Signal Processing 18 (1989) 327–339
Elsevier Science Publishers B.V.

MINIMAX DESIGN OF FACTORABLE NYQUIST FILTERS FOR DATA TRANSMISSION SYSTEMS*

Ravi P. RAMACHANDRAN and Peter KABAL**

Department of Electrical Engineering, McGill University, 3480 University, Montreal, Quebec, Canada H3A 2A7

Received 23 November 1988
Revised 8 March 1989 and 2 June 1989

Abstract. The use of Nyquist filters in data transmission systems is important in avoiding intersymbol interference. Moreover, the Nyquist filters should be factorable into lowpass transmitter/received filter pairs. Here, the design problem is formulated so as to generate zero-phase FIR lowpass Nyquist filters that can be split into minimum and maximum phase parts. Two factorable minimax design methods are given. These methods use the McClellan–Parks algorithm as a first step to control the stopband behaviour. The time domain constraints, imposed by solving a linear system of equations, determine the passband response. The final filter exhibits equiripple stopband behaviour. The advantages of these methods are that the minimum and maximum phase parts are obtained without direct factorization and that arbitrary frequency weighting can be easily incorporated to allow for a nonequiripple behaviour. Design examples depict both equiripple and nonequiripple magnitude responses. The new design approach is compared with other methods in terms of both magnitude and group delay behaviour. Finally, a practical design that conforms to a CCITT voice band modem specification is shown.

Zusammenfassung. Die Verwendung von Nyquist-Filtern ist für die Datenübertragung wichtig, wenn Intersymbol-Interferenzen vermieden werden sollen. Darüberhinaus wünscht man sich, daß die Nyquist-Filter in Tiefpaß-Filterpaare für die Sende und Empfangsseite zerlegt werden können. Im Folgenden wird das Entwurfsproblem so formuliert, daß nullphasige FIR-Nyquist-Tiefpässe so bestimmt werden, daß sie in einen minimalphasigen und einen maximalphasigen Teil aufgeteilt werden können. Zwei faktorisierbare Minimaxentwürfe werden angegeben. Die Methoden verwenden den McClellan–Parks-Algorithmus als ersten Schritt, um das Sperrverhalten zu bestimmen. Das Durchlaßverhalten wird durch Zeitbereichsvorschriften bestimmt, die in Form eines linearen Gleichungssystems eingebracht werden. Das fertige Filter zeigt Tschebyscheff-Verhalten im Sperrbereich. Die Vorzüge dieser Methoden bestehen darin, daß die minimal- und maximalphasigen Teilfilter ohne explizite Faktorisierung erhalten werden und daß eine beliebige Frequenzgewichtung auf einfache Weise eingefügt werden kann, mit der auch eine nicht-gleichmäßige Approximation erzielt werden kann. Entwurfsbeispiele sowohl mit als auch ohne Tschebyscheff-Verhalten der Dämpfung werden gezeigt. Der neue Entwurfsansatz wird mit anderen Methoden hinsichtlich des Dämpfungs- wie des Gruppenlaufzeit-Verhaltens verglichen. Zum Schluß wird ein praktisches Entwurfsbeispiel gezeigt, das die Anforderungen für ein CCITT-Sprachband-Modem erfüllt.

Résumé. Il est important d'utiliser des filtres de Nyquist dans les systèmes de transmission numériques en vue d'éviter l'interférence entre symboles. De plus, la factorisation des filtres de Nyquist en paires de filtres de transmission et de réception devra être possible. Le problème de conception de ces filtres est formulé de façon à générer des filtres pass-bas RIF à phase zéro qui peuvent être séparés en une section à phase minimale et une section à phase maximal. Deux méthodes de conception à factorisation minimax sont présentées. Dans un premier temps, ces méthodes font usage de l'algorithme de McClellan–Parks pour contrôler la réponse dans la bande d'arrêt. Les contraintes sur la réponse temporelle, imposées par la solution d'un système d'équations, déterminent la réponse dans la bande passant. Le filtre résultant possède une bande d'arrêt possédant des ondulations uniformes. Les avantages de ces méthodes sont, d'un part, que les sections à phase minimale et à phase maximale sont obtenues sans factorisation directe, et d'autre part, qu'une pondération fréquentielle arbitaire peut être façilement incorporée pour permettre des ondulations non uniformes. Des examples montrent des réponses en amplitude dans les cas uniformes et non uniformes. La nouvelle méthode de conception est comparée à d'autres méthodes en termes de réponse en emplitude et de délai de groupe. Enfin, un exemple pratique de conception de filtre se conformant aux spécifications du CCITT quant à la conception d'un modem à bande vocale est présenté.

Keywords. Minimax filter design, Nyquist filters.

* This work was supported by the Natural Sciences and Engineering Research Council of Canada.

** Also at INRS-Télécommunications, Université du Québec, 3 Place du Commerce, Verdun, Quebec, Canada H3E 1H6.

1. Introduction

Intersymbol interference occurs when a received data symbol is influenced by a combination of several transmitted data symbols. Intersymbol interference is avoided through the use of Nyquist filters. Nyquist filters have an impulse response with regular zero crossings. These zero crossings result in the condition that the output of the data receiver taken at symbol intervals depends *only* on its corresponding transmitted symbol. Furthermore, the filters for bandwidth efficient data transmission systems are based on lowpass prototypes. The design problem incorporates both time and frequency domain constraints.

This paper focuses on the design of zero-phase lowpass FIR Nyquist filters. The Nyquist filters are split into a transmitter/receiver filter pair. It is desirable that the transmitter and received filters have identical lowpass magnitude responses. One approach is to use the same filter at the transmitter and receiver. However, this strategy leads to nonlinear constraints on the filter coefficients. In fact, only a trivial solution can be obtained for the case in which the regular zero crossings occur for every second sample. In this paper, we proceed to split the Nyquist filter into a pair having identical magnitude responses but allow different phase responses. Specifically, the filter is split into minimum and maximum phase parts. The Nyquist filter designed by our approach is fundamentally equiripple. A nonequiripple filter can be obtained by applying an additional frequency weighting factor.

