A CIRCUIT ANALOGY METHOD FOR THE DESIGN OF RECURSIVE TWO-DIMENSIONAL DIGITAL FILTERS

E. Dubois and M.L. Blostein
Department of Electrical Engineering
McGill University, Montreal, Canada.

ABSTRACT

A technique is presented for the design of stable two-mensional recursive digital filters in which stability of the resulting filters is guaranteed, eliminating the need for the repeated application of stability tests. Essentially, the method is an extension of the one-dimensional bilinear transformation technique. The filter function is obtained by applying a double bilinear transformation to a passive network which is a function of two frequency variables. The method is illustrated by the design of lowpass filters containing circularly symmetric magnitude specifications.

1. INTRODUCTION

Techniques for designing two-dimensional digital filters are of great interest in the fields of picture processing and geophysics[1,2]. Design methods for such filters are currently under heavy investigation, but unlike the one-dimensional case, the theory is largely incomplete. Outstanding problems exist both in the areas of approximation and realization. In approximation, they arise in at least two important ways: first, many aspects of Chebyshev approximation do not extend easily from one to several dimensions[3]. In particular, the characterization of the best approximation no longer possess the simple alternation property so that an optimal filter is very difficult to identify. Secondly, whenever a recursive realization is needed, stability considerations must be taken into account and these place additional constraints on the approximation process. In general, once a filter characteristic has been determined, a check is required to determine whether or not stability requirements are met. Although stability tests[4,5] for two-variable polynomials have been devised, they are very time consuming and usually seriously inhibit a design algorithm if they must be applied repeatedly.

In this paper we present a design method for recursive two-dimensional filters which does not require a stability test. However, it is not clear as to whether the corresponding filter designs are in any sense optimal.

2. THE PROBLEM

The transfer function of a causal recursive two-dimensional digital filter is given by

$$H(z_1,z_2) = \frac{A(z_1,z_2)}{B(z_1,z_2)} = \frac{\sum_{i=1}^{M_A}\sum_{j=1}^{N_A} a_{ij}\, z_1^{i-1} z_2^{j-1}}{\sum_{i=1}^{M_B}\sum_{j=1}^{N_B} b_{ij}\, z_1^{i-1} z_2^{j-1}} \qquad (1)$$

It can be shown[4] that the filter is bounded-input, bounded-output stable if and only if there exist no values of z_1 and z_2 such that $B(z_1, z_2) = 0$ for $|z_1| \le 1$ and $|z_2| \le 1$ simultaneously. The frequency domain design problem is to choose the various a_{ij} and b_{ij} so that $H(e^{-j\omega_1 T}, e^{-j\omega_2 T})$ approximates an ideal response $\hat{H}(\omega_1, \omega_2)$ while $B(z_1, z_2)$ remains stable.

We denote the approximating function by $H_1(\underline{x}, \underline{\omega})$ where $\underline{x}$ represents a vector of design parameters and $\underline{\omega} = (\omega_1, \omega_2)^T$. Since the transfer function is periodic in $\underline{\omega}$, only values of $\underline{\omega}$ in the region $W = \{\omega_1, \omega_2 : -\frac{\pi}{T} \le \omega_1, \omega_2 \le \frac{\pi}{T}\}$ need be considered. We also define the error function $r(\underline{x}, \underline{\omega}) = \hat{H}(\underline{\omega}) - H_1(\underline{x}, \underline{\omega})$ and the stability region $S_x = \{\underline{x} : B \text{ is stable}\}$.

