THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 9, November 1977
Printed in U.S.A.

Optimum Digital Filters for Interpolative A/D Converters

By A. N. NETRAVALI

(Manuscript received March 31, 1977)

Interpolative analog-to-digital (A/D) converters allow a fine representation of signals by making many coarse representations and averaging them using a digital filter. In this paper, we give a method of optimizing the characteristics of this digital filter under two different criteria. The first criterion is the well-known signal-to-noise (S/N) ratio, whereas the second criterion is the weighted sum of the signal power, the quantization noise power, and the noise power within a given band of frequencies. We design optimum digital filters and simulate their performance on the computer. We show that the theoretically predicted S/N ratio is in good agreement with the performance obtained by computer simulation. It is seen that about 23 dB improvement in S/N ratio over the S/N ratio attainable by a constant-weight digital filter is possible when the number of coarse quantizations is 256. We also study the effects of changing various parameters of the A/D converter on the S/N ratio.

I. INTRODUCTION

Interpolative A/D converters[1–3] achieve a fine quantization of signals by making several coarse quantizations and averaging them. This requires high-speed operation of that part of the A/D converter which obtains the coarse quantizations. Higher and higher speeds are required for finer and finer ultimate quantization. This trade-off between the speed of operation and amplitude resolution is particularly relevant and important with present-day integrated circuit technology, which provides high-speed operation but no high-amplitude precision.

Several well known methods of obtaining the many coarse quantizations exist. Goodman,[1] and Goodman and Greenstein,[2] have considered the ordinary delta modulator which gives a two-level representation of the signal at a rate many times higher than the Nyquist rate. The output of the delta modulator is filtered by a digital filter and resampled at

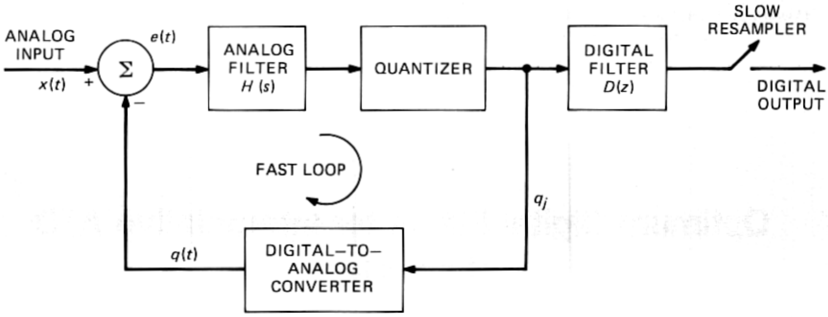

Fig. 1—An interpolative A/D converter.

Nyquist rate to obtain the PCM output. The performance of such an A/D converter depends upon the speed of the delta modulator and the characteristics of the digital filter.

Another method of obtaining the coarse quantization has been proposed recently by Candy.[3] In this method, the coarse quantizations are obtained by a direct feedback encoder shown in Fig. 1. In this encoder a difference between the analog input and a coarsely quantized representation is filtered by an analog filter with characteristics $H(s)$, and then quantized. This is done at speeds higher than the Nyquist rate. The output of the quantizer is represented by binary words and filtered by a digital filter having characteristics $D(z)$. The output of the digital filter is resampled at a slower rate to obtain the final digital output at the Nyquist rate. Use of direct feedback encoding allows shaping of the quantization noise in such a way that the digital filter can be made very simple. Candy[3] has shown that when the analog filter is taken to be a pure integrator the simple digital filter corresponding to "accumulate-and-dump" performs adequately.

