A Digital Signal Processing Approach to Interpolation

RONALD W. SCHAFER AND LAWRENCE R. RABINER

Abstract—In many digital signal processing systems, e.g., vocoders, modulation systems, and digital waveform coding systems, it is necessary to alter the sampling rate of a digital signal. Thus it is of considerable interest to examine the problem of interpolation of bandlimited signals from the viewpoint of digital signal processing. A frequency domain interpretation of the interpolation process, through which it is clear that interpolation is fundamentally a linear filtering process, is presented.

An examination of the relative merits of finite duration impulse response (FIR) and infinite duration impulse response (IIR) digital filters as interpolation filters indicates that FIR filters are generally to be preferred for interpolation. It is shown that linear interpolation and classical polynomial interpolation correspond to the use of the FIR interpolation filter. The use of classical interpolation methods in signal processing applications is illustrated by a discussion of FIR interpolation filters derived from the Lagrange interpolation formula. The limitations of these filters lead us to consider optimum FIR filters for interpolation that can be designed using linear programming techniques. Examples are presented to illustrate the significant improvements that are obtained using the optimum filters.

II. DIGITAL SAMPLING RATE ALTERATION

A. Sampling Continuous-Time Signals

Consider a continuous-time signal \( \hat{x}(t) \) with Fourier transform

\[
\hat{X}(\omega) = \int_{-\infty}^{\infty} \hat{x}(t)e^{-j\omega t} dt.
\]

The signal \( \hat{x}(t) \) is sampled to produce the sequence

\[
x(n) = \hat{x}(nT), \quad -\infty < n < \infty
\]

where \( T \) is the sampling period. The \( z \) transform of the sequence \( x(n) \) is defined as

\[
\hat{X}(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}.
\]

The \( z \) transform evaluated on the unit circle \( \hat{X}(e^{j\omega}) \) will be called the Fourier transform of the sequence \( x(n) \). It is well known that the Fourier transform of the sequence \( x(n) \) is related to the Fourier transform of \( \hat{x}(t) \) by [5]

\[
X(e^{j\omega}T) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \hat{X}(\omega + k\frac{2\pi}{T}).
\] (1)

If \( \hat{x}(t) \) is bandlimited, i.e., \( \hat{X}(\omega) = 0 \) for \( |\omega| \geq \Omega \), and if \( T \leq \pi/\Omega \), then it can be seen from (1) that

\[
X(e^{j\omega}T) = \frac{1}{T} \hat{X}(\omega), \quad -\frac{\pi}{T} \leq \omega \leq \frac{\pi}{T}
\]

as depicted in Fig. 1, where \( T = \pi/\Omega \).

Assuming that \( \hat{x}(t) \) is bandlimited, the original continuous-time signal can be obtained uniquely from the samples \( x(n) \) through the interpolation formula

\[
\hat{x}(t) = \sum_{k=-\infty}^{\infty} x(k) \frac{\sin \left[ \frac{\pi}{T} (t - kT) \right]}{\pi (t - kT)}.
\] (2)

In many digital signal processing problems, we are given a sequence \( x(n) \), corresponding to sampling period \( T \), and we must obtain from the sequence \( x(n) \) a sequence \( y(n) = \hat{x}(nT) \); i.e., the sequence \( y(n) \) corresponds to sampling \( \hat{x}(t) \) at a different sampling rate. If we evaluate (2) for \( t = nT \), we obtain

Manuscript received December 6, 1972; revised March 7, 1973.

The authors are with Bell Laboratories, Murray Hill, N.J. 07974.

Fig. 1. Illustration of the relationship between the Fourier transform of a continuous-time signal and the Fourier transform of the sequence obtained by sampling with period T.

A direct relationship between \( y(n) \) and \( x(n) \), but it is clear that such an expression is impossible to evaluate because the functions \( \sin \left( \frac{\pi}{T} (t-kT) \right) / \left( \sin \left( \frac{\pi}{T} (t-kT) \right) \right) \) are of infinite duration. Rather than simply truncate these functions, it is more reasonable to design finite duration interpolators. To understand how such interpolators can be designed and understand the limitations of classical interpolators, it is useful to consider the frequency-domain representation of the process of changing the sampling rate.

B. Sampling Rate Reduction—Integer Factors

Suppose that the desired sampling period is \( T' = MT \). If \( M \) is an integer, this simply implies that the new sequence is

\[
y(n) = \hat{x}(nT') = \hat{x}(nMT) = x(Mn).
\]

That is, the sequence \( x(n) \) is "sampled" by retaining only one out of each group of \( M \) consecutive samples. The values of the sequence \( y(n) \) are samples of the original sequence \( \hat{x}(t) \); however, these samples will uniquely determine \( \hat{x}(t) \) if and only if \( T' \leq \pi / \Omega \). This is clearly just a consequence of the sampling theorem as expressed by (1).