In the past, FIR Nyquist filters have been designed using linear programming techniques [7, 13], by the eigenfilter approach [11, 15], and by the use of the McClellan–Parks algorithm [8] as an intermediate step [9, 10, 14, 16]. The methods in [10, 11, 13] allow for the splitting of the filter into its minimum and maximum phase parts. Salazar and Lawrence [13] set up the design as a linear programming problem incorporating the time domain constraints. In addition, the frequency response of the filter is forced to be nonnegative in order that the minimum and maximum phase factorization be possible. Mintzer [10] deals exclusively with the case when the zero crossings occur for every second sample. In that paper, the frequency response of an unconstrained filter is offset to ensure that it becomes nonnegative. In [11], the eigenfilter concept is applied to obtain a Nyquist filter that is factorable into minimum and maximum phase parts.

Nyquist filters with Chebyshev stopband behaviour have been designed in [14] using a multistage structure. The focus in [14] is on a computationally efficient multistage implementation. However, the resulting filters are not necessarily factorable. One can make these filters factorable by adding a positive constant to the frequency response [4, 10] to make it nonnegative. However, this fixup excessively reduces the stopband attenuation for nonequiripple filters.

This paper proposes two design methods which use the McClellan–Parks algorithm as a first step to control the stopband response and achieve a Chebyshev stopband behaviour. The subsequent step incorporates the time domain constraints and automatically generates the passband response. A few iterations of the above steps produces a factorable Nyquist filter with Chebyshev stopband response. We refer to the proposed approaches as factorable minimax design methods.

In the factorable minimax methods, we directly achieve a nonnegative frequency response with controlled stopband characteristics. Furthermore, the polynomial factorization problem for the determination of the minimum phase part is considerably eased. The complexity of polynomial factorization is directly related to the order and hence, to the length of the designed Nyquist filter. We reduce this complexity by determining a partial factorization of the transfer function of the Nyquist filter as a byproduct of the design procedure. The remaining factorization involves a polynomial of much lower order than the overall transfer function.

The paper first discusses the concept of factorable Nyquist filters. Then, the two factorable minimax design procedures are described in Section 3. The factorization problem is examined in Section 4. A discussion of the methods in this paper along with design examples are given in Section 5. Comparisons with other design approaches are presented in Section 6.

2. Factorable Nyquist filters

Zero-phase FIR Nyquist filters $H(z)$ have the impulse response characteristic

$$h(iN)=\begin{cases}\dfrac{1}{N}, & \text{for } i=0,\\ 0, & \text{for } i\neq 0.\end{cases} \tag{1}$$

The parameter N is the zero crossing interval in the time response $h(n)$. In the frequency domain, this corresponds to

$$\sum_{i=0}^{N-1} H(e^{j(\omega-2\pi i/N)})=1. \tag{2}$$

The minimum bandwidth solution is an ideal lowpass filter bandlimited to π/N. For practical filters, we allow an excess bandwidth $\beta\pi/N$ to bring the overall bandwidth to $(1+\beta)\pi/N$. The parameter β is the roll-off factor of $|H(e^{j\omega})|$. Furthermore, bandwidth efficient systems use less than 100 percent excess bandwidth thereby imposing $\beta\leqslant 1$.[1] In this case, only adjacent replicas of the spectrum of $H(e^{j\omega})$ (located at center frequencies that are multiples of $2\pi/N$) overlap. Also, the upper edge of the passband is $\omega_p=(1-\beta)\pi/N$ and the lower edge of the stopband is $\omega_s=(1+\beta)\pi/N$. The ideal frequency characteristic is

$$|H(e^{j\omega})|=\begin{cases}1, & \text{for } 0\leqslant|\omega|\leqslant\omega_p \text{ Passband},\\ 0, & \text{for } \omega_s\leqslant|\omega|\leqslant\pi \text{ Stopband}.\end{cases} \tag{3}$$

The response of an ideal filter makes a symmetrical transition from the passband to the stopband passing through the value 0.5 at $\omega=\pi/N$. The factorable minimax design methods generate filters that approximate this ideal magnitude characteristic. First, the approximation is made in the stopband region. Then, forcing zero crossings in the impulse response leads to a spectrum that exactly satisfies (2). With a response satisfying (2), an approximately zero stopband assures an approximately constant passband (assuming $\beta<1$). In addition, the transition band will be inherently symmetrical.

The fact that $H(z)$ is zero-phase means that the time domain response is symmetric ($h(n)=h(-n)$). In addition, real axis zeros occur in pairs at $z=z_0$ and z_0^{-1}. Unit circle zeros occur in complex conjugate pairs. The general complex zeros of $H(z)$ occur in groups of four at $z=z_0$, z_0^*, z_0^{-1} and $(z_0^{-1})^*$. For $H(z)$ to be factorable into minimum and maximum phase parts, the additional constraint is that all of its zeros on the unit circle must occur as double zeros. The overall design problem includes the time and frequency domain approximations along with the additional requirement of double order unit circle zeros. The next section describes the design procedures.

3. Minimax factorable design procedures

The Nyquist filter $H(z)$ must have an odd number of coefficients in order to be factorable into minimum and maximum phase parts. As in [11], we factorize $H(z)$ as $H(z)=H_0(z)H_1^2(z)$ where $H_1^2(z)$ contains all the double zeros of $H(z)$ on the unit circle and $H_0(z)$ contains the other zeros of $H(z)$. The double zeros of $H_1^2(z)$ on the unit circle imply that it has an odd number of coefficients and that it is a zero-phase function. The zeros of $H_0(z)$ must occur in mirror image pairs reflected about the unit circle. Hence, $H_0(z)$ also has an odd number of coefficients and is a zero-phase function.

Let the lengths of $H_0(z)$ and $H_1^2(z)$ be $2l_0+1$ and $2l_1+1$ respectively. The number of coefficients of the overall Nyquist filter $H(z)$ is $M=$

[1] Note that for $N>2, \beta\leqslant 1$ is not a requirement for the design.

$2(l_0+l_1)+1$.[2] The inverse z-transforms of $H(z)$, $H_0(z)$ and $H_1^2(z)$ are defined to be $h(n)$, $h_0(n)$ and $f(n)$, respectively.

