12-1-1973

Considerations of the Pade Approximant Technique in the Synthesis of Recursive Digital Filters

Francis Brophy

Andres C. Salazar

Reprinted by permission from IEEE TRANSACTIONS ON AUDIO AND ELECTROACOUSTICS
Vol. AU-21, No. 6, December 1973, pp. 500-505

Considerations of the Padé Approximant Technique in the Synthesis of Recursive Digital Filters

FRANCIS BROPHY and ANDRES C. SALAZAR, Member, IEEE

Abstract—The Padé approximant technique provides a quick design of recursive digital filters. An added advantage of the technique lies in that spectrum shaping requirements as well as linear phase constraints can be handled easily, even for higher order filters. This is important in supplying initial guesses of the filter parameters to iterative routines that would then seek a locally optimal design solution. These advantages are among those discussed in a partly tutorial presentation of the technique that relates to filter needs found in data transmission systems. In addition, the question of stability is treated and a new criterion is presented. The criterion provides sufficient conditions in establishing stability for a filter designed by using the Padé approximant technique.

I. Introduction

The design of spectrum-shaping recursive digital filters in the z-plane often requires the use of a routine that calculates the extremum of an object function of several variables. The function is generally nonlinear and positive definite and indicates the "closeness" of the designed spectrum to the desired spectrum. In some cases, depending on the complexity of the function, the number of iterations or even convergence to an extremum is dependent on the initial guess for the α and β (feedforward and feedback) parameters.

By working in the time domain the degrees of freedom available can be used to match a set of time samples exactly, thus reducing the design to the solution of a linear system of equations. While this approach, call the Padé approximant,[1] does not lead to a locally optimal solution as an iterative technique would, it nevertheless provides a viable solution in a fraction of the time.

In the following we show how the Padé approximant technique can yield a simple digital filter design for spectrum shaping networks with linear (or nonlinear, if so desired) phase constraints often required in data transmission systems. The problem of stability is discussed and sufficient conditions are given to ensure that the design procedure will not lead to an unstable filter.

II. Padé Approximate in Digital Filter Design

Let $H(\omega)$ denote in the interval $[-2\pi W, 2\pi W]$ the bounded filter amplitude characteristic that is to be synthesized. Since $H(\omega) \in L_p[-2\pi W, 2\pi W]$, $p \geqslant 1$ it has a Fourier series expansion

$$H(\omega) = \sum h_n e^{-jn\omega/2W} \qquad (1)$$

Manuscript received December 28, 1972; revised June 29, 1973.
The authors are with Bell Laboratories, Holmdel, N.J. 07733.

[1] "Prony's method" is related to the Padé approximant technique through a transformation of variables (see [1] for more details).

where h_n are the samples of the impulse response of $H(\omega)$ at the appropriate sampling period $1/2W$ s.

Also, given $\epsilon > 0$, there exists an integer K such that

$$\frac{1}{2W} \sum_{-K}^{K} h_n^2 \geq \frac{1}{2\pi} \int_{-2\pi W}^{2\pi W} |H(\omega)|^2 \, d\omega - \epsilon.$$

In fact, the associated Fourier expansion of $H(\omega)$, viz.,

$$\sum_{-K}^{K} h_n e^{-jn\omega/2W},$$

is the best approximation in the mean (L_2) to $H(\omega)$ on the interval $[-2\pi W, 2\pi W]$ using only $2K + 1$ coefficients.

By adding the constant delay $K/2W$ to $H(\omega)$ we can reindex the samples $\{h_n\}_{-K}^{\infty}$ to $\{h_n\}_0^{\infty}$. Also, normalization on $H(\omega)$ can be induced by dividing $\{h_n\}_0^{\infty}$ by $h_0 \neq 0$ so that $\{h_n\}_0^{\infty}$ now represents the samples of a time function whose Fourier transform closely approximates a delayed and normalized version of $H(\omega)$.