Since we are interested in direct relationships between the sequences \( y(n) \) and \( x(n) \), it is instructive to derive an equation similar to (1) that relates the Fourier transforms of the two sequences. The derivation of the equation is facilitated by the definition of a new sequence \( w(n) \) which is nonzero only at integer multiples of \( M \); that is

\[
w(n) = \begin{cases} x(n), & n = 0, \pm M, \pm 2M, \cdots \\ 0, & \text{elsewhere} \end{cases}
\]

where the sampling period is assumed to be \( T \) for both sequences. A convenient representation of \( w(n) \) is

\[
w(n) = x(n) \left\{ \sum_{\ell=-\infty}^{M-1} e^{j2\pi \ell n / M} \right\}, \quad -\infty < n < \infty
\]

where the term in brackets may be recognized as a discrete Fourier series representation of a periodic sequence that is one at integer multiples of \( M \) and zero otherwise. The sequence \( y(n) \), corresponding to sampling period \( T' = MT \), is

\[
y(n) = w(Mn), \quad -\infty < n < \infty.
\]

Therefore

\[
Y(z) = \sum_{n=-\infty}^{\infty} w(Mn)z^{-n}.
\]

Since \( w(n) \) is zero except at integer multiples of \( M \), we obtain

\[
Y(z) = \sum_{n=-\infty}^{\infty} w(n)z^{-n/M} = \sum_{n=-\infty}^{\infty} x(n) \sum_{\ell=0}^{M-1} e^{j2\pi \ell n / M}z^{-n/M} = \frac{1}{M} \sum_{\ell=0}^{M-1} \sum_{n=-\infty}^{\infty} x(n) e^{j2\pi \ell n / M}z^{-n/M} = \frac{1}{M} \sum_{\ell=0}^{M-1} X(e^{j2\pi \ell / M}) z^{\ell / M}.
\]

(3)

If we evaluate \( Y(z) \) on the unit circle, with normalization appropriate for the new sampling rate, we obtain

\[
Y(e^{j\omega T'}) = \frac{1}{M} \sum_{\ell=0}^{M-1} X(e^{j(\omega T'-2\pi \ell / M)}).
\]

(4)

There is a clear similarity between (4) and (1).

An example of sampling rate reduction by a factor of 2 is shown in Fig. 2. The Fourier transform of \( x(n) \) is depicted in Fig. 2(a) for the case when \( \pi / 2\Omega < T < \pi / 2 \) so that

\[
X(e^{j\omega T}) = \frac{1}{T} \hat{X}(\omega), \quad -\frac{\pi}{T} < \omega < \frac{\pi}{T}.
\]

Fig. 2(b) shows \( Y(e^{j\omega T'}) \) for \( T' = 2T \). In this case, aliasing occurs and it is clear that, in general, aliasing will occur in the process of digital sampling rate reduction unless the original sampling period satisfies
where $\Omega$ is the Nyquist frequency of $\hat{x}(t)$. If this inequality is satisfied, however, then
\[
Y(\epsilon^{j\omega T}) = \frac{1}{M} X(\epsilon^{j\omega T})
\]
\[
= \frac{1}{MT} \hat{Y}(\omega)
\]
\[
= \frac{1}{T'} \hat{X}(\omega), \quad -\frac{\pi}{T'} < \omega < \frac{\pi}{T'}.
\]
If the original sampling period does not satisfy (5), aliasing distortion can be avoided only by passing the sequence $x(n)$ through an ideal low-pass digital filter with cutoff frequency $\pi/T'$. It can be seen that the filter must have unit gain, since (4) provides the factor $1/M$ needed to correct the amplitude for the new sampling rate. This, of course, results in a sequence $y(n)$ corresponding to a continuous time signal $\gamma(t)$ which is a low-pass filtered version of the original signal $\hat{x}(t)$.

C. Sampling Rate Increase—Integer Factors

If the sampling rate is increased by an integer factor $L$, then the new sampling period will be $T' = T/L$. Since the sequence $x(n)$ provides samples of the desired sequence only at intervals of $L$ samples at the new sampling rate, the remaining samples must be filled in by interpolation. To see how this can be done using a digital filter, consider the sequence
\[
v(n) = x(n/L), \quad n = 0, \pm L, \pm 2L, \cdots
\]
\[
= 0, \quad \text{otherwise}.
\]
The $z$ transform of $v(n)$ is
\[
V(z) = \sum_{n=-\infty}^{\infty} x(n/L)z^{-n}
\]
\[
= \sum_{n=-\infty}^{\infty} x(n)z^{-Ln} = X(z^L).
\]
The Fourier transform of this sequence is
\[
Y(\epsilon^{j\omega T'}) = X(\epsilon^{j\omega L})
\]
Thus $Y(\epsilon^{j\omega T'})$ is periodic with period $2\pi/T' = 2\pi/LT'$, rather than $2\pi/T$ as is the case in general for sequences associated with a sampling period $T$. Fig. 3(a) shows $Y(\epsilon^{j\omega T'})$ for the case $T' = T/3$. If we wish to obtain a sequence $y(n)$ such that
\[
y(n) = \hat{x}(nT')
\]
then we must insure that
\[
Y(\epsilon^{j\omega T'}) = \frac{1}{T'} \hat{Y}(\omega), \quad -\frac{\pi}{T'} \leq \omega \leq \frac{\pi}{T'}.
\]
Assuming that
\[
X(\epsilon^{j\omega T}) = \frac{1}{T} \hat{X}(\omega), \quad -\frac{\pi}{T} \leq \omega \leq \frac{\pi}{T}
\]
then it is clear from Fig. 3(a), that the images of $(1/T)\hat{X}(\omega)$

![Fig. 3. Sampling rate increase ($T' = T/3$). (a) Fourier transform of sequences $x(n)$ and $v(n)$. (b) Fourier transform of desired output interpolation process.

\[
v(\epsilon^{j\omega T'}) = \frac{1}{T} H(\epsilon^{j\omega T'}) Y(\epsilon^{j\omega T'})
\]
\[
= \frac{1}{T} H(\epsilon^{j\omega T'}) \hat{X}(\omega),
\]
where $H(\epsilon^{j\omega T'})$ is periodic with period $2\pi/T'$ and
\[
H(\epsilon^{j\omega T'}) = L, \quad |\omega| \leq \frac{\pi}{T'}
\]
\[
= 0, \quad \frac{\pi}{T'} < |\omega| < \frac{\pi}{T'}.
\]
Thus the ideal interpolation scheme for increasing the sampling rate requires the creation of a sequence of $L$-zero-valued samples between each value of the original sequence, which is then filtered with an ideal low-pass filter in (8).

