XV. DIGITAL SIGNAL PROCESSING

Academic and Research Staff

Prof. Alan V. Oppenheim
Prof. Jonathan Allen
Dr. Tomonori Aoyama
Dr. Russell M. Mersereau

Graduate Students

Ronald E. Crochiere
Dan E. Dudgeon
Ronald H. Frazier
Harlan S. Hersey
Gary E. Kopec
Michael R. Portnoff
José M. Tribolet
Ernest Vincent
Robert J. Wenzel

A. EXISTENCE OF CEPSTRA FOR TWO-DIMENSIONAL RATIONAL POLYNOMIALS

U. S. Navy Office of Naval Research (Contract N00014-67-A-0204-0064)

Dan E. Dudgeon

This report is based on work submitted in May 1974 to the Department of Electrical Engineering in partial fulfillment of the requirements of the degree of Doctor of Science. The work was supported in part by the M. I. T. Lincoln Laboratory (which is sponsored by the U. S. Air Force).

1. Introduction

In signal-processing problems where two signals are combined by convolution, homomorphic filtering techniques[1] may be useful. Assume that we are given a signal s(n) which is the convolution of x(n) and h(n).

$$s(n) = \sum_k h(k)\, x(n-k). \tag{1}$$

If we take the Fourier transform, complex logarithm, and inverse Fourier transform of both sides of Eq. 1, we get the relation $\hat{s}(n) = \hat{x}(n) + \hat{h}(n)$. The variables $\hat{s}$, $\hat{x}$, and $\hat{h}$ are the complex cepstra of s, x, and h. Linear filtering may be performed on $\hat{s}(n)$ in the cepstral domain, and the result Fourier transformed, exponentiated, and inverse Fourier transformed to complete the homomorphic filtering operation.

Such a technique is useful for inverse filtering problems or problems where a signal and noise are combined by convolution. To make the definition of the complex cepstrum tractable mathematically, it is important to ensure that the phase function associated with s(n) be continuous, odd, and periodic. Defining the phase by using the principal part of the complex logarithm is usually not satisfactory.[2] Therefore, to ensure continuity, phase must be defined as an integral.[1] If the phase so defined is not periodic, it may be made so by subtracting a linear phase component. This simply corresponds

to a shift in the time origin of the original signal.

We must restrict the set of signals for which cepstra can be defined to exclude signals whose Fourier transforms go to zero or infinity at some frequency. This corresponds to their z-transforms having a pole or a zero on the unit circle resulting in a phase discontinuity that cannot be removed by the integral formulation. Equivalently, the logarithms of their Fourier transforms become infinite at such frequencies.

One class of signals for which cepstra may be defined is the class whose z-transforms are rational polynomials and are analytic and nonzero on the unit circle. For these signals the cepstra also have z-transforms that are analytic on the unit circle.[1]

2. Two-Dimensional Cepstrum

In an exactly analogous manner, we can demonstrate that any 2-D array having a rational z-transform will also have a well-defined 2-D complex cepstrum, provided that (i) the Fourier transform is not equal to zero or infinity at any frequency, and (ii) we are careful to eliminate linear phase components by an appropriate shift of the original array.

This can easily be seen if we first consider finite-extent 2-D arrays. These arrays have Fourier transforms that are two-dimensional polynomials in $\exp[j\mu]$ and $\exp[j\nu]$. We shall now show that if a finite-extent array b(m, n) has a Fourier transform $B(\mu, \nu)$ that is nonzero for all μ and ν, then the phase function associated with $B(\mu, \nu)$ is the sum of a linear component plus a continuous, odd, and periodic component. First, let $z = \exp[j\nu]$, and consider

$$B_\mu(z) = \sum_n \left[\sum_m b(m, n)\, e^{-j\mu m} \right] z^{-n}$$

as a one-dimensional polynomial in z with a parameter μ. Since b(m, n) is of finite extent, $B_\mu(z)$ will only have poles inside the unit circle at $z = 0$. Now let us define the phase function $\phi(\mu, \nu)$ as a contour integral

$$\phi(\mu, \nu) = \mathrm{Im} \left\{ \oint \frac{B'_\mu(z)}{B_\mu(z)}\, dz \right\} + \phi(\mu, 0), \tag{2}$$

where the prime denotes differentiation with respect to z. The contour of integration starts at $z = 1$ and proceeds around the unit circle to $z = \exp[j\nu]$. It is necessary to define the constant $\phi(\mu, 0)$. We do this by considering $\phi(\mu, 0)$ to be a one-dimensional phase function given by

$$\phi(\mu_o, 0) = \int_0^{\mu_o} \frac{\frac{\partial B_I}{\partial \mu} B_R - \frac{\partial B_R}{\partial \mu} B_I}{B_R^2 + B_I^2}\, d\mu,$$

where B_R is the real part of $B(\mu, \nu)$, and B_I is the imaginary part. This formula is the phase as a function of μ for $\nu = 0$, and may be derived as Oppenheim and Schafer[2] have done.

By constructing $\phi(\mu, \nu)$ in this manner, we are assured that $\phi(\mu, \nu)$ is continuous and odd, and

$$\phi(\mu, \nu) = -\phi(-\mu, -\nu).$$

When $\nu = 2\pi$, the contour of integration in Eq. 2 is a closed curve. Using Cauchy's Residue theorem and Marden's[3] Theorem (1, 2), we have

$$\phi(\mu, 2\pi) = 2\pi r - 2\pi N,$$

where r is the number of roots (including multiplicities) of $B_\mu(z)$ inside the unit circle, and N is the number of poles at $z = 0$. If we let $k_\nu = r - N$, then it is clear that

$$\phi(\mu, \nu+2\pi) = \phi(\mu, \nu) + 2\pi k_\nu.$$

Similarly, we can derive the relation

$$\phi(\mu+2\pi, \nu) = \phi(\mu, \nu) + 2\pi k_\mu.$$

It remains to be shown that k_ν is not a function of μ, and k_μ is not a function of ν. If we examine the roots of $B_\mu(z)$ as we continuously vary the parameter μ from zero to 2π, we discover that the roots move about in a continuous manner.[4] Thus, for a root to move from inside to outside the unit circle (or vice versa), it must lie on the unit circle for some value of μ. But this violates the hypothesis that $B(\mu, \nu) \neq 0$. Therefore the number of roots r inside the unit circle is not a function of μ, and hence k_ν is not a function of μ. A similar argument can be made to show that k_μ is not a function of ν.