An approach often used in this type of problem is to minimize the L_p norm of $r(\underline{x}, \underline{\omega})$ over S_x. The optimal solution, denoted by $\underline{x}_p$, is given by

$$\|r(\underline{x}_p, \underline{\omega})\|_p = \inf_{\underline{x}\in S_x} \|r(\underline{x}, \underline{\omega})\|_p = \inf_{\underline{x}\in S_x} \left[\int_W |r(\underline{x}, \underline{\omega})|^p \, d\omega \right]^{1/p} \qquad (2)$$

The limiting case as $p \to \infty$ gives the minimax approximation

$$\|r(\underline{x}_\infty, \underline{\omega})\|_\infty = \inf_{\underline{x}\in S_x} \|r(\underline{x}, \underline{\omega})\|_\infty = \inf_{\underline{x}\in S_x} \max_{\underline{\omega}\in W} |r(\underline{x}, \underline{\omega})| \qquad (3)$$

If $\underline{x}_\infty$ is unique, then $\lim_{p\to\infty} \underline{x}_p = \underline{x}_\infty$. However, in the two-dimensional case uniqueness is not guaranteed.[3] Also, there is no effective characterization of the minimax approximation, as in the one-dimensional case where the error curve possesses simple alternation properties.

For many applications, $\hat{H}(\omega)$ has the form

$$\hat{H}(\omega) = \begin{cases} 1 & \underline{\omega} \in P \\ 0 & \underline{\omega} \in S \end{cases}$$

where P and S are pass and stop regions which may be separated by a transition region. The problem can then be

E. Dubois and M.L. Blostein, "A circuit analogy method for the design of recursive two-dimensional digital filters," in Proc. IEEE International Symposium on Circuits and Systems (ISCAS 1975), Newton, MA, April 1975, pp. 451-454.

stated as a constrained optimization, namely to minimize

$$\max_{\underline{\omega} \in S} |H_1(\underline{x}, \underline{\omega})|$$

subject to the constraint (4)

$$1 - \epsilon \le |H_1(\underline{x}, \underline{\omega})| \le 1 + \epsilon \qquad \underline{\omega} \in P$$

3. A "TWO-DIMENSIONAL CIRCUIT" ANALOGY

In one-dimensional recursive digital filter design a bilinear transformation is often applied to an analog filter transfer function with the desired characteristics to obtain the digital filter transfer function. The rationale for this approach is that it is desirable to exploit the considerable body of knowledge which has been built up in analog filter theory. Although little work has been carried out in two-dimensional analog frequency domain design, studies have been directed at developing a two-dimensional circuit theory[6]. Bearing in mind that it would be very desirable to have a class of network functions which are guaranteed to be stable, thus avoiding repeated stability tests, the following "two-dimensional circuit" analogy can be proposed: design a two-dimensional analog filter and apply a two-dimensional bilinear transformation to obtain the two-dimensional recursive digital filter. This method is analogous to the one-dimensional case, but the underlying motivation is different (guaranteed stability rather than previous experience).

Definition[6]: A finite, passive network of two variables is a network composed of finite numbers of two-terminal elements whose impedances are proportional to p_1, $1/p_1$, p_2, $1/p_2$ with positive coefficients, positive resistors, ideal transformers, and ideal gyrators.

Given a passive, two-variable n-port network N, the usual network functions such as driving point impedance and voltage transfer can be defined. Given such a function $H'(p_1, p_2)$, the transfer function of a digital filter $H(z_1, z_2)$ can be obtained by means of the bilinear transformation

$$p_i = \frac{1 - z_i}{1 + z_i} \qquad i = 1, 2 \tag{5}$$

The following statement can then be made about the resulting digital filter.

Assertion: The digital filter with transfer function $H(z_1, z_2)$ obtained by performing a bilinear transformation on a network function $H'(p_1, p_2)$ of a two-variable passive network is at worst marginally stable.

Proof: From Huang[4], $H(z_1, z_2)$ is stable if and only if

1) $H'(j\omega, p_2)$ has no poles in $\mathrm{Re}(p_2) \ge 0$ for all positive ω_1.
2) $H'(p_1, 1)$ has no poles in $\mathrm{Re}(p_1) \ge 0$.

