Design and Implementation of Transmitter and Receiver Filters with Periodic Coefficient Nulls for Digital Systems

A. C. Salazar
V. B. Lawrence
Bell Laboratories
Holmdel, NJ 07733

ABSTRACT

The problem of designing nonrecursive digital filters to perform baseband and passband pulse shaping in digital systems amounts to determining a discrete sequence with preassigned zero crossings. This approach yields a filter frequency response which permits minimal stop band noise enhancement. Several methods exist for designing transmitter and receiver filters for digital systems. An interesting technique using a minimax criterion in the design procedure employs linear programming techniques in the frequency domain, in addition to ensuring periodic nulls in the time domain. We derive a bound which gives insight into the frequency resolution or frequency sampling grid density problem, and also provides an analytical basis for determining an estimate of the filter length. Problems encountered with efficient transmitter implementation are discussed along with methods found useful in overcoming these problems.

I. Introduction

With the advent of inexpensive digital processing components, an interest has developed [1]-[5] in implementing pulse shaping techniques digitally for linear modulation systems. The basic problem in pulse shaping can be seen in its simplest form with the equation:

\[ s(t) = \sum_{n=-\infty}^{\infty} a_n p(t-nT) \]  

(1)

The output or line signal, \( s(t) \), consists of a modulated pulse stream. The data or signaling information \( \{a_n\} \) simply pulses a filter, at frequency \( 1/T \), whose impulse response is \( p(t) \). The data sequence \( \{a_n\} \) is formed from a finite alphabet \( \{a_1, a_2, \ldots, a_L\} \), where \( L \) is usually an even integer. Examples of data sequences include encoded speech samples, computer data, and digital image processing signals. We will confine our investigation of the formation of (1) to be independent of any specific application's requirement.

The concept of intersymbol interference (ISI) is best seen pictorially in Figure 1. The pulse shape \( p(t) \) yields samples \( p(nT); n \neq 0 \) which represent the strength of interference pulses from symbols or data sent at other symbol intervals, and which affect the symbol at \( 0 \). It is clear that if \( p(nT) = 0 \) for \( n \neq 0 \), then zero ISI would result. The spectrum of a particular \( p(t) \) (canonical in some sense) which would yield zero ISI is shown in Figure 2a. It can be easily demonstrated [1] that any \( p(t) \) with the spectral property that \( \sum_{m=-\infty}^{\infty} |P(f)|^2 = 1 \), \( |f| \leq \frac{1}{2T} \) exhibits zero ISI, where \( P(f) \) is the spectrum of \( p(t) \). This property is most easily understood when it is a convolution of an even spectral function \( A(f) \)

\[ A(f) = A(-f) \]

(see Figure 2b), with \( P_e(f) \). Such a convolution preserves the property of \( p(nT) = 0 \), \( n \neq 0 \). That is,

\[ P(f) = P_e(f) * A(f) \]

yields a whole family of \( P(f) \) (using \( A(f) \) as a parameter spectrum) which exhibit \( p(nT) = 0 \), \( n \neq 0 \), \( n \) an integer.

Having determined the nature of the spectrum \( P(f) \) which yields ISI-free pulse shaping, it is possible to now consider the means for implementing the transmitter output \( s(t) \) of equation (1). In Figure 3 is shown the system diagram for the transmitter. The observation of the fact that \( \{a_n\} \) is a numerical sequence, i.e., its elements are realized from a finite number of numerical values \( \{a_L\} \), has led many researchers to propose transmitter implementation using digital signal processing techniques [5]. The idea here is to form a digital or sampled version of \( s(t) \) and then convert the samples to analog form via a digital to analog converter and lowpass analog filter. The analog filter is presumed to do no shaping and its only purpose is to eliminate the higher periodic spectra of the sampling system output. Should the effect of the analog filter be considered non-negligible it would then be considered in the design of the pulse shape \( p(t) \). It is important to note that the sampling frequency used for the digital filter implementation is assumed to be a harmonic of the signaling frequency. This requirement leads to a simpler design criteria and easier implementation without a need of complex interpolators.

