R.J. Marks II, "Noise sensitivity of bandlimited signal derivative interpolation", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-31, pp.1029-1032 (1983).

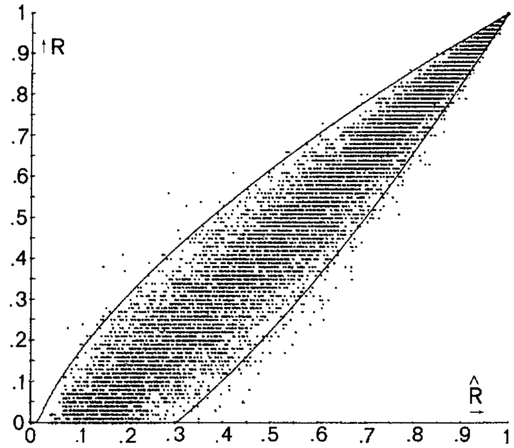

Fig. 1. Estimates $\hat{R}$ with confidence intervals for relative linear power contribution R from 100 simulations each of 100 values of R.

Equation (34) was simulated for $R = 0, 0.01, \cdots, 1$. For each value of R, 100 independent realizations of u and w of 64 samples during $T = 10$ s each were used. From each resulting pair of realizations of u and y, $\hat{R}$ was estimated using a Papoulis correlation window [4] with maximal window lag of 15 samples, yielding a window bandwidth [1], [4] $W = 1.896 \times 64/15T = 0.80$ Hz for $F_1 = 0.1$ and $F_2 = 3.2$; therefore $F = 3.1$ Hz.

Fig. 1 shows $\hat{R}(R)$ with 95 percent confidence interval approximations calculated according to (30)-(33). For the 10 000 values of $\hat{R}$ obtained, 34 of the actual R were over and 152 under the calculated 95 percent confidence limit approximations.

IV. Conclusions

Equation (21) shows that the variance of relative linear power contribution estimate $\hat{R}$ is lower than that of squared coherence function estimate $\hat{C}(f)$. This is accomplished at the cost of a larger bias in $\hat{R}$ compared to $\hat{C}(f)$, according to (28). Such bias might be corrected for.

Fig. 1 shows that for white signals, confidence intervals derived according to (30)-(33) are reasonable bounds on the simulated R for resulting estimate $\hat{R}$. The number of occurrences of R under the lower and over the upper 95 percent confidence limit did not reach the mathematical expectation 250 each, however. These deviations may be due to approximation errors like no linearity of $\hat{Z}$ required for (29) over the wide range of $\hat{R}$ for 64 samples or nonnormality of $\hat{Z}$. Similar analysis could be carried out for arbitrary spectra by using cocoloration factors according to (19).

Acknowledgment

The author is indebted to Dr. L.L.W. Hoogstraten for carrying out the simulations, to Prof. H. Kwakernaak for critically reading, and to G.G.M. Steijlen for typing the manuscript.

References

[1] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications.* San Francisco, CA: Holden-Day, 1968.

[2] P. Löhnberg, "Improved approximation of bias in squared coherence estimates for weakly smooth spectra," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 172-174, Apr. 1978.

[3] A. Hald, *Statistical Theory with Engineering Applications.* New York: Wiley, 1952.

[4] A. Papoulis, "Minimum-bias windows for high-resolution spectral estimates," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 9-12, Jan. 1973.

Noise Sensitivity of Band-Limited Signal Derivative Interpolation

ROBERT J. MARKS II

Abstract—**The sensitivity of interpolation of the *p*th derivative of a band-limited signal directly from the signal's samples in the presence of additive stationary noise is considered. Oversampling and filtering generally decrease the interpolation noise level when the data noise is not band-limited. A lower bound on the interpolation noise level can be approached arbitrarily closely by increasing the sampling rate. The lower bound is equivalent to the noise level obtained by low-pass filtering and *p*th-order differentiation of the unsampled additive input noise.**

Introduction

Given the sufficiently closely spaced samples of a band-limited signal, we can directly generate the pth derivative of the signal through appropriate interpolation functions [1].[1] Digital filters can generate good approximations of the samples of the pth derivative, given the signal samples as inputs [2]-[4]. The effects of filter design have been considered under the assumption of noiseless data [5]. Similarly, digital filters for sample interpolation ($p = 0$) have also been considered [6]-[7].

