Nonlinear Phase FIR Filter Design according to the L_2 Norm with Constraints for the Complex Error

Markus Lang
Institute of Communication Theory
University of Erlangen
Cauerstr. 7, 91058 Erlangen, Germany
e-mail: lang@nt.e-technik.uni-erlangen.de

Joachim Bamberger
Haagstr. 24
91054 Erlangen, Germany

November 18, 1994

Abstract

We examine the problem of approximating a complex frequency response by a real-valued FIR filter according to the L_2 norm subject to additional inequality constraints for the complex error function. Starting with the Kuhn-Tucker optimality conditions which specialize to a system of nonlinear equations we deduce an iterative algorithm. These equations are solved by Newton's method in every iteration step. The algorithm allows arbitrary tradeoffs between an L_2 and an L_∞ design. The L_2 and the L_∞ solution result as special cases.

Wir untersuchen das Problem der Approximation eines komplexen Frequenzganges mittels eines reellwertigen nichtrekursiven Filters nach der L_2-Norm mit zusätzlichen Ungleichungsbedingungen für die komplexe Fehlerfunktion. Ausgehend von den Kuhn-Tucker Optimalitätsbedingungen, die auf ein nichtlineares Gleichungssystem führen, leiten wir einen iterativen Algorithmus her. Diese Gleichungen werden in jedem Iterationsschritt mittels des Newton-Verfahrens gelöst. Der Algorithmus erlaubt beliebige Kompromisse zwischen einem L_2- und einem L_∞-Entwurf. Die L_2- und die L_∞-Lösung ergeben sich als Spezialfälle.

List of Tables

List of Figures

14 manuscript pages (including references)

Keywords: Complex FIR Filter, Filter Design, Constrained Optimization

1 Introduction

Although FIR filter design has been treated in the past by many publications only few consider the case of a non symmetric impulse response or an arbitrary nonlinear phase complex desired function (NLPFIR filter), which we will do in this paper. The design of FIR filters according to the L_∞ norm with an (anti) symmetric impulse response (LPFIR filters) can be carried out by using either the efficient Remez' algorithm [8] or Linear Programming [11]. If the L_2 norm is used, the problem is reduced to solving a system of linear equations. The latter approach can also be applied to arbitrary complex desired functions which can be interesting for equalization problems [10, 12] or nearly LPFIR filters with reduced complex error [5]. The generalization of the algorithms minimizing the L_∞ norm to this case is more complicated. There are three approaches: A generalization of Remez' algorithm [10, 12] which is fast but not guaranteed to converge, a Linear Programming approach [5, 6] yielding suboptimal results with a high demand of computation time and computer storage and an algorithm of Tang [2, 14] which is not guaranteed to converge theoretically but does so in nearly all practical applications [12]. As can be seen the optimum filter can be designed if just the L_2 or the L_∞ norm is used. Based on an iterative reweighted least squares algorithm Burrus and Barreto gave an L_p approach for LPFIR filters [4] which can be interpreted as a compromise between an L_2 and an L_∞ solution. With the methods mentioned above the minimization of the error energy (L_2 norm) keeping a prescribed maximum error at the same time is not possible. Adams [1] brought up this problem and gave an approach for the solution for LPFIR filters.

Considering a discrete frequency grid we present a new algorithm which can be seen as a generalization of Adams' approach to the case of NLPFIR filters. With this new method it is possible for the first time to design NLPFIR filters according to the constrained L_2 error criterion described above. The weighting and desired functions can be arbitrarily chosen. The usual L_2 and L_∞ design and the design of LPFIR filters result as specializations of our algorithm. Possible applications are the same as for the traditional L_2 or L_∞ approach (equalization, nearly LPFIR filter). But with our new algorithm we can also solve problems which do not fit into a L_2 or a L_∞ formulation. An example is the design of filters used in Σ-Δ converters [6] where we have the problem of minimizing the error energy of an NLPFIR filter in one band keeping a given tolerance in the other.

Furthermore we give a simple method to essentially improve the convergence behaviour of Adam's approach. Although we do not prove convergence our algorithm converges in most practical applications and we give a short discussion of convergence problems which is also valid for Adam's method. Convergence of the new algorithm is tantamount to optimality of the solution because of our approach.

2 Description of the Procedure

In this section we consider the problem of determining a real-valued NLPFIR filter approximating a desired function, which keeps a prescribed error tolerance δ and minimizes the absolute quadratic error. Both the L_∞ and the L_2 solution result as specializations if the minimax error is prescribed for δ on the one hand and a very large value for δ on the other hand. By continuously increasing δ beginning from the minimax error arbitrary compromises between an L_∞ and an L_2 solution are possible.

We formulate the filter design as the problem of minimizing a quadratic function subject to quadratic constraints. The general Kuhn-Tucker (KT) conditions, which are necessary and sufficient for the optimality of the solution, yield a set of nonlinear equations which are solved by Newton's Method. These KT conditions generalize the approach of finding the minimum of a function subject to *equality* constraints by use of Lagrange multipliers to the case of *inequality* constraints. Just as in the case of Remez' Algorithm the local extremal frequencies play an important role. So the two essential steps of our new iterative algorithm are: determination of the extremal frequencies of the error function and solving a system of nonlinear equations.

We introduce the causal and real-valued impulse response $h(k)$ of length $n+1$, the corresponding frequency response

$$H(e^{j\Omega}) = \sum_{k=0}^{n} h(k) e^{-jk\Omega}, \tag{1}$$

a desired function $D(e^{j\Omega})$ with the property $D(e^{j\Omega}) = D^*(e^{-j\Omega})$, and a complex error function

$$E(e^{j\Omega}) = H(e^{j\Omega}) - D(e^{j\Omega}) = E^*(e^{-j\Omega})\,. \tag{2}$$