A requirement implicit in every pulse shape design is that of bandwidth. We have seen that the frequency \( \frac{1}{2T} \) plays an important role in marking the point of symmetry in the spectral domain. However, as we show in Figure 2c the frequency point \( \frac{1}{2T} + \Delta \) marks the total bandwidth of the objective spectrum. The width \( 2\Delta \) Hertz represents a “transition band” or rolloff band. We refer to the ratio of \( \Delta/(1/2T) \) as the rolloff factor. The point to be remembered is that a design procedure should include consideration of a stop band past the point \( \frac{1}{2T} + \Delta \) Hertz and concentration of spectral shaping up to that same point. We begin our treatment of the design procedure by first reviewing a least squares procedure to provide contrast to the minimax design criterion.

II. The Minimax Design of Digital Pulse Shapers

A. Perspective: Least squares design

To gain a better perspective of how minimax design differs from least squares design of digital pulse shaping filters, let us review the latter procedure briefly. First, the digital spectral domain objective function is defined as \( D(f) \), \( |f| \leq 1/2T \) using normalized sampling frequency. The symbol signaling frequency is \( 1/T \) where \( T \) is an

\[*\] \( P(f) \) satisfies the Nyquist criterion if it produces a pulse which is ISI-free.
integer and $1/T$ is the subharmonic of 1, the sampling frequency. 
The filter transfer function is

$$H_N(f) = \sum_{n=-N}^{N} h_n \exp(\frac{j2\pi fn}{\Delta f})$$

(2)

where a delay factor of $e^{j2\pi f\Delta f}$ has been extracted and is removed from subsequent considerations. The symmetric coefficient property $h_n = h_{-n}$

(3)

leads to portraying $H_N(f)$ as a real function

$$H_N(f) = h_0 + 2 \sum_{n=1}^{N} h_n \cos n2\pi f\Delta f$$

(4)

and linear phase is evident in those bands where $H_N(f)$ does not change sign. To include the bandwidth constraint we form the following optimization problem:

$$\min \left\{ E_N \left( \frac{1}{2\Delta f} \right) \right\}$$

(5)

then seek

$$\max \left\{ \left| \int \left[ [W(f)]^2 \right] \right| \right\}$$

(6)

where $[W(f)]^2$ is a given weighting function. Without the constraints on the zero coefficients this formulation leads to the discrete form of the prolate spheroidal function for $H_N(f)$ as described in [6]. Adding the constraint of a spectral zero generates a minor modification of the problem. The zero-forced coefficient application to pulse shaping networks is given in [2]. The maximization of (6) represents an attempt to concentrate spectral energy in the band $\left[ -\frac{1}{2\Delta f}, \frac{1}{2\Delta f} \right]$ while the constraints of periodic zero coefficients keep ISI zero. The solutions given for (6) or similar expressions in references [1], [6] are typical least squares solutions of spectral optimization problems in that no sidelobe control is obtained and the in-band spectrum is not esthetically pleasing in that a sort of Gibbs phenomenon is observed to occur in the neighborhood of the frequency $1/2\Delta f$ as illustrated in Figure 5.

B. Minimax Design

In contrast to the least square design of Section IIIA let us now consider an alternative ISI-free design using a minimax criterion. The design problem can be stated in the simple form using the notation of (4):

$$E_N = \min \left\{ E_N \left( \frac{1}{2\Delta f} \right) \right\}$$

subject to the constraint

$$h_N = 0$$

(7a)

