Course Report

Digital Signal Processing

Assoc. Prof. Lin Jingran

Filter specifications and Design of digital filter based on the given specifications.

Umer Khalid (Pakistan)
201424020108
School of Automation Engineering

1- Introduction:

Imagine your friend is calling you from a place where there is a lot of background noise. Off course you will have problems in understanding him but thanks to the development of filters which allow for the cancellation of these kinds of unwanted signals. This is just one example of the millions of applications in which they are used. So filters are everywhere…

There are two main types of filters:

- Analog Filters
- Digital Filters

1.1-Why Digital Filters:

Digital Filters have various advantages over analogue ones such as:

- Higher accuracy
- Easier Testing (as software based)
- Cheap (compared to analog ones in complex systems)
- No temperature or humidity effects

Though there are a lot of advantages but it does not mean that analog filters are abandoned now. They are still used especially in high frequency applications.

2- Digital Filter Design:

A digital filter can be totally characterized by its transfer function which tells us how it will respond to a particular input. So designing a filter means to find a realizable transfer function according to the given frequency response specifications. According to Sanjit K.Mitra "The process of deriving the transfer function is called digital filter design".

2.1-Filter Specifications:

When we talk about specifications, we generally design only for magnitude specs as phase can be compensated by all pass filters. As the filters impulse response is infinite and non-causal so it can't be implemented practically unless we cut down the response leading to non-ideal behavior i.e. there will be gradual roll off rather than a sharp transition and there will be ripples. So the specs are given with some acceptable tolerances. In figure 1, δ_p indicates the magnitude of the pass band ripple whereas δ_s indicates the magnitude of the stop band ripple. Pass band ripple is generally specified as peak to peak whereas it is not the case in stop band ripples as the maximum magnitude is more important than how jerky it is in this band.

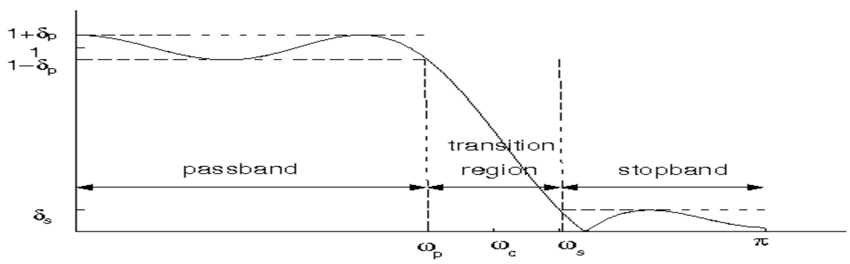

Figure 1

Moreover, ω_p is the pass band edge frequency, ω_s is the stop band edge frequency and ω_c is the cutoff frequency.

In the pass band defined by 0≤$\omega \leq \omega_c$, we want the magnitude to be approximately unity with an error of ±δ_p i.e.

$$1-\delta_p \leq |G(e^{j\omega})| \leq 1+\delta_p \quad \text{for } |\omega| \leq \omega_p$$

In the stop band we want the magnitude to be approximately 0 with a maximum peak /error of δ_s i.e.

$$|G(e^{j\omega})| \leq \delta_s \qquad \text{for } \omega_c \leq |\omega| \leq \pi$$

You may note in figure 1 that the response is given for frequencies 0≤ $|\omega| \leq \pi$.This is because the frequency response is a periodic function of ω and the magnitude response is even function of ω.