In this paper, an ideal pth-order differentiator is assumed and its operation in the presence of additive input data is investigated. We demonstate that the noise level, in general, can be reduced by increasing the sampling rate. The reduction, however, is sometimes insignificant. A lower bound for the noise level is shown to be that resulting from passing the unsampled input noise through a cascaded low-pass filter and pth-order differentiator.

In the next section, preliminary concepts are introduced. General formulas for the interpolation noise level are then derived, followed by establishment of a corresponding lower bound. In the final two sections, the specific cases of Laplace and triangular autocorrelations are considered. When appropriately parameterized, both degenerate to the special case of white noise samples.

Preliminaries

Let $\mathcal{B}_W$ denote the class of L_2 band-limited signals with bandwidth $2W$. That is, if $x(t) \in \mathcal{B}_W$, then

$$x(t) = \int_{-W}^{W} X(f) \exp(j2\pi ft)\, df$$

where

$$X(f) = \mathcal{F}x(t) = \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft)\, dt$$

and $\mathcal{F}$ denotes the Fourier transform operator. Then the pth derivative of $x(t)$ is

$$x^{(p)}(t) = (2B)^p \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2B}\right) d_p[2Bt - n] \qquad (1)$$

where the sampling rate $2B$ exceeds $2W$,

Manuscript received September 13, 1982; revised January 25, 1983.

The author is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195.

[1] The interested reader is referred to the introduction of [1] for a discussion of additional motivation for this work.

0096-3518/83/0800-1028$01.00

$$d_p(t) = \left(\frac{d}{dt}\right)^p \operatorname{sinc} t$$

and $\operatorname{sinc} t = \sin(\pi t)/(\pi t)$. For $p = 0$, (1) becomes the conventional Shannon sampling theorem [8]. Define the sampling rate parameter

$$r \equiv \frac{2W}{2B} \leqslant 1.$$

Since $x^{(p)}(t) \in \mathcal{B}_W$, it is unaltered by low-pass filtering. Passing (1) through a filter unity on $|f| \leqslant W$ and zero elsewhere gives

$$x^{(p)}(t) = r(2W)^p \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2B}\right) d_p[2Wt - rn]. \tag{2}$$

Note that (1) can be considered as a special case for $r = 1$.

Effects of Additive Noise

Let $\xi(t)$ denote a real zero-mean wide-sense stationary random process with autocorrelation

$$R_\xi(t-\tau) = E[\xi(t)\,\xi(\tau)]$$

where E denotes the expectation operator. If the noise samples $\xi(n/2B)$ are added to the signal samples in (2), the result is $x^{(p)}(t) + \eta_p(t)$ where

$$\eta_p(t) = r(2W)^p \sum_{n=-\infty}^{\infty} \xi\left(\frac{n}{2B}\right) d_p[2Wt - rn].$$

Thus

$$R_{\eta_p}(t-\tau) = E[\eta_p(t)\,\eta_p(\tau)]$$
$$= r^2(2W)^{2p} \sum_{m=-\infty}^{\infty} R_\xi\left(\frac{m}{2B}\right) \cdot \sum_{n=-\infty}^{\infty} d_p[2W\tau - (n-m)r]\, d_p[2Wt - rn].$$

Since $x(t) = d_p[2W(\tau - t) + mr] \in \mathcal{B}_W$, we can use (2) to evaluate the n sum above. Furthermore, since

$$d_p(t) = (-1)^p d_p(-t)$$

and

$$\left(\frac{d}{dt}\right)^p d_p(t) = d_{2p}(t)$$

it follows that

$$R_{\eta_p}(\tau) = (-1)^p r(2W)^{2p} \sum_{m=-\infty}^{\infty} R_\xi\left(\frac{m}{2B}\right) d_{2p}(2W\tau - rm). \tag{3}$$

Since $\eta_p(t)$ is zero mean, the corresponding interpolation noise level is

$$\overline{\eta_p^2} = R_{\eta_p}(0)$$
$$= (-1)^p r(2W)^{2p} \sum_{m=-\infty}^{\infty} R_\xi\left(\frac{m}{2B}\right) d_{2p}(rm). \tag{4}$$

A spectral density description of the process can be obtained by first transforming (3)

$$S_{\eta_p}(f) = \mathcal{F} R_{\eta_p}(t)$$
$$= \frac{1}{2B}(2\pi f)^{2p} \sum_{m=-\infty}^{\infty} R_\xi\left(\frac{m}{2B}\right) e^{-j\pi m f/B} \operatorname{rect}\left(\frac{f}{2W}\right)$$

