Computer-Aided Design of Digital Filters

- The FIR filter design techniques discussed so far can be easily implemented on a computer
- In addition, there are a number of FIR filter design algorithms that rely on some type of optimization techniques that are used to minimize the error between the desired frequency response and that of the computer-generated filter
- Basic idea behind the computer-based iterative technique
- Let $H(e^{j\omega})$ denote the frequency response of the digital filter $H(z)$ to be designed approximating the desired frequency response $D(e^{j\omega})$, given as a piecewise linear function of ω, in some sense

- Objective - Determine iteratively the coefficients of $H(z)$ so that the difference between between $H(e^{j\omega})$ and $D(e^{j\omega})$ over closed subintervals of $0 \le \omega \le \pi$ is minimized
- This difference usually specified as a weighted error function

$$\mathcal{E}(\omega) = W(e^{j\omega})[H(e^{j\omega}) - D(e^{j\omega})]$$

where $W(e^{j\omega})$ is some user-specified weighting function

- **Chebyshev** or **minimax criterion** - Minimizes the peak absolute value of the weighted error:

$$\boxed{\varepsilon = \max_{\omega \in R} |\mathcal{E}(\omega)|}$$

where R is the set of disjoint frequency bands in the range $0 \le \omega \le \pi$, on which $D(e^{j\omega})$ is defined

Design of Equiripple Linear-Phase FIR Filters

- The linear-phase FIR filter obtained by minimizing the peak absolute value of

$$\varepsilon = \max_{\omega \in R} |\mathcal{E}(\omega)|$$

is usually called the **equiripple FIR filter**

- After ε is minimized, the weighted error function $\mathcal{E}(\omega)$ exhibits an equiripple behavior in the frequency range R
- The general form of frequency response of a causal linear-phase FIR filter of length $2M+1$:

$$\boxed{H(e^{j\omega}) = e^{-jM\omega} e^{j\beta} \breve{H}(\omega)}$$

where the amplitude response $\breve{H}(\omega)$ is a real function of ω

- Weighted error function is given by

$$\mathcal{E}(\omega) = W(\omega)[\breve{H}(\omega) - D(\omega)]$$

where $D(\omega)$ is the desired amplitude response and $W(\omega)$ is a positive weighting function

- **Parks-McClellan Algorithm** - Based on iteratively adjusting the coefficients of $\breve{H}(\omega)$ until the peak absolute value of $\mathcal{E}(\omega)$ is minimized
- If peak absolute value of $\mathcal{E}(\omega)$ in a band $\omega_a \le \omega \le \omega_b$ is ε_o, then the absolute error satisfies

$$\left|\breve{H}(\omega) - D(\omega)\right| \le \frac{\varepsilon_o}{|W(\omega)|}, \quad \omega_a \le \omega \le \omega_b$$

- For filter design,

$$D(\omega) = \begin{cases} 1, & \text{in the passband} \\ 0, & \text{in the stopband} \end{cases}$$

- $\breve{H}(\omega)$ is required to satisfy the above desired response with a ripple of $\pm\delta_p$ in the passband and a ripple of δ_s in the stopband
- Thus, weighting function can be chosen either as

$$\boxed{W(\omega) = \begin{cases} 1, & \text{in the passband} \\ \delta_p/\delta_s, & \text{in the stopband} \end{cases}}$$

or

$$\boxed{W(\omega) = \begin{cases} \delta_s/\delta_p, & \text{in the passband} \\ 1, & \text{in the stopband} \end{cases}}$$