$H'(j\omega, p_2)$ represents the network function of a one-dimensional passive filter with imaginary elements, and thus has no poles in $\mathrm{Re}(p_2) \ge 0$. Similarly $H'(p_1, 1)$ has no poles in $\mathrm{Re}(p_1) \ge 0$. Hence, only marginal instability can occur, namely if $H'(j\omega, p_2)$ or $H'(p_1, 1)$ has j-axis poles. This can generally be avoided by choosing N to be a lossy network.

Two-variable networks have been used in the study of networks consisting of both lumped and distributed elements. Koga has shown that an arbitrarily prescribed n by n positive real matrix of two-variables is realizable as the impedance or admittance matrix of a finite passive n-port of two variables. However, it is not known if all stable transfer functions can be obtained as a network function of some two-variable circuits. This is a much more difficult problem, and its solution would indicate whether or not the class of transfer functions obtained in this way is restrictive.

4. THE DESIGN PROCEDURE

The two-dimensional circuit analogy can be used to generate a frequency domain design algorithm for recursive digital filters in a straightforward manner. The basic idea is to select a two-variable passive network which possesses qualitatively an approximation to the desired response. The corresponding digital transfer function can then be generated using a double bilinear transformation and the magnitude and/or phase can be optimized over a finite point set. In this particular procedure the members of $\underline{x}$, the set of design parameters, are the parameters of the two variable network. Thus, stability is assured as long as these values are positive, i.e. S_x is a very simple set. If the optimization were applied directly to the coefficients of $H(z_1, z_2)$, a complex stability test would be needed after each iteration, since S_x can no longer be described simply.

Near minimax solutions can be obtained by using an L_p norm with large p as an objective function. A suitable performance function has the form

$$J(\underline{x}) = \{ \sum_{\omega \in W} | r(\underline{x}, \underline{\omega})\, u(\underline{\omega}) |^p \}^{1/p} \tag{6}$$

where $u(\underline{\omega})$ is a weighting function. If $\nabla J(\underline{x})$ is needed for the minimization process, the gradient of the analog transfer function with respect to the designable parameters must be computed. The corresponding adjoint[7] to the two-variable network can be used to enhance the efficiency of such computations.

$J(\underline{x})$ can be modified for greater flexibility by restricting W to those points where the error function is larger than some threshold $\xi(\omega)$.[8] That is, we can write

$$J(\underline{x}) = M \Big[\sum_{\omega \in W'} \Big[\frac{[|r(\underline{x}, \underline{\omega})| - \xi(\underline{\omega})]\, u(\omega)}{M} \Big]^p \Big]^{1/p} \tag{7}$$

where

$$W' = \{ \underline{\omega} \in W : \; |r(\underline{x}, \underline{\omega})| > \xi(\omega) \}$$

$$M = \max_{\underline{\omega} \in W'} (|r(\underline{x}, \underline{\omega})| - \xi(\underline{\omega}))\, u(\omega)$$

M is introduced to reduce ill-conditioning which may arise for large values of p. If $M < 0$ at some stage in the optimization $\xi(\omega)$ must be attenuated by a constant factor and the procedure restarted.

5. APPLICATION TO LOW-PASS CIRCULARLY SYMMETRIC FILTERS

As an example of the application of the circuit analogy method, consider the design of low-pass filters with circularly symmetric magnitude specifications, i.e. filters which are insensitive to the relative orientation of data. The simple doubly-terminated LC ladder can be used as the two-variable, passive network (Figure 1). The designable parameter set then consists of the various L's and C's along with the terminating resistance and each reactive element can be associated with either a p_1 or a p_2 frequency variable.

Experimental designs were carried out using this approach assuming pass and stopband edges of $0.14\pi/T$ and $0.26\pi/T$, respectively. Sample points of W were chosen along radials and were sufficiently dense to adequately represent the response surface. Since the response of Figure 1 decreases monotonically in the stopband, only one stopband radial was included in W. Minimization of $J(\underline{x})$ was carried out by a quasi-Newton method using the complementary DFP update[9]. $p = 20$ was used as $\|r(\underline{x}_p, \underline{\omega})\|_\infty$ did not decrease significantly for $p > 20$.