Given a continuous odd phase function $\phi(\mu, \nu)$ such that

$$\phi(\mu, \nu+2\pi) = \phi(\mu, \nu) + 2\pi k_\nu$$

$$\phi(\mu+2\pi, \nu) = \phi(\mu, \nu) + 2\pi k_\mu$$

$$\phi(\mu, \nu) = -\phi(-\mu, -\nu),$$

we can subtract the linear phase component

$$\phi_L(\mu, \nu) = k_\nu \nu + k_\mu \mu$$

to obtain a term $\phi_A(\mu, \nu)$ that can be shown to be continuous, odd, and periodic.

This result may be extended to arrays with rational Fourier transforms. The phase function for such an array can be defined as the phase function of the numerator polynomial minus the phase function of the denominator polynomial. Clearly, this difference can also be written as a linear phase component plus a continuous, odd, and

periodic component.

It is easy to verify that subtracting the linear phase component is equivalent to a shift of the array. Therefore, if we are given an array h(m, n) whose Fourier transform is a ratio of polynomials

$$H(\mu, \nu) = \frac{\sum_k \sum_\ell a(k, \ell) \exp[-j\mu k - j\nu \ell]}{\sum_r \sum_s b(r, s) \exp[-j\mu r - j\nu s]},$$

then its phase function can be written in the form

$$\phi(\mu, \nu) = \phi_A(\mu, \nu) + k_\nu \nu + k_\mu \mu.$$

We can form a new sequence $g(m, n) = h(m-k_\mu, n-k_\nu)$ whose phase function is $\phi_A(\mu, \nu)$, which is continuous, odd, and periodic. Since

$$\ln |G(\mu, \nu)| = \ln |H(\mu, \nu)|$$

is continuous, even, and periodic, we may form the function

$$\hat{G}(\mu, \nu) = \ln |G(\mu, \nu)| + j\phi_A(\mu, \nu)$$

which has a real inverse Fourier transform denoted by $\hat{g}(m, n)$ and called the cepstrum of g(m, n). Since $G(\mu, \nu)$ is a ratio of polynomials, the properties of $\hat{G}(\mu, \nu)$ imply that the z-transform of $\hat{g}(m, n)$, denoted by F(w, z),

$$F(w, z) = \sum_m \sum_n \hat{g}(m, n)\, w^{-m} z^{-n}$$

is an analytic function for $|w| = |z| = 1$. This may be seen by considering $\hat{G}(\mu, \nu)$ to be a one-dimensional Fourier transform in ν with a parameter μ. Then the function $F(e^{j\mu}, z)$ is analytic for all values of μ, and $|z| = 1$. Similarly, the function $F(w, e^{j\nu})$ is analytic for all values of ν, and $|w| = 1$. Therefore F(w, z) is analytic for $|w| = |z| = 1$.

3. Summary

Two conclusions can be drawn from this derivation. First, any real array whose Fourier transform has a log-magnitude that is continuous, even, and periodic, and whose phase is continuous, odd, and periodic will have a well-defined cepstrum that is real. Second, one class of arrays with well-defined cepstra is the class whose z-transforms are the ratio of two polynomials (with the restriction that the z-transforms not be zero

or infinite for $|w| = |z| = 1$), provided that care has been taken first to shift the arrays to eliminate linear phase components. In addition, for this class of arrays we can argue that the logarithm of the z-transform is analytic on the surface $|w| = |z| = 1$.

The author would like to express his thanks to Professor Alan V. Oppenheim for his suggestions and encouragement during the course of this work.

References

1. A. V. Oppenheim, R. W. Schafer, and T. G. Stockham, Jr., "Nonlinear Filtering of Multiplied and Convolved Signals," Proc. IEEE 56, 1264-1291 (1968).
2. A. V. Oppenheim and R. W. Schafer, Digital Signal Processing (to be published by Prentiss-Hall, Inc., Englewood Cliffs, N. J.).
3. M. Marden, The Geometry of the Zeros of a Polynomial in a Complex Variable (American Mathematical Society, New York, 1949).
4. Ibid., p. 3.

B. AN ALGORITHM TO PERFORM MINIMAX APPROXIMATION IN THE ABSENCE OF THE HAAR CONDITION

U. S. Navy Office of Naval Research (Contract N00014-67-A-0204-0064)
National Science Foundation (Grant GK-31353)

Harlan S. Hersey, Russell M. Mersereau

In the minimax or Tchebycheff approximation problem the task is to approximate a function D(x) on the interval [a, b] by a linear combination of basis functions $\{\phi_i(x)\}$. Thus

$$D(x) \approx \sum_{i=0}^{N-1} c_i \phi_i(x),$$

where the coefficients $\{c_i\}$ are chosen to minimize an error defined as

$$\max_{a \leq x \leq b} \left| D(x) - \sum_{i=0}^{N-1} c_i \phi_i(x) \right| = \max_{a \leq x \leq b} |E(x)|.$$

If the set of basis functions satisfies a strong nondegeneracy condition (Haar condition), the problem can be solved very efficiently by using the second algorithm of Remez.[1] This condition not only implies that the basis functions be linearly independent on the interval [a, b], but also that they be linearly independent on any set of N samples chosen from this interval.

The Parks-McClellan algorithm[2] for the design of linear-phase finite-impulse response (FIR) filters performs an approximation of the form

$$D(e^{j\omega}) \approx H(e^{j\omega}) = \sum_{n=1}^{N-1} 2h(n) \cos \omega_n + h(0),$$

where the $\{h(n)\}$ are the impulse response coefficients of the filter. Since the set of basis functions $\{1, \cos\omega, \ldots, \cos(N-1)\omega\}$ satisfy the Haar condition on the interval $[0, \pi]$, the second algorithm of Remez plays a central role in the Parks-McClellan algorithm.

If equality constraints are applied to the impulse response coefficients such as $h(3) = 0$, $h(4) = h(5)$, or $h(17) = 3$, then the resulting approximation problem, although superficially very similar, differs from the unconstrained problem in the important respect that the new set of basis functions will generally not satisfy the Haar condition. Since the resulting basis functions are still linearly independent, we can apply the ascent algorithm described by Cheney.[1] Stiefel[3] has shown that this is equivalent to linear programming applied to the dual formulation of the problem.

It has been shown that when the optimum minimax solution has been obtained the error function $E[x]$ will oscillate between $\pm E_{max}$, $N+1$ times over the interval $[a, b]$. The ascent algorithm starts with an initial guess of these $N+1$ extremal locations. It then formulates and solves a set of $N+1$ linear equations in $N+1$ unknowns. The unknowns represent the N coefficients $\{c_i\}$ and the deviation σ. Next, a search algorithm uses the calculated coefficients to determine the location of the worst error over a dense grid of samples of x. This point is exchanged with one of the $N+1$ extremal locations and the ascent algorithm then solves for a new set of coefficients and repeats. The process continues until the worst error found over the entire grid is already a member of the extremal set. At each step the calculated derivation increases.