A realizable digital filter transfer function has the form of a rational function[2] in z^{-1}:

$$G(z^{-1}) = \frac{1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \cdots + \alpha_M z^{-M}}{1 - \beta_1 z^{-1} - \beta_2 z^{-2} - \cdots - \beta_N z^{-N}}. \quad (2)$$

If $z^{-1} = 0$ is not a point of singularity then

$$G(z^{-1}) = \sum_{n=0}^{\infty} g_n(\vec{\alpha}, \vec{\beta}) z^{-n} \quad (3)$$

is the Maclaurin series where $\vec{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_M)$ and $\vec{\beta} = (\beta_1, \beta_2, \cdots, \beta_N)$ for a positive radius of convergence $0 < |z^{-1}| < r_0$. If stability constraints[3] are placed on the β's we then have that $r_0 \geq 1$. The term $g_n(\vec{\alpha}, \vec{\beta})$ is easily found recursively from

$$g_n(\vec{\alpha}, \vec{\beta}) = \beta_1 g_{n-1}(\vec{\alpha}, \vec{\beta}) + \cdots + \beta_N g_{n-N}(\vec{\alpha}, \vec{\beta}) + \alpha_n \quad (4)$$

where $g_0(\vec{\alpha}, \vec{\beta}) = 1$ and $\alpha_n = 0$ if $n > M$.

If we represent h_n as the response samples of the digital filter we wish to synthesize we can equate (1), ($z = e^{jn\omega/2W}$), and (4) and find equations involving the α and β parameters:

$$h_n = \beta_1 g_{n-1} + \beta_2 g_{n-2} + \cdots + \beta_N g_{n-N} + \alpha_n, \quad 1 < n < M \quad (5)$$

$$h_n = \beta_1 g_{n-1} + \beta_2 g_{n-2} + \cdots + \beta_N g_{n-N}, \quad n > M \quad (6)$$

where the arguments of g_n have been dropped. Clearly there are not enough degrees of freedom in $G(z^{-1})$ upon fixing M and N to generate any response $\{h_n\}_{n=0}^{\infty}$ that satisfies $\Sigma |h_n| < \infty$. However, there is a possibility of minimizing $\mathcal{E}$, the weighted squared difference between the two responses through the Tth sample, $T \geq M + N$:

$$\mathcal{E} = \sum_{n=1}^{T} W_n (h_n - g_n)^2, \; W_n \geq 0, \quad n = 1, 2, \cdots T.$$

Since g_n is a nonlinear function of the α and β parameters the minimization of $\mathcal{E}$ for $T \geq M + N$ can be achieved only through iterative means. In one case, however, viz., $T = M + N$, the α and β parameters which minimize $\mathcal{E}$ can be found by solving $M + N$ linear equations.[4] By assuming $g_n = h_n$, $n = 1, \cdots M$, we can then solve (6) for the β's that produce $g_n = h_n$, $n = M + 1, M + 2, \cdots, M + N$. Further we note that for a given set of $\{\beta_k\}_{k=1}^{N}$ we can solve (5) for α's so that $g_n = h_n$, $n = 1, 2, \cdots, M$. This procedure amounts to equating the truncated power series of (1), i.e.,

$$\sum_{n=0}^{M+N} h_n z^{-n},$$

to the first $M + N + 1$ terms of the MacLaurin series (3). This approximation of a power series by a rational function is commonly referred to as the Padé approximate procedure. This technique generates the $M + N + 1$ samples $\{h_n\}_0^{M+N}$ exactly but produces a tail $\{g_n\}_{M+N+1}^{\infty}$ that can serve to yield a better approximation to $H(\omega)$.

Time domain synthesis of *analog* filters (i.e., s plane) utilizing this method is well documented [3]. By approximating the desired digital filter response by reproducing the first $M + N + 1$ samples (h_0 is given free by initial normalization) it is hoped that the overall time or frequency response from the approximating filter will not deviate unreasonably from the desired one. There is, however, no guarantee that the β's compose a stable filter. In Section III we will discuss stability considerations of the Padé technique.