Since all computations have to be carried out on a finite frequency grid $\Omega_i\,,\ i = 1\ldots L$, we restrict the considerations to such a formulation and use equidistant spacing. The continuous formulation in frequency is straightforward: the error sums are simply replaced by integrals. The structure of the essential equations remains unchanged.
By use of discretization, Eq. (1) can be written as

$$\mathbf{H} = \mathbf{A}\mathbf{h} \tag{3}$$

with the vector $\mathbf{H} = [H(e^{j\Omega_1}) \, \ldots \, H(e^{j\Omega_L})]^T$, the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & e^{-j\Omega_1} & \ldots & e^{-jn\Omega_1} \\ \vdots & \vdots & & \vdots \\ 1 & e^{-j\Omega_L} & \ldots & e^{-jn\Omega_L} \end{bmatrix}, \tag{4}$$

and the vector $\mathbf{h} = [h(0) \, \ldots \, h(n)]^T$. The error vector

$$\mathbf{E} = \mathbf{A}\mathbf{h} - \mathbf{D} \tag{5}$$

with $\mathbf{E} = [E(e^{j\Omega_1}) \, \ldots \, E(e^{j\Omega_L})]^T$ and $\mathbf{D} = [D(e^{j\Omega_1}) \, \ldots \, D(e^{j\Omega_L})]^T$ results. Since we want to minimize the L_2 norm we introduce the absolute weighted squared error

$$\begin{aligned} \varepsilon(\mathbf{h}) &= \sum_{i=1}^{L} W_E(\Omega_i)|E(e^{j\Omega_i})|^2 = \mathbf{E}^H \mathbf{W}_E \mathbf{E} \\ &= \mathbf{h}^T \mathbf{A}^H \mathbf{W}_E \mathbf{A} \mathbf{h} - 2\,\mathrm{Re}\,\{\mathbf{h}^T \mathbf{A}^H \mathbf{W}_E \mathbf{D}\} + \mathbf{D}^H \mathbf{W}_E \mathbf{D} \end{aligned} \tag{6}$$

where $\mathbf{W}_E = \mathrm{diag}\,(W_E(\Omega_1) \, \ldots \, W_E(\Omega_L))$ is a positive semidefinite diagonal matrix and the superscript H means conjugate transposed. As is well-known the L_2 solution vector can be determined by searching the minimum of ε which is equivalent to solving

$$\mathrm{grad}_{\mathbf{h}}(\mathbf{h}^T \mathbf{R} \mathbf{h} - 2\mathbf{h}^T \mathbf{d} + d) = 0, \tag{7}$$

with the abbreviations

$$\mathbf{R} = \mathrm{Re}\{\mathbf{A}^H \mathbf{W}_E \mathbf{A}\}, \tag{8}$$

$$\mathbf{d} = \mathrm{Re}\{\mathbf{A}^H \mathbf{W}_E \mathbf{D}\}, \tag{9}$$

$$d = \mathbf{D}^H \mathbf{W}_E \mathbf{D}. \tag{10}$$

Eq. (7) leads to the well-known “Normal Equation”

$$\mathbf{R}\mathbf{h} = \mathbf{d}\,. \tag{11}$$

Until here we considered only the unconstrained L_2 filter design problem. Additional inequality constraints concerning an error tolerance can be incorporated by prescribing

$$W_A^2(\Omega_i)|E(e^{j\Omega_i})|^2 \leq \delta^2\,, \; i = 1 \ldots L\,, \tag{12}$$

with the nonnegative weighting function $W_A(\Omega_i)$ and the error tolerance δ. Relation (12) can be equivalently written as

$$B(\mathbf{h}, \Omega_i) := W_A^2(\Omega_i)[\mathbf{h}^T \tilde{\mathbf{R}}(\Omega_i)\mathbf{h} - 2\mathbf{h}^T \tilde{\mathbf{d}}(\Omega_i) + \tilde{d}(\Omega_i)] - \delta^2 \leq 0 \tag{13}$$

with

$$
\begin{aligned}
\tilde{\mathbf{R}}(\Omega_i) &= \mathrm{Re}\,\{\mathbf{A}_i^H \mathbf{A}_i\}, &(14)\\
\tilde{\mathbf{d}}(\Omega_i) &= \mathrm{Re}\,\{\mathbf{A}_i^H D(e^{j\Omega_i})\}, &(15)\\
\tilde{d}(\Omega_i) &= |D(e^{j\Omega_i})|^2, &(16)
\end{aligned}
$$

where $i = 1 \ldots L$. $\mathbf{A}_i$ denotes the ith row of the matrix $\mathbf{A}$. The filter design problem can now be formulated as a quadratic minimization problem with quadratic inequality constraints:

$$
\begin{aligned}
&\underset{\mathbf{h}}{\text{Minimize}} \quad \varepsilon(\mathbf{h}) = \mathbf{h}^T\mathbf{R}\mathbf{h} - 2\mathbf{h}^T\mathbf{d} + d \\
&\text{subject to} \quad B(\mathbf{h}, \Omega_i) \le 0,\; i = 1 \ldots L.
\end{aligned}
\tag{17}
$$

The pivotal point of our further consideration are the necessary and sufficient KT conditions [13]. They hold for problems of a structure as in (17) where each function (ε and B) nonlinearly depends on $\mathbf{h}$. Furthermore we have to suppose that there exists a solution and both functions are differentiable and real and convex functions of the real vector $\mathbf{h}$. Then the optimum solution is given if and only if the following conditions are fulfilled:

$$
\begin{aligned}
\mathrm{grad}_{\mathbf{h}}\varepsilon(\mathbf{h}) + \sum_{i=1}^{L} \mu_i \mathrm{grad}_{\mathbf{h}} B(\mathbf{h}, \Omega_i) &= \mathbf{0}, &&(18)\\
B(\mathbf{h}, \Omega_i) &\le 0, \quad i = 1 \ldots L &&(19)\\
\mu_i &\ge 0, \quad i = 1 \ldots L &&(20)\\
\mu_i\, B(\mathbf{h}, \Omega_i) &= 0, \quad i = 1 \ldots L &&(21)
\end{aligned}
$$

The interpretation of these conditions (18)-(21) is analogous to the case of linear inequalities $B(\mathbf{h}, \Omega)$. We omit a detailed discussion here since Adams gave one in [1]. We only want to point out some statements. The so-called KT coefficients μ_i must be nonnegative (20). This property guarantees that we get a minimum of ε. For any positive μ_i the corresponding constraints $B(\mathbf{h}, \Omega_i) \le 0$ must be met with equality and vice versa because of Eq. (21). Those frequency points Ω_i with $B(\mathbf{h}, \Omega_i) = 0$ correspond to local maxima of the weighted error function (12). The set of all such constraints is called the set of active constraints and the corresponding set of M indices is called $I = \{\nu_1 \ldots \nu_M\}$. Because of Eq. (21) only the KT coefficients of this set can be nonzero.

