Design of Non-recursive Digital Filters

Contents

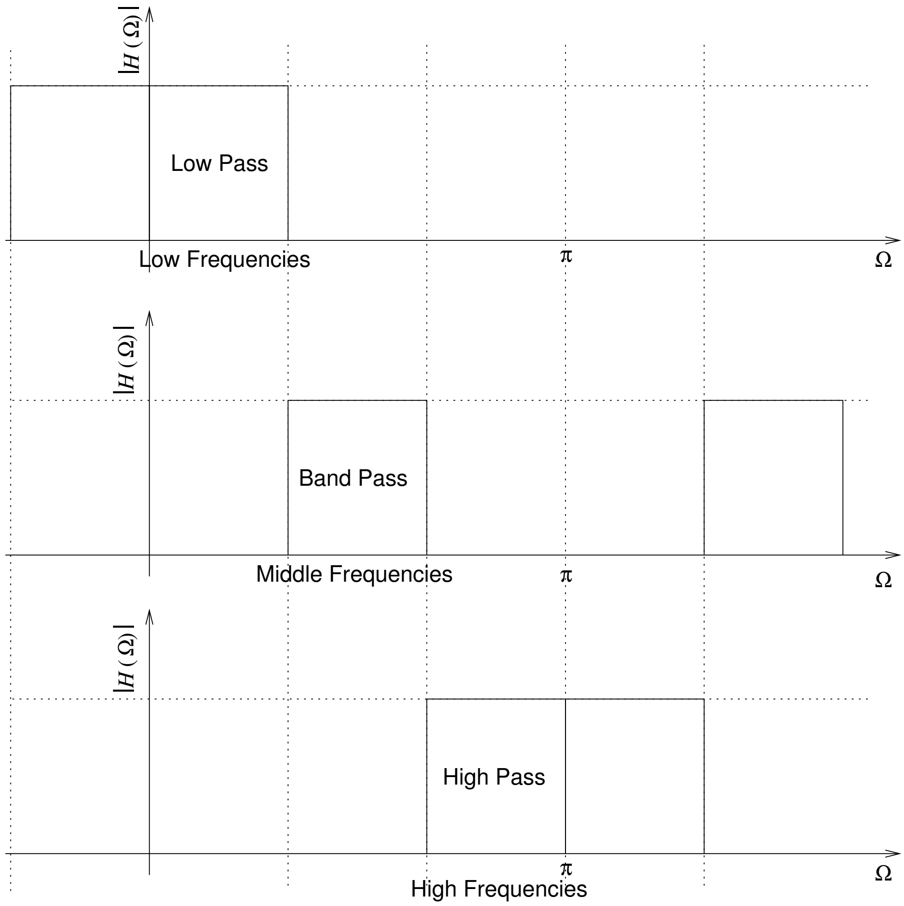
1 Introduction

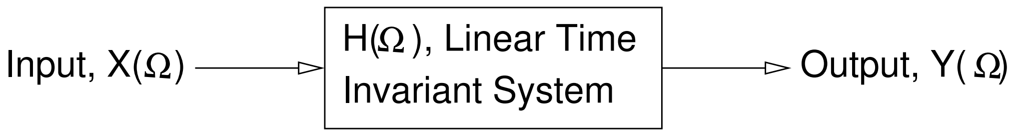
1.1 What is a Digital Filter?

Any digital system can be described as a digital filter. The word "*filter*" means to remove a part of a signal and allow another part to pass through.
The verb "*to filter*" is used in many areas of English language.
Examples

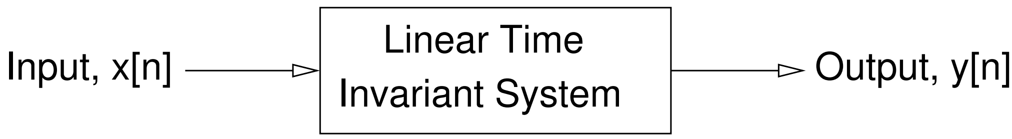
- The water filter cleans the water for drinking.
- The cook filtered the bad fruit from the good for cooking later.
- The air conditioning filters dust from the air.

Often used to remove some frequencies from a signal $X(\Omega)$ and to allow other frequencies to pass through to the output $Y(\Omega)$.





1.2 Non-recursive digital filters



What is a *non-recursive digital filter*?

- "*Recursive*" comes from the word "to recur"; Meaning: to repeat

Examples
A recursive filter uses past output values ($y[n-i]$) for the current output $y[n]$:

- *Recursive Filter Example*

$$y[n] = 0.5y[n-1] + 0.5x[n].$$

A non-recursive filter only uses input values $x[n-i]$:

- *Non-recursive Filter Example*

$$y[n] = 0.5x[n-1] + 0.5x[n].$$

Non-recursive digital filters are often known as **Finite Impulse Response (FIR)** Filters as a non-recursive digital filter has a finite number of coefficients in the impulse response $h[n]$.
Recursive digital filters are often known as **Infinite Impulse Response (IIR)** Filters as the impulse response of an IIR filter has an infinite number of coefficients.

FIR Filters Vs IIR Filters

- FIR filters Have linear phase characteristics (*i.e.* no phase distortion);
- But FIR filters typically require a higher number of computations in comparison to IIR filters because IIR filters can rely on performing operations on past output values as well as the operations on the input and past input values.

1.3 Generalised Non-Recursive Difference Equation

Recall the generalised difference equation for causal LTI systems:

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{N} b_k x[n-k]$$

So a non-recursive digital filter in a causal LTI system is given by:

$$y[n] = \sum_{k=0}^{N} b_k x[n-k]$$

2 Moving Average Filters

Moving average filters are usually implemented non-recursively. Moving average filters are interesting as they:

- Are useful for some applications
- But the frequency response is not ideal

Typically a moving average filter will weight all inputs the same and just average over the current input in combination with past input values and divide by the number of terms being averaged over. If it is a non-causal moving average filter then it will average over future inputs as well. A non-causal moving average filter has impulse response of the form:

$$h[n] = \begin{cases} \frac{1}{k} & \text{if} \quad -k/2 \le n \le k/2 \\ 0 & \text{otherwise} \end{cases} \text{; where } k \text{ is odd.}$$

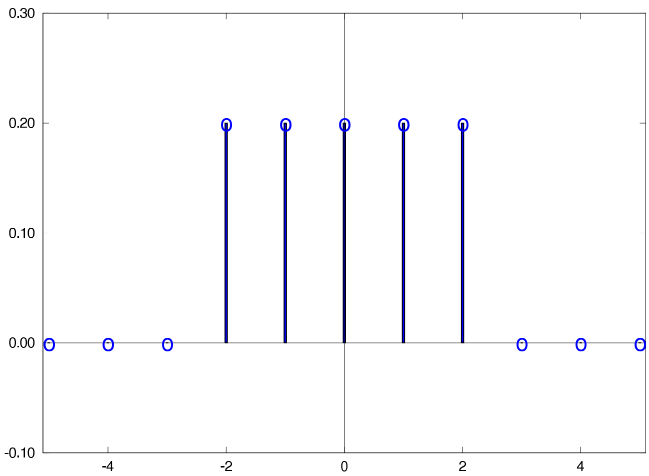
A causal moving average filter has impulse response of the form:

$$h[n] = \begin{cases} \frac{1}{k} & \text{if} \quad 0 \le n \le k-1 \\ 0 & \text{otherwise} \end{cases} \text{; where } k \text{ is odd.}$$

Example
Consider the impulse response (in the time domain) for a non-causal moving average filter that averages over 5 input values, *i.e.* $k = 5$. The equation for such a filter is given by:

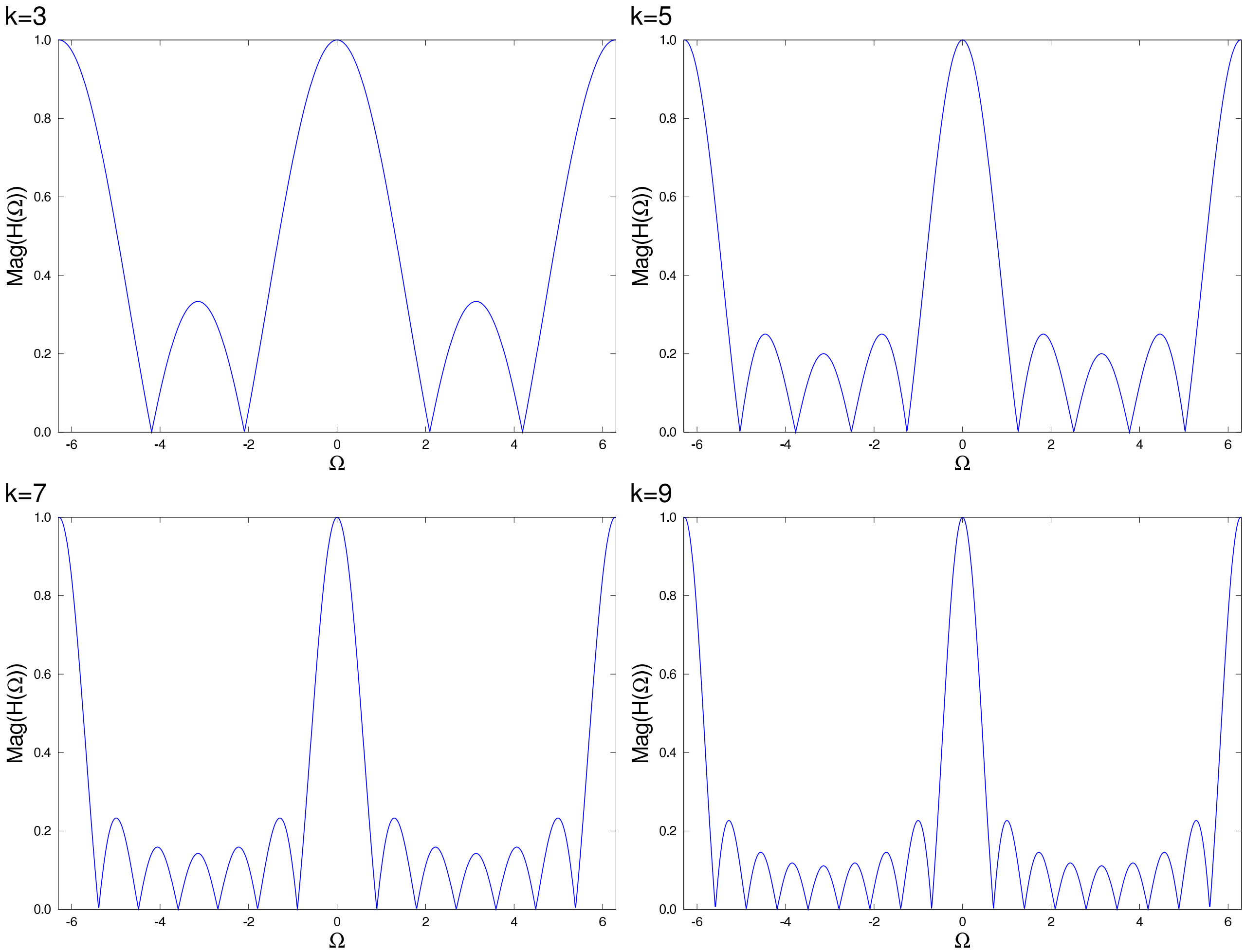
$$h[n] = \begin{cases} 0.2 & \text{if} \quad -2 \le n \le 2 \\ 0 & \text{otherwise} \end{cases} .$$

We can visualize this moving average filter too:

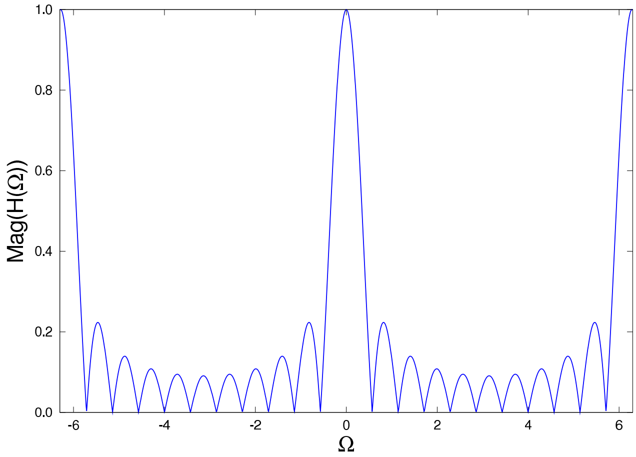


2.1 Moving Average Frequency Response

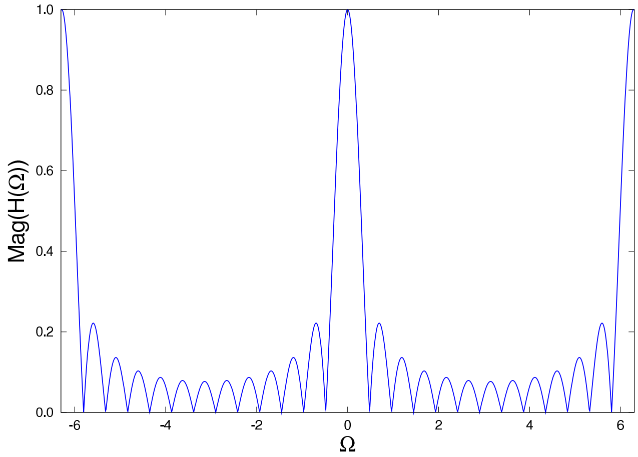
The frequency response of a moving average filter is also of interest. The frequency response tells us the range of frequencies that a filter will affect if it receives an input signal. If the frequency response is close to zero for a range of frequencies then the output of the filter will greatly attenuate the signal for an input signal consisting of those frequencies. If the frequency response is close to unity for another range of frequencies then the input signal will not be changed for those frequencies. This is useful because a signal may consist of useful signal components at some ranges of frequencies but other frequencies may not be useful or may be a problem. For example, a reconstruction filter after Digital to Analogue conversion (DAC) is often required. A Moving average filter could filter the output of a DAC to smooth out the steps associated with the quantisation levels of the signal when it was previously in digital form.



