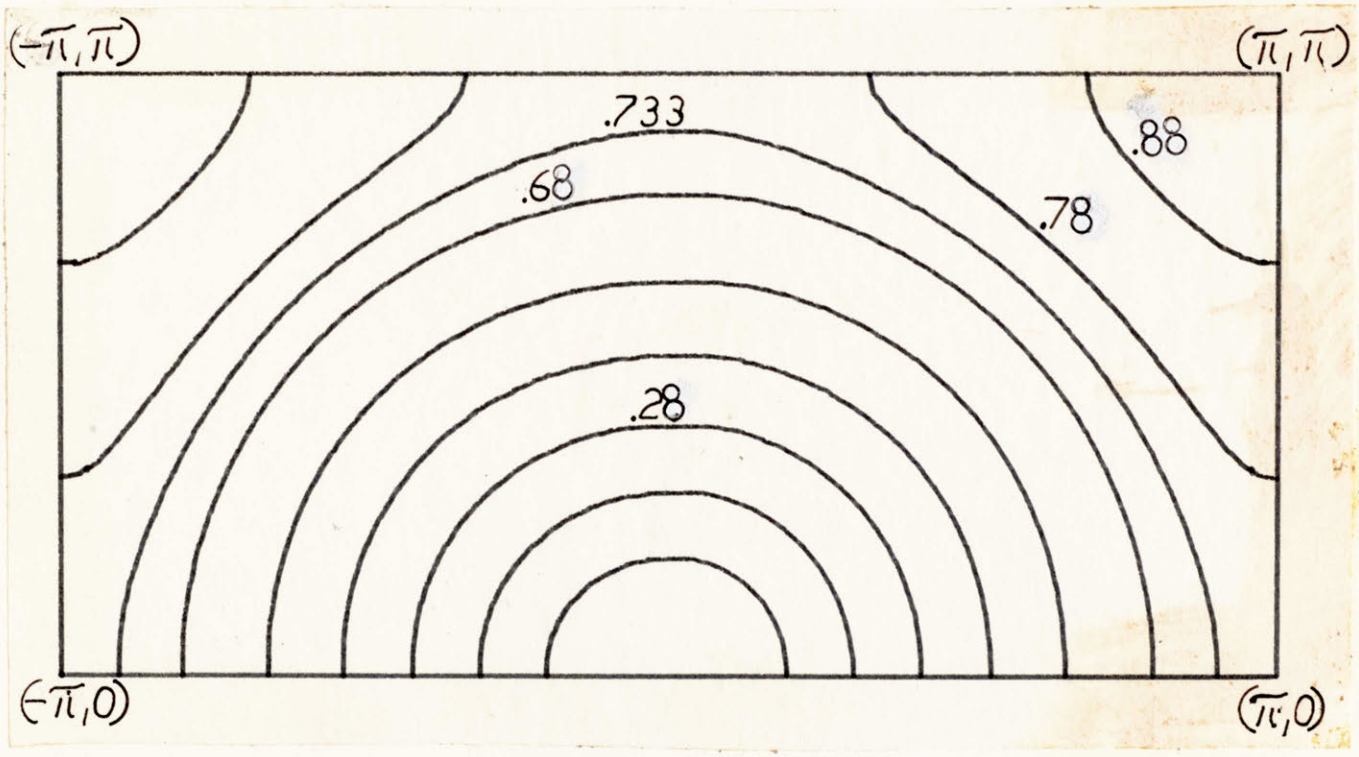

Fig. 2.7 The contours of the first order transformation with cos ω free using the least squares criterion for the ideal radius 0.8 π.

From equation (2.11),

$$\omega = \arccos [\cos \omega_o - e(\omega_1)] \tag{2.29}$$

where ω_o is the frequency which is mapped to the ideal curve. Traversing the curve $g(\omega_1)$ will yield the actual ω (given by 2.28, 2.29), which is mapped to the ideal curve as a function of ω_1. The frequency deviation can be written

$$\Delta\omega = \omega(\omega_1) - \omega_o . \tag{2.30}$$

The effects of this deviation are clearly seen by an example. Suppose the passband edge $\omega_o = \omega_p$ of a one-dimensional lowpass filter is mapped to an ideal curve using the minimax approximation procedure. $\Delta\omega$ (2.30) is a measure of the variation in frequency (about ω_p) which maps to the ideal curve $g(\omega_1)$. Figure 2.8 shows that a frequency response error of one is possible for a large $\Delta\omega$.

Returning to the general case, we let $z(\omega_1) = \cos \omega_o - e(\omega_1)$. Figure 2.9 shows graphically the function $\omega = \arccos (z(\omega_1))$. We see that the variation in ω is more sensitive to error deviations in $e(\omega_1)$ for mappings where ω_o is close to zero or π.

2.4 Implementation

Conventional techniques in implementing two-dimensional FIR filters involve either direct convolution or the FFT. Briefly presented in this section is a novel implementation, taking advantage of the unique properties of the McClellan transformation. The technique was recently developed by Mecklenbräuker and Mersereau.

The z transform of a zero phase FIR filter is defined as

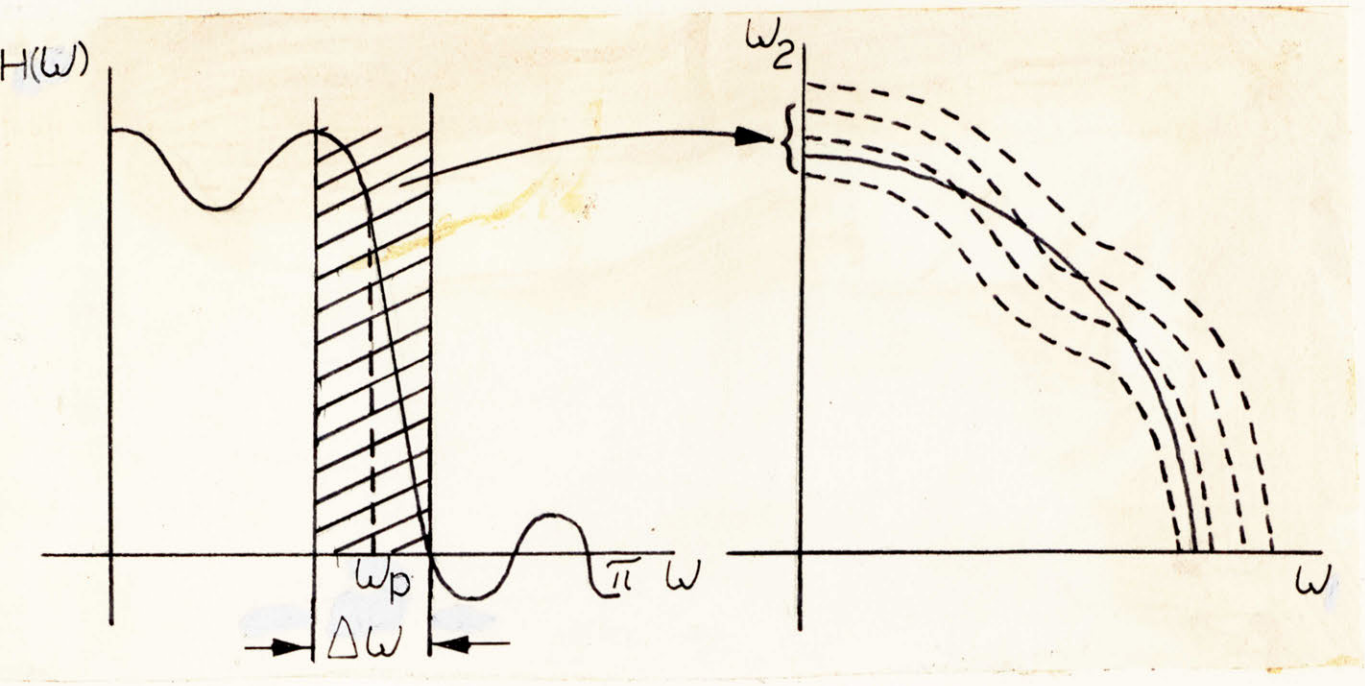

Fig. 2.8 The frequency variation about the ω_p mapping.

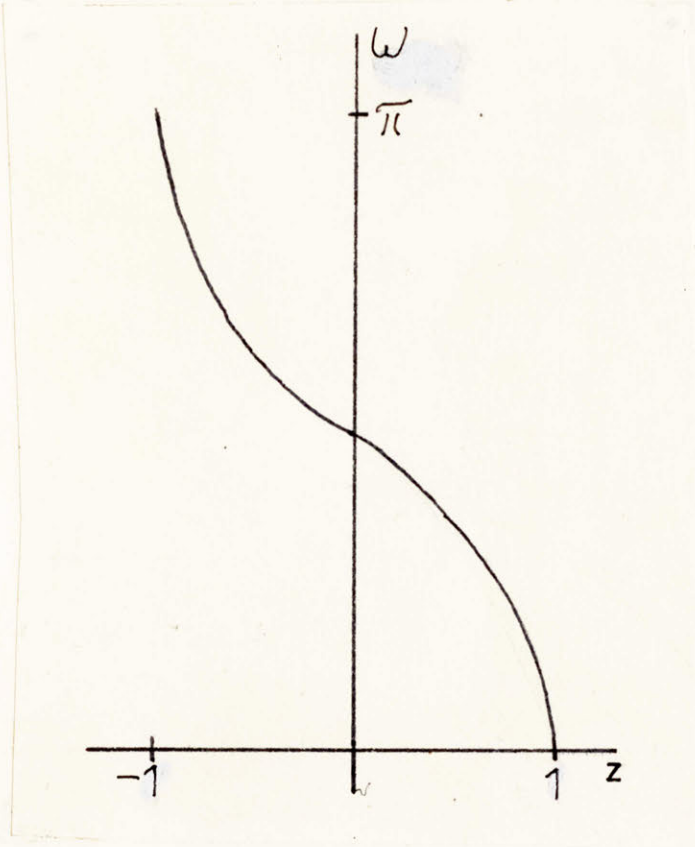

Fig. 2.9 The circularity deviation measured by the change in ω.

$$H(z) = \sum_{m=-n}^{n} h(m)\, z^{-m}$$

$$= \sum_{m=0}^{n} b(m) \left[\frac{z + z^{-1}}{2}\right]^m \tag{2.31}$$

where $z = \exp(j\omega)$. Since

$$\left[\frac{z + z^{-1}}{2}\right] = \sum_{\ell=0}^{M} \sum_{k=0}^{M} q(\ell, k)\; z_1^{\ell}\, z_2^{k} \triangleq F(z_1, z_2) \tag{2.32}$$

where $z_1 = \exp(j\omega_1)$ and $z_2 = \exp(j\omega_2)$, substituting in equation (2.31),

$$H(z_1, z_2) = \sum_{m=0}^{n} b(m)\, [F(z_1, z_2)]^m . \tag{2.33}$$

Equation (2.33) describes the digital structure of Fig. 2.9.

For the first order mapping

$$F(z_1, z_2) = t(0,0) + t(1,0)\left[\frac{z_1 + z_1^{-1}}{2}\right] + t(0,1)\left[\frac{z_2 + z_2^{-1}}{2}\right] + t(1,1)\left[\frac{z_1 + z_1^{-1}}{2}\right]\left[\frac{z_2 + z_2^{-1}}{2}\right]. \tag{2.34}$$

Therefore $f(n_1, n_2)$, the impulse response of $F[z_1, z_2]$, can be represented by the nonrecursive 3 x 3 filter of Fig. 2.11. Due to the symmetry of $f(n_1, n_2)$, implementation of each stage of Fig. 2.10 requires in general only four multiplies for each output point. Since there are n stages, each output point requires 5n + 1 total multiplies (n + 1 multiplies to implement b(m)). This can be reduced with the introduction of equality constraints. For example, the transformation of section 2.3

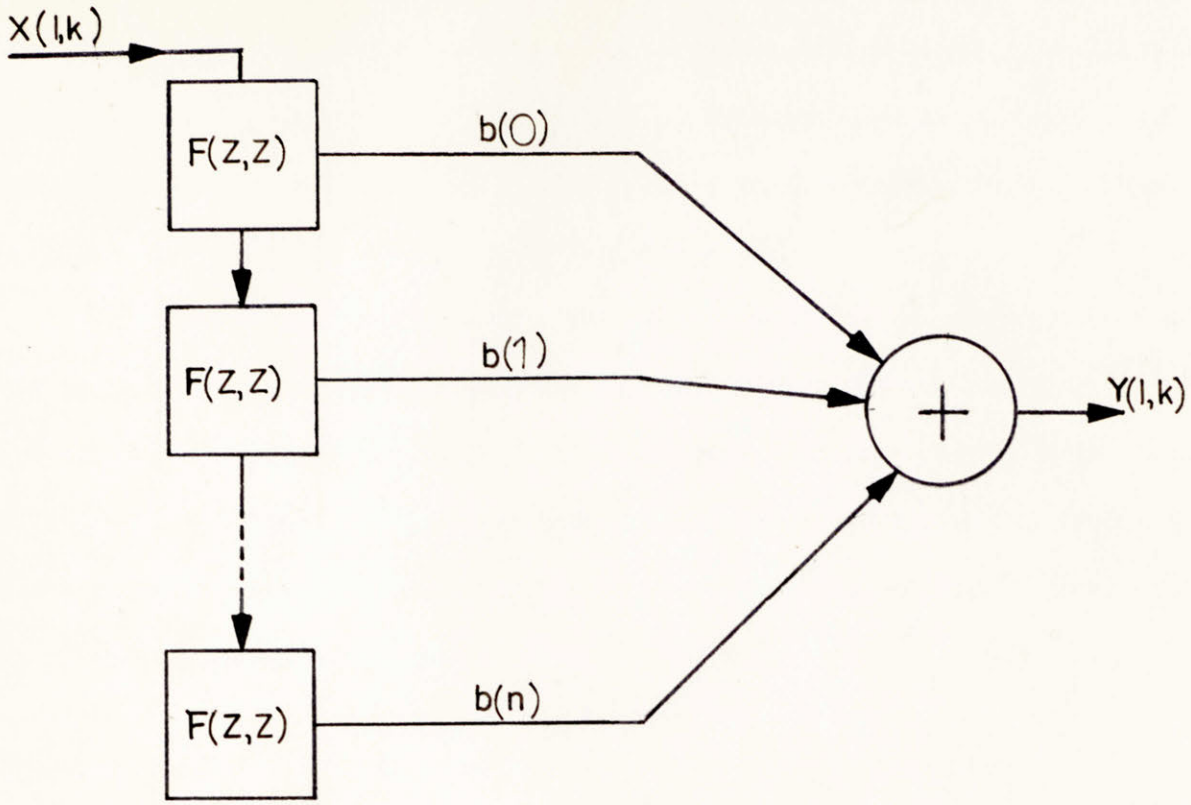

Fig. 2.10 The digital filter implementation of the first order transformation.

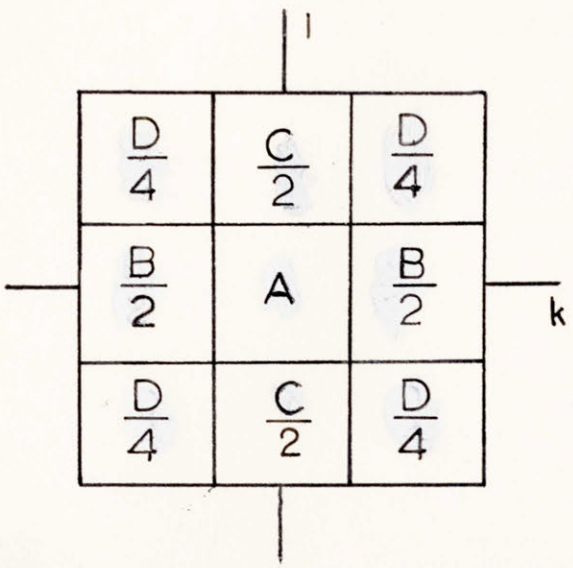

Fig. 2.11 The 3 x 3 nonrecursive filter for the first order transformation implementation.

with $x = \cos \omega$ a free variable, requires only 1 multiply per output point for each stage. The reduction occurs because $t(1, 0) = t(0, 1) = 0.5$ and $t(1, 1) = -t(0, 0)$, and so multiplies can be avoided with the proper hardware implementation (logical shifts, complements, etc.). This implies $2n + 1$ total multiplies per output point.

In general, for any order transformation the number of multiplies for each output point is on the order of kn, where k is a constant dependent on the number of degrees of freedom and the constraints in the transformation. This is a significant reduction from conventional techniques* for a certain range of input and impulse response duration. Research is presently going on to quantify these results.

---

*Direct convolution requires about $(2n + 1)^2$ multiplies per output point. The FFT requires greater than kn multiplies up to approximately $n = 21$ for the first order transformation.

CHAPTER 3

THE CONSTRAINT PROBLEM WITH APPLICATIONS TO CIRCULARLY SYMMETRIC CONTOURS

3.1 Constraints for a Well-Defined Mapping

A number of constraints are derived for guaranteeing a well-defined mapping, a condition which prevents the mapping of points from off the unit circle in the z-plane, into the two-dimensional region $[-\pi, \pi] \times [-\pi, \pi]$. Also, we present conditions for controlling the general shape of contours; that is, we derive constraints to insure contours of increasing or decreasing monotonicity of a given shape.

The definition of a well-defined mapping can be developed by expressing the one-dimensional frequency response as a function of $x = \cos \omega$:

$$H(x) = \sum_{m=0}^{n} a(m)\, T_m(x) \qquad (3.1)$$

Figure 3.1 shows a typical lowpass filter as a sum of Chebyshev polynomials. We see that outside the region $|x| \leq 1$ the polynomial $H(x)$ becomes quite large in magnitude, a characteristic of Chebyshev polynomials. An alternate viewpoint is that the region $|x| > 1$ corresponds to points in the z plane off the unit circle. This implies that the frequency response of points off the unit circle maps onto the two-dimensional unit surface. The problem becomes more serious as the order of the one-dimensional filter increases, because higher order Chebyshev polynomials are involved. Therefore, there appears to be no control over the magnitude of $H(x)$ for $|x| > 1$. We see from the three-step mapping of section 2.2 that points outside the interval $|x| \leq 1$

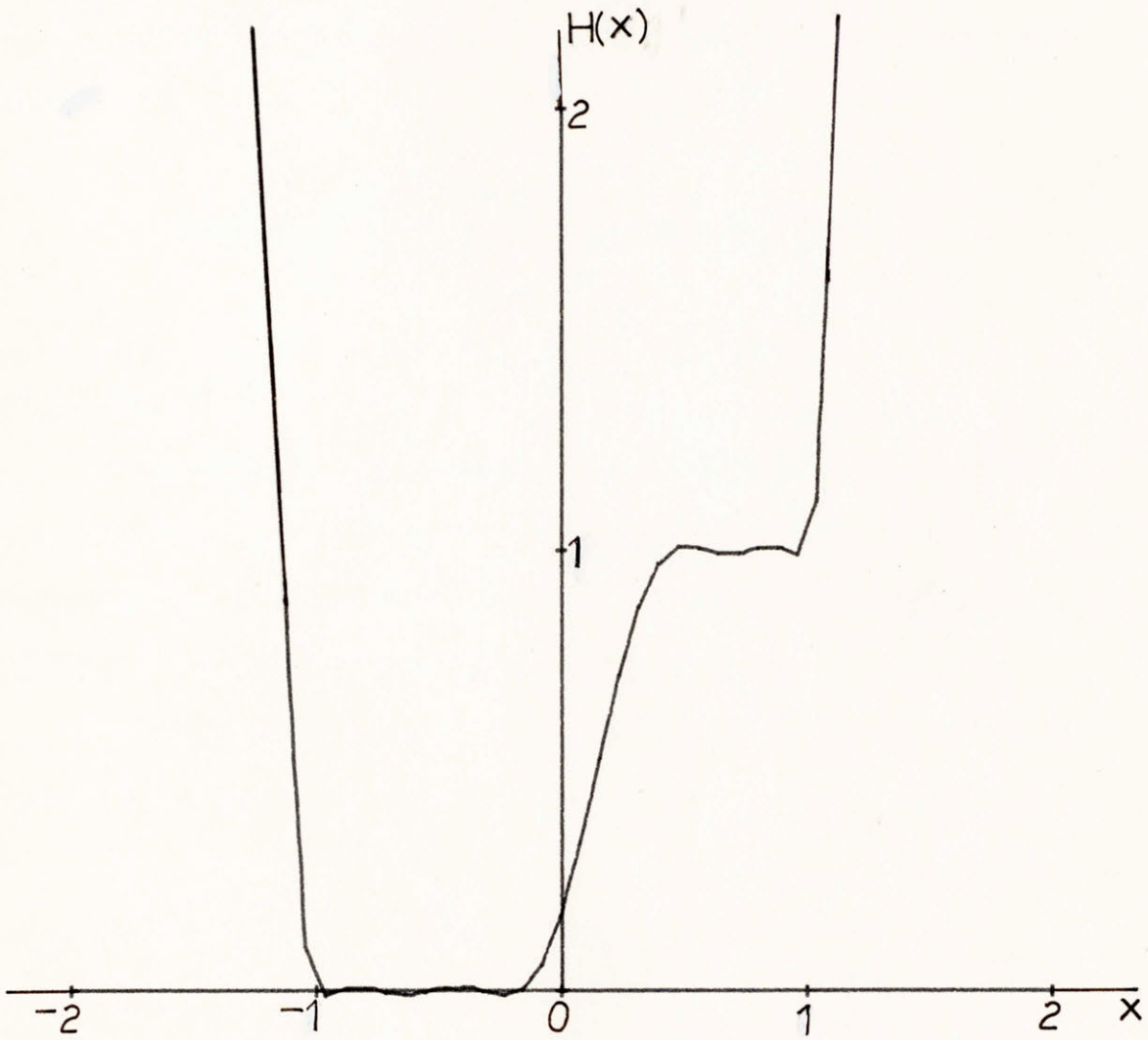

Fig. 3.1 A lowpass filter expressed as a sum of Chebyshev polynomials.

can map to the u, v region (-1, 1) x (-1, 1). Therefore, the constraint

$$|x| = \left| \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k)\, R_{\ell}(u)\, S_{k}(v) \right| \leq 1 \qquad (3.2)$$

over (-1, 1) x (-1, 1) must be made on the transformation coefficients $t(\ell, k)$ to avoid an ill-defined mapping. This implies the equivalent constraint

$$|\cos \omega| = \left| \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k)\ \cos(\ell \omega_1) \cos(k\omega_2) \right| \leq 1 \qquad (3.3)$$

over $(0, \pi)$ x $(0, \pi)$.

One possible way to implement this constraint is to force $|x| \leq 1$ on a densely spaced grid over (-1, 1) x (-1, 1). This requires a very large number of inequality constraints on $t(\ell, k)$ and so also the tedious and time consuming use of linear programming.

A more elegant approach is the following. Consider the first order transformation,

$$\cos \omega = t(0,0) + t(1,0) \cos \omega_1 + t(0,1) \cos \omega_2 + t(1,1) \cos \omega_1 \cos \omega_2$$

which can be written

$$x = t(0,0) + t(1,0)u + t(0,1)v + t(1,1)uv\ . \qquad (3.4)$$

Differentiating,

$$\frac{\partial x}{\partial u} = t(1,0) + t(1,1)v \tag{3.5a}$$

$$\frac{\partial x}{\partial v} = t(0,1) + t(1,1)u \,. \tag{3.5b}$$

Along the ω_1, ω_2 axes,

$$\frac{\partial x}{\partial u} = t(1,0) + t(1,1) \tag{3.6a}$$

$$\frac{\partial x}{\partial v} = t(0,1) + t(1,1) \,. \tag{3.6b}$$

Imposing the constraint $0 \rightarrow (0,0)$ and octagonal symmetry on $H(\omega_1, \omega_2)$, we see from Fig. 3.2 that the mapping remains well defined for $t(1,0) + t(1,1) \leq 1$ along the lines $v = 1$ and $u = 1$. If the additional constraint $\pi \rightarrow (\pi, \pi)$ is applied with the above equality constraints and inequality constraint, $t(1,0) + t(1,1) \leq 1$, then a well-defined mapping is guaranteed over the entire boundary of the region $(-1,1) \times (-1,1)$. Because of the linear nature of the mapping, a well-defined mapping on the boundary implies a well-defined mapping within the boundary.

Another problem of interest is contour monotonicity. For any order transformation, we can show that

$$\frac{d\omega_2}{d\omega_1} = \frac{dv}{du}\left(\frac{1 - u^2}{1 - v^2}\right)^{\frac{1}{2}}\Bigg|_{\substack{u = \cos\omega_1 \\ v = \cos\omega_2}} \,. \tag{3.7}$$

Thus $\frac{d\omega_2}{d\omega_1}$ and $\frac{dv}{du}$ have the same sign over the region $(0,\pi) \times (0,\pi)$. In particular it can be shown that for the first order transformation

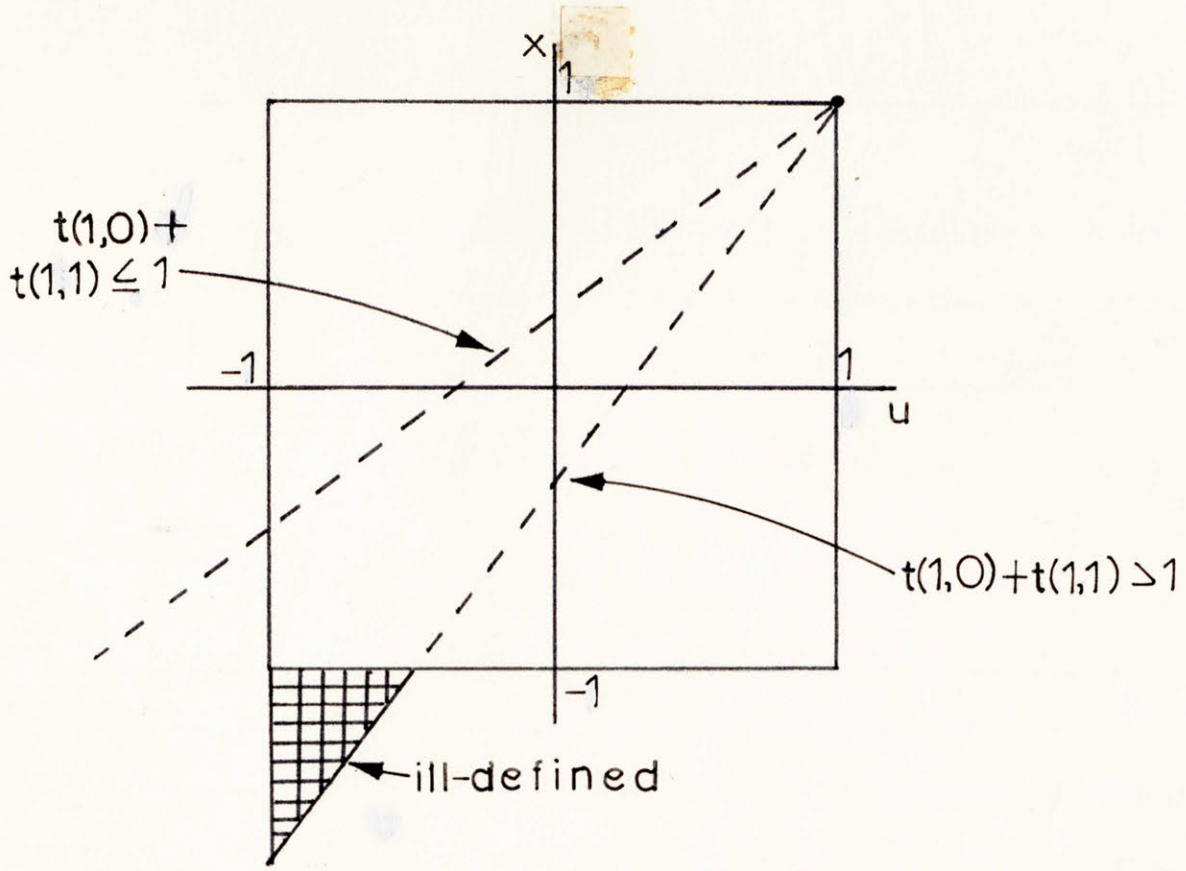

Fig. 3.2 The first order transformation along the u axis under the constraint $0 \rightarrow (0, 0)$.