Because of the absence of the Haar condition, uniqueness is not guaranteed. Nonuniqueness at any step is indicated by the fact that the matrix of equations used to solve for the coefficients is singular. By perturbing the elements slightly, however, we may arrive at one of the infinite number of best approximations.

To give the flavor of the computations involved, we shall describe some of the details of the ascent algorithm. Further details can be found in Cheney[1] or in Kamp and Thiran.[4] First, we must find the signs of the errors at the $N+1$ extremal locations or, in the filter design case, extremal frequencies. We must choose the signs so that the origin of $N+1$ space lies in the convex hull of the set of basis functions evaluated at the $N+1$ extremal frequencies, which we refer to as the basis vectors. Thus we need only find a nontrivial solution to the set of equations

$$\sum_{j=1}^{N+1} t_j A^{i_j} = 0,$$

where A^{i_j} is the basis vector at the j^{th} extremal point or the i_j^{th} grid point. Then the

sign of the error associated with the j^{th} extremal point is $s_j = \operatorname{sgn} t_j$. Now we may solve the set of equations

$$e \cdot s_j + \sum_{k=1}^{N} c_k A_k^{i_j} = D(i_j) \qquad j = 1, 2, \ldots, N+1,$$

where $A_k^{i_j}$ is the k^{th} coordinate of the basis vector and D(i) is the ideal function at the i^{th} grid point. Having calculated the coefficients c_k, we must now find the location of the worst error on the entire grid. If this error is no larger than the error calculated at the extremal frequencies, we are finished. Otherwise, we must pick a member of the extremal set with which to exchange this point. We choose the point so that the origin of N+1 space again lies in the convex hull of the basis vectors exchanged one at a time with the new basis vector. This can be effectively accomplished by taking the dot product of the new basis vector with each of the old basis vectors and picking the resulting maximum. These points are exchanged and the ascent algorithm continues. Also, we need not reinvert the matrix for the new solution, since we can modify the inverse matrix itself to perform the exchange. Further, if the new matrix (before inversion) is nonsingular, the new error e is guaranteed to ascend in magnitude.

In the event that the matrix should be singular (which the Haar condition does not permit), it can be made nonsingular by perturbing its elements slightly. The error e may not strictly ascend, however, and this is referred to as a static exchange. It is indicative that at this stage the solution will not be unique. The condition can be detected and the nonuniqueness indicated if we have the final solution. Other techniques can be applied to arrive at solutions that optimize the solution according to some other consideration. The extremal error e for each solution will be the same, however. For further information on the question of nonuniqueness see Kamp and Thiran.[4]

The most time-consuming part of the ascent algorithm is in calculating the error function at all points of the grid. The result of this computation is one number: the index of the location of the worst error. We have explored two techniques that speed up the design algorithm. The first uses the principle stated in the first algorithm of Remez. The second is an attempt at a multiple exchange algorithm.

The first algorithm of Remez (see Cheney[1]) states that we may solve the approximation problem on a thin subset of grid points containing at least N+1 points. Then the solution is used to find the location of the worst error over the entire grid. This point is then appended to the thin subset and the ascent algorithm is called for the solution on these N+2 points. The algorithm continues in this fashion until the location of the worst error is already a member of the thin set. The speedup apparently occurs because (i) the ascent algorithm need only use a few well-chosen points, and (ii) old peak locations are immediately available to the ascent algorithm should they again become

locations of worst error. The obvious disadvantage is that the subset grows larger at each iteration.

The multiple exchange algorithm seeks to keep the thin subset relatively thin by replacing old (and, we hope, unusable) points with local maxima of the current error function. In this way we provide many peaks rather than only one to the ascent routine. This is known to provide increased efficiency in Remez' second algorithm when the Haar condition is present. As the approximation approaches convergence, each peak will be significant. Preliminary results indicate that a considerable speedup can result. We shall continue work in this area in the hope that a highly efficient algorithm can be developed. An application for this algorithm is presented in Section XV-C.

References

1. E. W. Cheney, Introduction to Approximation Theory (McGraw-Hill Book Company, Inc., New York, 1966).
2. T. W. Parks and J. H. McClellan, "Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase," IEEE Trans., Vol. CT-19, No. 2, p. 189, March 1972.
3. E. Stiefel, "Note on Jordan Elimination, Linear Programming and Tchebycheff Approximation," Numer. Math., Vol. 2, p. 1, November 1960.
4. Y. Kamp and J. P. Thiran, "Chebyshev Approximation for Two-Dimensional Nonrecursive Digital Filters," Research Laboratory Report No. R245, Manufacture Belge des Lampes et de Matériel Électronique, S. A., Brussels, March 1974.

C. DESIGN TECHNIQUES FOR TWO-DIMENSIONAL FINITE IMPULSE RESPONSE CIRCULARLY SYMMETRIC DIGITAL FILTERS

U. S. Navy Office of Naval Research (Contract N00014-67-A-0204-0064)
National Science Foundation (Grant GK-31353)

Harlan S. Hersey, Russell M. Mersereau

1. Introduction

In this report we are concerned with the design of circularly symmetric two-dimensional (2-D) digital filters whose impulse response is finite in extent (FIR). The choice of circular symmetry (or more precisely "octal" symmetry) is appealing from practical, as well as theoretical, considerations. Theoretically, it allows us to limit the number of free parameters in the approximation problem. Practically, there is a natural tendency in some applications to do 2-D filtering without regard to direction. For the most part, however, the notions of the new algorithm can be generalized to other types of symmetry. Hu and Rabiner[1] have explored the 2-D extension of the

1-D technique known as "frequency sampling." They also applied linear programming (a very expensive solution) to obtain a minimax approximation. A more subtle approach was taken by Fiasconaro[2] in which linear programming is applied to a "thin" set of points in the 2-D domain. In more recent work, Kamp and Thiran[3] present a detailed algorithm in which the multiple exchange properties of a Remez-type algorithm are exploited.

All of these optimal design techniques are rather costly with respect to computer design time. Another technique proposed by McClellan[4] transforms 1-D optimal filters into certain optimal 2-D filters. This approach is extremely efficient and perhaps is suitable for most applications. We propose, however, by exploiting the notions of projections,[5] that more nearly optimal 2-D filters can be designed by transforming the problem into a one-dimensional design problem. To this end, we shall explore the techniques of least-squares approximation, the Parks-McClellan algorithm[6] in 1-D, and approximation in 1-D subject to certain symmetry constraints implied by 2-D circular symmetry. We shall also compare the last with the approximation problem in 2-D.