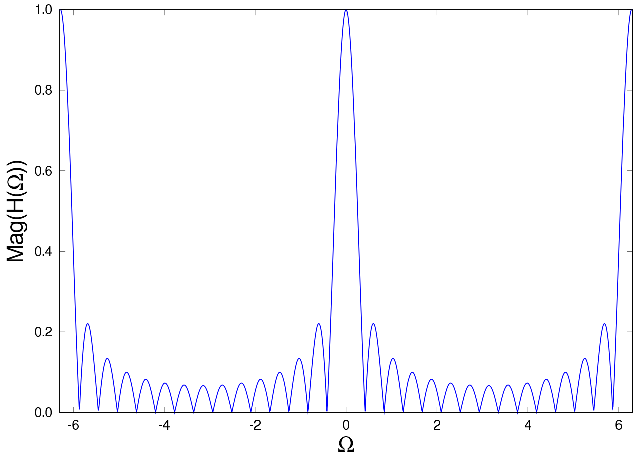
k=11



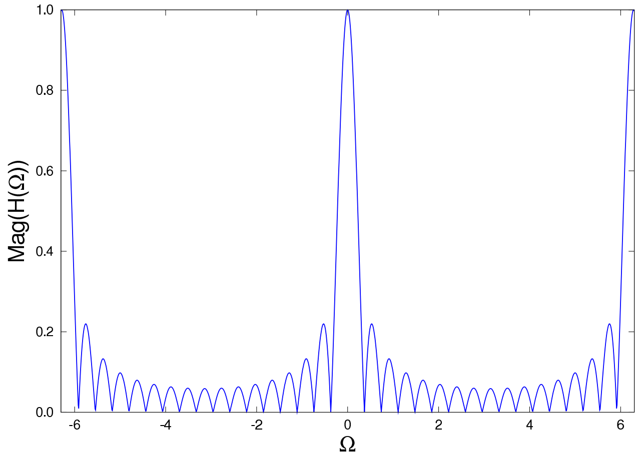
k=13



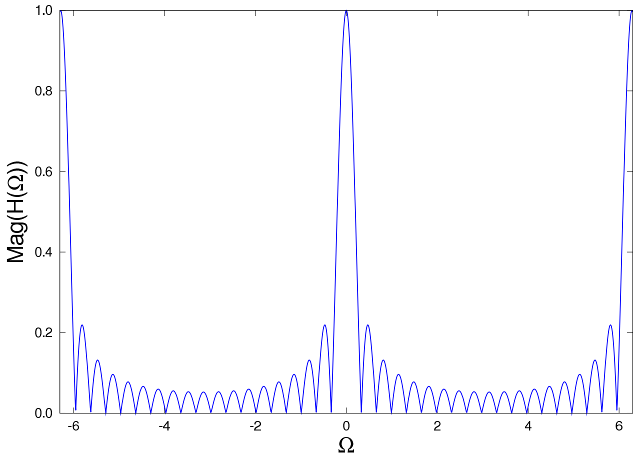
k=15



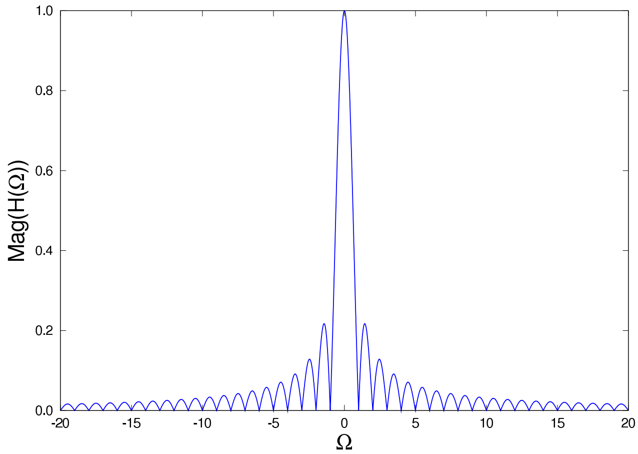
k=17



k=19



sinc function



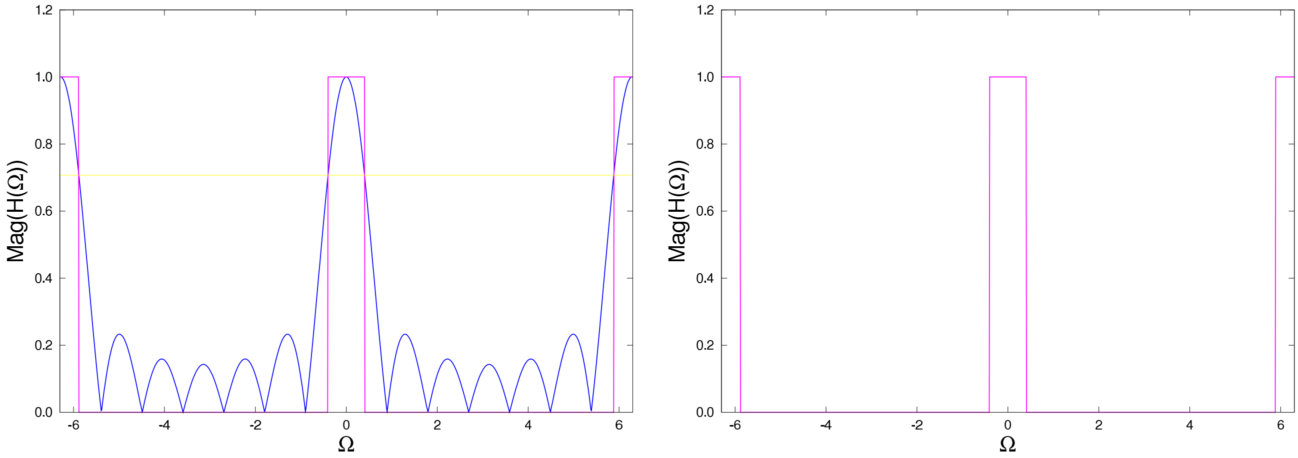
Here in the last plot the Sinc function is shown. The frequency response tends to a Sinc function which is a sine function divided by the argument of the sine, defined here by

$$\operatorname{sinc}(x) = \frac{1}{\pi x}\sin(\pi x).$$

The above frequency plots for a variety of different moving average filters are not ideal.

2.2 Moving Average and Ideal Frequency Responses (Low Pass)

Moving average ($k = 7$) frequency response (blue) and ideal low pass frequency response (magenta). Ideal low pass frequency response only.



The moving average frequency response is not ideal. The ripples allow many other frequencies to pass through the filter. The ideal low pass frequency response is perfect, stopping all unwanted frequencies. But it is rarely achievable.

3 Ideal Frequency Response

3.1 Ideal Frequency Response Time Domain Representation

What is the time domain representation for the ideal low pass frequency response?
Recall the inverse Fourier Transform is given by:

$$x[n] = \frac{1}{2\pi}\int_{2\pi} X(\Omega)\exp(j\Omega n)\mathrm{d}\Omega. \tag{1}$$

So for an impulse response $h[n]$:

$$h[n] = \frac{1}{2\pi}\int_{2\pi} H(\Omega)\exp(j\Omega n)\mathrm{d}\Omega; \tag{2}$$

or sometimes simpler to use

$$h[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(\Omega)\exp(j\Omega n)\mathrm{d}\Omega; \tag{3}$$

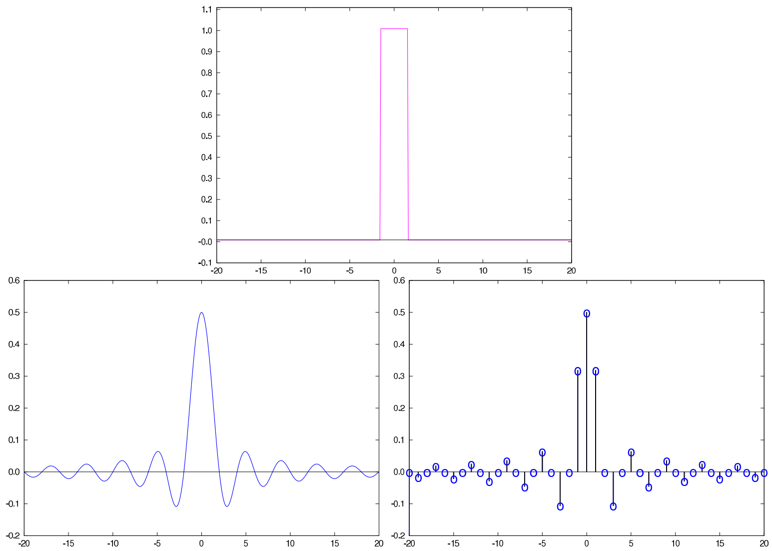
which is possible due to periodicity.
The ideal low pass frequency response is given by:

$$H(\Omega) = \begin{cases} 1 & \text{for} \quad -\Omega_1 \le \Omega \le \Omega_1, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

This equation states that any frequency Ω between $-\Omega_1$ and $+\Omega_1$ can pass without attenuation ($H(\Omega) = 1$). All other frequencies are stopped altogether. The time domain representation of (4) is given by

$$\begin{aligned} h[n] =& \frac{1}{2\pi}\int_{-\pi}^{\pi} H(\Omega)\exp(j\Omega n)\mathrm{d}\Omega = \frac{1}{2\pi}\left(\int_{-\pi}^{-\Omega_1} 0\mathrm{d}\Omega + \int_{-\Omega_1}^{\Omega_1} 1\times\exp(j\Omega n)\mathrm{d}\Omega + \int_{\Omega_1}^{\pi} 0\mathrm{d}\Omega\right) \\ =& \frac{1}{2\pi}\left([0]_{-\pi}^{-\Omega_1} + \left[\frac{\exp(j\Omega n)}{jn}\right]_{-\Omega_1}^{\Omega_1} + [0]_{\Omega_1}^{\pi}\right) = \frac{1}{2\pi}\left[\frac{\exp(j\Omega n)}{jn}\right]_{-\Omega_1}^{\Omega_1} \\ =& \frac{1}{2\pi jn}\left(\exp(j\Omega_1 n) - \exp(-j\Omega_1 n)\right) = \frac{1}{2\pi jn}2j\sin(\Omega_1 n) = \frac{1}{\pi n}\sin(\Omega_1 n) \\ =& \frac{\Omega_1}{\pi}\mathrm{sinc}(n\Omega_1) \end{aligned}$$

where $\mathrm{sinc}(x)$ is known as the sinc function.