We often measure the pass band ripple and stop band ripple in decibels, as shown in the following equations:

$$\text{Pass band ripple} = -20\log_{10}(1-\delta_p)\ \text{dB}$$

$$\text{Stop band ripple} = -20\log_{10}(\delta_s)\ \text{dB}$$

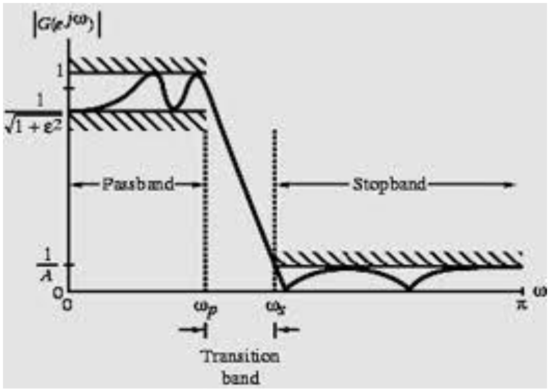

Figure 2

We can also give the specifications in normalized form by considering the maximum value of the pass band magnitude as unity as in figure 2. Here, $1/\sqrt{1+\varepsilon^2}$ denotes maximum pass band deviation and $1/A$ denotes the maximum stop band magnitude. The maximum pass band attenuation in dBs can be given by:

$$\propto_{max} = 20log_{10}(\sqrt{1+\varepsilon^2})\ \text{dB}$$

As all the filters are designed using normalized angular frequencies, the specified critical frequencies need to be normalized before using the corresponding design technique.

2.2-Important Considerations:

Before we start to design based on the given specifications, we need to select the type of filter among the two famous types:

- Infinite Impulse Response Filters
- Finite Impulse Response Filters

Each has its own pros and cons and we need to make the decision based on the problem at hand. Here the advantages of both are compared to make the decision easy.

IIR Filters	FIR Filters
They can be built with less order hence cheaper and lesser hardware complexity.	They can achieve exactly linear phase
As they are built using analog prototypes, we can make use of the highly advanced techniques developed for design of analog filters.	They are always stable
Some design tables are already available for analog filters.	

Now after we have understood the specifications and selected the filter type we have to design it using a variety of available techniques.

3- IIR Filter Design:

IIR digital filters are also known as recursive filters because of the presence of feedback. This feedback is the reason for instability. So, while designing an IIR filter we have to take care of the stability issue. The design problem then becomes finding a realizable as well as stable transfer function or the difference equation. The equation is then implemented using various filter structures.

The approach which we use is as follows:

- Define digital filter specifications
- Select a prototype analog filter
- Select a transformation to be used
- Convert the digital specs into analog ones
- Design analog filter
- Convert it into digital one using the selected transformation
- If a filter other than low pass is desired, use the frequency transformations to convert them into required filter

We have already discussed the specifications, the analog filter can be chosen based on whether we want to minimize ripples or the transition gap. If we want to minimize ripple we can choose Butterworth as it has the flattest response in the pass band whereas if we want the minimal transition period we can use the Elliptic one and if we want minimum peak ripple in the stop band we can use the Chebyshev approximation. The four types of analog filters used are shown in figure 3.