2. Least-Squares Design in Two Dimensions

The design of 2-D circularly symmetric lowpass filters is a particularly straightforward problem when we apply the least-squares error criterion. The solution corresponds to sampling the inverse Fourier transform of the ideal lowpass filter. If

$$H\left(e^{j\omega_m}, e^{j\omega_n}\right) = \begin{cases} 1 & \omega_m^2 + \omega_n^2 \leq \omega_p^2 \\ 0 & \text{elsewhere} \end{cases} \tag{1}$$

then

$$h(m, n) = \frac{\omega_p J_1\left(\omega_p \sqrt{m^2 + n^2}\right)}{2\pi\sqrt{m^2 + n^2}}, \qquad m \cdot n \neq 0$$

$$h(0, 0) = \frac{\omega_p^2}{4\pi}, \tag{2}$$

where $J_1(x)$ is the first-order Bessel function of the first kind. Note that $h(m, n)$ possesses "octal symmetry." That is, $h(m, n) = h(\pm m, \pm n) = h(\pm n, \pm m)$ for any choice of signs. Specifically, for $0 \neq m \neq n \neq 0$, there are 8 points in the impulse response that have the same value. These points all lie on a circle of radius $\sqrt{m^2 + n^2}$. The case of a 5×5 filter is

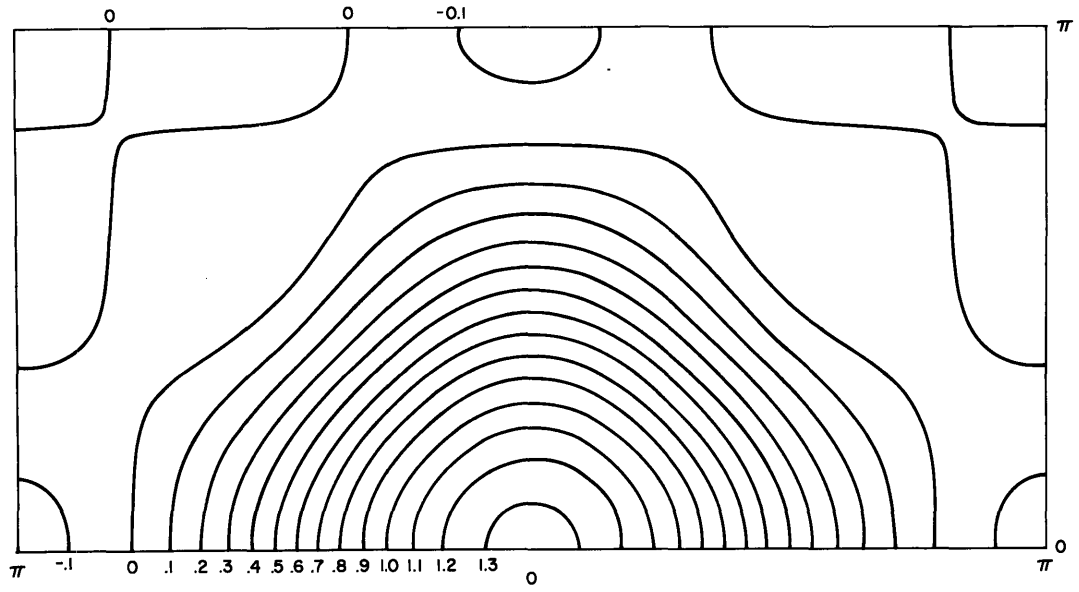

Fig. XV-1. 5 × 5 least-squares 2-D lowpass cutoff = π/2.

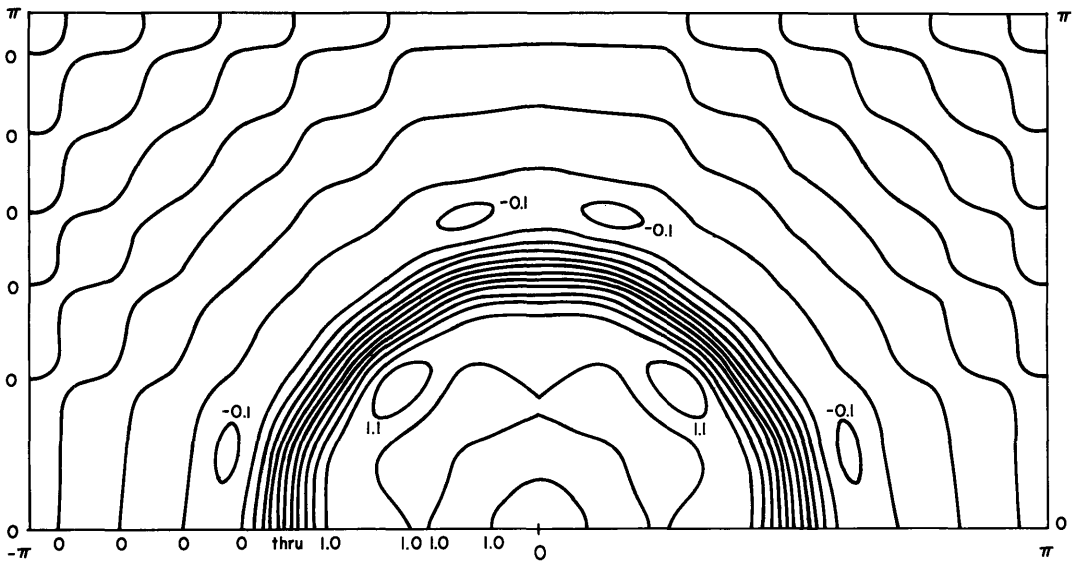

Fig. XV-2. 15 × 15 least-squares 2-D lowpass cutoff = π/2.

f e d e f
e c b c e
d b a b d
e c b c e
f e d e f

Here, h(0, 0) = a, h(0, 1) = h(0, −1) = h(−1, 0) = h(1, 0) = b, etc.

We shall now show that this symmetry produces symmetry about the two axes and the diagonals in the frequency plane. We write the 2-D frequency response and collect like factors of the impulse response to obtain for N × N = (2M+1) × (2M+1):

$$H\left(e^{j\omega_m}, e^{j\omega_n}\right) = h(0,0) + \sum_{m=1}^{M} 2h(m,0)(\cos\omega_m m + \cos\omega_n m)$$

$$+ \sum_{m=1}^{M} 4h(m,m)\cos\omega_n m \cos\omega_m m$$

$$+ \sum_{m=2}^{M}\sum_{n=1}^{m-1} 4h(m,n)(\cos\omega_n n\cos\omega_m m + \cos\omega_n m\cos\omega_m n). \qquad (3)$$

Thus we see that $H\left(e^{\pm j\omega_n}, e^{\pm j\omega_m}\right) = H\left(e^{\pm j\omega_m}, e^{\pm j\omega_n}\right) = H\left(e^{j\omega_m}, e^{j\omega_n}\right)$. We also see that $H\left(e^{j\omega_m}, e^{j\omega_n}\right)$ is purely real. A similar result can be shown when N is even, although we do not consider it here.