3.1. First method

The design procedure for the first method is as follows:
(1) Initialization: Fix l_0, l_1, N, and ω_s. Set $H_0(z)=1$. The weighting is given as $W(e^{j\omega})$.
(2) Design $H_1(z)$ using the McClellan-Parks algorithm such that it has zeros only on the unit circle in the stopband region $[\omega_s, \pi]$.
(3) Impose the time domain constraints by solving for the coefficients of $H_0(z)$ through a linear system of equations.
(4) Form the Nyquist filter $H(z)$. If the design warrants improvement, go back to Step 2.
(5) Split $H(z)$ into its minimum and maximum phase parts.

We now describe Steps 2-5 in more detail.

3.1.1. Step 2: Frequency domain specifications

The McClellan-Parks algorithm is used to get the coefficients of $H_1(z)$. The specifications are that the frequency response must be one at $\omega=0$ and must approximate zero in the stopband region $[\omega_s, \pi]$. The weighting function applies to $H_1^2(z)$. The weighting function is $W(e^{j\omega})|H_0(e^{j\omega})|$. Initially, it is $W(e^{j\omega})$ since $H_0(z)=1$. Subsequent iterations involve an update of the weighting function as $H_0(z)$ is recomputed. For the design of $H_1(z)$, tabulated values of the square root of the weighting function are inputs to the algorithm.

In the stopband, the frequency resoonse of $H_1(z)$ exhibits a ripple-like behaviour with local minima and maxima occurring at the extremal frequencies. If l_1 is even, $H_1(z)$ has an odd number of coefficients (l_1+1). Two of the extremal frequencies are 0 and π [12]. However, the total number of zeros is a multiple of two, all occurring in complex conjugate pairs (no zero at $z=-1$). At $\omega=\pi$, either a local maximum or a local minimum occurs. If l_1 is odd, $H_1(z)$ has an even number of coefficients. In this case, a zero occurs at $z=-1$. However, π is not an extremal frequency. The other zeros occur in complex conjugate pairs bringing the total number of zeros to l_1.

3.1.2. Step 3: Time domain constraints

Given $H_1(z)$, we form $H_1^2(z)$ and solve for the coefficients of $H_0(z)$ such that $H(z)$ has the Nyquist property. Since $h(n)$ has samples for $n=-(l_0+l_1)$ to l_0+l_1, the number of zero-valued samples that occur as n goes from 1 to l_0+l_1 is $\lfloor (l_0+l_1)/N \rfloor$. The same holds true as n goes from -1 to $-(l_0+l_1)$. Since, the sample for $n=0$ is also known, the number of known coefficients of $H(z)$ is[3]

$$L=2\left\lfloor \frac{l_0+l_1}{N} \right\rfloor+1. \tag{4}$$

The coefficients of $H(z)$ are found by performing the convolution $h_0(n)*f(n)$. By expanding the convolution sum, one can uniquely determine $H_0(z)$ such that the time domain constraints are satisfied [11] if the number of unknown coefficients of $H_0(z)$ equals the number of known coefficients of $H(z)$. This results in a system of linear equations of dimension $2l_0+1$. By further exploiting the time domain symmetry of each filter, the problem is reduced to that of a system of dimension l_0+1. The system of equations can be expressed as $\boldsymbol{Fh}=\boldsymbol{c}$ where $\boldsymbol{h}^{\mathrm{T}}=[h_0(0),\ldots,h_0(l_0)]$, $\boldsymbol{c}^{\mathrm{T}}=[1/N, 0, \ldots, 0]$ and

$$\boldsymbol{F}=\begin{bmatrix} f(0) & 2f(1) & \cdots & 2f(l_0) \\ f(N) & f(N-1)+f(N+1) & \cdots & f(N-l_0)+f(N+l_0) \\ \vdots & \vdots & & \vdots \\ f(Nl_0) & f(Nl_0-1)+f(Nl_0+1) & \cdots & f(Nl_0-l_0)+f(Nl_0+l_0) \end{bmatrix}. \tag{5}$$

[2] The case $l_0+l_1=kN$ for any integer k renders a Nyquist filter with $h(-l_0-l_1)=h(l_0+l_1)=0$ thereby reducing the effective length by two.

[3] This formula is a corrected version of the formula given in [11].

The constraint that $L=2l_0+1$ is equivalent to $l_0=\lfloor(l_0+l_1)/N\rfloor$ which in turn translates to constraints on l_0 and l_1 given by

$$l_0(N-1)\leq l_1<l_0(N-1)+N. \quad (6)$$

Appendix A gives the derivation of closed form expressions for l_0 and l_1 in terms of N and M,

$$l_0=\left\lfloor\frac{M-1}{2N}\right\rfloor, \qquad l_1=\frac{M-1}{2}-\left\lfloor\frac{M-1}{2N}\right\rfloor. \quad (7)$$

This method of satisfying the Nyquist property automatically takes care of the passband response of $H(z)$. Note that $H_0(z)$ is a highpass function that primarily controls the passband characteristic and hence has no zeros on the unit circle.

3.1.3. Step 4: Convergence

The coefficients of $H(z)$ are found from $H_0(z)$ and $H_1^2(z)$. Steps 2 and 3 are iterated if the design warrants improvement. For Step 2, the weighting function $W(e^{j\omega})|H_0(e^{j\omega})|$ is updated to include a new $|H_0(e^{j\omega})|$ calculated from the coefficients of $H_0(z)$ formed in Step 3 of the previous iteration. The application of this weighting factor significantly influences the stopband behaviour of $H(z)$ through the design of $H_1(z)$. In the weighting function, the factor $|H_0(e^{j\omega})|$ leads to a stopband behaviour of $H_1^2(z)$ that compensates for the highpass response of $H_0(z)$. The stopband behaviour of $H(z)$ is either equiripple or nonequiripple depending on the other factor $W(e^{j\omega})$ in the weighting function. The iterations are terminated when the extremal frequencies obtained by designing $H_1(z)$ do not change by more than a given threshold.

3.1.4. Step 5: Final filter

This step factors $H(z)$ into minimum and maximum phase parts. Let the minimum phase part of $H(z)$ be $H^-(z)=H_0^-(z)H_1(z)$ where $H_0^-(z)$ is the minimum phase part of $H_0(z)$. The factor $H_1(z)$ is known as a byproduct of the design procedure. Only $H_0(z)$ needs to be factored in order to derive its minimum phase part. The maximum phase part of $H(z)$ is obtained by time reversing the coefficients of the minimum phase filter.