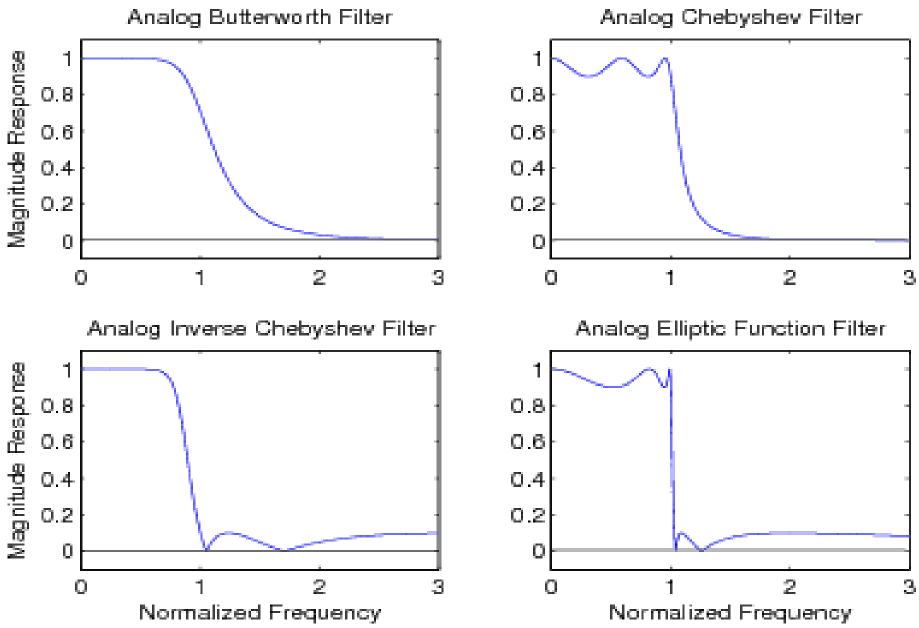

Figure 3

Now the next step is to choose the transformation. We will analyze two of them:

- Impulse Invariance Transform
- Bilinear Transform

Transformation is actually mapping from the s-plane to the z-plane. Any valid transformation must satisfy two requirements:

- The imaginary axis is mapped onto the unit circle.
- The left half of the s-plane should be mapped to inside of the unit circle.

3.1- Impulse Invariance Method:

The mapping in this transformation is governed by the following formula:

$$z=e^{sT}$$

Now we check whether it satisfies the two conditions:

Condition 1:

Here, s=jΩ as we are checking for imaginary axis mapping. Putting in the transformation formula we have,

$$z=e^{sT} \Rightarrow z=e^{j\Omega T} \Rightarrow z=e^{j\omega} \Rightarrow |z|=1$$

Hence imaginary axis is mapping onto unit circle.

Condition 2:

Here s=$\sigma + j\Omega$. Putting in transformation formula we have,

$$z=e^{sT} \Rightarrow z=e^{(\sigma+j\Omega)T} \Rightarrow z=e^{\sigma T}e^{j\Omega T} \Rightarrow |z|=e^{\sigma T}<1 \text{ for negative } \sigma.$$

As in left half of s-plane σ is negative, and the z-plane magnitude is less than 1, hence proved the satisfaction of 2nd condition.

So, this transformation can be used to convert the analog designed filter into digital one or for conversion of digital specifications into analog ones. Figure 4 describes this mapping.

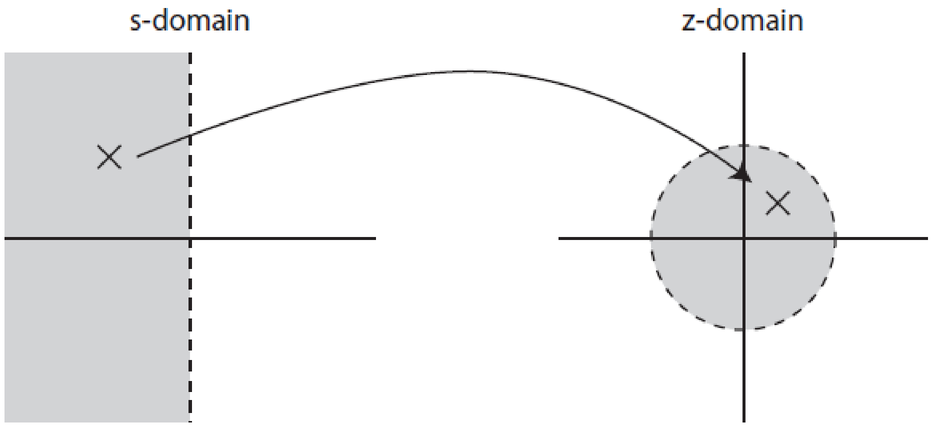

Figure 4

This method though simple also has its disadvantages. The main problem is that the whole inside of unit circle is mapped till $2\pi/T$ in s domain and if we want to map a point above $2\pi/T$, overlap occurs. Aliasing can occur in this method of transformation. To avoid these drawbacks we analyze another method which is commonly used i.e. Bilinear Transformation.

3.2-Bilinear Transformation:

The mapping in this transformation is governed by the formula:

$$z= \left(1+\frac{sT}{2}\right) \Big/ \left(1-\frac{sT}{2}\right)$$

Using the simplified version,

$$z= 1+s/1-s \quad \text{considering T=2}$$

Now we will check the whether it satisfies our mapping conditions or not:

Condition 1:

Using s=jω ,

$$z= 1+j\omega / 1-j\omega \Rightarrow |z|=1$$

Hence imaginary axis is mapped onto the unit circle.

