2

# Filter Design

# Filter Requirements and Specification

The Signal Processing Toolbox provides functions that support a range of filter design methodologies. This chapter explains how to apply the filter design tools to *Infinite Impulse Response* (IIR) and *Finite Impulse Response* (FIR) filter design problems.

The goal of filter design is to perform frequency dependent alteration of a data sequence. A possible requirement might be to remove noise above 30 Hz from a data sequence sampled at 100 Hz. A more rigorous specification might call for a specific amount of passband ripple, stopband attenuation, or transition width. A very precise specification could ask to achieve the performance goals with the minimum filter order, or it could call for an arbitrary magnitude shape, or it might require an FIR filter.

Filter design methods differ primarily in how performance is specified. For "loosely specified" requirements, as in the first case above, a Butterworth IIR filter is often sufficient. To design a fifth-order 30 Hz lowpass Butterworth filter and apply it to the data in vector `x`,

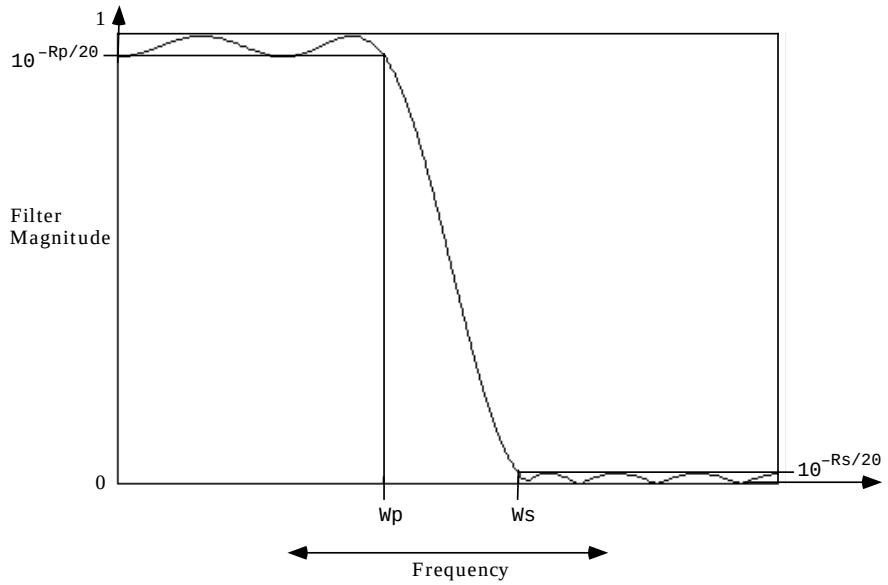
```
[b,a] = butter(5,30/50);
y = filter(b,a,x);
```

The second input argument to `butter` indicates the cutoff frequency, normalized to half the sampling frequency (the Nyquist frequency).

**Frequency Normalization in the Signal Processing Toolbox**

All of the filter design functions operate with normalized frequencies, so they do not require the system sampling rate as an extra input argument. This toolbox uses the convention that unit frequency is the Nyquist frequency, defined as half the sampling frequency. The normalized frequency, therefore, is always in the interval $0 \le f \le 1$. For a system with a 1000 Hz sampling frequency, 300 Hz is 300/500 = 0.6. To convert normalized frequency to angular frequency around the unit circle, multiply by $\pi$. To convert normalized frequency back to Hertz, multiply by half the sample frequency.

More rigorous filter requirements traditionally include passband ripple (`Rp`, in decibels), stopband attenuation (`Rs`, in decibels), and transition width (`Ws-Wp`, in Hertz).




You can design Butterworth, Chebyshev type I, Chebyshev type II, and elliptic filters that meet this type of performance specification. The toolbox order selection functions estimate the minimum filter order that meets a given set of requirements.

To meet specifications with more rigid constraints like linear phase or arbitrary filter shape, use the FIR and direct IIR filter design routines.

# IIR Filter Design

The primary advantage of IIR filters over FIR filters is that they typically meet a given set of specifications with a much lower filter order than a corresponding FIR filter. Although IIR filters have nonlinear phase, data processing within MATLAB is commonly performed "off-line," that is, the entire data sequence is available prior to filtering. This allows for a noncausal, zero-phase filtering approach (via the `filtfilt` function), which eliminates the nonlinear phase distortion of an IIR filter.

The classical IIR filters, Butterworth, Chebyshev types I and II, elliptic, and Bessel, all approximate the ideal "brickwall" filter in different ways. This toolbox provides functions to create all these types of classical IIR filters in both the analog and digital domains (except Bessel, for which only the analog case is supported), and in lowpass, highpass, bandpass, and bandstop configurations. For most filter types, you can also find the lowest filter order that fits a given filter specification in terms of passband and stopband attenuation, and transition width(s).

The direct filter design function `yulewalk` finds a filter with magnitude response approximating a desired function. This is one way to create a multiband bandpass filter.

You can also use the parametric modeling or system identification functions to design IIR filters. These functions are discussed in the "Parametric Modeling" section of Chapter 4.

The generalized Butterworth design function `maxflat` is discussed in the section "Generalized Butterworth Filter Design" on page 2-15.

The following table summarizes the various filter methods in the toolbox and lists the functions available to implement these methods:

| Method | Description | Functions |
|---|---|---|
| Analog Prototyping | Using the poles and zeros of a classical lowpass prototype filter in the continuous (Laplace) domain, obtain a digital filter through frequency transformation and filter discretization. | Complete design functions:<br>`besself`, `butter`, `cheby1`, `cheby2`, `ellip`<br>Order estimation functions:<br>`buttord`, `cheb1ord`, `cheb2ord`, `ellipord`<br>Lowpass analog prototype functions:<br>`besselap`, `buttap`, `cheb1ap`, `cheb2ap`, `ellipap`<br>Frequency transformation functions:<br>`lp2bp`, `lp2bs`, `lp2hp`, `lp2lp`<br>Filter discretization functions:<br>`bilinear`, `impinvar` |
| Direct Design | Design digital filter directly in the discrete domain by approximating a piecewise linear magnitude response. | `yulewalk` |
| Parametric Modeling* | Find a digital filter that approximates a prescribed time or frequency domain response. | Time-domain modeling functions:<br>`lpc`, `prony`, `stmcb`<br>Frequency-domain modeling functions:<br>`invfreqs`, `invfreqz` |
| Generalized Butterworth Design | Design lowpass Butterworth filters with more zeros than poles. | `maxflat` |

** See the System Identification Toolbox for an extensive collection of parametric modeling tools.*

## Classical IIR Filter Design Using Analog Prototyping

The principal IIR digital filter design technique this toolbox provides is based on the conversion of classical lowpass analog filters to their digital equivalents. The following sections describe how to design filters and summarize the characteristics of the supported filter types. See “Special Topics in IIR Filter Design” on page 2-37 for detailed steps on the filter design process.

### Complete Classical IIR Filter Design

You can easily create a filter of any order with a lowpass, highpass, bandpass, or bandstop configuration using the filter design functions:

| Filter Type | Design Function |
|---|---|
| Butterworth | `[b,a] = butter(n,Wn,`*options*`)`<br>`[z,p,k] = butter(n,Wn,`*options*`)`<br>`[A,B,C,D] = butter(n,Wn,`*options*`)` |
| Chebyshev type I | `[b,a] = cheby1(n,Rp,Wn,`*options*`)`<br>`[z,p,k] = cheby1(n,Rp,Wn,`*options*`)`<br>`[A,B,C,D] = cheby1(n,Rp,Wn,`*options*`)` |
| Chebyshev type II | `[b,a] = cheby2(n,Rs,Wn,`*options*`)`<br>`[z,p,k] = cheby2(n,Rs,Wn,`*options*`)`<br>`[A,B,C,D] = cheby2(n,Rs,Wn,`*options*`)` |
| Elliptic | `[b,a] = ellip(n,Rp,Rs,Wn,`*options*`)`<br>`[z,p,k] = ellip(n,Rp,Rs,Wn,`*options*`)`<br>`[A,B,C,D] = ellip(n,Rp,Rs,Wn,`*options*`)` |
| Bessel (analog only) | `[b,a] = besself(n,Wn,`*options*`)`<br>`[z,p,k] = besself(n,Wn,`*options*`)`<br>`[A,B,C,D] = besself(n,Wn,`*options*`)` |

By default, each of these functions returns a lowpass filter; you need only specify the desired cutoff frequency `Wn` in normalized frequency (Nyquist frequency = 1 Hz). For a highpass filter, append the string `'high'` to the function’s parameter list. For a bandpass or bandstop filter, specify `Wn` as a two-element vector containing the passband edge frequencies, appending the string `'stop'` for the bandstop configuration.

Here are some example digital filters:

```
[b,a] = butter(5,.4);                % lowpass Butterworth
[b,a] = cheby1(4,1,[.4 .7]);         % bandpass Chebyshev type I
[b,a] = cheby2(6,60,.8,'high');      % highpass Chebyshev type II
[b,a] = ellip(3,1,60,[.4 .7],'stop');% bandstop elliptic
```

To design an analog filter, perhaps for simulation, use a trailing `'s'` and specify cutoff frequencies in radians/second:

```
[b,a] = butter(5,.4,'s'); % analog Butterworth filter
```

All filter design functions return a filter in the transfer function, zero-pole-gain, or state-space linear system model representation, depending on how many output arguments are present.