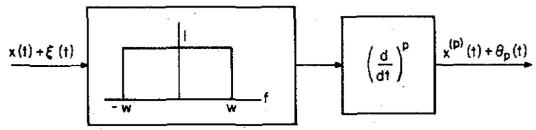

Fig. 1. Cascaded low-pass filter and pth-order differentiator. The output noise level $\overline{\theta_p^2}$ is a lower bound for pth-order derivative interpolation from the input samples.

where $\operatorname{rect}(y)$ is unity for $|y| \leqslant \frac{1}{2}$, and zero elsewhere. Application of the Poisson sum formula to the m sum gives

$$S_{\eta_p}(f) = (2\pi f)^{2p} \sum_{n=-\infty}^{\infty} S_\xi(f - 2nB) \operatorname{rect}\left(\frac{f}{2W}\right)$$

where $S_\xi(f) = \mathcal{F} R_\xi(t)$ is the input noise power spectral density. An alternate expression for the output noise level follows as

$$\overline{\eta_p^2} = \int_{-\infty}^{\infty} S_{\eta_p}(f)\, df$$
$$= (2\pi)^{2p} \sum_{n=-\infty}^{\infty} \int_{-W}^{W} f^{2p} S_\xi(f - 2nB)\, df. \tag{5}$$

For the unfiltered case ($r = 1$), the integration interval in (5) is over $|f| \leqslant B$. Since $S_\xi \geqslant 0$, filtering always results in a noise level equal to or better than the unfiltered case.

A Lower Bound on Output Noise Level

Consider $\xi(t)$ input into the cascaded low-pass filter and pth-order differentiator in Fig. 1. Let $\theta(t)$ denote the output. Recall that, in general, the output spectral density $S_0(f)$ due to a spectral density input $S_i(f)$ into a system with transfer function $H(f)$ is $S_0(f) = |H(f)|^2 S_i(f)$ [9]. Thus

$$S_\theta(f) = (2\pi f)^{2p} S_\xi(f) \operatorname{rect}\left(\frac{f}{2W}\right)$$

and

$$\overline{\theta_p^2} = (2\pi)^{2p} \int_{-W}^{W} f^{2p} S_\xi(f)\, df. \tag{6}$$

Compare to (5). Since $S_\xi \geqslant 0$, it follows that $\overline{\theta_p^2}$ is a lower bound for the output noise level

$$\overline{\eta_p^2} \geqslant \overline{\theta_p^2}.$$

Equality is achieved when $\xi(t)$ has band-limited spectral density (say over the interval $|f| \leqslant \Omega$) and the sampling rate is sufficiently high to avoid aliasing (i.e., $2B - \Omega > W$). A lower sampling rate would result in aliasing and a higher output noise level.

For finite $\overline{\xi^2}$, $S_\xi(f) \to 0$ as $|f| \to \infty$. We see from (5) that, as $2B \to \infty$, $\overline{\eta_p^2} \to \overline{\theta_p^2}$. Hence, the bound can be approached arbitrarily close by appropriate increase in sampling rate. Note we can guarantee from (5) that $\overline{\eta_p^2}$ decreases with r if $S_\xi(u)$ decreases with $u > 0$. This spectral density property, however, is not applicable for the triangular autocorrelation.

Triangular Autocorrelation

Consider the triangle autocorrelation parameterized by $a > 0$.

$$R_\xi(\tau) = \overline{\xi^2}\, \Lambda(\tau/a) \tag{7}$$

where

$$\Lambda(t) = (1 - |t|) \operatorname{rect}(t/2).$$

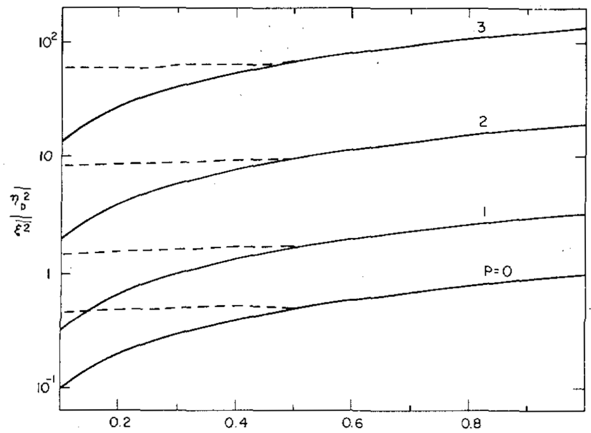

Fig. 2. Normalized interpolation noise level $\overline{\eta_p^2}/\overline{\xi^2}$ for the triangle autocorrelation with $2W = 1$. The solid curve is for $a = 0.1$ and the dashed $a = 0.5$. The curves are identical for $r > 1/2$ where the noise samples are white. For $a = 0.1$, the noise samples are white for $r \geqslant 0.1$.