Candy et al.[4] have described a method of optimizing the weights of the digital filter when the analog filter in the "fast loop" is a pure integrator. They have shown that the optimum weights can be approximated by a set of triangularly distributed weights and evaluated the improvements in S/N ratio by using these weights. Their results are applicable only when the integrator in the "fast loop" is not reset to zero at the beginning of each slow cycle. In this paper, we first show that when the analog filter in the "fast loop" is reset to zero at the beginning of every slow cycle an advantage in the S/N ratio is obtained when the uniformly distributed weights are used. We then give a different method of optimizing the digital filter characteristics under the assumption that the analog filter is reset. Our method of optimization is applicable to the case of any arbitrary analog filter in the place of the integrator in the fast loop. The resulting optimum weights when the integrator is reset every slow cycle have a different shape than the optimum weights given by Candy[4] which are applicable when the integrator is not reset. We compare our optimum weights with the triangular weights proposed by Candy as an approximation to his optimum weights. Optimization of the digital filters using a different criterion, which includes a deviation of the digital filter characteristics from desired characteristics is also discussed. In this case it is possible, for example, to shape the discrete Fourier transform of the digital filter weights so that it resembles, as far as possible, an ideal low-pass filter. We evaluate the performance of the A/D converter in terms of S/N ratio by computer simulation for several typical cases.

II. SUMMARY OF RESULTS

Our computer simulations indicate that there is about 3 dB improvement in S/N ratio by resetting the integrator at the beginning of each slow cycle when uniform weights are used for the digital filters. This improvement is independent of the coarseness of the quantizer in the fast loop, the number of fast cycles and the correlations present in the input signal. The use of optimum weights for the digital filter leads to significant improvements in S/N ratio over that obtained by a digital filter with constant weights. This improvement although independent of the coarseness of the quantizers depends on the number of fast cycles; for 32 fast cycles, there is about a 14 dB improvement, whereas for 256 fast cycles, there is a 23 dB improvement. Also, the optimum weights outperform the "triangular" weights used by Candy et al.[4] by about 7.30 dB when the number of fast cycles is 32 and by about 8.80 dB when the number of fast cycles is 256. We also show that there is a good agreement between the theoretically predicted S/N ratio and that obtained from computer simulations of the A/D converter. Changing the analog filter from an integrator to a general analog filter with a given characteristic indicates that there is a gain of a few dB in S/N ratio by choosing the dc gain and the cutoff frequencies judiciously. Our second method of optimizing the digital filter characteristic allows us to minimize the deviation of its frequency characteristics from a given characteristic. Using the desired characteristic to be ideal low-pass, we are able to decrease the noise power in a given band of frequencies. This decrease is about 0.5 to 1.0 dB, but it comes at the expense of an increase in the overall noise power of about 1.0 to 1.5 dB. Thus the digital filter suppresses the noise power in one band of frequencies, but enhances the noise in the rest of the frequency band, resulting consequently in an overall increase in the noise power.

III. DERIVATION OF OPTIMUM DIGITAL FILTER WEIGHTS

In this section, we derive the weights of the optimum digital filter. First we concern ourselves with those digital filters which minimize the S/N ratio, and then derive those weights which can be spectrally shaped.

Let $x(t)$ be the analog input to the A/D converter shown in Fig. 1. Also let $h(\cdot)$ be the impulse response of the time-invariant analog filter in the fast loop; N, the number of fast cycles; T, the fast sampling period; and q_j, the output of the quantizer at the jth fast cycle. We assume that the output of the digital-to-analog converter is given by

$$q(t) = q_j \quad jT \leq t < (j+1)T \tag{1}$$

The equation for the fast loop can be written as:

$$\int_0^t h(\tau)[x(t-\tau) - q(t-\tau)]d\tau = q(t+T) + n(t+T) \tag{2}$$

Here we have assumed that the analog filter is reset at the beginning of each slow cycle and that the quantization distortion can be represented by additive random noise $n(\cdot)$. Assuming that $x(t)$ is constant $(=x)$ over a slow cycle, then at $t = (i+1)T$,

$$x \int_0^{iT} h(t)dt - \int_0^{iT} h(\tau)q(iT-\tau)d\tau = q[(i+1)T] + n[(i+1)T] \tag{3}$$