### Designing IIR Filters to Frequency Domain Specifications

This toolbox provides order selection functions that calculate the minimum filter order that meets a given set of requirements:

| Filter Type | Order Estimation Function |
|---|---|
| Butterworth | `[n,Wn] = buttord(Wp,Ws,Rp,Rs)` |
| Chebyshev type I | `[n,Wn] = cheb1ord(Wp,Ws,Rp,Rs)` |
| Chebyshev type II | `[n,Wn] = cheb2ord(Wp,Ws,Rp,Rs)` |
| Elliptic | `[n,Wn] = ellipord(Wp,Ws,Rp,Rs)` |

These are useful in conjunction with the filter design functions. Suppose you want a bandpass filter with a passband from 1000 to 2000 Hz, stopbands starting 500 Hz away on either side, a 10 kHz sampling frequency, at most 1 dB

of passband ripple, and at least 60 dB of stopband attenuation. To meet these specifications with the `butter` function:

```
[n,Wn] = buttord([1000 2000]/5000,[500 2500]/5000,1,60)

n =

    12

Wn =

    0.1951    0.4080

[b,a] = butter(n,Wn);
```

An elliptic filter that meets the same requirements is given by

```
[n,Wn] = ellipord([1000 2000]/5000,[500 2500]/5000,1,60)

n =

    5

Wn =

    0.2000    0.4000

[b,a] = ellip(n,1,60,Wn);
```

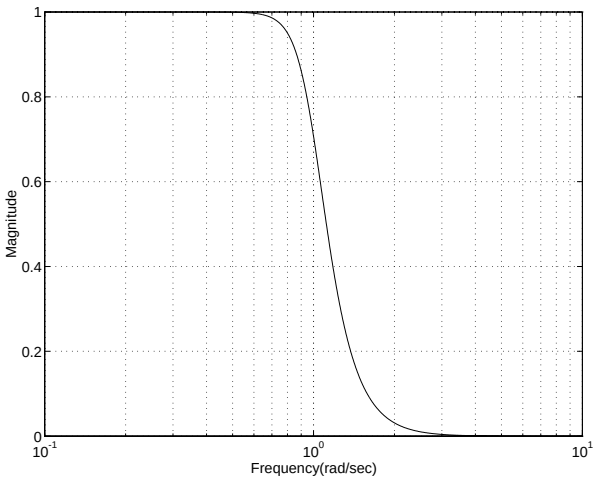
These functions also work with the other standard band configurations, as well as for analog filters; see Chapter 6 for details.

## Comparison of Classical IIR Filter Types

The toolbox provides five different types of classical IIR filters, each optimal in some way. This section shows the basic analog prototype form for each and summarizes major characteristics.

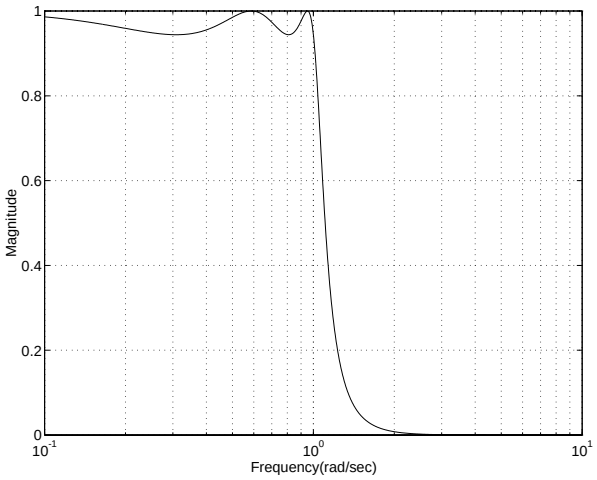
### Butterworth Filter

The Butterworth filter provides the best Taylor Series approximation to the ideal lowpass filter response at $\Omega = 0$ and $\Omega = \infty$; for any order $N$, the magnitude squared response has $2N-1$ zero derivatives at these locations (*maximally flat* at $\Omega = 0$ and $\Omega = \infty$). Response is monotonic overall, decreasing smoothly from $\Omega = 0$ to $\Omega = \infty$. $|H(j\Omega)|$ = `sqrt(1/2)` at $\Omega = 1$.



### Chebyshev Type I Filter

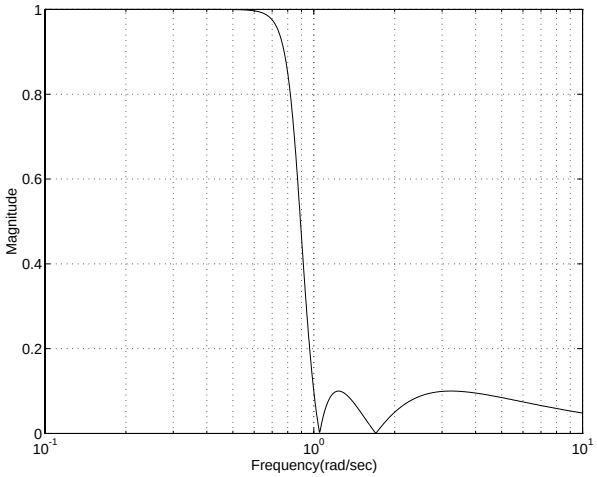
The Chebyshev type I filter minimizes the absolute difference between the ideal and actual frequency response over the entire passband by incorporating an equal ripple of Rp db in the passband. Stopband response is maximally flat. The transition from passband to stopband is more rapid than for the Butterworth filter. $|H(j\Omega)| = 10^{-Rp/20}$ at $\Omega = 1$.



### Chebyshev Type II Filter

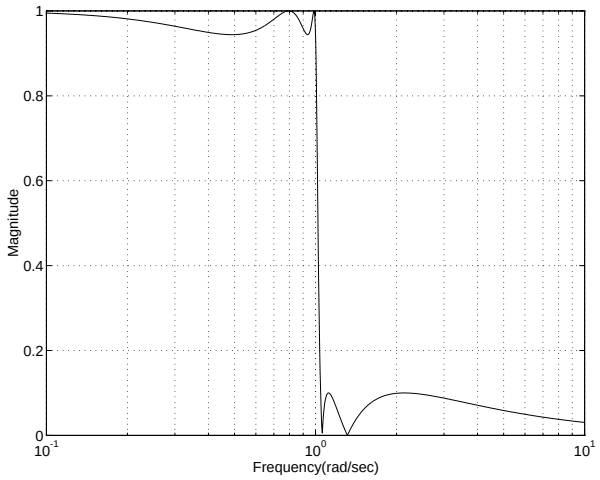
The Chebyshev type II filter minimizes the absolute difference between the ideal and actual frequency response over the entire stopband, by incorporating an equal ripple of Rs db in the stopband. Passband response is maximally flat.

The stopband does not approach zero as quickly as the type I filter (and does not approach zero at all for even-valued n). The absence of ripple in the passband, however, is often an important advantage. $|H(j\Omega)| = 10^{-Rs/20}$ at $\Omega = 1$.



### Elliptic Filter

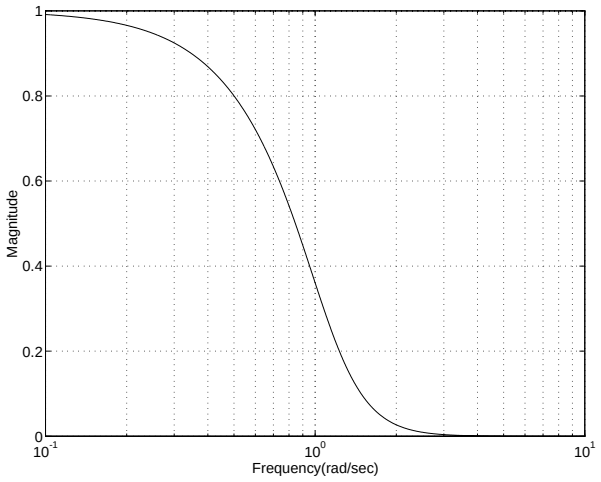
Elliptic filters are equiripple in both the passband and stopband. They generally meet filter requirements with the lowest order of any supported filter type. Given a filter order `n`, passband ripple `Rp` in decibels, and stopband ripple `Rs` in decibels, elliptic filters minimize transition width. $|H(j\Omega)| = 10^{-Rp/20}$ at $\Omega = 1$.



### Bessel Filter

Analog Bessel lowpass filters have maximally flat group delay at zero frequency and retain nearly constant group delay across the entire passband. Filtered signals therefore maintain their waveshapes in the passband frequency range. Frequency mapped and digital Bessel filters, however, do not have this maximally flat property; this toolbox supports only the analog case for the complete Bessel filter design function.

Bessel filters generally require a higher filter order than other filters for satisfactory stopband attenuation. $|H(j\Omega)| < 1/\sqrt{2}$ at $\Omega = 1$ and decreases as n increases.




**NOTE** The lowpass filters shown above were created with the analog prototype functions `besselap`, `buttap`, `cheb1ap`, `cheb2ap`, and `ellipap`. These functions find the zeros, poles, and gain of an order n analog filter of the appropriate type with cutoff frequency of 1 rad/sec. The complete filter design functions (`besself`, `butter`, `cheby1`, `cheby2`, and `ellip`) call the prototyping functions as a first step in the design process. See “Special Topics in IIR Filter Design” on page 2-37 for details.