Perfect circular symmetry cannot be obtained by FIR filters except in the trivial 1 × 1 case. The octal symmetry represents a reasonable approximation to it. As Fig. XV-1 shows, for low-order filters the approximation may not be too good. Since the least-squares design merely truncates the impulse response in Eq. 2, it is possible to design filters of large extent. Figure XV-2 shows the frequency response contour for a 15 × 15 lowpass filter with cutoff at $\pi/2$.

3. Design Techniques in 1-D Involving a Projection

We now show how, by using the technique of Mersereau and Dudgeon,[5] the 2-D impulse response can be invertibly projected onto a line. The frequency response of this 1-D sequence is the 2-D frequency response evaluated along a series of lines (called "slice" lines) in the frequency plane. Since the projection is invertible, we may map the 2-D ideal filter specifications into 1-D, design the 1-D, and then back-project the impulse response to obtain the 2-D frequency response.

We start with an N × N FIR 2-D filter, as before. We project this array onto a line at angle $\tan^{-1} 1/N$ with the horizontal. This is equivalent to reading out the array

$$
\begin{array}{ccccc}
h(-2,2) & h(-1,2) & h(0,2) & h(1,2) & h(2,2) \\
h(-2,1) & h(-1,1) & h(0,1) & h(1,1) & h(2,1) \\
h(-2,0) & h(-1,0) & h(0,0) & h(1,0) & h(2,0) \\
h(-2,-1) & h(-1,-1) & h(0,-1) & h(1,-1) & h(2,-1) \\
h(-2,-2) & h(-1,-2) & h(0,-2) & h(1,-2) & h(2,-2)
\end{array}
$$

Fig. XV-3. 2-D 5 × 5 impulse response.

$$
h(-2,-2)\ h(-2,-2)\ h(-2,0)\ h(-2,1)\ h(-2,2)\ h(-1,-2)
$$
$$
h(-1,-1)\ \ldots\ h(0,-1),\ h(0,0)\ h(0,1)\ \ldots\ h(2,2)
$$

Fig. XV-4. 1-D projection of 5 × 5 impulse response.

by columns starting at the lower left. For example, the 2-D sequence in Fig. XV-3 would be projected onto the 1-D sequence in Fig. XV-4.

The z-transform of the 2-D sequence is

$$
H(z_m, z_n) = \sum_{m=-M}^{M} \sum_{n=-M}^{M} h(m,n)\, z_m^{-m} z_n^{-n},
$$

and the z-transform of the 1-D projection h(Nm+n) = h(m, n) is

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

and thus

$$
H(z) = H(z^N, z) \quad \text{or} \quad H(e^{j\omega}) = H(e^{jN\omega}, e^{j\omega}). \tag{4}
$$

The line $\omega_n = N\omega_m$ for the case N = 5 is plotted in Fig. XV-5. This figure also shows the ideal filter and the sample points.

The mapping of the ideal filter specifications onto the slice line is mainly a problem in trigonometry and axis scaling. A cutoff circle may cut the slice line more than once. Thus a lowpass circularly symmetric filter (Fig. XV-5), would map into a multiband filter of one pass and two stop bands. The projected 2-D filter is shown in Fig. XV-6.

The problem with the straightforward approach of Mersereau and Dudgeon is that the frequency response of the back-projected filters may deviate wildly from the desired

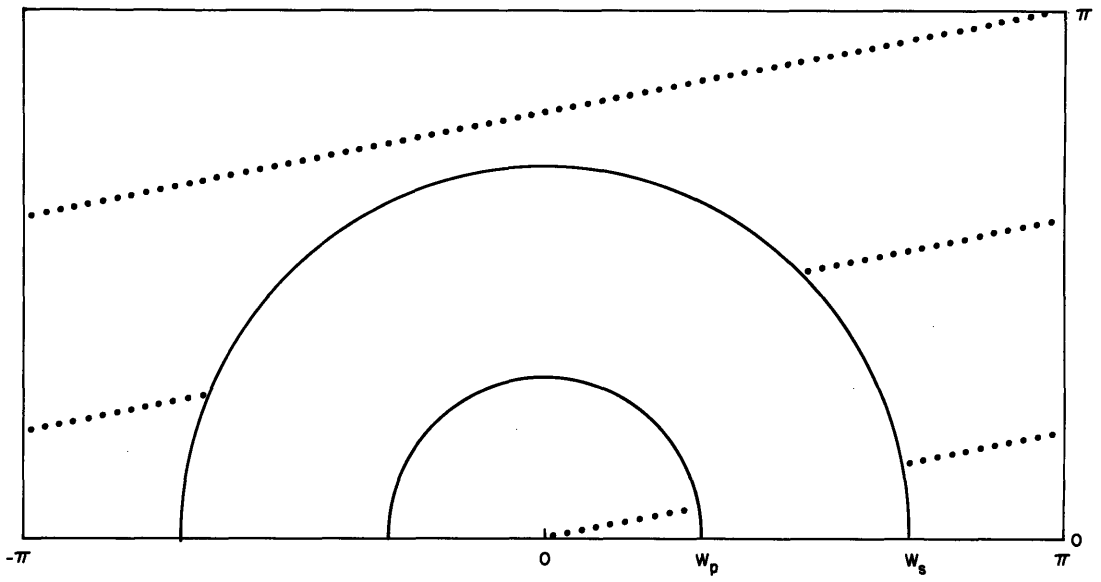

Fig. XV-5. 5 × 5 slice lines with ideal filter specifications. $\omega_p = .3\pi$ and $w_s = .7\pi$.

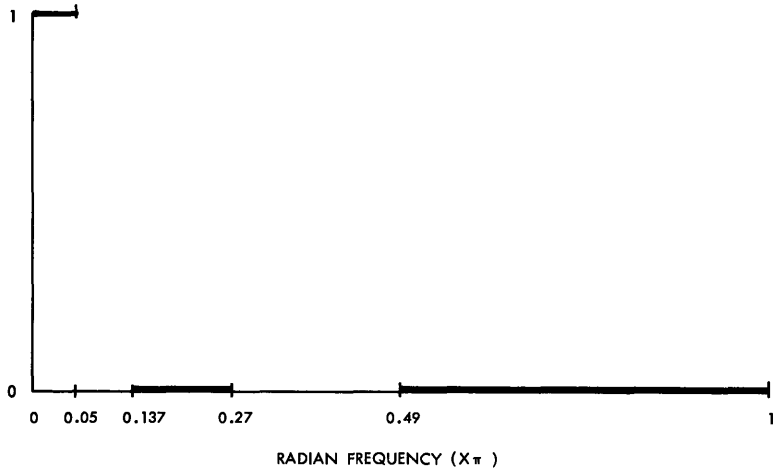

Fig. XV-6. Ideal frequency response of Fig. XV-5 taken along the 5 × 5 slice line.

response between the slice lines. We propose that the symmetry constraints previously discussed can be imposed on the projections to yield equality constraints among the 1-D FIR samples. In this way, we hope to improve the interpolation of the frequency response between the slice lines.