3.2. Second method

The difference between the second method and the previous approach lies in Step 2 in which a constrained form of the McClellan-Parks algorithm is used to directly compute the coefficients of $H_1^2(z)$ rather than to first design $H_1(z)$. The specifications are that the frequency response must be one at $\omega=0$ and must approximate zero in the stopband region $[\omega_s,\pi]$. As before, the weighting function is $W(e^{j\omega})|H_0(e^{j\omega})|$. Tabulated values of the weighting function are supplied as inputs. Since double zeros on the unit circle are required, we constrain the frequency response to be nonnegative in the stopband region. We implement the procedure in [3] (see also [6]) to obtain a minimax approximation to a desired response that satisfies given upper and lower constraints.

In the stopband, the frequency response of $H_1^2(z)$ exhibits a ripple-like behaviour with local minima and maxima occurring at the extremal frequencies. The local minima correspond to the frequencies at which the response touches zero. It is these frequencies which determine the double zeros of $H_1^2(z)$ on the unit circle. Given that $H_1^2(z)$ has $2l_1+1$ coefficients, a total of l_1+1 extremal frequencies result [12]. Two of the extremal frequencies are 0 and π regardless of the value of l_1. If l_1 is odd, the extremum at π is a local minimum thereby producing a double zero at $z=-1$. The other zeros occur in groups of four in the stopband region bringing the total number of zeros to $2l_1$. If l_1 is even, the extremum at π is a local maximum (no zero at $z=-1$). The total number of zeros is a multiple of four and occur in groups of four in the stopband region.

Steps 3 and 4 are identical to the first approach. In splitting $H(z)$ into its minimum and maximum phase parts, we take advantage of the fact that the frequencies corresponding to the double zeros of $H_1^2(z)$ are available as a byproduct of the modified McClellan-Parks algorithm (similar to the approach used in [6] to generate minimum phase filters). Given these frequencies and hence, the locations of the zeros on the unit circle, $H_1(z)$ can be formed without directly factoring $H_1^2(z)$. As before, only $H_0(z)$ must be factored to form $H^-(z) = H_0^-(z)H_1(z)$.

The next section discusses the merits of factoring only $H_0(z)$ as opposed to $H(z)$ in determining the minimum phase part. Also, observations concerning the relative orders of $H_0(z)$ and $H(z)$ are given.

4. The factorization problem

Polynomial factorization can be an ill-conditioned problem [17]. There is an advantage to substantially lowering the order of the polynomial to be factored. A general zero plot of $H(z)$ includes stopband zeros, passband zeros and extra real zeros [1]. The double order stopband zeros on the unit circle contribute to the stopband ripples and the passband zeros that occur in mirror-image pairs reflected about the unit circle contribute to the passband response. The extra real zeros insert a compensating spectral tilt. If $H(z)$ were to be factored, the double zeros on the unit circle and the other zeros would be determined through one factorization procedure. Note that finding the double zeros can be an ill-conditioned problem [17]. Furthermore, the use of polynomial deflation can be troublesome since the zeros of the resulting polynomial may in some cases diverge from those of the original polynomial [17]. In our approach, both factorization and deflation of $H(z)$ are avoided. In particular, the knowledge of $H_1(z)$ ensures that any errors that would normally occur in locating the unit circle zeros are absent and do not affect the zeros of $H_0(z)$. Furthermore, the factorization of $H_0(z)$ does not involve multiple zeros since $H_0(z)$ has only the simple passband and extra zeros of $H(z)$.

Since only the zeros of $H_0(z)$ have to be determined, the extent to which the factorization problem is eased depends on the ratio l_1/l_0. The ratio l_1/l_0 is both a measure of the proportion of unit circle zeros to the other zeros of $H(z)$ and of the degrees of $H(z)$ and $H_0(z)$. The higher the value of l_1/l_0, the lower the relative orders of $H_0(z)$ and $H(z)$. Appendix B shows that l_1 is greater than l_0 by a factor of at least $N-1$. Therefore, the inherent advantage in terms of polynomial factorization increases as N increases. However, even for the lowest value, $N=2$, the degree of $H(z)$ is at least twice the degree of $H_0(z)$. Note that the lower bound for $l_1/l_0 = N-1$ is satisfied when the end points of the impulse response are zero-valued (shown in Appendix B). We discard this artificial case because the values of l_0 and M can be reduced by 1 and 2 respectively thereby giving a new value of l_1/l_0.

A typical designed Nyquist response $h(n)$ is depicted in Fig. 1. The time response consists of a main lobe between $n=-N$ and $n=N$ and a series of sidelobes each occurring between the zero crossings. The value of l_0 is a measure of the number of sidelobes. As the number of coefficients M increases, l_1 also increases. For a fixed number of lobes (constant value of l_0), increasing M results in a higher stopband attenuation while maintaining the same factorization complexity. Hence, for a fixed number of lobes, one can maximize l_1/l_0 by increasing M. The largest disparity in the relative orders of $H_0(z)$ and $H(z)$ results by choosing the filter lengths to be of the form $M=2kN-1$.

Given that the filter lengths are constrained to be of the form $M=2kN-1$, the ratio l_1/l_0 is

$$\frac{l_1}{l_0} = \frac{k(N-1)}{k-1}. \tag{8}$$

This ratio is a maximum for $k=2$.[4] As k increases, a tradeoff results in that a higher stopband attenuation due to a longer filter is obtained at the expense of both a lower l_1/l_0 and a higher l_0. The subsequent examples show that a value of $k=5$

[4] This is also a unique maximum for a general M (see Appendix B).