To create similar plots, use `n = 5` and, as needed, `Rp = 0.5` and `Rs = 20`. For example, to create the elliptic filter plot:

```
[z,p,k] = ellipap(5,.5,20);
w = logspace(-1,1,1000);
h = freqs(k*poly(z),poly(p),w);
semilogx(w,abs(h)), grid
```

### Direct IIR Filter Design

This toolbox uses the term *direct methods* to describe techniques for IIR design that find a filter based on specifications in the discrete domain. Unlike the analog prototyping method, direct design methods are not constrained to the standard lowpass, highpass, bandpass, or bandstop configurations. Rather, these functions design filters with an arbitrary, perhaps multiband, frequency response. This section discusses the `yulewalk` function, which is intended specifically for filter design; "Parametric Modeling" in Chapter 4 discusses other methods that may also be considered direct, such as Prony's method, Linear Prediction, the Steiglitz-McBride method, and inverse frequency design.

`yulewalk` designs recursive IIR digital filters by fitting a specified frequency response. `yulewalk`'s name reflects its method for finding the filter's denominator coefficients: it finds the inverse FFT of the ideal desired power spectrum and solves the "modified Yule-Walker equations" using the resulting autocorrelation function samples. The statement

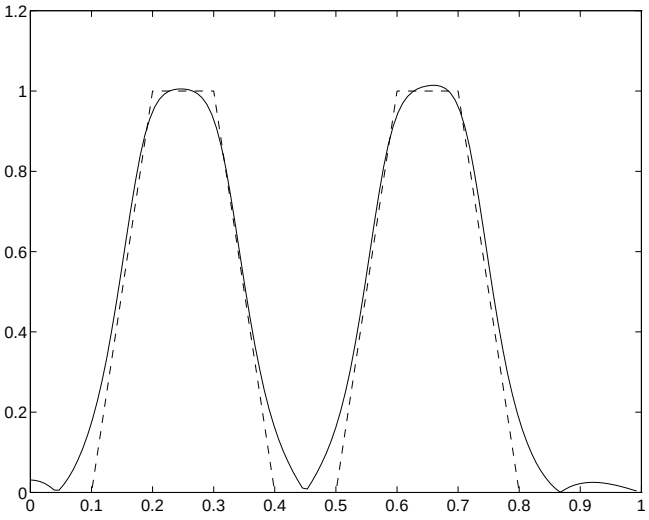
```
[b,a] = yulewalk(n,f,m)
```

returns row vectors `b` and `a` containing the `n+1` numerator and denominator coefficients of the order `n` IIR filter whose frequency-magnitude characteristics approximate those given in vectors `f` and `m`. `f` is a vector of frequency points ranging from 0 to 1, where 1 represents the Nyquist frequency. `m` is a vector containing the desired magnitude response at the points in `f`. `f` and `m` can describe any piecewise linear shape magnitude response, including a multiband response. The FIR counterpart of this function is `fir2`, which also designs a filter based on an arbitrary piecewise linear magnitude response. See "FIR Filter Design" on page 2-17 for details.

Note that `yulewalk` does not accept phase information, and no statements are made about the optimality of the resulting filter.

Design a multiband filter with `yulewalk`, and plot the desired and actual frequency response:

```
m = [0   0   1   1   0   0   1   1   0 0];
f = [0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 1];
[b,a] = yulewalk(10,f,m);
[h,w] = freqz(b,a,128);
plot(f,m,w/pi,abs(h))
```




### Generalized Butterworth Filter Design

The toolbox function `maxflat` enables you to design generalized Butterworth filters, that is, Butterworth filters with differing numbers of zeros and poles. This is desirable in some implementations where poles are more expensive computationally than zeros. `maxflat` is just like the `butter` function, except that it you can specify *two* orders (one for the numerator and one for the denominator) instead of just one. These filters are *maximally flat*. This means that the resulting filter is optimal for any numerator and denominator orders, with the maximum number of derivatives at 0 and the Nyquist frequency $\omega=\pi$ both set to 0.

For example, when the two orders are the same, `maxflat` is the same as `butter`:

```
[b,a] = maxflat(3,3,0.25)
b =
    0.0317    0.0951    0.0951    0.0317
a =
    1.0000   –1.4590    0.9104   -0.1978
[b,a] = butter(3,0.25)
b =
    0.0317    0.0951    0.0951    0.0317
a =
    1.0000   –1.4590    0.9104   -0.1978
```

However, `maxflat` is more versatile, because you can design a filter with more zeros than poles:

```
[b,a] = maxflat(3,1,0.25)
b =
    0.0950    0.2849    0.2849    0.0950
a =
    1.0000   -0.2402
```

The third input to maxflat is the *half-power frequency*, a frequency between 0 and 1 with a desired magnitude response of $1/\sqrt{2}$.

You can also design linear phase filters that have the maximally flat property using the `'sym'` option:

```
maxflat(4,'sym',0.3)
ans =
    0.0331    0.2500    0.4337    0.2500    0.0331
```

For complete details of the `maxflat` algorithm, see Selesnick and Burrus [2].

# FIR Filter Design

Digital filters with finite-duration impulse response (all-zero, or FIR filters) have both advantages and disadvantages compared to infinite-duration impulse response (IIR) filters.

FIR filters have the following primary advantages:

- They can have exactly linear phase.
- They are always stable.
- The design methods are generally linear.
- They can be realized efficiently in hardware.
- The filter startup transients have finite duration.

The primary disadvantage of FIR filters is that they often require a much higher filter order than IIR filters to achieve a given level of performance. Correspondingly, the delay of these filters is often much greater than for an equal performance IIR filter.

| Method | Description | Functions |
|---|---|---|
| Windowing | Apply window to truncated inverse Fourier transform of desired “brickwall” filter | `fir1,fir2,kaiserord` |
| Multiband with Transition Bands | Equiripple or least squares approach over sub-bands of the frequency range | `firls,remez,remezord` |
| Constrained Least Squares | Minimize squared integral error over entire frequency range subject to maximum error constraints | `fircls,fircls1` |
| Arbitrary Response | Arbitrary responses, including nonlinear phase and complex filters | `cremez` |
| Raised Cosine | Lowpass response with smooth, sinusoidal transition | `firrcos` |

## Linear Phase Filters

Except for `cremez`, all of the FIR filter design functions design linear phase filters only. The filter coefficients, or "taps," of such filters obey either an even or odd symmetry relation. Depending on this symmetry, and on whether the order *n* of the filter is even or odd, a linear phase filter (stored in length *n*+1 vector `b`) has certain inherent restrictions on its frequency response:

| Linear Phase Filter Type | Filter Order n | Symmetry of Coefficients | Response *H*(f), f = 0 | Response *H*(f), f = 1 (Nyquist) |
|---|---|---|---|---|
| Type I | Even | Even: $b(k) = b(n+2-k), \quad k = 1,\ldots,n+1$ | No restriction | No restriction |
| Type II | Odd | | No restriction | $H(1) = 0$ |
| Type III | Even | Odd: $b(k) = -b(n+2-k), \quad k = 1,\ldots,n+1$ | $H(0) = 0$ | $H(1) = 0$ |
| Type IV | Odd | | $H(0) = 0$ | No restriction |

The phase delay and group delay of linear phase FIR filters are equal and constant over the frequency band. For an order *n* linear phase FIR filter, the group delay is *n*/2, and the filtered signal is simply delayed by *n*/2 time steps (and the magnitude of its Fourier transform is scaled by the filter's magnitude response). This property preserves the wave shape of signals in the passband, that is, there is no phase distortion.

The functions `fir1`, `fir2`, `firls`, `remez`, `fircls`, `fircls1`, and `firrcos` all design type I and II linear phase FIR filters by default. Both `firls` and `remez` design type III and IV linear phase FIR filters given a `'hilbert'` or `'differentiator'` flag. `cremez` can design any type of linear phase filter, and nonlinear phase filters as well.

**NOTE** Because the frequency response of a type II filter is zero at the Nyquist rate ("high" frequency), `fir1` does not design type II highpass and bandstop filters. For odd-valued `n` in these cases, `fir1` adds 1 to the order and returns a type I filter.