We now apply the octal symmetry constraints to design some filters, by using the least-squares error criterion on the 1-D slice line. From Eq. 3, we obtain the slice response

$$H(e^{j\omega}) = H(e^{jN\omega}, e^{j\omega}) = h(0,0) + \sum_{m=1}^{M} 2h(m,0)(\cos m\omega + \cos mN\omega)$$

$$+ \sum_{m=1}^{M} 4h(m,m) \cos m\omega \cos mN\omega$$

$$+ \sum_{m=2}^{M} \sum_{n=1}^{m-1} 4h(m,n)(\cos n\omega \cos mN\omega + \cos m\omega \cos nN\omega). \qquad (5)$$

Applying the cosine product law, we get

$$H(e^{j\omega}) = h(0,0) + \sum_{m=1}^{M} 2h(m,0)(\cos m\omega + \cos mN\omega)$$

$$+ \sum_{m=1}^{M} 2h(m,m)(\cos(N+1)m\omega + \cos(N-1)m\omega)$$

$$+ \sum_{m=2}^{M} \sum_{n=1}^{m-1} 2h(m,n)(\cos(mN+n)\omega + \cos(mN-n)\omega + \cos(nN+m)\omega$$

$$+ \cos(nN-m)\omega). \qquad (6)$$

In this form, the least-squares result may be obtained by integration to yield h in terms of the lowpass cutoff ω_p:

$$h(0,0) = \frac{\omega_p}{\pi}$$

$$h(m,0) = \frac{1}{2\pi}\left[\frac{\sin m\omega_p}{m} + \frac{\sin mN\omega_p}{mN}\right] \qquad m = 1, 2, \ldots, M$$

$$h(m,m) = \frac{1}{2\pi}\left[\frac{\sin(N+1)m\omega_p}{(N+1)m} + \frac{\sin(N-1)m\omega_p}{(N-1)m}\right] \qquad m = 1, 2, \ldots, M$$

$$h(m,n) = \frac{1}{4\pi}\left[\frac{\sin(mN+n)\omega_p}{(mN+n)} + \frac{\sin(mN-n)\omega_p}{(mN-n)} + \frac{\sin(nN+m)\omega_p}{(nN+m)}\right.$$

$$\left. + \frac{\sin(nN-m)\omega_p}{(nN-m)}\right] \qquad m = 2, 3, \ldots, M \qquad n = 1, 2, \ldots, m-1. \qquad (7)$$

The contour plots in Figs. XV-7 and XV-8, which are to be compared with Figs. XV-1 and XV-2, show that good agreement is obtained by using the two design techniques. As we shall see, the symmetry constraint is a powerful factor in obtaining good back-projected 2-D filters.

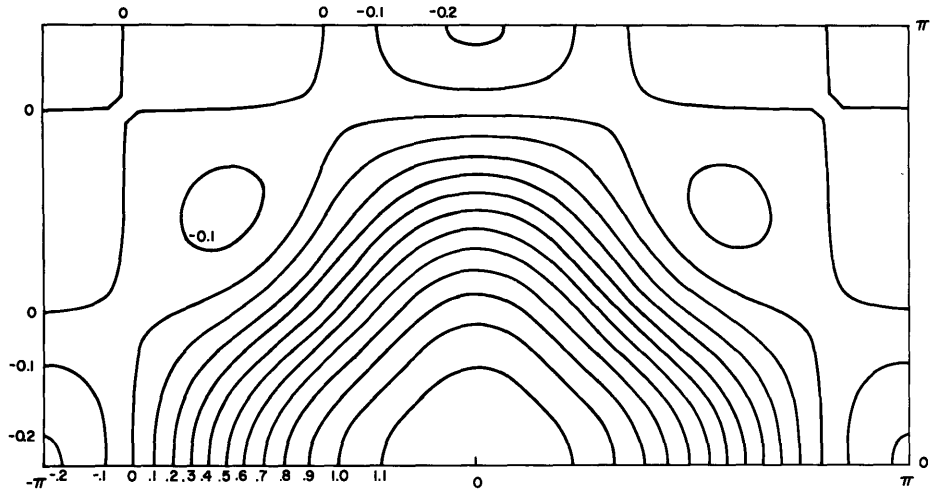

Fig. XV-7.

2-D frequency response designed on 5×5 slice line using least squares.

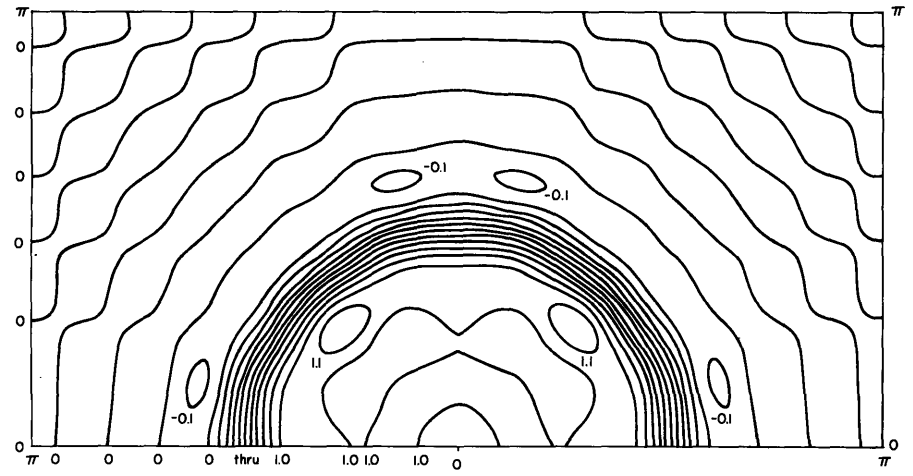

Fig. XV-8.

2-D frequency response designed on 15×15 slice line using least squares.