III. Stability of the Padé Approximant

Since the Padé approximant produces a rational function approximation, we know how the stability of the approximation can be related to the moduli of the poles of the approximant. We would like to know, however, if we can relate the stability of the proposed Padé approximant to some measure other than the poles.

We can state the matrix equation of the Padé approximant in the following form by reindexing $\{g_n\}_0^{\infty}$ and $\{h_n\}_0^{\infty}$ to $\{g_n\}_{-N+1}^{\infty}$ and $\{h_n\}_{-N+1}^{\infty}$, respectively.

[2] For notational convenience G is normalized to 1 at $z^{-1} = 0$.
[3] A filter is stable if bounded output results from bounded input.

[4] This special case has been discussed in [2] and [6] for synthesis of a "time" sequence $\{h_n\}$.

$$(\beta_1 \cdots \beta_N)\begin{bmatrix} h_0 & h_1 & \cdots & h_{N-1} \\ h_{-1} & h_0 & & h_{N-2} \\ \vdots & h_{-1} & & \vdots \\ & \vdots & & \\ h_{-N+1} & h_{-N+2} & \cdots h_{-1} & h_0 \end{bmatrix} = \begin{bmatrix} h_1 \\ \vdots \\ h_N \end{bmatrix}^t.$$

We define

$$Q_i = \begin{bmatrix} q_i & q_{i+1} & \cdots & q_{i+N-1} \\ q_{i-1} & q_i & & \vdots \\ \vdots & & & \\ q_{i-N+1} & & \cdots & q_i \end{bmatrix} \quad (9)$$

where $q_i = h_i$ or g_i and $Q_i = H_i$ or G_i. Note that for the Toeplitz forms G_i there is a recurrence relation:

$$G_{i+2} = BG_{i+1} \text{ or } G_{i+2} = B^{i+2} H_0$$

since $G_0 = H_0$ and $G_1 = H_1$, where B is the $N \times N$ companion[5] matrix, $B = H_1 H_0^{-1}$.

$$B = \begin{bmatrix} \beta_1 \; \beta_2 & & & \beta_N \\ 1 \;\; 0 & & & 0 \\ & 1 & & 0 \\ & & 1 & \cdot \\ & & \ddots & \cdot \\ & & & \cdot \\ & & & 1 \;\; 0 \end{bmatrix}. \quad (10)$$

We note that

$$H_1 = BH_0 \quad (11)$$

where B has eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_N\}$.[6] We know, in addition, that the eigenvector corresponding to λ_i is $(1, \lambda_i^{-1}, \lambda_i^{-2}, \cdots, \lambda_i^{-N+2}, \lambda_i^{-N+1})$. We can decompose B in terms of this eigensystem to get

$$B = P \Lambda P^{-1} \quad \text{and} \quad G_n = P \Lambda^n P^{-1} H_0 \quad (12)$$

where Λ is a diagonal eigenvalue matrix and P is an $N \times N$ matrix whose columns are the eigenvectors of B.

From (12) we can see that a bounded impulse response $\{g_n\}$ is possible only if $|\lambda_k| \leq 1$, $k = 1, 2, \cdots N$. Since $g_n \to 0$ necessarily for a stable filter, this means $|\lambda_k| < 1$. The converse is also true, i.e., $|\lambda_k| < 1$ implies filter stability. This can be most easily seen by noting that the eigenvalues of B are precisely the roots of the denominator polynomial of $G(z^{-1})$ since B is its companion matrix. This development, of course, is well known in linear system theory. The point we make here, however, is that the expression for B, namely $B = H_1 H_0^{-1}$, involves only the first $2N + 1$ samples of the target sequence $\{h_n\}$. Hence, we can test $H_1 H_0^{-1}$ for its eigenvalue of maximum modulus. If it is less than one, the Padé procedure yields a stable filter. Hence, a valuable condition for stability of the synthesized Padé filter is easily obtained in this fashion.