Now letting

$$\int_{kT}^{(k+1)T} h(t)dt = h_k \tag{4}$$

eq. (3) can be written as:

$$x \sum_{k=0}^{i-1} h_k - \sum_{k=0}^{i-1} q_{i-k-1}h_k = q_i + n_i \qquad i = 1, \ldots, N \tag{5}$$

where

$$n_i = n(iT)$$

Equation (5) can be written for $i = 1, \ldots, N$, in a matrix form

$$x \cdot A = HQ + N_0 + q_0 \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{N-1} \end{bmatrix} \tag{6}$$

where

$$A = \text{col}\left(h_0, h_0 + h_1, \ldots, \sum_{i=0}^{N-1} h_i\right)$$

$$H = \begin{bmatrix} 1 & 0 & & 0 \\ h_0 & 1 & 0 & 0 \\ h_1 & h_0 & 1 & 0 \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ h_{N-2} & h_{N-3} \cdots & h_1 h_0 & 1 \end{bmatrix}$$

$$N_0 = \text{col}\,(n_1, n_2, \ldots, n_N)$$

$$Q = \text{col}\,(q_1, q_2, \ldots, q_N)$$

Observe that matrix H has an inverse and therefore eq. (6) can be rewritten as:

$$Q = xH^{-1}A - H^{-1}N_0 - q_0 H^{-1} \begin{bmatrix} h_0 \\ \vdots \\ h_{N-1} \end{bmatrix} \tag{7}$$

The digital filter will process vector Q every slow cycle by multiplying it by a weight vector D, and thus the PCM output will be

$$D^TQ = xD^TH^{-1}A - D^TH^{-1}N_0 - q_0 D^TH^{-1} \begin{bmatrix} h_0 \\ \vdots \\ h_{N-1} \end{bmatrix} \tag{8}$$

Here we assume that $D^TU = 1$, where $U = \text{col}\,(1, 1, \ldots, 1)$. The first term on the right-hand side of eq. (8) is the signal component, whereas the second term is the noise component. The third term results from the initial condition on the D/A, q_0. We assume $q_0 = 0$. In order to maximize the ratio of signal energy to the noise energy, we maximize the following expression:

$$(S/N) \triangleq (D^TG^{-1}A)^2 / \overline{(D^TH^{-1}N_0)^2} \tag{9}$$

where $\overline{(\cdot)}$ denotes expectation. Assuming† that the noise components n_i are independent, identically distributed with variance σ^2, we can write eq. (9) as:

$$(S/N) = \frac{1}{\sigma^2}\,[D^TH^{-1}AA^T(H^{-1})^TD] / D^TH^{-1}(H^{-1})^TD \tag{10}$$

We note that since H has an inverse H^{-1}, $(H^{-1})^T$ is positive definite and therefore the denominator of the right hand side of eq. (10) will not be zero unless $D \equiv 0$, a case which we rule out. This implies that (S/N) will

† This assumption is not required. It is easy to extend the following analysis to colored noise.

be bounded from above. Since (S/N) is a ratio of two quadratic forms generated by two symmetric matrices, we can write the optimum D, D^*, as a solution of the eigenvalue problem

$$(H^{-1}A)(H^{-1}A)^T D^* = \lambda_{\max}(H^{-1})(H^{-1})^T D^* \qquad (11)$$

or

$$H^T A A^T (H^{-1})^T D^* = \lambda_{\max} D^* \qquad (12)$$

where $\lambda_{\max}$ is the maximum eigenvalue. It is easy to see that the only eigenvector for eq. (12) corresponding to a nonzero eigenvalue is given by

$$D^* = H^T A \qquad (13)$$

for which

$$\begin{aligned}(S/N) &= \frac{A^T A}{\sigma^2} \\ &= \frac{1}{\sigma^2}\left[\sum_{i=1}^{N}\left(\sum_{k=0}^{i-1} h_k\right)^2\right]\end{aligned} \qquad (14)$$