D. Changing by Noninteger Factors

In the previous two subsections we have discussed methods for increasing or decreasing the sampling rate using a linear time-invariant digital filter. However, because of the fact that we required that the process be discrete-time and that we have a relationship between the two integer multiple or submultiple of the original sampling period $T$. This restriction can be eased somewhat by a two-step process involving a sampling rate increase followed by a decrease.

Suppose the desired sampling period is $T' = (M/L)T$, where $M$ and $L$ are integers. This does not seem to be a significant limitation in practice since any factor can be approximated as closely as desired by proper choice of $M$ and $L$. The two-step process for interpolating to a sampling period $T' = (M/L)T$ is as follows:

1. Increase the sampling rate by a factor of $L$ by interpolation as in Section II-C. Let the resulting sequence be denoted $y_L(n)$ with sampling period $T_1 = T/L$.

2. Decrease the sampling rate by a factor of $M$ as in Section II-B to obtain the desired sequence $y(n)$ with sampling period $T' = MT_1 = (M/L)T$.
This process is illustrated in Fig. 4 for \( T' = \frac{3}{4} T \); i.e., for a set increase in sampling rate. Fig. 4(a) shows the Fourier transform of the original sequence \( x(n) \). Fig. 4(b) shows the Fourier transform of the intermediate sequence \( y_1(n) \) which results from filtering the sequence

\[
\begin{align*}
\gamma(n) &= x(n), \quad n = 0, \pm 3, \pm 6, \cdots \\
&= 0, \quad \text{elsewhere}
\end{align*}
\]

with an ideal low-pass filter having gain 3 and cutoff frequency \( \pi/7 \). The result of reducing the sampling rate of \( y_1(n) \) by a factor of 2 is shown in Fig. 4(c).

If \( M < L \), i.e., there is a net increase in sampling rate as in Fig. 4, no aliasing can occur, and the interpolation filter can have a cutoff frequency \( \pi/7 \). However, if \( M > L \), there is a possibility of aliasing. In this case aliasing can be avoided by making the cutoff frequency of the interpolation filter equal to \( \pi/7 \). As in Section II-B, the resulting output sequence \( y(n) \) will correspond to samples of a low-pass filtered version of the original continuous time signal.

III. INTERPOLATION USING FIR DIGITAL FILTERS

The previous section makes it abundantly clear that the process of changing the sampling rate requires a low-pass filter. Since it is impossible to realize the ideal low-pass filter that is required for exact results, we must consider digital filters that approximate this ideal behavior. As in all filtering problems, there are many important considerations. A basic consideration is the choice between filters from the class of FIR filters and from the class of IIR filters. Given the type of filter to be used, the problem of approximating the ideal low-pass filter must be solved. Finally, there are important considerations in how the filter is realized as software or hardware. All of these facets of the problem are interrelated—resulting in arbitrary tradeoffs between accuracy of interpolation and efficiency of realization. In this section we present some observations on filter design and realization that seem to imply that FIR filters are the proper choice for most interpolation problems.

A. Phase Distortion

The ideal interpolator has zero phase or at most a linear phase corresponding to an integer number of samples of delay. IIR filters cannot have precisely linear phase [6]. In contrast, there currently exist several techniques for designing optimal FIR digital filters with precisely linear phase. These filters are optimal in the sense that the width of transition band between passband and stopband is minimum for given values of passband and stopband rolloff and specified passband and stopband roll-off frequencies [7]–[11]. These filters can be designed with arbitrarily small values for passband ripple, stopband ripple, and transition bandwidth, at the cost of increased impulse response duration. Thus with FIR filters, the interpolation error due to phase nonlinearity can be zero and the error due to amplitude distortion can be made arbitrarily small. In the case of IIR filters, although extremely good amplitude characteristics can be achieved, there will always be an interpolation error due to phase nonlinearity.

B. Filter Realization

IIR filters have recursive realizations that are very economical in terms of computational complexity. Leaving phase considerations aside, FIR filters in general require more computation to achieve a given accuracy of approximation to the desired amplitude response than do IIR filters. However, the particular nature of the interpolation problem makes FIR filters computationally competitive with IIR filters.

Consider for convenience a zero-phase FIR filter with impulse response \( h(n) \) which is nonzero in the interval

\[
\left( -\frac{N - 1}{2} \right) \leq n \leq \left( \frac{N - 1}{2} \right)
\]

where \( N \) is an odd integer. In reducing the sampling rate by an integer factor, i.e., \( T' = MT \), we may need a unity gain low-pass filter to insure that no aliasing occurs in retaining only every \( M \)th sample of the sequence \( x(n) \). In this case computation is reduced because of the nature of the desired output sequence. The filtered output sequence at the original sampling rate, defined as \( \gamma(n) \), is

\[
\gamma(n) = \sum_{k=-\left(\frac{N-1}{2}\right)}^{\left(\frac{N-1}{2}\right)} h(k)x(n-k) \tag{9}
\]

where all sequences in (9) are associated with sampling period \( T \). Clearly, all values of the sequence \( \gamma(n) \) need not be computed since the desired output is

\[
\gamma(n) = \sum_{k=-\left(\frac{N-1}{2}\right)}^{\left(\frac{N-1}{2}\right)} h(k)x(nM-k) \tag{10a}
\]

where \( \gamma(n) \) is associated with a sampling period \( T' = MT \). This is in contrast to a comparable IIR filter where the computations required to realize the poles of the system function would have to be carried out at the original sampling rate even

\[1 \text{ See Section III-C for a comment on why } N \text{ should be odd.} \]
though \( M - 1 \) out of \( M \) samples of \( y(n) \) would be discarded. A further simplification results from the fact that zero-phase FIR filters have the property
\[
h(n) = h(-n).
\]