The method was applied to second, fourth and eighth order low-pass ladders. In all cases acceptable results were obtained only when the p_1 and p_2 variables were cascaded (Figure 2). Other distributions of p_1 and p_2 yielded poor responses. A Butterworth or Chebyshev design along $p_1 = p_2$ was usually selected to initiate the procedure.

Significantly, each design yielded a limiting solution in which some elements vanished, thereby producing a separable transfer function. For example, in the 4th order case

$$H(z_1, z_2) = \frac{(1/34.8)^2 (1 + z_1)^2 (1 + z_2)^2}{(1 - 1.62 z_1 + 0.705 z_1^2)(1 - 1.62 z_2 + 0.705 z_2^2)}$$

with $\|r(\underline{x}, \omega)\|_\infty = 0.2545$. In the 8th order case $\|r(\underline{x}, \omega)\|_\infty = 0.0708$. Contour plots of the responses are shown in Figures 3 and 4, respectively (the contours are displayed in steps of 0.1 relative to a desired passband of unity).

Since the numerator of (1) does not affect stability, transfer functions obtained by the preceeding approach can be multiplied by

$$A'(p_1, p_2) = \sum_{m=1}^{M_A} \sum_{n=1}^{N_A} a'_{mn}\, p^{m-1}\, p^{n-1}$$

where the coefficients a'_{mn} become part of the parameter vector $\underline{x}$. This can introduce transmission zeros in the stopband and a more dense set of stopband points must be used. Figure 5 shows the response of the 4th order filter with the modified numerator. The response represents a substantial improvement over Figure 3 in both pass and stopbands. Note that the transfer function no longer possesses a separable numerator, although the denominator remains separable.

5. CONCLUSION

A method for the design of stable two-dimensional recursive digital filters which does not require an explicit stability test has been presented. In the examples tested, the algorithm converged to solutions as good or better than previously published results. Moreover, it offers the opportunity to design higher order optimal filters than have been reported to date. Further testing of the method is required to obtain a proper evaluation of its effectiveness and a number of substantial questions remain to be answered. Two of the most important are: 1) What synthesis procedures can be used to generate the seed analog two-variable, passive network, and 2) does the method constrain the nature of the resulting magnitude and/or phase response functions?

REFERENCES

1. H.C. Andrews, Computer Techniques in Picture Processing. New York : Academic Press, 1970.
2. E.G. Zurflueh, Applications of Two-Dimensional Linear Wavelength Filtering, Geophysics, Vol. 32, pp. 1015 - 1035.
3. J.R. Rice, The Approximation of Functions - Vol. II. Reading, Mass. : Addison-Wesley, 1969.
4. T.S. Huang, Stability of Two-Dimensional Recursive Filters, IEEE Transactions, Vol. AU-20, pp. 158-163, June 1972.
5. B.D. Anderson and E.I. Jury, Stability Test for Two-Dimensional Recursive Filters, IEEE Transactions, Vol. AU-21, pp. 366-372, August 1973.
6. T. Koga, Synthesis of Finite Passive N-ports with Prescribed Positive Real Matrices of Several Variables, IEEE Transactions, Vol. CT-15, pp. 2-23, March 1968.
7. S. Director and R. Rohrer, Automated Network Design - Frequency Domain, IEEE Transactions, Vol. CT-16, pp. 330-337, August 1969.
8. J.W. Bandler and C. Charlambous, Practical Least-pth Optimization of Networks, IEEE Transactions, Vol. MTT-20, pp. 834-840, December 1972.
9. P.E. Gill and W. Murraw, Quasi-Newton Methods for Unconstrained Optimization, J. Inst. Math. and Appl., Vol. 9, pp. 91-108, February 1972.