Writing out H and A, we get

$$D^* = \begin{bmatrix} \sum_{j=-1}^{N-2} h_j \left(\sum_{k=0}^{j+1} h_k\right) \\ \vdots \\ \sum_{j=-1}^{N-3} h_j \left(\sum_{k=0}^{j+2} h_k\right) \\ \vdots \\ \sum_{j=1}^{N-(N+1)} h_j \left(\sum_{k=0}^{j+N} h_k\right) \end{bmatrix} \qquad (15)$$

where we have assumed for notational convenience that $h_{-1} = 1$. If the filter in the fast loop is a pure integrator, then $h_k = T$, and the optimum digital filter can be written as:

$$D = \text{col}\,(D_1, \ldots, D_j, \ldots, D_N)$$

where

$$D_j = \frac{T(N - j + 1)(N + j)}{2} \qquad j = 1, \ldots, N \qquad (16a)$$

and for large N the S/N ratio is given (except for a proportionality constant) by

$$S/N = \frac{N(N + 1)(2N + 1)}{6} \qquad (16b)$$

3.1 *Optimum digital filter with spectral shaping*

Let $D(\omega)$ be the discrete Fourier transform of the samples $\{D_k\}_{k=0,\ldots,N-1}$ and $C(\omega)$ be the transform of the desired response that is obtained from the filter weights $\{C_k\}_{k=0,\ldots,N-1}$. The shaping of the digital filter in the Fourier domain can thus be accomplished by proper choice of $C(\omega)$. We use the following expression for the error between the two:

$$\begin{aligned} E_{RR} &= \int |[D(\omega) - C(\omega)]|^2 \, d\omega \\ &= \int [D(\omega) - C(\omega)][D(\omega) - C(\omega)]^* \, d\omega \\ &= \int D(\omega)D^*(\omega) \, d\omega - \int D(\omega)C^*(\omega) \, d\omega \\ &- \int D^*(\omega)C(\omega) \, d\omega + \int C(\omega)C^*(\omega) \, d\omega \end{aligned} \tag{17}$$

where $(\cdot)^*$ is the complex conjugate. In minimizing E_{RR} with respect to D, we can drop the third term of eq. (17) and rewrite (17) as

$$\begin{aligned} E_{RR} &= \sum_{k=0}^{N-1} D_k^2 - \int \left(\sum_{k=0}^{N-1} D_k e^{-j2\pi\omega k/N} \right) C^*(\omega) \, d\omega \\ &- \int \left[\sum_{k=0}^{N-1} C_k e^{-j2\pi\omega k/N} \right] D^*(\omega) \, d\omega \\ &= \sum_{k=0}^{N-1} D_k^2 - 2 \sum_{k=0}^{N-1} D_k C_k \\ &= D^T D - 2D^T C \end{aligned} \tag{18}$$

The performance function (PF) that we want to maximize can be written as:

$$(\mathrm{PF}) = D^T H^{-1} A (H^{-1} A)^T D - \lambda D^T H^{-1} (H^{-1})^T D - \gamma (D^T D - 2D^T C) \tag{19}$$

where the first term on the right-hand side corresponds to signal energy, second term corresponds to noise energy and the last term is the E_{RR} from eq. (18), and λ and γ are positive constants. Equation (19) can be rewritten as:

$$(\mathrm{PF}) = D^T [H^{-1} A (H^{-1} A)^T - \lambda H^{-1} (H^{-1})^T - \gamma I] D + 2\gamma D^T C \tag{20}$$

The best D which maximizes (PF) is given by

$$D^* = 2\gamma \, [H^{-1} A (H^{-1} A)^T - \lambda H^{-1} (H^{-1})^T - \gamma I]^{-1} C \tag{21}$$

IV. RESULTS OF COMPUTER SIMULATION

In our computer simulations we used uniformly distributed pseudo-random noise as the input signal $x(t)$ to the A/D converter. This was held