Condition 2:

Using s= $\sigma + j\Omega$ in the transformation formula, we have

$$z= 1+\sigma+j\Omega / 1-\sigma-j\Omega \Rightarrow |z|= \sqrt{(1-\sigma)^2+\Omega^2} \Big/ \sqrt{(1+\sigma)^2+\Omega^2} \leq 1$$

Using negative σ, so signs changed. Moreover, the above term will always be smaller than the below one, so a point in left half s-plane will always be mapped inside the unit circle.

For this transformation, the relation between the analog and digital frequencies is given by

$$\Omega = \tan(\omega/_2)$$

In this case the positive imaginary axis is mapped to the upper half of the unit circle and negative axis is mapped to the lower half but this mapping is highly non-linear. The cut-off frequencies of a digital filter will therefore be tangentially warped compared with those of the analog filter from which it was designed. To eliminate this effect we need to pre-warp the frequencies before designing the analog filter. The above relation will be used for pre-warping.

After designing the analog filter we need to convert into the digital one using one of the above methods. Then, if we want a different filter than the low pass, we can convert it using the transformations developed for that purpose.

4- FIR Filter Design:

FIR Filters are also known as linear phase filters. It means that we get the same time delay in different frequency components which is very important in certain applications.
FIR Filters can be designed by two methods:

- Windowing
- Frequency Sampling

But we will analyze just one i.e. windowing in this report.

4.1- Windowing:

In this case as there is no analytical formula rather all the formulas are empirical ones so we will use the method of trying to approach the ideal/desired frequency response $H_d e^{j\omega}$ and then if N (the order) is not giving the required specs we increase the order and perform the steps again. Hence it is an iterative process. So our ideal response is

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega\tau} & when\ \ \omega \le \omega_p \\ 0 & otherwise \end{cases}$$

We will now use Fourier transform to expand it in frequency domain.

$$H_d(e^{j\omega}) = \sum_{n=-\infty}^{n=+\infty} h_d(n) e^{-jn\omega}$$

Since it is going to infinity on both sides, it will become an IIR Filter so we truncate it. As we want a realizable filter so we have to make it causal by changing the limits from 0 to N-1.
Now

$$H_d(e^{j\omega}) \cong H(e^{j\omega}) = \sum_{n=0}^{n=N-1} h_d(n) e^{-jn\omega} = \sum_{n=-\infty}^{n=+\infty} h_d(n) w(n) e^{-jn\omega}$$

Which means we have used a rectangular window with properties as:

$$w(n) = \begin{cases} 1 & 0 \le N-1 \\ 0 & otherwise \end{cases}$$

Here the N which is used are not analytically calculated rather it is an empirical guess so there can be a lot of iterations required to reach a final design satisfying our specs. Due to the use of such window this design is called windowing technique.
Rectangular window is one method but it has several disadvantage the most important of them is the Gibbs Phenomenon due to sharp discontinuity. Hence, to minimize this phenomenon, we can use a window with a gradual truncation. So there are a lot of window functions proposed each with its own set of advantages and disadvantages:

- Hamming Window
- Hanning Window
- Blackman Window

Figure 5 shows such windows:

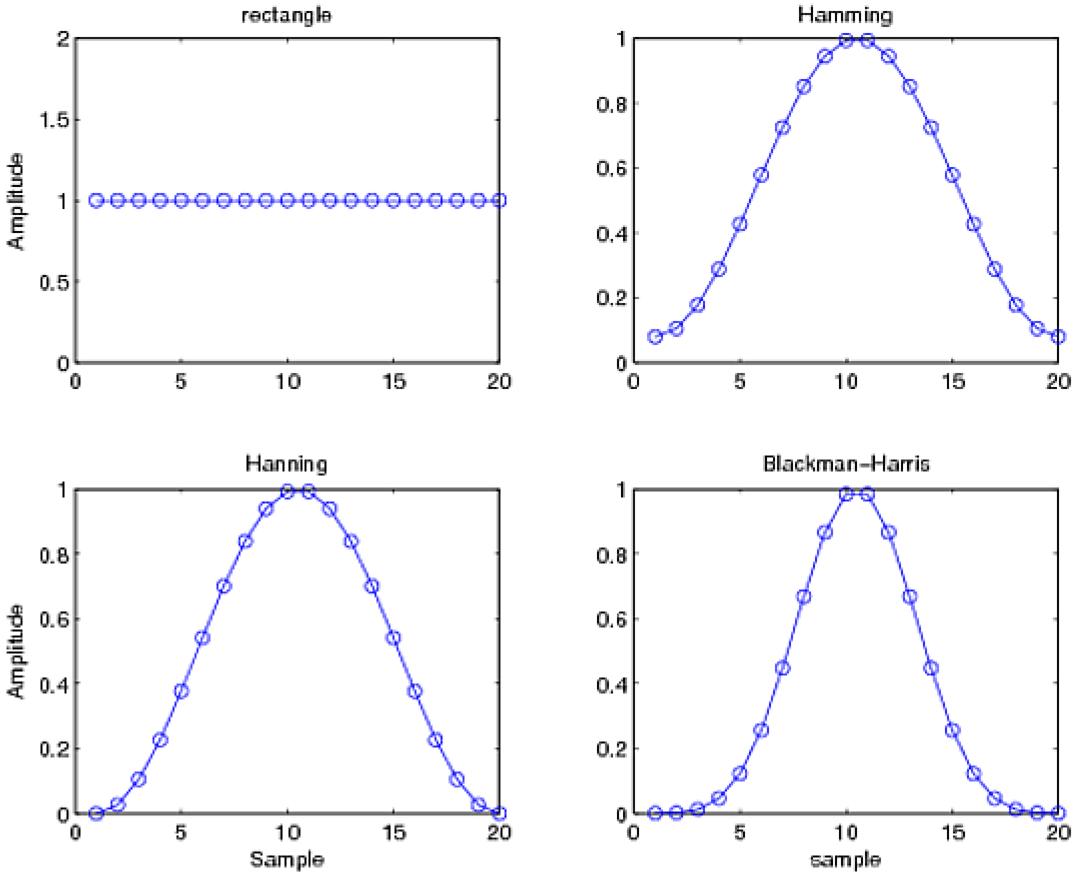

Figure 5

We want to approximate an impulse in the frequency domain so the main lobe width should be as small as possible and the side lobe height should also be as small as possible but this is a contradictory thing as when we try to lessen main lobe width the side lobe height increases hence it is a compromise.

Moreover, all the windows take care of symmetry as we have to design it for linear phase.

5- Conclusion:

Design of IIR and FIR filter methods are totally different. In IIR filter design we make use of the already advanced techniques for the design of analog filters whereas in FIR filters, instead of starting with specifications, we directly try to approximate order using some empirical formulas. Then, if the specs are not satisfied we can do the process again until we get the desired results.

References:

1- Digital Signal Processing Lecture Notes

2- Digital Signal Processing : A computer based approach by Sanjit K.Mitra

3- http://www.mikroe.com/chapters/view/73/chapter-3-iir-filters/

4- www.emba.uvm.edu/~mirchand/ee275ps/pdf_files/ch07(1).pdf

5- https://engineering.purdue.edu/...filter_design/1.7.3_iir

6- en.wikipedia.org/wiki/Digital_filter

7- www.robots.ox.ac.uk/~sjrob/Teaching/SP/l6.pdf

8- en.wikipedia.org/wiki/Bilinear_transform

9- https://ccrma.stanford.edu/~jos/fp/Frequency_Warping.html

10- http://en.wikipedia.org/wiki/Infinite_impulse_response

11- http://www.dspguru.com/dsp/faqs/fir/basics