Example

Q. What is the time domain impulse response for an ideal low pass frequency response filter with $\Omega_1 = \pi/2$?

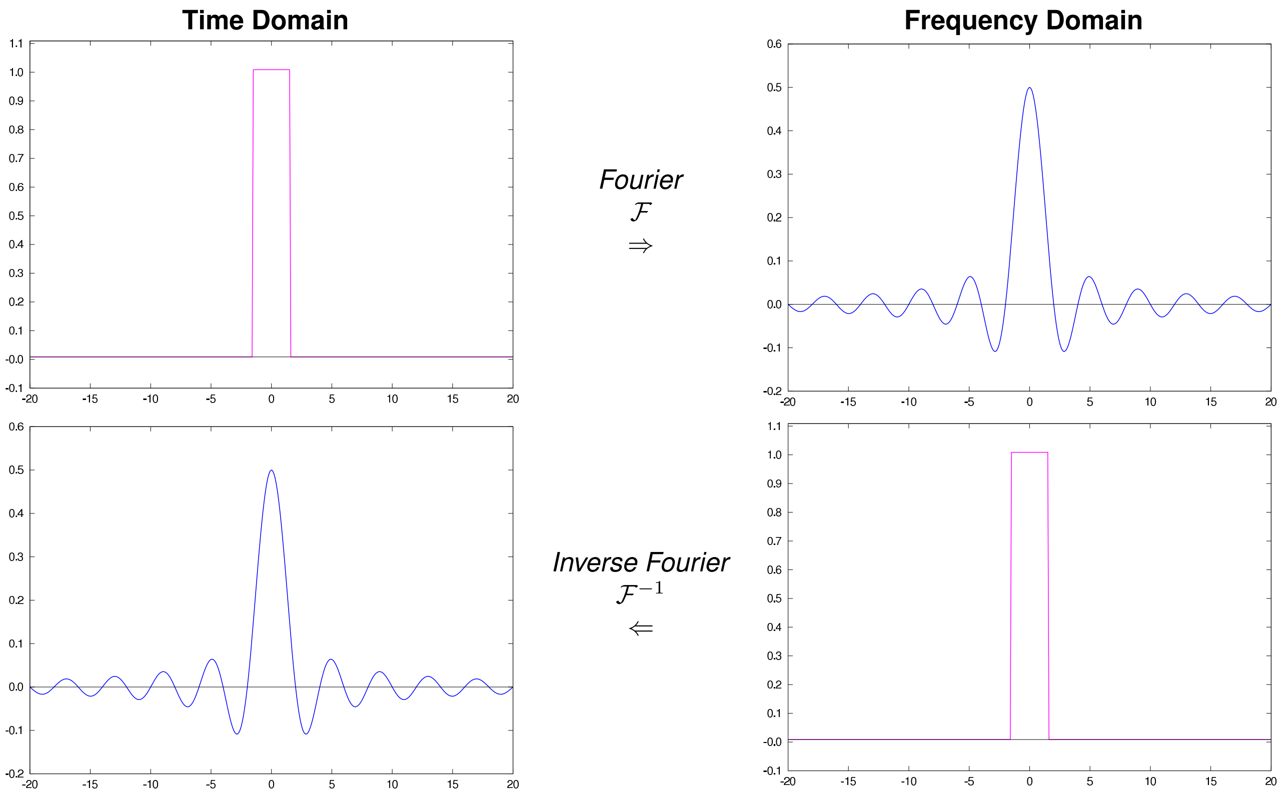
A.

$$
\begin{aligned}
h[n] =& \frac{\pi}{2\pi}\mathrm{sinc}\left(n\frac{\pi}{2}\right) = \frac{1}{2}\mathrm{sinc}\left(\frac{n\pi}{2}\right) \\
=& \frac{2}{n\pi}\frac{1}{2}\sin\left(\frac{n\pi}{2}\right) \\
=& \frac{1}{n\pi}\sin\left(\frac{n\pi}{2}\right).
\end{aligned}
$$

□

3.2 Overview of the Sinc Function

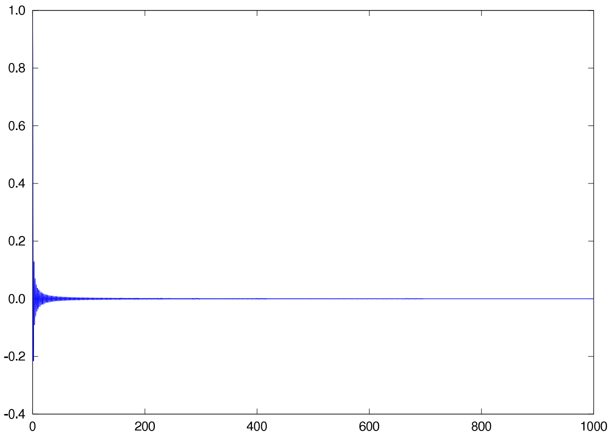
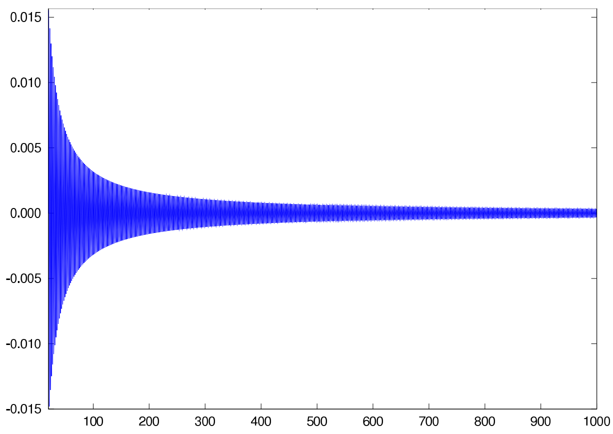
- The Sinc function is the Fourier transform of a square wave or impulse response.



- The Sinc function is also the inverse Fourier transform of a square wave.

The frequency response of a moving average filter with uniform weights is important. It helps us to understand that if we want to remove frequencies above a certain cut-off point then some filter design process is required to be followed to enable us to design a filter with a suitable specification. The converse is also true. We can see that the result of an ideal frequency response results in a sinc function.

However the Sinc function continues forever.

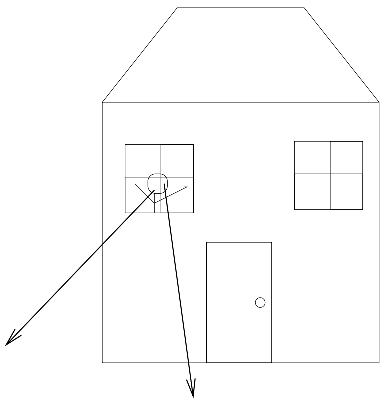



Therefore the time domain representation has to be stopped early which also leads to a non-ideal frequency response. So there are a number of design trade offs to be made. Windowing is the technique used to remove part of the sinc function.

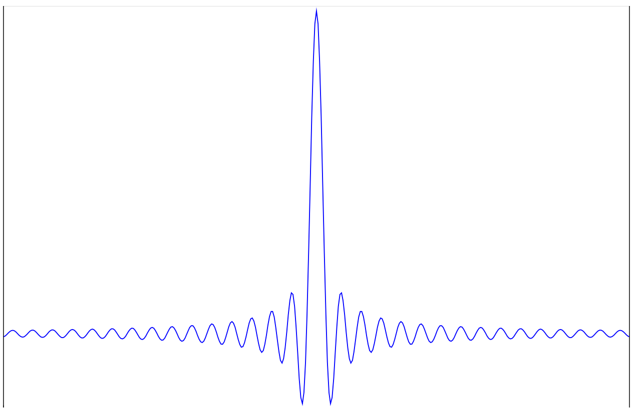
4 Windowing

4.1 How to Stop the Sinc function early?

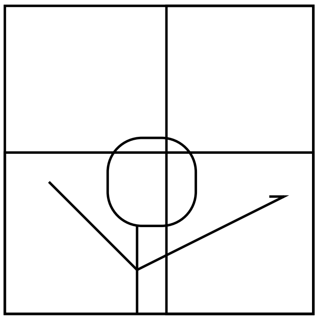
A window can be used...