Substituting into (4) gives the normalized error.

$$\overline{\eta_p^2}/\overline{\xi^2} = (-1)^p\, r(2W)^{2p}\left[d_{2p}(0) + 2\sum_{m=1}^{N}\left(1 - \frac{m}{T}\right) d_{2p}(rm)\right] \tag{8}$$

where

$$T \equiv 2Ba$$

and N is the greatest integer not exceeding T. Plots of (8) are shown in Fig. 2 for $2W = 1$ and $a = 0.5$ and 0.1.

Of specific interest is the case where sampling is performed such that $T \leqslant 1$. The noise samples are then white. That is,

$$R_\xi\left(\frac{n}{2B}\right) = \overline{\xi^2}\,\delta_n \tag{9}$$

where δ_n denotes the Kronecker delta. Since

$$d_{2p}(0) = \int_{-1/2}^{1/2} (j2\pi f)^{2p}\, df = \frac{(-1)^p \pi^{2p}}{2p + 1} \tag{10}$$

we have for white samples

$$\overline{\eta_p^2}/\overline{\xi^2} = \frac{r(2\pi W)^{2p}}{2p + 1}. \tag{11}$$

For the conventional sampling theorem, $p = 0$ and the noise level is improved by a factor of r.

Since $2W = 1$ in Fig. 2, the plots there are equivalent to (11) for $r \geqslant a$. Note (11) is independent of the a parameter—thus the merging of the $a = 0.1$ and 0.5 plots at $r = 0.5$. For the domain shown, all of the $a = 0.1$ samples are white.

Note there can exist a point whereupon a further increase of sampling rate results in an insignificant improvement in the interpolation noise level.

Since $\mathfrak{F}\Lambda(t) = \text{sinc}^2 f$, from (6), the normalized lower bound for the triangle autocorrelation is

$$\overline{\theta_p^2}/\overline{\xi^2} = 2a(2\pi)^{2p} \int_0^W f^{2p}\, \text{sinc}^2\,(af)\, df. \tag{12}$$

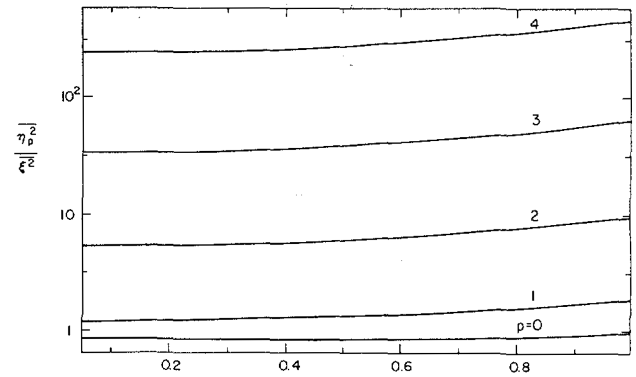

Fig. 3. Normalized interpolation noise level $\overline{\eta_p^2}/\overline{\xi^2}$ for Laplace autocorrelation with $\alpha = 2W = 1$.

In Appendix A, we show that

$$\overline{\theta_p^2}/\overline{\xi^2} = \begin{cases} \dfrac{2}{\pi}\{Si(2\pi aW) - \sin(\pi aW)\,\text{sinc}\,(aW)\}\,; & p = 0 \\ \dfrac{4^p W^{2p-1}}{a}\left\{\dfrac{\pi^{2p-2}}{2p-1} + (-1)^p d_{2p-2}(2aW)\right\}; & p \geqslant 1 \end{cases} \tag{13}$$

where

$$Si(t) = \int_0^t \frac{\sin\tau}{\tau}\, d\tau$$

is the sine integral function. Lower bounds for each of the plots in Fig. 2 are graphically indistinguishable from the $r = 0.1$ values.