Since the Hessian matrices of $\varepsilon(\mathbf{h})$, $B(\mathbf{h}, \Omega_i)$ are the positive semidefinite matrices $\mathbf{R}$ and $\tilde{\mathbf{R}}(\Omega_i)$, respectively, we have convex functions [13]. The other requirements for the Kuhn-Tucker conditions are also fulfilled. Consequently, the optimum solution of our filter design problem is uniquely characterized by the KT conditions.

If the active set and the set of indices I were known in advance the relations (18) and (19) would yield a system of $n+1+M$ nonlinear equations for the $n+1+M$ unknown parameters combined in the vector

$$\mathbf{x}^T = \left[\mathbf{h}^T \ \boldsymbol{\mu}^T\right], \qquad \text{with} \quad \boldsymbol{\mu}^T = [\mu_{\nu_1} \dots \mu_{\nu_M}]. \tag{22}$$

In this case we have to determine the minimum of ε subject to equality constraints and the KT conditions reduce to the well-known method of Lagrangian multipliers [13]. The problem of course is that we do *not* a priori know the active constraints. Similar to Remez' algorithm we proceed iteratively using each local error maximum of $W_A(\Omega_i)|E(e^{j\Omega_i})|$ larger than δ as an indicator for an active constraint. Let us suppose we are given an initial solution $\mathbf{h}^{(0)}$ and a corresponding set of KT coefficients $\boldsymbol{\mu}^{(0)}$. This can be computed by use of Eq. (11) and a vector $\boldsymbol{\mu}^{(0)} = [1 \dots 1]^T$. Then our algorithm consists of two major steps:

1. Determine all M local maxima $W_A(\Omega_i)|E(e^{j\Omega_i})| \geq \delta$ and the corresponding set of indices $\{\nu_1 \dots \nu_M\}$. Note that the value of M usually varies from one iteration step to the next.

2. Solve the system of $n+1+M$ nonlinear equations

$$\text{grad}_{\mathbf{h}}\varepsilon(\mathbf{h}) + \sum_{m=1}^{M} \mu_{\nu_m} \, \text{grad}_{\mathbf{h}} B(\mathbf{h}, \Omega_{\nu_m}) = \mathbf{0} \tag{23}$$

$$B(\mathbf{h}, \Omega_{\nu_m}) = 0\,, \quad m = 1 \dots M \tag{24}$$

   for $\mathbf{h}$ and μ_{ν_m}. If the solution meets the KT conditions (18)-(21) then stop. Repeat otherwise.

Since the first step is easily implemented by essentially determining the frequency response $H(e^{j\Omega_i})$ we concentrate on the second step. From Equations (23) and (24) it is seen that this second step can be interpreted as the method of Lagrangian multiplier where the actual equality constraints (24) change from one iteration step to the next due to the change of the extremal frequencies (active constraints). By use of (13) and (17) Eqs. (23) and (24) can be written as

$$\mathbf{g}_1(\mathbf{x}) = \mathbf{R}\mathbf{h} - \mathbf{d} + \sum_{m=1}^{M} \mu_{\nu_m} W_A^2(\Omega_{\nu_m})[\tilde{\mathbf{R}}(\Omega_{\nu_m})\mathbf{h} - \tilde{\mathbf{d}}(\Omega_{\nu_m})] = \mathbf{0} \tag{25}$$

$$\mathbf{g}_2(\mathbf{x}) = \begin{bmatrix} W_A^2(\Omega_{\nu_1})[\mathbf{h}^T\tilde{\mathbf{R}}(\Omega_{\nu_1})\mathbf{h} - 2\mathbf{h}^T\tilde{\mathbf{d}}(\Omega_{\nu_1}) + \tilde{d}(\Omega_{\nu_1})] - \delta^2 \\ \vdots \\ W_A^2(\Omega_{\nu_M})[\mathbf{h}^T\tilde{\mathbf{R}}(\Omega_{\nu_M})\mathbf{h} - 2\mathbf{h}^T\tilde{\mathbf{d}}(\Omega_{\nu_M}) + \tilde{d}(\Omega_{\nu_M})] - \delta^2 \end{bmatrix} = \mathbf{0}\,. \tag{26}$$

This system of nonlinear Eqs. (25) and (26) can be solved using Newton's method [13]

$$\mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} - \mathbf{J}^{-1}(\mathbf{x}^{(q)})\mathbf{f}(\mathbf{x}^{(q)}) \qquad (27)$$

with

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \mathbf{g}_1(\mathbf{x}) \\ \mathbf{g}_2(\mathbf{x}) \end{bmatrix}, \qquad (28)$$

the ith row of the Jacobian matrix

$$\mathbf{J}_i(\mathbf{x}) = [\mathrm{grad}_{\mathbf{x}} f_i(\mathbf{x})]^T , \qquad (29)$$

and $f_i(\mathbf{x})$ denoting the ith component of the column vector $\mathbf{f}(\mathbf{x})$. The iteration counter of Newton's method is indicated by (q) in Eq. (27). The Jacobian matrix can be written as

$$\mathbf{J} = \begin{bmatrix} \underset{n+1\times n+1}{\mathbf{J}_{11}} & \underset{n+1\times M}{\mathbf{J}_{12}} \\ & \\ \underset{M\times n+1}{\mathbf{J}_{21}} & \underset{M\times M}{\mathbf{J}_{22}} \end{bmatrix}, \qquad (30)$$

with

$$\mathbf{J}_{11} = \mathbf{R} + \sum_{m=1}^{M} \mu_{\nu_m} W_A^2(\Omega_{\nu_m})\tilde{\mathbf{R}}(\Omega_{\nu_m}), \qquad (31)$$