## Windowing Method

Consider the ideal, or "brick-wall," digital lowpass filter with a cutoff frequency of $\omega_0$ rad/sec. This filter has magnitude 1 at all frequencies with magnitude less than $\omega_0$, and magnitude 0 at frequencies with magnitude between $\omega_0$ and $\pi$. Its impulse response sequence $h(n)$ is

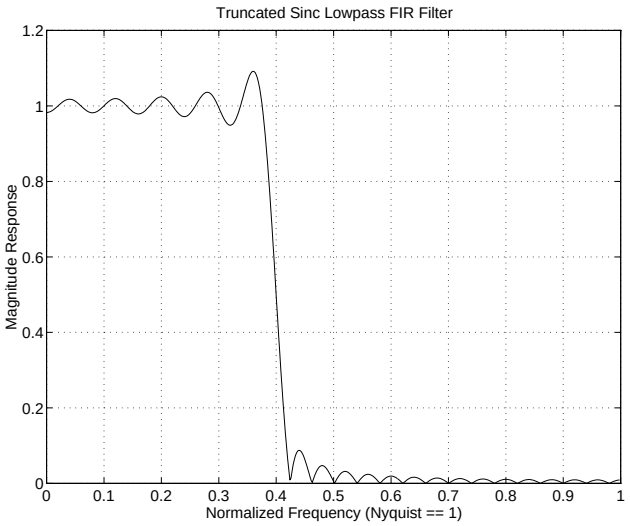
$$h(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(\omega)e^{j\omega n}d\omega = \frac{1}{2\pi}\int_{-\omega_0}^{\omega_0} e^{j\omega n}d\omega = \frac{\omega_0}{\pi}\,\mathrm{sinc}\left(\frac{\omega_0}{\pi}n\right)$$

This filter is not implementable since its impulse response is infinite and non-causal. To create a finite-duration impulse response, truncate it by applying a window. By retaining the central section of impulse response in this truncation, you obtain a linear phase FIR filter. For example, a length 51 filter with a lowpass cutoff frequency $\omega_0$ of $0.4\pi$ rad/sec is

```
b = 0.4*sinc(0.4*(-25:25));
```

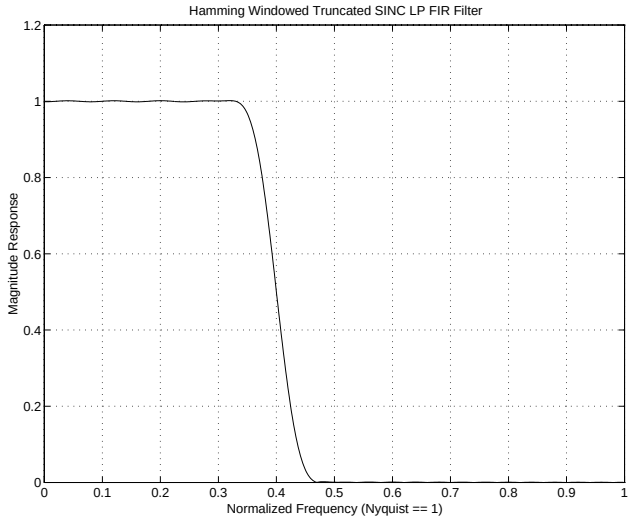
The window applied here is a simple rectangular or "boxcar" window. By Parseval's theorem, this is the length 51 filter that best approximates the ideal lowpass filter, in the integrated least squares sense. To view its frequency response,

```
[H,w] = freqz(b,1,512,2);
plot(w,abs(H)), grid
```



Note the ringing and ripples in the response, especially near the band edge. This "Gibbs effect" does not vanish as the filter length increases, but a nonrectangular window reduces its magnitude. Multiplication by a window in the time domain causes a convolution or smoothing in the frequency domain. Apply a length 51 Hamming window to the filter:

```
b = b.*hamming(51)';
[H,w] = freqz(b,1,512,2);
plot(w,abs(H)), grid
```




As you can see, this greatly reduces the ringing. This improvement is at the expense of transition width (the windowed version takes longer to ramp from passband to stopband) and optimality (the windowed version does not minimize the integrated squared error).

The functions `fir1` and `fir2` are based on this windowing process. Given a filter order and description of an ideal desired filter, these functions return a windowed inverse Fourier transform of that ideal filter. Both use a Hamming window by default, but they accept any window function. See the "Windows" section of Chapter 4 for an overview of windows and their properties.

### Standard Band FIR Filter Design: fir1

`fir1` implements the classical method of windowed linear phase FIR digital filter design. It resembles the IIR filter design functions in that it is formulated

to design filters in standard band configurations: lowpass, bandpass, highpass, and bandstop.

The statements

```
n = 50;
Wn = 0.4;
b = fir1(n,Wn);
```

create row vector `b` containing the coefficients of the order `n` Hamming-windowed filter. This is a lowpass, linear phase FIR filter with cutoff frequency `Wn`. `Wn` is a number between 0 and 1, where 1 corresponds to the Nyquist frequency, half the sampling frequency. (Unlike other methods, here `Wn` corresponds to the 6 dB point.) For a highpass filter, simply append the string `'high'` to the function's parameter list. For a bandpass or bandstop filter, specify `Wn` as a two-element vector containing the passband edge frequencies; append the string `'stop'` for the bandstop configuration.

`b = fir1(n,Wn,window)` uses the window specified in column vector `window` for the design. The vector `window` must be `n+1` elements long. If you do not specify a window, `fir1` applies a Hamming window.

**Kaiser Window Order Estimation.** The `kaiserord` function estimates the filter order, cutoff frequency, and Kaiser window beta parameter needed to meet a given set of specifications. Given a vector of frequency band edges and a corresponding vector of magnitudes, as well as maximum allowable ripple, `kaiserord` returns appropriate input parameters for the `fir1` function. For details on `kaiserord`, see the reference description in Chapter 6.

### Multiband FIR Filter Design: fir2

The function `fir2` also designs windowed FIR filters, but with an arbitrarily shaped piecewise linear frequency response. This is in contrast to `fir1`, which only designs filters in standard lowpass, highpass, bandpass, and bandstop configurations.

The commands

```
n = 50;
f = [0 .4 .5 1];
m = [1  1  0 0];
b = fir2(n,f,m);
```

return row vector `b` containing the `n+1` coefficients of the order `n` FIR filter whose frequency-magnitude characteristics match those given by vectors `f` and `m`. `f` is a vector of frequency points ranging from 0 to 1, where 1 represents the Nyquist frequency. `m` is a vector containing the desired magnitude response at the points specified in `f`. (The IIR counterpart of this function is `yulewalk`, which also designs filters based on arbitrary piecewise linear magnitude responses. See "IIR Filter Design" for details.)

## Multiband FIR Filter Design with Transition Bands

The `firls` and `remez` functions provide a more general means of specifying the ideal desired filter than the `fir1` and `fir2` functions. These functions design Hilbert transformers, differentiators, and other filters with odd symmetric coefficients (type III and type IV linear phase). They also let you include transition or "don't care" regions in which the error is not minimized, and perform band dependent weighting of the minimization.

`firls` is an extension of the `fir1` and `fir2` functions in that it minimizes the integral of the square of the error between the desired frequency response and the actual frequency response.

`remez` implements the Parks-McClellan algorithm, which uses the Remez exchange algorithm and Chebyshev approximation theory to design filters with optimal fits between the desired and actual frequency responses. The filters are optimal in the sense that they minimize the maximum error between the desired frequency response and the actual frequency response; they are sometimes called *minimax* filters. Filters designed in this way exhibit an equiripple behavior in their frequency response, and hence are also known as *equiripple* filters. The Parks-McClellan FIR filter design algorithm is perhaps the most popular and widely used FIR filter design methodology.

The syntax for `firls` and `remez` is the same; the only difference is their minimization schemes. The next example shows how filters designed with `firls` and `remez` reflect these different schemes.

### Basic Configurations

The default mode of operation of `firls` and `remez` is to design type I or type II linear phase filters, depending on whether the order you desire is even or odd,

respectively. A lowpass example with approximate amplitude 1 from 0 to 0.4 Hz, and approximate amplitude 0 from 0.5 to 1.0 Hz is

```
n = 20;           % filter order
f = [0 .4 .5 1];  % frequency band edges
a = [1  1  0 0];  % desired amplitudes
b = remez(n,f,a);
```

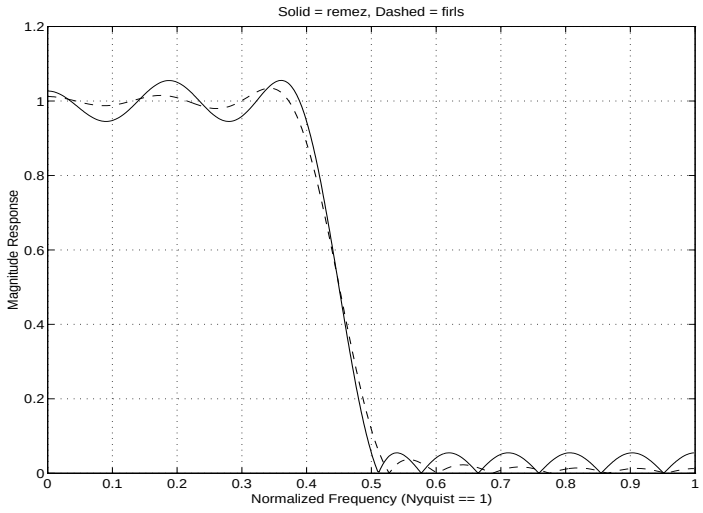
From 0.4 to 0.5 Hz, `remez` performs no error minimization; this is a transition band or "don't care" region. A transition band minimizes the error more in the bands that you do care about, at the expense of a slower transition rate. In this way, these types of filters have an inherent trade-off similar to FIR design by windowing.

To compare least squares to equiripple filter design, use `firls` to create a similar filter:

```
bb = firls(n,f,a);
```

and compare their frequency responses:

```
[H,w]=freqz(b);
[HH,w]=freqz(bb);
plot(w/pi,abs(H),w/pi,abs(HH),'- -'), grid
```



You can see that the filter designed with `remez` exhibits equiripple behavior. Also note that the `firls` filter has a better response over most of the passband and stopband, but at the band edges (`f = 0.4` and `f = 0.5`), the response is further away from the ideal than the `remez` filter. This shows that the `remez` filter's *maximum* error over the pass- and stopbands is smaller and, in fact, it is the smallest possible for this band edge configuration and filter length.