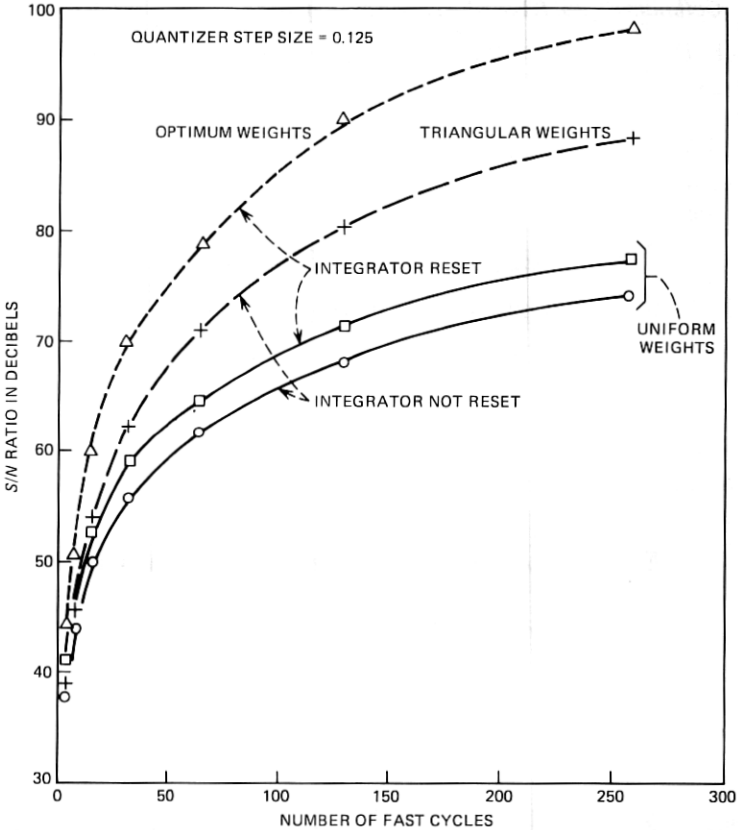

Fig. 2—Performance of an interpolative A/D converter with various digital filter weights.

constant throughout each slow cycle. We also considered cases when the input signal was filtered by an appropriate filter before going into the A/D converter. Simulations were carried out with the quantizer having two different step sizes, namely 0.125 and 0.0625 (signal range 0–1). The quantizer was assumed to have an unlimited number of levels and thus the effects of saturation were neglected. This assumption becomes more restrictive when the gain of the analog filter in the fast loop is increased. To evaluate the dependence of S/N on the number of fast cycles, several (4, 8, 16, . . . ,256) values of fast cycles were used. For the purpose of comparison, we also considered the following cases:

(*i*) Uniform weights, i.e., $D_j = 1, j = 1, \ldots, N$, with integrator not reset.

(*ii*) Triangular weights, i.e., $D_j = \min(j, N + 1 - j), j = 1, \ldots, N$, with integrator not reset. Both these weights have been investigated previously.[3,4]

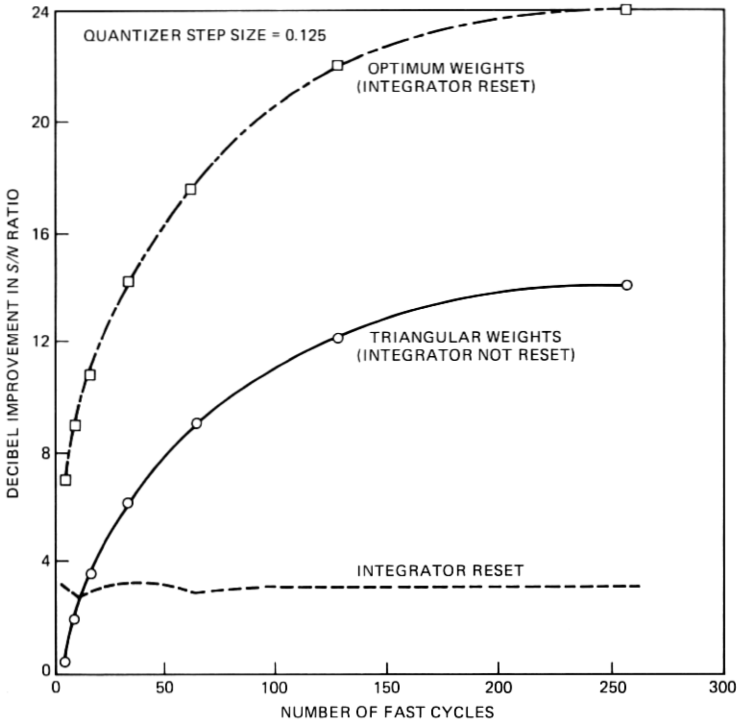

Fig. 3—Improvements in S/N ratio over constant weight digital filter.