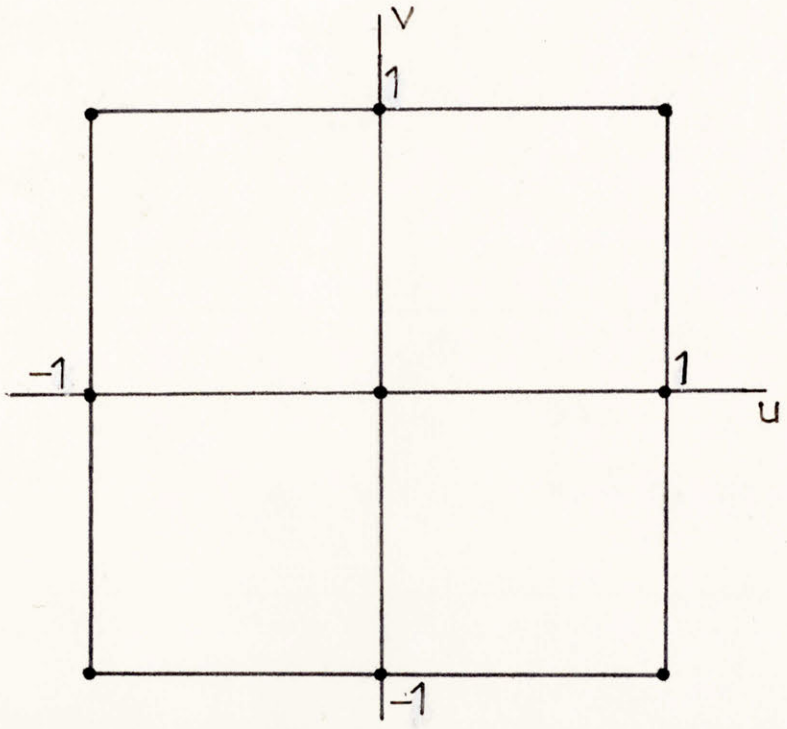

Fig. 3.3 A sparse grid over which increasing monotonicity and a well-defined mapping is guaranteed with the use of linear programming.

$$\frac{dv}{du} = \frac{t(0,0)\,t(1,1) - t(1,0)\,t(0,1) - t(1,1)x}{(t(0,1) + t(1,1)u)^2} \tag{3.8}$$

and thus v is either a monotonically increasing or decreasing function of u for a given x [1]. Furthermore, we can show increasing monotonicity for

$$x < \frac{[t(0,0)\,t(1,1) - t(1,0)\,t(0,1)]}{t(1,1)} \tag{3.9}$$

and decreasing monotonicity otherwise. For example, McClellan's original mapping coefficients are such that $[t(0,0)\,t(1,1) - t(1,0)\,t(0,1)]/t(1,1) = -1$. Since x in the region $|x| \leq 1$ is always greater than -1, all contours are thus guaranteed to be monotonically decreasing.

A set of constraints for a well-defined mapping can also be derived for the second order transformation, which is expressed as

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(\ell, k)\, u^{\ell} v^{k} . \tag{3.10}$$

Differentiating,

$$\frac{\partial x}{\partial u} = \sum_{\ell=1}^{2} \sum_{k=0}^{2} q(\ell, k)\, v^{k} (\ell u^{\ell-1}) . \tag{3.11}$$

Hence $\frac{\partial x}{\partial u}$ is linear in u and quadratic in v. Imposing the constraints,*

$$0 \rightarrow (0,0) \tag{3.12a}$$

---

* Compatible with circularly symmetric lowpass filters.

$$\pi \rightarrow (\omega_1, \pi) \tag{3.12b}$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1) \tag{3.12c}$$

a well-defined mapping is guaranteed when

$$\frac{\partial x}{\partial u} \geq 0 \tag{3.13a}$$

$$\frac{\partial^2 x}{\partial u \partial v} \geq 0 \tag{3.13b}$$

at the points (-1, -1), (1, 1), (-1, 1), and (1, -1) (see Appendix 2). These inequality constraints are naturally satisfied in designing certain circularly symmetric contours as shown in a later section. Thus linear programming can be avoided. Also we can show that the above constraints imply decreasingly monotonic contours and that $\frac{\partial x}{\partial u} \geq 0$ along any horizontal line in the region (-1, 1) x (-1, 1). Thus the constraints are compatible with the design of lowpass filters.

Similarly a set of constraints can be obtained for a well-defined mapping with increasingly monotonic contours, a necessity in designing fan filters, for example. Because of the lack of symmetry and few equality constraints ($0 \rightarrow (\pi, 0)$, $\pi \rightarrow (0, \pi)$), these constraints are rather complicated. Hence, a convenient sparse grid was used where $|x| \leq 1$ and where increasing monotonicity was guaranteed. See Fig. 3.3. Chapter 4 discusses this in greater depth with the use of linear programming.

As a final note, a recent development by Mersereau and Mecklenbräuker has led to a very efficient technique for avoiding an ill-defined mapping. Suppose coefficients $t(\ell, k)$ are found such that

$|\cos \omega| > 1$ (see equation (2.7)). The error deviation is then scaled by the maximum value of $\cos \omega$ (c_m) over the region $(0, \pi) \times (0, \pi)$:

$$e'(\omega_1) = \frac{e(\omega_1)}{c_m} = \frac{\cos \omega_o}{c_m} - \sum_{\ell=0}^{M} \sum_{k=0}^{M} \frac{t(\ell, k)}{c_m} \cos(\ell \omega_1) \cos(kg(\omega_1)) \tag{3.14}$$

where ω_o maps to the original approximating contour. It can be seen that the new error $e'(\omega_1)$ and the corresponding contours remain optimal (either in the minimax or least squares sense). However, the frequency from which this contour is mapped has changed ($\cos \omega_o' = \frac{\cos \omega_o}{c_m}$). Scaling the coefficients $t(\ell, k)$ by c_m has made the mapping well-defined. However, there is no guarantee on monotonicity as with the previous methods. This technique is presently under experimental investigation.

3.2 The Design of Circularly Symmetric Lowpass Filters with Transition Width Compression

The problem is to find the best possible transformation to satisfy a set of two-dimensional lowpass filter specifications. The one-dimensional passband cutoff frequency must map closely to a curve described by $\omega_2 = (R_p^2 - \omega_1^2)^{\frac{1}{2}}$ where R_p is the two-dimensional passband edge radius.

It is well known that the transition width of a one-dimensional lowpass filter affects the amount of ripple (error) which must be tolerated in the passband and stopband of such a filter. Thus, the one-dimensional transition width should be made as large as possible while satisfying the two-dimensional specifications. A mapping which produces a dense packing of contours in the two-dimensional transition width is desirable, because then a wide one-dimensional transition width can be mapped into

a narrow two-dimensional transition width.

We can show that

$$\left.\frac{\partial \omega}{\partial \omega_1}\right|_{\omega_2} = \left.\frac{\partial x}{\partial u}\left(\frac{1 - u^2}{1 - x^2}\right)^{\frac{1}{2}}\right|_{\substack{v = \cos \omega_2 \\ u = \cos \omega_1}} \tag{3.15}$$

where $\frac{\partial \omega}{\partial \omega_1}$ represents the slope (compression) of contours along a horizontal line in ω_2. In particular for $\omega_2 = 0$, $\frac{\partial \omega}{\partial \omega_1}$ is a measure of contour compression along the ω_1 and ω_2 axes (assuming octagonal symmetry). One possibility is to find linear constraints such that $\frac{\partial \omega}{\partial \omega_1}$ is large in the transition region. The remaining degrees of freedom can be used to approximate circularity along the passband edge with the hope that circularity will be maintained in the transition width while preserving the density of contours. Maximizing $\frac{\partial \omega}{\partial \omega_1}$ at a particular ω_1 generates nonlinear constraints and is thus undesirable, while maximizing $\frac{\partial x}{\partial u}$ at $u = \cos \omega_1$ generates linear constraints. Although $\frac{\partial x}{\partial u}$ is distorted by the nonlinear term $\left(\frac{1 - a^2}{1 - x^2}\right)^{\frac{1}{2}}$, compression has nonetheless been achieved as shown in the following sections.

The theory of transition width compression is now presented for the first, second, and third order transformations. Finally, in section 3.3 a technique for finding a good filter which matches a given set of specifications is developed and comparisons are made between the different order transformations.

We saw in section 2.3 the constraints which led to McClellan's original coefficients $-t(0,0) = t(1,0) = t(0,1) = t(1,1) = 0.5$. Since $t(1,0) + t(1,1) = 1$ and $(t(0,0)\, t(1,1) - t(1,0)\, t(0,1)/t(1,1) = -1$, the mapping is well defined and all contours are monotonically decreasing. However,

there are no degrees of freedom remaining and so no approximation problem. Fortunately it can be shown (Fig. 2.3) that these contours are nearly circular, at least for small to moderate radii. It is also true that these constraints preserve the one-dimensional filter along the ω_1, ω_2 axes. This is because the above constraints imply the constraint $\omega \rightarrow (\omega, 0), (0, \omega)$. Hence the transition width is determined by that of the one-dimensional filter.

Suppose $\omega = 0.5\,\pi$ is closely mapped to the circle with radius $0.5\,\pi$. Then,

$$\frac{\partial \omega}{\partial \omega_1} \approx \frac{\partial x}{\partial u} \left(\frac{1 - 0}{1 - 0}\right)^{\frac{1}{2}} = \frac{\partial x}{\partial u} \tag{3.16}$$

in the region of the $0.5\,\pi$ radius along the ω_1, ω_2 axes. To obtain compression in this region, $t(1, 0) + t(1, 1)$ must be made large as shown in Fig. 3.4. However, this condition leads to an ill-defined mapping. In fact the maximum slope $t(1, 0) + t(1, 1)$ is one for any set of constraints. Likewise, it can be argued that the actual contour compression $\frac{\partial \omega}{\partial \omega_1} \leq 1$. Figure 3.5 shows $\frac{\partial \omega}{\partial \omega_1}$ as a function of $\frac{\partial x}{\partial u}$ for $u = 0$ and $u = \pm 0.8$, where a well-defined mapping is guaranteed. Since in the original McClellan mapping $u = v = x$ along the axes, $\frac{\partial \omega}{\partial \omega_1} = \frac{\partial x}{\partial u} = 1$. Hence, the McClellan mapping yields the densest packing along the axes.

Under constraints 2.21a, b, c with $x = \cos \omega$ a free variable, almost perfect circularity was achieved for circles of large radius, whereas McClellan's original mapping yields box-like contours for large radii. Figure 2.6 shows a circular approximation on the $0.8\,\pi$ radius for the mapping $x = \cos(0.685\,\pi)$. We see that a disadvantage is that the transition width will be wider due to the decrease of the slope $t(1, 0) + t(1, 1)$. Furthermore, this technique tends to give ill-defined mappings

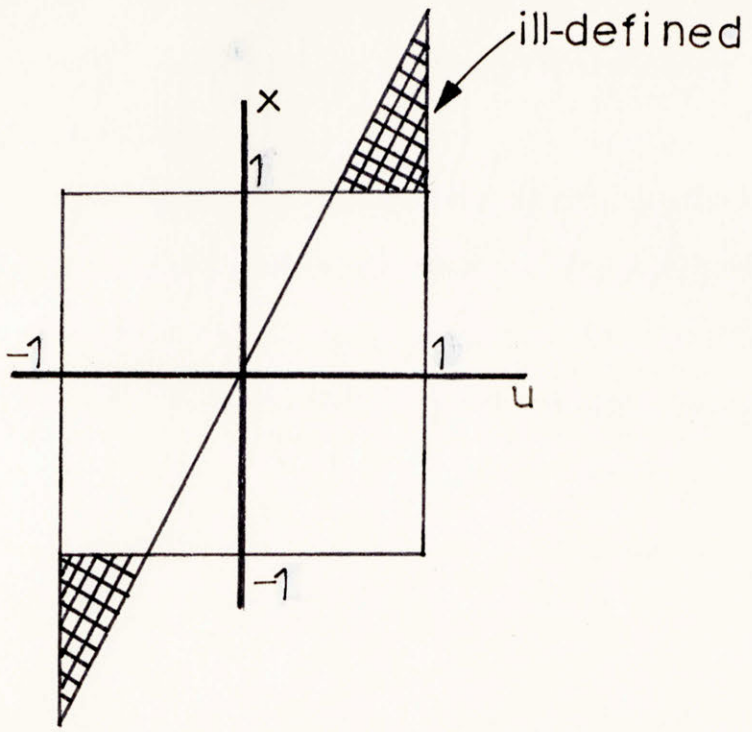

Fig. 3.4 Compression for the first order transformation.

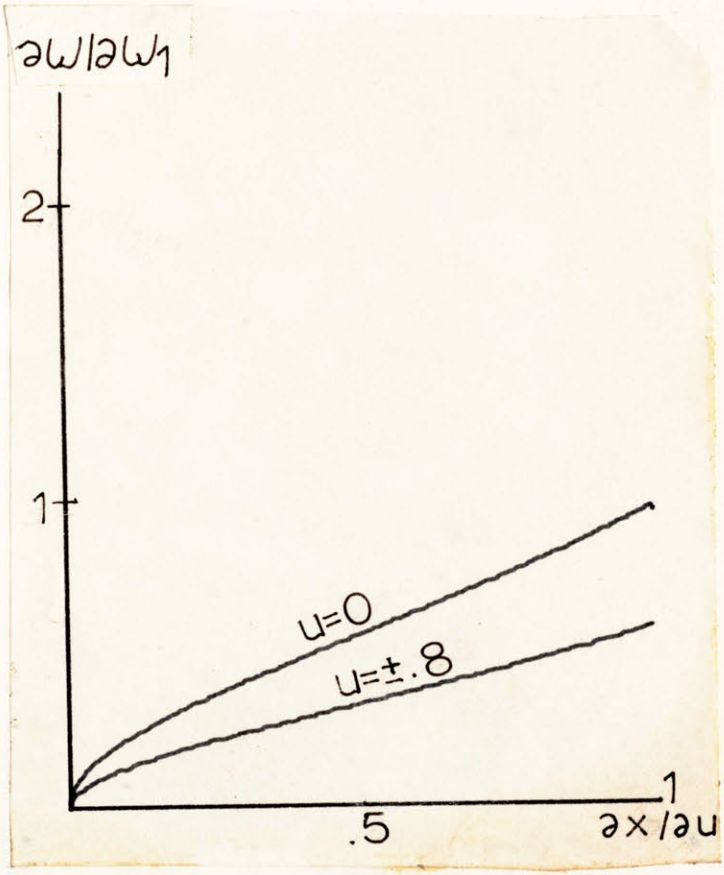

Fig. 3.5 The contour compression for the first order transformation as a function of $\partial x/\partial u$.

for circles of radius less than 0.2π, since x is such that $t(1,0) + t(1,1) > 1$.* Nevertheless, the original McClellan mapping yields almost perfect circles in this region, making the first order transformation useful for any radius (see section 3.4).

Since the first order transformation gives nearly perfect circularity on any radius, higher order transformations will probably prove useful only for transition width compression or for non-circular contour approximation. The second order transformation can be expressed as

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(u,v)\, u^{\ell} v^{k}, \tag{3.17}$$

a quadratic in u and v. Along the u axis,

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(\ell,k)\, u^{\ell}. \tag{3.18}$$

Imposing constraints 3.12a, b, c yields three free variables. These constraints fix the quadratic at the endpoints along the u axis. Therefore, two of the three degrees of freedom of the quadratic are consumed as demonstrated in Fig. 3.6.

If the constraint $\omega \rightarrow (\omega, 0)$ is imposed, the one-dimensional ω axis maps to the ω_1, ω_2 axes exactly. This implies that $x = u = v$ along the u, v axes. Thus the one-dimensional filter is preserved along the ω_1, ω_2 axes and no compression is obtained. This will be referred to as the basic second order transformation. The approximation problem can

---

* This can be avoided with the new technique of Mersereau and Mecklenbräuker (see section 2.1).

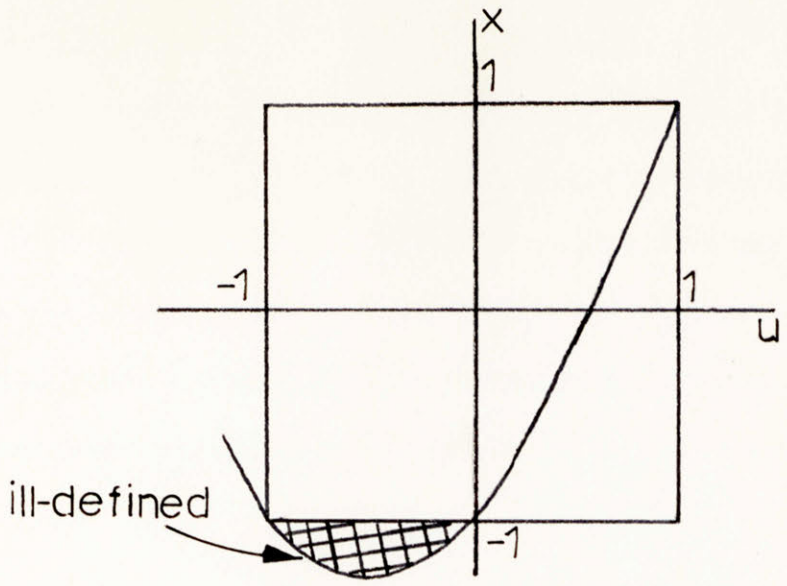

Fig. 3.6 The second order transformation along the u axis.

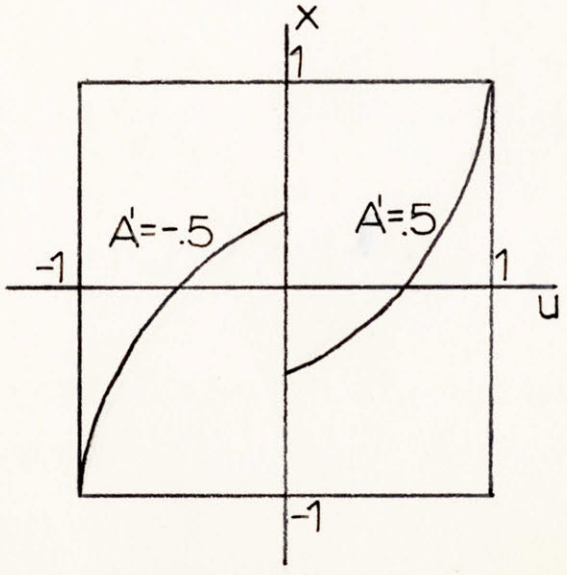

Fig. 3.7 The quadratic for maximum compression with the second order transformation.

be formulated similarly to that of the compression constraint problem to follow.

The possible ill-defined mapping of Fig. 3.6 can be prevented by constraining $\frac{\partial x}{\partial u}\Big|_{u = \pm 1} \geq 0$. Under this constraint and constraints (3.12a, b, c), an additional constraint was derived to maximize $\frac{\partial x}{\partial u}$ as a function of u (see Appendix 3). Letting A' equal the coefficient of the squared term in the quadratic expression for x along the u axis, then we see maximum compression occurs when

$$A' = \begin{cases} \frac{1}{2}, & u \geq 0 \\ -\frac{1}{2}, & u < 0 \end{cases} . \qquad (3.19)$$

This is shown graphically in Fig. 3.7. The slope at $u = \pm 1$ is two and decreases monotonically to one at u = 0. Thus compression is possible for small and large radii. However, the weighting factor $\left(\frac{1 - u^2}{1 - x^2}\right)^{\frac{1}{2}}$ will reduce the slope. At $u = \pm 1$, v = 1 using L'Hospital's rule,

$$\frac{\partial \omega}{\partial \omega_1} = \left(\frac{\partial x}{\partial u}\right)^{\frac{1}{2}} = \sqrt{2} . \qquad (3.20)$$

Therefore there is compression of $\sqrt{2}$ for very large and very small circles. Figure 3.8 shows an example of transition width compression with $R_p = 0.8\pi$. The approximation problem is presented in Appendix 3.

Recall that the second order transformation doubles the order of the two-dimensional filter achieved by the first order transformation, for a fixed length one-dimensional filter. Therefore the second order transformation must compress the transition width by at least a factor of two for circles of small radius for a fixed order two-dimensional filter.

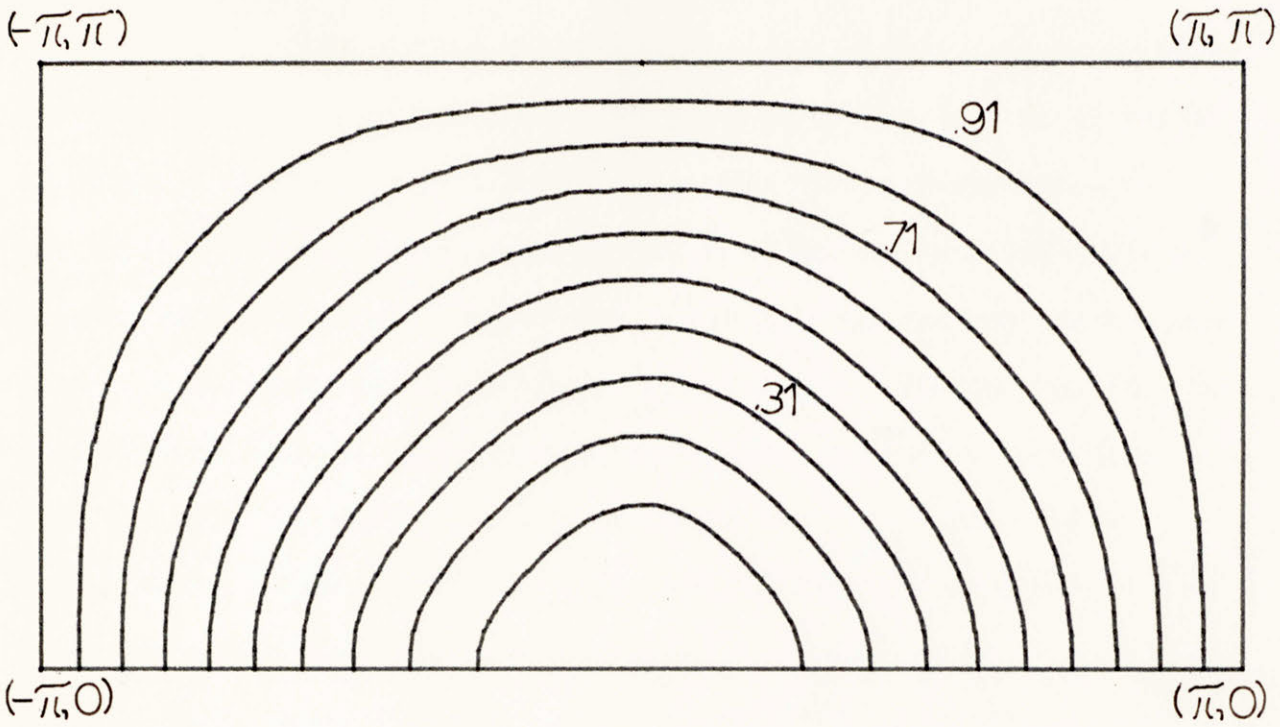

Fig. 3.8 The compressed contours of the second order transformation with an ideal radius of 0.8 π.

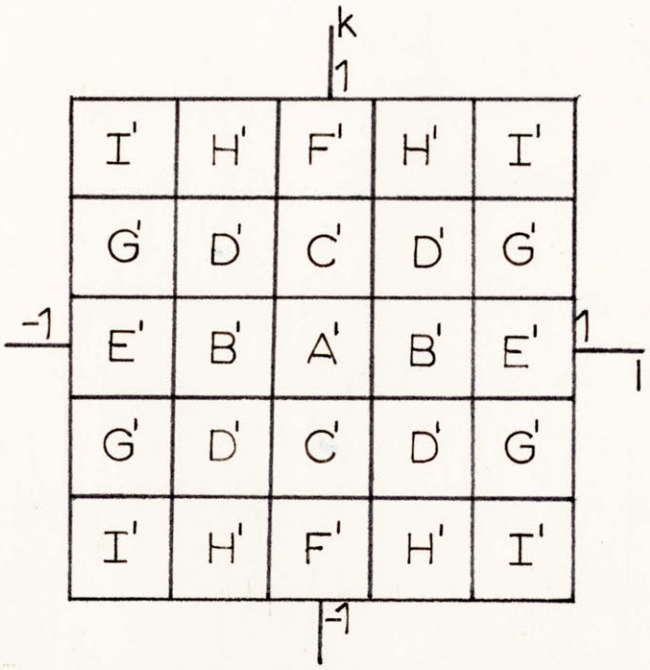

Fig. 3.9 The 5 x 5 nonrecursive filter for the second order transformation implementation.

For large circles the necessary compression is somewhat less due to the widening of the transition width with the first order transformation.

In conventional implementation schemes (direct convolution or the FFT) the number of multiplies per output point grows with the order of the filter. This is not necessarily true with the implementation of section 2.4. The number of multiplies is on the order of kn where k depends on the number of degrees of freedom in the transformation and (2n+1) is the order of the one-dimensional filter. We can show for the above second order transformation that the 5 x 5 nonrecursive filter of Fig. 3.9 must be implemented at each stage of Fig. 2.10. Each coefficient is a linear function of the transformation coefficients. The constraints (see Appendix 3) imply about two multiplies per output point at each stage, generating a total of 3n+1 multiplies per output point. The first order transformation of section 2.4 required 2n+1 multiplies per output point while yielding nearly perfect circularity for circles of large radius, but with a widening of the transition width. We will see in section 3.4 that the compression in this region for the second order transformation is nearly twice that of the first order transformation with a comparable number of multiplies (2n+1 versus 3n+1). Hence, the second order transformation would be preferred when multiplies are of primary cost and not filter order.

One consideration remaining is that of guaranteeing a well-defined mapping. The constraint (3.13b) is not necessarily satisfied. Nevertheless, as shown in Fig. 3.7, constraint (3.13a) is satisfied due to the nature of the quadratic along the u, v axes. It was found empirically that these constraints are sufficient for guaranteeing a well-defined mapping.

Compression can also be achieved with the third order transformation which can be expressed as

$$x = \sum_{\ell=0}^{3} \sum_{k=0}^{3} g(\ell, k)\, u^{\ell} v^{k} \tag{3.21}$$

a cubic in u and v. Along the v axis

$$x = \sum_{\ell=0}^{3} \sum_{k=0}^{3} g(\ell, k)\, u^{\ell}. \tag{3.22}$$

Imposing the constraints (3.12a, b, c) with the slope constraint

$$\left.\frac{\partial x}{\partial u}\right|_{u = \pm 1} = 0 \tag{3.23}$$

led to a maximum slope at the origin of value 3/2 (see Appendix 4). Figure 3.10a shows the cubic nature of x as a function of u along the v axis. The term $\left(\frac{1 - u^2}{1 - x^2}\right)^{\frac{1}{2}}$ is negligible in the region of contour compression, since $u \approx x \approx 0$ (see Fig. 3.10b). Figure 3.10c demonstrates this compression near a radius of $0.5\,\pi$. Although compression is not three or greater, the same implementation consideration can be made as with compression for the second order transformation. The number of multiplies per output point is about $5n+1$. Thus, if filter order is not of primary cost, this technique may be preferable for circles of radius about $0.5\,\pi$.

An attempt was made to achieve compression for large and small radii (using a technique similar to that of Appendix 4). With constraints (3.12a, b, c) and the additional constraint

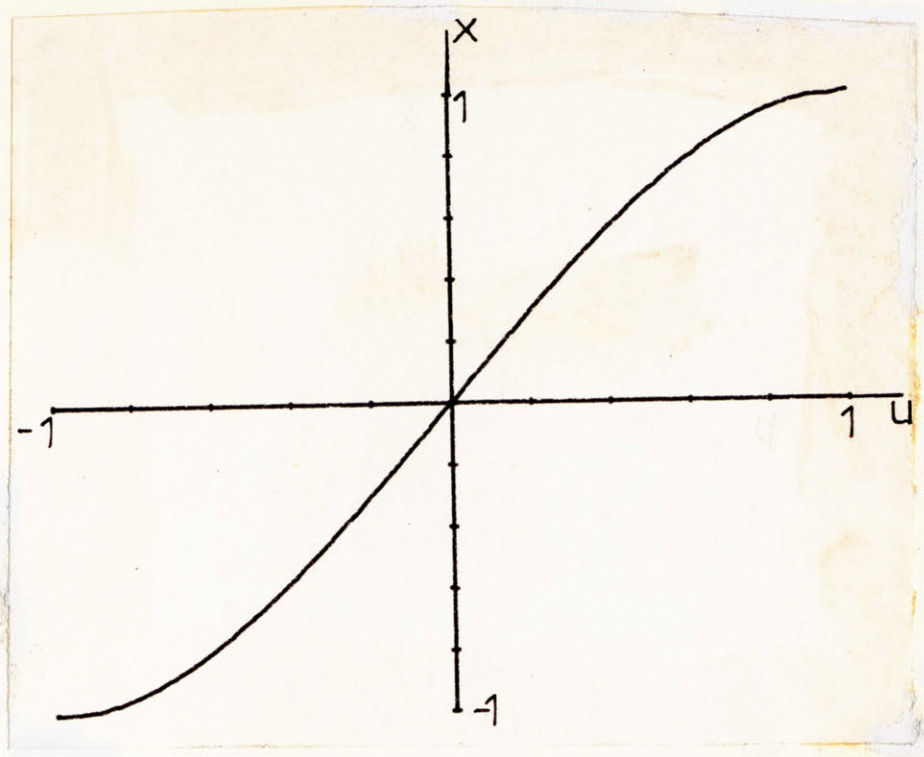

Fig. 3.10a The cubic along the u axis for the third order transformation with maximum compression on the 0.5 π radius.