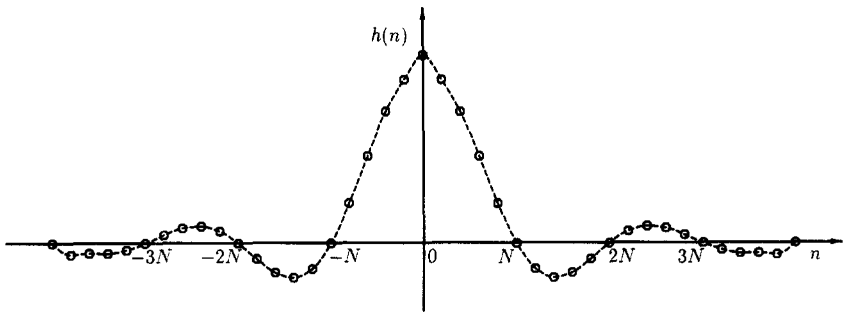


Fig. 1. Typical Nyquist response $h(n)$ (shown for $N=5$, $M=39$ and $\beta=0.2$).

results in about an 80 dB stopband attenuation for a roll-off factor of 0.52. Then, $l_1/l_0=5(N-1)/4$ and $l_0=4$. Only an eighth order polynomial with simple zeros needs to be factored. Smaller roll-off factors require a larger number of taps (larger value of k) and hence, a lower value of l_1/l_0 and a higher value of l_0 for an 80 dB stopband attenuation.

5. Discussion of the design techniques

Given that two factorable minimax design methods are proposed, their relative merits are considered. Design examples are given. Finally, the group delay behaviour of the minimum phase part is examined.

5.1. Comparison of the two proposed methods

The two methods in this paper can be used to design factorable Nyquist filters with Chebyshev stopband behaviour. An equiripple stopband is obtained when $W(e^{j\omega})=1$. A nonequirriple design is achieved by specifying a nonconstant $W(e^{j\omega})$. In the first method, we design an unconstrained $H_1(z)$. When this $H_1(z)$ is squared, the resulting nonnegative frequency response has extremal frequencies that include those obtained in the design of $H_1(z)$. These are augmented by another set at which the response is zero. In the second method, we design $H_1^2(z)$ directly. The error is minimized over the same closed region as in the first method while maintaining the same total number of extremal frequencies. Since the constrained minimax approximation is unique [3], $H_1^2(z)$ is the same for both methods.

Despite the theoretical equivalence of the two methods, numerical differences do arise. The coefficients of $H_1^2(z)$ obtained by the two methods differ slightly in practice. Although these small differences lead to more pronounced differences in the coefficients of $H_0(z)$, the coefficients of the overall Nyquist filters formed by the two methods show only small differences. These differences manifest themselves mostly in the stopband region of the frequency response. An equiripple characteristic is more closely approached by the first method.

5.2. Design examples

Examples are presented to demonstrate both equiripple and nonequiripple designs. The design computations were done using double precision floating point arithmetic. Four iterations were necessary to resolve the coefficients. The following examples are generated by the first of our methods.

Example 1. We generate an equiripple design with parameters $N=6$, $l_0=4$, $l_1=25$, $\omega_p=0.08\pi$ and $\omega_s=0.254\pi$. This results in a filter with 59 coefficients having a roll-off factor $\beta=0.52$ whose magnitude response is shown in Fig. 2. The passband response is flat to within 0.003 dB. The filter length is of the form $M=2kN-1$ with $k=5$.

Example 2. The parameters used in this example are $N=4$, $l_0=4$, $l_1=15$, $\omega_p=0.12\pi$ and $\omega_s=0.38\pi$.

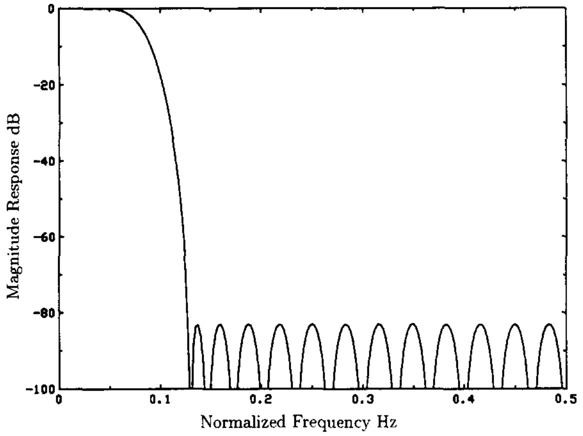


Fig. 2. Magnitude response of the Nyquist filter: Example 1.

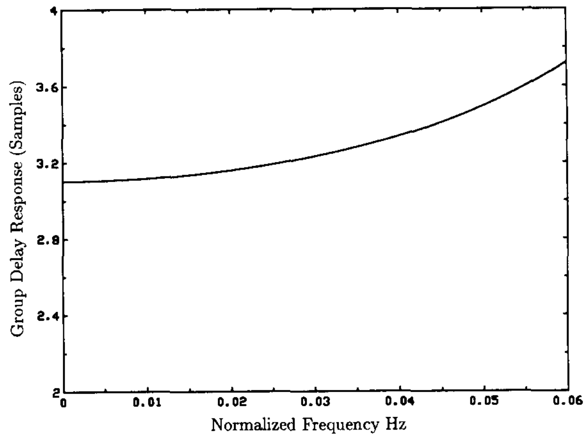


Fig. 4. Group delay response of the minimum phase part: Example 2.

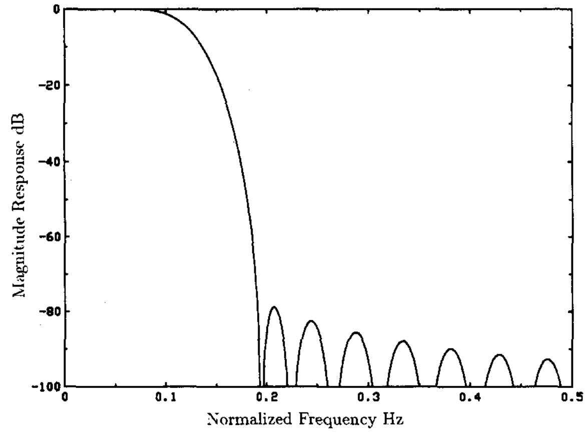
The weighting is

$$W(e^{j\omega}) = \begin{cases} 1, & \text{for } \omega = 0, \\ \dfrac{20}{2\pi}(\omega - \omega_s) + 1, & \text{for } \omega_s \leq \omega \leq \pi. \end{cases} \tag{9}$$

This gives a nonequiripple Nyquist filter with 39 coefficients and a roll-off factor $\beta = 0.52$. The filter length is of the form $M = 2kN - 1$ with $k = 5$. Fig. 3 shows the magnitude response of the filter. The passband response is flat to within 0.002 dB. Fig. 4 shows the group delay response of the minimum phase part of the filter.