- **Type 1 FIR Filter** - $\boxed{\breve{H}(\omega) = \sum_{k=0}^{M} a[k]\cos(\omega k)}$

  where

  $a[0] = h[M], \ a[k] = 2h[M-k], \ 1 \le k \le M$
- **Type 2 FIR filter** -

  $$\boxed{\breve{H}(\omega) = \sum_{k=1}^{(2M+1)/2} b[k]\cos\left(\omega(k-\tfrac{1}{2})\right)}$$

  where

  $b[k] = 2h[\tfrac{2M+1}{2} - k], \ 1 \le k \le \tfrac{2M+1}{2}$
- **Type 3 FIR Filter** - $\boxed{\breve{H}(\omega) = \sum_{k=1}^{M} c[k]\sin(\omega k)}$

  where

  $c[k] = 2h[M-k], \ 1 \le k \le M$
- **Type 4 FIR Filter** -

  $$\boxed{\breve{H}(\omega) = \sum_{k=1}^{(2M+1)/2} d[k]\sin\left(\omega(k-\tfrac{1}{2})\right)}$$

  where

  $d[k] = 2h[\tfrac{2M+1}{2} - k], \ 1 \le k \le \tfrac{2M+1}{2}$

- Amplitude response for all 4 types of linear-phase FIR filters can be expressed as

$$\breve{H}(\omega) = Q(\omega)A(\omega)$$

where

$$\boxed{Q(\omega) = \begin{cases} 1, & \text{for Type 1} \\ \cos(\omega/2), & \text{for Type 2} \\ \sin(\omega), & \text{for Type 3} \\ \sin(\omega/2), & \text{for Type 4} \end{cases}}$$

and

$$\boxed{A(\omega) = \sum_{k=0}^{L} \tilde{a}[k]\cos(\omega k)}$$

where

$$\tilde{a}[k] = \begin{cases} a[k], & \text{for Type 1} \\ \tilde{b}[k], & \text{for Type 2} \\ \tilde{c}[k], & \text{for Type 3} \\ \tilde{d}[k], & \text{for Type 4} \end{cases}$$

with

$$L = \begin{cases} M, & \text{for Type 1} \\ \frac{2M-1}{2}, & \text{for Type 2} \\ M-1, & \text{for Type 3} \\ \frac{2M-1}{2}, & \text{for Type 4} \end{cases}$$

$\tilde{b}[k]$, $\tilde{c}[k]$, and $\tilde{d}[k]$, are related to $b[k]$, $c[k]$, and $d[k]$, respectively

- Modified form of weighted error function

$$\begin{aligned} \mathcal{E}(\omega) &= W(\omega)[Q(\omega)A(\omega) - D(\omega)] \\ &= W(\omega)Q(\omega)[A(\omega) - \tfrac{D(\omega)}{Q(\omega)}] \\ &\boxed{= \tilde{W}(\omega)[A(\omega) - \tilde{D}(\omega)]} \end{aligned}$$

where we have used the notation

$$\tilde{W}(\omega) = W(\omega)Q(\omega)$$

$$\tilde{D}(\omega) = D(\omega)/Q(\omega)$$

- Optimization Problem - Determine $\tilde{a}[k]$ which minimize the peak absolute value ε of

$$\boxed{\mathcal{E}(\omega) = \tilde{W}(\omega)[\sum_{k=0}^{L} \tilde{a}[k]\cos(\omega k) - \tilde{D}(\omega)]}$$

over the specified frequency bands $\omega \in R$

- After $\tilde{a}[k]$ has been determined, corresponding coefficients of the original $A(\omega)$ are computed from which $h[n]$ are determined

- Alternation Theorem - $A(\omega)$ is the best unique approximation of $\tilde{D}(\omega)$ obtained by minimizing peak absolute value ε of

$$\boxed{\mathcal{E}(\omega) = W(\omega)[Q(\omega)A(\omega) - D(\omega)]}$$

if and only if there exist at least $L+2$ extremal frequencies, $\{\omega_i\}$, $0 \le i \le L+1$, in a closed subset R of the frequency range $0 \le \omega \le \pi$ such that $\omega_0 < \omega_1 < \cdots < \omega_L < \omega_{L+1}$ and $\mathcal{E}(\omega_i) = -\mathcal{E}(\omega_{i+1})$, $|\mathcal{E}(\omega_i)| = \varepsilon$ for all i