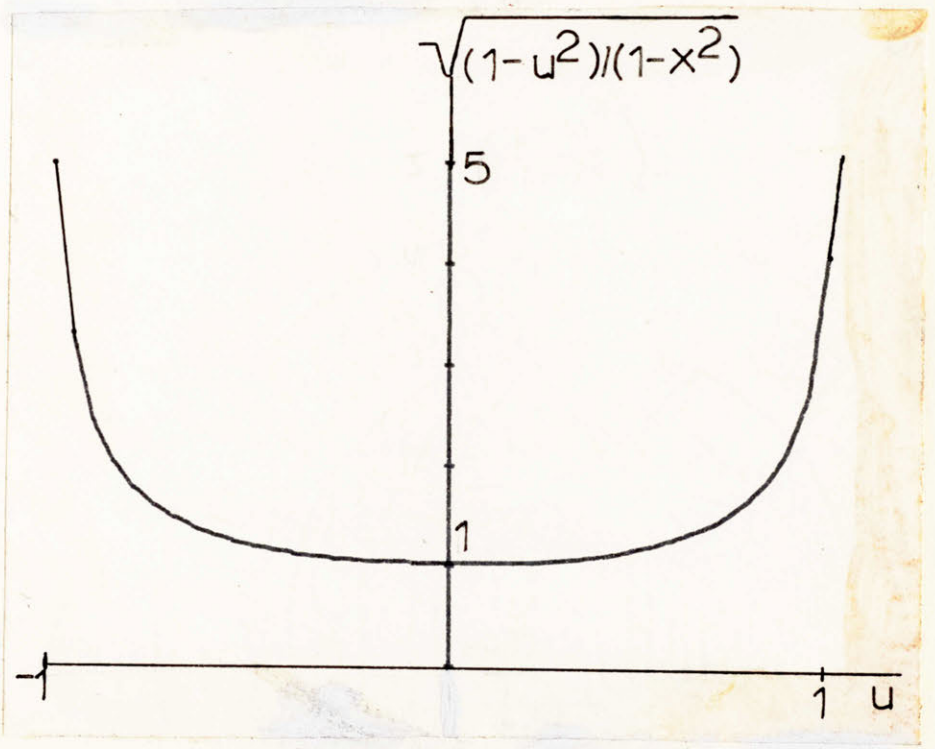

Fig. 3.10b The weighting function for the cubic of Fig. 3.10a.

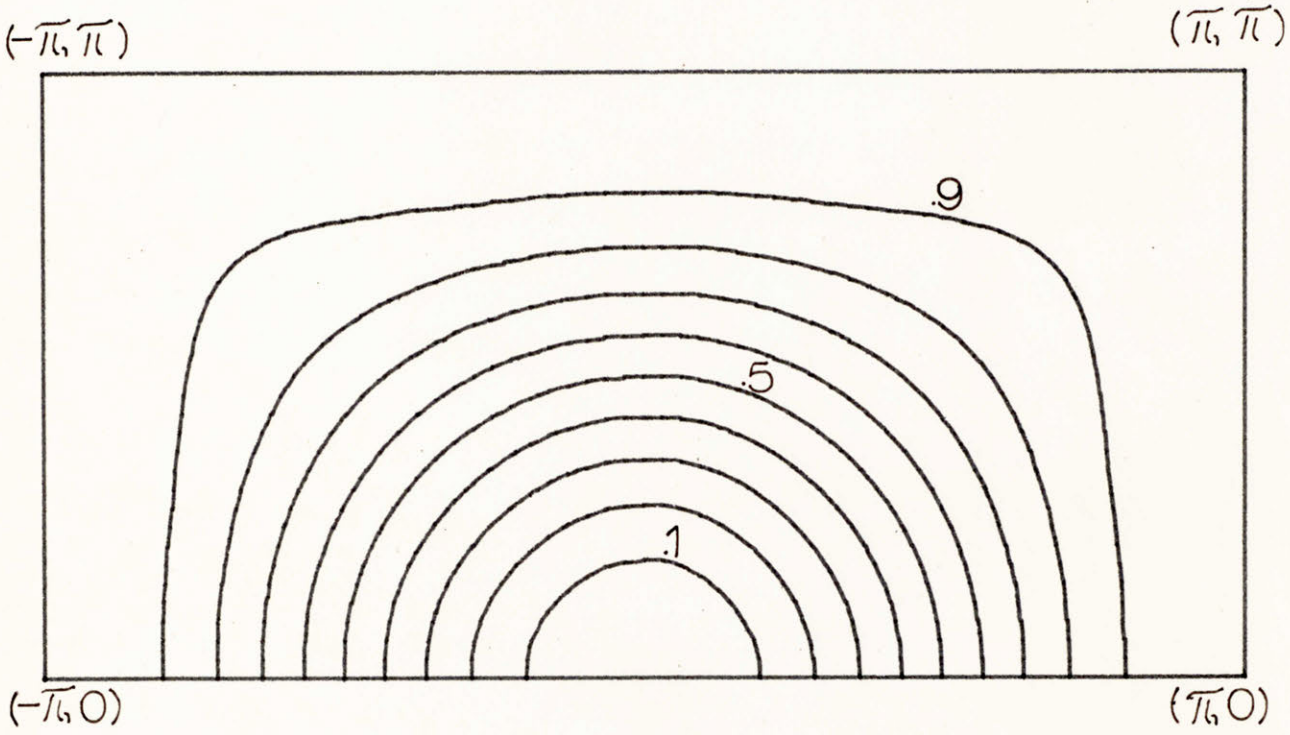

Fig. 3.10c The compressed contours of the third order transformation with an ideal radius of 0.5π.

$$\left.\frac{\partial x}{\partial u}\right|_{u = 0} = 0 \qquad (3.24)$$

the maximum possible slope at the endpoints is three. Using L'Hospital's rule $\frac{\partial \omega}{\partial \omega_1} = \sqrt{3}$ at zero and π. Figures 3.11a, b show x as a function of u along the u axis and the resulting contours. As seen in Fig. 3.11b, the mapping is ill-defined and lacks monotonicity. Hence, linear programming may be necessary to guarantee a well-defined mapping.

For the second and third order transformations, π was mapped to the outer boundaries. Replacing this constraint with the constraint $\pi \rightarrow (\pi, \pi)$ may lead to better circularity for large radii. Compression constraints for this mapping and other mappings have yet to be investigated for the higher order cases.

3.3 An Algorithm for Meeting Two-Dimensional Specifications

An algorithm will be presented for meeting two-dimensional specifications in terms of one-dimensional passband and stopband cutoff frequencies. The algorithm is designed for circularly symmetric lowpass filters, but can be generalized.

A specification problem for circularly symmetric lowpass filters is the following. Find the transformation and one-dimensional filter such that the filter satisfies a given passband cutoff tolerance R_p and stopband cutoff R_s with the smallest passband and stopband deviation. The two-dimensional filter should be of size $(2n+1) \times (2n+1)$.

The specifications are met when the one-dimensional passband and stopband frequencies ω_p and ω_s map to contours tangent to the specified cutoff radii (see Fig. 3.12).

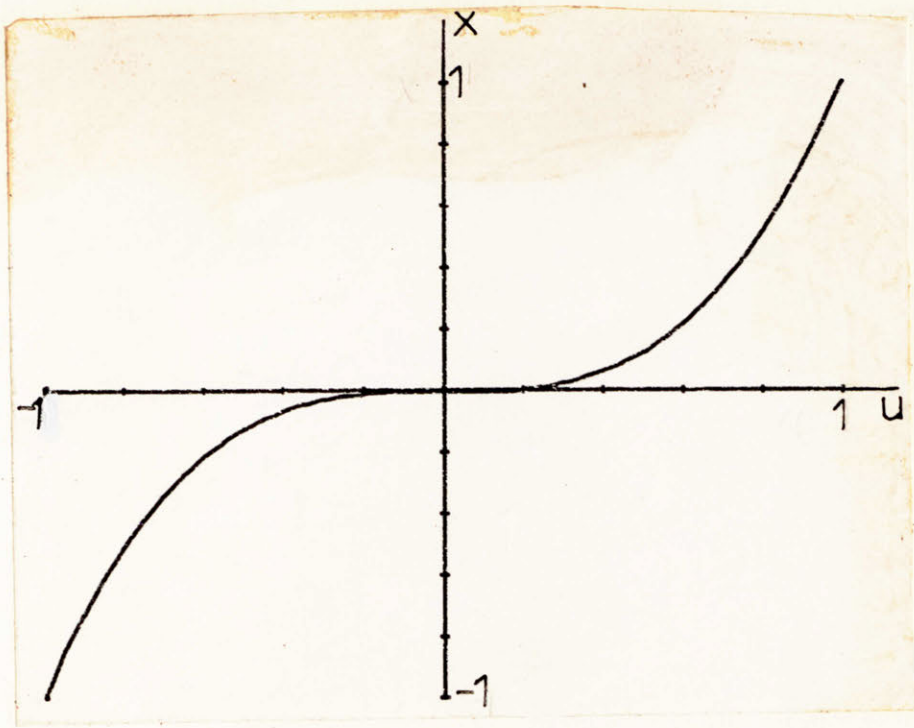

Fig. 3.11a The cubic along the u axis for the third order transformation with maximum compression on the 0 and π radii.

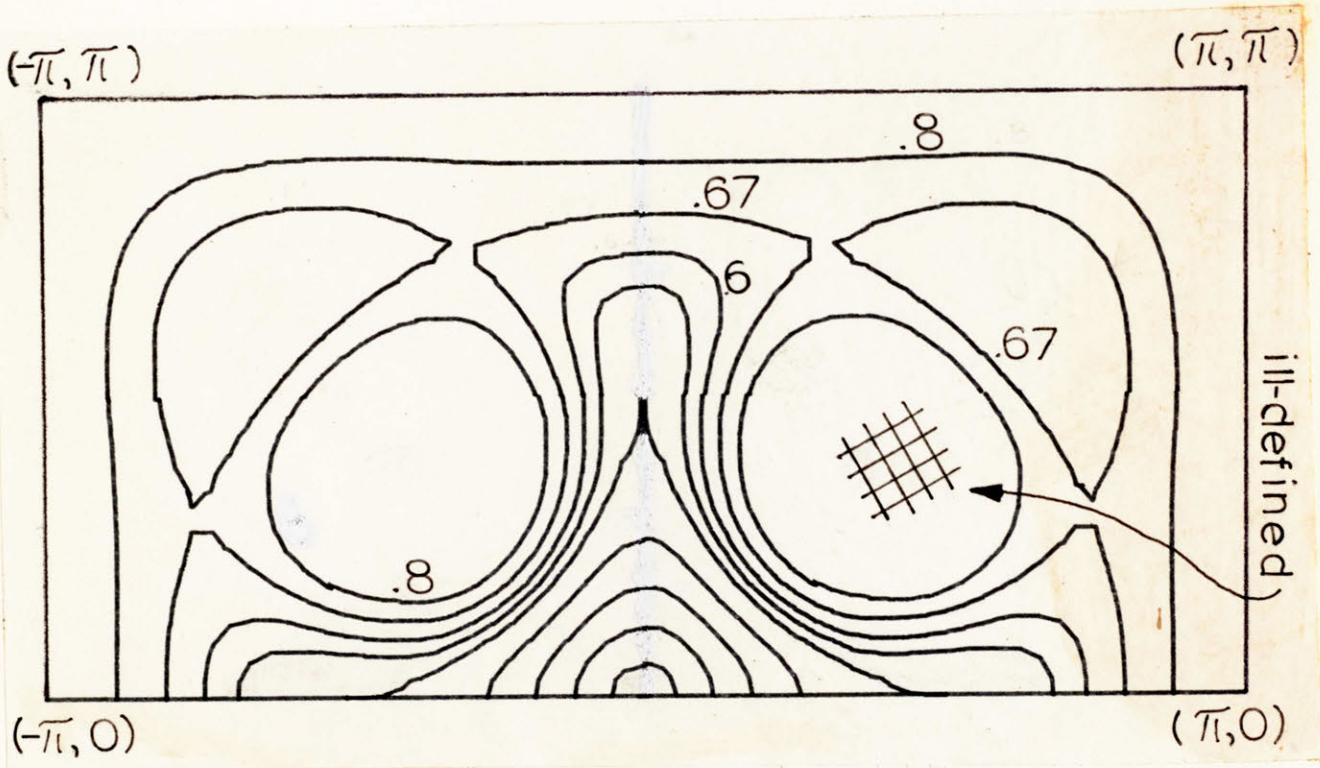

Fig. 3.11b The contours for the third order transformation with compression on the 0 and π radii, and an ideal radius of 0.8 π.

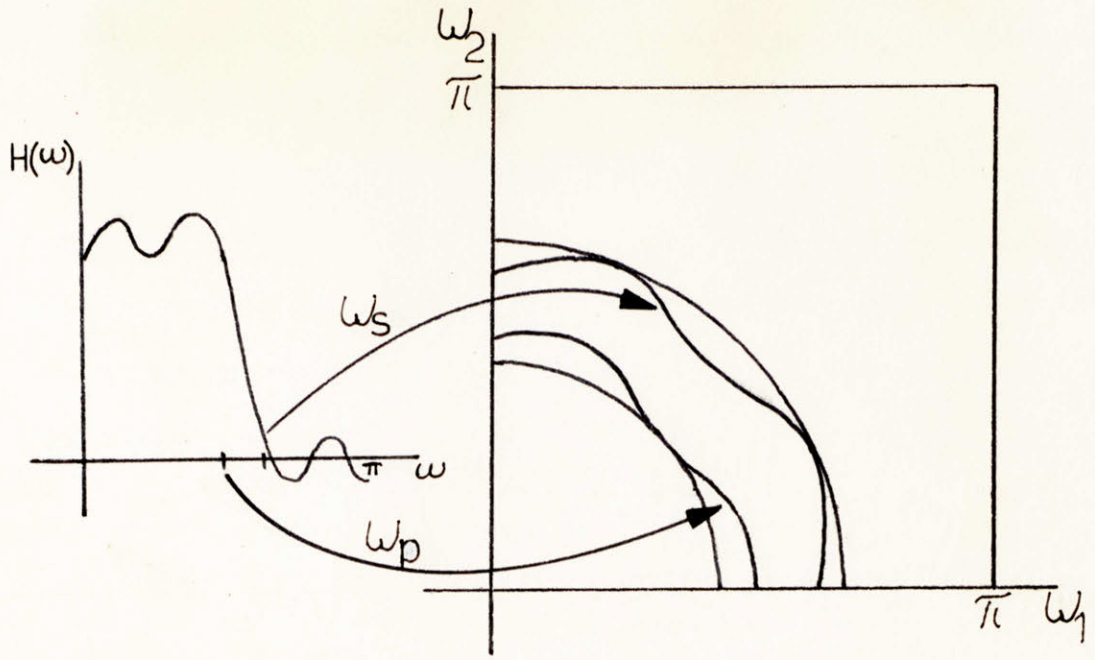

Fig. 3.12 Specification matching.

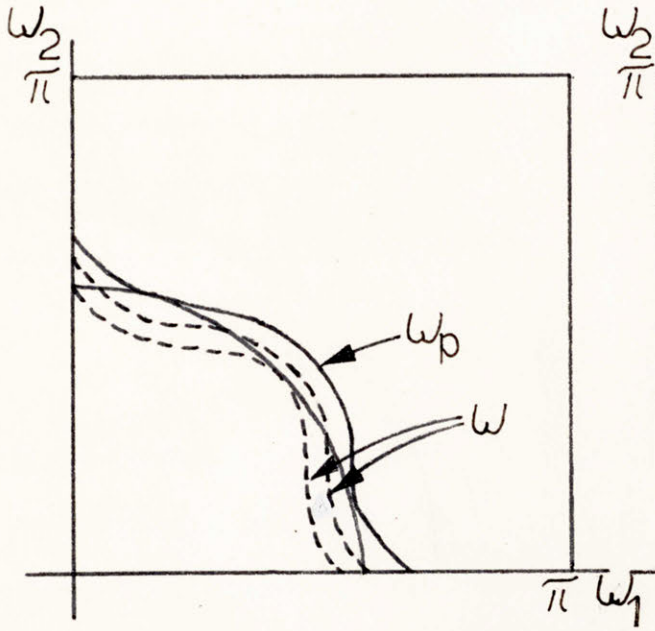

Fig. 3.13 The test for a tangent passband edge contour.

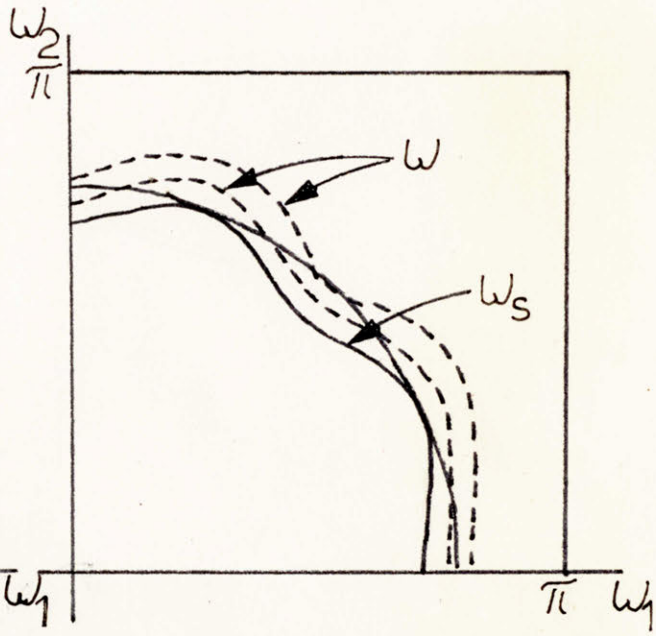

Fig. 3.14 The test for a tangent stopband edge contour.

The algorithm to perform the specification matching can be described in the following way. In general, the approximation routine yields a contour oscillating about the ideal contour. Thus the passband specification is not met. A test for a tangent contour is made. If the test fails within a certain error criterion, the ideal radius is incremented and the approximation repeated for a new contour.

Since the value of ω increases radially ($\cos \omega$ decreases), a very simple test for a tangent contour is whether the following holds:

$$\cos \omega = \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \cos(\ell\omega_1) \cos(kg(\omega_1)) < \cos \omega_p \qquad (3.25)$$

That is, traversing the ideal curve $g(\omega_1)$ yields a different inverse frequency image ω (different from ω_p) for each ω_1. If the contour mapped from ω_p is tangent to the ideal curve $g(\omega_1)$, then ω is less than or equal to ω_p ($\cos \omega \geq \cos \omega_p$) everywhere along $g(\omega_1)$. See Fig. 3.13 for a graphical interpretation. If $\cos \omega$ is less than $\cos \omega_p$ along $g(\omega_1)$, the radius for the ideal approximation is incremented by the tangent error

$$\epsilon = \max [\omega - \omega_p] . \qquad (3.26)$$

If the radius is incremented so that the new passband contour goes beyond the given ideal passband circle, then the radius will be decremented by ϵ on the next iteration. Eventually the passband contour mapped from ω_p will converge to a curve almost tangent (within some allowable error) to the given passband radius.

The one-dimensional cutoff frequency ω_s, which maps to a contour that meets the stopband radius specification, can be found quite simply.

The maximum value of cos ω along $g_s(\omega_1) = \sqrt{R_s^2 - \omega_1^2}$ corresponds to $\omega = \omega_s$, which is tangent to the ideal stopband edge. See Fig. 3.14 for a graphical interpretation. ω_s can be expressed as

$$\omega_s = \arccos\left[\min_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \cos(\ell\omega_1) \cos(k g_s(\omega_1))\right]. \qquad (3.27)$$

The maximum deviation from circularity along the ideal passband edge can be approximated by the expression

$$\epsilon \approx 2(R_p' - R_p) \qquad (3.28)$$

where R_p is the original ideal passband radius and R_p' is the new (incremented) ideal radius over which the approximation is made. Although the approximating contour is not exactly equiripple about the ideal, ϵ is a good indication of the maximum circularity deviation.

The flow graph for the algorithm is shown in Fig. 3.15. For example, the second order compression mapping in Fig. 3.16 has an ideal passband radius 0.8π and stopband radius 0.9π. The one-dimensional cutoff frequencies necessary for tangent contours are $f_p = 0.727$ and $f_s = 0.82$.

Alternatively, we could approximate the stopband ideal cutoff first and then find the tangent passband contour. In certain cases this may yield better circularity over the transition band since inner contours tend more toward circularity than the outer contours. Another possibility is that the design could be performed on the two contours simultaneously.

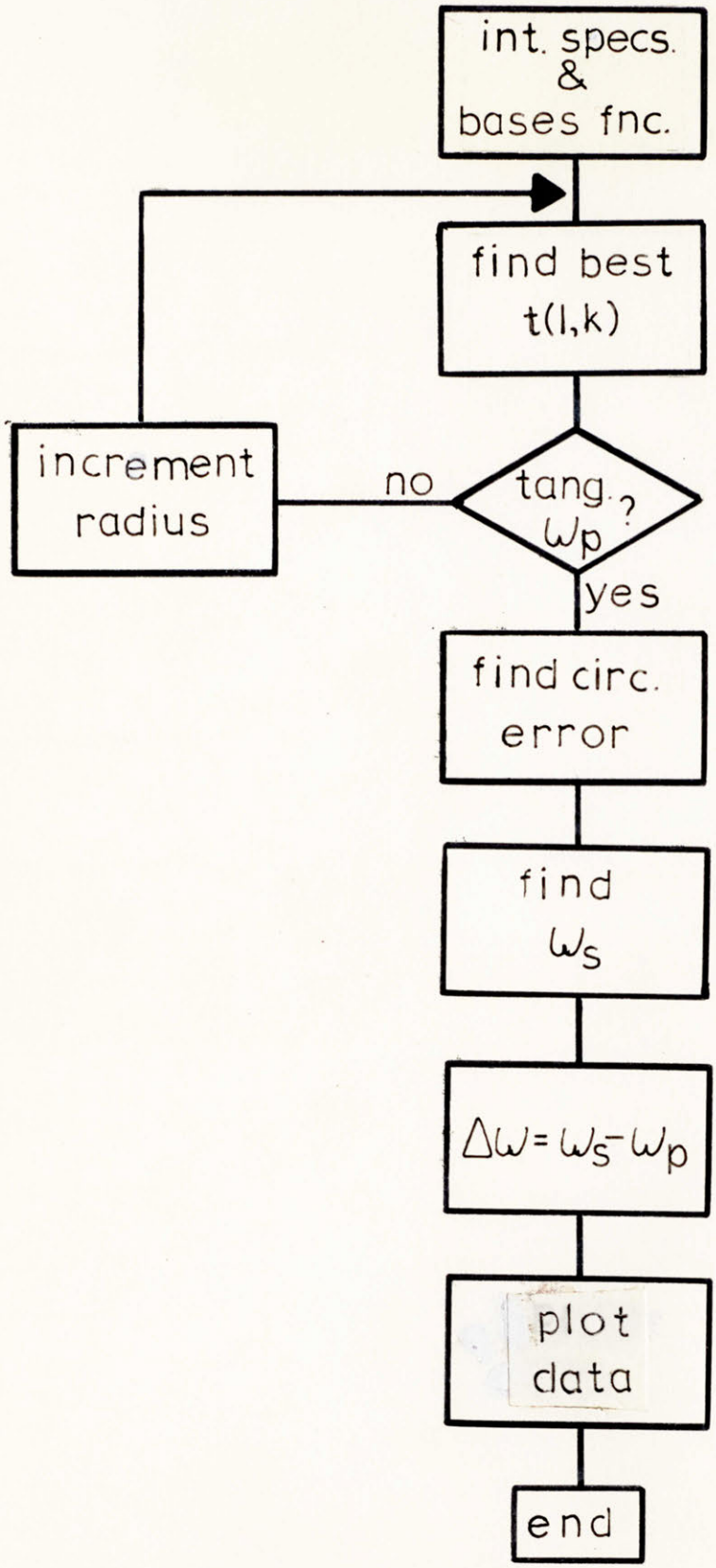

Fig. 3.15 A skeleton flow graph for the specification matching algorithm.

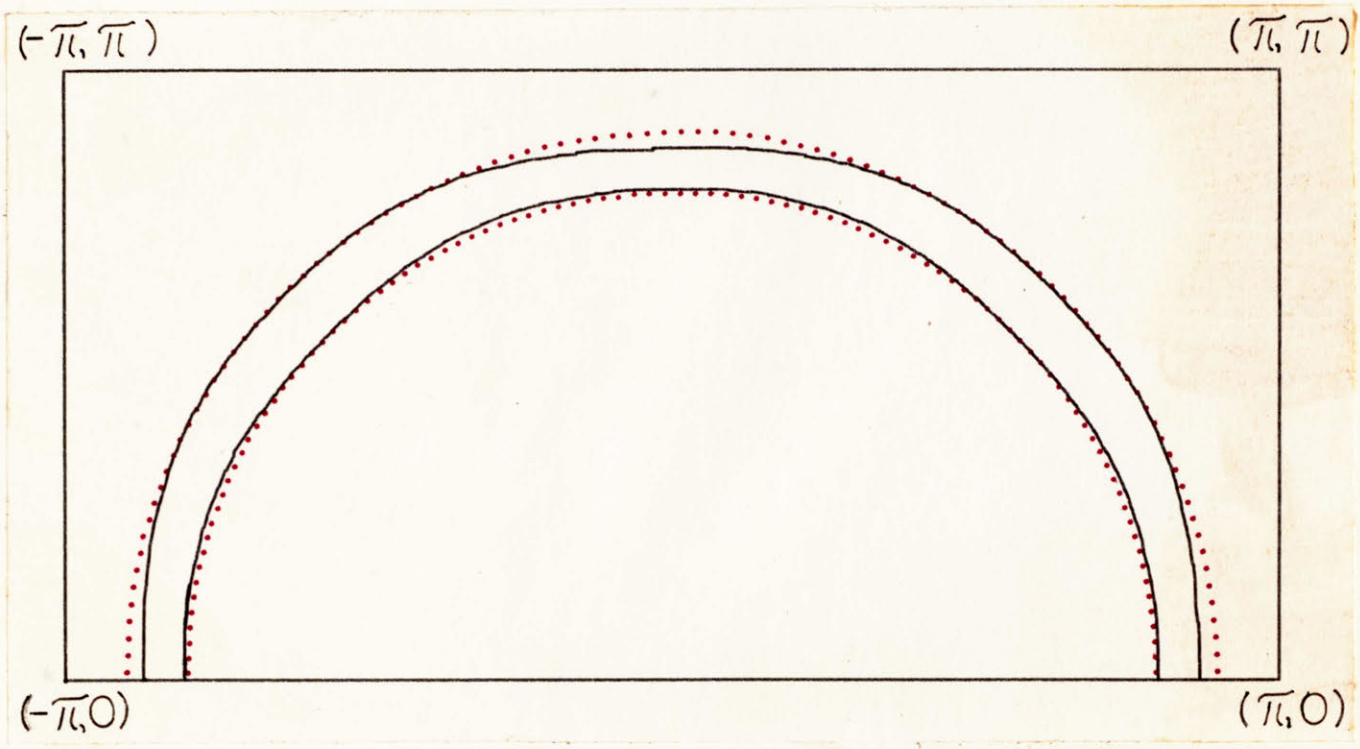

Fig. 3.16 An example of specification matching. The dotted lines are the ideal passband and stop-band edges.

3.4 Comparisons in Transition Width Compression

Figures 3.17a, b show the relative advantages of the methods discussed in section 3.2 for transition width compression. All curves refer to the two-dimensional specifications which require a passband radius scanning from 0.1π to 0.9π and a transition width of 0.05π.

The first graph gives the allowable one-dimensional transition width. The second graph shows the one and hence two-dimensional passband and stopband deviations given by the appropriate one-dimensional formula due to Rabiner et al.:

$$14(N-1)\Delta\omega_t + 15 \approx -20 \log (\delta_p \cdot \delta_s) \tag{3.29}$$

where

N = filter order

δ_p = passband deviation

δ_s = stopband deviation

$\Delta\omega_t$ = transition width.

We assume $N = 21$ and $\delta_p = \delta_s = \delta$.

The third graph plots the circularity deviation described in section 3.3. Since McClellan's original first order mapping does not involve an approximation, there is no ripple about the ideal curve. The deviation, maximum on the 45^o line, can be approximated by

$$\epsilon \approx \omega_p - \left[\sum_{\ell=0}^{1} \sum_{k=0}^{1} t(\ell, k) \cos(\ell\omega_1) \cos\left(k (R_p^2 - \omega_1^2)^{\frac{1}{2}}\right) \right] \tag{3.30}$$

where $\omega_1 = \cos(\pi/4)$.

