FIR Filter Design in the Frequency Domain

Window design and minimax (equiripple) design + spectral transformations

Digitial Signal Processing with a focus on audio signals

Fabio Antonacci

February 18

Design specs: what we require in frequency domain

We design a **linear-phase FIR** lowpass filter with magnitude constraints:

$$|H(e^{j\omega})| \approx \begin{cases} 1, & \omega \in [0, \wp] \quad \text{(bright band / passband)} \\ 0, & \omega \in [\omega_s, \pi] \quad \text{(dark band / stopband)} \end{cases}$$

with a transition band $\omega \in (\wp, \omega_s)$.

Typical ripple/attenuation constraints:

$$\big|\,|H(e^{j\omega})| - 1\,\big| \le \delta_p, \quad \omega \in [0, \wp] \qquad |H(e^{j\omega})| \le \delta_s, \quad \omega \in [\omega_s, \pi].$$

Equivalently, in dB:

$$A_p = -20\log_{10}\left(\frac{1-\delta_p}{1+\delta_p}\right) \approx 20\log_{10}\left(\frac{1+\delta_p}{1-\delta_p}\right), \qquad A_s = -20\log_{10}(\delta_s).$$

Cutoff frequency and transition width

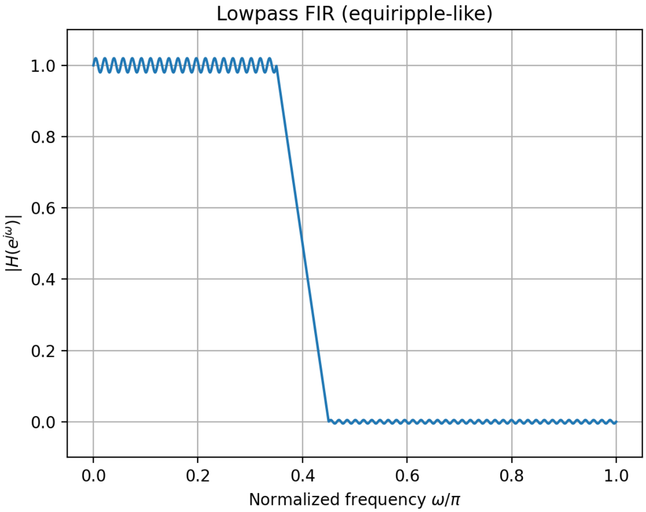
Given $\wp$ and ω_s, a common nominal cutoff is

$$\omega_c = \frac{\wp + \omega_s}{2}, \qquad \Delta\omega = \omega_s - \wp \quad \text{(transition width).}$$

For **window-based** design, $\Delta\omega$ is mainly controlled by the FIR length $N + 1$ and by the chosen window (rectangular, Hann, Hamming, Blackman, ...).

However, window design typically **does not optimize** ripple in the minimax sense: it provides a convenient, closed-form approximation but not the best equiripple solution.

Lowpass design



Recall: linear-phase FIR parameterization

Assume an FIR of length $N+1$, impulse response $h[n]$, $n = 0, \ldots, N$.

Linear phase is obtained with symmetric coefficients:

$$h[n] = h[N-n].$$

Then the frequency response can be written as

$$H(e^{j\omega}) = e^{-j\omega N/2} A(\omega),$$

where $A(\omega)$ is real and even.

For **Type I** (odd length, N even), we can express

$$A(\omega) = a_0 + 2\sum_{k=1}^{M} a_k \cos(k\omega), \qquad M = \frac{N}{2}.$$

Designing H reduces to approximating the desired magnitude using a cosine series.

Window-based lowpass design (frequency-domain view)

Start from the **ideal** brick-wall response:

$$H_d(e^{j\omega}) = \begin{cases} 1, & |\omega| \le \omega_c, \\ 0, & \omega_c < |\omega| \le \pi. \end{cases}$$

Its impulse response is the (non-causal, infinite) sinc:

$$h_d[n] = \frac{\omega_c}{\pi} \operatorname{sinc}\left(\frac{\omega_c}{\pi} \left(n - \frac{N}{2} \right) \right), \quad \text{with } \operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}.$$

A finite FIR is obtained by truncation / windowing:

$$h[n] = h_d[n]\, w[n], \qquad n = 0, \ldots, N,$$

where $w[n]$ is the chosen window.

Windowing interpreted in frequency domain

Multiplication in time corresponds to convolution in frequency:

$$H(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\theta})\, W(e^{j(\omega-\theta)})\, d\theta.$$

Consequences:

- The sharp discontinuity of H_d is **smoothed** by W, producing a transition band.
- Ripples are shaped by sidelobes of W (Gibbs phenomenon).
- Window choice trades **mainlobe width** (transition width) vs **sidelobe level** (stopband attenuation).