5.3. Group delay

An important question concerns the delay distortion of the minimum phase part. The group delay of the minimum phase part is only important in the passband and is primarily influenced by the passband zeros which are within the unit circle. For a given number of taps and a given N, the group delay tends to be more constant as the roll-off factor increases. Also, for a given roll-off factor and a given N, a larger number of taps produces a group delay with a greater deviation. The minimum phase filters generated in Examples 1 and 2 that achieve about a 40 dB stopband attenuation have a relatively small passband group delay variation (approximately 0.15 zero crossing intervals).



If factorization of $H(z)$ into two constant group delay functions $A(z)$ and $B(z)$ is a requirement, a procedure is possible as follows. First, the double zeros of $H_1^2(z)$ are allocated one each to $A(z)$ and to $B(z)$. Then, we classify the zeros of $H_0(z)$ in polar form $re^{j\theta}$ and only consider $0 \leq \theta \leq \pi$. The zeros of $H_0(z)$ are taken in ascending order of θ and the mirror-image pairs are alternately assigned to $A(z)$ and $B(z)$. This ensures that both $A(z)$ and $B(z)$ have constant group delay. Note that if l_0 is odd, the number of taps of $A(z)$ and $B(z)$ differ by two. Otherwise, they have the same number of taps. Due to the presence of identical stopband zeros in $A(z)$, $B(z)$ and $H(z)$, the stopband response of both $A(z)$ and $B(z)$ is good. However, the passband responses can deviate significantly

6. Comparison with other approaches

This section discusses the relative merits of the factorable minimax design methods when compared with other approaches.

6.1. Linear programming technique

In [13], a linear programming approach that is also based on a minimax criterion is used to design a factorable Nyquist filter. For comparison, we generate a filter with the same parameters as the example in [13] using our factorable minimax approach. It is observed that the magnitude and group delay responses of the filters given by the two designs are very similar. The equiripple magnitude characteristic is more exactly given by our approach. Arbitrary weighting can be easily applied in both the factorable minimax approach and a linear programming formulation [7].

6.2. Eigenfilter formulation

The eigenfilter approach [11] also simplifies the factorization problem and meets the time domain constraints by solving a linear system of equations. The differences between the factorable minimax approach and the eigenfilter method are as follows. First, our approach is based on a minimax criterion as opposed to a least squares design achieved by the eigenfilter method. The factorable minimax approach naturally generates an equiripple behaviour whereas the eigenfilter method naturally renders nonequiripple filters. However, weighting can be applied in both methods to alter the stopband characteristic. For the factorable minimax method, the McClellan-Parks algorithm can easily incorporate arbitrary weighting, whereas, the incorporation of an arbitrary weighting factor into the eigenfilter formulation involves the use of numerical integration techniques.

A design example illustrates the differences in performance of the two methods. Identical parameters to the ones in [11] are used. In particular, $N=3$, $l_0=10$, $l_1=21$, $\omega_p=0.233\pi$, $\omega_s=0.433\pi$ and $W(e^{j\omega})=1$. This gives a Nyquist filter with 63 coefficients and a roll-off factor $\beta=0.3$. Fig. 5 shows the magnitude response of the minimum phase part generated by our factorable minimax method. The stopband attenuation of the minimum phase filter achieved by our method is about 48 dB whereas the first stopband ripple of its counterpart generated by the eigenfilter method shows an attentuation of approximately 45 dB. For higher frequencies, the ripples of the filter designed by the eigenfilter method show an attenuation that is more than that achieved by our method.

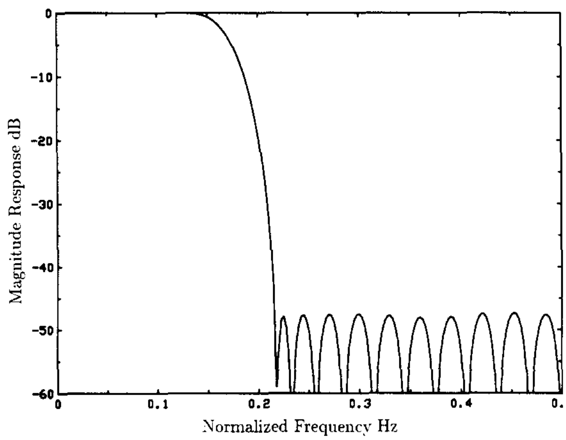


Fig. 5. Magnitude response of the minimum phase part of the Nyquist filter designed with the same parameters as in [11] ($N=3$, $l_0=10$, $l_1=21$, $\omega_p=0.233\pi$, $\omega_s=0.433\pi$ and $W(e^{j\omega})=1$).

6.3. Direct use of the McClellan-Parks algorithm

Factorable Nyquist filters can also be designed by invoking the constrained form of the McClellan-Parks algorithm [3] to get a nonnegative response that approximates a raised cosine characteristic. Another approach is to design a linear phase filter to approximate the square root of a raised cosine response using the McClellan-Parks algorithm (no constraints required) and convolve it with itself to produce a Nyquist filter. In this case, no factorization is required. The basic drawback of these two approaches is that exact zero crossings in the impulse response are not guaranteed as compared to other design methods.

The direct approaches used to design approximations to raised cosine Nyquist filters can be used for modem design. The CCITT recommendation V.22 [2] includes the specification of a pair of transmitter/receiver filters which should approximate the square root of a raised cosine response. The specified roll-off factor is 0.75. Upper and lower bounds in the frequency response in both the passband, transition band and a small portion of the stopband must be met. In addition, the group delay variation should be below a prescribed limit in the passband and a portion of the transition band.

We design Nyquist filters with a roll-off factor of 0.75 and with $N=4$ using the first factorable minimax method and the direct approaches that use the McClellan–Parks algorithm. The approaches are described in slightly more detail as follows:

(1) Design a filter that approximates a raised cosine response by invoking the constrained form of the McClellan–Parks algorithm [3] such that the response is nonnegative and its minimum and maximum phase parts have a frequency response that satisfies the upper and lower bounds specified by V.22.