4.1 *Effect of integrator reset with uniform weights*

The effect of resetting the integrator in the fast loop was evaluated by using uniform weights for two cases: (*i*) integrator reset, (*ii*) integrator not reset. The resulting S/N ratios are plotted in Fig. 2. The improvement in S/N ratio by resetting the integrator is plotted in Figs. 3 and 4, for two quantizer step sizes. It is seen that there is about 3 dB improvement by resetting the integrator, and this improvement is somewhat independent of the quantizer step size and the number of fast cycles. This can be easily explained by rewriting eq. (8), for $h_i = 1$ and $D_i = 1$, as:

$$\sum_{i=1}^{N} q_i = Nx - n_N - q_0 \tag{22}$$

Thus there is an extra term on the right-hand side, q_0, if there is no reset. Assuming that it is comparable to n_N, and that it is not correlated with n_N, the S/N ratio would decrease by about 3 dB due to its presence. We also simulated the effects of correlations in the input data, by filtering the pseudorandom noise, and then putting it through the A/D converter. Several low-pass filters were tried, and it was observed that the im-

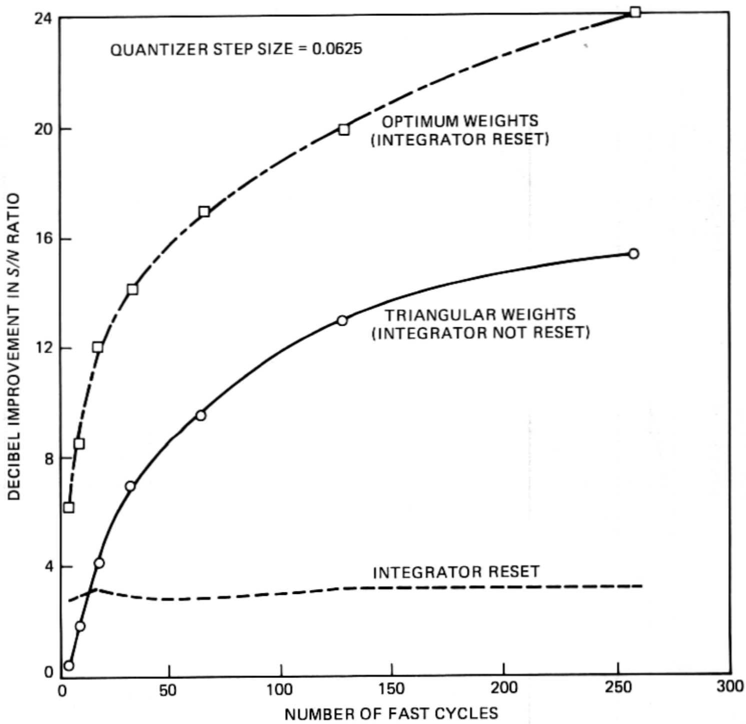

Fig. 4—Improvements in S/N ratio over constant weight digital filter.

provement in S/N ratio was still 3 dB regardless of the amount of low-pass filtering.

4.2 Effect of optimum digital filter weights

Figure 2 shows the effects of optimum digital filter weights on the S/N ratio. The analog filter in the loop is assumed to be a pure integrator with unity gain and it is reset at the beginning of each slow cycle. Figure 2 also shows the advantages of using the triangular weights, proposed by Candy et al., when the integrator is not reset. As observed by Candy et al., triangular weights are significantly better than the uniform weights, and the optimum weights allow a further increase in S/N ratio over the triangular weights. Figures 3 and 4 show the improvements in S/N ratio over those obtainable by the uniform weights when the integrator is not reset. It is seen that the rate of change of S/N ratio depends upon the number of fast cycles and is in close agreement with that predicted by eq. (16b). The S/N ratio using uniform weights when the integrator is not reset is given by (except for a proportionality constant)