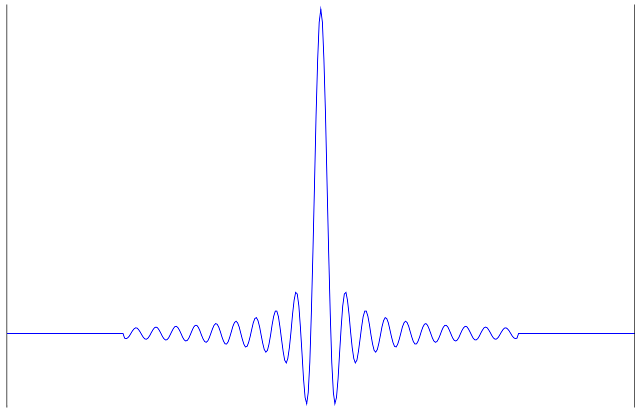
... to limit the impulse response.



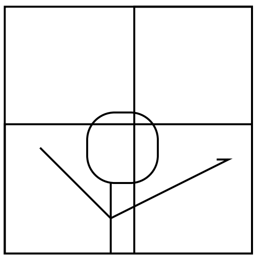
Big Window



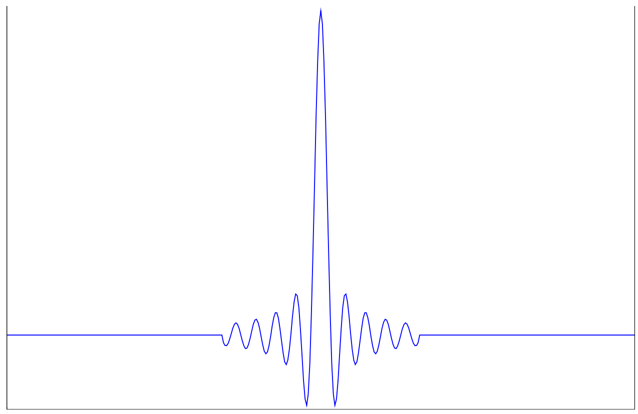
Windowing
$w \times \mathrm{sinc}(x)$
$\Rightarrow$



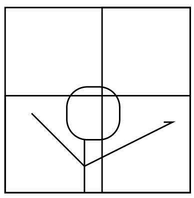
Medium Window



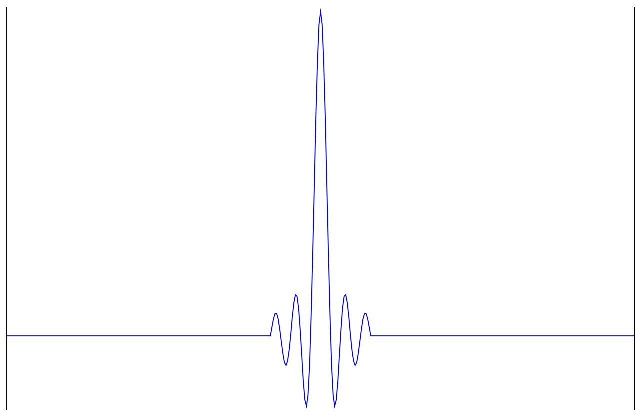
Windowing
$w \times \mathrm{sinc}(x)$
$\Rightarrow$



Small Window



Windowing
$w \times \mathrm{sinc}(x)$
$\Rightarrow$



The window here $w[n]$ is rectangular and constant, *i.e.*

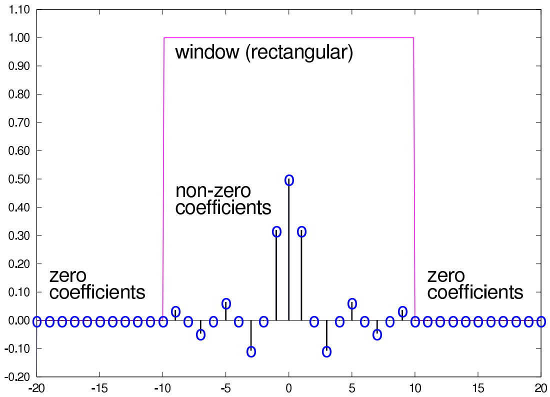
$$w[n] = \begin{cases} 1 & \text{for } -\frac{\text{window width}}{2} \leq n \leq \frac{\text{window width}}{2} \\ 0 & \text{every where else} \end{cases}$$

usually known as a rectangular window.
The sinc function filter coefficients are multiplied by this window (in the time domain):

$$h_2[n] = h[n] \times w[n].$$

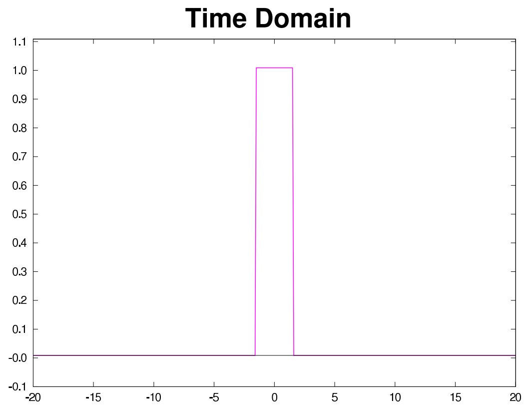
So that

$$h_2[n] = \begin{cases} h[n] & \text{for } -\frac{\text{window width}}{2} \leq n \leq \frac{\text{window width}}{2}, \\ 0 & \text{every where else.} \end{cases}$$

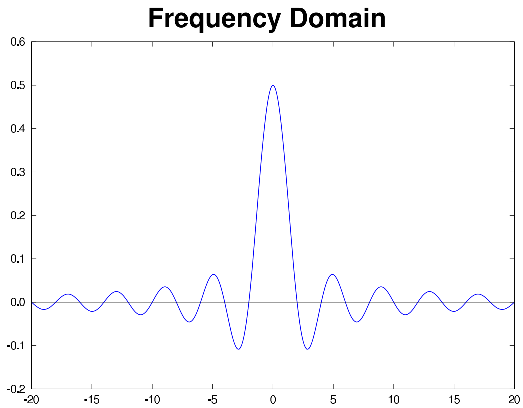


4.2 Rectangular Window Problems!

HOWEVER! As we already know, the frequency response of this window is not ideal.
Rectangular pulse in time domain = sinc function in frequency domain:



Fourier
$\mathcal{F}$
$\Rightarrow$



- The original ideal low pass filter in the frequency domain is corrupted by the windowing[1].

Time Domain		**Frequency Domain**
Convolution $x[n] * y[n]$	$\Longleftrightarrow$	*Multiplication* $X(\Omega) \times Y(\Omega)$
Multiplication $x[n] \times y[n]$	$\Longleftrightarrow$	*Convolution* $X(\Omega) * Y(\Omega)$

Time domain multiplication of $w[n]$ with $h[n]$ is the same as convolution in the frequency domain, *i.e.*

$$\mathcal{F}(w[n] \times h[n]) = W(\Omega) * H(\Omega).$$

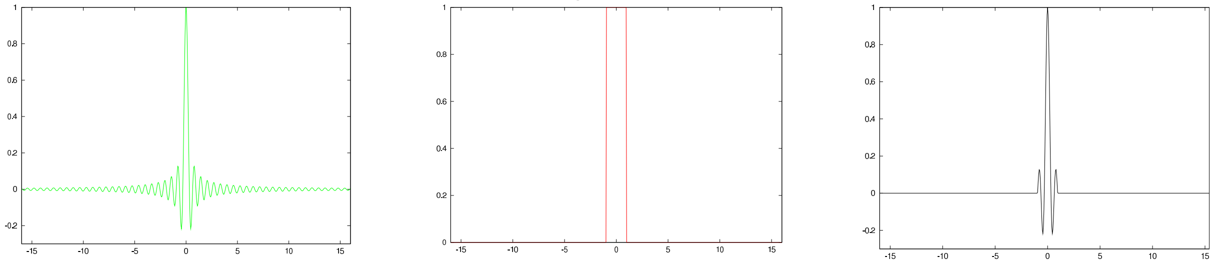
Time domain		**Frequency Domain**
$w[n]$	$\Longleftrightarrow$	$W(\Omega)$
$h[n]$	$\Longleftrightarrow$	$H(\Omega)$
$h_2[n]$	$\Longleftrightarrow$	$H_2(\Omega)$

$$H_2(\Omega) = \mathcal{F}(h_2(n)) = W(\Omega) * H(\Omega).$$

[1] Windowing describes multiplication by the window function.