where $T$ is an integer and $J_i$ is defined to be the union of subintervals $J_i = \left[ 0, \frac{1}{2\Delta f} - \frac{1}{T} \right] \cup \left[ \frac{1}{2\Delta f} + \Delta \right]$, where $\Delta$ is a rolloff parameter and $D(f)$ is the desired spectral shape which for our purposes is 1.0 on $J_0$ (passband) and 0.0 on $J_t$ (stopband). The design problem posed in (7) is the basic problem discussed in reference [7] except for the zero coefficient constraints. The significance of the zero coefficients becomes apparent when one asks whether a unique solution to (7) even exists. We know that without these constraints $H_N(f)$ has been formed from a Haar set [8], p. 74 and thus a unique continuous solution exists to the problem stated in (7). Upon insertion of the zero coefficient constraints, the Haar condition no longer applies to $H_N(f)$ unless something is said about the character of $D(f)$. As it is shown in [10], if we constrain the problem posed in (7) so that $D(f)$ has the property

$$\sum \left[ D\left( f + \frac{1}{N} \right) \right] = 1$$

(8)

then a solution to (7) exists.*

A distinction between our formulation and that in reference [12] is the incorporation of the constraint given in equation 7a. This constraint guarantees that the impulse response has exact zeros at the signal sampling times (symbol rate). This will result in zero intersymbol interference. The formulation given in reference [12] is probable but not guaranteed to produce a filter that has an impulse response with exact zeros at the signal sampling times (i.e., symbol points).

1. Discretization

We can now consider the discrete form of (7) for which a unique solution always exists regardless of the condition imposed on $D(f)$. However, the uniqueness of the continuous form of the problem allows one to consider any discrete form of the problem in reference to a single limiting approximation. To see this better, let us consider the problem of (7) in discrete form:

Minimize $w$ subject to

$$D(f_i) - h_0 - 2 \sum_{n=1}^{N} h_n \cos n2\pi f_i \leq w$$

for $i = 1, 2, ..., N$. The set of frequencies $f_i$, $i = 1, 2, ..., N = F_D$ has been chosen from J and the obvious intent is to have $F_D$ "dense" in J. That is, no point in J is very far from some $f_k$ in $F_D$. Define

$$d(f_i, f_k) = \frac{1}{\pi} \left| \cos^2 f_i - \cos^2 f_k \right|$$

(10)

for two points $f_i$ and $f_k$ in $F_D$. The "density" $d_M$ of $F_D$ a discrete subset is defined as

$$d_M = \max_{\forall f} d_M(f, f_k)$$

(11)

Paraphrasing a theorem (18), p. 93 in approximation theory, we have that if $F_D$ is chosen such that

$$d_M \leq \lambda \theta^{-1}$$

(12)

when the truncated trigonometric series $H_N(f)$ of best approximation to $D(f)$ on $F_D$ converges uniformly to $D(f)$ on J as $M \rightarrow \infty$. As it is shown in [10] the following inequality can be derived for $E_N$ of equation (7):

$$E_N = \left[ \frac{24n^3}{\pi^2} \right] E_N$$

(13)

where $M = nN$.

The expression $(E_N)_{F_D}$ is that minimax error realized in the solution to (9). The utility of a bound like that of (13) comes from the fact that the gain in approximation accuracy has now been related to the number of points $M$ taken to form $F_D$. This relationship takes on a degree of importance when one can reduce, by use of this bound, the number of constraints in (9) to obtain the level of accuracy in the approximation problem. Previously, a rule of thumb was that of using $M$ equal to 8 to 10 times the number of approximating functions. It may be that in some cases such a number is high and computer time to solve (9) becomes excessive.

* We have been unable to prove uniqueness.
It is important to note that no formula for the \( M \) points has been
given although the bound in (13) holds for \( F_P \) chosen so that (10)
is used to define the distance metric. A distribution of points in \( F_P \), viz.
\[
f_k = \text{m} \{(2k-1)\pi/2M\} \quad k = 1, 2, \ldots, M
\]
achieves
\[
h_M = \pi/2M.
\]  
(15)
This distribution of points was used to derive (13).

2. Filter Splitting
Having acquired familiarity with the basic problem of (9) and
assuming for the moment that we have obtained a solution, let us
consider a step known as "filter splitting."