To see this, let us index the column vectors of H_0 by $\{\vec{h}_0, \vec{h}_1, \cdots, \vec{h}_{N-1}\}$ where $\vec{h}_k^t = (h_{k0}, h_{k1}, \cdots, h_{kN-1})$ for $k = 0, 1, 2, \cdots, N-1$. Since H_0 was assumed invertible we have then that $\{\vec{h}_k\}_{k=0}^{N-1}$ form a linearly independent set of vectors. Let $\vec{x}$ be a nonzero vector in R^N (Euclidean N space). Then there exist constants $\gamma_0, \gamma_1, \cdots, \gamma_{N-1}$ (not all zero) such that $\vec{x} = \sum_{k=0}^{N-1} \gamma_k \vec{h}_k$. Hence, the squared Euclidean norm of $\vec{x}$ is

$$\|\vec{x}\|^2 = \left\| \sum_{k=0}^{N-1} \gamma_k \vec{h}_k \right\|^2 = \sum_{l=0}^{N-1} \left\{ \sum_{k=0}^{N-1} \gamma_k h_{kl} \right\}^2$$

$$\|\vec{x}\|^2 = \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} \gamma_k \gamma_m \sum_{l=0}^{N-1} h_{kl} h_{ml}. \quad (13)$$

Now define

$$c_{km} = \sum_{l=0}^{N-1} h_{kl} h_{ml},$$

which forms a strictly positive definite matrix, C_0, Similarly, we can form

$$\|B\vec{x}\|^2 = \left\| \sum_{k=0}^{N-1} \gamma_k B\vec{h}_k \right\|^2. \quad (14)$$

But we recall from (11) that $B\vec{h}_k = \vec{h}_{k+1}$, $k = 0, 1, \cdots, N-1$, [where $\vec{h}_N^t = (h_{N0}, h_{N1}, \cdots, h_{NN-1})$]. Thus, we can now write

$$\|B\vec{x}\|^2 = \sum_{l=0}^{N-1} \left\{ \sum_{k=0}^{N-1} \gamma_k h_{k+1,l} \right\}^2$$

$$= \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} \gamma_k \gamma_m \sum_{l=0}^{N-1} h_{k+1,l} h_{m+1,l}. \quad (15)$$

Form

$$c_{k+1,m+1} = \sum_{l=0}^{N-1} h_{k+1,l} h_{m+1,l},$$

$$k, m = 0, 1, \cdots, N-1, \quad (16)$$

a positive definite matrix, C_1. We are now ready to formulate the following.

Stability Criterion: The digital filter formed by the

[5] H_0 is invertible since it is nonsingular for cases of interest, (e.g., $H(z^{-1})$ is not a rational function). Fast techniques for inverting H_0 are available [4].

[6] It does not seem overly restrictive to assume simple roots here.

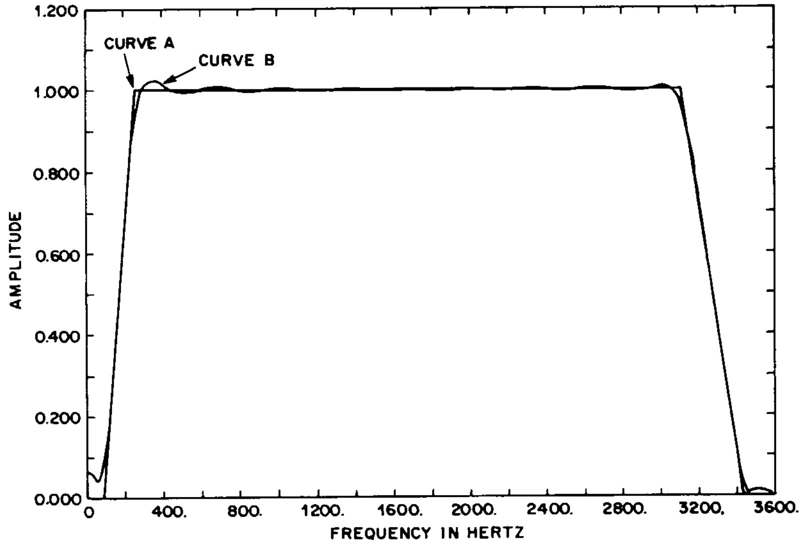

Fig. 1. Amplitude spectrum for bandpass filter (24th order).