Figures 3.17a, b demonstrate a number of important points. The original McClellan mapping (solid line in Fig. 3.17b) makes specification

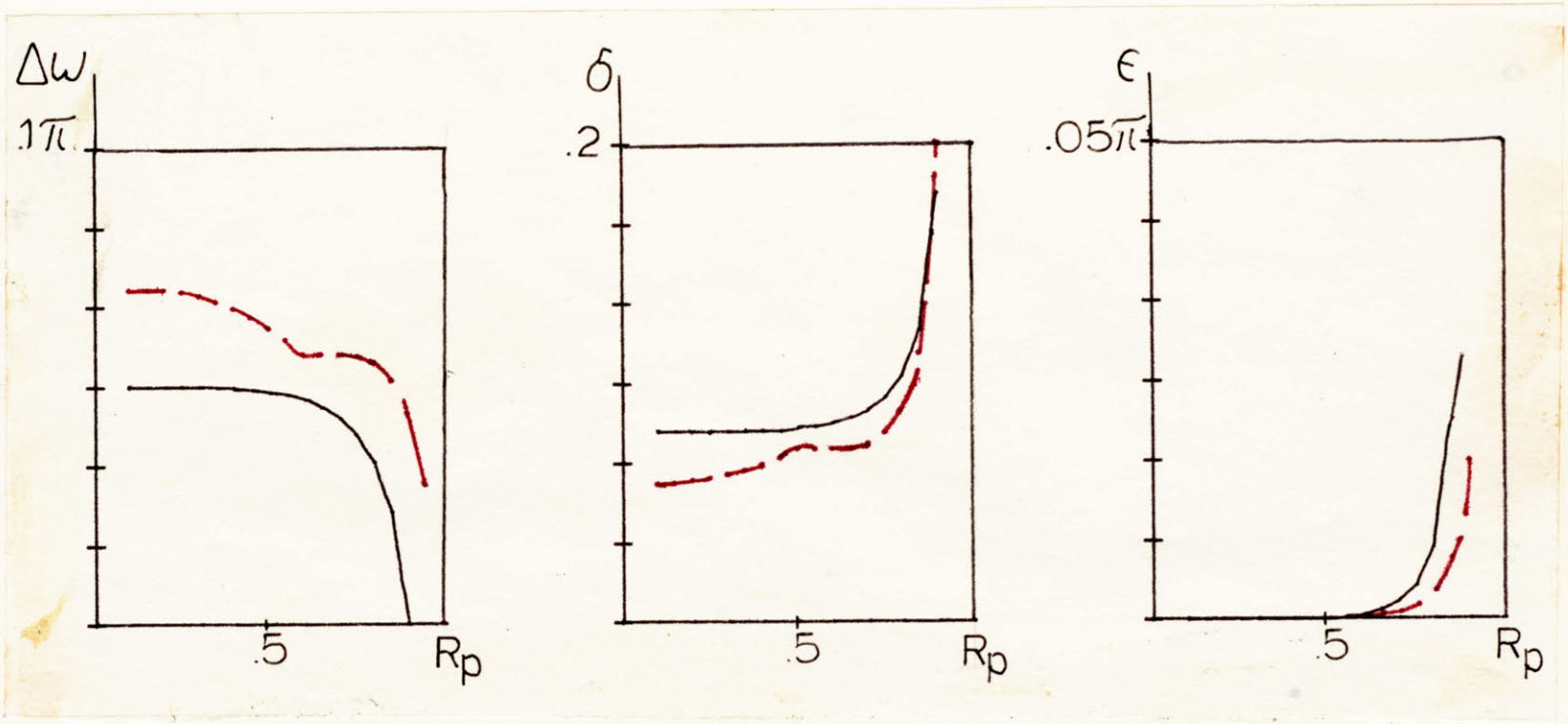

Fig. 3.17a Parameters for the meeting of specifications with the second order transformation. The two-dimensional transition width is fixed at 0.05 π.

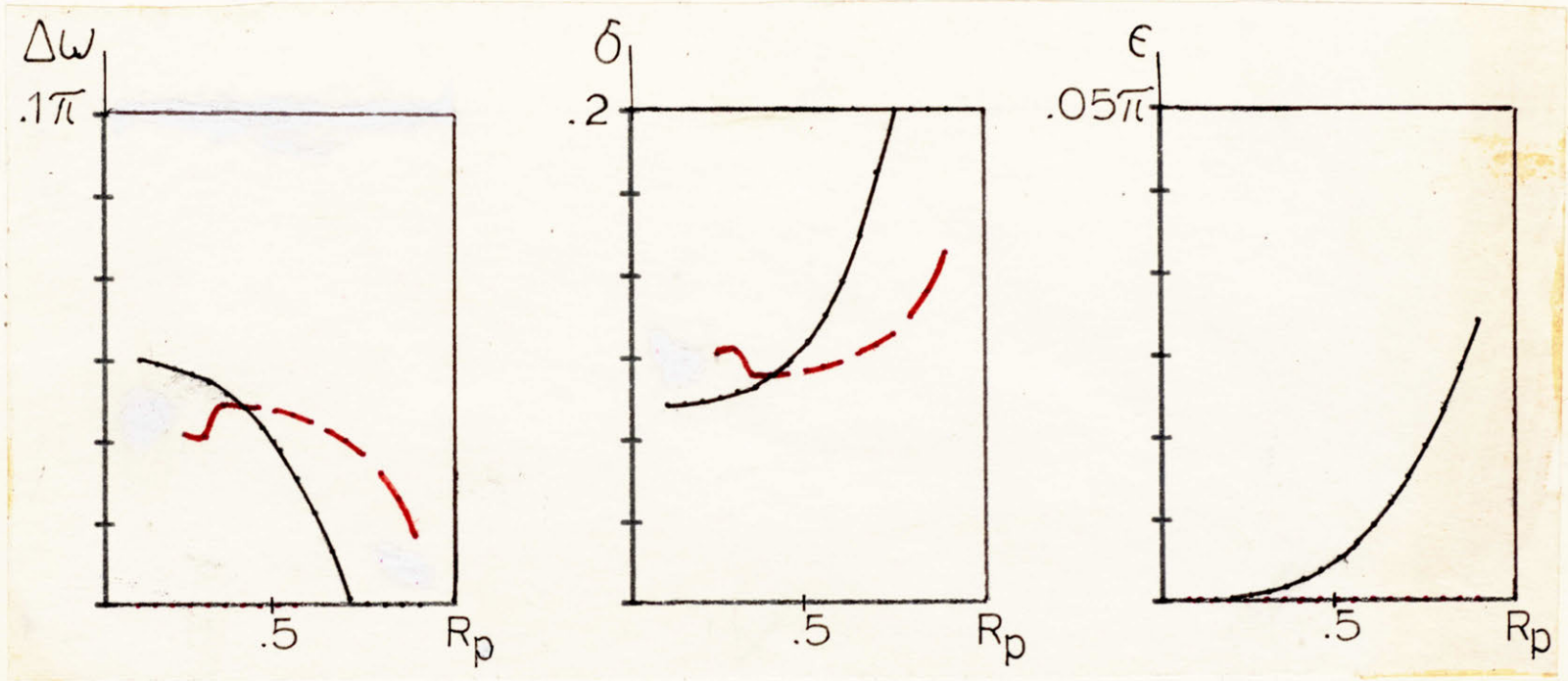

Fig. 3.17b Parameters for the meeting of specifications with the first order transformation. The two-dimensional transition width is fixed at 0.05 π.

matching impossible for radii greater than about 0.7π. This is due to the box-like nature of the contours in this region. Thus passband and stopband deviations as well as the circularity deviation grow arbitrarily large for large radii.

The first order mapping with $x = \cos\omega$ free (dotted line in Fig. 3.17b), yields ill-defined regions for radii less than about 0.2π; that is, x is such that $t(1,0) + t(1,1)$, the slope of the mapping along the u, v axes, is greater than one. Nevertheless, there is approximately zero circularity deviation for any radius.

Figure 3.17a compares the second order transformation with compression constraints (dotted line) and without compression constraints (solid line). We see the compression mapping yields an improvement in passband and stopband deviation over all other three techniques. However, the circularity is degraded for contours of very large radii. Of course, this advantage is recognized when the number of multiplies and not the filter order is of primary cost.

In considering a fixed order filter, we see the required transition width for the second order compression mapping is not twice that of either first order mapping. The original first order mapping is best up to about a radius of 0.45π, after which the alternative first order mapping yields a higher allowable transition width and lower deviations. Figure 3.18 gives a plot of the best value of $t(0,0)$ and its respective first order transformation for a fixed two-dimensional filter order of 21 x 21.

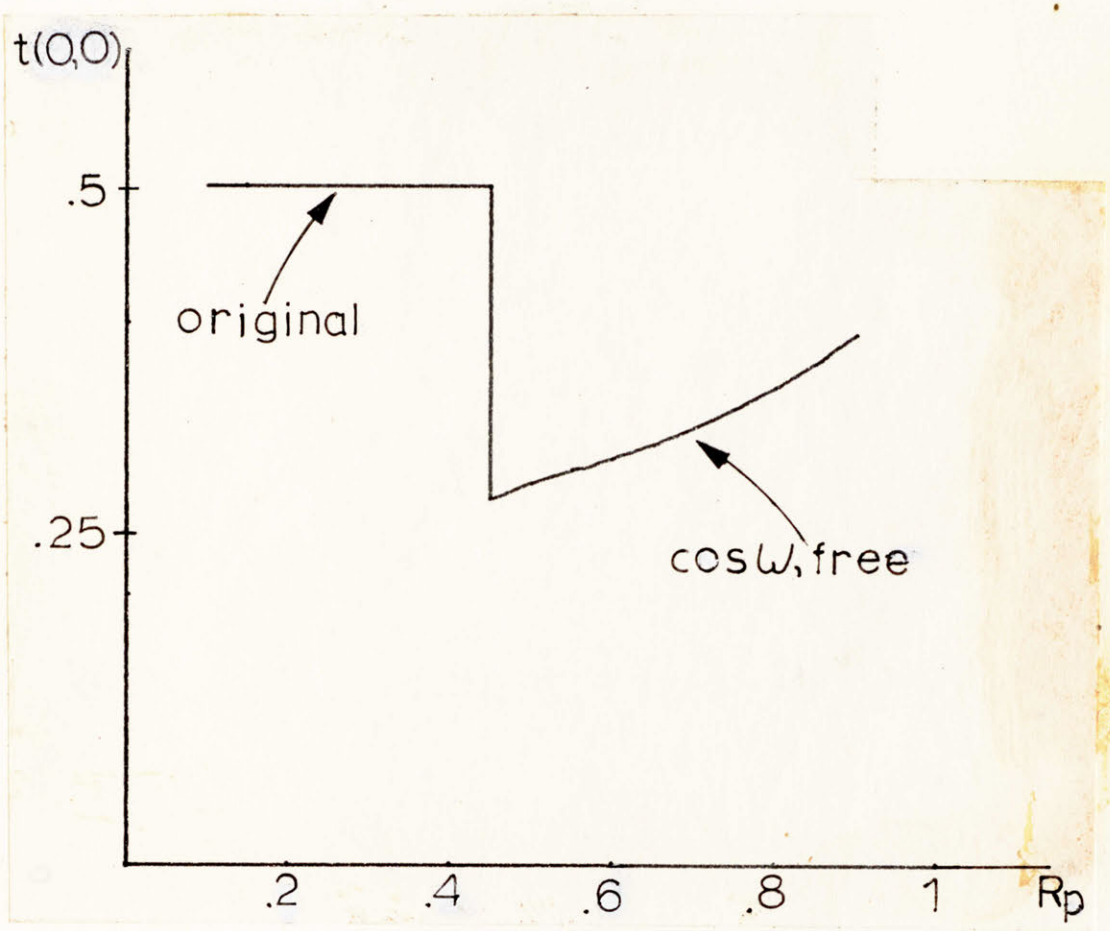

Fig. 3.18 The best choice of t(0, 0) for a fixed order 21 x 21 two-dimensional filter with a transition width of 0.05 π.

CHAPTER 4

THE APPROXIMATION OF ARBITRARY CONTOURS

4.1 The Lowpass to Highpass Transformation

In Chapter 2 we discussed the general approximation problem where the function $\omega_2 = g(\omega_1)$ described the ideal curve in the two-dimensional region $(0,\pi) \times (0,\pi)$. In this chapter we shall apply this idea together with new constraints to obtain a variety of novel two-dimensional filters.

As a starting point consider the design of a highpass filter. The ideal highpass filter can be defined to have a value of one outside a radius R_p and zero inside R_s with some allowable transition width. One design procedure is to transform a one-dimensional highpass filter into a two-dimensional highpass filter. Another method which demonstrates the flexibility of the algorithm is the transformation of a one-dimensional lowpass filter into a two-dimensional highpass filter.

Applying the constraints

$$0 \rightarrow (\omega_1, \pi) \tag{4.1a}$$

$$\pi \rightarrow (0, 0) \tag{4.1b}$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1), \tag{4.1c}$$

the constraint equations for the first order transformation become

$$t(1,0) - t(1,1) = 0 \tag{4.2a}$$

$$t(0,0) - t(1,0) = 1 \tag{4.2b}$$

$$t(0,0) + t(1,0) + t(0,1) + t(1,1) = -1 \tag{4.2c}$$

$$t(1,0) = t(0,1). \tag{4.2d}$$

The four constraints leave no free variables. Hence solving we obtain

$$-t(0,0) = t(1,0) = t(0,1) = t(1,1) = -0.5. \qquad (4.3)$$

Along the u axis

$$x = (t(0,0) + t(0,1)) + (t(1,0) + t(1,1))u = -u. \qquad (4.4)$$

Thus the slope of x has changed sign from that of the original lowpass to lowpass transformation (x = +u). Figure 4.1 shows the frequency contour plot for the lowpass to highpass transformation.

4.2 The Elliptical Contour

The ideal passband cutoff contour for an elliptically shaped filter can be expressed as

$$\omega_2 = g(\omega_1) = b(1 - (\omega_1/a)^2)^{\frac{1}{2}}$$

where a and b are the values of the major and minor axes, respectively. One possible constraint set is

$$0 \rightarrow (0,0) \qquad (4.5a)$$

$$\pi \rightarrow (\pi,\pi). \qquad (4.5b)$$

For the ideal ellipse with major axis in the horizontal direction, these constraints compress the contours around the ω_1 axis. Although almost zero approximation error is obtained, ill-defined regions exist near the boundary $\omega_2 = \pi$.

Linear programming can be avoided by applying the alternate constraint set

$$\omega \rightarrow (0,\omega) \qquad (4.6a)$$

$$\pi \rightarrow (\omega_1,\pi). \qquad (4.6b)$$

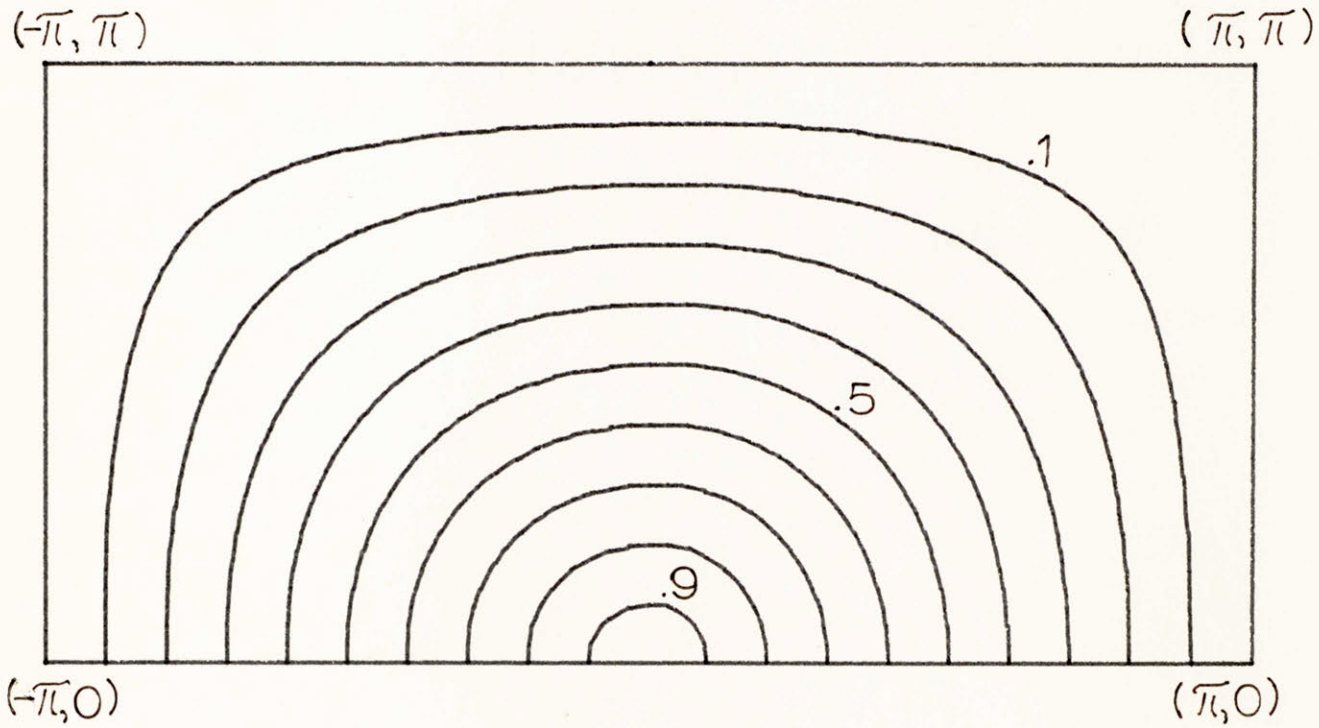

Fig. 4.1 The contours of the lowpass to highpass first order transformation.

For a > b the first constraint controls the spacing of the contours. The second constraint tends to spread out the contours in the vertical direction. This serves as a pseudo-guarantee for a well-defined mapping. For the first order transformation the constraint equations are:

$$t(1,0) + t(1,1) = 1 \qquad (4.7a)$$

$$t(0,0) + t(0,1) = 0 \qquad (4.7b)$$

$$t(0,0) - t(1,0) = -1 \qquad (4.7c)$$

$$t(0,1) - t(1,1) = 0 \qquad (4.7d)$$

Constraints (4.6a, b) determine the mapping along the ω_2 axis and the upper boundary $\omega_2 = \pi$. Hence, these four equations are not independent in the sense that any three imply the fourth. Solving in terms of the free variable t(0, 0) yields the three equations

$$t(1,0) = -t(0,0) \qquad (4.8a)$$

$$t(0,1) = 1 + t(0,0) \qquad (4.8b)$$

$$t(1,1) = -t(0,0). \qquad (4.8c)$$

Since the mapping of the vertical axis $\omega_1 = 0$ is fixed, making $x = \cos \omega$ a free variable is of no value. The approximation routine will force x = b, the length of the minor axis. Thus, the approximation problem is formulated to minimize $e(\omega_1)$ where

$$e(\omega_1) = H_d(\omega_1) - t(0,0)\,\varphi_1(\omega_1) \qquad (4.9)$$

and

$$H_d(\omega_1) = \cos b - \cos \omega_2 \qquad (4.10a)$$

$$\varphi_1(\omega_1) = 1 - \cos\omega_1 + \cos\omega_2 - \cos\omega_1\cos\omega_2 \qquad (4.10b)$$

$$\omega_2 = b(1 - (\omega_1/a)^2)^{\frac{1}{2}}. \qquad (4.10c)$$

See Appendix 3 for the technique involved in setting up the approximation problem. Figure 4.2 gives the frequency contours for the approximation of an ideal ellipse with major axis 0.5π and minor axis 0.25π. The approximation error from the ideal is effectively zero for $t(0,0) \approx -0.146$.

A length 63 one-dimensional filter was transformed using the above mapping. Figures 4.3a, b, c show the two-dimensional elliptical frequency response, impulse response, and transition band where the innermost contour represents a value of 0.9 and the outermost contour represents a value of 0.1.

The design of the elliptical contour under the above constraints, using the second order transformation, yields four free variables. Although the approximation cannot be improved since it was essentially perfect for the first order mapping, it is perhaps possible to compress the transition width using the additional free variables.

4.3 The Square Contour

With octagonal symmetry the ideal curve for the square can be expressed as

$$\omega_2 = g(\omega_1) = k \qquad (4.11)$$

where k is the length of a side of the square and ω_1 goes from zero to k.

The original first order mapping, having no free variables, is not suitable for this approximation problem. Alternatively, the first

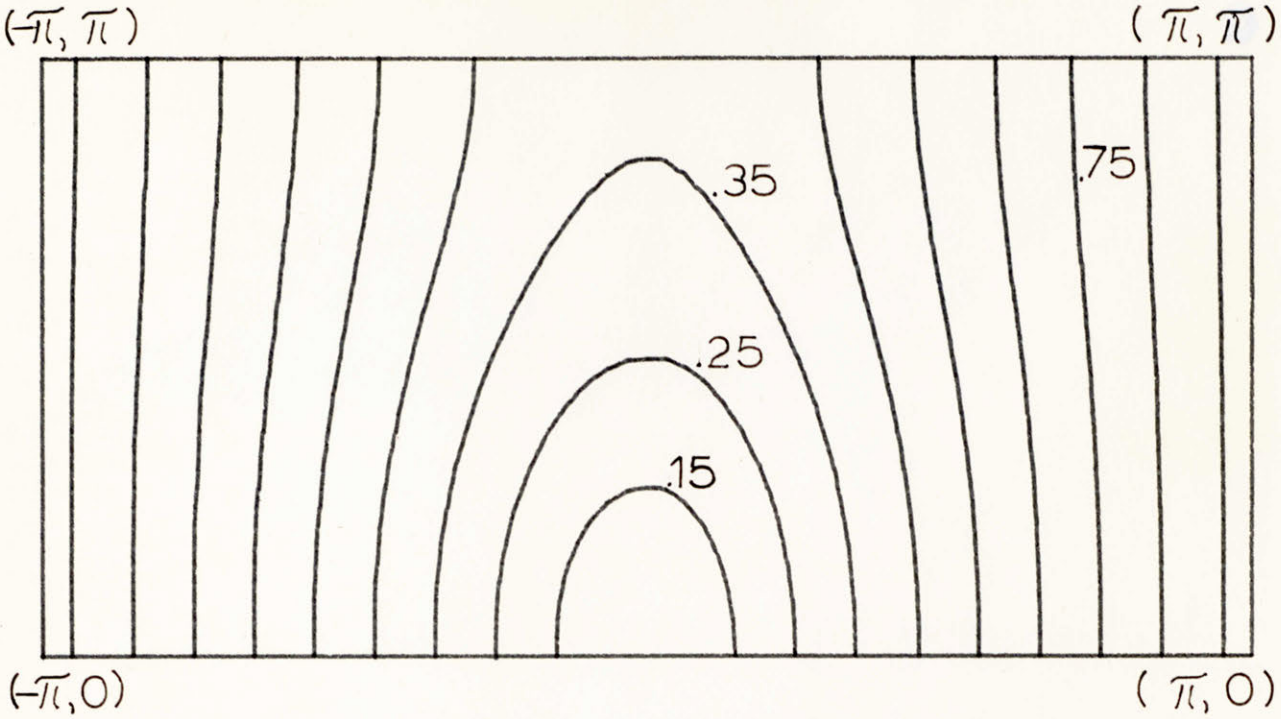

Fig. 4.2 The contours of the lowpass to elliptical filter first order transformation.

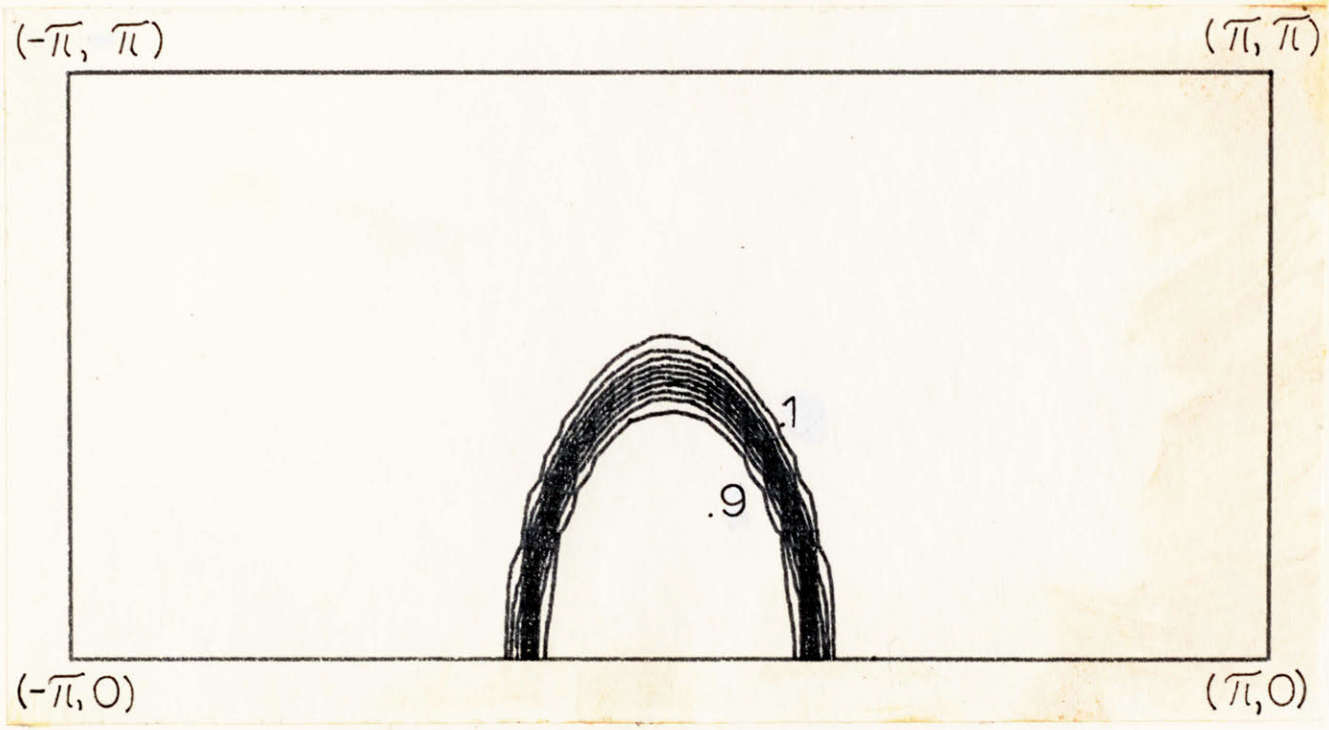

Fig. 4.3a The transition band gradient of the 63 x 63 elliptical filter corresponding to the contours of Fig. 4.2.

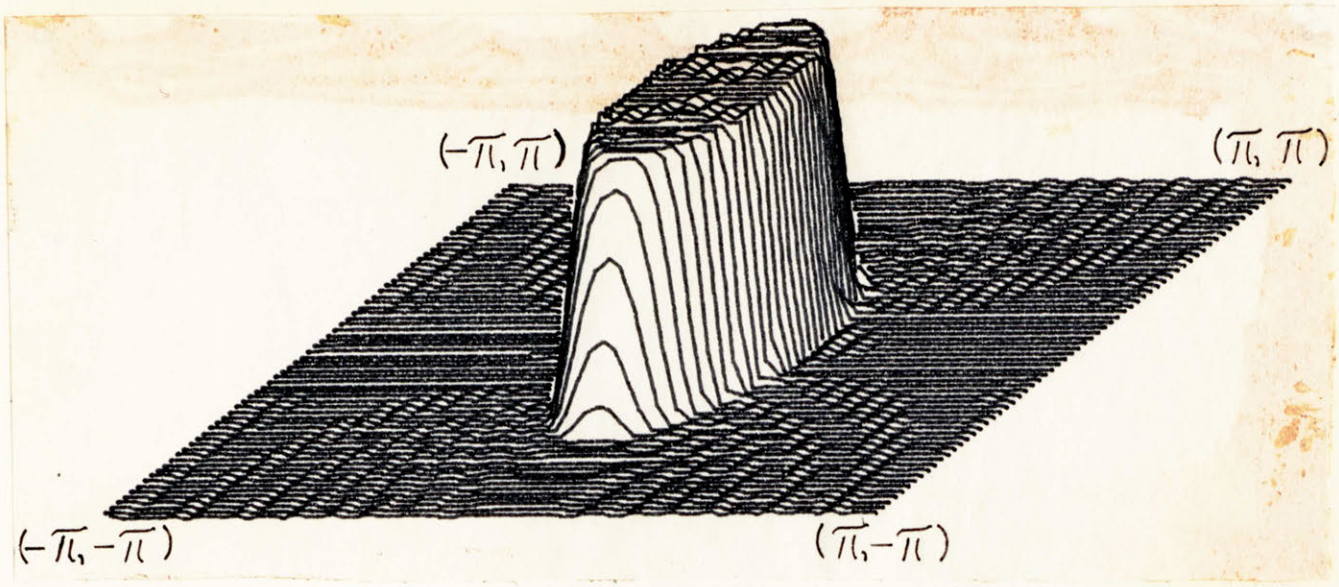

Fig. 4.3b The frequency response of the 63 x 63 elliptical filter.