Thus (10a) becomes
\[
y(n) = \sum_{k=1}^{((N-1)/2)} h(k) [x(nM - k) + x(nM + k)] + h(0)x(nM). \tag{10b}
\]

In the case of sampling rate increase, the interpolated output is obtained with sampling period \( T' = T/L \) by filtering the sequence
\[
v(n) = x(n/L), \quad n = 0, \pm L, \pm 2L, \ldots = 0, \text{ otherwise.}
\]

In this case it is convenient to write
\[
y(n) = \sum_{k=-(N-1)/2}^{N-1/2} v(k)h(n - k). \tag{11}
\]

Substituting for \( v(k) \) results in
\[
y(n) = \sum_{k=-(N-1)/2}^{N-1/2} x(k/L)h(n - k), \quad k/L \text{ an integer}
\]
\[
= \sum_{k=-(N-1)/2}^{N-1/2} x(k/L)h(n - kL) \tag{12}
\]

where \( \lfloor a \rfloor \) means “the largest integer contained in \( a \).” Although (11) suggests that computation required for each output sample is proportional to \( N \), we note that only one out of every \( L \) samples of \( v(n) \) is nonzero. Thus we see from (12) that the actual computation required is proportional to \( N/L \). Note that in this case the symmetry of the impulse response cannot be exploited to reduce computation.

If an IIR filter were used, relatively little saving could be achieved. In fact, in a cascade realization, almost no computational saving could be attained by considering the zeroes in the input sequence.

Changing the sampling period according to \( T' = (M/L)T \), requires that we first increase the sampling rate by a factor \( L \) and then reduce it by a factor \( M \). Clearly, the savings previously discussed could be incorporated into both steps of this process.

C. Impulse Response Constraints

The previous discussion has presented compelling reasons for the use of linear-phase FIR filters in interpolation. To conclude this section we discuss some constraints on the impulse response that are specific to the problem of interpolation. Recall that the output \( y(n) \) is given by (12). A reasonable requirement on the interpolation filter is that the values of the output at the original sampling times be the same as the original samples. That is, for \( r \) an integer,
\[
y(rL) = x(rL/L) = x(r), \quad -\infty < r < \infty.
\]

Equation (12) implies that
\[
y(rL) = x(r) = \sum_{k=-(N-1)/2}^{(N-1)/2} x(k)h(rL - kL) \tag{13}
\]

for all integer values of \( r \). From this equation, we see that
\[
h(0) = 1
\]
\[
h(n) = 0, \quad n = \pm L, \pm 2L, \ldots, \left\lfloor \frac{N - 1}{2L} \right\rfloor L. \tag{14}
\]

Constraints of this type are rather difficult to impose on an IIR filter design procedure.

Some final comments on the choice of the length of the impulse response \( N \) are in order. First, we have assumed that \( N \) is an odd integer; however, in general \( N \) can be either even or odd for the FIR filter. It can be shown [6], [12] that for \( N \) even, a linear-phase FIR filter must have a delay of at least one one-half sample. This one-half sample delay itself corresponds to interpolation between samples; thus such a filter could not preserve the samples of the original sequence. Although there may be instances where the half-sample delay may not be objectionable, or may even be desirable, odd values of \( N \) appear to be appropriate for most applications.

A second comment regarding the choice of \( N \) concerns the fact that we have asserted that by increasing \( N \), we achieve a better approximation in the frequency domain to the ideal interpolating filter. In the time domain, increasing \( N \) implies that more values of the original sequence are involved in the computation of a given interpolated sample \( y(n) \) thus we should expect increasingly better results as \( N \) increases. We have observed that because the sequence \( v(n) \) has mostly zero samples, the computation is proportional to \( N/L \) since the impulse response always spans approximately \( N/L \) nonzero samples. Suppose that we wish to always have \( \lfloor N/L \rfloor \) samples of the original input sequence involved in the computation of each interpolated sample. Using (12) it is easily shown that the length of the impulse response should be
\[
N = QL, \quad \text{if } Q \text{ and } L \text{ are odd}
\]
\[
= QL - 1, \quad \text{if either } Q \text{ or } L \text{ are even.} \tag{15}
\]

If \( N \) is slightly larger than this value, the computation of some interpolated samples will involve \( Q+1 \) original samples while the rest will involve only \( Q \). If \( N \) is slightly smaller than the value given by (15), then the interpolation will involve either \( Q \) or \( Q-1 \) samples. Thus to insure that \( Q \) samples of the original sequence are always involved, \( N \) should satisfy (15).

This constraint is satisfied, as we will see, for filters derived from classical interpolation formulas. It is also important in hardware and software realizations of interpolation filters where it allows the computations to be structured so that 

C. Classical Polynomial Interpolation

In this section we apply the previous discussion of the interpolation process to a study of classical polynomial interpolation methods. Our aim is to give a frequency domain interpretation of these formulas that will shed some light on their applicability in interpolation problems arising in digital signal processing. In the course of this discussion, we shall indicate how to derive linear phase interpolation filters from tables of Lagrange coefficients. We begin with a discussion of linear interpolation.