Window design is easy and intuitive, but it is **not minimax-optimal** for given band edges and weights.

From specs to an optimization problem

We now pose an **optimal** design problem.

Define a desired *amplitude* response $D(\omega)$ and a weight $W(\omega) \geq 0$:

$$D(\omega) = \begin{cases} 1, & \omega \in [0, \wp], \\ 0, & \omega \in [\omega_s, \pi]. \end{cases} \qquad W(\omega) = \begin{cases} W_p, & \omega \in [0, \wp], \\ W_s, & \omega \in [\omega_s, \pi]. \end{cases}$$

Design variables: cosine-series coefficients $\{a_k\}$ (hence $A(\omega)$).

Define weighted error:

$$E(\omega) = W(\omega) \left(A(\omega) - D(\omega) \right).$$

Minimax (Chebyshev) criterion:

$$\min_{\{a_k\}} \max_{\omega \in [0,\wp] \cup [\omega_s, \pi]} |E(\omega)|.$$

This yields **equiripple** (equal-ripple) FIR filters.

Connecting ripple specs to weights

If the design is equiripple, the weighted error magnitude is (approximately) constant:

$$\max |E(\omega)| \approx \epsilon.$$

Hence, in each band:

$$|A(\omega) - 1| \le \frac{\epsilon}{W_p} \approx \delta_p, \qquad |A(\omega) - 0| \le \frac{\epsilon}{W_s} \approx \delta_s.$$

Therefore, weight ratio controls ripple ratio:

$$\frac{\delta_p}{\delta_s} \approx \frac{W_s}{W_p}.$$

Practical rule:

$$W_p = 1, \qquad W_s = \frac{\delta_p}{\delta_s}.$$

(or scaled versions; only ratios matter).

Why minimax yields equiripple: alternation principle

For a linear approximation space (cosine series) and Chebyshev norm, the optimal solution satisfies the **alternation theorem**:

There exist at least $M + 2$ frequencies $\{\omega_i\}$ in the union of design bands such that

$$E(\omega)(\omega_i) = (-1)^i \epsilon, \qquad i = 0, 1, \ldots, M + 1,$$

i.e., the weighted error **alternates in sign** and hits the same magnitude ϵ.

Interpretation:

- ripples have equal height in the **weighted** sense,
- the solution is the best possible under the minimax criterion for the chosen order.

Remez exchange algorithm (Parks–McClellan)

The standard way to compute the minimax FIR is the **Remez exchange** algorithm (also known as Parks–McClellan for FIR equiripple filters).

High-level iteration:

1. Choose an initial set of extremal frequencies $\{\omega_i\}$ (size $M + 2$).
2. Solve for $\{a_k\}$ and ϵ such that $E(\omega)(\omega_i) = (-1)^i \epsilon$.
3. Evaluate $E(\omega)(\omega)$ on the bands and find new extremal points where $|E(\omega)|$ is maximal.
4. Replace the old $\{\omega_i\}$ with the new set and repeat until convergence.

Output: an **equiripple** linear-phase FIR meeting the minimax optimum for the chosen order.

Solving the interpolation step

With

$$A(\omega) = a_0 + 2\sum_{k=1}^{M} a_k \cos(k\omega),$$

and constraints at extremal frequencies ω_i:

$$W(\omega)_i \left(A(\omega_i) - D(\omega_i)\right) = (-1)^i \epsilon,$$

we get a linear system in unknowns $\{a_k\}$ and ϵ.

This step is **linear** (given ω_i), and the exchange step updates the ω_i.

Numerical note: Many implementations use a barycentric/interpolation form to improve stability, but the underlying idea remains: enforce equal alternating weighted error at $M + 2$ points.

Design workflow (recommended in practice)

1. Translate amplitude specs to $\wp, \omega_s, \delta_p, \delta_s$ (or A_p, A_s).
2. Choose an FIR length $N+1$ (or iterate on N until specs are met).
3. Set weights (e.g., $W_p = 1$, $W_s = \delta_p/\delta_s$).
4. Run minimax design (Remez / Parks–McClellan) to obtain $h[n]$.
5. Verify:
   - passband ripple $\leq \delta_p$,
   - stopband magnitude $\leq \delta_s$,
   - transition band between $\wp$ and ω_s.

If specs are violated, increase order (larger N) or relax $\Delta\omega$ / ripple constraints.