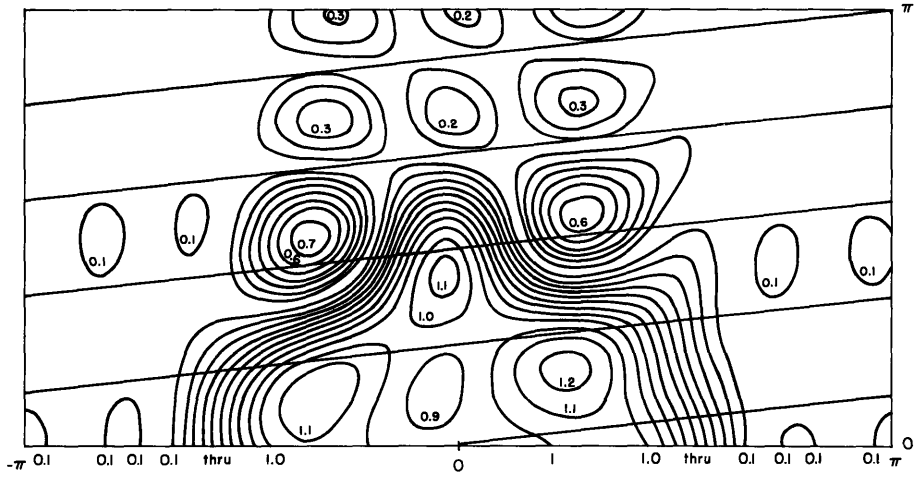

Fig. XV-9. 2-D frequency response of 9×9 Hu filter #1 designed on the slice by the Parks-McClellan algorithm[6] without symmetry constraints.

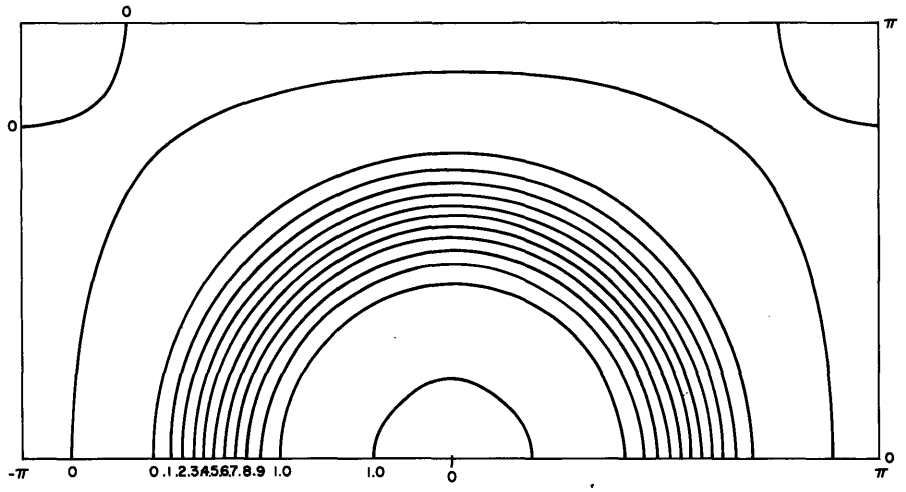

Fig. XV-10. 2-D frequency response of 9 × 9 Hu[1] filter #1 designed by linear programming. $w_p = 4\pi/9$, $w_s = 6\pi/9$.

4. Minimax Approximation on the Slice

We now approximate the ideal filter response on the slice, by using the minimax error criterion. First, we shall require only that the 2-D frequency response be real; that is, h(m, n) = h(−m, −n). The frequency response then becomes

$$H(e^{j\omega}) = H(e^{jN\omega}, e^{j\omega}) = \sum_{k=0}^{(N^2-1)/2} h(k) \cos k\omega, \tag{8}$$

where h(mN+n) = h(m, n). This is precisely the problem solved by the Parks-McClellan algorithm.[6] Second, we shall impose the full octal symmetry constraints. We may not apply the Parks-McClellan algorithm to the latter problem because the set of basis functions does not satisfy the required Haar condition; instead, we have developed in Section XV-B a new algorithm that does not require that the set of basis functions satisfy the Haar condition.

We compare a Parks-McClellan slice-designed lowpass filter to a minimax filter designed in the 2-D frequency plane by linear programming (see Hu and Rabiner[1]). The contour plots are shown in Figs. XV-9 and XV-10, respectively. In these we see that the lack of octal symmetry allows wild variations of the frequency response between the slice lines. In Fig. XV-11 we show that we have designed the same filter by imposing the octal symmetry constraints on the impulse response of the 2-D filter. Computation of the 9 × 9 filter in Fig. XV-11 (15 basis functions) required 20 hours of

HP 9830A calculator time (roughly equivalent to 10 seconds of IBM 370 CPU time). The cost reduction was estimated to be roughly 100:1 and the error was within 7% of the error in the Hu filter. Indeed, the worst error on the 2-D plane occurred at coordinates $(\pi, 0)$, a point through which a slice line will never pass no matter how high the order.

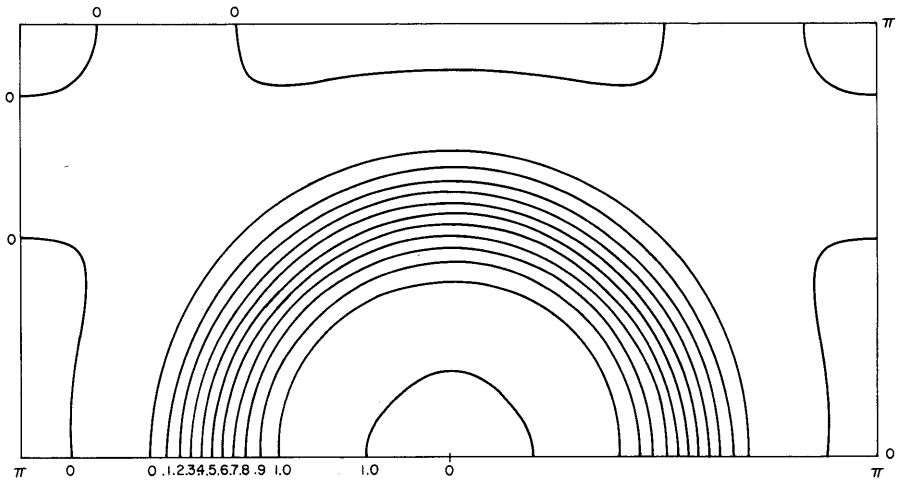

Fig. XV-11. 2-D frequency response of 9 × 9 Hu[1] filter #1 designed on the slice by the new algorithm exploiting the symmetry constraints of the impulse response coefficients.

We have seen that by imposing octal symmetry constraints among the impulse response coefficients, we are able to control to a large extent the frequency response between the slice lines.