(2) Design a linear phase filter that approximates a square root raised cosine characteristic by invoking the modified form of the McClellan–Parks algorithm such that its frequency response satisfies the upper and lower bounds specified by V.22.

(3) Use the first factorable minimax method to design a Nyquist filter such that its minimum and maximum phase parts satisfy the V.22 specifications of the frequency response.

In all cases, the smallest number of taps that satisfy the constraint $M=2kN-1$ is used. This leads to 15 tap Nyquist filters for the three methods. The weighting, $W(e^{j\omega})=1$, is used in the factorable minimax method.

Before considering the relative performance of the three methods, some general comments are in order. In the first method, factoring the z-transform of the Nyquist filter can be avoided since the unit circle zeros can be extracted from the extremal frequencies. However, the other zeros would have to be determined by first deflating the original polynomial. It is observed in [17] that deflation is more stable is the zeros of smaller magnitude were extracted first. This further discourages the division of the original polynomial by a polynomial that has the unit circle zeros since they have a larger magnitude than the zeros within the unit circle which should be extracted first to enhance the stability of the deflation process. A remedy to this problem is to use interpolation as in [6] to obtain a polynomial that represents the passband zeros and then factor it to obtain the zeros inside the unit circle. An alternative is to use a modified Newton's iteration [1] on the original polynomial to obtain the zeros inside the unit circle. The second method imposes no factorization problems.

The third approach using the first factorable minimax method does not guarantee a filter that satisfies any prescribed specifications of the frequency response. However, filters that satisfy the V.22 specifications can be designed by choosing the number of taps, carrying out the design and finally verifying that the constraints are met. We find that the constraints are met with 15 taps. It is observed that increasing the number of taps will cause the frequency response constraints to be violated since the transition band becomes more steep and lies outside the acceptable region.

In comparing the performance of the three methods, we consider the stopband attenuation of the Nyquist filter, the group delay of the factorized filter in the region considered in the V.22 specifications and quantitative measures of the intersymbol interference. The measures of the intersymbol interference are the peak distortion D_P and the RMS distortion D_{RMS} defined by

$$D_P=\sum_{\substack{n\\ n\neq 0}} |h(nN)|/|h(0)| \tag{10}$$

and

$$D_{RMS}=\sqrt{\sum_{\substack{n\\ n\neq 0}} h^2(nN)/h^2(0)}. \tag{11}$$

The stopband attenuations produced for

methods 1, 2 and 3 are about 45, 42 and 50 dB respectively. The allowable variation in group delay as specified by V.22 is 0.18 zero crossing intervals. Method 1 generates a minimum phase filter whose group delay variation is slightly under the prescribed 0.18 zero crossing intervals. Method 2 generates a filter with no group delay variation. Only Method 3 does not meet the group delay requirement in that the filter it produces has a variation of 0.24 zero crossing intervals.[5] In terms of peak and RMS distortion, Method 3 assures exact zero crossings and hence, produces no such distortion. Method 1 produces peak and RMS distortions of 0.0004 and 0.0003, respectively. Method 2 leads to much higher peak and RMS distortions of 0.186 and 0.132, respectively. This coupled with the lower stopband attenuation achieved by the second method leads us to choose Method 1 over Method 2. Method 3 gives a higher stopband attenuation than Method 1 and produces exact zero crossings in the impulse response. This enhanced stopband attenuation comes at the expense of a larger group delay variation.

7. Summary and conclusions

This paper describes two factorable minimax methods to design zero-phase FIR lowpass factorable Nyquist filters. Both use the McClellan–Parks algorithm as a first step in establishing the unit circle zeros and achieving a Chebyshev stopband response. The time domain constraints and the passband response are found by solving a system of linear equations. Both methods are iterative and four iterations are found to be sufficient in our examples to resolve the coefficients. The main advantages of the design techniques are that the polynomial factorization complexity in finding the minimum phase part is considerably eased and that arbitrary frequency weighting can be applied without additional computational overhead. Although the two design approaches should theoretically give the same filter, the first of our methods is numerically more accurate and hence, renders a slightly better frequency response. No numerical difficulties were encountered when using this technique. Both the appropriate zero crossings and expected weighted stopband behaviour are achieved.

The transfer function of the Nyquist filter is split into a product of a relatively low order polynomial having zeros that are not on the unit circle and a relatively high order polynomial having all its zeros on the unit circle. The orders of these polynomials are unique given the number of filter coefficients and the zero crossing interval N. This decomposition of the transfer function of the filter is advantageous in that the minimum and maximum phase parts are found by only factoring the low order polynomial as opposed to factorizing the overall polynomial representing the Nyquist filter. The relative factorization complexity depends on the proportion of the number of unit circle zeros to the number of other zeros. Upper and lower bounds for this ratio are derived. For practical filter lengths, the polynomial factorization remains simple and can be done by a general purpose routine.

Design examples include both equiripple and nonequiripple filters. Comparisons with both a linear programming approach and the eigenfilter formulation show that the proposed methods are good in terms of both magnitude response and group delay variation. A practical design that conforms to the CCITT V.22 specifications is compared with approaches that directly invoke the McClellan–Parks algorithm to approximate a raised cosine response. The factorable minimax method and the use of the McClellan–Parks algorithm differ in performance in that the former achieves a higher stopband attenuation but a group delay with a greater variation. Furthermore, the direct approaches do not guarantee exact zero crossings in the impulse response thereby leading to residual intersymbol interference.

[5] A simple second order allpass equalizer brings the group delay within specifications. However, the use of such equalizers sacrifices the exact zero crossing property of the original design.

The ideas in this paper are generalizable to the design of factorable partial response filters. Only two of the nine partial response systems described in [5] can be factored into equal magnitude parts. The transfer functions of these two filters can be obtained by first designing the Nyquist filter $H(z)$ and cascading it with the appropriate polynomial in z^N. This introduces the controlled amount of intersymbol interference. The minimum phase factor of the partial response filter is the minimum phase factor of $H(z)$ multiplied by the minimum phase factor of the polynomial in z^N as given in [5].