Think of frequency bands as lines over short frequency intervals. `remez` and `firls` use this scheme to represent any piecewise linear desired function with any transition bands. `firls` and `remez` design lowpass, highpass, bandpass, and bandstop filters; a bandpass example is

```
f = [0 .3  .4  .7  .8  1];  % band edges in pairs
a = [0  0   1   1   0  0];  % bandpass filter amplitude
```

Technically, these `f` and `a` vectors define five bands:

- Two stopbands, from 0.0 to 0.3 and from 0.8 to 1.0
- A passband from 0.4 to 0.7
- Two transition bands, from 0.3 to 0.4 and from 0.7 to 0.8

Example highpass and bandstop filters are

```
f = [0 .7  .8  1];  % band edges in pairs
a = [0  0   1  1];  % highpass filter amplitude

f = [0 .3  .4  .5  .8  1];  % band edges in pairs
a = [1  1   0   0   1  1];  % bandstop filter amplitude
```

An example multiband bandpass filter is

```
f = [0 .1 .15 .25 .3 .4 .45 .55 .6 .7 .75 .85 .9 1];
a = [1  1  0   0   1  1  0   0   1  1   0   0  1 1];
```

Another possibility is a filter that has as a transition region the line connecting the passband with the stopband; this can help control "runaway" magnitude response in wide transition regions:

```
f = [0 .4 .42 .48 .5  1];
a = [1  1  .8  .2  0  0]; % passband, linear transition, stopband
```

### The Weight Vector

Both `firls` and `remez` allow you to place more or less emphasis on minimizing the error in certain frequency bands relative to others. To do this, specify a weight vector following the frequency and amplitude vectors. An example low-pass equiripple filter with 10 times less ripple in the stopband than the pass-band is

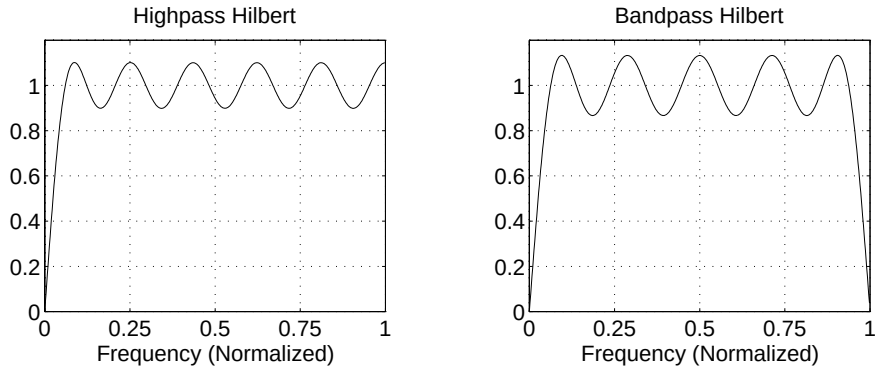
```
n = 20;           % filter order
f = [0 .4 .5 1];  % frequency band edges
a = [1  1  0 0];  % desired amplitudes
w = [1 10];       % weight vector
b = remez(n,f,a,w);
```

A legal weight vector is always half the length of the `f` and `a` vectors; there must be exactly one weight per band.

### Anti-Symmetric Filters / Hilbert Transformers

When called with a trailing `'h'` or `'Hilbert'` option, `remez` and `firls` design FIR filters with odd symmetry, that is, type III (for even order) or type IV (for odd order) linear phase filters. An ideal Hilbert transformer has this anti-symmetry property and an amplitude of 1 across the entire frequency range. Try the following approximate Hilbert transformers:

```
b = remez(21,[.05 1],[1 1],'h');    % highpass Hilbert
bb = remez(20,[.05 .95],[1 1],'h'); % bandpass Hilbert
```



You can find the delayed Hilbert transform of a signal x by passing it through these filters:

```
Fs = 1000;           % sampling frequency
t = (0:1/Fs:2)';     % two second time vector
x = sin(2*pi*300*t); % 300 Hz sine wave example signal
xh = filter(bb,1,x); % Hilbert transform of x
```

The analytic signal corresponding to x is the complex signal that has x as its real part and the Hilbert transform of x as its imaginary part. For this FIR method (an alternative to the `hilbert` function), you must delay x by half the filter order to create the analytic signal:

```
xd = [zeros(10,1); x(1:length(x)-10)]; % delay 10 samples
xa = xd + j*xh;  % analytic signal
```

This method does not work directly for filters of odd order, which require a noninteger delay. In this case, the `hilbert` function, described in the "Specialized Transforms" section in Chapter 4, estimates the analytic signal. Alternately, use the `resample` function to delay the signal by a noninteger number of samples.

### Differentiators

Differentiation of a signal in the time domain is equivalent to multiplication of the signal's Fourier transform by an imaginary ramp function. That is, to differentiate a signal, pass it through a filter that has a response $H(w) = jw$. Approximate the ideal differentiator (with a delay) using `remez` or `firls` with a `'d'` or `'differentiator'` option:

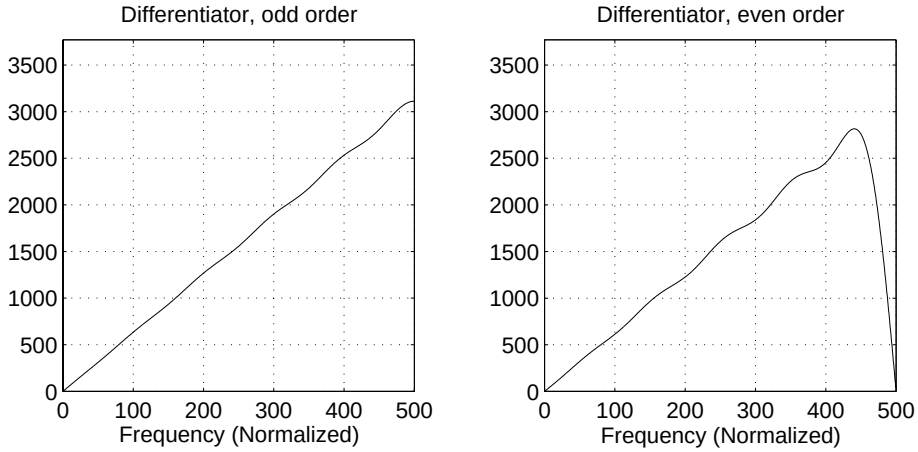
```
b = remez(21,[0 1],[0 pi*Fs],'d');
```

To obtain the correct derivative, scale by `pi*Fs` rad/sec, where `Fs` is the sampling frequency in Hertz. For a type III filter, the differentiation band should stop short of the Nyquist frequency, and the amplitude vector must reflect that change to ensure the correct slope:

```
bb = remez(20,[0 .9],[0 .9*pi*Fs],'d');
```

In the `'d'` mode, `remez` weights the error by $1/w$ in nonzero amplitude bands to minimize the maximum *relative* error. `firls` weights the error by $(1/w)^2$ in nonzero amplitude bands in the `'d'` mode.

The magnitude response plots for the differentiators shown above are




## Constrained Least Squares FIR Filter Design

The Constrained Least Squares (CLS) FIR filter design functions implement a technique that enables you to design FIR filters without explicitly defining the transition bands for the magnitude response. The ability to omit the specification of transition bands is useful in several situations. For example, it may not be clear where a rigidly defined transition band should appear if noise and signal information appear together in the same frequency band. Similarly, it may make sense to omit the specification of transition bands if they appear only to control the results of Gibbs phenomena that appear in the filter's response. See Selesnick, Lang, and Burrus [2] for discussion of this method.

Instead of defining passbands, stopbands, and transition regions, the CLS method accepts a cutoff frequency (for the highpass, lowpass, bandpass, or bandstop cases), or passband and stopband edges (for multiband cases), for the desired response. In this way, the CLS method defines transition regions implicitly, rather than explicitly.

The key feature of the CLS method is that it enables you to define upper and lower thresholds that contain the maximum allowable ripple in the magnitude response. Given this constraint, the technique applies the least square error minimization technique over the frequency range of the filter's response, instead of over specific bands. The error minimization includes any areas of discontinuity in the ideal, "brick wall" response. An additional benefit is that

the technique enables you to specify arbitrarily small peaks resulting from Gibbs' phenomena.

There are two toolbox functions that implement this design technique:

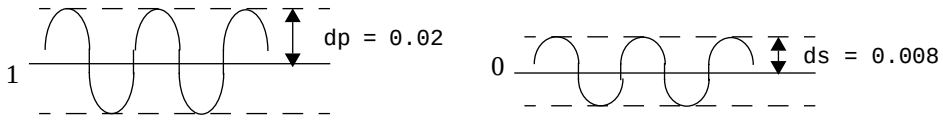
| Description | Function |
|---|---|
| Constrained least square multiband FIR filter design. | `fircls` |
| Constrained least square filter design for lowpass and highpass linear phase filters | `fircls1` |

For details on the calling syntax for these functions, see their reference descriptions in Chapter 6.