In this step we take the shaping filter response
\[
H(z^{-1}) = \sum_{n=-N}^{N} h_n z^{-n},
\]
(16)
and factor it into the form
\[
H(z^{-1}) = K \prod_{k=1}^{L_1} S_k(z^{-1}) \prod_{p=0}^{L_2} S_p(z) (z^{-1} + u_0 + z^-2)
\]
(17)
where \( S_k(z^{-1}) \) is a quadratic factor (or a single real root) inside the
unit circle, \( S_p(z) \) is the quadratic factor formed from inverting the
roots of \( S_0(z) \) and finally \( z^{-1} + u_0 + z^-2 \) represents a quadratic
factor on the unit circle, \( L_1 + L_2 = N \) if no real roots exist.
Specifically if
\[
S_0(z^{-1}) = a_0 z^{-2} + a_1 z^{-1} + 1
\]
then
\[
S_0(z) = a_0 + a_1 z + z^2
\]
We see then that the shaping filter transfer function can almost be
split into two mirror image filters and a possible discrepancy factor
coming about from roots on the unit circle. These unit circle roots are
not uncommon since they arise from the stopband specification.

Let us stop for the moment and review the reasons for even
considering filter splitting. Figure 4 illustrates the system implementation
for the digital signaling applications we discussed in
Section 1. The optimum partitioning of the shaping filter for the
case where additive white noise is the only transmission impairment
is that of putting the square root of the filter transfer function at the
transmitter and the square root at the receiver.\(^*\)

We have seen that \( H(z^{-1}) \) is not a "perfect square" since the roots
on the unit circle are not guaranteed to be double roots. A method
by which these roots can be forced to be double is easily
implemented in the problem of (9) by augmenting another block of
constraints of the form
\[
-H(f_k) \leq 0 \quad f_k \in J_k
\]
The solution then would yield a factorization
\[
H(z^{-1}) = K \prod_{k=1}^{L_1} S_k(z^{-1}) \prod_{p=0}^{L_2} S_p(z)
\]
(18)
where
\[
P_1(z^{-1}) = \prod_{k=1}^{L_1} S_k(z^{-1})
\]
is a minimum delay function and
\[
P_2(z) = \prod_{p=0}^{L_2} S_p(z)
\]
is a maximum delay function. The significance of
\(^*\) In this way a "matched" filter pair is provided.

is that \( p_0 \) is the largest coefficient in absolute magnitude of all \( p_i \). It
should be pointed out that the factorization of which quadratic
factor of (18) goes into the transmitter transfer function and which
 goes into the receiver does not have to be along the lines of (19) and
(20). For example, we could mix quadratic factors from the
\( S_k(z^{-1}) \) group with those from the \( S_p(z) \) class. We know that the
counterpart quadratic factor has the same magnitude function on
the unit circle. For every bipartite sectioning of \( H(z^{-1}) \) we end up
with the same magnitude function for both the transmitter and
receiver but different phase functions.

III. Implementation
We will now review several aspects of implementation which
significantly affect the procedure used in the design of shaping filters.
Before we begin, let us mention three important features of the
system diagram of Figure 3.

1. The pulse shape parameters \( \{h_n\}^N_{n=0} \) are symmetric, i.e.,
\[
h_k = h_{-k} \quad k = 1, 2, \ldots, N.
\]
The reason for the symmetry stems from the fact that linear phase is obtained from this type of
filter in the passband region. The filter output can be written
\[
P_m = \sum_{k=-N}^{N} h_k a_{m-k} = h_0 + \sum_{k=1}^{N} h_k (a_{m-k} + a_{m+k})
\]
(21)
Observation of (21) tells us that only \( N + 1 \) multiplications
are required for computing \( P_m \). If shift scaling is used to
normalize the filter coefficients, then only \( N \) multiplications
are required. Minimizing the number of multiplications has
the effect of minimizing roundoff noise error sources in a
digital signal processor implementation of this type of
filter.