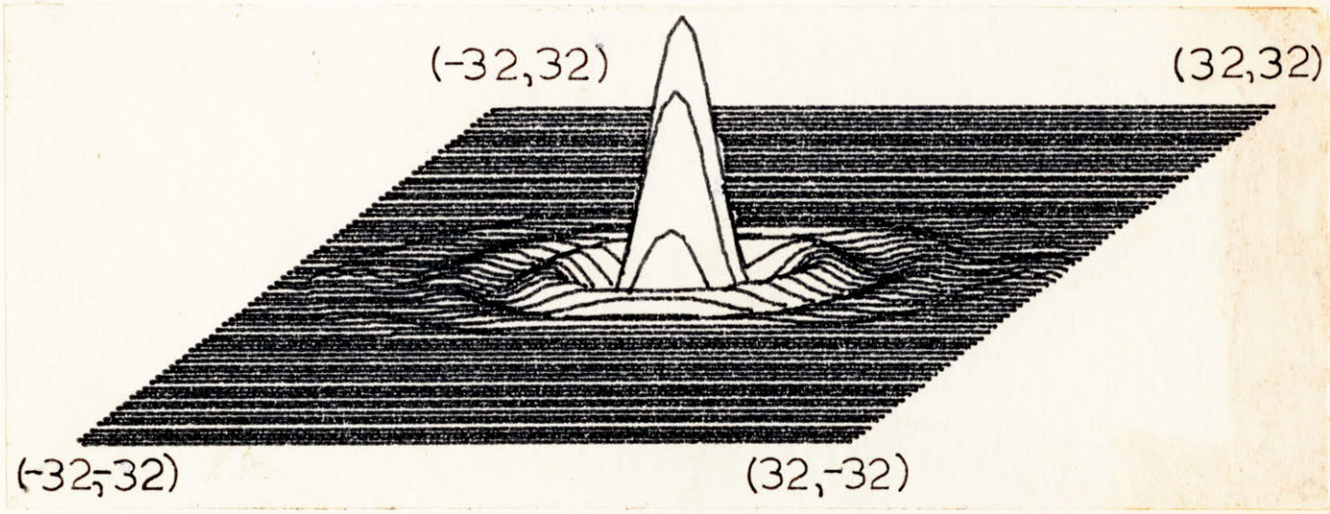

Fig. 4.3c The impulse response of the 63 x 63 elliptical filter.

order mapping under the constraints

$$0 \rightarrow (0, 0) \tag{4.12a}$$

$$\pi \rightarrow (\pi, \pi) \tag{4.12b}$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1) \tag{4.12c}$$

with cos ω fixed yields huge errors. For example, for $k = 0.6\pi$, a maximum error of $e(\omega_1) = 0.25$ was incurred. Allowing cos ω to be free led to degeneracies in both the least squares and minimax approximation problem.

The second order transformation under the constraints

$$0 \rightarrow (0, 0) \tag{4.13a}$$

$$\pi \rightarrow (\omega_1, \pi) \tag{4.13b}$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1) \tag{4.13c}$$

generates the constraint equations ($A = t(0, 0)$, $B = t(1, 0) \cdots I = t(2, 2)$)

$$C = B \tag{4.14a}$$

$$D = 0.5 \tag{4.14b}$$

$$E = F = -1 - A + B \tag{4.14c}$$

$$G = H = 0.5 - B \tag{4.14d}$$

$$I = 1.5 = A - 2B\,. \tag{4.14e}$$

Hence there are three free variables in the approximation. The ideal and basis functions are derived similarly to those of Appendix 3. Finally, with $x = \cos(0.6\pi)$ (that is, the ideal square has sides of length 0.6π), the contours of Fig. 4.4 were obtained. The ideal curve is drawn in dotted lines. A maximum error of $e(\omega_1) \approx 0.093$ was incurred. This is due to the incompatibility of the differentiable

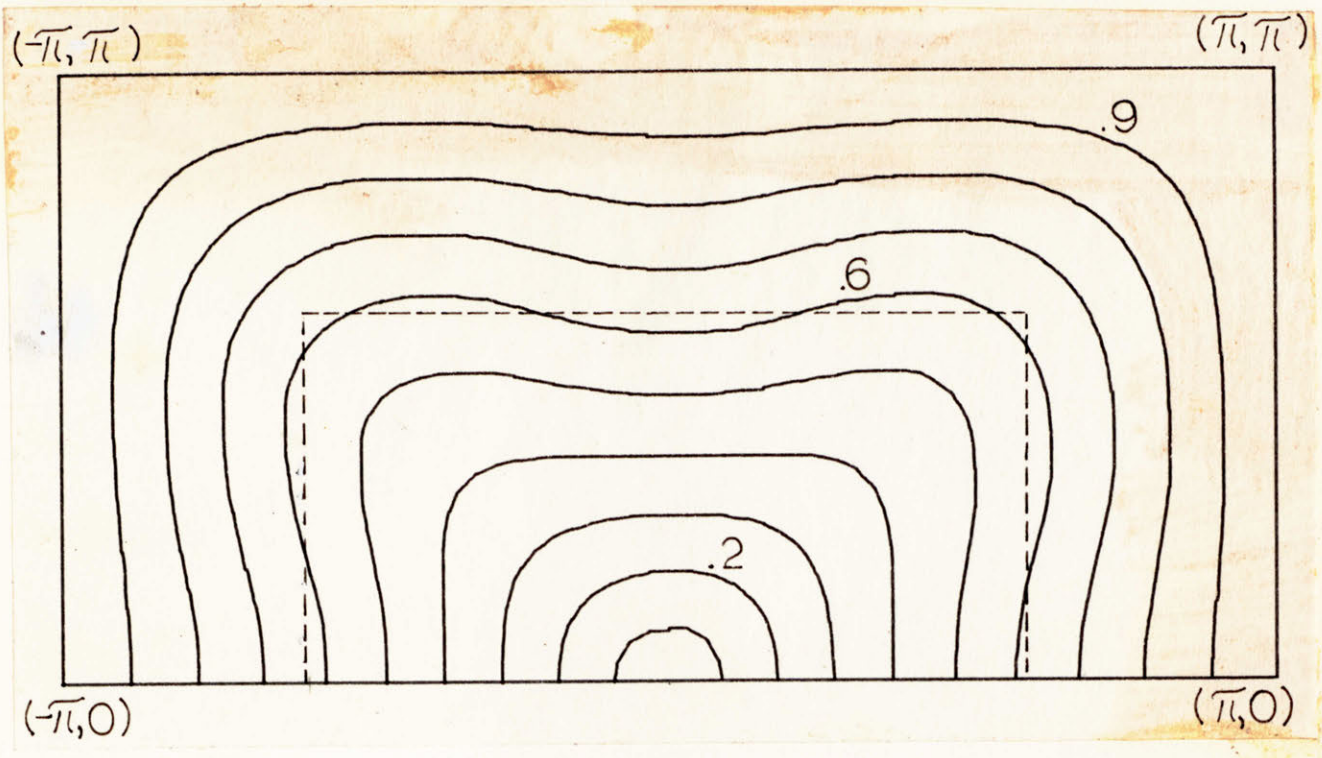

Fig. 4.4 The contours of the lowpass to square filter second order transformation.

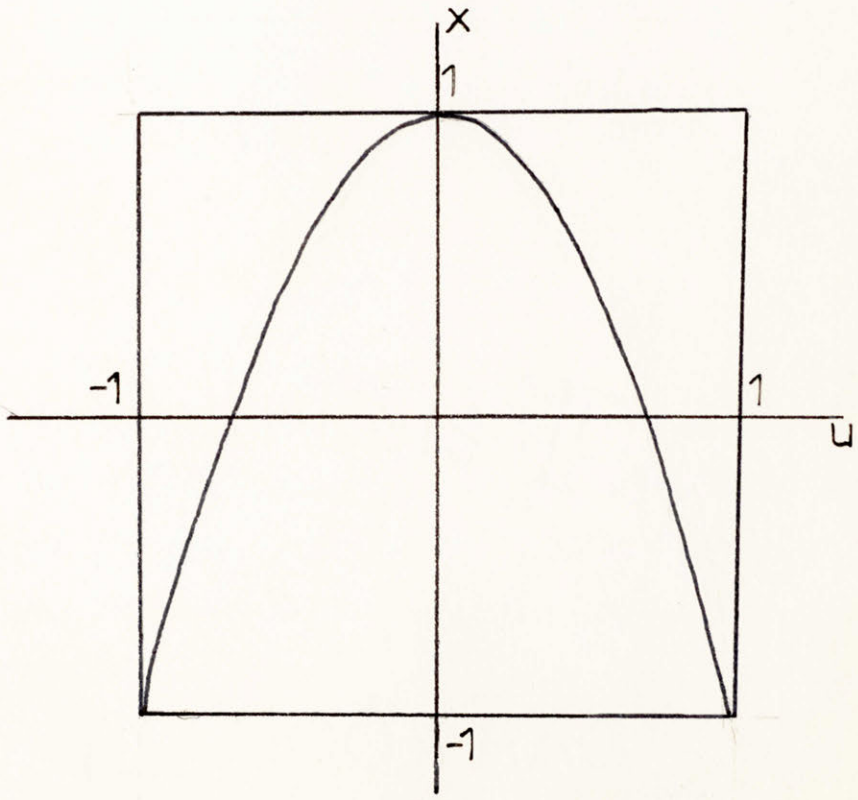

Fig. 4.5 The desired quadratic along the u axis for the annulus filter centered on the 0.5 π radius.

cosine-like basis functions and the non-differentiable ideal contour.

4.4 The Lowpass to Annulus Transformation

The second order transformation can be expressed as

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(\ell, k)\, u^{\ell} v^{k}.$$

Along the line $v = 1$, x can be written as the quadratic,

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(\ell, k)\, u^{\ell}. \tag{4.15}$$

If the coefficients $q(\ell, k)$ are constrained such that the quadratic in Fig. 4.5 results, then the two-dimensional filter takes on the appearance of a ring or annulus centered on the $0.5\,\pi$ radius. The polynomial in Fig. 4.5 can be determined by the constraints

$$0 \rightarrow (0.5\,\pi, 0) \tag{4.16a}$$

$$\pi \rightarrow (\omega_1, \pi),\ (0, 0) \tag{4.16b}$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1)\ . \tag{4.16c}$$

These constraints have the overall effect of mapping a lowpass filter to two regions, one region from the $0.5\,\pi$ radius to the outer boundaries, and the other from the $0.5\,\pi$ radius to the origin.

The second and third constraints can be implemented directly. The resulting constraint equations are ($A = t(0, 0)$, $B = t(1, 0), \cdots I = t(2, 2)$)

$$C = B \tag{4.17a}$$

$$D = 0 \tag{4.17b}$$

$$E = F = -1 - A + B \tag{4.17c}$$

$$G = H = -B \tag{4.17d}$$

$$I = 1 + A - 2B\,, \tag{4.17e}$$

where A and B are free variables. Constraint (4.16a) is made indirectly in approximating a circle of radius $0.5\,\pi$ for the mapping $x = \cos(0)$. Under the above constraint equations the approximation problem can be formulated to minimize

$$e(\omega_1) = H_d(\omega_1) - A\varphi_1(\omega_1) - B\varphi_2(\omega_1) \tag{4.18}$$

where $\varphi_1(\omega_1)$ and $\varphi_2(\omega_2)$ are linear functions of $\cos\omega_1$, $\cos\omega_2$, $\cos(2\omega_1)$, and $\cos(2\omega_2)$, and where $\omega_2 = [(0.5\,\pi)^2 - \omega_1^2]^{\frac{1}{2}}$ (see Appendix 3 for the derivation technique).

The resulting frequency contours are shown in Fig. 4.6a. A 61 x 61 ring-shaped filter was designed by transforming a length 31 one-dimensional filter with cutoff frequencies $\omega_p \approx 0.16\,\pi$ and $\omega_s \approx 0.3\,\pi$. The transition band contours and the surface plot are shown in Figs. 4.6b and 4.6c. The square-like nature of the outer contours suggests a better approximation might have been obtained if the contour design were performed on the outer stopband cutoff. This is because the contours inside the ideal radius tend toward circularity more than the contours outside this radius. This can be accomplished by finding the value of x in Fig. 4.5 which maps to the outer stopband radius ($u = \cos R_s$). The approximation problem is then reformulated to be compatible with this mapping.

We see in Fig. 4.6a that the number of contours in the region $(0,\pi) \times (0,\pi)$ has been doubled over that of the lowpass to lowpass transformation. Thus the transition width is reduced by approximately one

half. An optimum one-dimensional bandpass filter must be approximately twice as long as a lowpass with the same deviation and transition width. Since the second order lowpass to bandpass transformation compresses the transition width by about a factor of two, this approach seems preferable to transforming a one-dimensional bandpass filter. The advantage is not as great, however, for a bandpass centered at a larger radius. Here a quadratic sufficient to maintain a well-defined mapping is shown in Fig. 4.7. This can be accomplished with the constraints

$$\pi \rightarrow (0, 0), (\pi, \pi) \tag{4.19a}$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1) \tag{4.19b}$$

$$\left.\frac{\partial x}{\partial u}\right|_{u = u_0} = 0 \quad . \tag{4.19c}$$

It is evident from Fig. 4.7 that the contours will be spread out due to the increase in width of the quadratic function. Under the above constraints with $u_0 = \cos(0.7\pi)$ (an ideal radius of 0.7π), the frequency contours of Fig. 4.8 resulted. Here the bandpass to bandpass transformation might be preferred.

4.5 Fan Filters

The Fan filter of Chapter 2 (Fig. 2.1) can be derived for the first order transformation under the constraints

$$0 \rightarrow (\pi, 0) \tag{4.20a}$$

$$\pi \rightarrow (0, \pi) \tag{4.20b}$$

$$0.5\pi \rightarrow \omega_2 = \omega_1 \, . \tag{4.20c}$$

The constraints yield the equations

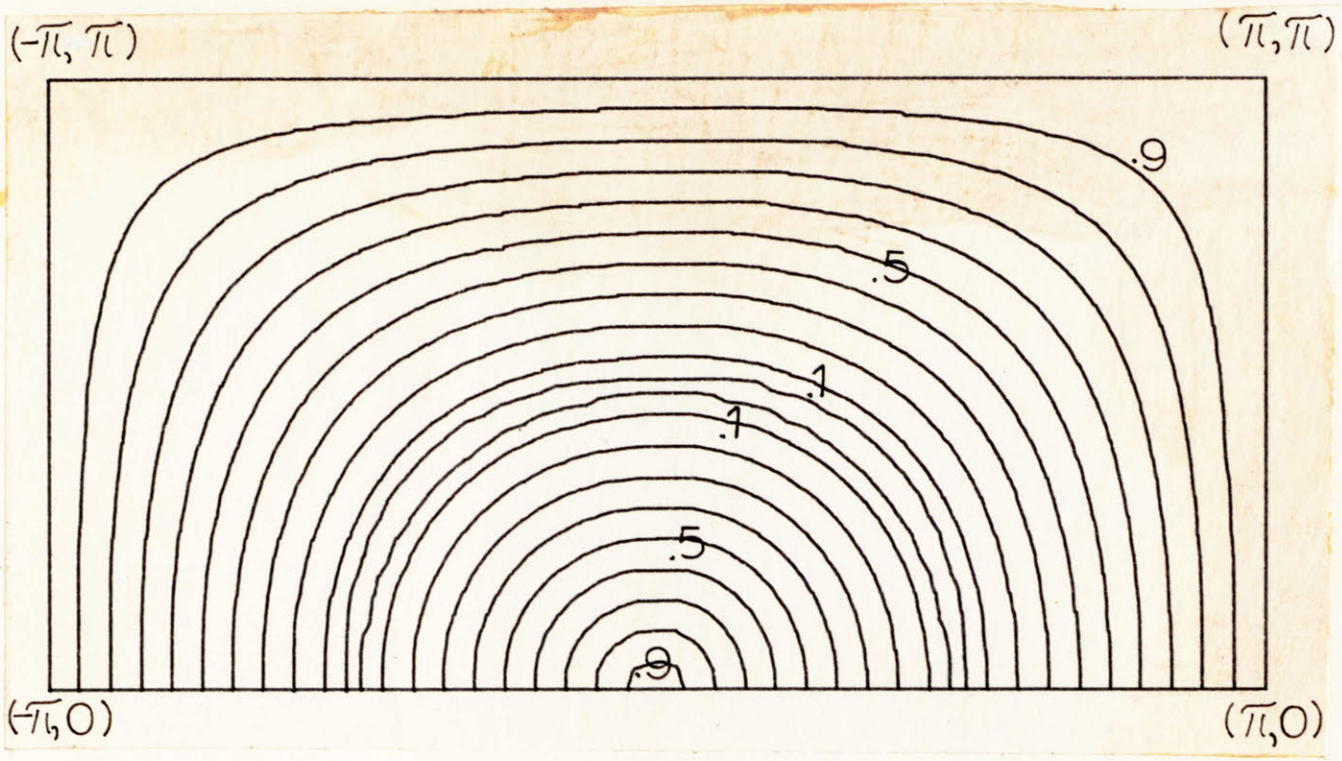

Fig. 4.6a The contours of the lowpass to annulus second order transformation (centered on the 0.5 π radius).

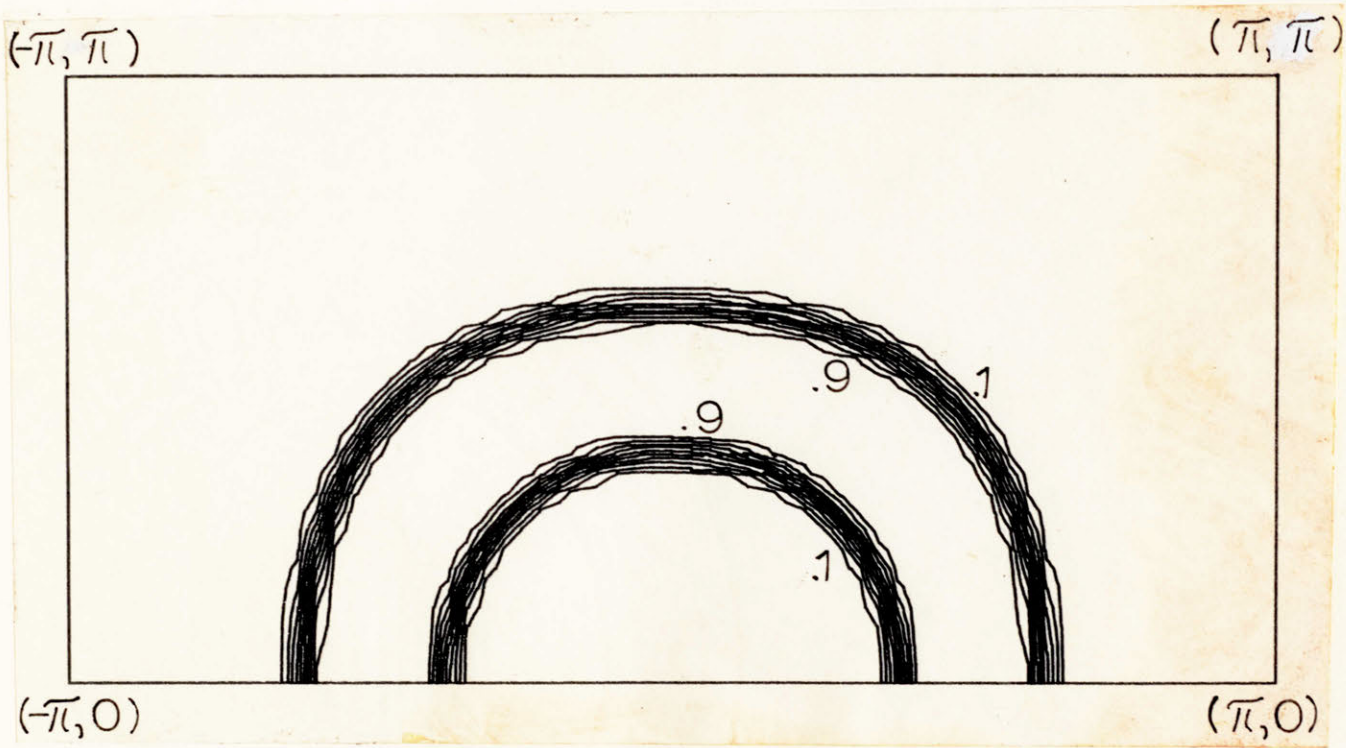

Fig. 4.6b The transition band gradients of a 61 x 61 annulus filter.

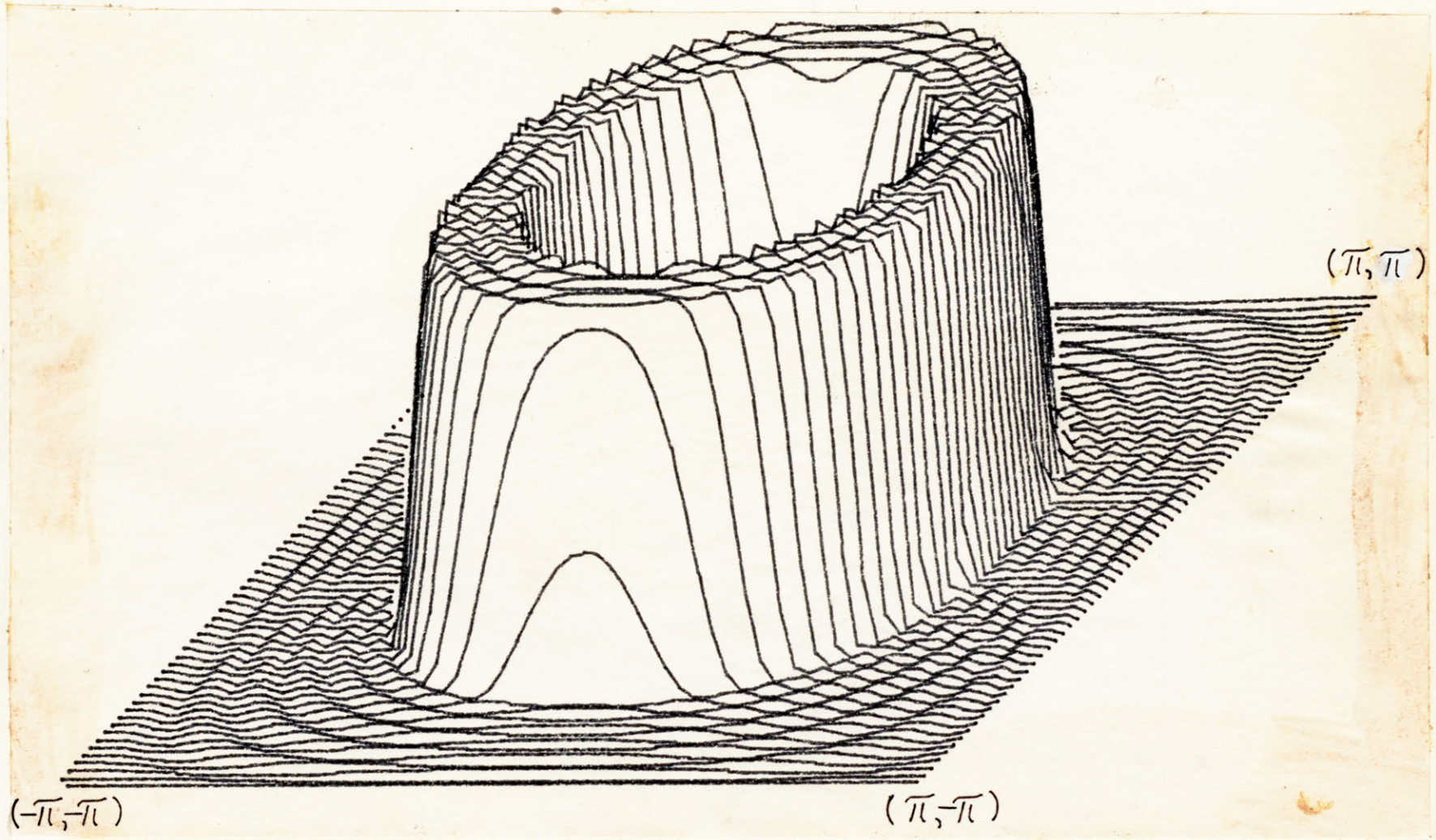

Fig. 4.6c The frequency response of the 61 x 61 annulus filter of Fig. 4.6b.

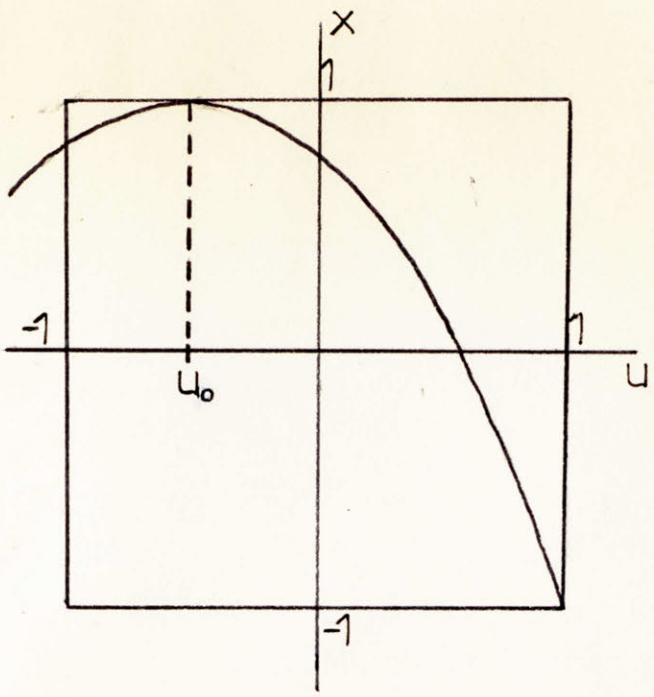

Fig. 4.7 The desired quadratic along the u axis for the general annulus filter.

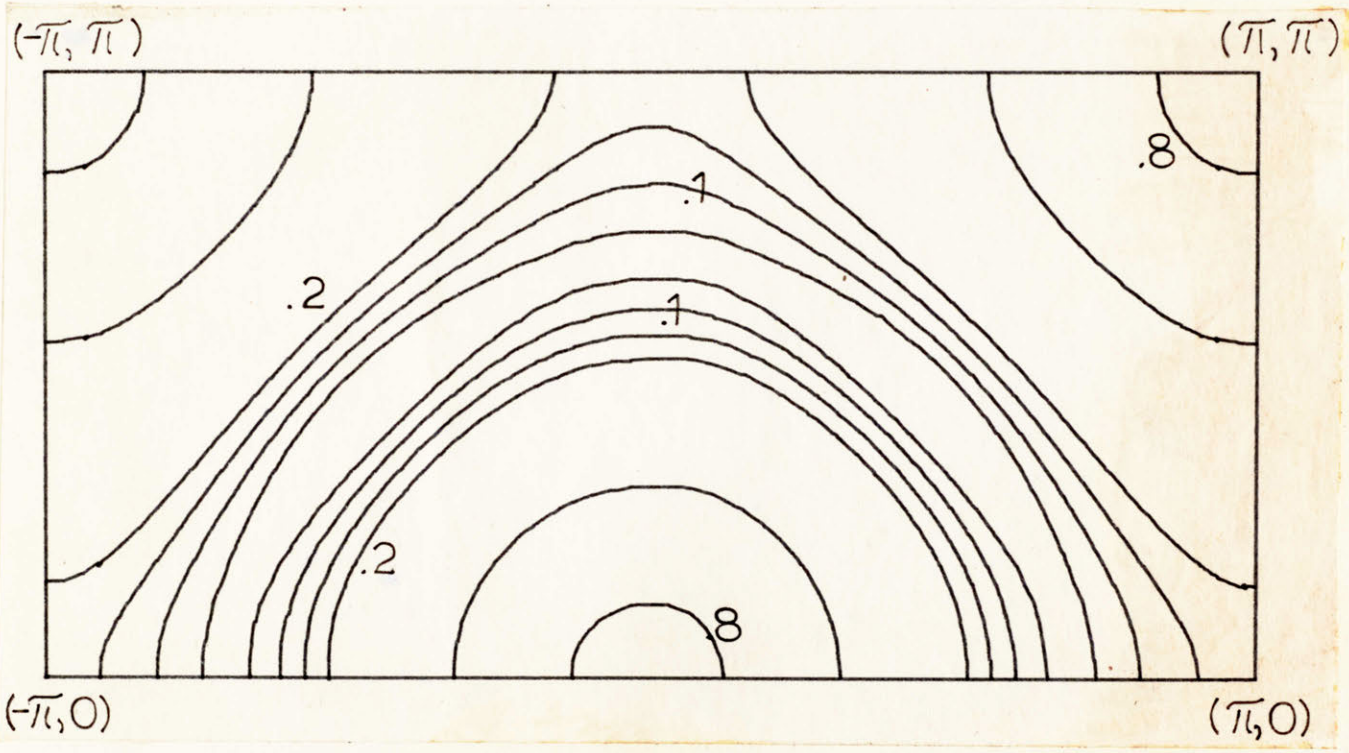

Fig. 4.8 The contours of the lowpass to annulus filter second order transformation (centered on the 0.7 π radius).

$$t(0,0) - t(1,0) + t(0,1) - t(1,1) = 1 \qquad (4.21a)$$

$$t(0,0) + t(1,0) - t(0,1) - t(1,1) = -1 \qquad (4.21b)$$

$$t(0,0) + t(1,0)\cos\omega_1 + t(0,1)\cos\omega_1 + t(1,1)\cos^2\omega_1 = 0 \qquad (4.21c)$$

Equation 4.21c implies that $t(0,0) = t(1,1) = 0$ and that $t(1,0) + t(0,1) = 0$. Hence two equations with two unknowns develop:

$$t(0,1) - t(1,0) = 1 \qquad (4.22a)$$

$$t(1,0) + t(0,1) = 0 \qquad (4.22b)$$

Thus, $t(1,0) = -t(0,1) = -0.5$.