4.3 Frequency Domain Effect of (Rect.) Windowing Ideal Low Pass Filter Example

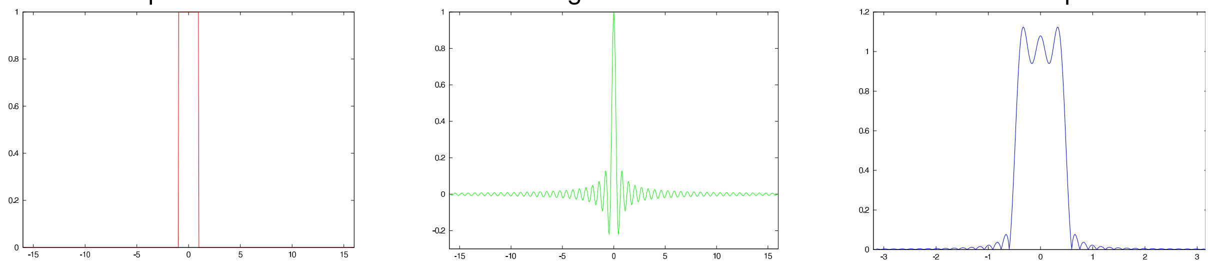
Ideal Time Domain Filter | Rectangular Window | Windowed Time Domain



$h[n]$	$\times$ (*multiplication*)	$w[n]$	=	$h_2[n]$

We can view the effect of cutting the sinc function short in the frequency domain by performing a Fourier transform on the windowed time domain $h_2[n]$ function.

Ideal Freq. Domain Filter | Rectangular Window | Windowed Freq. Domain

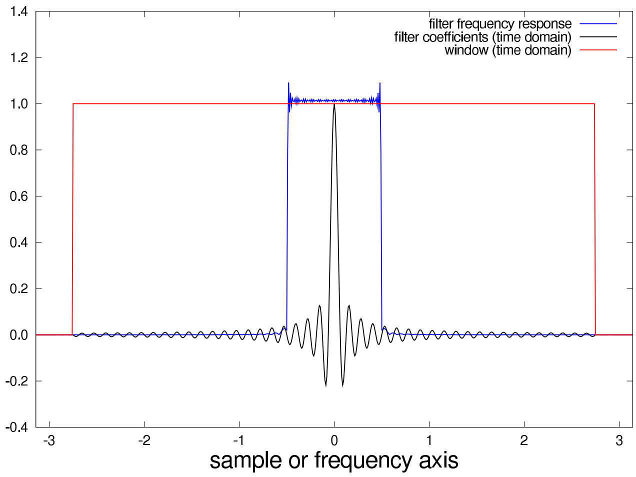


$H(\Omega)$	$*$ (*convolution*)	$W(\Omega)$	=	$H_2(\Omega)$

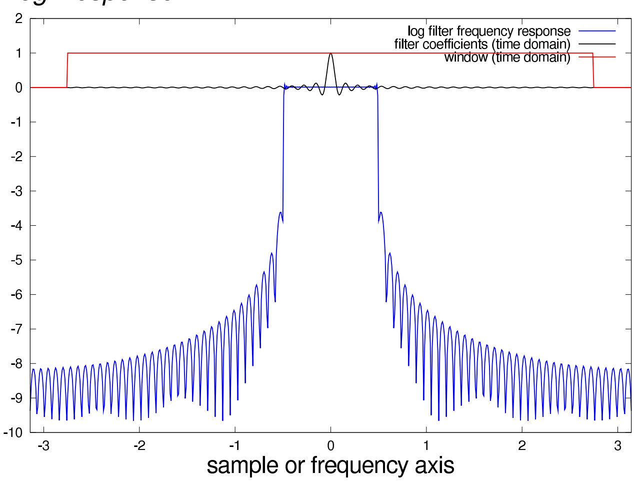
4.4 Frequency Domain Effect of (Rect.) Windowing Ideal Low Pass Filter Further Examples

Here are some further examples of the effect of cutting the sinc function off with varying sizes of rectangular window. Note here that the log response is used. The log of the frequency response enables the side lobes in the transition and stop bands to be emphasised.