$$S/N = \frac{N^2}{2} \tag{23a}$$

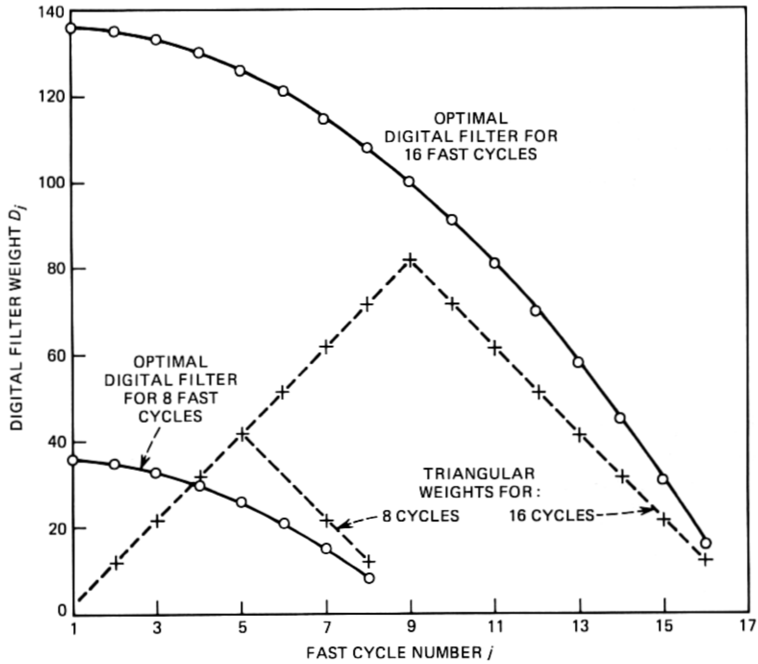

Fig. 5—Weights of optimum- and triangular-weight digital filters.

and with triangular weights

$$S/N = \frac{(N+1)^3}{16} \tag{23b}$$

These are derived by Candy et al.[4] Our simulations are in close agreement with the above equations. Thus when $N = 32$, the improvement in S/N ratio by using triangular weights over uniform weights is 6.55 dB, which is close to 6.40 dB predicted by the above equations. Similar agreement is found at other values of N. Also for large values of N the improvement obtained by our optimum weights, in the presence of integrator being reset, over the triangular weights without resetting the integrator is about 7.25 dB as predicted by eqs. (23b) and (16b). Our simulations indicate that this improvement varies between 5.80 and 10.0 dB with a mean of 7.57 dB. This is a little higher than that predicted by the equations, however the agreement is satisfactory.

The weights of the optimum digital filters are shown in a graphical form in Fig. 5 along with the triangular weights used by Candy et al.[4] It is seen that they have a parabolic shape. Although we have not considered the effect of approximations, for implementational simplicity, the filter shape could be approximated by piecewise straight lines.

4.3 *Effect of variation of the analog filter in the fast loop*

We attempted to evaluate the effect of varying the analog filter in the fast loop on the S/N ratio. It is known that certain types of analog filters tend to make the fast loop unstable; however, we did not consider questions of stability. Two types of transfer functions for the analog filter were considered:

$$H_1(s) = \frac{\overline{\alpha}}{s + \overline{\beta}}$$

and

$$H_2(s) = \frac{\overline{\alpha}}{s(s + \overline{\beta})}$$

The first case resulted in $h_i = \alpha e^{-i\beta}$ (using eq. 4) and the second case gave $h_i = \alpha - \sigma e^{-i\beta}$, where constants α, β, and σ are related to $\overline{\alpha}$, $\overline{\beta}$. Several simulations were run by varying α, β, and σ. For each of these simulations, the optimum weights were computed by eq. (15), and the resulting S/N ratio was compared with that obtained by using a pure integrator in the fast loop and the optimum digital filter. We considered only 32 fast cycles and a quantizer step size of .0625. In the first case, it was found that larger α and smaller β generally gave better S/N ratio. At $\alpha = 1.2$ and $\beta = 0.01$, the improvement in S/N ratio was about 3.0 dB. For many other cases studied, the improvement was somewhat marginal. For the second case, again, larger values of α, smaller values of β and σ around 1.0 gave the best results. At $\alpha = 1.8$, $\beta = 0.01$, and $\sigma = 1.0$, the improvement in S/N ratio was about 4.2 dB over that obtained by pure integrator in the fast loop. Thus it appears that S/N ratio can be further improved by a proper choice of the analog filter in the loop.