$$\mathbf{J}_{12} = \left[W_A^2(\Omega_{\nu_1})[\tilde{\mathbf{R}}(\Omega_{\nu_1})\mathbf{h} - \tilde{\mathbf{d}}(\Omega_{\nu_1})] \ \ldots \ W_A^2(\Omega_{\nu_M})[\tilde{\mathbf{R}}(\Omega_{\nu_M})\mathbf{h} - \tilde{\mathbf{d}}(\Omega_{\nu_M})]\right], \qquad (32)$$

$$\mathbf{J}_{21} = \mathbf{J}_{12}^T, \qquad (33)$$

$$\mathbf{J}_{22} = \mathbf{0} \qquad (34)$$

Considering Eqs. (31) and (32) and the fact that $\mathbf{J}$ has to be built up for every iteration step (q) the computational load and the storage requirements only for determining $\mathbf{J}$ seem to be huge. Both difficulties can be reduced by using the fact that $\mathbf{R}$ and $\tilde{\mathbf{R}}(\Omega_i)$ are symmetrical Toeplitz matrices, so only one row of each matrix is necessary for a complete description. A further improvement is possible since the product $W_A^2(\Omega_{\nu_m})\tilde{\mathbf{R}}(\Omega_{\nu_m})$ in Eq. (31) remains unchanged during Newton's method.

Assuming our algorithm converges, only the first two KT conditions are fulfilled by the procedure described above. The KT coefficients μ_{ν_m} are not guaranteed to be nonnegative. This problem also occurs at Adams' algorithm [1]. In both cases it can simply be solved by checking the signs of μ_{ν_m}, $m = 1 \ldots M$ when a solution vector $\mathbf{x}$ of the Eqs. (23) and (24) is found. Those equations $B(\mathbf{h}, \Omega_{\nu_m}) = 0$ which correspond to negative values of μ_{ν_m} are removed from the set of active constraints. Newton's method which corresponds to the solution of a system of linear equations at Adam's algorithm is started again. The procedure is repeated until no negative KT coefficients result. This step considerably reduces convergence problems and leads to the optimum solution in the case of convergence.

The latter is due to the fact that all KT conditions (18)-(21) are taken into account by the algorithm described above.

3 Design Examples

Example1: We consider the design of an approximately linear phase lowpass filter with $n = 30$,

$$D(e^{j\Omega}) = \begin{cases} e^{-j12\Omega} & \Omega \in B_p \\ 0 & \Omega \in B_s \end{cases}, \tag{35}$$

$$W_A(\Omega) = \begin{cases} 1 & \Omega \in B_p \\ 10 & \Omega \in B_s \end{cases}, \tag{36}$$

with the passband $B_p = [0, \Omega_p]$, the stopband $B_s = [\Omega_s, \pi]$, $\Omega_p = 0.12\pi$, and $\Omega_s = 0.24\pi$, which was introduced by Chen and Parks in [8]. We used $L_p + 1 = 51$ frequency points $\Omega_i = i \cdot \Omega_p/L_p$, $i = 0 \ldots L_p$ in the passband and $L_s + 1 = 401$ frequency points $\Omega_i = \Omega_s + i \cdot (\pi - \Omega_s)/L_s$, $i = 0 \ldots L_s$ in the stopband.

In the following we present some examples which illustrate the effect of weighting and the choice of δ. The different examples are marked by Ex. 1-1 ... Ex. 1-7. With $W_E(\Omega) \equiv 1$ and the choice $\delta = 1$ (or any other "large" value) or $\delta = 0.0439252$ we get the L_2 and approximately the L_∞ solution (Ex. 1-1 and Ex. 1-2), respectively. In the first case we have only to solve the "Normal Equation" (11) once, whereas in the latter case $k = 12$ iteration steps are necessary as can be seen from Table 1. The corresponding magnitude responses and error functions $W_A(\Omega) \cdot |E(e^{j\Omega})|$ are depicted in Fig. 1 showing the well-known Gibbs phenomenon of the L_2 solution and an equiripple behaviour of the L_∞ solution. Obviously the L_∞ norm

$$\varepsilon_\infty = \max_{\Omega \in B} \left\{ W_A(\Omega) \cdot |E(e^{j\Omega})| \right\}, \quad B = B_p \cup B_s \tag{37}$$

of the L_2 solution is much larger than that of the L_∞ solution and vice versa the same is true for the error energy

$$\varepsilon_2 = \sum_{\Omega_i \in B} |E(e^{j\Omega_i})|^2, \quad B = B_p \cup B_s \tag{38}$$

as can be seen in Table 1. By our new algorithm it is now possible to design arbitrary compromises between these solutions. If we choose for example $\delta = 0.07$, which is far less than ε_∞ in the case of the L_2 solution, and

$$W_E(\Omega) = \begin{cases} 1, & \Omega \in B_p \\ 10^4, & \Omega \in B_s \end{cases} \tag{39}$$

and

$$W_E(\Omega) \equiv 1, \quad \Omega \in B, \tag{40}$$

respectively we get an equiripple behaviour in the passband and minimum stopband error energy

$$\varepsilon_{2s} = \sum_{\Omega_i \in B_s} |E(e^{j\Omega_i})|^2 \tag{41}$$

in the first case (Ex. 1-3) and minimum overall error energy ε_2 (with respect to $\delta = 0.07$) in the second (Ex. 1-4). The resulting values can be taken from Table 1. The corresponding magnitude responses and error functions are shown in Fig. 2.

Next we want to compare our algorithm to known methods as Tang's algorithm [14], Adams' algorithm for LPFIR filters [1] and Pei's and Shyu's Eigenfilter method [9]. The latter is equivalent to the solution of the "Normal Equation" (11) if we choose a continuous formulation (integrals) where our formulation seems to have some advantages. It is easier to understand and to implement and applies to arbitrary desired and weighting functions in opposite to [9] because one has to determine a new matrix for each new filter type (lowpass, differentiator, ...) and each new weighting function there. To get comparable results to [9], Example 1, we have chosen

$$W_E(\Omega) = \begin{cases} \Omega_p / L_p\,, & \Omega \in B_p \\ 5(\pi - \Omega_s)/L_s, & \Omega \in B_s \end{cases}, \tag{42}$$

which yields a good approximation to the error integrals in [9]. Unfortunately Pei and Shyu gave neither ε_∞ nor ε_2 of their solution but the maximum deviation in the passband with 0.0658 and the error magnitude of the first sidelobe in the stopband with 0.0068. Using (42) we get corresponding values 0.0652 and 0.0069 with our algorithm (Ex. 1-5) which seems to be near the result of [9]. Further values resulting for this design can be found in Table 1.