- Consider a Type 1 FIR filter with an amplitude response $A(\omega)$ whose approximation error $\mathcal{E}(\omega)$ satisfies the Alternation Theorem
- Peaks of $\mathcal{E}(\omega)$ are at $\omega = \omega_i$, $0 \le i \le L+1$ where $d\mathcal{E}(\omega)/d\omega = 0$
- Since in the passband and stopband, $\tilde{W}(\omega)$ and $\tilde{D}(\omega)$ are piecewise constant,

$$\frac{d\mathcal{E}(\omega)}{d\omega} = \frac{dA(\omega)}{d\omega} = 0 \quad \text{at } \omega = \omega_i$$

- Using $\cos(\omega k) = T_k(\cos\omega)$, where $T_k(x)$ is the k-th order Chebyshev polynomial

$$T_k(x) = \cos(k\cos^{-1} x)$$

- $A(\omega)$ can be expressed as

$$\boxed{A(\omega) = \sum_{k=0}^{L} \alpha[k](\cos\omega)^k}$$

which is an Lth-order polynomial in $\cos\omega$

- Hence, $A(\omega)$ can have at most $L-1$ local minima and maxima inside specified passband and stopband

- At bandedges, $\omega = \omega_p$ and $\omega = \omega_s$, $|\mathcal{E}(\omega)|$ is a maximum, and hence $A(\omega)$ has extrema at these points
- $A(\omega)$ can have extrema at $\omega = 0$ and $\omega = \pi$
- Therefore, there are at most $L+3$ extremal frequencies of $\mathcal{E}(\omega)$
- For linear-phase FIR filters with K specified bandedges, there can be at most $L+K+1$ extremal frequencies

---

- The set of equations

$$\boxed{\tilde{W}(\omega_i)[A(\omega_i) - \tilde{D}(\omega_i)] = (-1)^i \varepsilon, \; 0 \le i \le L+1}$$

is written in a matrix form

$$\begin{bmatrix} 1 & \cos(\omega_0) & \cdots & \cos(L\omega_0) & -1/\tilde{W}(\omega_0) \\ 1 & \cos(\omega_1) & \cdots & \cos(L\omega_1) & 1/\tilde{W}(\omega_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \cos(\omega_L) & \cdots & \cos(L\omega_L) & (-1)^{L-1}/\tilde{W}(\omega_L) \\ 1 & \cos(\omega_{L+1}) & \cdots & \cos(L\omega_{L+1}) & (-1)^{L}/\tilde{W}(\omega_{L+1}) \end{bmatrix} \begin{bmatrix} \tilde{a}[0] \\ \tilde{a}[1] \\ \vdots \\ \tilde{a}[L] \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \tilde{D}(\omega_0) \\ \tilde{D}(\omega_1) \\ \vdots \\ \tilde{D}(\omega_L) \\ \tilde{D}(\omega_{L+1}) \end{bmatrix}$$

- The matrix equation can be solved for the unknowns $\tilde{a}[i]$ and ε if the locations of the $L+2$ extremal frequencies are known a priori
- The **Remez exchange algorithm** is used to determine the locations of the extremal frequencies

Remez Exchange Algorithm

- Step 1: A set of initial values of extremal frequencies are either chosen or are available from completion of previous stage
- Step 2: Value of ε is computed using

$$\boxed{\varepsilon = \frac{c_0\tilde{D}(\omega_0) + c_1\tilde{D}(\omega_1) + \cdots + c_{L+1}\tilde{D}(\omega_{L+1})}{\frac{c_0}{\tilde{W}(\omega_0)} - \frac{c_1}{\tilde{W}(\omega_1)} + \cdots + \frac{(-1)^{L+1}c_{L+1}}{\tilde{W}(\omega_{L+1})}}}$$

where $c_n = \prod_{\substack{i=0\\ i\neq n}}^{L+1} \frac{1}{\cos(\omega_n) - \cos(\omega_i)}$ By solving the matrix eq.