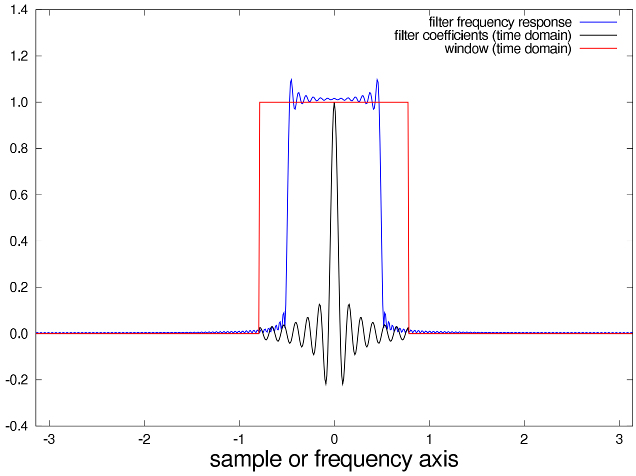
Big Window



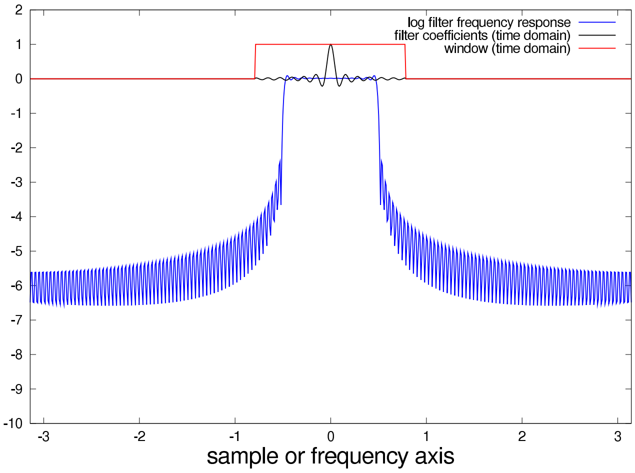
Log Response



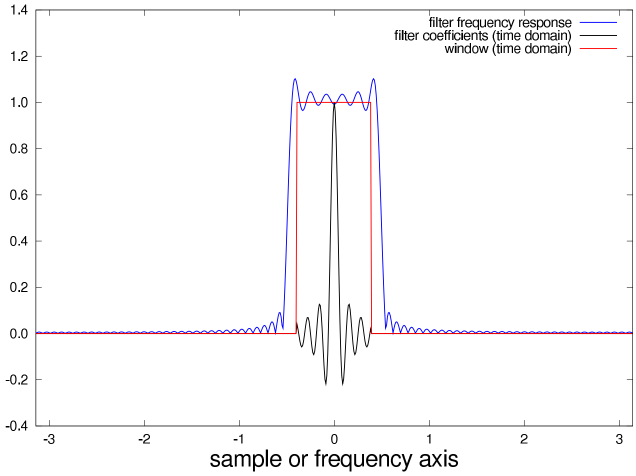
Smaller Window



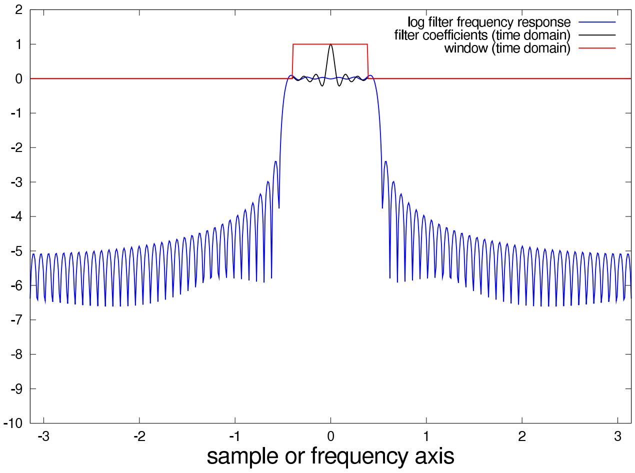
Log Response



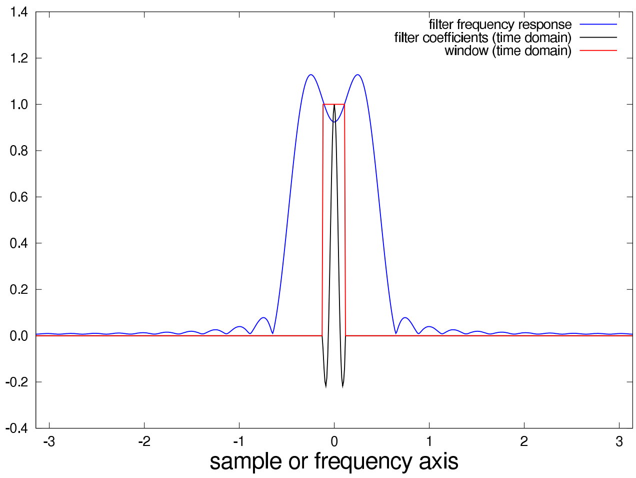
Even Smaller Window



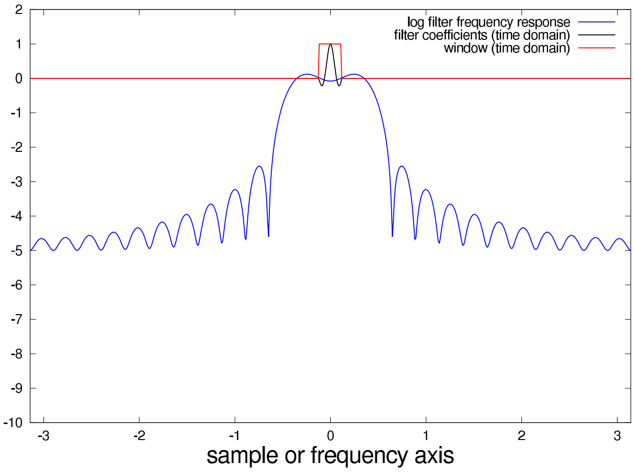
Log Response



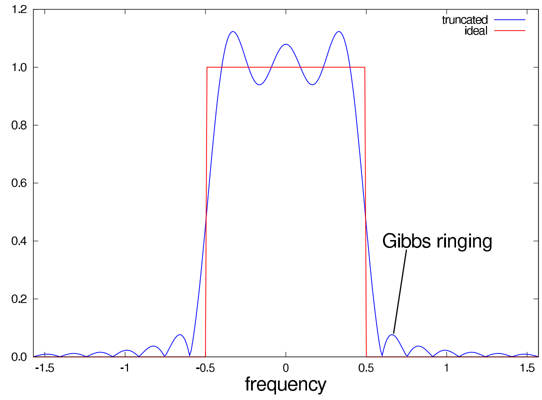
Very Small Window



Log Response



Gibbs Ringing or Truncation Artifact The side lobes or bumps either side of the main lobe are often referred to as a Gibbs artifact or truncation artifact.



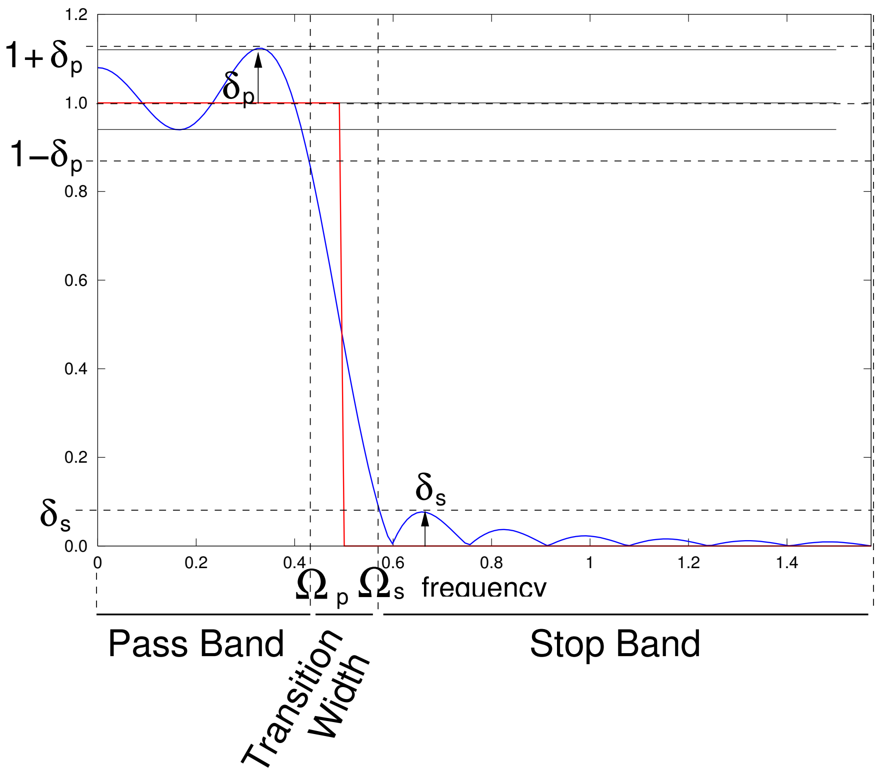
5 Filter Parameters

5.1 Filter Parameters

A filter can be described by a number of *parameters*:

- δ_p: pass band ripple
- δ_s: stop band ripple
- Pass Band (little attenuation)
- Transition width
- Stop Band (highest attenuation)

These parameters are illustrated by means of an example here:



In the example shown here we have:

- A pass band ripple value of $\delta_p = 0.12$ and
- A stop band ripple value of $\delta_s = 0.08$

Therefore the gain at the end of the pass band is $1 - \delta_p = 0.88$. Gain and attenuation in decibels (dB):

$$\text{gain}_{\text{dB}} = 20 \log(\text{ gain }).$$

So the gain at the end of pass band

$$20 \log(1 - \delta_p) = -1.11\text{dB}.$$

The gain at the end of the stop band is $1 - \delta_s = 0.08$ or -21.94dB.

The pass band frequencies are given by:

$$\Omega = 0 \text{ to } \Omega_p$$

where $\Omega_p = 0.43$ radians.

The transition width is given by:

$$\Omega = \Omega_p \text{ to } \Omega_s =$$

$$0.43 \text{ to } 0.56 \text{ radians.}$$

And the stop band starts at:

$$\Omega = \Omega_s = 0.56 \text{ radians.}$$

5.2 Filter Bandwidth

The filter bandwidth is defined by:

"Range of frequencies the filter gain is greater than -3dB".

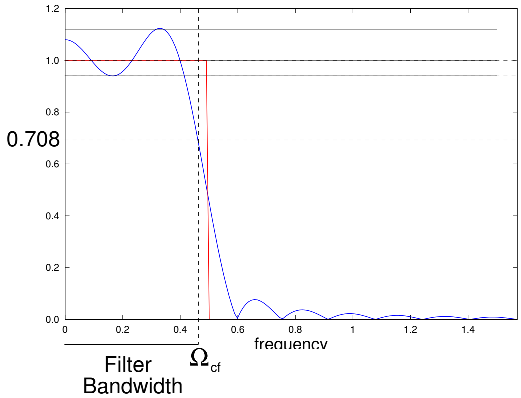
As

$$\text{gain}_{\text{dB}} = 20 \log(\text{ gain }).$$

Then

$$\text{gain} = 10^{\left(\frac{\text{gain}_{\text{dB}}}{20}\right)}.$$

So for $\text{gain}_{\text{dB}} = -3\text{dB}$ then $\text{gain} = 0.708$.



- The cut-off frequency Ω_{cf} corresponds to when the gain falls below -3dB.