5. Optimal 2-D Design

Figure XV-12 shows the sample points of the now familiar slice line for the 9 × 9 filter that we have just designed. The octal symmetry constraint implies that each of these points has a counterpart in the lower triangular portion of the upper frequency plane. At the counterpart, the response is the same as at the original point. The precise form of the sampling of this triangle is shown in Fig. XV-13. It is not surprising, therefore, that the filter is close to optimal (in the minimax sense). Because the sampling is not dense in places, we would be surprised it it were the optimal.

By making two changes in the design program (Fig. XV-14), we were able to design 2-D filters in 2-D. First, the basis functions were generalized to the form of Eq. 3. Second, the set of sample points (an array of 2-D coordinates) was specified as in

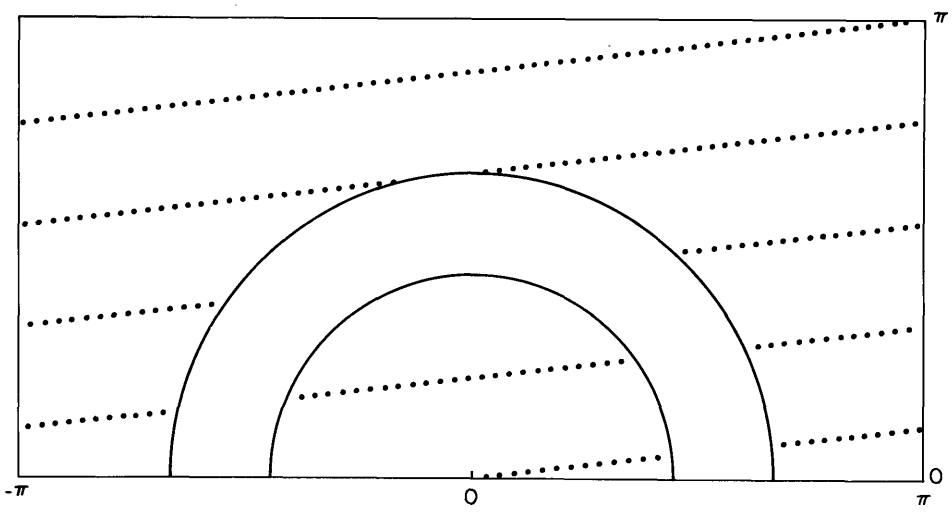

Fig. XV-12. Sample placement along the 9 × 9 slice lines for the Hu[1] filter #1 design problem.

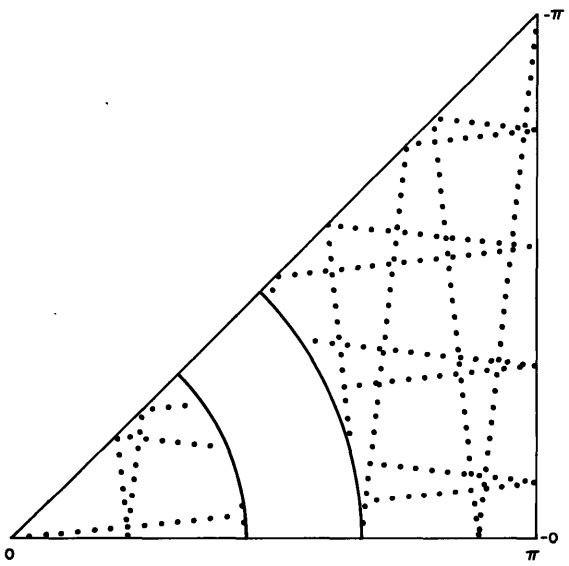

Fig. XV-13.

Sample placement of Fig. XV-12 when the octal symmetry is taken into account.

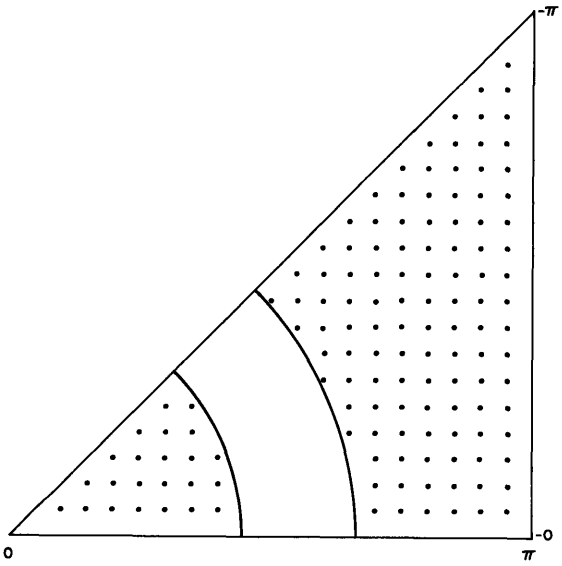

Fig. XV-14.

Placement of sample points for the design problem in 2-D. Hu[1] filter #1 with the total number of sample points approximately the same as in Fig. XV-13.

Fig. XV-12. The number of sample points is roughly the same as the 1-D slice case. These samples emphasize the edge regions, as well as the outline of the triangle.

The final concept of the design of the 2-D filters in 2-D bears a marked similarity to the work of Kamp and Thiran.[3] In addition, however, they have dealt with the problem of nonuniqueness sufficiently to have incorporated it into the design algorithm.

6. Conclusion

We have shown that the slice projection can be used to design nearly optimal 2-D filters when symmetry constraints are imposed. The only time advantage that is apparent is that we need not sample the slice as often as the 2-D plane. In effect, this means that we may design larger filters in the same amount of storage. The price for these larger filters is in the more random placement of points by the 1-D algorithm, rather than the controlled placement by the 2-D algorithm.

References

1. J. V. Hu and L. R. Rabiner, "Design Techniques for Two-Dimensional Digital Filters," IEEE Trans., Vol. AU-20, No. 4, pp. 249-257, October 1972.
2. J. G. Fiasconaro, "Two-Dimensional Non-recursive Digital Filters," Ph. D. Thesis, Department of Electrical Engineering, M. I. T., May 1973.
3. Y. Kamp and J. P. Thiran, "Chebyshev Approximation for Two-Dimensional Nonrecursive Digital Filters," Research Laboratory Report No. R245, Manufacture Belge de Lampes et de Matériel Electronique, S. A., Brussels, March 1974.
4. J. H. McClellan, "The Design of Two-Dimensional Digital Filters by Transformations," Proc. Princeton Conference on Information Sciences and Systems, 1973.
5. R. M. Mersereau and D. E. Dudgeon, "Representation of Two-Dimensional Sequences as One-Dimensional Sequences" (submitted to IEEE Trans. (ASSP)).
6. T. W. Parks and J. H. McClellan, "Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase," IEEE Trans., Vol. CT-19, No. 2, p. 189, March 1972.