Padé approximant procedure outlined in Section II for $M = N$ and invertible H_0 is stable if $C_0 - C_1 > 0$, that is, the matrix is a strictly positive definite matrix.

Proof: Let us begin by first defining $\vec{\gamma} = (\gamma_0, \gamma_1, \cdots, \gamma_{N-1})$. Then $(C_0\vec{\gamma}, \vec{\gamma}) = \|\vec{x}\|^2$ from (13) and the definition of C_0 while $(C_1\vec{\gamma}, \vec{\gamma}) = \|B\vec{x}\|^2$ from (15) and (16). Hence, $((C_0 - C_1)\vec{\gamma}, \vec{\gamma}) > 0$, for nontrivial $\vec{x}$. So $\|\vec{x}\|^2 - \|B\vec{x}\|^2 > 0$ or $\|B\vec{x}\| < \|\vec{x}\|$ for arbitrary $\vec{x}$ so conclude $\|B\| < 1$. But we know that the eigenvalue of maximum modulus of B, namely $|\lambda_{\max}(B)|$ is bounded above by $\|B\|$ so we have the desired conclusion.

Armed with this stability criterion we check for the possibility of instability in a Padé designed filter by simply following three steps.

Step 1: Obtain the time sequence $\{h_k\}_{-N+1}^{N-1}$.

Step 2: Form the matrices C_0 and C_1, (total of $N(N+3)/2$ inner products are necessary).

Step 3: Calculate the principal minors of $C_0 - C_1$ and check for positivity.

The last step is the well-known test for positive definiteness of a matrix [5, p. 59]. The usefulness of following the procedure is appreciated when the alternatives for checking stability are considered. For example, stability can be determined by simply solving for the set of β's and then applying a root finding routine to the resulting polynomial. Equivalently, we can compute $H_1 H_0^{-1} = B$ and then apply an eigenvalue-eigenvector routine to the matrix. Last, we consider solving for the β's and then computing the response $\{g_n\}_{N+1}^{T}$ of (4). If the response is seen to decrease the filter is stable. In all three cases, a sizable amount of computation is involved when compared to the three step stability algorithm discussed above.

To avoid misrepresentation, it must be pointed out that stability is still possible even if $C_0 - C_1 \not> 0$. In fact, if h_N of $\{h_n\}_0^{\infty}$ is chosen to be the peak of the impulse response of $H(\omega)$, then $C_0 - C_1$ can not be expected to be strictly positive definite. However, we improve the chance that $C_0 - C_1 > 0$ if we choose the index of the peak closer to $N/2$.[7]

Last, we mention that the condition that H_0 be invertible has always been met easily in all the filter designs we have attempted using the Padé technique.[8]

V. Examples

Realistic examples have been chosen to illustrate the development in the previous sections. They constitute two filter designs that are typical of data transmission systems employing digital signal processing.

As the first example[9] consider the bandpass shape depicted in Fig. 1 and labeled by curve A. Suppose that design criteria for this spectrum specify 3-dB loss points at 200 Hz and 3200 Hz and no more than a 0.25-dB peak to peak ripple over the passband. Linear phase over the passband as well as slope loss greater than 12 dB/octave in the rejection bands are other requirements. The classical approach to designing a

[7]The intention of our work on the stability question has been to employ the full symmetry of $B = H_1 H_0^{-1}$. The classical approach to determining conditions on B so that $B^n \to 0$ has not been especially useful in our case in spite of this symmetry. (See [7] and [8] for a summary of results in this area.)

[8]We offer this statement in response to the warnings given by a theorem (commonly found in studies of Toeplitz forms) that says that the minimum eigenvalue of H_0, $|\lambda_{\min}(H_0)|$, monotonically approaches $\inf_{\omega} |H(\omega)|$ (as $N \to \infty$) which, of course, is zero for most filters of interest—low pass, bandpass; etc. Hence, H_0 would not be invertible if $|\lambda_{\min}(H_0)| = 0$. More importantly, machine computation of H_0 inverse becomes more difficult as N increases even though $|\lambda_{\min}(H_0)| \neq 0$. (See [9, p. 147] and [10, pp. 201-202] for a precise statement of this theorem.)