A. Linear Interpolation

Linear interpolation involves only two consecutive samples of the original sequence \( x(n) \) in the computation of \( y(n) \).
interpolated sample. Specifically, the values interpolated between two samples \( x(0) \) and \( x(1) \) lie on a straight line connecting the two original samples, with the original samples of course being preserved. Thus the equation relating the output \( y(n) \), having sampling period \( T' = T/L \), to the input sequence \( x(n) \), having sampling period \( T \), is

\[
y(n) = x(0) + \frac{x(1) - x(0)}{LT'} (nT' - 0)
\]

\[
= x(0) \left( 1 - \frac{n}{L} \right) + x(1) \left( \frac{n}{L} \right), \quad 0 \leq n < L.
\]  

(16)

In order to interpret linear interpolation as a linear filtering process, we must derive an impulse response for the linear interpolation filter. This can be done by comparing (16) to (12). To begin, it is clear that the length of the impulse response must be

\[
N = 2L - 1
\]

as discussed in Section III-C. If \( N \) were larger, more than 2 samples of \( x(n) \) would be involved in the computation of some of the interpolated values. Likewise, if \( N \) were smaller, only one sample of the input would enter into the computation of some of the interpolated values. Using this information, we can write (12) as

\[
y(n) = x(0)h(n) + x(1)h(n - L), \quad 0 \leq n < L.
\]  

(17)

Comparing (16) and (17) we see that

\[
h(n) = 1 - \frac{n}{L}, \quad 0 \leq n < L
\]

\[
h(n - L) = \frac{n}{L}, \quad 0 \leq n < L.
\]

Thus \( h(n) \) is seen to be

\[
h(n) = 1 - \left| \frac{n}{L} \right|, \quad \left| \frac{n}{L} \right| < L
\]

\[
= 0, \quad \text{otherwise.}
\]  

(18)

Clearly \( h(n) \) satisfies the requirements of (14) since \( h(0) = 1 \) and \( h(n) = 0 \) for \( \left| \frac{n}{L} \right| > L \). The length of the impulse response is

\[
N = 2L - 1
\]

consistent with (15) for \( Q = 2 \).

Fig. 5 depicts linear interpolation as a convolution process. The sequence \( v(k) \) and the triangular envelope of the impulse response \( h(n-k) \) are shown for the case of \( T' = T/5 \). It is clear that because the impulse response has duration \( N = 5(2) - 1 = 9 \), only two nonzero samples of \( v(k) \) are ever coincident with \( h(n-k) \). Also it is clear that

\[
y(n) = x(n/L), \quad n = 0, \pm L, \pm 2L, \cdots.
\]

The system function corresponding to linear interpolation is

\[
H(e^{j\omega T'}) = \frac{1}{L} \left( \sin \left[ \omega LT'/2 \right] \right)^2.
\]  

(19)

This system function is plotted in Fig. 6 for \( L = 5 \). (Curve labelled \( Q = 2 \).) We recall that the purpose of the interpolation filter is to remove the images of the signal spectrum that are centered at integer multiples of \( 2\pi/T \), while leaving the frequencies below \( \pi/T \) unaltered. It can be seen from Fig. 6, that linear interpolation achieves significant attenuation in only a very small region around each integer multiple of \( 2\pi/T \). Specifically, attenuation is \( 40 \text{ dB} \) or greater in a band of width \( 0.35\pi/T \) centered at \( 2\pi/T, 4\pi/T, \) etc. Thus it seems reasonable to note that linear interpolation is appropriate only if the original sampling rate is many times the Nyquist rate.

B. Lagrange Interpolation

Clearly linear interpolation will not be satisfactory in many digital signal processing applications. In classical numerical analysis, the inadequacies of linear interpolation lead to the use of higher order polynomials; i.e., in contrast to connecting two points by a straight line, one finds a polynomial of degree \( Q - 1 \) that passes through \( Q \) original samples. The interpolated values then, are samples of this polynomial. A variety of formulas have been derived for obtaining samples of the polynomial directly from the samples \( x(n) \), however since we are only interested in the interpolation filter corresponding to polynomial interpolation, we shall use the most convenient form; namely, the Lagrange interpolation formula [13], [14]. The form corresponding to interpolation with equal spacing from sampling period \( T \) to \( T' = T/L \) is
\[ y(n) = \sum_{k=-(Q-2)/2}^{Q/2} A_k^Q(n/L)x(k), \quad Q \text{ even} \] (20a)

\[ y(n) = \sum_{k=-(Q-1)/2}^{(Q-1)/2} A_k^Q(n/L)x(k), \quad Q \text{ odd} \] (20b)

where we have again chosen to consider interpolation around some arbitrary time labelled \( O \). In (20), the quantities \( A_k^Q(\cdot) \) are called the Lagrange coefficients and are given by the equations

\[ A_k^Q(t) = \frac{(-1)^{k+Q/2}}{\left(\frac{Q-2}{2}\right)!(\frac{Q}{2}-k)!(t-k)} \prod_{\omega=1}^{Q/2} \left( t + \frac{Q}{2} - 1 \right), \quad Q \text{ even} \] (21a)

\[ A_k^Q(t) = \frac{(-1)^{k+(Q-1)/2}}{\left(\frac{Q-1}{2}\right)!(\frac{Q-1}{2}-k)!(t-k)} \prod_{\omega=1}^{Q-1/2} \left( t + \frac{Q-1}{2} - 1 \right), \quad Q \text{ odd}. \] (21b)

Extensive tabulations of these coefficients are available in tables of mathematical functions [15]. It is interesting to note that

\[ A_k^Q(t) = 0, \quad t \text{ an integer, and } k \neq k. \]

\[ = 1, \quad t = k. \]

This is a result of the fact that the interpolation polynomial passes through the \( Q \) original data samples. That is,

\[ y(kL) = x(k), \quad \text{for } k \text{ an integer}. \]

Thus the condition of (14) is satisfied for Lagrange interpolation filters.