### Basic Lowpass and Highpass CLS Filter Design

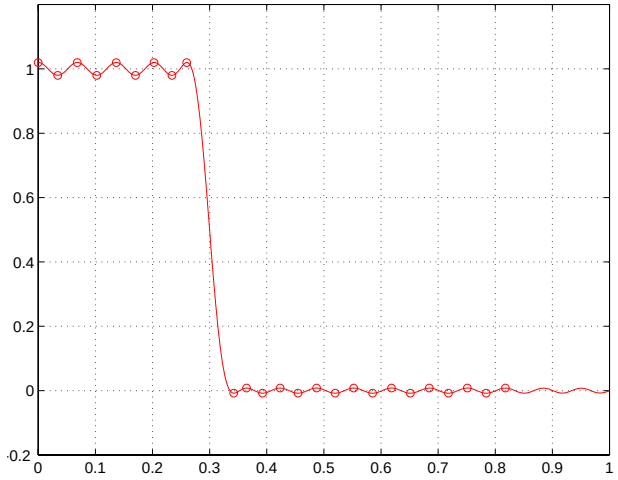
The most basic of the CLS design functions, `fircls1`, uses this technique to design lowpass and highpass FIR filters. As an example, consider designing a filter with order 61 impulse response and cutoff frequency of 0.3 (normalized). Further, define the upper and lower bounds that constrain the design process as

- Maximum passband deviation from 1 (passband ripple) of 0.02.
- Maximum stopband deviation from 0 (stopband ripple) of 0.008.



To approach this design problem using `fircls1`:

```
n = 61;
wo = 0.3;
dp = 0.02;
ds = 0.008;
h = fircls1(n,wo,dp,ds,'plot');
```




### Multiband CLS Filter Design

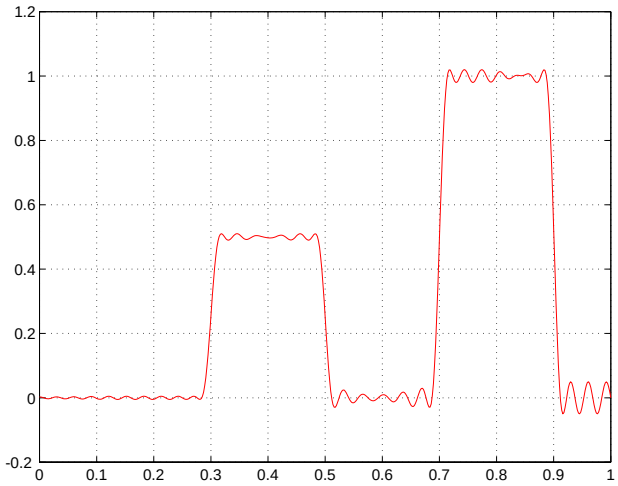
`fircls` uses the same technique to design FIR filters with a desired piecewise constant magnitude response. In this case, you can specify a vector of band edges and a corresponding vector of band amplitudes. In addition, you can specify the maximum amount of ripple for each band.

For example, assume the specifications for a filter call for

- From 0 to 0.3 (normalized): amplitude 0, upper bound 0.005, lower bound -0.005
- From 0.3 to 0.5: amplitude 0.5, upper bound 0.51, lower bound 0.49
- From 0.5 to 0.7: amplitude 0, upper bound 0.03, lower bound -0.03
- From 0.7 to 0.9: amplitude 1, upper bound 1.02, lower bound 0.98
- From 0.9 to 1: amplitude 0, upper bound 0.05, lower bound -0.05

Design a CLS filter with impulse response order 129 that meets these specifications:

```
n = 129;
f = [0 0.3 0.5 0.7 0.9 1];
a = [0 0.5 0 1 0];
up = [0.005 0.51 0.03 1.02 0.05];
lo = [-0.005 0.49 -0.03 0.98 -0.05];
h = fircls(n,f,a,up,lo,'plot');
```




### Weighted CLS Filter Design

Weighted CLS filter design lets you design lowpass or highpass FIR filters with relative weighting of the error minimization in each band. The `fircls1` function enables you to specify the passband and stopband edges for the least squares weighting function, as well as a constant `k` that specifies the ratio of the stopband to passband weighting.

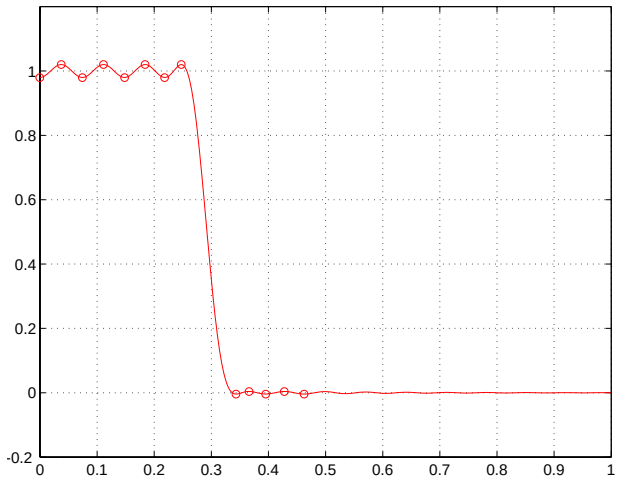
For example, consider specifications that call for an FIR filter with impulse response order of 55 and cutoff frequency of 0.3 (normalized). Also assume max-

imum allowable passband ripple of 0.02 and maximum allowable stopband ripple of 0.004. In addition, add weighting requirements:

- Passband edge for the weight function of 0.28 (normalized)
- Stopband edge for the weight function of 0.32
- Weight error minimization 10 times as much in the stopband as in the passband

To approach this using `fircls1`:

```
n = 55;
wo = 0.3;
dp = 0.02;
ds = 0.004;
wp = 0.28;
ws = 0.32;
k = 10;
h = fircls1(n,wo,dp,ds,wp,ws,k,'plot');
```




## Arbitrary-Response Filter Design

The `cremez` filter design function provides a tool for designing FIR filters with arbitrary complex responses. It differs from the other filter design functions in how the frequency response of the filter is specified: it passes the name of a function to `cremez`, which returns the filter response calculated over a grid of

frequencies. This capability makes `cremez` a highly versatile and powerful technique for filter design.

This design technique may be used to produce nonlinear-phase FIR filters, asymmetric frequency-response filters (with complex coefficients), or more symmetric filters with custom frequency responses.

The design algorithm optimizes the Chebyshev (or minimax) error using an extended Remez-exchange algorithm for an initial estimate. If this exchange method fails to obtain the optimal filter, the algorithm switches to an ascent-descent algorithm that takes over to finish the convergence to the optimal solution.

For details on the calling syntax for `cremez`, see the reference description in Chapter 6.

### Multiband Filter Design

Consider a multiband filter with the following special frequency-domain characteristics:

| Band | Amplitude | Optimization Weighting |
|---|---|---|
| [-1 -0.5] | [5 1] | 1 |
| [-0.4 +0.3] | [2 2] | 10 |
| [+0.4 +0.8] | [2 1] | 5 |

A linear-phase multiband filter may be designed using the predefined frequency-response function `multiband`, as follows:

```
b = cremez(38, [–1 –0.5 –0.4 0.3 0.4 0.8], ...
               {’multiband’, [5 1 2 2 2 1]}, [1 10 5]);
```

For the specific case of a multiband filter, we can use a shorthand filter design notation similar to the syntax for `remez`:

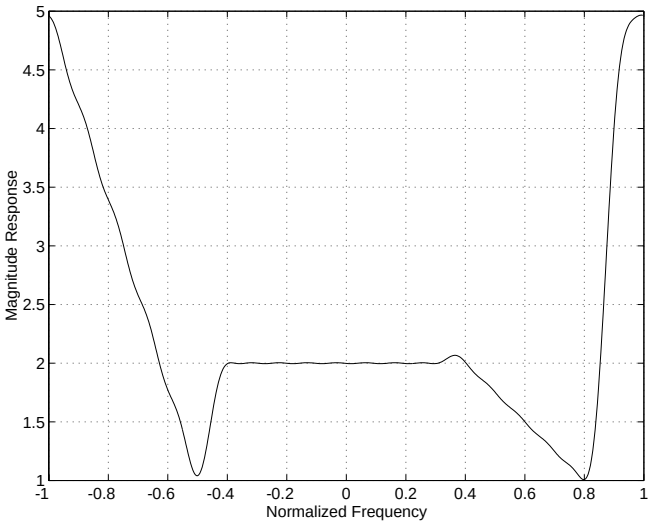
```
b = cremez(38,[–1 –0.5 –0.4 0.3 0.4 0.8], ...
              [5 1 2 2 2 1], [1 10 5]);
```

As with `remez`, a vector of band edges is passed to `cremez`. This vector defines the frequency bands over which optimization is performed; note that there are two transition bands, from -0.5 to -0.4 and from 0.3 to 0.4.