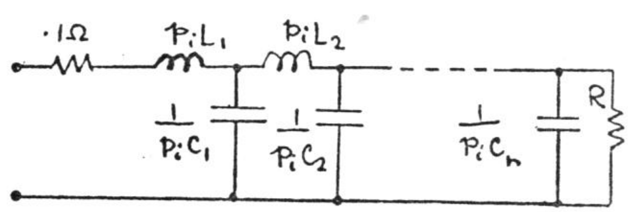

Figure 1

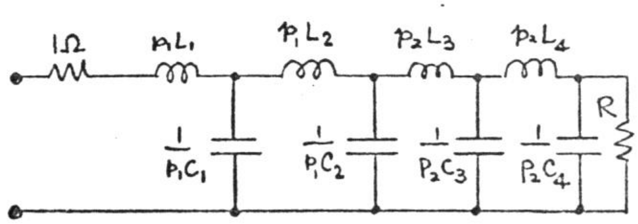

Figure 2

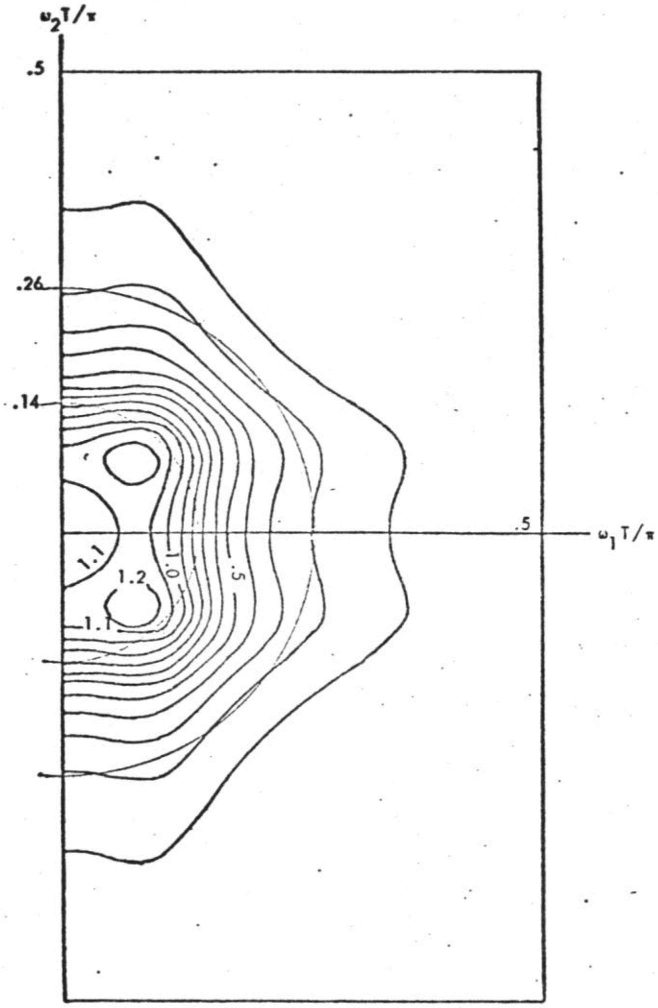