An alternative method is the formulation of an approximation problem where the ideal curve is described by $\omega_2 = g(\omega_1) = k\omega_1$. k is the slope of the ideal passband edge; hence, variable angle filters can be approximated. In addition, the free variables of the second order transformation allow transition width compression at the same time the approximation is made.

With the second order transformation under the constraints

$$0 \rightarrow (\pi, 0) \qquad (4.23a)$$

$$\pi \rightarrow (\omega_1, \pi) \qquad (4.23b)$$

the ideal curve $\omega_2 = 0.5\,\omega_1$ was approximated with the hope that the constraint 4.23b would spread the contours to produce a well-defined mapping. This approximation formulation is similar to that of previous sections. The resulting contours are shown in Fig. 4.9. Although the approximation is very good, ill-defined regions are generated.

To avoid this problem, the increasing monotonicity constraint

$$\frac{\partial v}{\partial u} \geq 0 \qquad (4.24)$$

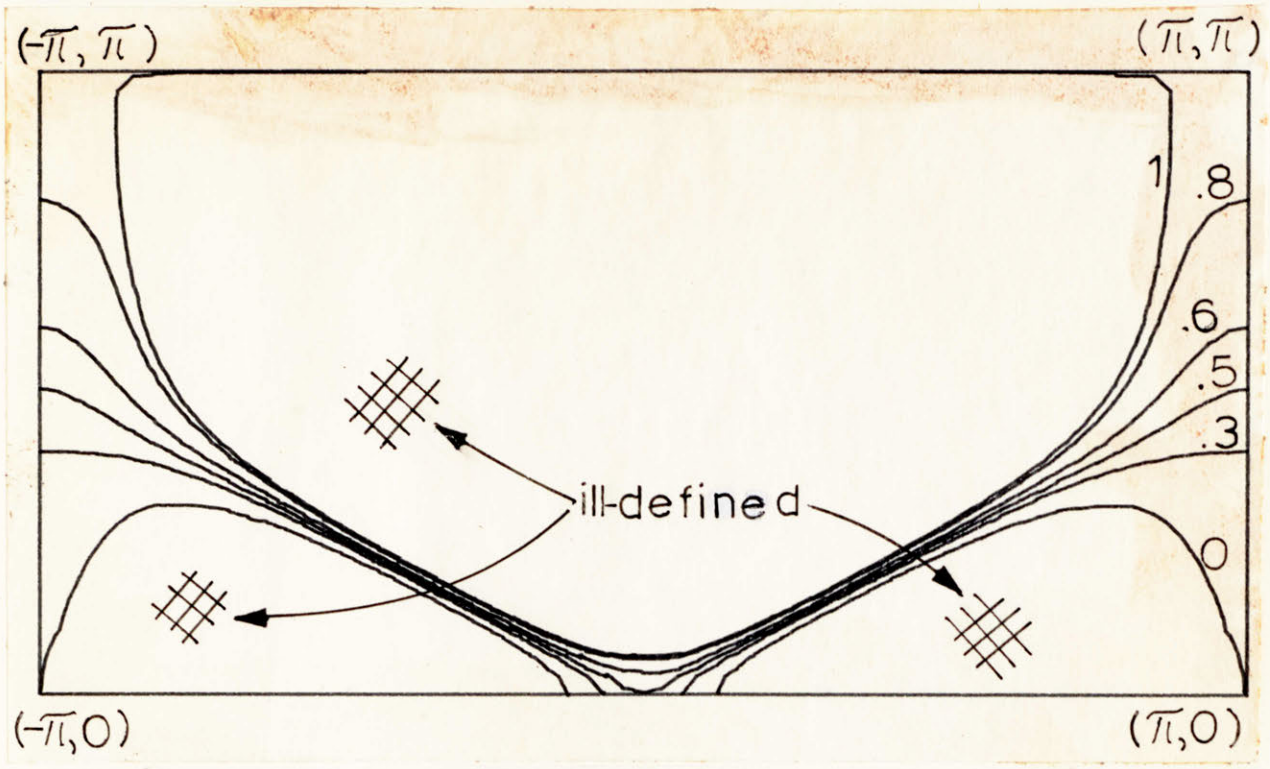

Fig. 4.9 The fan filter contours designed without constraints for a well-defined mapping.

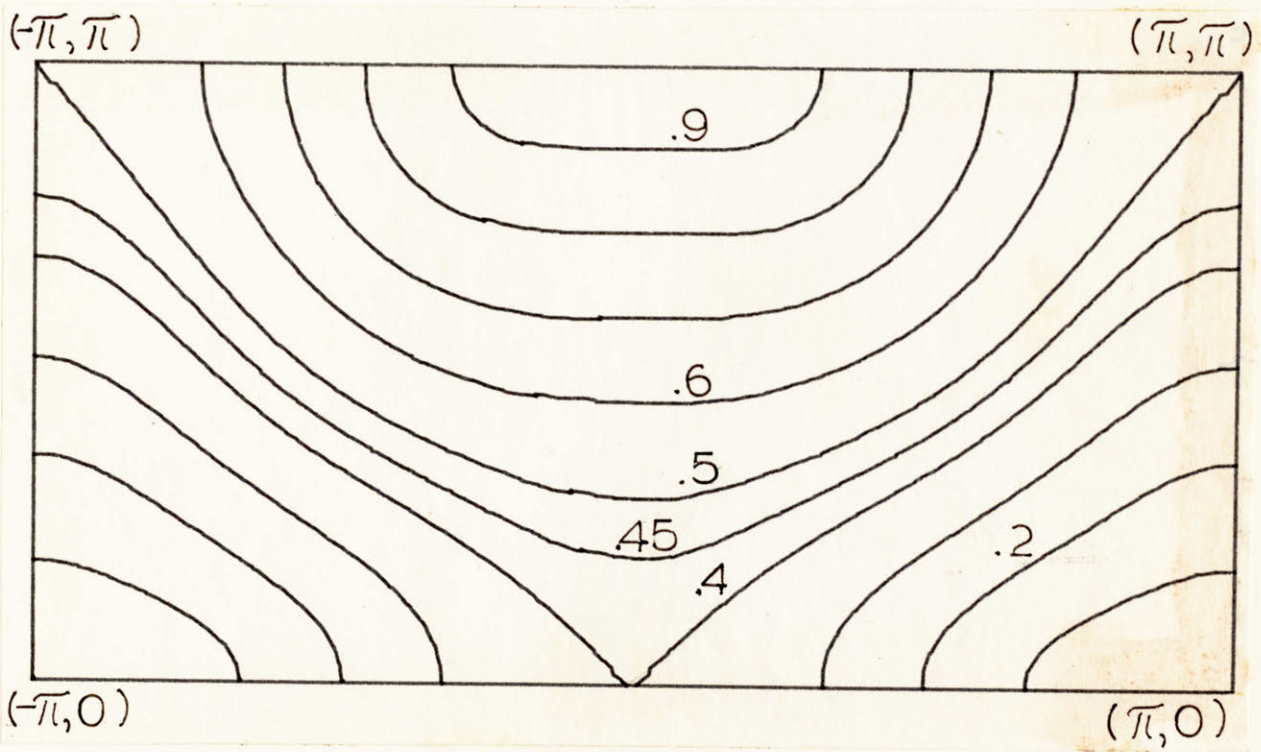

Fig. 4.10 The fan filter contours designed by linear programming with constraints for a well-defined mapping.

was imposed over the sparse grid of Fig. 3.3. We can argue that these constraints imply a pseudo-guarantee for a well-defined mapping when the additional constraints (4.23a, b) are made. The utility of linear programming was demonstrated for this example (see Appendix 5). The constraint

$$0.5\,\pi \rightarrow (0, 0.3\,\pi), (\pi, \pi) \tag{4.25}$$

was also imposed to restrain the transition width. The transition width was found to be very large without this constraint. Figure 4.10 shows the resulting contours for an ideal passband edge $\omega_2 = 0.7\,\omega_1$. Hence, although the approximation has been slightly degraded, the ill-defined regions of Fig. 4.9 have been eliminated. More study might be made in the application of inequality constraints to the approximation problem.

4.6 The Bandpass Filter

A two-dimensional bandpass filter is generated by modulating a two-dimensional narrowband lowpass filter. The modulation can be expressed as

$$g(n, m) = 4h(n, m) \cos(\omega_x n) \cos(\omega_y m)\,. \tag{4.26}$$

This implies the frequency spectrum

$$\begin{aligned} G(\omega_1, \omega_2) = {} & H(\omega_1 + \omega_x, \omega_2 + \omega_y) + H(\omega_1 + \omega_x, \omega_2 - \omega_y) \\ & + H(\omega_1 - \omega_x, \omega_2 + \omega_y) + H(\omega_1 - \omega_x, \omega_2 - \omega_y)\,. \end{aligned} \tag{4.27}$$

A narrowband lowpass 63 x 63 filter was designed by performing the original first order transformation on a length 63 one-dimensional filter with cutoff frequencies $\omega_p \approx 0.1\,\pi$ and $\omega_s \approx 0.2\,\pi$. The surface

plot is shown in Fig. 4.11.

4.7 Miscellaneous Contours

To further show the versatility of the technique, a number of other examples were designed.

With the ideal curve $\omega_2 = 0.6\pi - \omega_1$, the diamond contours of Fig. 4.12 were designed. The first order transformation coefficients are $t(0,0) = -t(1,1) = 0.3548$ and $t(1.0) = t(0,1) = 0.5$. Constraints (2.21) were used.

Interchanging the signs of the coefficients $t(0,0)$ and $t(1,1)$ of the original McClellan mapping generates the convex contours of Fig. 4.13a. An example is the 63 x 63 order filter of Fig. 4.13b.

Two additional contour designs are shown in Figs. 4.14 and 4.15.

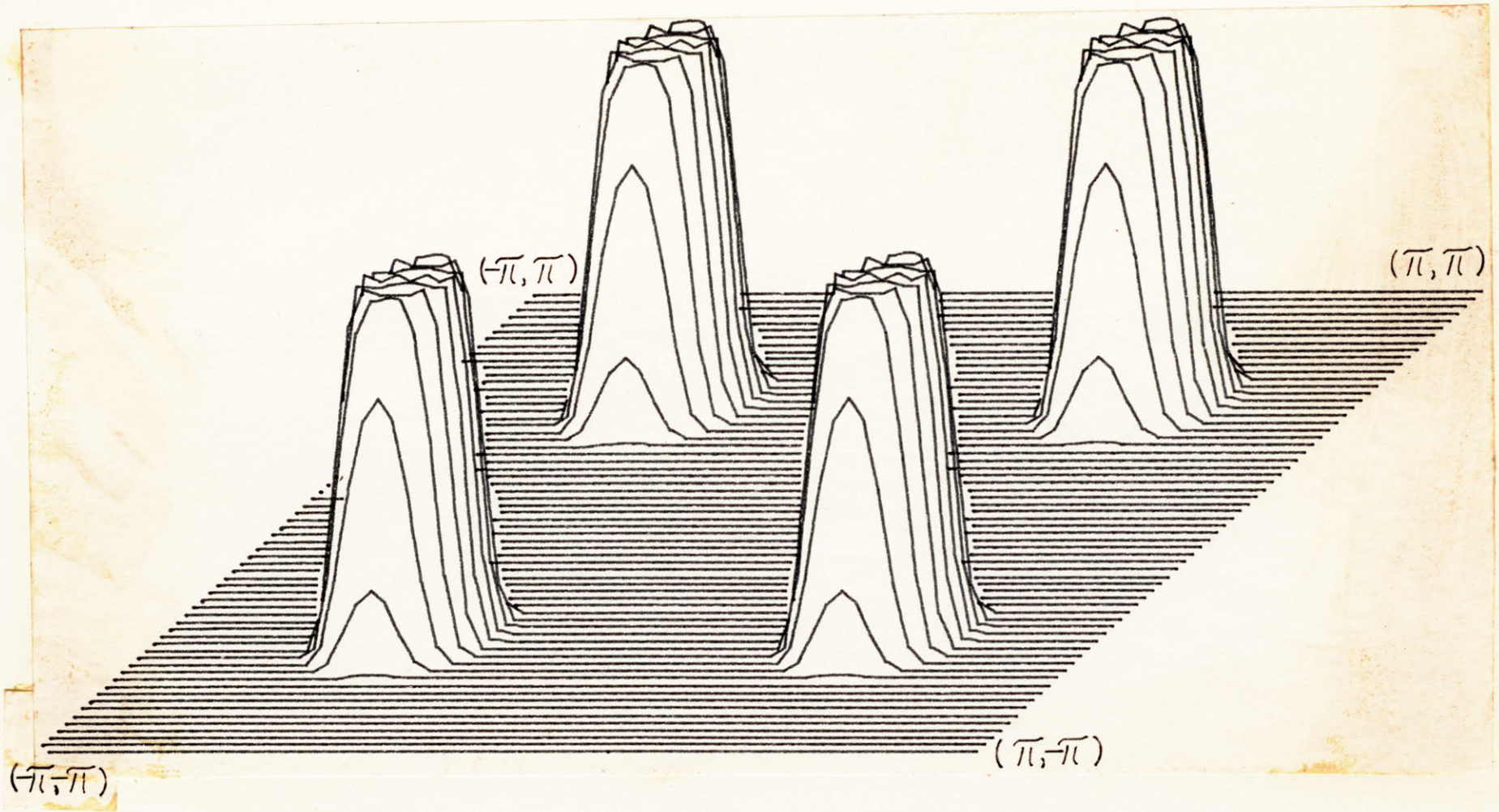

Fig. 4.11 The frequency response of the modulated lowpass filter.

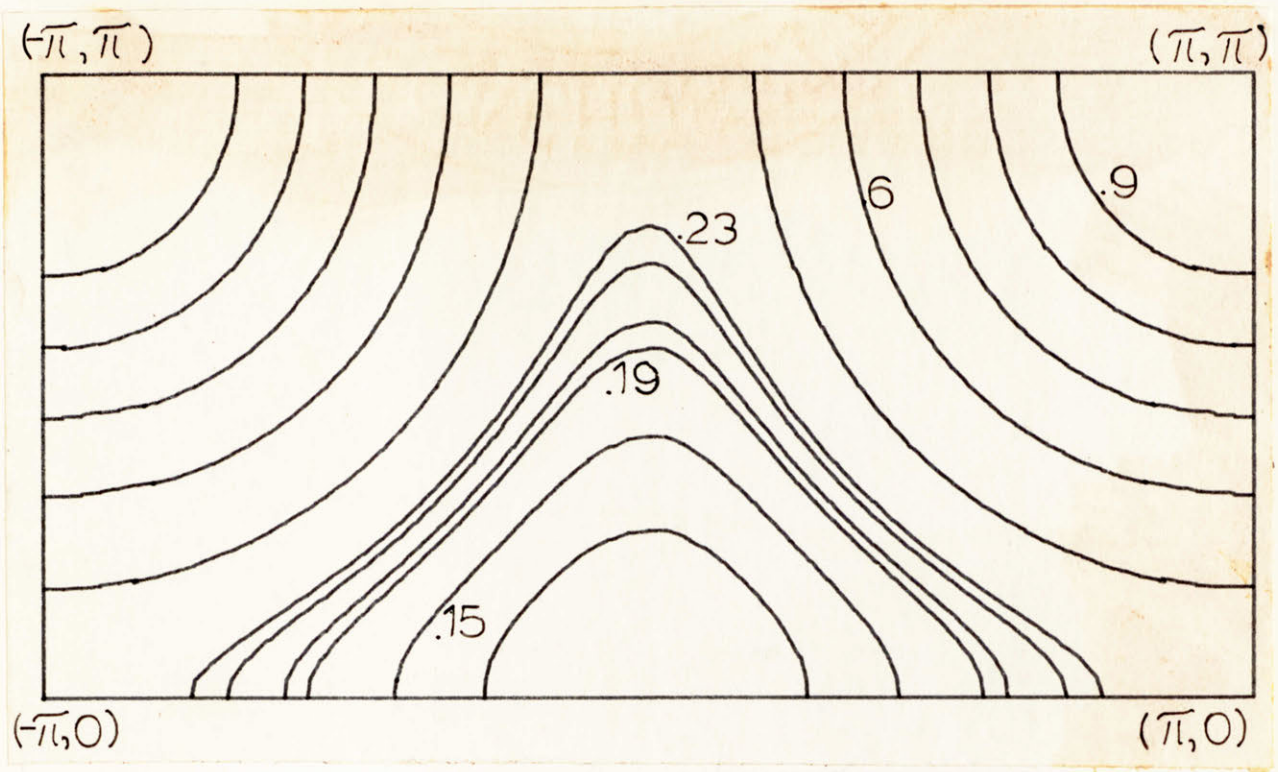

Fig. 4.12 The contours of a diamond-like filter.

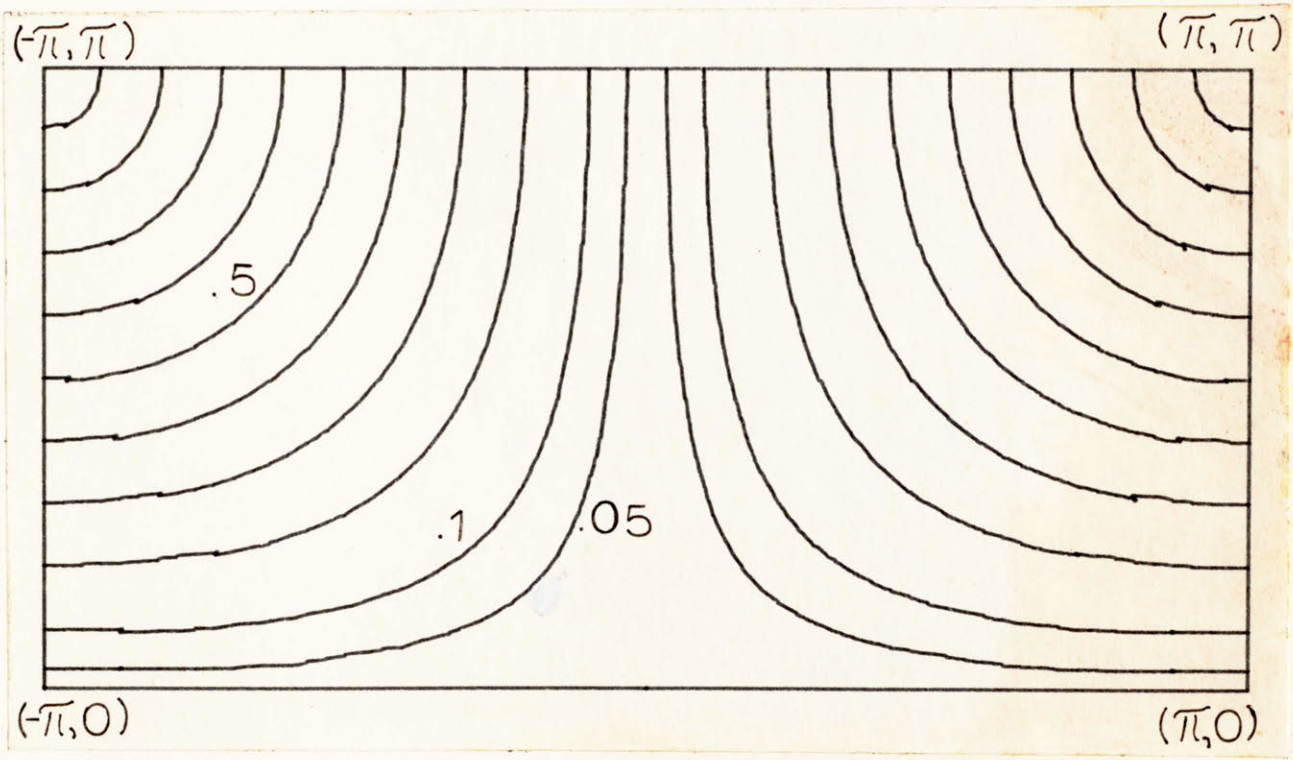

Fig. 4.13a The convex contours generated from a modification of the original McClellan transformation.

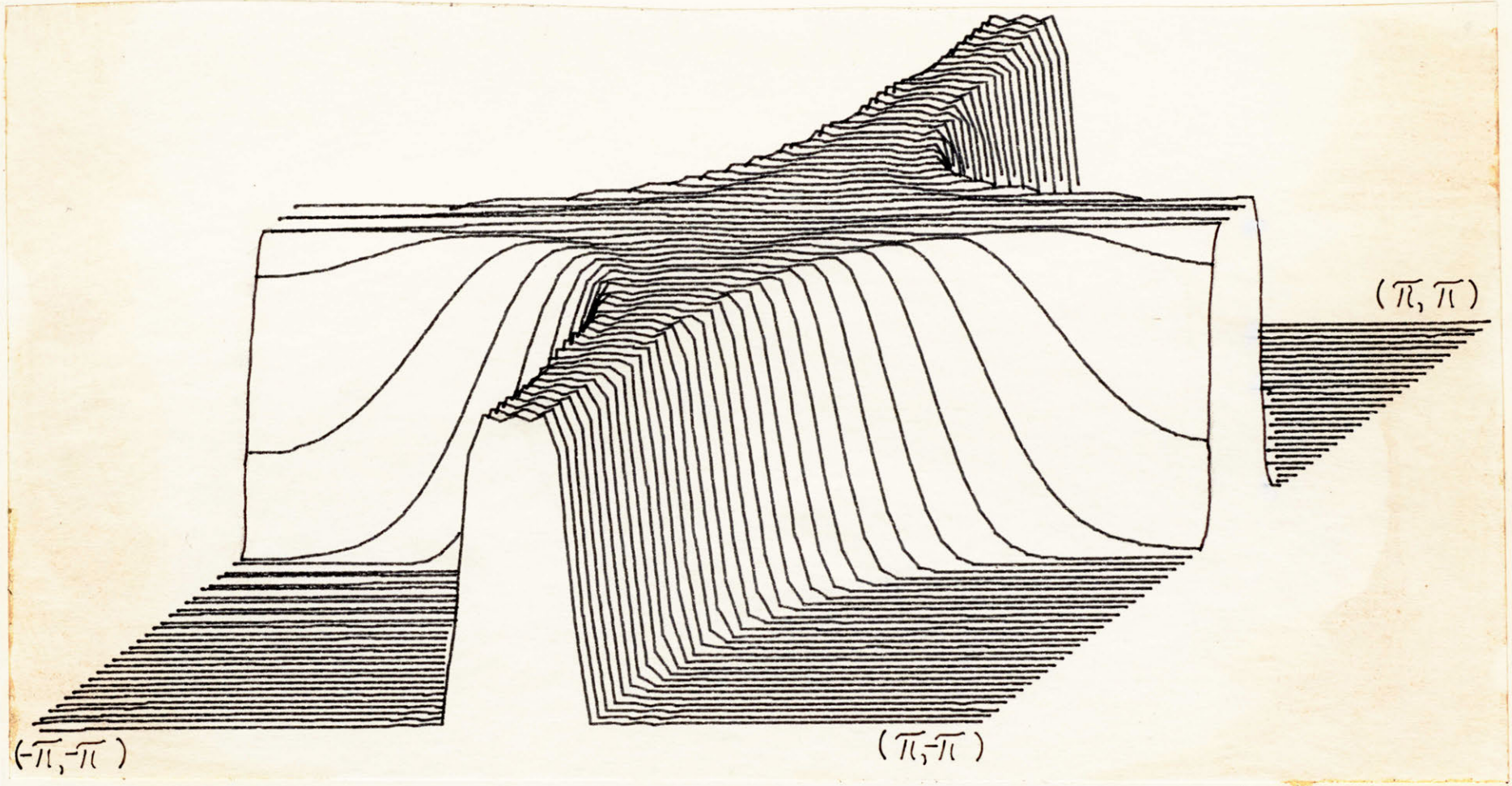

Fig. 4.13b The frequency response of a 63 x 63 filter corresponding to the contours of Fig. 4.13a.

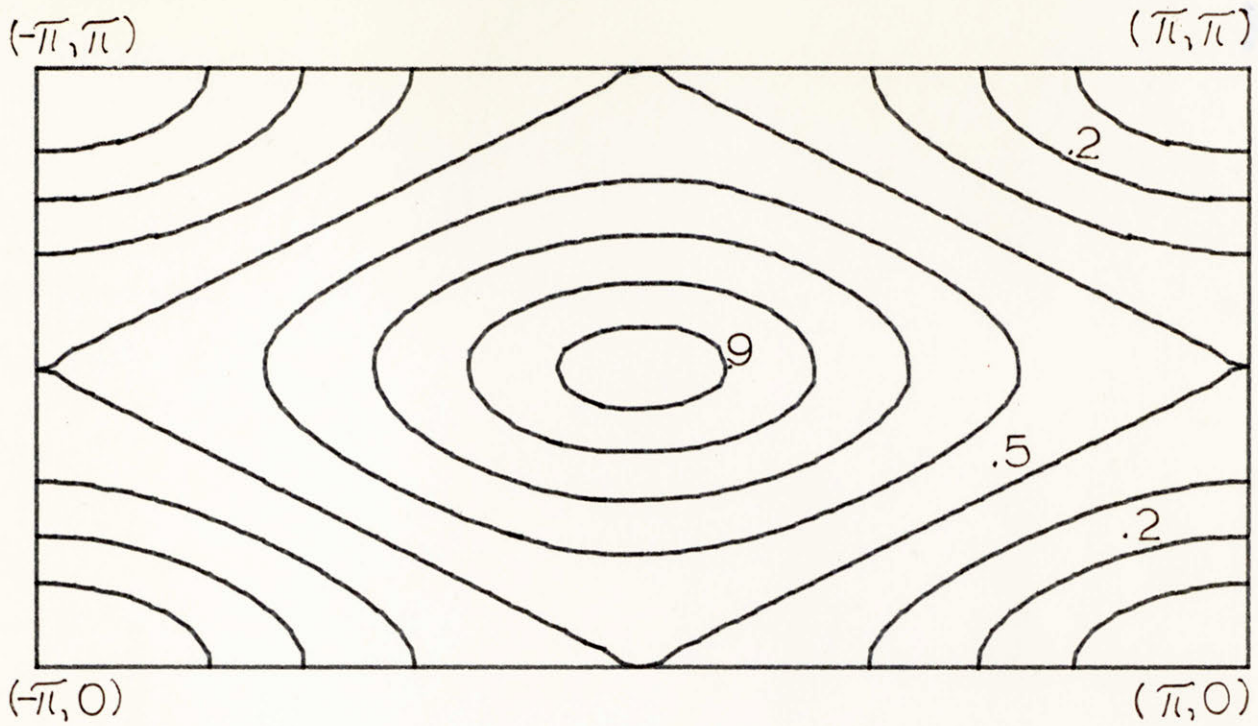

Fig. 4.14 Arbitrary contours.