In general, the formulas in (20) may be evaluated for any value of \( n \) such that

\[ -\left(\frac{Q-2}{2}\right)L < n < \frac{Q}{2}L, \quad Q \text{ even} \]

\[ -\left(\frac{Q-1}{2}\right)L < n < \left(\frac{Q-1}{2}\right)L, \quad Q \text{ odd}. \]

In this case the same \( Q \) original samples are involved in the computation of all the interpolated samples in the time interval spanned by the original samples. If we perform the interpolation with a linear filter, we can choose the length of the filter impulse response so that \( Q \) original samples are always involved in the computation; however, a given set of \( Q \) original samples is used only to compute \( L-1 \) interpolated samples. Thus we can interpret the Lagrange formulas in (20) in terms of \( Q-1 \) different impulse responses—corresponding to the \( Q-1 \) different interpolation intervals between the \( Q \) original samples. As an example, consider the case for \( Q=3 \).

In comparing (12) and (20b), we note that the three samples \( x(-1), x(0), \) and \( x(1) \) can enter into the interpolation in two intervals. Thus

\[ y(n) = \sum_{k=-1}^{1} A_k^Q(n/L)x(k) \]

\[ = \sum_{k=-1}^{1} x(k)h(n - kL) \]

for values of \( n \)

\[ -L < n \leq 0 \]

\[ 0 \leq n < L. \]

If we compare the above two equations, we obtain

\[ h(n + L) = A_{-1}^Q(n/L) \]

\[ h(n) = A_0^Q(n/L) \]

\[ h(n - L) = A_1^Q(n/L). \]

If these equations are evaluated for the first interval we obtain the impulse response of Fig. 7(a) where \( L = 5 \). Likewise, Fig. 7(b) shows the impulse response corresponding to Lagrange interpolation in the second interval. Clearly neither of these impulse responses has linear phase, since they do not satisfy the symmetry condition \( h(n) = h(-n) \). Indeed, it is easy to show that whenever \( Q \) is odd, none of the impulse responses corresponding to Lagrange interpolation can have linear phase. However, if \( Q \) is even, one of the \( Q-1 \) impulse responses does have linear phase.

As a second example suppose \( Q=4 \). Using (20a) we obtain

\[ y(n) = \sum_{k=-1}^{3} A_k^4(n/L)x(k). \] (22)
phase impulse response derived from a Q point Lagrange interpolation formula is obtained from the equations

\[ h\left(n + \left(\frac{Q-2}{2}\right) L\right) = A_{-\left(\frac{Q-2}{2}\right)}Q(n/L) \]

\[ h(n + L) = A_{-1}Q(n/L) \]
\[ h(n) = A_{0}Q(n/L) \]
\[ h(n - L) = A_{1}Q(n/L) \]

\[ h\left(n - \frac{Q}{2} L\right) = A_{Q/2}Q(n/L) \quad (23) \]

where \(0 \leq n < L\).

In Fig. 6 we have plotted the frequency response of the zero-phase interpolators derived in this manner for \(Q = 2, 4, 6, 8,\) and \(L = 5.\) Since the impulse response duration is \(N = QL - 1, N = 9, 19, 29, 39\) for these cases. Clearly, the effect of increasing \(Q\) is to improve the frequency response of the interpolator. Whereas the linear interpolator has 40-dB attenuation in a bandwidth of 0.35\(\pi/T\) centered around each integer multiple of \(2\pi/T,\) for the 4-point and 6-point interpolators, the bandwidths are 0.7\(\pi/T\) and 0.9\(\pi/T,\) respectively, for at least 40-dB attenuation. Thus at the expense of increased computation, we can achieve a significantly better interpolation by using a filter derived using (23).

The interpolation filters derived from the Lagrange interpolation formula achieve high attenuation in a narrow band around integer multiples of \(2\pi/T\) because the zeroes of the system function tend to be clustered around those frequencies. For example, in the case of linear interpolation \((Q = 2),\) by looking at (19) we see that \(H(e)\) has a double zero on the unit circle at integer multiples of \(2\pi/T.\) For \(Q = 4,\) we have found that there are clusters of 4 zeroes not precisely on the unit circle but close in the vicinity of \(\omega = 2\pi/T, 4\pi/T,\) etc. As a result, the attenuation close to frequencies \(2\pi/T\) and \(4\pi/T\) is very high. This is clear from Fig. 9, where the system function for the case \(Q = 6\) is plotted on a log scale. However, it is also clear that between \(2\pi/T\) and \(4\pi/T,\) the response of the Lagrange interpolation filter leaves much to be desired. Clearly, as \(Q\) gets larger, the impulse response gets longer and there are more zeroes to distribute so as to increase the attenuation and broaden the attenuation bands. This raises the question as to how we might design FIR digital filters so as to make the best use of the filter zeroes.

V. OPTIMUM FIR INTERPOLATION FILTERS

In practical situations, signals are often sampled at a rate that is only slightly higher than twice the Nyquist frequency in order to minimize the computation required for digital filtering and other signal processing procedures. In this case, the ideal interpolation filter for increasing the sampling rate has constant gain in the frequency range \(0 \leq |\omega| < \pi/T,\) and zero gain elsewhere. For such signals we are clearly interested in the best possible approximations to the ideal low-pass filter.

On the other hand, in situations where the original sequence \(x(n)\) is derived by sampling at a rate considerably higher than twice the Nyquist frequency, we require a relatively narrow passband of constant gain and a number of stop-
bands of zero gain, with the frequency response being somewhat arbitrary elsewhere. This means that high-order polynomial interpolation filters may be quite satisfactory for signals that are sufficiently oversampled. However, it is generally possible to achieve significantly better interpolation filters using optimization techniques.