[9]In each of the examples the sampling rate is 7200 samples/s.

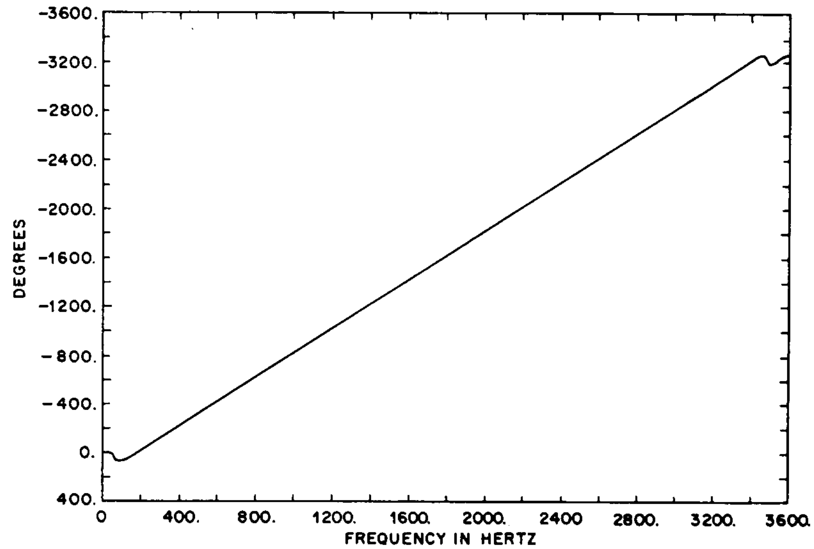

Fig. 2. Phase for bandpass filter (24th order).

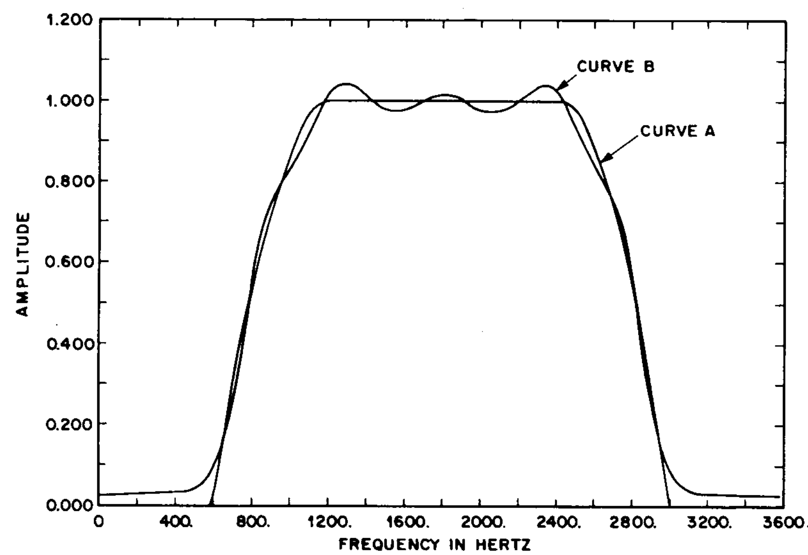

Fig. 3. Amplitude spectrum for cosine rolloff filter (10th order).

filter of this type would be to use several degrees of freedom to satisfy the amplitude spectrum requirements and then attempt phase equalization with the remaining degrees of freedom. Curve *B* represents the amplitude spectrum of a Padé realization of the desired spectrum in 12 second-order sections (24th order filter). The resultant phase is shown in Fig. 2. The largest absolute error for the time samples in this realization is 0.0018 where $\max |h_n| = 1$.