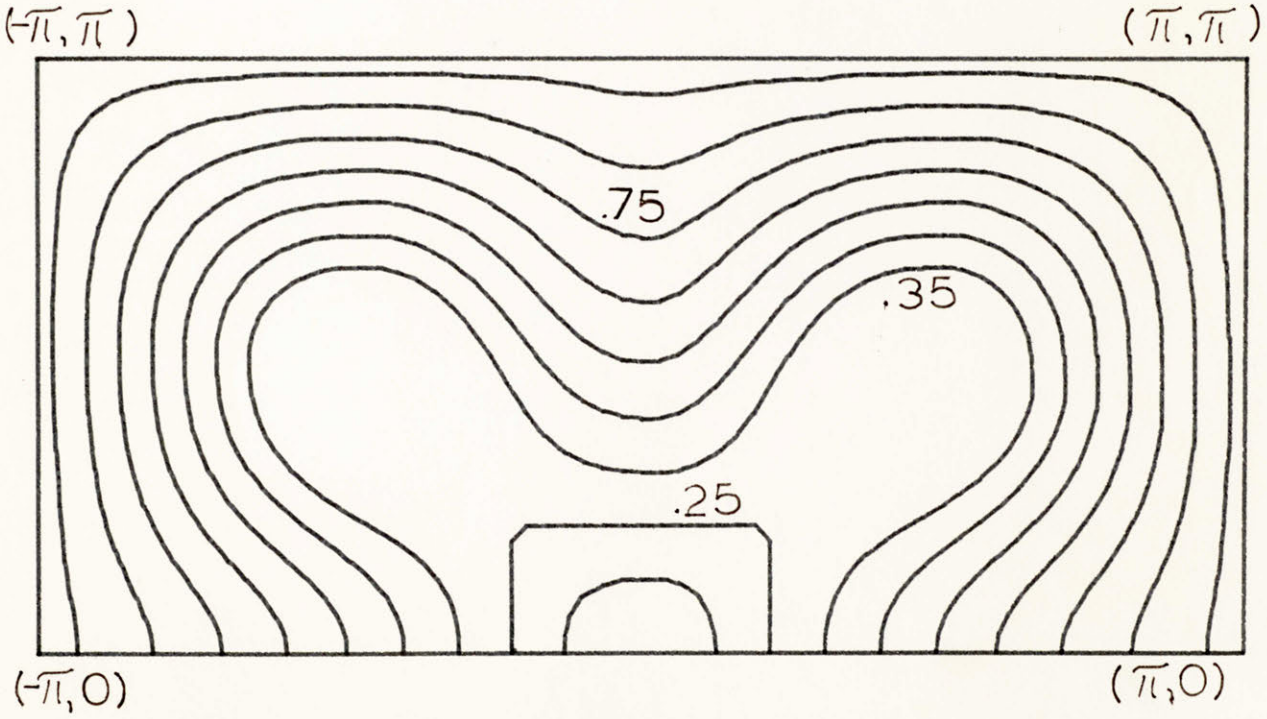

Fig. 4.15 Arbitrary contours.

CHAPTER 5

OTHER TECHNIQUES AND TRANSFORMATIONS

5.1 Comparisons with the Optimal Design

McClellan has suggested that the original first order transformation yields optimal minimax two-dimensional designs when transformed from optimal minimax one-dimensional designs. The argument is that along the ω_1 axis

$$H(\omega_1, \omega_2) = \sum_{\ell=0}^{n} \sum_{k=0}^{n} b(\ell, k) \cos(\omega_1 \ell) \qquad (5.1)$$

and thus since the number of degrees of freedom here equals that of the one-dimensional design, and because the frequency response along the ω_1 axis is indeed the one-dimensional filter, the error deviation along this axis is the smallest possible. Hence, the minimax error over the region $(0, \pi) \times (0, \pi)$ is the smallest possible when McClellan's cutoff contours match those of the optimum. The theoretical validity of this assertion is still in question. This section nevertheless presents some empirical comparisons with optimal designs.

The above argument suggests that higher order mappings are not necessarily optimum in the minimax sense. Along the ω_1 axis

$$H(\omega_1, \omega_2) = \sum_{\ell=0}^{nM} \sum_{k=0}^{nM} b(\ell, k) \cos(\omega_1 \ell) \, . \qquad (5.2)$$

If $H(\omega)$ is mapped to the ω_1 axis, then the number of ripples along ω_1 is not sufficient to guarantee an optimal design along this axis. We see this is because equation (5.2) has, in general, nM degrees of freedom, whereas the original one-dimensional design has n degrees of freedom.

A comparison was made with optimal two-dimensional circularly symmetric lowpass filters of order 5 x 5, 7 x 7, 9 x 9, and 11 x 11 designed by Harris with a novel technique under the minimax error criterion. The two-dimensional specifications are $R_p = 0.4\pi$ and $R_s = 0.6\pi$. The comparisons are made with the original McClellan transformation and the first order transformation with $\cos\omega$ a free variable.

The results are summarized in the tables shown in Figs. 5.1a, b, c, d. The second row in each table (Mc. 1) gives the one-dimensional passband and stopband frequencies necessary for tangent cutoff contours (see section 3.4) for the original McClellan transformation. We see that in all four cases the one-dimensional and hence two-dimensional passband and stopband deviation ($\delta_p = \delta_s = \delta$) is greater than that of the optimal two-dimensional deviation. The maximum difference in deviation is approximately 0.02 or about 35% greater for the large order filters.

If, instead of meeting the two-dimensional specifications exactly, we use the two-dimensional radii 0.4π and 0.6π as the one-dimensional passband and stopband cutoff frequencies, the parameters of row three (Mc. 2) are generated. The deviation is actually less than the optimal in all four cases. As we see in Fig. 2.3, the contours 0.4π and 0.6π are near circular. If this circularity deviation is tolerated, better than the optimal deviation can be achieved. However, a fairer comparison is to design the optimal two-dimensional filter over the contours of the McClellan transformation and then compare the resulting deviation with the deviation of the transformed filter. This technique would serve to test McClellan's hypothesis as well as make a more valid comparison.

	ω_p	ω_s	δ
optimal	-	-	0.2670
Mc. 1	0.4000	0.5760	0.2852
Mc. 2	0.4000	0.6000	0.2639
A. 1	0.4000	0.5758	0.2854

Fig. 5.1a Deviation comparison for 5 x 5 two-dimensional filters.

	ω_p	ω_s	δ
optimal	-	-	0.1269
Mc. 1	0.4000	0.5760	0.1278
Mc. 2	0.4000	0.6000	0.1129
A. 1	0.4000	0.5758	0.1279

Fig. 5.1b Deviation comparison for 7 x 7 two-dimensional filters.

	ω_p	ω_s	δ
optimal	-	-	0.1141
Mc.1	0.4000	0.5760	0.1334
Mc.2	0.4000	0.6000	0.1129
A.1	0.4000	0.5758	0.1336

Fig. 5.1c Deviation comparison for 9 x 9 two-dimensional filters.

	ω_p	ω_s	δ
optimal	-	-	0.0569
Mc.1	0.4000	0.5760	0.0704
Mc.2	0.4000	0.6000	0.0508
A.1	0.4000	0.5758	0.0705

Fig. 5.1d Deviation comparison for 11 x 11 two-dimensional filters.

We can also test optimality in a more formal way with the Characterization Theorem (10).

Row four (A. 1) presents the one-dimensional parameters for the first order transformation with cos ω a free variable. We see that in all four cases the deviation is slightly larger (approximately 0.0001 or 0.07%) than the original McClellan transformation with specifications satisfied (Mc. 1). This is due to the widening of the transition band, which is done to improve the contour shape.

Finally, a comparison of implementation requirements shows a very significant improvement. For example, an 11 x 11 optimal filter requires approximately 21 multiplies per output point, while the original McClellan transformation generates a filter of the same order requiring only 11 multiplies per output point. This is a considerable reduction in multiplies for a negligible increase in deviation.

5.2 Windowing

In this section we present comparisons between two-dimensional filters designed from transforming one-dimensional windowed filters and filters designed by windowing directly in two dimensions.

Figure 5.2 shows a surface plot of a 63 x 63 two-dimensional filter generated from the original McClellan mapping. The one-dimensional prototype was designed by applying a rectangular window to the impulse response of an ideal lowpass filter with cutoff frequency 0.45 π. The one-dimensional impulse response is expressed as

$$h(n) = \begin{cases} \dfrac{\sin(0.45\,\pi n)}{\pi n}, & |n| \leq 31 \\ 0, & |n| > 31 . \end{cases} \qquad (5.3)$$

Note the Gibbs phenomenon behavior near the transition edges in Fig. 5.2.

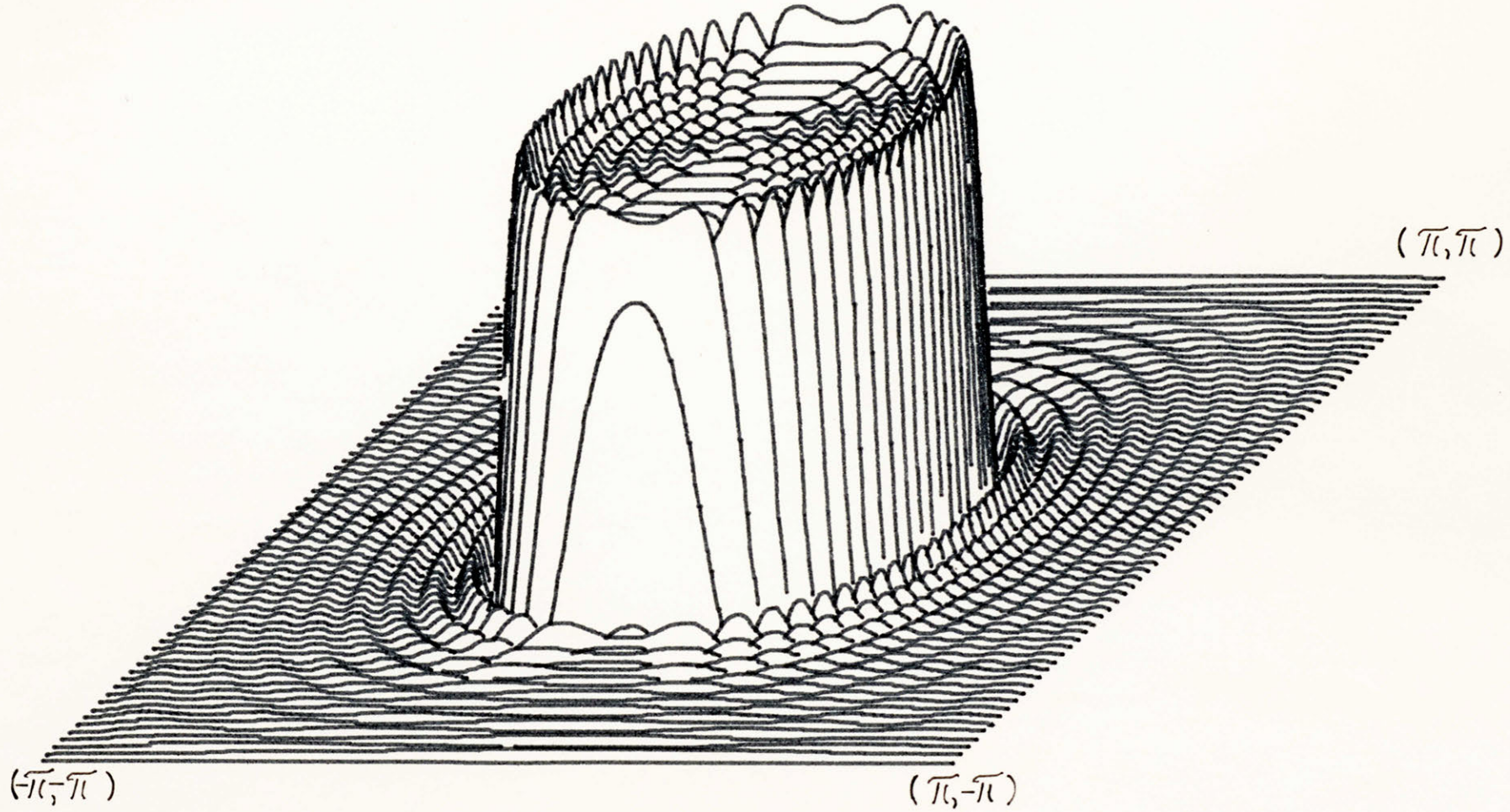

Fig. 5.2 The frequency response of a 63 x 63 filter generated from the transformation of a filter designed with a rectangular window.

The mean squared error for this filter from the ideal is defined as

$$\epsilon = \sum_{\ell=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |h(\ell,k) - h_d(\ell,k)|^2$$

$$= \sum_{\ell=-n}^{n} \sum_{k=-n}^{n} |h(\ell,k) - h_d(\ell,k)|^2 \qquad (5.4)$$

$$+ \sum_{|\ell|>n} \sum_{|k|>n} |h_d(\ell,k)|^2 .$$

We find that the mean squared error is approximately 3.2×10^{-3}. The mean squared error for the filter designed from windowing the two-dimensional impulse response directly is simply the second term in the sum of equation (5.4). This yields about 2.0×10^{-3}.

We can show therefore that the error performance of the transformed 63 x 63 filter is comparable to a 39 x 39 filter designed by windowing directly in two dimensions. The number of multiplies to implement the 39 x 39 by direct convolution is approximately 211, whereas the number of multiplies required for the transformed 63 x 63 is approximately 63. If the FFT is used the 39 x 39 is, however, more efficient.

The comparison here is not entirely fair since contour deviations introduce large errors in the transition bands. This error probably outweighs the passband and stopband deviation error due to the very high order of the filter. Perhaps a fairer comparison can be made with an ideal filter of order 21 x 21 and passband cutoff of radius 0.3π. Both the transformed and the directly windowed filters were designed. The mean squared error for the direct window design is 4.407×10^{-3}. The mean

squared error for the transformed window design is 5.117×10^{-3}. We can show that the error performance of this filter is equivalent to a direct window design of order 9 x 9. The number of required multiplies per output point for the 9 x 9 is about 15, whereas the transformation implementation requires 21 multiplies per output point. These two results suggest that it is more efficient to perform the windowing operation in two dimensions rather than in one dimension.

Similarly, the continuous Kaiser window

$$\omega(t) = \begin{cases} \dfrac{I_o[\alpha(1-(t/T)^2)]}{I_o(\alpha)}, & |t| < T \\ 0 & |t| > T \end{cases} \qquad (5.5)$$

was sampled and applied to the impulse response of an ideal lowpass filter of cutoff frequency 0.3π. α was chosen to provide a transition width of about 0.1π. The frequency response of order 41 x 41 designed from the original transformation of this filter is shown in Fig. 5.3. The mean squared error is 2.729×10^{-3}. Figure 5.4 shows the frequency response due to windowing the impulse response of the ideal lowpass filter in two dimensions with cutoff radius 0.3π. The window used was the Cartesian product of the above one-dimensional Kaiser window. The mean squared error is 2.355×10^{-3}.

Hence the mean squared errors are comparable. However, we can see that the ripple of the filter in Fig. 5.4 is greater than that of Fig. 5.3. This perhaps is explained by noting that the peaks of the transformed filter are determined by the one-dimensional convolution in the windowing process. These peaks are given by the area under the

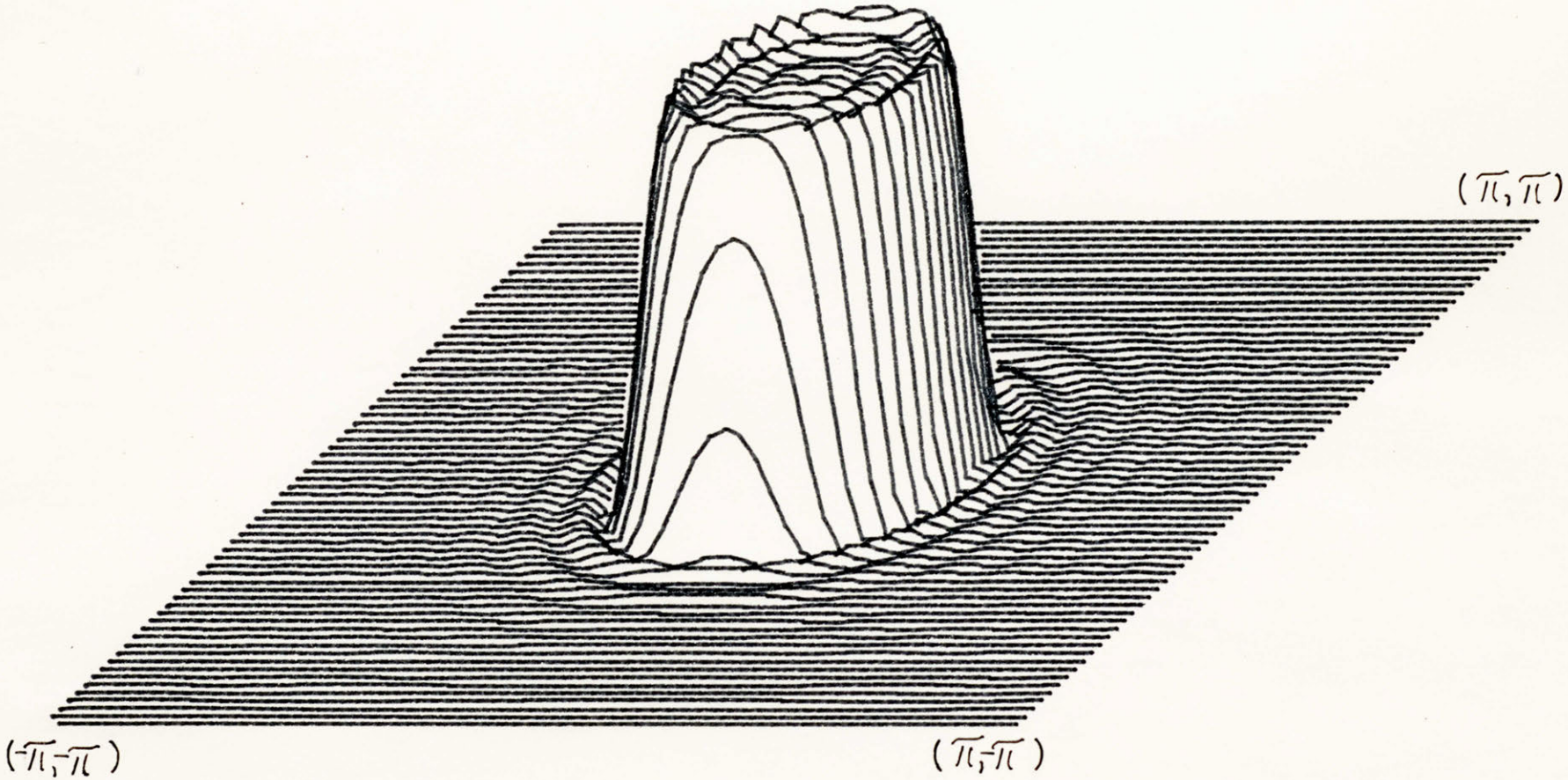

Fig. 5.3 The frequency response of a 41 x 41 filter generated from the transformation of a filter designed with a Kaiser window.

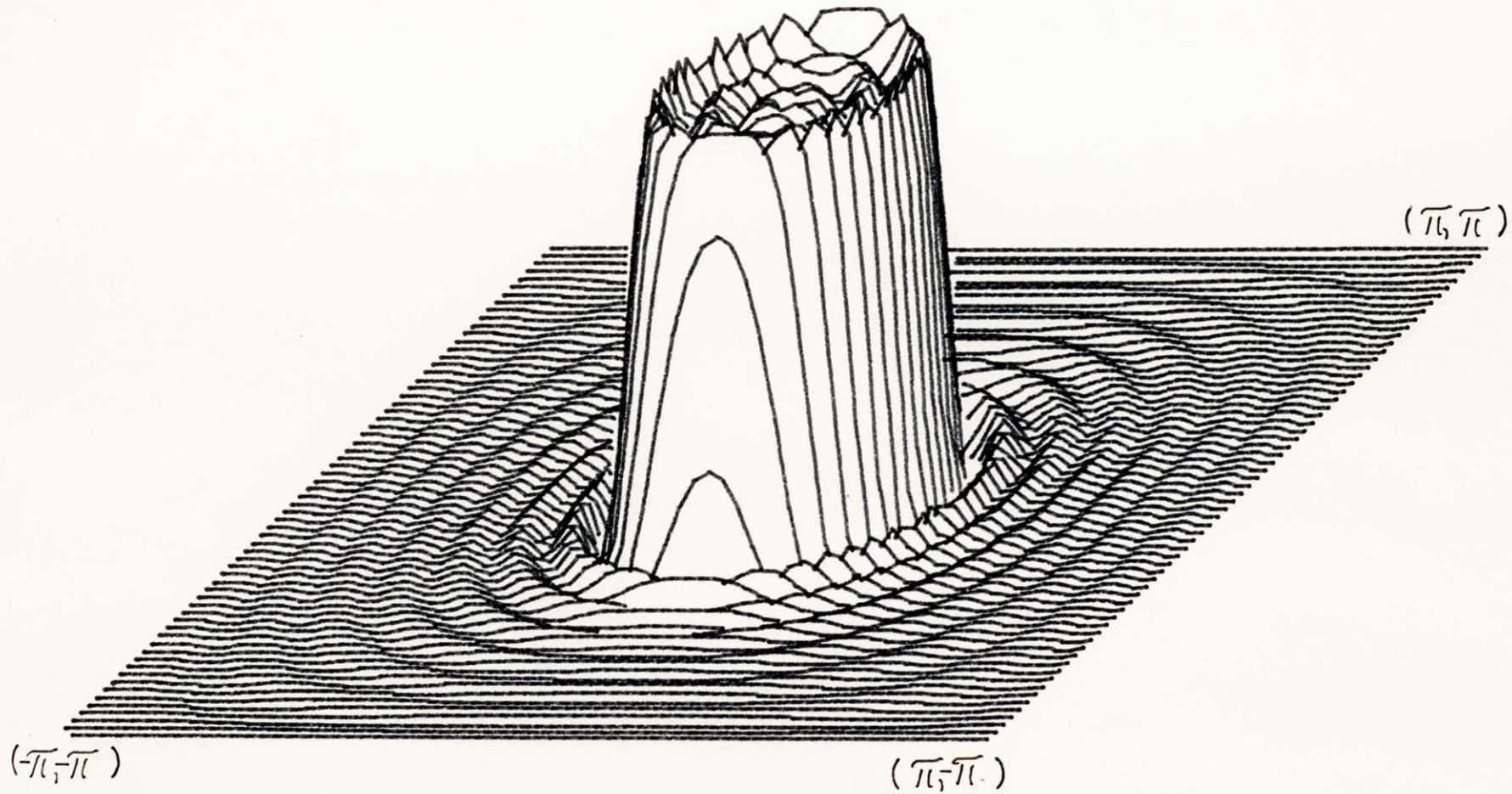

Fig. 5.4 The frequency response of a 41 x 41 filter designed with a two-dimensional Kaiser window.

product of the ideal frequency response and the transform of the window. The direct convolution in two dimensions, however, involves a volume integration. Thus, since the frequency response of the two-dimensional window is the Cartesian product of the frequency response of the one-dimensional window, larger peaks are possible. This suggests, but does not prove, that windowing in one dimension is preferred when peak ripple is of primary concern.

5.3 Alternative Transformations to Design Arbitrary Phase Two-Dimensional FIR Filters

In this section we present alternative transformations and their properties. Also, we show that with these transformations any zero phase FIR filter with a real impulse response can be designed. That is, we do not restrict the zero-phase FIR filter to have quadrant to quadrant positive symmetry, as done previously. Finally, we show that any arbitrary phase FIR filter with a real impulse response can be designed with a set of four different transformations.*

Suppose we perform the following transformation

$$\cos \omega = \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \sin (\ell\omega_1) \sin (k\omega_2) . \qquad (5.6)$$

Then making the substitution in the expression for the one-dimensional zero-phase filter (equation [2.3]), we obtain

$$H(\omega_1, \omega_2) = \sum_{p=0}^{nM} \sum_{q=0}^{nM} b(p, q) \sin (p\omega_1) \sin (q\omega_2). \qquad (5.7)$$

---

* These concepts were recently shown to be in error since sin $n\omega$ cannot in general be written as a sum of terms of the form $\sin^n \omega$ (see equations 5.7, 5.10, and 5.12).

$H(\omega_1, \omega_2)$ represents a zero-phase negative symmetric two-dimensional FIR filter. This is a filter where $h(n, m) = h(n, -m) = h(-n, m) = h(-n, -m)$.

It can be shown that any odd length zero-phase FIR two-dimensional filter with a real impulse response can be written as a sum of a positive symmetric and negative symmetric FIR filter. In the frequency domain this implies that $H(\omega_1, \omega_2)$ can be expressed as

$$H(\omega_1, \omega_2) = \sum_{p=0}^{P} \sum_{q=0}^{Q} a(p, q) \cos(p\omega_1) \cos(q\omega_2) + \sum_{p=0}^{P} \sum_{q=0}^{Q} b(p, q) \sin(p\omega_1) \sin(q\omega_2) . \tag{5.8}$$

Hence we can specify any ideal zero-phase filter with arbitrary magnitude, separate this ideal filter into its postive symmetric and negative symmetric components, and then perform two individual transformations, the cosine-to-cosine transformation for the positive symmetric ideal and the cosine-to-sine transformation for the negative symmetric ideal.

Also, we can show that any FIR filter with real impulse response (analogous to Hermitian symmetry for one-dimensional filters and odd duration) can be written as a sum of four terms

$$H(\omega_1, \omega_2) = \sum_{p=0}^{P} \sum_{q=0}^{Q} a(p, q) \cos(p\omega_1) \cos(q\omega_2) + \sum_{p=0}^{P} \sum_{q=0}^{Q} b(p, q) \sin(p\omega_1) \sin(q\omega_2)$$

$$+ \; j \sum_{p=0}^{P} \sum_{q=0}^{Q} c(p, q) \sin (p\omega_1) \cos (q\omega_2)$$

$$+ \; j \sum_{p=0}^{P} \sum_{q=0}^{Q} d(p, q) \cos (p\omega_1) \sin (q\omega_2). \qquad (5.9)$$

This is quite suitable for the transformation problem. A one-dimensional negative symmetric FIR filter of length $2n+1$ can be expressed as

$$H(\omega) = \sum_{m=1}^{n} (-j2h(m)) \sin (m\omega)$$

$$= j \sum_{m=1}^{n} a(m) \sin (m\omega) \qquad (5.10)$$

$$= j \sum_{m=1}^{n} \hat{a}(m) \sin^m (\omega) .$$

Suppose we perform the transformations

$$\sin \omega = \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \sin (\ell\omega_1) \cos (k\omega_2) \qquad (5.11a)$$

and

$$\sin \omega = \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \cos (\ell\omega_1) \sin (k\omega_2) . \qquad (5.11b)$$

Substituting into equation (5.10) we obtain:

$$H(\omega_1, \omega_2) = j \sum_{p=0}^{nM} \sum_{q=0}^{nM} c(p, q) \sin (p\omega_1) \cos (q\omega_2) \qquad (5.12a)$$

and

$$H(\omega_1, \omega_2) = j \sum_{p=0}^{nM} \sum_{q=0}^{nM} d(p, q) \cos(p\omega_1) \sin(q\omega_2). \qquad (5.12b)$$

This is precisely what we need to represent the arbitrary phase FIR filter of equation (5.9).

Hence, given any arbitrary phase FIR filter, we can perform different transformations (cosine-to-cosine, cosine-to-sine, and sine-to-cosine, sine) to obtain all four components of equation (5.9).

These results offer many possibilities of generalizing the class of filters previously designed.

CHAPTER 6

CONCLUSIONS

6.1 Summary

We have seen that the generalized McClellan transformation offers a very efficient technique for both designing and implementing two-dimensional filters of a large class of frequency responses and sizes. This technique can yield filters very close to optimal, as we saw in Chapter 5. Chebyshev optimization in two dimensions is slow and has been limited to the design of filters of highest order 15 x 15 because of the large amount of computer time required. Hence, not only is the transformation technique more efficient than the direct approximation, but, perhaps, offers the only practical method of designing nearly optimal two-dimensional FIR filters of large order.

We have seen that the contour approximation problem is solvable using either the least squares or minimax criterion. The minimax solution can be found with the efficient algorithm due to Hersey and Mersereau. Linear programming can be utilized when inequality constraints are needed in guaranteeing a well-defined mapping.

In Chapter 4 we saw that arbitrary contour shapes can be readily designed with the appropriate constraints and ideal curve $g(\omega_1)$. In many cases these constraints offer a pseudo-guarantee for a well-defined mapping. Hence, the time consumption of linear programming can be avoided if the scaling technique of section 3.1 fails to give the desired contours.

Chapter 3 presented methods for controlling the shape and increasing the density of the contours with the use of the second and third

order transformations. Also, an algorithm was developed to find the one-dimensional passband and stopband frequencies to satisfy two-dimensional specifications.

We have seen that the second big advantage over the optimal design is the speed of implementation. A transformed filter requires on the order of n multiplies per output point where $2n + 1$ is the order of the one-dimensional filter. Since our current technology dictates that speed is a function of multiplies, this implementation offers another strong justification for the transformation technique.

In Chapter 5 we further generalized the transformation for the design of arbitrary phase FIR filters. New transformations involving sine terms were presented for accomplishing this.

In essence, an entire design and implementation package has been developed with many potentialities remaining.

6.2 Suggestions for Further Research

One obvious extension of this thesis is the design of additional arbitrary contours with higher order transformations. Since multiplies are not of great concern due to the efficient implementation, high order transformations yielding high order filters do not pose a real implementation problem. These transformations may improve compression and present greater flexibility as in mapping a lowpass filter to a multiband-pass filter.

The scaling technique of section 3.1 needs further study. The effects of scaling to provide a well-defined mapping have not been experimentally shown.

A recent idea is the design of two contours simultaneously. Mersereau has designed circular and elliptical contours simultaneously,

using the least squares criterion. Designing a circular passband and stopband simultaneously should result in a more circular transition width than designing one of the two cutoff edges separately.