For comparison with Tang's algorithm [14] we designed the lowpass filter according to the L_∞ norm by use of his method (Ex. 1-6) and the weighting function $W_A(\Omega)$ given in Eq. (36). A comparison with Ex. 1-2 shows that we get the same error maximum up to 3 decimal places as with our algorithm when we use the value δ equal to ε_∞ of the L_∞ solution. Of course the resulting value ε_∞ of our algorithm is slightly worse since it is not possible to improve the L_∞ solution. On the other hand ε_2 of Ex. 1-6 is larger which is also expected as our method minimizes the L_2 norm in opposite to Tang's algorithm. The only disadvantage of our method with regard to the determination of an L_∞ solution is the fact that we must know the value δ in advance. This is a nontrivial problem which could be solved starting with the L_2 solution and by iteratively running our algorithm with decreasing δ until we are close enough to the L_∞ solution.

Finally we want to demonstrate the advantages of the NLPFIR filter design over the LPFIR filter design according to Adams [1]. For this Ex. 1-7 we modify our desired function to

$$D(e^{j\Omega}) = \begin{cases} e^{-j15\Omega} & \Omega \in B_p \\ 0 & \Omega \in B_s \end{cases} \tag{43}$$

and choose the same weighting functions and δ as in Ex. 1-3. The resulting magnitude response and the modulus of the weighted error function are shown in Fig. 3 together with the corresponding functions of Ex. 1-3. Table 1 contains the values of the different error norms. We get the same maximum value but larger values ε_2 and ε_{2s} than with our method.

As can be seen from Table 1, the algorithm typically converges within $k = 4$ iteration steps of the main loop as long as we are not very close to the L_∞ solution which is the case for Ex. 1-2. Table 1 additionally contains a column with the necessary number of floating point operations (Flops) for Ex. 1-1 ... Ex. 1-7.

Example2: To demonstrate the flexibility of our algorithm we choose a "phantasy" design

$$D(e^{j\Omega}) = \begin{cases} \cos(\Omega)e^{-j(a\cdot\Omega+c\cdot\Omega^2)}, & \Omega \in B_p \\ 0, & \Omega \in B_s \end{cases}, \tag{44}$$

$$W_A(\Omega) = \begin{cases} 1+9\,\dfrac{\Omega}{\Omega_p}, & \Omega \in B_p \\ 1+9\,\dfrac{\pi-\Omega}{\pi-\Omega_s}, & \Omega \in B_s \end{cases}, \tag{45}$$

with $n = 49,\ a = 20,\ c = 25/3,\ \Omega_p = 0.3\pi,\ \Omega_s = 0.4\pi$, and $L_p = L_s = 200$ equidistantly spaced frequency points in the passband and stopband. We used $W_E(\Omega) \equiv 1$, $\delta = 1$, for an L_2 design on the one hand (Ex. 2-1) and

$$W_E(\Omega) = \begin{cases} 1, & \Omega \in B_p \\ 1000, & \Omega \in B_s \end{cases}, \tag{46}$$

$\delta = 0.05$ to get an equiripple behaviour in the passband on the other hand (Ex. 2-2). Fig. 4(a) and 4(b) show the magnitude response with a cosine shape and the nearly linear group delay, respectively, of both filter designs in the passband. The corresponding error functions with and without weighting $W_A(\Omega)$ are depicted in Fig. 4(c) and 4(d). The weighting with the piecewise linear function $W_A(\Omega)$ means a prescription of a tolerance

$$\delta(\Omega) = \frac{\delta}{W_A(\Omega)}, \tag{47}$$

for the error function. This can be seen in Fig. 4(d) where the error function of Ex. 2-2 keeps the tolerance $\delta(\Omega)$ which is additionally given there. Consequently, the error magnitude at the beginning of the passband is ten times as large as at its end Ω_p. This explains the large deviations of the magnitude response and the group delay from the desired functions in Fig. 4(a) and 4(b) near $\Omega = 0$.

Again the error ε_∞ is considerably reduced by an increase of ε_2. The results are summarized in Table 2.

4 Discussion

The examples in the previous section show the flexibility of our new algorithm. The first example shows that nearly an L_∞ or an L_2 solution result as a special case of our method. It is a possible alternative to the methods described in [2, 5, 9, 10, 12]. Furthermore filters with less ε_∞ than in the L_2 case can be designed by an increase of ε_2 and vice versa in the L_∞ case. Our method can be seen as a generalization of Adams' algorithm [1] to the NLPFIR filter case with improved convergence behaviour (KT coefficients) and the additional possibility to choose arbitrary weighting and desired functions due to the choice of a discrete frequency grid. The presented nonlinear phase approach has advantages compared to the linear phase case [1] in a similar way as was shown in [5] for the L_∞ design, which is verified by Ex. 1-7.

Convergence problems arise if the Toeplitz matrices $\mathbf{R}$, $\tilde{\mathbf{R}}$ in the Jacobian matrix $\mathbf{J}$ which are known to be able to cause numerical problems [13] have a large condition number. This case especially occurs when the ratio $\max(W_E(\Omega))/\min(W_E(\Omega))$ is large. In many cases this problem can be circumvented realizing the fact that an increase of $W_E(\Omega)$ often is not necessary since it cannot improve the design as can be seen with Ex. 1-3. The passband shows equiripple behaviour, so that larger values of $W_E(\Omega)$ in the stopband can only marginally reduce ε_2 in the stopband. A MATLAB version of the described algorithm can be obtained from the authors.

Acknowledgement

The authors thank Dr. T. Laakso for his careful proofreading and help in preparing the final version of this paper.

References

[1] Adams, John W. Fir digital filters with least-squares stopbands subject to peak-gain constraints. *IEEE Transactions on Circuits and Systems*, 39:376–388, April 1991.

[2] Alkhairy, A., K. Christian, and J. Lim. Design of fir filters by complex chebyshev approximation. In *International Conference on ASSP*, pages 1985–1988, 1991.