Similarly, a tenth ($N = 10$) order recursive digital filter has been synthesized to match the spectral shape of the bandpass function with cosine rolloff labeled *A* in Fig. 3. (This receiving filter shape can be used for baseband data transmission on channels with no frequency offset.) A time sequence comparison reveals $\max_n |h_n - g_n| = 0.024$ where $\max_n |h_n| = 1$. Its approximately linear phase is shown in Fig. 4. This example provides us with an opportunity to discuss the stopband behavior of Padé-realized filters. At times a minimum loss of 40, 50, or even 80 dB is specified for a filter with a stopband. This type of requirement cannot be easily worked into the time domain terms of the Padé synthesis procedure. Hence, only iterative routines, working in the frequency domain, from the Padé initial α and β values, could work these requirements into the final design. From our experience in obtaining initial designs by the Padé

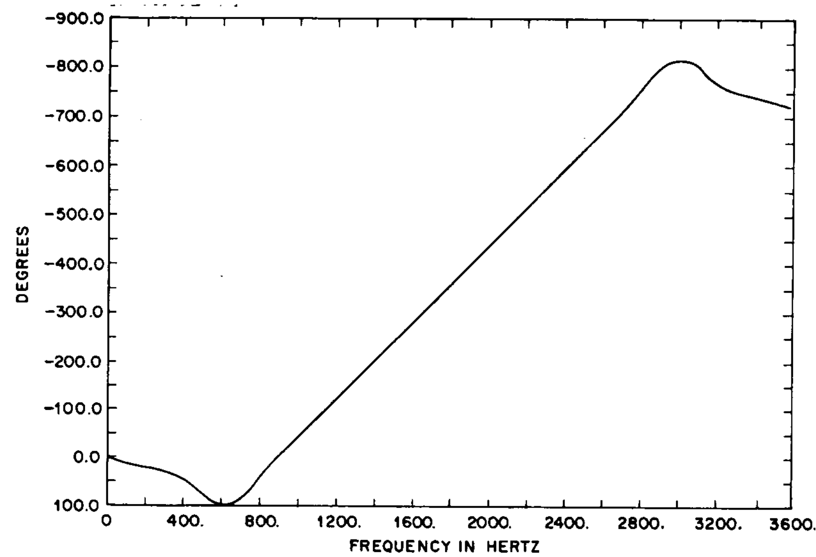

Fig. 4. Phase for cosine rolloff filter (10th order).

technique we have observed that the stopband loss ranges from 20 to 40 dB, which in many cases may be satisfactory.

References

[1] L. Weiss and R. N. McDonough, "Prony's method, z-transforms, and Padé approximation," *SIAM Rev.*, vol. 5, no. 2, p. 145, 1963.
[2] C. S. Burrus and T. W. Parks, "Time domain design of recursive digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 137-141, June 1970.
[3] K. L. Su, *Time Domain Synthesis of Linear Networks.* Englewood Cliffs, N.J.: Prentice-Hall, 1971, ch. 2.
[4] W. F. Trench, "An algorithm for the inversion of finite toeplitz matrices," *J. Appl. Math.*, vol. 12, pp. 515-522, Sept. 1964.
[5] E. F. Beckenbach and R. Bellman, *Inequalities.* New York: Springer Verlag, 1965.
[6] J. L. Shanks, "Recursion filters for digital processing," *Geophysics*, vol. 32, pp. 33-51, Feb. 1967.
[7] P. Stein, "Some general theorems on iterants," *J. Res. Nat. Bur. Stand.*, vol. 48, pp. 82-83, Jan. 1952.
[8] O. Taussky (Todd), "Matrices C with $C^n \rightarrow 0$," *J. Algebra*, vol. 1, pp. 5-10, 1964.
[9] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications.* Berkeley, Calif: Univ. California Press, 1958.
[10] H. Widom, "Toeplitz matrices," in *Studies in Mathematics*, vol. 3, I. I. Hirschman, Jr., Ed. Math. Assoc. of Amer. Englewood Cliffs, N.J.: Prentice-Hall, 1965, pp. 179-209.