Figure 3

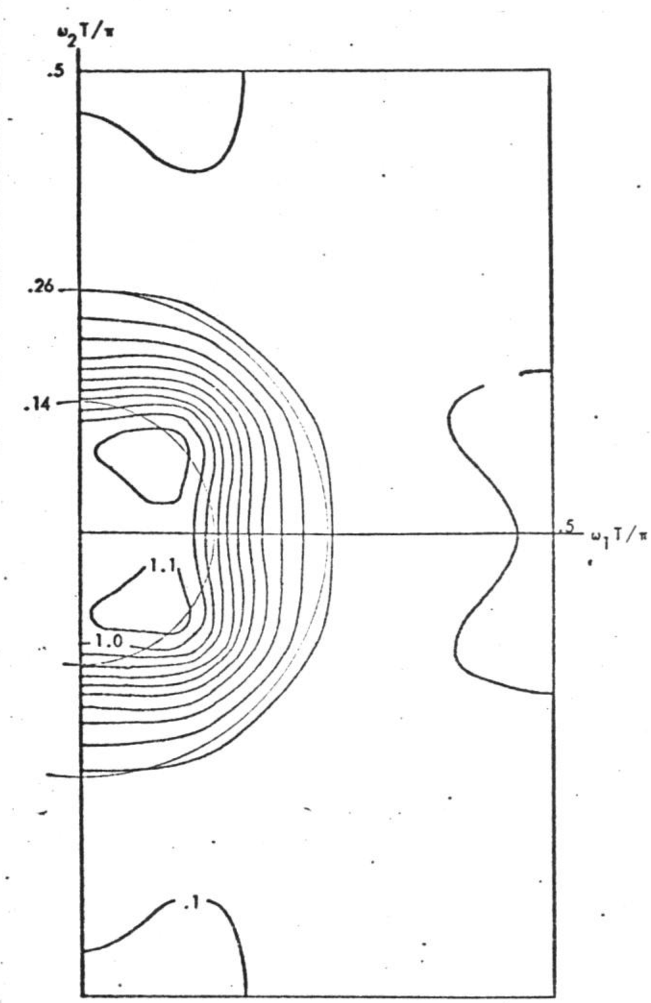

Figure 5

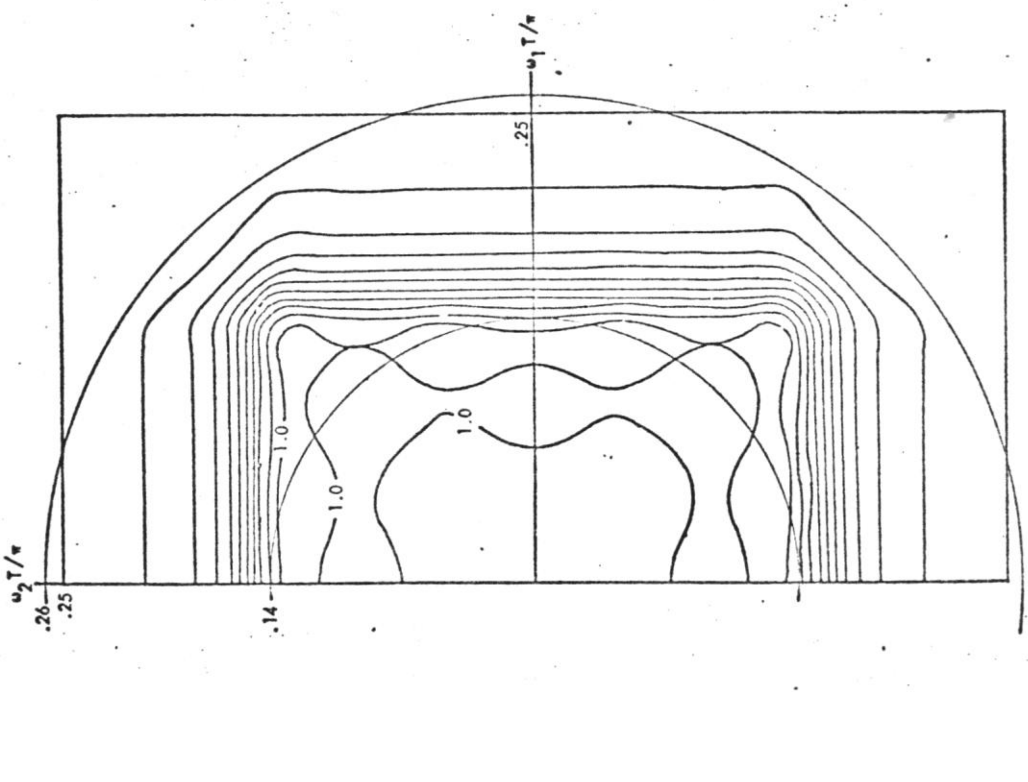

Figure 4