Worked example (specification table)

Example lowpass FIR (normalized rad/sample):

Parameter	**Symbol**	**Value**
Passband edge	$\wp$	0.35π
Stopband edge	ω_s	0.45π
Passband ripple	δ_p	0.01
Stopband ripple	δ_s	0.001
Weights	(W_p, W_s)	$(1,\ \delta_p/\delta_s = 10)$

Choose N (e.g., $N = 60$) and compute an equiripple FIR using minimax optimization.

(In the notebook: show magnitude response and ripple zoom in pass/stop bands.)

Comparing window vs minimax (qualitative)

- **Window method**:
  - simple closed-form $h[n] = h_d[n]w[n]$,
  - ripple/attenuation are indirect consequences of window sidelobes,
  - not optimal for a fixed order.
- **Minimax (equiripple)**:
  - directly optimizes worst-case error with weights,
  - typically achieves **smallest** maximum ripple for a given order,
  - yields an **equiripple** error pattern (alternating extrema).

Key takeaway: minimax design is the go-to when you need the **best worst-case performance** given strict ripple constraints.

Implementation note (Python/Matlab/Octave)

Common functions:

- Python (SciPy): `scipy.signal.remez` / `fir_pm` (depending on version)
- MATLAB: `firpm` (Parks–McClellan)
- Octave: `remez` / `firpm` (package dependent)

Typical call structure (conceptually):

$$h = \mathrm{remez}(N,\ [0, \wp, \omega_s, \pi],\ [1, 1, 0, 0],\ \text{weights}).$$

(Then verify: $|H(e^{j\omega})|$ and ripple bounds.)

From lowpass prototype to other filters

A lowpass FIR is a convenient **prototype**. Many useful responses (highpass, bandpass, bandstop) can be obtained by **spectral transformations**.

Two common approaches (FIR, linear phase preserved):

- **Frequency shifting / modulation**: move the spectrum to a new center.
- **Spectral inversion / complement**: swap passband and stopband.

We focus on transformations that are easy to implement in time domain and keep linear phase.

Highpass from lowpass via spectral inversion

Let $H_{\mathrm{LP}}(e^{j\omega})$ be a lowpass prototype with cutoff ω_c. A highpass with the same transition width can be obtained by

$$H_{\mathrm{HP}}(e^{j\omega}) = 1 - H_{\mathrm{LP}}(e^{j\omega}).$$

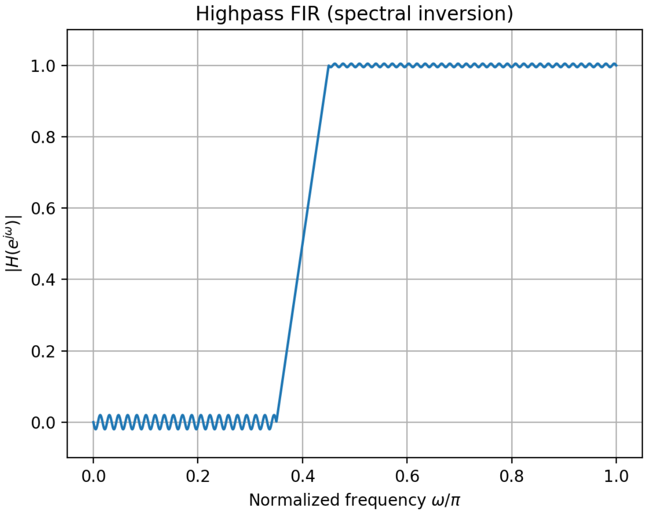
In time domain:

$$h_{\mathrm{HP}}[n] = \delta\left[n - \frac{N}{2}\right] - h_{\mathrm{LP}}[n].$$

Notes:

- Requires $\frac{N}{2} \in \mathbb{Z}$ (Type I linear phase, N even), so the centered delta is on a sample.
- Passband/stopband ripples are inherited (with the same weights, if the prototype is equiripple).

Highpass by spectral inversion



Adjusting the highpass cutoff

If the lowpass prototype has passband edge $\wp$ and stopband edge ω_s, then the inverted highpass has:

$$\text{HP passband: } [\omega_s, \pi], \qquad \text{HP stopband: } [0, \wp].$$

So the *same* band edges are reused, simply swapping bright/dark bands.

In minimax design, you may directly design the highpass by setting

$$D(\omega) = \begin{cases} 0, & \omega \in [0, \wp] \\ 1, & \omega \in [\omega_s, \pi] \end{cases}$$

with the same weighting logic. Spectral inversion is a fast way when a good LP prototype already exists.

Bandpass from lowpass via modulation (frequency shift)

A bandpass can be obtained by **modulating** a lowpass prototype in time domain.