Appendix A. Constraints on the parameters l_0 and l_1

Let the zero crossing interval be N and the number of filter coefficients be $M=2(l_0+l_1)+1$. The parameters l_0 and l_1 satisfy

$$l_0(N-1)\leqslant l_1 < l_0(N-1)+N. \tag{A.1}$$

Since $l_1=\frac{1}{2}(M-1)-l_0$, the inequality reduces to

$$l_0 \leqslant \frac{M-1}{2N} < l_0+1. \tag{A.2}$$

This new inequality is satisfied by a unique l_0 given by

$$l_0=\left\lfloor \frac{M-1}{2N} \right\rfloor. \tag{A.3}$$

Then, l_1 is given by

$$l_1=\frac{M-1}{2}-\left\lfloor \frac{M-1}{2N} \right\rfloor. \tag{A.4}$$

Appendix B. The ratio l_1/l_0: Lower and upper bounds

This appendix derives lower and upper bounds for l_1/l_0 and shows how to fix the filter length to achieve these bounds. This ratio is only finite for $l_0 \neq 0$ which is a reasonable assumption. If $l_0=0$, the filter length $M<2N-1$ thereby giving an impulse response with no zero crossings and hence, an insufficient length for an acceptable stopband attenuation.

B.1. Lower bound

The lower bound for l_1/l_0 is given by the lefthand side of (A.1),

$$\frac{l_1}{l_0} \geqslant N-1.$$

The lower bound is achieved if and only if l_0 and l_1 are given by

$$l_0=\frac{M-1}{2N}, \qquad l_1=\frac{(M-1)(N-1)}{2N}. \tag{B.1}$$

In this case, the filter length is of the form $M=2l_0N+1$ thereby giving an impulse response with the two end coefficients equal to zero.

If l_0 and l_1 are chosen as above, the system of equations $\boldsymbol{Fh}=\boldsymbol{c}$ that solve for the coefficients of $H_0(z)$ can be decoupled into a reduced system of dimension l_0 and the additional equation $f(l_1)h_0(l_0)=0$. Hence, $h_0(-l_0)=h_0(l_0)=0$ thereby reducing the effective values of l_0 and M by 1 and 2 respectively. Such a choice of parameters gives results that are identical to the case when l_0 is reduced by 1.

B.2. Upper bound

The upper bound for l_1/l_0 is obtained by examining the righthand side of (A.1),

$$\frac{l_1}{l_0} < N-1+\frac{N}{l_0}.$$

Since the minimum value of l_0 is 1, an upper bound is $2N-1$. Achieving a ratio equal to a value of $2N-2$ is possible if and only if $l_0=1$ and $l_1=2N-2$. If $l_0>1$, the upper bound $N-1+N/l_0 \leqslant 2N-2$ for every $N\geqslant 2$. Hence, $l_1/l_0<2N-2$ for every $l_0>1$. The final conclusion is that for a given N, there exists only one filter length, namely, $M=4N-1$ that achieves the maximum value $l_1/l_0=2N-2$.

References

[1] X. Chen and T.W. Parks, "Design of optimal minimum phase FIR filters by direct factorization", *Signal Process.*, Vol. 10, June 1986, pp. 369-383.

[2] CCITT Recommendation V.22, "Data communication over the telephone network", Red Book, V Series, October 1984.

[3] F. Grenez, "Design of linear and minimum phase FIR filters by constrained Chebyshev approximation", *Signal Process.*, Vol. 5, July 1983, pp. 325-332.

[4] O. Herrmann and W. Schussler, "Design of nonrecursive digital filters with minimum phase", *Electron. Letts.*, Vol. 6, April 1970, pp. 329-330.

[5] P. Kabal and S. Pasupathy, "Partial-response signaling", *IEEE Trans. Commun.*, Vol. COM-23, September 1975, pp. 921-934.

[6] Y. Kamp and C.J. Wellekens, "Optimal design of minimum phase FIR filters", *IEEE Trans. Acoust. Speech, Signal Process.*, Vol. ASSP-31, August 1983, pp. 922-926.

[7] J.K. Liang, R.J.P. DeFigueiredo and F.C. Lu, "Design of optimal Nyquist, partial response, *N*th band and nonuniform tap spacing FIR digital filters using linear programming techniques", *IEEE Trans. Circuits Syst.*, Vol. CAS-32, April 1985, pp. 386-392.

[8] J.H. McClellan, T.W. Parks and L.R. Rabiner, "A computer program for designing optimum FIR linear phase digital filters", *IEEE Trans. Audio Electroacoust.*, Vol. AU-21, December 1973, pp. 506-526.

[9] F. Mintzer, "On half-band, third-band and *N*th-band FIR filters and their design", *IEEE Trans. Acoust. Speech, Signal Process.*, Vol. ASSP-30, October 1982, pp. 734-738.

[10] F. Mintzer, "Filters for distortion-free two-band multirate filter banks", *IEEE Trans. Acoust. Speech, Signal Process.*, Vol. ASSP-33, June 1985, pp. 626-630.

[11] T.Q. Nguyen, T. Saramaki and P.P. Vaidyanathan, "Eigenfilters for the design of special transfer functions with applications in multirate signal processing", *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, New York, NY, April 1988, pp. 1467-1470.

[12] L.R. Rabiner and B. Gold, *Theory and Applications of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.

[13] A.C. Salazar and V.B. Lawrence, "Design and implementation of transmitter and receiver filters with periodic coefficients nulls for digital systems", *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Paris, France, May 1982, pp. 306-310.

[14] T. Saramaki and Y. Neuvo, "A class of FIR Nyquist (*N*th band) filters with zero intersymbol interference", *IEEE Trans. Circuits Syst.*, Vol. CAS-34, October 1987, pp. 1182-1190.

[15] P.P. Vaidyanathan and T.Q. Nguyen, "Eigenfilters: A new approach to least-squares FIR filter design and applications including Nyquist filters", *IEEE Trans. Circuits Syst.*, Vol. CAS-34, January 1987, pp. 11-23.

[16] P.P. Vaidyanathan and T.Q. Nguyen, "A TRICK for the design of FIR half-band filters", *IEEE Trans. Circuits Syst.*, Vol. CAS-34, March 1987, pp. 297-300.

[17] J.H. Wilkinson, *Rounding Errors in Algebraic Processes*, Prentice-Hall, Englewood Cliffs, NJ, 1963.