A. Design Specifications

A scheme for approximating the ideal interpolator is shown in Fig. 10. There is a band of frequencies, \(0 \leq \omega \leq \omega_p\), called the passband in which the frequency response should be close to \(1\). In practice, we allow for an error of \(\pm \delta\). In addition, there are one or more stopbands in which the frequency response is required to be within \(\pm \delta\) of zero. The choice of the parameters in Fig. 10 depends upon the desired accuracy of interpolation and to some extent upon the nature of the input sequence. The choice of passband and stopbands depends upon the Nyquist frequency of the original signal. If the original sampling period is such that \(\pi/T = \Omega\), then \(\omega_p\) must be very close to \(\pi/T\), and the first stopband, beginning at \(\omega_m\), must also be close to \(\pi/T\). Thus the transition band \(\omega_m \leq \omega_1 \leq \omega_2\), must be very narrow, which implies that a large value of \(N\) will be required.

In such cases, it is reasonable to define only one stopband \(\omega_1 \leq \omega \leq \pi/T\). However, in cases where \(\pi/T\) is significantly higher than the Nyquist frequency, the transition band between passband and stopband can be wider, and it makes sense to define stopbands around each integer multiple of \(2\pi/T\), with transition bands \(\omega_n + \Delta\omega_1 \leq \omega \leq \omega_n + \Delta\omega_2\), etc., in which the frequency response is unconstrained. As can be seen from Figs. 4 and 9, this is the form of frequency response that characterizes the Lagrange interpolation filters.

B. Design Techniques

The frequency response of a zero-phase FIR filter can be expressed as

\[
H(e^{j\omega T}) = h(0) + 2 \sum_{n=1}^{(N-1)/2} h(n) \cos(\omega n T).
\]

(24)

The tolerance scheme depicted in Fig. 10 can be expressed in the following set of inequalities:

\[
1 - K\delta_1 \leq H(e^{j\omega T}) \leq 1 + K\delta_1, \quad 0 \leq \omega \leq \omega_p
\]

(25)

\[
- \delta_1 \leq H(e^{j\omega T}) \leq \delta_1, \quad \omega_1 \leq \omega \leq \omega_1 + \Delta\omega_1
\]

(26)

\[
\omega_2 \leq \omega \leq \omega_2 + \Delta\omega_2
\]

where we have defined \(K = \delta_1/\delta_2\) as the ratio of the passband to stopband error. For a given \(N\) and \(K\), these equations may be evaluated on a dense set of frequencies in the specified passband and stopbands and may be solved for \(\{h(n)\}\) and \(h_0\) using either linear programming techniques [11] or discrete Chebyshev approximation techniques [10]. It has been shown [10] that the filters designed by these techniques are optimum in the sense of having the narrowest transition bands for given \(\delta_1\) and \(\delta_2\). For these filters, the error in the passband and stopbands exhibits an equiripple behavior. An example of a low-pass filter designed by linear programming is shown in Fig. 11. In this case \(L = 5\), \(N = 29\) \((Q = 6)\), \(\omega_p = 0.6\pi/T\), \(\omega_m = 4\pi/T\), \(K = 1.0\). The resulting filter has \(\delta_1 = \delta_2 = 0.00586\). (The filter gain has been normalized to 0 dB for convenience in plotting.) The equiripple behavior of the stopband is readily apparent.

The filter in this example does not satisfy the constraints on \(h(n)\) given in (14); however, these constraints are linear and can be added to the constraints of (25) and (26). Thus in the optimization procedure \(h(0)\) is constrained to 1 and \(h(\pm L), h(\pm 2L), \text{etc.}\), are set to 0. The filter performance is only slightly degraded by the addition of these constraints. Fig. 11 shows an example where all the fixed-design parameters were the same as in the previous example. In this case the resulting value of \(\delta_1 = \delta_2 = 0.00599\). It can be seen from Fig. 12 that the equiripple nature of the frequency response is destroyed, but with little sacrifice in performance.

If the original signal bandwidth is much less than \(\pi/T\), a bandstop filter may provide superior performance to a comparable low-pass filter. Fig. 13 shows an example for \(L = 5\) where \(N = 29\), \(\omega_p = 0.5\pi/T\), \(\omega_m = 1.7\pi/T\), \(\omega_1 = 0.6\pi/T\), \(\omega_2 = 3.7\pi/T\), and \(K = 1.0\). In this case \(\delta_1 = \delta_2 = 0.0000681\). The filter does a much better job of attenuating the image signal spectrum around the frequencies \(2\pi/T\) and \(4\pi/T\), but only if the original signal bandwidth was much less than \(\pi/T\). It can be seen that the extra attenuation around \(2\pi/T\) and \(4\pi/T\) is obtained at the expense of the regions around...
Fig. 11. Frequency response of a typical low-pass interpolation filter designed by linear programming. \( L = 5 \), \( Q = 6 \), \( N = 29 \), \( \omega_0 = 0.6\pi/T \), \( \omega_0 = 1.4\pi/T \), \( \delta_1 = \delta_2 = 0.00586 \). (No impulse response constraints.)

Fig. 12. Frequency response for the example of Fig. 11, with impulse response constraints. In this case \( \delta_1 = \delta_2 = 0.00599 \).

Fig. 13. Frequency response of a typical bandstop interpolation filter designed by linear programming. \( L = 5 \), \( Q = 6 \), \( N = 29 \), \( \omega_p = 0.3\pi/T \), \( \omega_1 = 1.7\pi/T \), \( \omega_2 = 5.7\pi/T \), \( \Delta\omega_1 = \Delta\omega_2 = 0.6\pi/T \), and \( \delta_1 = \delta_2 = 0.0000681 \).