Let H_{LP} be a lowpass prototype with cutoff ω_b (bandwidth parameter), and choose a desired center frequency ω_0. Define:

$$h_{\text{BP}}[n] = 2\,h_{\text{LP}}[n]\cos\left(\omega_0\left(n - \frac{N}{2}\right)\right).$$

Then, approximately:

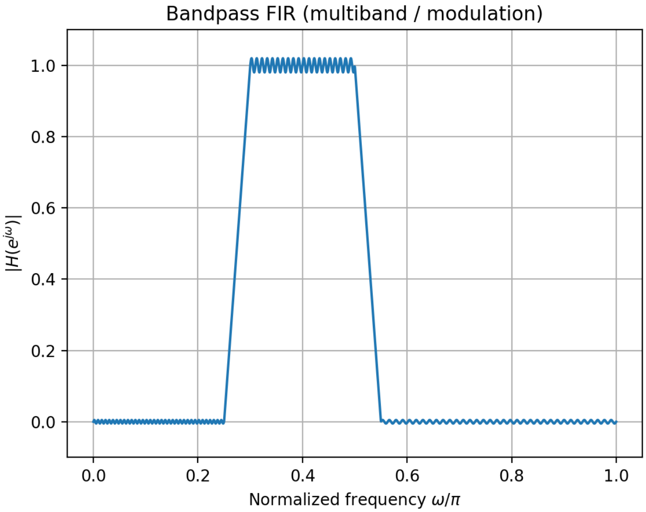
$$H_{\text{BP}}(e^{j\omega}) \approx H_{\text{LP}}(e^{j(\omega-\omega_0)}) + H_{\text{LP}}(e^{j(\omega+\omega_0)}),$$

i.e., two shifted copies (due to real modulation), producing a bandpass around $\pm\omega_0$.

Band edges (positive frequencies) are roughly:

$$\omega_{p1} \approx \omega_0 - \omega_b, \qquad \omega_{p2} \approx \omega_0 + \omega_b.$$

Bandpass FIR design



Bandstop from bandpass or via complementary response

If a bandpass H_{BP} is available, a bandstop (notch) follows by complement:

$$H_{\mathrm{BS}}(e^{j\omega}) = 1 - H_{\mathrm{BP}}(e^{j\omega}),$$

and in time domain:

$$h_{\mathrm{BS}}[n] = \delta\left[n - \frac{N}{2}\right] - h_{\mathrm{BP}}[n].$$

Alternatively, design a bandstop directly with minimax by specifying

$$D(\omega) = \begin{cases} 1, & \omega \in [0, \omega_{s1}] \cup [\omega_{s2}, \pi], \\ 0, & \omega \in [\omega_{p1}, \omega_{p2}], \end{cases}$$

with appropriate weights for each band.

Minimax design with multiple bands (generalization)

Remez / Parks–McClellan naturally supports **multiband** designs.

For a bandpass:

$$D(\omega) = \begin{cases} 0, & [0, \omega_{s1}] \\ 1, & [\omega_{p1}, \omega_{p2}] \\ 0, & [\omega_{s2}, \pi] \end{cases}$$

and piecewise weights $W(\omega)$ per band:

$$W(\omega) = \begin{cases} W_{s1}, & [0, \omega_{s1}] \\ W_{p}, & [\omega_{p1}, \omega_{p2}] \\ W_{s2}, & [\omega_{s2}, \pi]. \end{cases}$$

The minimax objective remains:

$$\min \max_{\omega \in \text{all design bands}} \left| W(\omega) \left(A(\omega) - D(\omega) \right) \right| .$$

Design workflow: prototype + transformation vs direct design

Two equivalent strategies:

(A) Prototype + transform

- Design lowpass minimax prototype with desired ripple.
- Use inversion (HP) or modulation (BP) to obtain target filter.
- Useful for intuition and fast exploration.

(B) Direct minimax multiband design

- Specify the final desired bands directly in $D(\omega)$.
- Run Remez with band edges and weights.
- Typically preferred for strict specifications (exact band edges / ripple allocation).

Summary

- Lowpass minimax FIR is a powerful **prototype**.
- **Highpass** via spectral inversion: $H_{\mathrm{HP}}(e^{j\omega}) = 1 - H_{\mathrm{LP}}(e^{j\omega})$, $h_{\mathrm{HP}}[n] = \delta[n - N/2] - h_{\mathrm{LP}}[n]$.
- **Bandpass** via modulation: $h_{\mathrm{BP}}[n] = 2h_{\mathrm{LP}}[n]\cos(\omega_0(n - N/2))$.
- **Bandstop** via complement: $h_{\mathrm{BS}}[n] = \delta[n - N/2] - h_{\mathrm{BP}}[n]$.
- Remez supports **direct** multiband minimax design with weights per band.