The ideas presented in section 5.3 may open up a whole new area of interest. Although these results were recently shown to be in error, the question still remains whether there exist transformations that allow the efficient design and implementation of arbitrary phase FIR filters.

APPENDIX 1

The following is a generalization of McClellan's recursive algorithm [2] to obtain the two-dimensional impulse response coefficients from the parameters of the transformation.

Define $T_k(x)$, $R_k(u)$, and $S_k(v)$ as the Chebyshev polynomials of degree k in the three different variables. With $x = \cos \omega$, the frequency response can be written in x as

$$H(x) = \sum_{k=0}^{n} a(k)\, T_k(x) .$$

The Chebyshev polynomials satisfy the identity

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x) .$$

With this substitution for $T_k(x)$, H(x) can be written as

$$H(x) = T_n(x)\,a(n) + T_{n-1}(x)\,a(n-1) + T_{n-2}(x)\,a(n-2) + \sum_{k=0}^{n-3} a(k)\, T_k(x)$$

$$= [2a(n)x + a(n-1)]T_{n-1}(x) + [a(n-2) - a(n)]T_{n-2}(x) + \sum_{k=0}^{n-3} a(k)\, T_k(x) .$$

We recall that the Mth order transformation is expressed as

$$x = \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k)\, R_{\ell}(u)\, S_{k}(v) .$$

$R_{\ell}(u)$ and $S_{k}(v)$ are Chebyshev polynomials up to degree M. We therefore define the functions

$$P_{M}(x) = 2a(n)x + a(n-1)$$

$$Q_{o}(x) = a(n-2) - a(n) .$$

Substituting the transformation relation for x in these equations, we obtain the new functions, $P_{M}(u, v)$ and $Q_{o}(u, v)$ in terms of u and v. $P_{M}(u, v)$ can also be considered a function of Chebyshev polynomials $R_{o}(u)$, $R_{1}(u) \cdots\cdots R_{M}(u)$ and $S_{o}(v)$, $S_{1}(v) \cdots\cdots S_{M}(v)$, or

$$P_{M}(u, v) = P_{M}(R_{o}, R_{1}, \cdots R_{M}; S_{o}, S_{1}, \cdots S_{M}) = P_{M}(\underline{R}, \underline{S})$$

where $\underline{R}$ and $\underline{S}$ represent Chebyshev polynomials up to order M (the subscript of P). Likewise, $Q_{o}(u, v)$ can be written as a function of Chebyshev polynomials R_{o} and S_{o}

$$Q_{o}(u, v) = Q_{o}(R_{o}, S_{o}) = Q_{o}(\underline{R}, \underline{S})$$

where $\underline{R}$ and $\underline{S}$ represent Chebyshev polynomials of order zero (the subscript of Q).

Hence we can write H(x) as a function of $\underline{R}$, $\underline{S}$, and x. However, x is a function of Chebyshev polynomials up to degree M. Thus we write

$$H(\underline{R}, \underline{S}) = P_M(\underline{R}, \underline{S})\, T_{n-1}(x) + Q_o(\underline{R}, \underline{S})\, T_{n-2}(x) + \sum_{k=0}^{n-3} a(k)\, T_k(x)$$

$$= \left[2xP_M(\underline{R}, \underline{S}) + Q_o(\underline{R}, \underline{S})\right] T_{n-2}(x) + \left[a(n-3) - P_M(\underline{R}, \underline{S})\right]$$

$$\cdot\ T_{n-3}(x) + \sum_{k=0}^{n-4} a(k)\, T_k(x)\ .$$

Now, let

$$P_{2M}(\underline{R}, \underline{S}) = 2xP_M(\underline{R}, \underline{S}) + Q_o(\underline{R}, \underline{S})$$

$$Q_M(\underline{R}, \underline{S}) = a(n-3) - P_M(\underline{R}, \underline{S})\ .$$

We can see in the following way that $P_{2M}(\underline{R}, \underline{S})$ is a function of Chebyshev polynomials, $\underline{R}$ and $\underline{S}$, up to degree 2M. Substituting for x,

$$P_M(\underline{R}, \underline{S}) = 2a(n) \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k)\, R_\ell(u)\, S_k(v) + a(n-1)$$

$$= \sum_{p=0}^{M} \sum_{q=0}^{M} b(p, q)\, R_p(u)\, S_q(v)$$

where

$$b(p, q) = \begin{cases} 2a(n)\, t(0, 0) + a(n-1), & p = q = 0 \\ 2a(n)\, t(p, q), & \text{otherwise} \end{cases}$$

Therefore,

$$P_{2M}(\underline{R}, \underline{S}) = 2\left[\sum_{\ell=0}^{M}\sum_{k=0}^{M} t(\ell, k)\, R_{\ell}(u)\, S_{k}(v)\right]$$

$$\cdot \left[\sum_{p=0}^{M}\sum_{q=0}^{M} b(p, q)\, R_{p}(u)\, S_{q}(v)\right] + Q_{o}(\underline{R}, \underline{S})$$

$$= \sum_{i=0}^{2M}\sum_{j=0}^{2M} d(i, j)\, R_{i}(u)\, S_{j}(v) .$$

It is obvious that $Q_{M}(\underline{R}, \underline{S})$ can be expressed as

$$Q_{M}(\underline{R}, \underline{S}) = \sum_{i=0}^{M}\sum_{j=0}^{M} c(i, j)\, R_{i}(u)\, S_{j}(v) .$$

For example, for M = 1

$$P_{2M}(\underline{R}, \underline{S}) = P_{2}(\underline{R}, \underline{S}) = 2\left[\sum_{\ell=0}^{1}\sum_{k=0}^{1} t(\ell, k)\, R_{\ell}(u)\, S_{k}(v)\right]$$

$$\cdot \left[\sum_{p=0}^{1}\sum_{q=0}^{1} b(p, q)\, R_{p}(u)\, S_{q}(v)\right] + Q_{o}(\underline{R}, \underline{S})$$

$$= 2\Big[t(0, 0) + t(1, 0)\, R_{1}(u)\, S_{o}(v) + t(0, 1)\, R_{o}(u)\, S_{1}(v)$$

$$+ t(1, 1)\, R_{1}(u)\, S_{1}(v)\Big]\left[\sum_{p=0}^{1}\sum_{q=0}^{1} b(p, q)\, R_{p}(u)\, S_{q}(v)\right]$$

$$+ \sum_{p=0}^{0}\sum_{q=0}^{0} q(p, q)\, R_{p}(u)\, S_{q}(v) .$$

Since $T_1(u)\, T_n(u) = \frac{1}{2}[T_{n-1}(u) + T_{n+1}(u)]$ for $n \neq 0$, we can write

$$P_2(\underline{R}, \underline{S}) = \sum_{p=0}^{1} \sum_{q=0}^{1} 2t(0,0)\, b(p,q)\, R_p(u)\, S_q(v)$$

$$+ \sum_{p=0}^{1} \sum_{q=0}^{1} 2t(1,0)\, b(p,q)\, \tfrac{1}{2}\left[R_{p+1}(u) + R_{p-1}(u)\right] S_q(v)$$

$$+ \sum_{p=0}^{1} \sum_{q=0}^{1} 2t(0,1)\, b(p,q)\, R_p(u)\, \tfrac{1}{2}\left[S_{p+1}(v) + S_{p-1}(v)\right]$$

$$+ \sum_{p=0}^{1} \sum_{q=0}^{1} 2t(1,1)\, b(p,q)\, \tfrac{1}{2}\left[R_{p+1}(u) + R_{p-1}(u)\right]$$

$$\cdot\ \tfrac{1}{2}\left[S_{p+1}(v) + S_{p-1}(v)\right] + \sum_{p=0}^{0} \sum_{q=0}^{0} q(p,q)\, R_p(u)\, S_q(v)$$

$$= \sum_{i=0}^{2} \sum_{j=0}^{2} d(i,j)\, R_i(u)\, S_j(v)$$

where, in general

$$d(i,j) = 2t(0,0)\, b(i,j) + t(1,0)\,[b(i+1,j) + b(i-1,j)]$$

$$+ t(0,1)\,[b(i,j+1) + b(i,j-1)]$$

$$+ \tfrac{1}{2} t(1,1)\,[b(i+1,j+1) + b(i-1,j+1) + b(i+1,j-1)$$

$$+ b(i-1,j-1)] + q(i,j)\,.$$

Note the special case where $n = 0$ in the formula $T_1(u)\, T_n(u) = \frac{1}{2}[T_{n-1}(u) + T_{n+1}(u)]$. This special case must be accounted for in the above formulation. Also, $Q_M(\underline{R}, \underline{S})$ can be written as

$$Q_M(\underline{R}, \underline{S}) = Q_1(\underline{R}, \underline{S})$$

$$= a(n-3) - P_1(\underline{R}, \underline{S})$$

$$= \sum_{i=0}^{1} \sum_{j=0}^{1} c(i, j)\, R_i(u)\, S_j(v) .$$

This procedure can be generalized for higher order transformations.

At the next step of recursion

$$P_{3M}(\underline{R}, \underline{S}) = 2xP_{2M}(\underline{R}, \underline{S}) + Q_M(\underline{R}, \underline{S})$$

$$= \sum_{i=0}^{3M} \sum_{j=0}^{3M} d(i, j)\, R_i(u)\, S_j(v)$$

and

$$Q_{2M}(\underline{R}, \underline{S}) = a(n-4) - P_{2M}(\underline{R}, \underline{S})$$

$$= \sum_{i=0}^{2M} \sum_{j=0}^{2M} c(i, j)\, R_i(u)\, S_j(v) .$$

At the kth step of recursion

$$H(\underline{R}, \underline{S}) = P_{(k+1)M}(\underline{R}, \underline{S})\, T_{n-k-1}(x) + Q_{kM}(\underline{R}, \underline{S})\, T_{n-k-2}(x)$$

$$+ \sum_{k=0}^{n-k-2} a(k)\, T_k(x) .$$

The degree of $T_k(x)$ is decreasing by one, while that of $\underline{R}(u)$ and $\underline{S}(v)$ in

H, P, and Q is increasing by M in each iteration.

Finally, after n iterations

$$H(\underline{R}, \underline{S}) = P_{nM}(\underline{R}, \underline{S}) + \sum_{i=0}^{nM} \sum_{j=0}^{nM} d(i, j)\, R_i(u)\, S_j(v) .$$

Therefore, with $u = \cos \omega_1$ and $v = \cos \omega_2$, we can write $H(\underline{R}, \underline{S})$ as

$$H(\omega_1, \omega_2) = \sum_{i=0}^{nM} \sum_{j=0}^{nM} d(i, j) \cos (i\omega_1) \cos (j\omega_2)$$

where the order of the two-dimensional filter is $(2nM + 1) \times (2nM + 1)$.

We note that due to the special cases ($T_1(u)\, T_n(u) \neq \frac{1}{2}[T_{n+1}(u) + T_{n-1}(u)]$ for $n = 0$), this method becomes very tedious for high order transformations. A simpler, more elegant approach is to apply a unit sample to the recently developed filter implementation of section 2.4. This generates the unit sample (impulse) response, the desired result.

APPENDIX 2

This appendix presents constraints for guaranteeing a well-defined second order transformation.

The second order transformation can be expressed as

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(\ell, k)\, u^{\ell}\, v^{k}.$$

Differentiating,

$$\frac{\partial x}{\partial u} = \sum_{\ell=1}^{2} \sum_{k=0}^{2} q(\ell, k)\, (\ell u^{\ell-1})\, v^{k}.$$

Thus, $\frac{\partial x}{\partial u}$ is linear in u and quadratic in v. The constraints

$$0 \rightarrow (0, 0)$$

$$\pi \rightarrow (\omega_1, \pi)$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1)$$

imply that $x = 0$ maps to $[1, 1]$ in the u, v space and $x = \pi$ maps to the lines $v = -1$ and $u = -1$. Making $\frac{\partial x}{\partial u} \geq 0$ in all four corners of the region $(-1, 1) \times (-1, 1)$ constrains the line $v = 1$ to be well defined. From octagonal symmetry the line $u = 1$ is likewise well defined. Therefore, the entire boundary is well defined.

Since, along any horizontal line, $\frac{\partial x}{\partial u}$ is linear in u, and since the boundaries are well defined, then constraining $\frac{\partial x}{\partial u} \geq 0$ along the lines $v = \pm 1$ guarantees a well-defined mapping. This can be seen in Fig. A.1. However, if this constraint is not satisfied, a well-defined mapping may still exist, as shown in Fig. A.2. Since $\frac{\partial x}{\partial u}$ is a quadratic in v, the

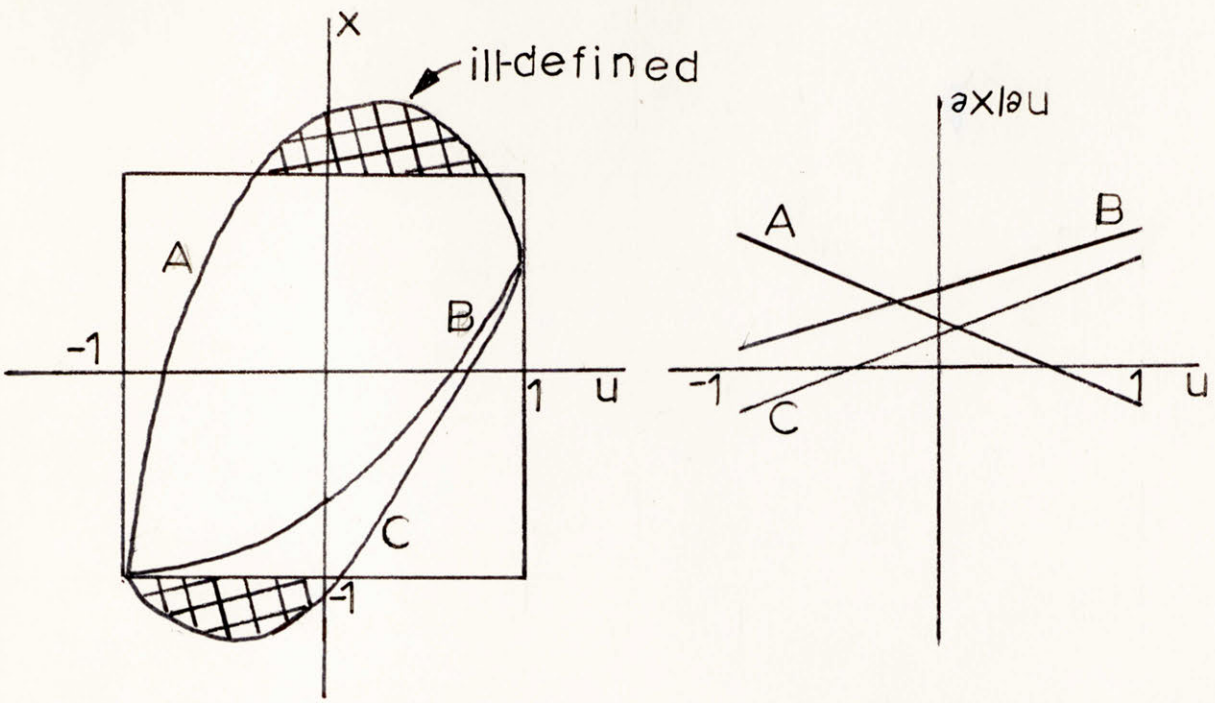

Fig. A.1 The possible quadratic functions along the u axis.

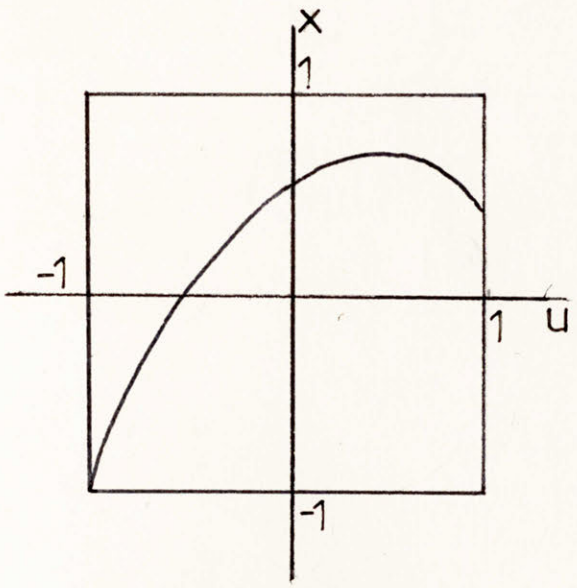

Fig. A.2 A well-defined, but nonmonotonic quadratic along the u axis.

constraint $\frac{\partial^2 x}{\partial u \partial v} \geq 0$ in all four corners is sufficient.

Hence, the constraints

$$\frac{\partial x}{\partial u} \geq 0$$

$$\frac{\partial^2 x}{\partial v \partial u} \geq 0$$

at (1, 1), (-1, -1), (-1, 1) and (1, -1) are sufficient but not necessary for a well-defined second order mapping.

APPENDIX 3

Here we present the derivation of the second order transformation approximation problem for transition width compression, for circularly symmetric lowpass filters.

The second order transformation can be written

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(\ell, k)\, u^{\ell} v^{k} .$$

Along the u axis

$$x = \sum_{\ell=0}^{2} \sum_{k=0}^{2} q(\ell, k)\, u^{\ell} .$$

We let $t(\ell, k) = t(k, \ell)$ and make the substitution $A = t(0,0)$, $B = t(1,0)$, $C = t(0,1) \cdots\cdots I = t(2,2)$. Then we can show that

$$x = (A + B - G - 2I) + (B + D + G)u + (2E + 2G + 2I)u^{2} .$$

Letting

$$A' = 2E + 2G + 2I$$

$$B' = B + D + G$$

$$C' = A + B - G - 2I$$

then

$$x = A'u^{2} + B'u + C' .$$

Imposing the additional constraints

$$0 \rightarrow (0,0)$$

$$\pi \rightarrow (\omega_1, \pi)$$

$$\left.\frac{\partial x}{\partial u}\right|_{u = \pm 1} \geq 0$$

we can show that

$$B' = 1$$

$$A' = -C'$$

and hence

$$x = A'u^2 + u - A' .$$

These constraints guarantee a well-defined mapping along the u, v axes. Continuing,

$$\left.\frac{\partial x}{\partial u}\right|_{u = \pm 1} = 2A'u + 1 \geq 0$$

implying that

$$A' \geq -\tfrac{1}{2}, \quad u = 1$$

$$A' \leq \tfrac{1}{2}, \quad u = -1 .$$

The objective is to maximize the slope at a particular $u = u_o$:

$$\max \left(\left.\frac{\partial x}{\partial u}\right|_{u_o}\right) = \max (2A'u_o + 1).$$

Thus, by inspection

$$A' = \begin{cases} \tfrac{1}{2}, & u_o \geq 0 \\ -\tfrac{1}{2}, & u_0 < 0 \end{cases}$$

to maximize the slope at u_o.

The approximation problem is formulated in the following way. The constraint equations can be expressed as

$$
\begin{aligned}
A' = \pm\tfrac{1}{2}: &\quad 2E + 2G + 2I = \pm\tfrac{1}{2} \\
0 \to (0,0): &\quad A + 2B + D + 2E + 2G + I = 1 \\
\pi \to (\omega_1, \pi): &\quad A - B + E = -1 \\
&\quad B - D + G = 0 \\
&\quad E - G + I = 0
\end{aligned}
$$

These equations can also be expressed in matrix form (with A the only free variable) as:

$$
\underbrace{\begin{vmatrix} 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 2 & 1 & 2 & 2 & 1 \end{vmatrix}}_{\underline{M}} \times \underbrace{\begin{vmatrix} B \\ D \\ E \\ G \\ I \end{vmatrix}}_{\underline{C}} = A \underbrace{\begin{vmatrix} 0 \\ -1 \\ 0 \\ 0 \\ -1 \end{vmatrix}}_{\underline{P}} + \underbrace{\begin{vmatrix} \pm\frac{1}{4} \\ -1 \\ 0 \\ 0 \\ 1 \end{vmatrix}}_{\underline{Q}}
$$

We can write this in the more compact form

$$\underline{C} = A\underline{M}^{-1}\underline{P} + \underline{M}^{-1}\underline{Q}.$$

Solving for $u > 0$

$$
\begin{aligned}
B &= 0.625 \\
D &= 0.5 \\
E &= -A - 0.375 \\
G &= -0.125
\end{aligned}
$$

$$I = A + 0.75.$$

Let

$$\alpha_1 = \cos \omega_1$$

$$\alpha_2 = \cos \omega_2$$

$$\beta_1 = \cos (2\omega_1)$$

$$\beta_2 = \cos (2\omega_2)$$

$$\omega_2 = (R_p^2 - \omega_1^2)^{\frac{1}{2}}.$$

The ideal and basis function for A can be found for $u > 0$ by writing $\cos \omega$ as

$$\begin{aligned}\cos \omega = {} & A + 0.625 (\alpha_1 + \alpha_2) + 0.5 \alpha_1\alpha_2 \\ & + (-A - 0.375)(\beta_1 + \beta_2) - 0.125 (\beta_1\alpha_2 + \alpha_1\beta_2) \\ & + (A + 0.25) \beta_1\beta_2.\end{aligned}$$

Therefore

$$e(\omega_1) = H_d(\omega_1) - A\varphi_1(\omega_1)$$

where

$$\begin{aligned}H_d(\omega_1) = {} & \cos \omega - (0.625 (\alpha_1 + \alpha_2)) \\ & + 0.5 \alpha_1\alpha_2 - 0.375 (\beta_1 + \beta_2) \\ & - 0.125 (\beta_1\alpha_2 + \alpha_1\beta_2) \\ & + 0.25 \beta_1\beta_2\end{aligned}$$

$$\varphi_1(\omega_1) = 1 - \beta_1 - \beta_2 + \beta_1\beta_2.$$

The approximation problem can be similarly formulated for $u < 0$.

It is important to note that since the quadratic along the u, v axes is completely determined, the value of ω (the inverse frequency image of the desired contour) must be such that

$$\cos \omega = x = A' u_o^2 + u_o + A'$$

where $R_p = \cos^{-1} u_o$, the ideal two-dimensional passband radius. For example, if we desire a circular contour of radius $0.8\,\pi$, then $\cos \omega$ of the ideal function $H_d(\omega_1)$ must be such that

$$\cos \omega = A'(0.8\,\pi)^2 + (0.8\,\pi) + A'$$

where $A' = -\frac{1}{2}$.

APPENDIX 4

In this appendix we derive the third order transformation approximation problem for transition width compression for a circularly symmetric lowpass filter with passband cutoff near 0.5π.

The third order transformation can be expressed as

$$x = \sum_{\ell=0}^{3} \sum_{k=0}^{3} q(\ell, k)\, u^{\ell} v^{k} .$$

As in Appendix 3 along the u axis we can write

$$x = A'u^3 + B'u^2 + C'u + D'$$

where A', B', C', and D' are linear functions of the transformation coefficients $t(\ell, k)$.

For compression near the 0.5π radius constraints on A', B', C', and D' must be found such that $\left.\frac{\partial x}{\partial u}\right|_{u = 0}$ is large. Constraining $\left.\frac{\partial x}{\partial u}\right|_{u = \pm 1} > 0$ and $\frac{\partial x}{\partial u}$ (a quadratic) to be concave downward guarantees a well-defined mapping on the u, v axes. We impose the additional equality constraints

$$0 \rightarrow (0, 0)$$

$$\pi \rightarrow (\omega_1, \pi)$$

$$H(\omega_1, \omega_2) = H(\omega_2, \omega_1) .$$

The total set of constraint equations on the cubic can then be written

$$\left.\frac{\partial x}{\partial u}\right|_{u = \pm 1} \geq 0 : \quad \begin{aligned} 3A' + 2B' + C' &\geq 0 \\ 3A' - 2B' + C' &\geq 0 \end{aligned}$$

$$\text{downward concavity: } A' \leq 0$$

equality constraints: $A' + C' = 1$

$$B' = -D'$$

where $\left.\frac{\partial x}{\partial u}\right|_{u = 0} = z$ is to be maximized. Under these constraints, conditions for the existence of a solution can be derived:

$$3(1 - C') - 2B' + C' \geq 0$$

or

$$B' \leq \frac{3}{2} - z .$$

Likewise,

$$B' \geq z - \frac{3}{2} .$$

Hence, for a solution to exist

$$1 \leq z \leq \frac{3}{2} .$$

Thus maximum compression at the origin is achieved when $\left.\frac{\partial x}{\partial u}\right|_{u = 0} = \frac{3}{2}$ where it can be shown that

$$\left.\frac{\partial x}{\partial u}\right|_{u = \pm 1} = 0$$

$$A' = -\tfrac{1}{2}$$

$$B' = 0$$

$$C' = \frac{3}{2} .$$

Finally, with the same technique as in Appendix 3, the approximation problem can be formulated. All the above constraints taken together yield three free variables $A = t(0, 0)$, $B = t(1, 0)$, and $D = t(1, 1)$. Therefore the error can be expressed as

$$e(\omega_1) = H_d(\omega_1) - A\varphi_1(\omega_1) - B\varphi_2(\omega_1) - D\varphi_3(\omega_1) .$$

Again (as in Appendix 3) cos ω, which is part of the ideal function, $H_d(\omega_1)$, must be such that

$$\cos \omega = x = A'u_o^3 + B'u_o^2 + C'u_o + D'$$

if a particular frequency mapping is desired. For $u_o = 0$ (ideal radius of $0.5\,\pi$) the curves and contours of Figs.3.10a, b, c were obtained.

APPENDIX 5

The following presents the procedure for the design of Fan filters under inequality constraints. Linear programming is utilized.

Using implicit differentiation, we can show that (A = t(0, 0), B = t(1, 0) $\cdots$ I = t(2, 2))

$$\frac{dv}{du} = \frac{(H - B) + 4(I - E)u - Dv - 4Guv - 2Hv^2 - 8Iuv^2}{(C - G) + Du + 4(F - I)v + 2Gu^2 + 4uv + 8Iu^2v} .$$

The increasing monotonicity constraint $\frac{dv}{du} \geq 0$ (over the sparse grid of Fig. 3.3) is imposed by constraining both the numerator and denominator to be greater than zero. This yields eighteen inequality constraints. Remarkably, the constraints $\frac{\partial x}{\partial u} \geq 0$ and $\frac{\partial x}{\partial v} \leq 0$ (implying, but not guaranteeing, a well-defined mapping) yield the same set of constraints over this grid.

Hence the linear programming problem can be formulated in the following way. Minimize δ_e under the constraints

$$0 \rightarrow (\pi, 0)$$

$$\pi \rightarrow (0, \pi)$$

$$\left.\frac{dv}{du}\right|_{u_i v_j} \geq 0, \text{ over a sparse grid}$$

$$\cos \omega - \sum_{\ell=0}^{M} \sum_{k=0}^{M} t(\ell, k) \cos(\ell\omega_1) \cos(kg(\omega_1)) \leq \delta_e$$

where ω_1 and $g(\omega_1) = b\omega_1$ are sampled over a fine grid. b represents the slope of the passband edge of the Fan filter. (See Fig. 4.10 for an example.)

BIBLIOGRAPHY

1. J. H. McClellan, "The Design of Two-Dimensional Digital Filters by Transformations", Special Paper, Proceedings of the 7th Annual Princeton Conference on Information Sciences and Systems, 1973.

2. J. H. McClellan and T. W. Parks, "Equiripple Approximation of Fan Filters", Geophysics, Vol. 37, pp. 573-583, August, 1972.

3. J. H. McClellan, T. W. Parks, "A Unified Approach to the Design of Optimum FIR Linear-Phase Digital Filters", IEEE Trans. Circuit. Theory, Vol. CT-20, No. 6, November, 1973.

4. T. S. Huang, "Two-Dimensional Windows", IEEE Trans. Audio Electroacoustics (Corresp.), Vol. AU-20, pp. 88-89, March, 1972.

5. J. U. Hu and L. Rabiner, "Design Techniques for Two-Dimensional Digital Filters", IEEE Trans. Audio Electroacoustics, Vol. AU-20, No. 4, October, 1974.

6. J. G. Fiasconaro, "Two-Dimensional Nonrecursive Digital Filters", Lincoln Lab. Technical Report, No. 505, May, 1973.

7. Y. Kamp, J. P. Thiran, "Chebyshev Approximation for Two-Dimensional Nonrecursive Digital Filters", M. B. L. E. Laboratoire de Recherches, Report R245, March, 1974.

8. R. M. Mersereau, D. B. Harris, H. S. Hersey, "An Efficient Algorithm for the Design of Equiripple Two-Dimensional FIR Digital Filters", Proceedings of 1975 IEEE Inter. Symp. Circuits Systems.

9. H. S. Hersey, R. M. Mersereau, "An Algorithm to Perform Minimax Approximation in the Absence of the Haar Condition", MIT Quarterly Progress Report, No. 114, July, 1974.

10. E. W. Cheney, Introduction to Approximation Theory, McGraw Hill Book Co., New York, 1966.