Fig. 14. Comparison of linear-phase Lagrange interpolators and optimum bandstop filters (\( L = 5 \)). Plots show minimum stopband attenuation as a function of half the stopband width. For the optimum bandstop filter (\( N = 29 \)), the curve is off scale for \( \omega_p < 0.3\pi/T \).

3\( \pi/T \) and 5\( \pi/T \) where the attenuation is much less. Thus optimum filters of this type (and classical filters) should only be used if one is certain of the bandwidth of the input sequence.

C. Comparison to Classical Interpolators

It is of considerable interest to compare the filters derived from classical interpolation formulas and those designed by linear programming. In the design of the optimum filters, the parameters of the tolerance scheme of Fig. 10 were set so that \( \delta_1 = \delta_2 \) with symmetrical stopbands of width \( 2\Delta\omega = \Delta\omega_1 = \Delta\omega_2 = 2\omega_p \) centered at integer multiples of \( 2\pi/T \). This is reasonable since in interpolation, preservation of the passband is generally as important as rejection of stopbands. In order to compare the Lagrange filters to the optimal filters, values of \( \delta_1 \) and \( \delta_2 \) were measured at the edges of the passband and the first stopband, respectively. Since for the Lagrange filters, \( \delta_1 \) is always greater than \( \delta_2 \) for the above definition of passband and stopbands, the comparison has been made on the basis of passband error.

Fig. 14 shows such a comparison for the case \( L = 5 \). The solid curves are for linear-phase filters derived from the Lagrange interpolation formula where \( N = 2 \), \( Q = 2 \), and 6.

These curves show the error in decibels at the edge of the passband, i.e., 20 \( \log_{10} \delta \), as a function of \( \omega_p \). As an example, for linear interpolation (\( Q = 2 \) and \( N = 9 \)) we see that the passband error is \(-42 \text{ dB} \) for \( \omega_p = 0.1\pi/T \). Furthermore, since \( \delta_1 > \delta_2 \) we can say that the attenuation is at least 42 dB in the region \( (2\pi/T - 0.1\pi/T) \leq \omega \leq (2\pi/T + 0.1\pi/T) \). For wider bandwidths, the performance is worse; however, for higher order interpolators the performance becomes appreciably better as is expected.

The dotted curves in Fig. 14 show passband error for bandstop filters designed by linear programming. By comparing corresponding curves, it can be seen that the optimum designs are always significantly better than the corresponding classical interpolator, with the improvement being most striking for narrow bandwidths and for the higher order filters.

Clearly, there are a variety of optimum designs corresponding to situations in which passband and stopband approximation errors are not treated as being of equal importance. For example if \( \delta_1 = K\delta_2 \), then the case \( K > 1 \) corresponds to placing more importance on stopband attenuation than on passband.

The comparisons are similar for larger values of \( L \).
error. This situation would be a more favorable situation for the classical filters, although it is always possible to design a better filter using the optimum design procedures.

Another interesting comparison is between optimum low-pass filters and optimum bandstop filters. To compare these filters, we set

\[ \omega_p = \frac{2\pi}{T} - \omega_1 = \frac{2\pi}{T} - \omega_2 \]

and \( \Delta \omega = \Delta \omega_1 = 2\omega_p \) for the bandstop filters and

\[ \omega_p = \frac{2\pi}{T} - \omega_1 \]

and \( \Delta \omega = \pi/T - \omega_2 \) for the low-pass filters. That is, both filters were designed to accommodate the same input signal bandwidth. Fig. 15 shows the difference between stopband attenuations for the bandstop filters and the low-pass filters as a function of bandwidth. From these curves we see that for narrow bandwidths the bandstop filters have significantly greater attenuation; however as the bandwidth approaches half the original sampling frequency, there is no difference between the two types of filters.

VI. CONCLUSIONS

In this paper we have discussed the process of interpolation as a problem in digital filtering. Most of our discussion has involved frequency-domain representations of the interpolation process and design criteria for digital interpolation filters. We have taken this approach because it is the most reasonable for digital signal processing applications where it is necessary to either raise or lower the sampling rate of a signal. This point of view is in contrast to that of interpolation in tables where one is concerned primarily with minimizing the error in a particular interpolated sample. Because of the variety of factors involved in the design of an interpolation filter, we have not tried to give design formulas and error bounds that would have limited value, but rather we have chosen to attempt to illuminate the important factors involved in the interpolation process and to discuss general design procedures that can be adapted to a variety of situations.

In particular, we have argued that linear-phase FIR filters have many attractive features for discrete-time interpolators and have shown how they may be efficiently utilized. Class polynomial interpolation has been discussed in the context of digital signal processing. Interpolation filters derived from polynomial interpolation formulas are attractive because their impulse response can be easily computed or looked up in a table. However, we have seen that the frequency response of such systems leaves much to be desired in digital signal processing applications where the original sampling rate may only slightly exceed the Nyquist rate.

As an alternative to filters based on classical interpolating formulas, we discussed optimum low-pass and bandstop FIR filters that were designed by linear programming. The bandstop filters have frequency responses that are very similar to the classical interpolators, but are always superior. The bandstop designs appear to be most important for cases where the original sampling rate is several times the Nyquist rate. The low-pass designs are appropriate when the original sampling rate is close to the Nyquist rate.

ACKNOWLEDGMENT

The authors would like to acknowledge the helpful comments and criticisms provided by Dr. O. Herrmann and Dr. J. F. Kaiser.

REFERENCES