- Step 3: Values of $A(\omega)$ at $\omega = \omega_i$ are then computed using

$$A(\omega_i) = \frac{(-1)^i\varepsilon}{\tilde{W}(\omega_i)} + \tilde{D}(\omega_i), \quad 0 \le i \le L+1$$

- Step 4: The polynomial $A(\omega)$ is determined by interpolating the above values at the L+2 extremal frequencies using the Lagrange interpolation formula

---

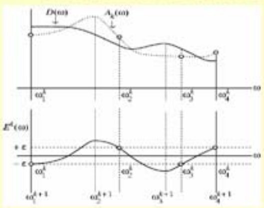
- Step 4: The new error function

$$\boxed{\mathcal{E}(\omega) = \tilde{W}(\omega)[A(\omega) - \tilde{D}(\omega)]}$$

is computed at a dense set S ($S \ge L$) of frequencies. In practice $S = 16L$ is adequate. Determine the L+2 new extremal frequencies from the values of $\mathcal{E}(\omega)$ evaluated at the dense set of frequencies.

- Step 5: If the peak values ε of $\mathcal{E}(\omega)$ are equal in magnitude, algorithm has converged. Otherwise, go back to Step 2.

- Illustration of algorithm



Iteration process is stopped if the difference between the values of the peak absolute errors between two consecutive stages is less than a preset value, e.g., 10^{-6}

- Example - Approximate the desired function $D(x) = 1.1x^2 - 0.1$ defined for the range $0 \le x \le 2$ by a linear function $a_1x + a_0$ by minimizing the peak value of the absolute error

$$\boxed{\max_{x\in[0,2]} \left|1.1x^2 - 0.1 - a_0 - a_1x\right|}$$

- Stage 1:

Choose arbitrarily the initial extremal points

$$x_1 = 0,\ x_2 = 0.5,\ x_3 = 1.5$$

- Solve the three linear equations

$$a_0 + a_1x_\ell - (-1)^\ell\varepsilon = D(x_\ell), \quad \ell = 1, 2, 3$$

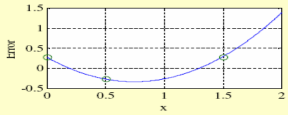
i.e.,

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0.5 & -1 \\ 1 & 1.5 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} -0.1 \\ 0.175 \\ 2.375 \end{bmatrix}$$

for the given extremal points yielding

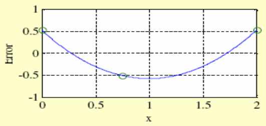
$$a_0 = -0.375,\ a_1 = 1.65,\ \varepsilon = 0.275$$

- Plot of $\mathcal{E}_1(x) = 1.1x^2 - 1.65x + 0.275$ along with values of error at chosen extremal points shown below



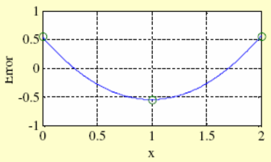
- Note: Errors are equal in magnitude and alternate in sign
- Stage 2:
- Choose extremal points where $\mathcal{E}_1(x)$ assumes its maximum absolute values
- These are $x_1 = 0,\ x_2 = 0.75,\ x_3 = 2$
- New values of unknowns are obtained by solving
$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0.75 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} -0.1 \\ 0.5188 \\ 4.3 \end{bmatrix}$$
yielding $a_0 = -0.6156,\ a_1 = 2.2,\ \varepsilon = 0.5156$

- Plot of $\mathcal{E}_2(x) = 1.1x^2 - 2.2x + 0.5156$ along with values of error at chosen extremal points shown below



- Stage 3:
- Choose extremal points where $\mathcal{E}_2(x)$ assumes its maximum absolute values
- These are $x_1 = 0,\ x_2 = 1,\ x_3 = 2$
- New values of unknowns are obtained by solving
$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} -0.1 \\ 1.0 \\ 4.3 \end{bmatrix}$$
yielding $a_0 = -0.65,\ a_1 = 2.2,\ \varepsilon = 0.55$

- Plot of $\mathcal{E}_3(x) = 1.1x^2 - 2.2x + 0.55$ along with values of error at chosen extremal points shown below



- Algorithm has converged as ε is also the maximum value of the absolute error