2. The pulse shape parameters \( \{h_n\}^N_{n=0} \) have periodic nulls, i.e.,
\[
h_{M+k} = 0; \quad M \text{ an integer} \quad k = 1, 2, \ldots,
\]
Nyquist shaping is desired for the filter in order to prevent
intersymbol interference. This means that every \( M^{th} \)
coefficient of the filter is identically zero. Conversely, if \( M^{th} \)
nulls are observed in the coefficients, then the filter has
Nyquist shaping according to the baud rate.

3. The input symbols \( \{a_m\} \) into the filters arriving at the baud
rate are interpolated into the sampling rate. Just as it occurs in
an interpolating environment \( [9] \) the filter \( \{h_n\}^N_{n=0} \) receives an
input once every baud interval. Hence, if we use an \( M \) (or
even) baud shaping filter we need only consider a maximum
\( L/2 + 1 \) multiplications to form the output because of the
interpolation process.

A. Filter Length Bounds
It is often useful to determine the approximate length of the
shaping filter required for achieving a given stopband loss. In
Figure 6 is shown the typical design parameters for the minimax
design of a shaping filter. We note that the transition region of
width \( 2A \) of the filter can be generated with the aid of the
parameter spectrum \( A(f) \) as we have seen in Section I. A
particular example of \( A(f) \) is
\[ A(f) = \begin{cases} \frac{1 + \cos \frac{x}{\Delta}}{2\Delta}, & |x| \leq \Delta \\ 0 & \text{otherwise} \end{cases} \]

Defining the total shaping spectrum as

\[ P(f) = P_r(f) * A(f) \]

as in Section I we have

\[ P(x) = 1, \quad x < \frac{x}{M} - \Delta x \]

\[ P(x) = -\frac{1}{2\Delta} \left[ (x-M-\Delta x) \cos \frac{x}{M} + \Delta \sin \left( x-\frac{x}{M} / \Delta \right) \right] \]

\[ \Delta x / M - \Delta x \leq x \leq \Delta x / M + \Delta x \]

\[ P(x) = 0, \quad x > \Delta x / M + \Delta x \]

We have then that the derivatives of \( P(x) \) take on their largest absolute values in the transition region. Namely, it is simple to calculate

\[ |P'(x)| \leq \frac{1}{\Delta x} \]

and

\[ |P''(x)| \leq \frac{1}{2\Delta^2} \]

Using Jackson's theorem (18), p. 145 we have that

\[ E_n(x) \leq n/2 \left( \frac{1}{N+1} \right)^2 \frac{1}{2\Delta^3} \]

or

\[ N \leq \frac{1}{2\Delta \sqrt{E_n(x)}} - 1 \quad \text{filter length} = 2N + 1 \]

Inequality (27) relates the length of the filter to the transition width parameter \( \Delta \) and the desired stopband ripple, \( E_n \). The reader should be aware of a good bound based upon empirical data by J. F. Kaiser which can be used to give a reasonable estimate of filter length. Our bound is an analytically derived bound applicable only to filters generated using the parameter spectrum \( A(f) \) defined in equation 25.

B. Filter Design on Time-Share System

The filter design algorithm has been programmed and tested on many designs. The program input consists of setting bits per symbol, number of symbols in filter length and rolloff factor. Table 1 lists the results of a filter design with the accompanying spectra. The spectra points shown is that of a four bits per symbol design with a 12.5 percent rolloff factor and an eight symbol length filter. This minimax spectra points can be compared to that of a least squares design described in Section II-A. The latter design also uses four bits per symbol, is of eight symbol length and has a 12.5 percent rolloff factor. The least squares design in this case has a larger passband ripple (1.2 dB compared to 0.79 dB) and a larger stopband ripple (16.3 dB loss compared to 21.37 dB). Characteristic of least squares designs, these large ripples occur near the discontinuity of the frequency point.

IV. Summary and Conclusions

We have treated the minimax design of digital pulse shaping filters. This design problem leads to basic mathematical questions as to existence of solution and sampling grid density. In the first case we probe existence of solution but show that uniqueness is still an open question. We have found bounds which give insight into the sampling grid density problems. In addition, an analytical basis has been provided for determining an estimate to the filter length as a function of stopband loss and transition width.

Implementation considerations including dividing the filter into the transmitter and receiver have also been treated. Design examples have been provided, and a comparison was made between a least squares design and a minimax shaping filter design which shows the overall smaller ripple advantage in the spectrum designed with the minimax criterion.

REFERENCES


\[ \text{Figure 1: Intersymbol Interference Samples of Pulse} \]

\[ P(n) \text{ at } n = \pm 1, \pm 2, \pm 3, \ldots \]
**Figure 2:** Parameterized Formation of Shaping Spectrum with Rolloff

\[ f(t) \rightarrow \text{Pulse Shaping Filter} \rightarrow s(t) \rightarrow D/A \rightarrow s'(t) \]

**Symbol Rate**: \( 1/T \)

\[ s(t) = \sum_{n=-\infty}^{\infty} a_n p(t-nT) \]

**Figure 3:** Baseband Transmitter

**System Noise**

**Figure 4:** Digital Signaling System Model

**Passband Ripple Height \( b_1 \)**

**Stopband Ripple Height \( b_2 \)**

**Figure 5:** Amplitude Spectrum of Baseband Shaping Filter (Least Squares Design)

**Figure 6:** Amplitude Spectrum of Baseband Shaping Filter (Minimax Design)

<table>
<thead>
<tr>
<th>Frequency in Hz</th>
<th>Amplitude in dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>-0.7</td>
</tr>
<tr>
<td>450</td>
<td>0.1</td>
</tr>
<tr>
<td>850</td>
<td>0.6</td>
</tr>
<tr>
<td>1050</td>
<td>0.8</td>
</tr>
<tr>
<td>1200</td>
<td>-6.2</td>
</tr>
<tr>
<td>1350</td>
<td>-21.4</td>
</tr>
<tr>
<td>1650</td>
<td>-22.7</td>
</tr>
<tr>
<td>2050</td>
<td>-47.0</td>
</tr>
<tr>
<td>2450</td>
<td>-26.2</td>
</tr>
<tr>
<td>2850</td>
<td>-60.2</td>
</tr>
<tr>
<td>3250</td>
<td>-68.2</td>
</tr>
<tr>
<td>3650</td>
<td>-22.6</td>
</tr>
<tr>
<td>4050</td>
<td>-32.4</td>
</tr>
<tr>
<td>4450</td>
<td>-50.4</td>
</tr>
<tr>
<td>4850</td>
<td>-27.4</td>
</tr>
</tbody>
</table>

**Stopband Ripple**: -21.36 dB

**Passband Ripple**: 0.8 dB

**Table 1**

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Minimax</th>
<th>Least Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.27446</td>
<td>0.25008</td>
<td></td>
</tr>
<tr>
<td>0.21583</td>
<td>0.22442</td>
<td></td>
</tr>
<tr>
<td>0.14773</td>
<td>0.15714</td>
<td></td>
</tr>
<tr>
<td>0.07308</td>
<td>0.07288</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>-0.03965</td>
<td>-0.04149</td>
<td></td>
</tr>
<tr>
<td>-0.04811</td>
<td>-0.04714</td>
<td></td>
</tr>
<tr>
<td>-0.03122</td>
<td>-0.02736</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>0.01933</td>
<td>0.91909</td>
<td></td>
</tr>
<tr>
<td>0.02662</td>
<td>0.02275</td>
<td></td>
</tr>
<tr>
<td>0.01888</td>
<td>0.01358</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>-0.01093</td>
<td>-0.00966</td>
<td></td>
</tr>
<tr>
<td>-0.02333</td>
<td>-0.01148</td>
<td></td>
</tr>
<tr>
<td>-0.02810</td>
<td>-0.00679</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1**