Laplace Autocorrelation

A second tractable solution results from the Laplace correlation parameterized by α.

$$R_\xi(\tau) = \overline{\xi^2} e^{-\alpha|\tau|}. \tag{14}$$

As is shown in Appendix B, the corresponding normalized output noise level can be written in integral form as

$$\overline{\eta_p^2}/\overline{\xi^2} = (2\pi B)^{2p} \sinh\frac{\alpha}{2B} \int_0^r u^{2p}\left[\cosh\frac{\alpha}{2B} - \cos\pi u\right]^{-1} du. \tag{15}$$

The well-behaved (strictly increasing) integrand in (15) provides for straightforward digital integration. Sample plots of (15) are shown in Fig. 3 for $2W = 1$.

Two special cases of (15) are worthy of note. a) The $p = 0$ case as plotted in Fig. 2 simply corresponds to conventional sampling theorem interpolation followed by filtering. For this case, (15) can be evaluated in closed form [10].

$$\overline{\eta_0^2}/\overline{\xi^2} = \frac{2}{\pi}\,\text{atan}\left[\frac{\sinh\dfrac{\alpha}{2B}\tan\dfrac{\pi r}{2}}{\cosh\dfrac{\alpha}{2B} - 1}\right].$$

b) If

$$\frac{\alpha}{2B} \gg 1 \tag{16}$$

then (15) approaches

$$\overline{\eta_p^2}/\overline{\xi^2} = (2\pi B)^{2p} \int_0^r u^{2p}\, du = \frac{r(2\pi W)^{2p}}{2p+1} \tag{17}$$

which is the same result as the discrete white noise case in (11). Indeed, when (16) is applicable, the noise samples are very nearly white and the plots for white samples in Fig. 2 can be used as excellent approximations.

Here, as in the previous example, one must be cautioned on comparing equally parameterized interpolation noise levels for differing p. The units of $\overline{\eta_p^2}/\overline{\xi^2}$ are $(s)^{-2p}$.

For a lower bound for the Laplace autocorrelation, we transform (14) and substitute into (6). The result is

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{4(2\pi)^{2p}}{\alpha} \int_0^W \frac{f^{2p}\, df}{1 + (2\pi f/\alpha)^2}. \tag{18}$$

As is shown in Appendix C

$$\overline{\theta_p^2}/\overline{\xi^2} = \begin{cases} \frac{2}{\pi} \operatorname{atan} \epsilon; & p = 0 \\ \frac{2\alpha^2}{\pi} [\epsilon - \operatorname{atan} \epsilon]; & p = 1 \\ \frac{2\alpha(2\pi W)^{2p-1}}{\pi(2p-1)} - \alpha^2 [\overline{\theta_p^2}/\overline{\xi^2}]; & p > 1 \end{cases} \tag{19}$$

where

$$\epsilon = 2\pi W/\alpha.$$

Graphically, these bounds are also indistinguishable from the corresponding smallest values on the plots in Fig. 3.

Appendix A
Derivation of (13) from (12)

We begin by rewriting (12) as

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{(2\pi)^{2p}}{\pi^2 a} \int_0^W f^{2p-2}\, [1 - \cos 2\pi a f]\, df. \tag{A1}$$

For even index

$$d_{2q}(t) = \int_{-1/2}^{1/2} (j2\pi\chi)^{2q} e^{-j2\pi\chi t}\, d\chi = 2(2\pi)^{2q} (-1)^q \int_0^{1/2} \chi^{2q} \cos(2\pi\chi t)\, d\chi.$$

For $f = 2W\chi$, it follows from (A1) that for $p \geqslant 1$

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{2(2W)^{2p}(-1)^{p-1}}{a} [d_{2p-2}(0) - d_{2p-2}(2aW)].$$

Using (10) gives (13) for $p \geqslant 1$. The $p = 0$ case follows immediately from (12) using integration by parts.