6 Alternative Windows for Truncation

6.1 Other Window Types for Truncation (Other than Rectangle)

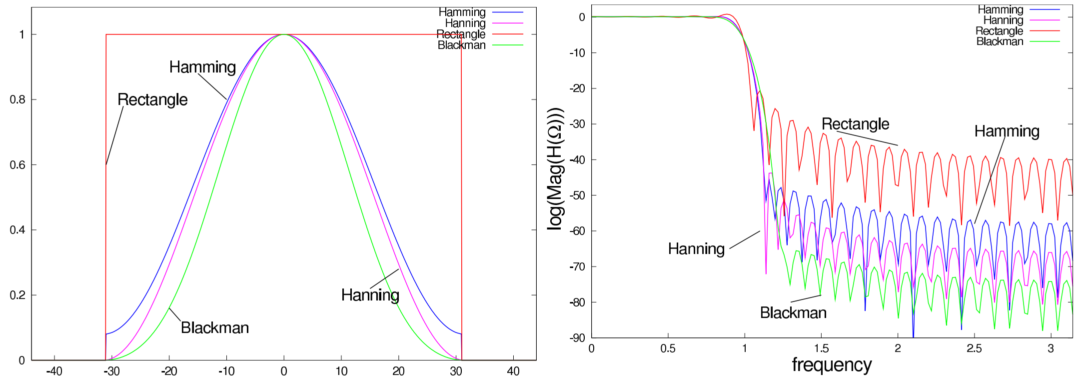
Other types of window functions can be used to truncate the ideal time domain response:

Window Name	Equation
Bartlett/ triangular	$w[n] = \frac{(N+1)-\|n\|}{(N+1)^2}$ for $\|n\| \leq (N-1)/2$
Hamming	$w[n] = 0.54 + 0.46 \cos\left(\frac{2\pi n}{N-1}\right)$ for $\|n\| \leq (N-1)/2$
Hanning	$w[n] = 0.5 + 0.5 \cos\left(\frac{2\pi n}{N-1}\right)$ for $\|n\| \leq (N-1)/2$
Kaiser	$w[n] = \frac{1}{I_0(\alpha)} I_0\left(\alpha\sqrt{1 - \left(\frac{2n}{N-1} - 1\right)^2}\right)$ for $\|n\| \leq (N-1)/2.$

Others include Blackman, Lanczos and Tukey windows.

6.2 Example Frequency Responses with Different Windows

Each window is slightly different in the time domain. Windowed truncation of a filter ideal frequency response.



- Each window type has different properties.
- *e.g.* Stop band attenuation for Blackman is highest but Blackman has long transition width.

Some example window parameters that can be used in the design of FIR filters can be seen below:

Window	**Highest side-lobe** (dB)	**side lobe fall off** (dB/oct)	**coherent gain**	**3dB B/W** (bins)	**6dB B/W** (bins)	**Scalloping loss** (dB)
Rectangular	-13	-6	1	0.89	1.21	3.92
Bartlett	-27	-12	0.5	1.28	1.78	1.82
von Hann	-32	-18	0.5	1.44	2	1.42
Hamming	-43	-6	0.54	1.3	1.81	1.78
Dolph-Chebyshev ($\alpha = 2.5$)	-50	0	0.53	1.33	1.85	1.7

7 Filter Design

There are a number of different commonly found techniques that enables a filter to be designed. The technique depends on the information to hand and the requirements of the filter. A relatively simple filter design technique is outlined below which involves the steps that we have followed already but extended to include windowing with one of the windows that we have compared.

7.1 FIR Low Pass Filter Design Steps

Here is a design strategy for designing a low pass FIR filter. A low pass FIR filter only allows low frequencies to pass. Shortly we will see how to create bandpass and high pass filters as well.

FIR Low Pass Filter Design Steps

1. Find the cut-off digital frequency, Ω_{cf}
   (a) It may be given directly,
      - *e.g.* $\Omega_{\text{cf}} = \pi/4$ radians
   (b) Or the sampling frequency and cut off frequencies may be given instead, calculated from $\Omega = 2\pi f/f_s$,
      - *e.g.* $f_s = 100$kHz and $f_{\text{cf}} = 12.5$kHz, so that $\Omega_{\text{cf}} = 2\pi 12500/100000 = \pi/4$ radians.
2. Calculate the appropriate sinc function for ideal low pass filter:

$$h[n] = \frac{\Omega_1}{\pi}\text{sinc}(n\Omega_1)$$

   where $\Omega_1 = \Omega_{\text{cf}}$ is the cut-off frequency.

3. Select a window with appropriate parameters. *e.g.* short transition width or high stop band attenuation.
4. Calculate non-causal time domain impulse response from $h_2[n] = w[n] \times h[n]$.
5. Shift the impulse response to make a causal version $h_3[n] = h_2[n - (N-1)/2]$.

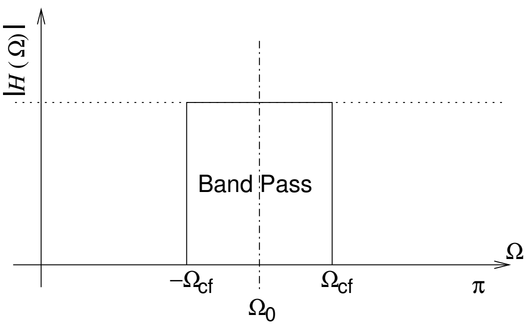
8 High Pass and Band Pass FIR Filters

8.1 Band Pass FIR Filter Design

Low pass filter $H(\Omega)$ can be converted to a bandpass filter by:

- Convolution in frequency domain with delta function δ_{Ω_0};
- At centre frequency Ω_0.

Frequency Domain		**Time Domain**
(*Convolution*)		(*Multiplication*)
$H(\Omega) * \delta_{\Omega_0}$	$\Longleftrightarrow$	$h[n] \times \cos(n\Omega_0)$



The resulting band pass impulse response $h'[n]$

- with bandwidth $2 \times \Omega_{\mathrm{cf}}$,
- using window function $w[n]$,
- and centre frequency Ω_0

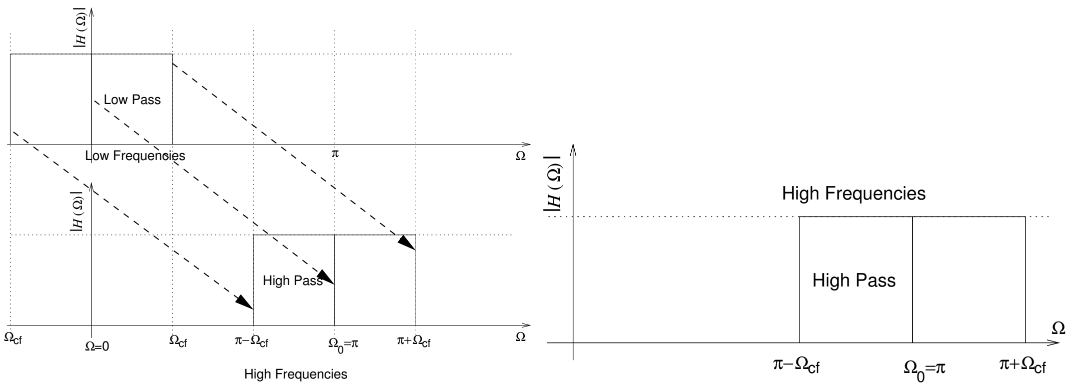
is given by

$$\begin{aligned} h'[n] &= h[n] \times w[n] \times \cos(n\Omega_0) \\ &= \frac{\Omega_{\mathrm{cf}}}{\pi} \mathrm{sinc}(n\Omega_{\mathrm{cf}}) \times w[n] \times \cos(n\Omega_0). \end{aligned}$$

8.2 High Pass FIR Filter

High pass filter can be achieved by

- Shifting the impulse response to $\Omega_0 = \pi$,
- Via multiplication by $\cos(n\Omega_0) = \cos(n\pi)$.
- This is the limit of the unique part of the digital spectrum.



The resulting high pass impulse response $h'[n]$

- with bandwidth Ω_{cf},
- using window function $w[n]$,

is given by

$$\begin{aligned} h'[n] &= h[n] \times w[n] \times \cos(n\pi) \\ &= \frac{\Omega_{\text{cf}}}{\pi} \text{sinc}(n\Omega_{\text{cf}}) \times w[n] \times \cos(n\pi). \end{aligned}$$

8.3 Other Topics in Filter Design

Band stop is another type of filter,

- Created from a combination of high and low pass filters.

Digital differentiators are common in DSP applications,

- To differentiate a signal, to calculate *e.g.* speed.

Other techniques for FIR filter design include:

- Equiripple filters
  - ◇ Optimization of passband and stopband ripples.
- Frequency sampling method
  - ◇ Optimization from specified (sampled) frequency response.

Other factors to consider

- Phase response of the filter, not just the magnitude.
  - ◇ *e.g. Hilbert transformer* places a 90^o phase shift on a signal.

9 Summary

The following topics have been covered:

- Moving average filters (and their frequency response)
- Ideal frequency response of FIR filters
- Windowing techniques
- Filter parameters
- Filter design techniques for
  - ◇ low pass,
  - ◇ band pass
  - ◇ and high pass FIR filters