4.4 *Effect of spectrally shaped digital filters*

Our final simulations used digital filters which resemble a given digital filter as far as possible. For our simulation we obtained the desired digital filter characteristics from an analog function $C(t)$ whose Fourier transform $C(f)$ was 0 outside $|f| > \Omega$ and was constant (=Mag) in the interval $|f| \leq \Omega$. Sampling such a function at N times the Nyquist rate (corresponding to the number of fast cycles) gave

$$\begin{aligned} C_i &= C(i/2\Omega N) \\ &= \frac{2 \cdot \text{Mag} \cdot \Omega \cdot \text{N}}{\pi i} \sin{(i\pi/N)} \qquad i = 0, \ldots, N-1 \end{aligned}$$

Using these weights for the desired filter characteristics and some values of λ, γ (of Section 3.1), optimum digital filters with spectral shaping were obtained for the case when the analog filter in the fast loop was a pure integrator. Computer simulations were carried out for various values of λ and γ, $N = 32$, and quantizer step size = .0625. Two quantities were measured: (*i*) S/N ratio as before, (*ii*) the noise power in frequency band $-\Omega$ to $+\Omega$. It was found that by giving a high value to γ (i.e., heavily penalizing any deviation of the filter characteristics from the desired characteristics), a decrease of about 1 dB in the noise power in frequency band $-\Omega$ to $+\Omega$ was possible. However, this resulted in a decrease of S/N ratio by about 1.5 dB. Thus it appears that the inband noise could be suppressed to some extent at the expense of decrease of overall S/N ratio.

SUMMARY AND CONCLUSIONS

In this paper, we have given two techniques for optimizing the digital filter characteristics of an interpolative A/D converter. Computer simulations showed that the optimum digital filters with the integrator reset increases the signal-to-noise ratio by as much as 23 dB over that obtainable by a digital filter with uniform weights and no resetting of the integrator. We also showed that by resetting the integrator a 3 dB advantage in signal-to-noise ratio is obtained when uniform weights are used. We varied the transfer function of the analog filter in the fast loop and found that a gain of a few decibels is possible by proper choice of the analog filter. Finally we considered digital filters whose characteristics could be made close to certain desirable characteristics, and found that it is possible to decrease the quantization noise power within a band, but only at the expense of decrease of the overall signal-to-noise ratio. We note that two important factors, which we have not paid attention to, are: (*i*) stability of the fast loop, and (*ii*) simplicity of implementation of the digital filters. These would be crucial in any practical implementation of the interpolative A/D converters.

ACKNOWLEDGMENT

The author wishes to thank J. C. Candy who, through many discussions, contributed to the development of this work.

REFERENCES

1. D. J. Goodman, "The Application of Delta Modulation to Analog-to-Digital PCM Encoding," B.S.T.J., *48*, No. 2, (February 1969), pp. 321–343.
2. D. J. Goodman and L. J. Greenstein, "Quantization Noise of DM/PCM Encoders," B.S.T.J., *52*, No. 2, (February 1973), pp. 183–204.
3. J. C. Candy, "A Use of Limit Cycle Oscillations to Obtain Robust Analog-to-Digital Converter," IEEE Trans. Commun., *COM-22*, March 1974, pages 298–305.
4. J. C. Candy, Y. C. Ching, and D. S. Alexander, "Using Double Interpolation to Get 13-Bit PCM from a Sigma-Delta Modulator," IEEE Trans. Commun., *COM-24*, November 1976, pp. 1268–1275.