Appendix B
Derivation of (15) from (14) and (4)

Note that we can write (4) as

$$\overline{\eta_p^2} = \frac{1}{2B} (2\pi)^{2p} \int_{-W}^{W} f^{2p} \sum_{n=-\infty}^{\infty} R_\xi\left(\frac{n}{2B}\right) e^{-j\pi n f/B}\, df. \tag{A2}$$

[An alternate derivation follows by application of the Poisson sum formula to (5).] For the Laplace autocorrelation in (14), the n sum becomes

$$\frac{1}{\overline{\xi^2}} \sum_{n=-\infty}^{\infty} R_\xi\left(\frac{n}{2B}\right) e^{-j\pi n f/B} = 1 + 2 \sum_{m=1}^{\infty} e^{-m\alpha/2B} \cos \pi m f/B = 1 + 2 \operatorname{Re} \sum_{m=1}^{\infty} [e^{-(\alpha + j2\pi f)/2B}]^m = \frac{\sinh \frac{\alpha}{2B}}{\cosh \frac{\alpha}{2B} + \cos \frac{\pi f}{B}} \tag{A3}$$

where Re denotes the "real part of" and we have used the geometric series

$$\sum_{m=0}^{\infty} Z^m = (1 - Z)^{-1}; \quad |Z| < 1. \tag{A4}$$

Substituting (A3) into (A2), recognizing the integrand is even and making the variable substitution $f = Bu$ results in (15).

Appendix C
Derivation of (19) from (18)

Set $u = 2\pi f/\alpha$ and (18) becomes

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{2}{\pi} \alpha^{2p} \int_0^\epsilon \frac{u^{2p}\, du}{1 + u^2}. \tag{A5}$$

The $p = 0$ case follows immediately. For $p > 0$, consider first the case where $\epsilon < 1$. With $Z = -u^2$, the denominator in (A5) can be expanded via (A4) and the resulting integral evaluated to give

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{2}{\pi} \alpha^{2p} \sum_{m=0}^{\infty} (-1)^m \frac{\epsilon^{2m+2p+1}}{2m + 2p + 1}.$$

Set $n = m + p$ and recall the Taylor series

$$\operatorname{atan} x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}. \tag{A6}$$

Thus, for $\epsilon < 1$,

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{2}{\pi} (-1)^p \alpha^{2p} \left[\operatorname{atan} \epsilon - \sum_{n=0}^{p-1} (-1)^n \frac{\epsilon^{2n+1}}{2n+1}\right]. \tag{A7}$$

For $\epsilon > 1$ we rewrite (A5) as

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{2}{\pi} \alpha^{2p} \left[\int_0^{1-} + \int_{1+}^{\epsilon}\right] \frac{u^{2p}\, du}{1 + u^2} = \frac{2}{\pi} \alpha^{2p} \left[(-1)^p \left\{\frac{\pi}{4} - \sum_{n=0}^{p-1} \frac{(-1)^n}{2n+1}\right\} + \int_{1+}^{\epsilon} \frac{u^{2p}\, du}{1 + u^2}\right].$$

From (A4), it follows that

$$\sum_{n=0}^{\infty} Z^{-n} = Z(Z - 1)^{-1}; \quad |Z| > 1.$$

Again, with $Z = -u^2$, we obtain

$$\overline{\theta_p^2}/\overline{\xi^2} = \frac{2}{\pi}\alpha^{2p}\left[(-1)^p\left\{\frac{\pi}{4} - \sum_{n=0}^{p-1}\frac{(-1)^n}{2n+1}\right\} + \sum_{m=0}^{\infty}(-1)^m\frac{\epsilon^{2p-2m-1}-1}{2p-2m-1}\right].$$

Set $n = m - p$ in the m sum and use (A6). Recognizing $\pi/2 - \text{atan}\ 1/\epsilon = \text{atan}\ \epsilon$ for $\epsilon > 0$ again yields (A7). The recursive formula for $p \geqslant 1$ in (19) follows easily from (A7).

ACKNOWLEDGMENT

The author expresses his appreciation to D. Radbel for many insightful discussions.

REFERENCES

[1] R. J. Marks II and M. W. Hall, "Differintegral interpolation from a bandlimited signal's samples," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 872-877, Aug. 1981.
[2] A. Papoulis, *Circuits and Systems: A Modern Approach.* New York: Holt, Rinehart and Winston, 1980, p. 222.
[3] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1975, p. 206.
[4] R. W. Hamming, *Numerical Methods for Scientists and Engineers*, 2nd ed. New York: McGraw-Hill, 1973, p. 599.
[5] J. Spriet and J. Bens, "Optimal design and comparison of wide-band on-line differentiators," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 46-52, Feb. 1979.
[6] R. W. Schafer and L. R. Rabiner, "A digital processing approach to interpolation," *Proc. IEEE*, vol. 61, pp. 692-702, 1973.
[7] E. B. Hogenauer, "An economical class of digital filters for decimation and interpolation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 155-162, Apr. 1981.
[8] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, pp. 1565-1596, 1977.
[9] A. Papoulis, *Probability, Random Variables and Stochastic Processes.* New York: McGraw-Hill, 1965, p. 317.
[10] I. S. Gradsteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, 4th ed. New York: Academic, 1965, p. 148, entry 2.553.3.

A New Design Method of Optimal Finite Wordlength Linear Phase FIR Digital Filters

J. P. MARQUES DE SÁ

Abstract—**The branch and bound technique is applied to the design of optimal finite wordlength linear phase FIR digital filters in a new efficient way. Compared to other reported methods there is a great saving of storage area used by the designing program; reductions in computation time are also to be expected.**

I. INTRODUCTION

The design of optimal finite wordlength linear phase FIR digital filters has been discussed by several authors. References [1] and [2] are recent reports of design methods containing detailed information on the subject.

The design problem can be formulated in the following way. Given a set $\{(w_k, F(w_k));\ w_k \in [0, \pi],\ F(w_k) \in \mathbb{R},\ k = 1, 2, \cdots, M\}$ and an N-length linear phase FIR filter characteristic function $H(w) = h_0 + \sum_{i=1}^{(N-1)/2} h_i \cos(iw)$—for simplicity symmetrical odd length only is considered—find the $N_c = (N+1)/2$ coefficients h_i that minimize

$$\|E(w_k)\| \equiv \max_{w_k} W(w_k)\cdot|F(w_k) - H(w_k)|$$

with h_i, $i = 0, 1, \cdots, N_c - 1$: D-digit numbers of base b. $E(w_k)$ are the deviations of $H(w_k)$ from the desired function values $F(w_k)$ weighted by $W(w_k)$.

The interest of optimal finite wordlength solutions for FIR filters is easily understood in connection with either hardware or software implementations. In the first case, finite wordlength is mandatory; in the second case, it is required for speeding up filtering operations (floating point substituted by fixed point operations). In both cases, it can be of interest to use number bases other than base two.

Coefficient quantization formulas have been found [3] which can guide the choice of a rounded solution wordlength for a given degradation of the infinite precision (continuous) solution. However, there is abundant evidence [1], [2] that the performance of rounded solutions can be quite inferior to the optimal one.

Substituting the expressions of the deviations by corresponding inequalities in the formulation of the problem given above, the problem clearly becomes a particular case of the mixed integer programming problem.

The most successful methods the author is aware of to solve this particular problem are the branch and bound algorithm for mixed integer programming (Land and Doig algorithm) and integer zero-one programming [1], [2]. These algorithms are available in commercial software packages and are described in several books [4], [5].

Two difficult aspects of the application of these methods follow.

1) Computation Time: The design can take a huge amount of computer time even when the coefficient values are restricted to a limited interval centered on the rounded values.

Reference [1] indicates that for a 33-point filter it took 5.3 times longer to find the best rounded solution than to find the continuous one (best rounding involves a $\pm b^{-1}$ neighborhood of the rounded values). The optimal solution for the same filter took about 439 times longer to find, and it took 870 times longer to prove its optimality.

2) Storage Area: As mentioned above, the suitable reformulation of the problem is

$$\text{minimize } E,$$

$$\text{with}\begin{cases} H(w_k) - E/W(w_k) \leqslant F(w_k) \\ -H(w_k) - E/W(w_k) \leqslant -F(w_k) \end{cases} \quad k = 1, 2, \cdots, M.$$

The algorithm therefore has to operate over an initial large and dense matrix of $2M(N_c + 1)$ elements. For a not too small problem, this represents a rather big amount of storage area. For instance, with a length $N = 33$ and a grid density $N_g = 10$ there is a maximum of $2 \times 170 \times 18 = 6120$ real numbers. This imposes a severe limitation on the use of such design methods in small systems, namely microcomputer systems.

II. PROPOSED METHOD

The proposed method is based on the following considerations.

1) The present mixed integer programming problem is, in fact, originally an approximation problem in the Chebyshev sense. Therefore, it should be advantageous to use the special case of linear programming algorithm suited for the Chebyshev approximation. This algorithm is characterized by a considerable saving of storage area since there is no need to store the whole matrix at all. Matrix row values are functionally defin-

Manuscript received February 8, 1982; revised October 11, 1982.
The author is with the Faculdade de Engenharia, 4099 Porto Codex, Portugal.

0096-3518/83/0800-1032$01.00