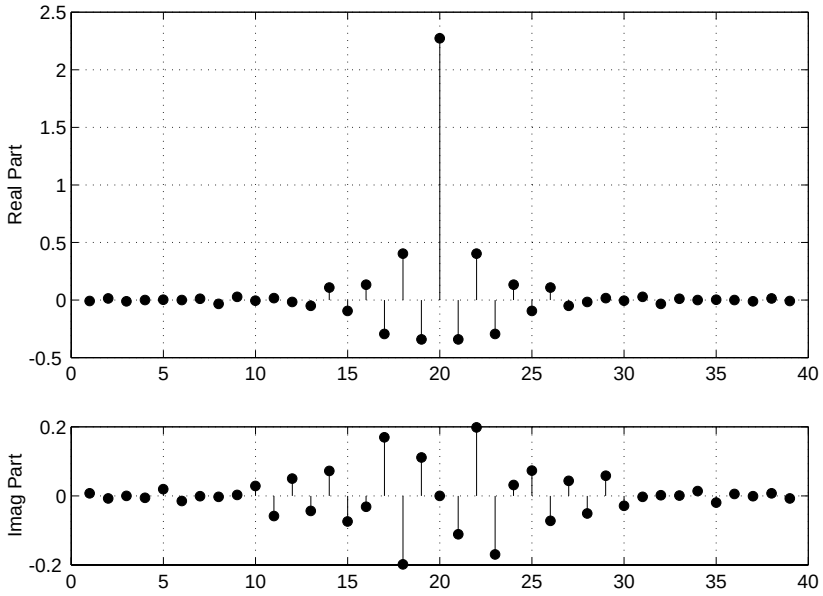
In either case, the frequency response is obtained and plotted using linear scale:

```
[h,w] = freqz(b,1,512,'whole');
plot(w/pi-1,fftshift(abs(h)));
```

Note that the frequency response has been calculated over the entire normalized frequency range [-1 +1] by passing the option `'whole'` to `freqz`. In order to plot the negative frequency information in a natural way, the response has been "wrapped," just as FFT data is, using `fftshift`:



The filter response for this multiband filter is complex, which is expected because of the asymmetry in the frequency domain. The filter response is



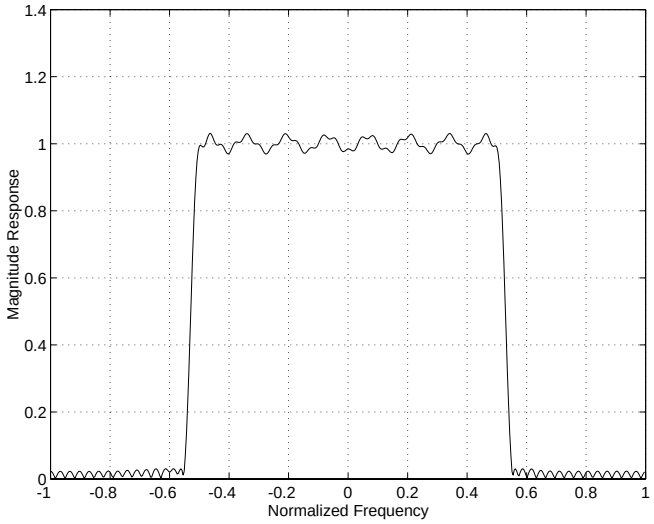

### Filter Design with Reduced Delay

Consider the design of a 62-tap lowpass filter with a half-Nyquist cutoff. If we specify a negative offset value to the `lowpass` filter design function, the group delay offset for the design is significantly less than that obtained for a standard linear-phase design. This filter design may be computed as follows:

```
b = cremez(61, [0 0.5 0.55 1], {'lowpass', -16});
```

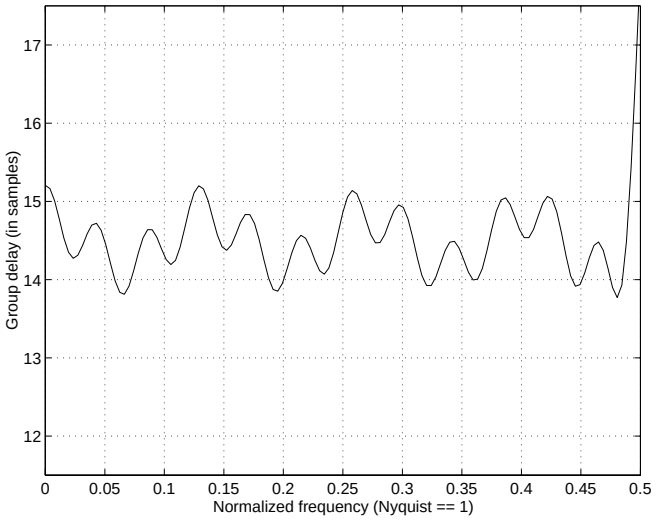
The resulting magnitude response is:

```
[h,w] = freqz(b,1,512,'whole');
plot(w/pi-1,fftshift(abs(h)));
```




The group delay of the filter reveals that the offset has been reduced from `N/2=30.5` to `N/2-16=14.5`. Now, however, the group delay is no longer flat in

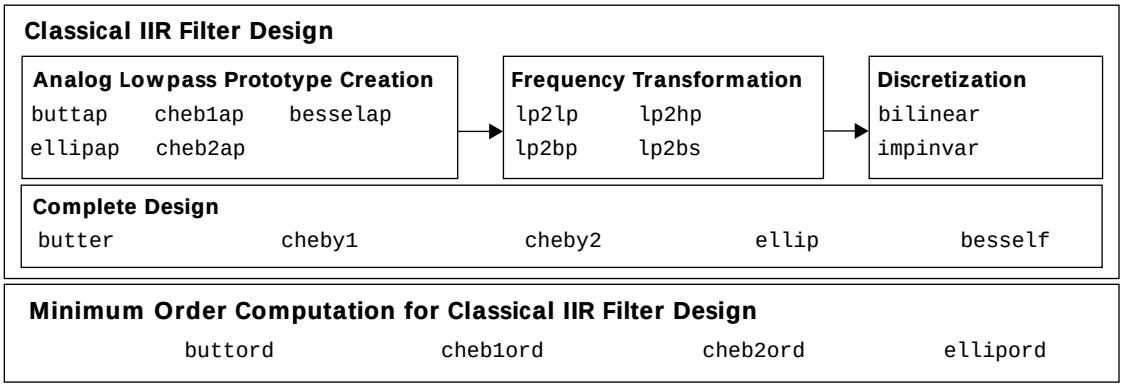
the passband region (plotted over the normalized frequency range 0 to 0.5 for clarity):




If we compare this nonlinear-phase filter to a linear-phase filter that has exactly 14.5 samples of group delay, the resulting filter is of order 2*14.5 or 29. Using `b = cremez(29,[0 0.5 0.55 1],'lowpass')`, the passband and stopband ripple is much greater for the order 29 filter. These comparisons can assist you in deciding which filter is more appropriate for a specific application.

# Special Topics in IIR Filter Design

The classic IIR filter design technique finds an analog lowpass filter with cutoff frequency of 1, translates this “prototype” filter to the desired band configuration, then transforms the filter to the digital domain. The toolbox provides functions for each step of this process:




The `butter`, `cheby1`, `cheby2`, and `ellip` functions are sufficient for many design problems, and the lower level functions are generally not needed. But if you do have an application where you need to transform the band edges of an analog filter, or discretize a rational transfer function, this section describes tools to do so.

## Analog Prototype Design

This toolbox provides a number of functions to create lowpass analog prototype filters with cutoff frequency of 1, the first step in the classical approach to IIR filter design. The table below summarizes the analog prototype design functions for each supported filter type; plots for each type are shown in the "IIR Filter Design" section above.

| Filter Type | Analog Prototype Function |
|---|---|
| Bessel | `[z,p,k] = besselap(n)` |
| Butterworth | `[z,p,k] = buttap(n)` |
| Chebyshev type I | `[z,p,k] = cheb1ap(n,Rp)` |
| Chebyshev type II | `[z,p,k] = cheb2ap(n,Rs)` |
| Elliptic | `[z,p,k] = ellipap(n,Rp,Rs)` |

## Frequency Transformation

The second step in the analog prototyping design technique is the frequency transformation of a lowpass prototype. The toolbox provides a set of functions to transform analog lowpass prototypes (with cutoff frequency of 1 rad/sec) into bandpass, highpass, bandstop, and lowpass filters of the desired cutoff frequency:

| Freq. Transformation | Transformation Function |
|---|---|
| Lowpass to lowpass $s' = s/\omega_0$ | `[numt,dent] = lp2lp(num,den,Wo)`<br>`[At,Bt,Ct,Dt] = lp2lp(A,B,C,D,Wo)` |
| Lowpass to highpass $s' = \frac{\omega_0}{s}$ | `[numt,dent] = lp2hp(num,den,Wo)`<br>`[At,Bt,Ct,Dt] = lp2hp(A,B,C,D,Wo)` |
| Lowpass to bandpass $s' = \frac{\omega_0}{B_w}\frac{(s/\omega_0)^2+1}{s/\omega_0}$ | `[numt,dent] = lp2bp(num,den,Wo,Bw)`<br>`[At,Bt,Ct,Dt] = lp2bp(A,B,C,D,Wo,Bw)` |
| Lowpass to bandstop $s' = \frac{B_w}{\omega_0}\frac{s/\omega_0}{(s/\omega_0)^2+1}$ | `[numt,dent] = lp2bs(num,den,Wo,Bw)`<br>`[At,Bt,Ct,Dt] = lp2bs(A,B,C,D,Wo,Bw)` |

As shown, all of the frequency transformation functions can accept two linear system models: transfer function and state-space form. For the bandpass and bandstop cases,

$$\omega_0 = \sqrt{\omega_1\omega_2}$$

and

$$B_w = \omega_2 - \omega_1$$

where $\omega_1$ is the lower band edge and $\omega_2$ is the upper band edge.

The frequency transformation functions perform frequency variable substitution. In the case of `lp2bp` and `lp2bs`, this is a second-order substitution, so the output filter is twice the order of the input. For `lp2lp` and `lp2hp`, the output filter is the same order as the input.

To begin designing an order 10 bandpass Chebyshev type I filter with a value of 3 dB for passband ripple:

```
[z,p,k] = cheb1ap(5,3);
```

`z`, `p`, and `k` contain the poles, zeros, and gain of a lowpass analog filter with cutoff frequency $\Omega_c$ equal to 1 rad/sec. Use the `lp2bp` function to transform this lowpass prototype to a bandpass analog filter with band edges $W_1 = \pi/5$ and $W_2 = \pi$. First, convert the filter to state-space form so the `lp2bp` function can accept it:

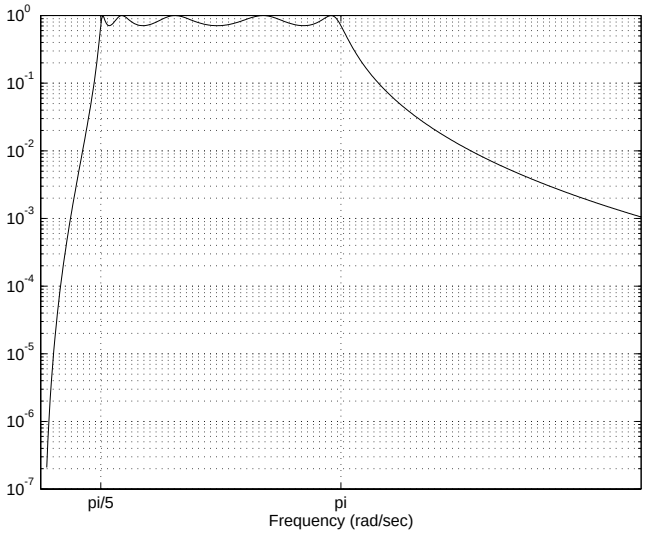
```
[A,B,C,D] = zp2ss(z,p,k); % Convert to state-space form.
```

Now, find the bandwidth and center frequency, and call `lp2bp`:

```
u1 = 0.1*2*pi;  u2 = 0.5*2*pi; % in radians per second
Bw = u2-u1;
Wo = sqrt(u1*u2);
[At,Bt,Ct,Dt] = lp2bp(A,B,C,D,Wo,Bw);
```

Finally, calculate the frequency response and plot its magnitude:

```
[b,a] = ss2tf(At,Bt,Ct,Dt);  % Convert to TF form.
w = linspace(.01,1,500)*2*pi;% Generate frequency vector.
h = freqs(b,a,w);            % Compute frequency response.
semilogy(w/2/pi,abs(h)),grid % Plot log magnitude vs. freq.
```




## Filter Discretization

The third step in the analog prototyping technique is the transformation of the filter to the discrete-time domain. The toolbox provides two methods for this: the impulse invariant and bilinear transformations. The filter design functions `butter`, `cheby1`, `cheby2`, and `ellip` use the bilinear transformation for discretization in this step.

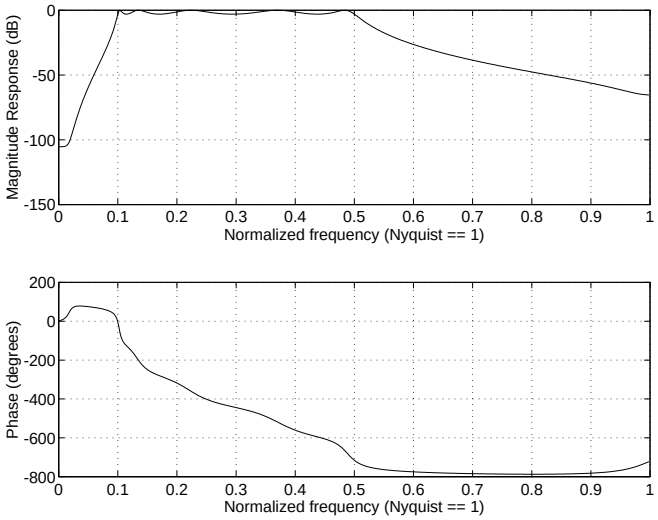
| Analog to Digital Transformation | Transformation Function |
|---|---|
| Impulse invariance | `[numd,dend] = impinvar(num,den,Fs)` |
| Bilinear transform | `[zd,pd,kd] = bilinear(z,p,k,Fs,Fp)`<br>`[numd,dend] = bilinear(num,den,Fs,Fp)`<br>`[Ad,Bd,Cd,Dd] = bilinear(At,Bt,Ct,Dt,Fs,Fp)` |

### Impulse Invariance

The toolbox function `impinvar` creates a digital filter whose impulse response is the samples of the continuous impulse response of an analog filter. This function only works on filters in transfer function form. For best results, the analog filter should have negligible frequency content above half the sampling frequency, because such high frequency content is aliased into lower bands upon sampling. Impulse invariance works for some lowpass and bandpass filters, but is not appropriate for highpass and bandstop filters.

Design a Chebyshev type I filter and plot its frequency response:

```
[bz,az] = impinvar(b,a,2);
freqz(bz,az)
```




Impulse invariance retains the cutoff frequencies of 0.1 Hz and 0.5 Hz.

### Bilinear Transformation

The bilinear transformation is a nonlinear mapping of the continuous domain to the discrete domain; it maps the *s*-plane into the *z*-plane by

$$H(z) = H(s)\Big|_{s=k\frac{z-1}{z+1}}$$

Bilinear transformation maps the jΩ axis of the continuous domain to the unit circle of the discrete domain according to

$$\omega = 2\tan^{-1}\left(\frac{\Omega}{k}\right)$$

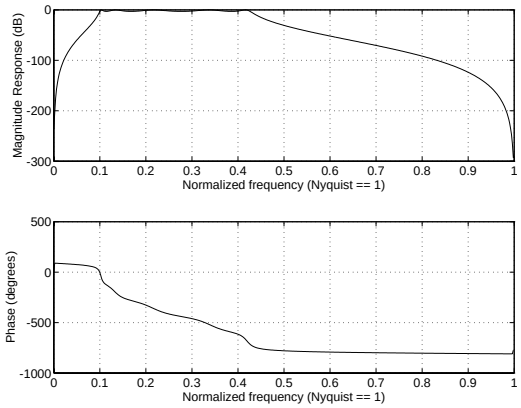
The toolbox function `bilinear` implements this operation, where the frequency warping constant $k$ is equal to twice the sampling frequency (`2*Fs`) by default and equal to `2*pi*Fp/tan(pi*Fp/Fs)` if you give `bilinear` a trailing argument that represents a "match" frequency `Fp`. If a match frequency `Fp` (in Hertz) is present, `bilinear` maps the frequency $\Omega = 2\pi f_p$ (in radians/second) to the same frequency in the discrete domain, normalized to the sampling rate: $\omega = 2\pi f_p / f_s$ (also in radians/second).

The `bilinear` function can perform this transformation on three different linear system representations: zero-pole-gain, transfer function, and state-space form. Try calling `bilinear` with the state-space matrices that describe the Chebyshev type I filter from the previous section, using a sampling frequency of 2 Hz, and retaining the lower band edge of 0.1 Hz:

```
[Ad,Bd,Cd,Dd] = bilinear(At,Bt,Ct,Dt,2,0.1);
```

The frequency response of the resulting digital filter is

```
[bz,az] = ss2tf(Ad,Bd,Cd,Dd); % convert to TF
freqz(bz,az)
```



The lower band edge is at 0.1 Hz as expected. Notice, however, that the upper band edge is slightly less than 0.5 Hz, although in the analog domain it was exactly 0.5 Hz. This illustrates the nonlinear nature of the bilinear transformation. To counteract this nonlinearity, it is necessary to create analog domain filters with "prewarped" band edges, which map to the correct locations upon bilinear transformation. Here the prewarped frequencies `u1` and `u2` generate `Bw` and `Wo` for the `lp2bp` function:

```
Fs = 2;  % sampling frequency (Hertz)
u1 = 2*Fs*tan(.1*(2*pi/Fs)/2); % lower band edge (radians/second)
u2 = 2*Fs*tan(.5*(2*pi/Fs)/2); % upper band edge (radians/second)
Bw = u2-u1;    % bandwidth
Wo = sqrt(u1*u2);  % center frequency
[At,Bt,Ct,Dt] = lp2bp(A,B,C,D,Wo,Bw);
```

A digital bandpass filter with correct band edges 0.1 and 0.5 times the Nyquist frequency is

```
[Ad,Bd,Cd,Dd] = bilinear(At,Bt,Ct,Dt,Fs);
```

The example bandpass filters from the last two sections could also be created in one statement using the complete IIR design function `cheby1`. For instance, an analog version of the example Chebyshev filter is

```
[b,a] = cheby1(5,3,[0.1 0.5]*2*pi,'s');
```

Note that the band edges are in radians/second for analog filters, whereas for the digital case, frequency is normalized (the Nyquist frequency is equal to 1 Hz):

```
[bz,az] = cheby1(5,3,[0.1 0.5]);
```

All of the complete design functions call `bilinear` internally. They prewarp the band edges as needed to obtain the correct digital filter. See Chapter 6 for more on these functions.

# References


1 Karam, L.J., and J.H. McClellan. “Complex Chebyshev Approximation for FIR Filter Design.” *IEEE Trans. on Circuits and Systems II*. March 1995.

2 Selesnick, I.W., and C.S. Burrus. “Generalized Digital Butterworth Filter Design.” *Proceedings of the IEEE Int. Conf. Acoust., Speech, Signal Processing*. Vol. 3 (May 1996).

3 Selesnick, I.W., M. Lang, and C.S. Burrus. “Constrained Least Square Design of FIR Filters without Specified Transition Bands.” *Proceedings of the IEEE Int. Conf. Acoust., Speech, Signal Processing*. Vol. 2 (May 1995). Pgs. 1260-1263.