[3] Bamberger, Joachim. *Entwurf von Fensterfunktionen und nichtrekursiven Filtern hinsichtlich eines quadratischen Fehlermaßes.* Diplomarbeit, Universität Erlangen-Nürnberg, 1992.

[4] Burrus, C. S. and J. A. Barreto. Least p-power error design of fir filters. In *IEEE International Symposium on Circuits and Systems*, pages 545–548, 1992.

[5] Chen, X. and T. W. Parks. Design of fir filters in the complex domain. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35:144–153, 1987.

[6] Horbach, Ulrich, and Markus Lang. Design and implementation of sigma-delta d/a-converters with optimized loop elements. In *International Symposium on Circuits and Systems*, pages 1633–1636, 1991.

[7] Lang, Markus, and Joachim Bamberger. Nonlinear phase fir filter design with minimum ls error and additional constraints. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages III–57–III–60, 1993.

[8] Parks, T. W. and J. H. McClellan. Chebyshev approximation for nonrecursive digital filters with linear phase. *IEEE Transactions on Circuit Theory*, 19:189–194, 1972.

[9] Pei, S. and J.Shyu. Eigen-approach for designing fir filters and all-pass phase equalizers with prescribed magnitude and phase response. *IEEE Transactions on Circuits and Systems II*, 39:137–146, March 1992.

[10] Preuss, Klaus. On the design of fir filters by complex chebyshev approximation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37:702–712, May 1989.

[11] Rabiner, L. R. Linear program design of finite impulse response (fir) digital filters. *IEEE Transactions on Audio and Electroacoustics*, 20:280–288, October 1972.

[12] Schulist, Matthias. *Ein Beitrag zum Entwurf nichtrekursiver Filter.* Dissertation, Universität Erlangen-Nürnberg, 1992.

[13] Strang, Gilbert. *Introduction to applied mathematics.* Wellesley, Wellesley, 1986.

[14] Tang, P. T. P. A fast algorithm for linear complex chebyshev approximations. *Mathematics of Computation*, 51:721–739, 1988.

	k	$\delta/10^{-2}$	$W_E(\Omega_s)/W_E(\Omega_p)$	Flops/10^5	$\varepsilon_2/10^{-2}$	$\varepsilon_\infty/10^{-2}$	$\varepsilon_{2s}/10^{-4}$
Ex. 1-1	1	100.00	1.00	7.22	2.34	35.458	92.4
Ex. 1-2	12	4.3952	1.00	93.2	5.58	4.3984	41.7
Ex. 1-3	4	7.0000	$1.00 \cdot 10^4$	18.8	10.0	7.0000	8.27
Ex. 1-4	4	7.0000	1.00	18.0	3.40	7.0000	24.5
Ex. 1-5	1	100.00	3.96	7.22	2.93	6.5179	21.0
Ex. 1-6	484	–	–	723.	5.58	4.3952	42.0
Ex. 1-7	4	7.0000	$1.00 \cdot 10^4$	19.2	13.6	7.0000	18.8

Table 1:

	k	$\delta/10^{-2}$	$W_E(\Omega_s)/W_E(\Omega_p)$	Flops$/10^6$	$\varepsilon_2/10^{-3}$	$\varepsilon_\infty/10^{-2}$	$\varepsilon_{2s}/10^{-4}$
Ex. 2-1	1	100.00	1.00	1.04	3.40	16.494	16.3
Ex. 2-2	7	5.0000	$1.00 \cdot 10^3$	9.22	23.4	5.0000	1.91

Table 2:

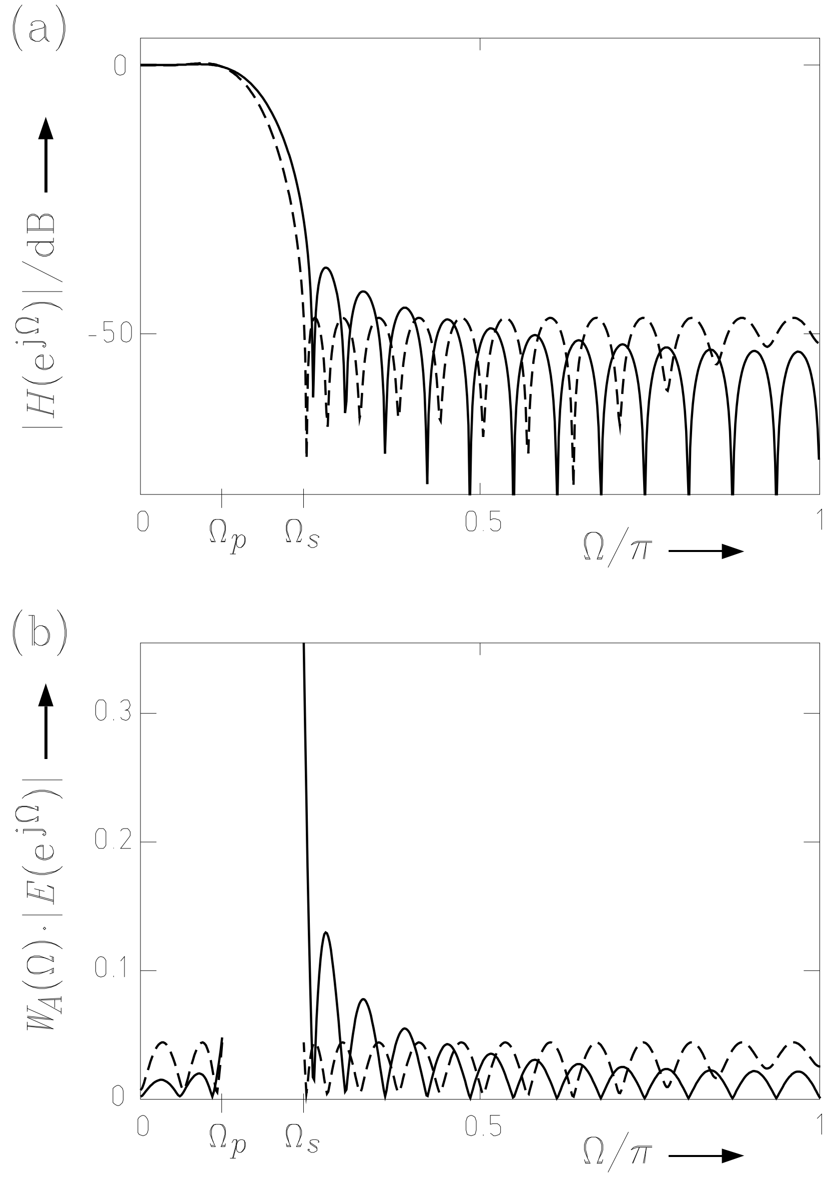
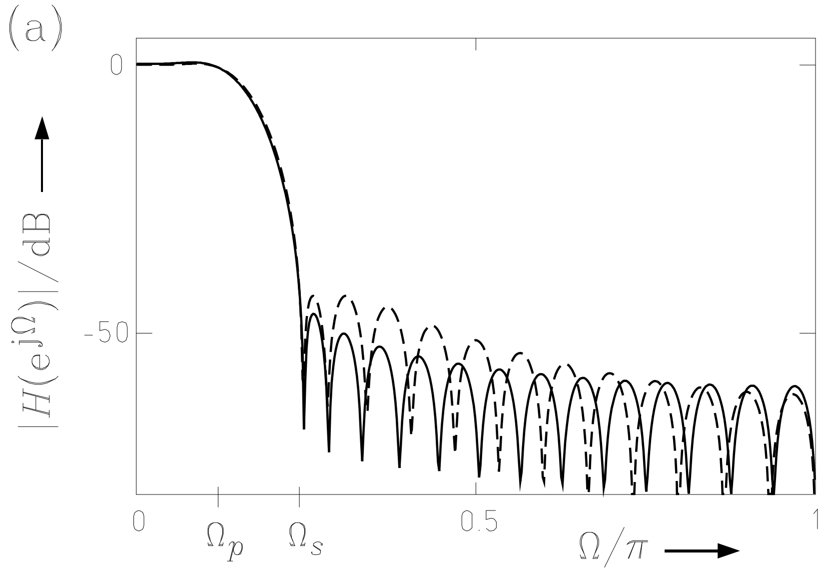
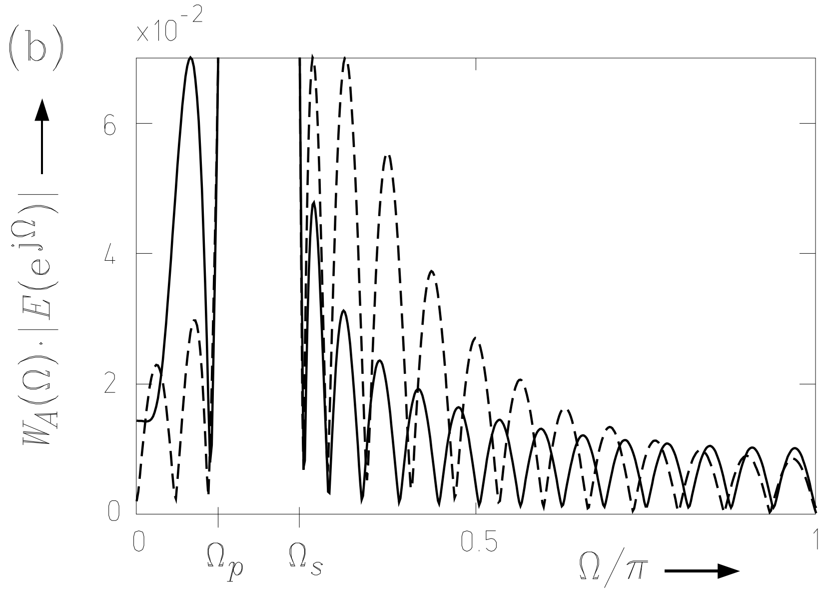

(a)
0
-50
$|H(e^{j\Omega})|$/dB
0
Ω_p
Ω_s
0.5
1
Ω/π
(b)
0.3
0.2
0.1
0
$W_A(\Omega) \cdot |E(e^{j\Omega})|$
0
Ω_p
Ω_s
0.5
1
Ω/π

Figure 1:



(a)
0
-50
$|H(e^{j\Omega})|$/dB
0
Ω_p
Ω_s
0.5
1
Ω/π



(b)
x10^-2
6
4
2
0
$W_A(\Omega)\cdot|E(e^{j\Omega})|$
0
Ω_p
Ω_s
0.5
1
Ω/π

Figure 2:

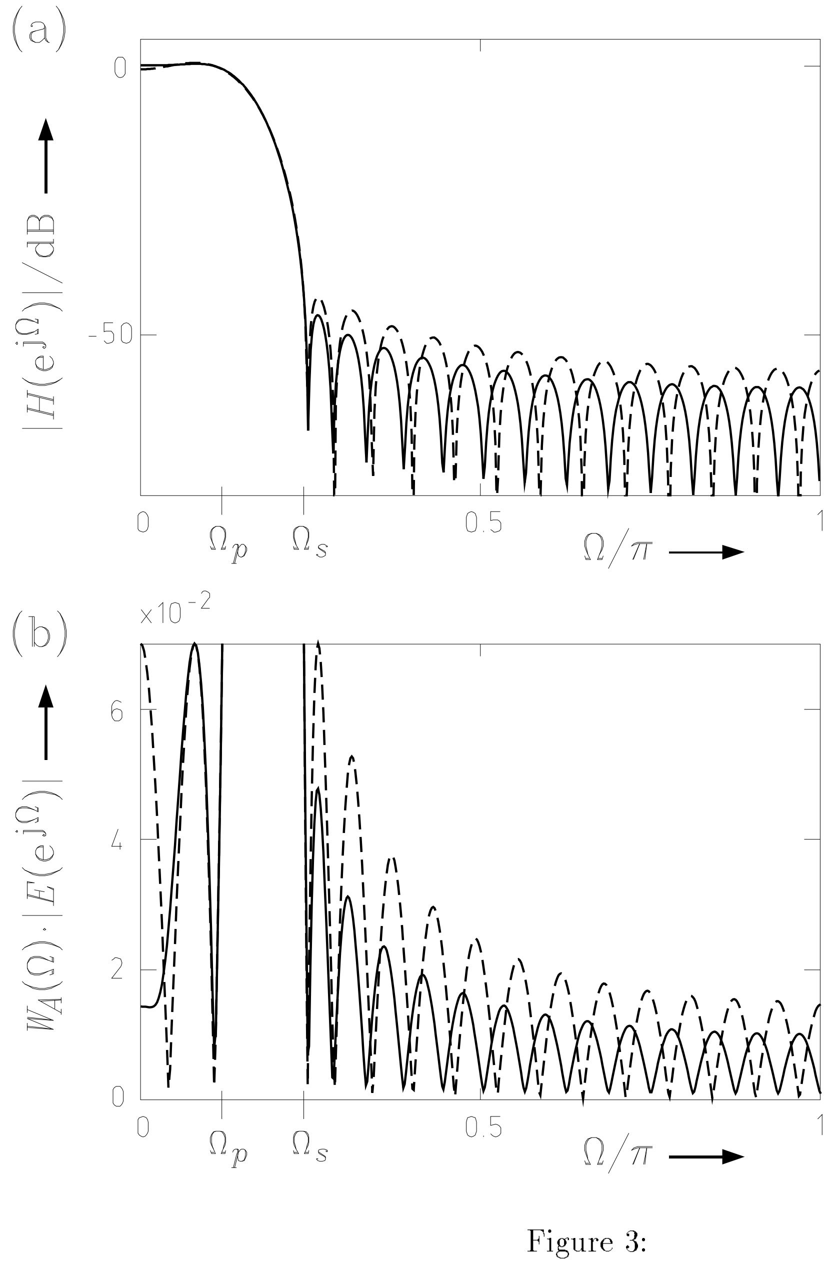

(a)
$|H(\mathrm{e}^{\mathrm{j}\Omega})|/\mathrm{dB}$
0
-50
0
Ω_p
Ω_s
0.5
1
Ω/π
(b)
x10^-2
$W_A(\Omega)\cdot|E(\mathrm{e}^{\mathrm{j}\Omega})|$
6
4
2
0
0
Ω_p
Ω_s
0.5
1
Ω/π

Figure 3:

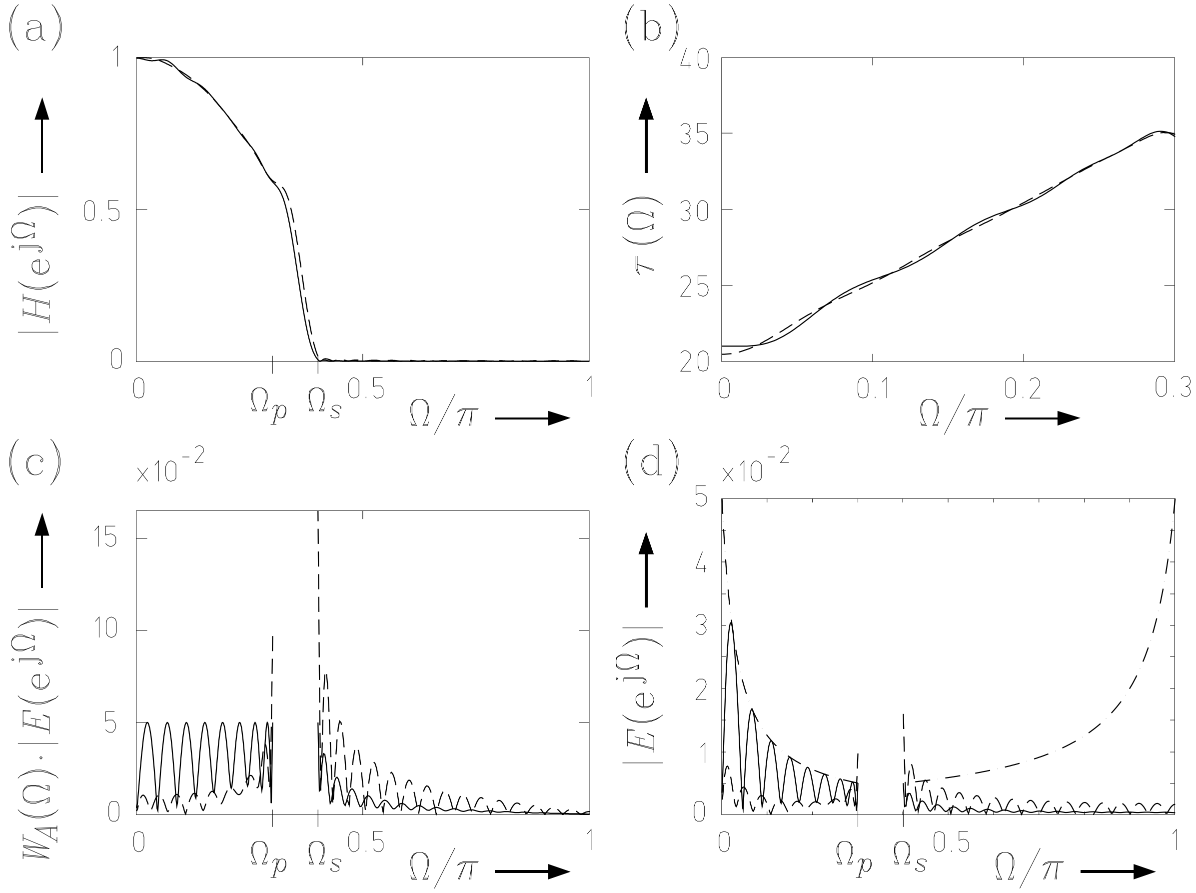

(a)
1
0.5
0
$|H(e^{j\Omega})|$
0
Ω_p
Ω_s
0.5
1
Ω/π
(b)
40
35
30
25
20
$\tau(\Omega)$
0
0.1
0.2
0.3
Ω/π
(c)
$\times 10^{-2}$
15
10
5
0
$W_A(\Omega)\cdot|E(e^{j\Omega})|$
0
Ω_p
Ω_s
0.5
1
Ω/π
(d)
$\times 10^{-2}$
5
4
3
2
1
0
$|E(e^{j\Omega})|$
0
Ω_p
Ω_s
0.5
1
Ω/π

Figure 4: