

Lab7 Digital Filter Design(part1)

张旭东 12011923

1.Introduction

In this lab, some basic concepts and examples of FIR and IIR filters will be covered. The effect of zeros and poles will be illustrated and two systematic methods of designing both FIR and IIR filters will be introduced.

2.Background on Digital Filters

This part will discuss the theoretical analysis for the classification of digital filters.

In general, a linear and time-invariant causal digital filter with input $x[n]$ and output $y[n]$ can be shown as the following difference equation:

$$y[n] = \sum_{i=0}^{N-1} b_i x[n-i] - \sum_{k=1}^{M-1} a_k y[n-k] \quad (1)$$

This filter has $N - 1$ zeros and $M - 1$ poles, and the impulse response $h[n]$, which is a recursive difference equation, is:

$$h[n] = \sum_{i=0}^{N-1} b_i \delta[n-i] - \sum_{k=1}^{M-1} a_k h[n-k] \quad (2)$$

For finite impulse response(FIR) digital filter, $a_k = 0$ for all k . Such a filter has no poles, only zeros. The impulse response of it, $h[n]$, has finite duration N , whose expression is:

$$h[n] = \sum_{i=0}^{N-1} b_i \delta[n-i] \quad (3)$$

For infinite impulse response(IIR) digital filter where $a_k \neq 0$, the impulse response of it has non-zero values as $n \rightarrow \infty$.

For such a linear difference equation in (1), the z -transform of it is:

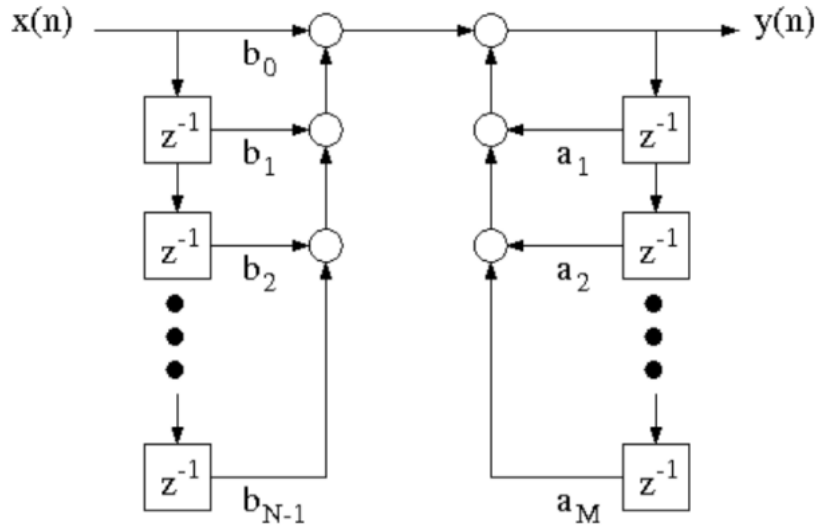
$$Y(z) = \sum_{i=0}^{N-1} b_i z^{-i} X(z) - \sum_{k=1}^{M-1} a_k z^{-k} Y(z) \quad (4)$$

The z -transform and frequency response of the digital filter are:

$$H(z) = \frac{\sum_{i=0}^{N-1} b_i z^{-i}}{1 + \sum_{k=1}^{M-1} a_k z^{-k}} \quad (5)$$

$$H(e^{j\omega}) = \frac{\sum_{i=0}^{N-1} b_i e^{-j\omega i}}{1 + \sum_{k=1}^{M-1} a_k e^{-j\omega k}} \quad (6)$$

The process for implementing a digital filter can be designed according to the difference equation above, which can be shown as the following system diagram.



3.Design of a Simple FIR Filter

In this part, the use of zeros in filter design will be discussed.

The filter in this part is a simple second order FIR filter with the two zeros on the unit circle. In order for the filter's impulse response to be real-valued, the two zeros must be complex conjugates of one another.

$$z_1 = e^{j\theta}, z_2 = e^{-j\theta} \quad (7)$$

The transfer function for this filter given by

$$H_f(z) = (1 - z_1 z^{-1})(1 - z_2 z^{-1}) = (1 - e^{j\theta} z^{-1})(1 - e^{-j\theta} z^{-1}) = 1 - 2 \cos \theta z^{-1} + z^{-2} \quad (8)$$

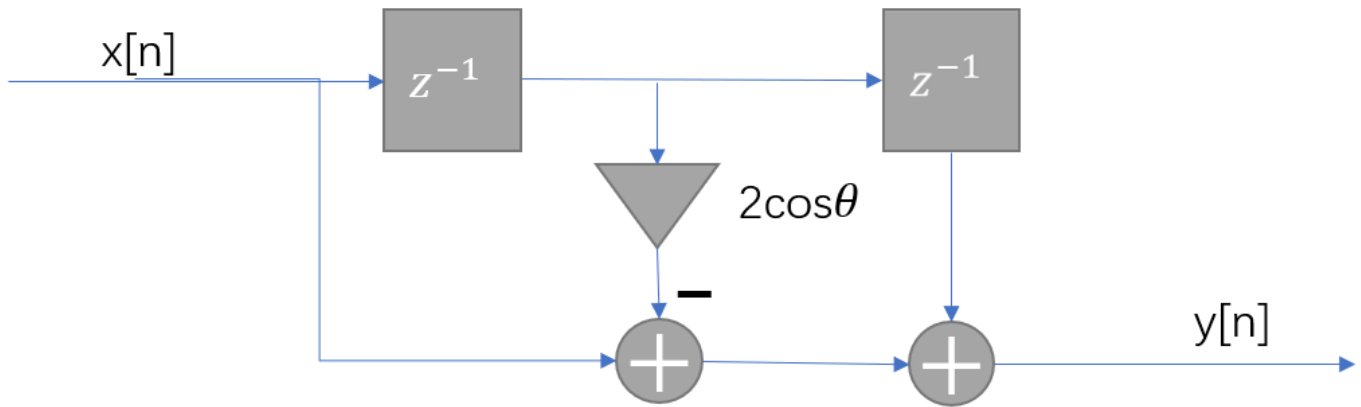
According to the transfer function, the difference equation can be derived as:

$$y[n] = x[n] - 2 \cos \theta x[n-1] + x[n-2] \quad (9)$$

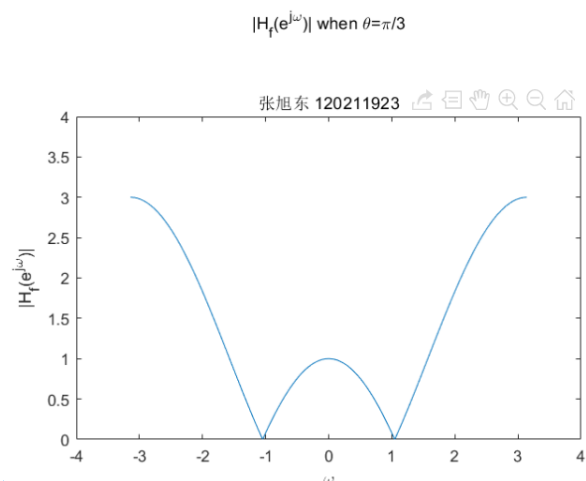
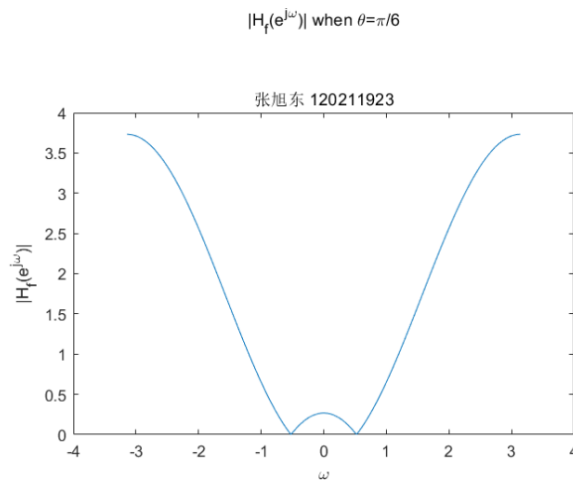
The impulse response:

$$h[n] = \delta[n] - 2 \cos \theta \delta[n-1] + \delta[n-2] \quad (10)$$

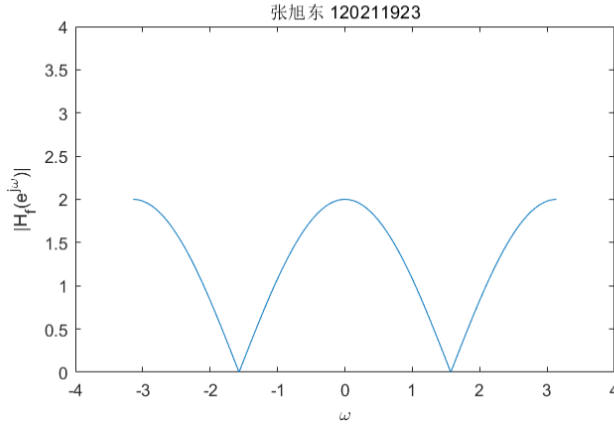
The system diagram:



According to the expression of impulse response of $h[n]$, it is obvious that $h[n]$ has finite duration. FIR filter with only zeros and no poles other than those at 0 and $\pm\infty$. The magnitude of the filter's frequency response $|H_f(e^{j\omega})|$ in three different θ values, $\theta = \frac{\pi}{6}$, $\theta = \frac{\pi}{3}$, $\theta = \frac{\pi}{2}$ are shown as following.



$|H_f(e^{j\omega})|$ when $\theta=\pi/2$



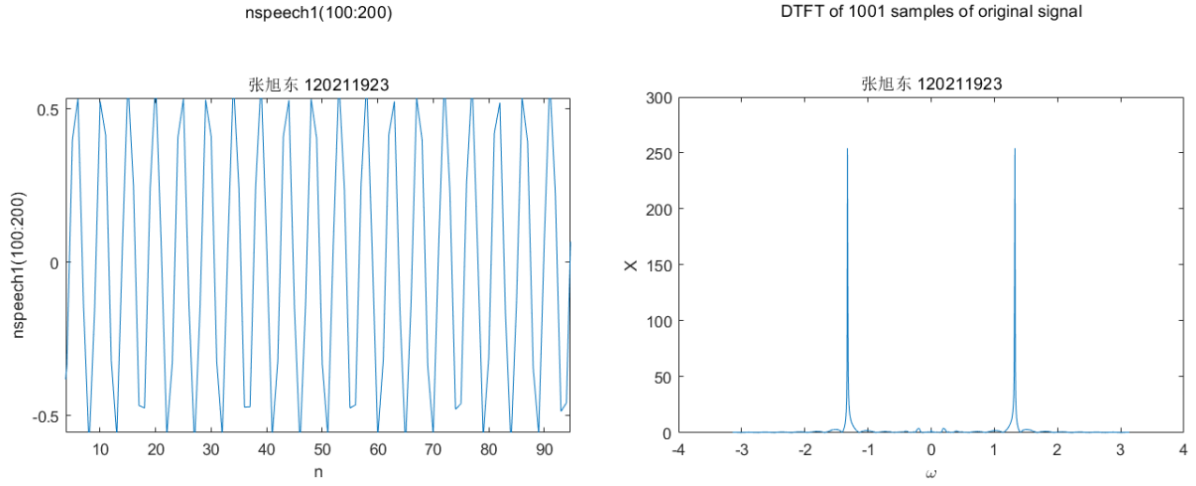
According to the above three pictures, the maximum value of magnitude of the filter's frequency response decrease with θ increasing in the interval $[0, \frac{\pi}{2}]$. The result can be derived from the following expression. The square magnitude of the filter's frequency response is

$$\begin{aligned} |H_f(e^{j\omega})|^2 &= H_f(z)H_f(z^{-1})|_{z=e^{j\omega}} \\ &= (1 - 2\cos\theta z^{-1} + z^{-2})(1 - 2\cos\theta z + z^2)|_{z=e^{j\omega}} \\ &= 2 + 4\cos^2\theta - 4\cos\theta(z^{-1} + z) + z^{-2} + z^2|_{z=e^{j\omega}} \\ &= 4(\cos\omega - \cos\theta)^2 \end{aligned}$$

According to the above expression, the number of points where the magnitude of the filter's frequency response is 0 is 2, $\omega = \theta$ and $\omega = -\theta$. What's more, it is obvious that the magnitude of the filter's frequency response $|H_f(e^{j\omega})|$ is an even function. When $\omega \in [0, \pi]$, the maximum value of the magnitude is taken at the farthest distance from θ . When $\theta = \frac{\pi}{6}$ and $\frac{\pi}{3}$, the maximum values are at $\omega = \pi$, and that is at $\omega = 0, \pi$ when $\theta = \frac{\pi}{2}$. The situation is similar when $\omega \in [-\pi, 0]$. So, the above figures are consistent with the analysis of theory.

Remove undesirable frequencies

In the next experiment, the filter $H_f(z)$ is used to remove an undesirable sinusoidal interference from a speech signal. The audio signal *nspeech1.nat* is loaded into MATLAB. Its time domain signal of indices from 100:200 and its DTFT samples of time indices from 100 to 1100 is shown below:



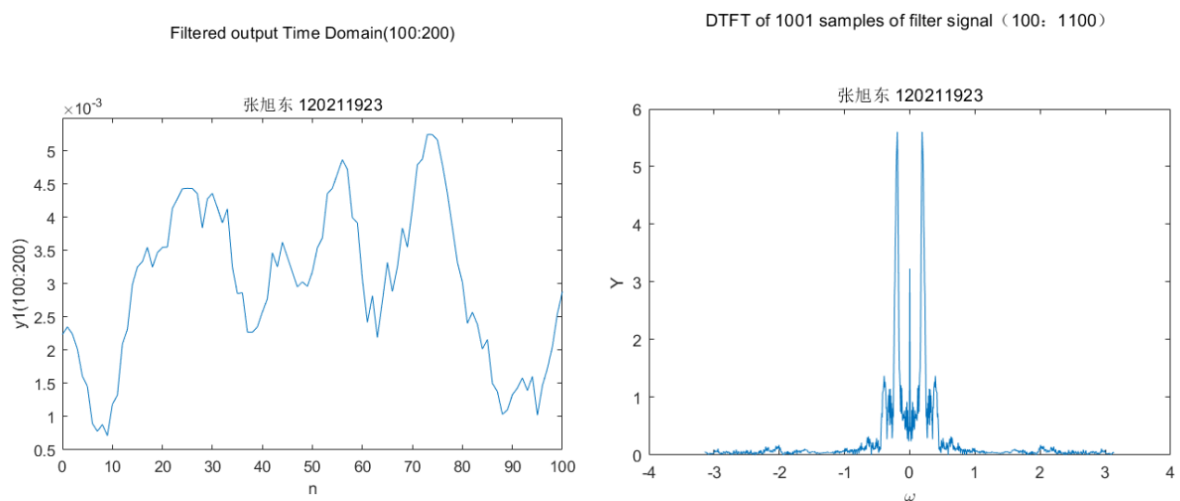
The exact frequency of peaks needs to be determined, which is the value of θ that will be used for filtering with $H_f(z)$. The command $[X_{max}, I_{max}] = \max(abs(X))$ is used to find the value and index of the maximum element in X . θ can be derived using this index.

$$\theta = -\pi + (I_{max} - 1) \frac{2\pi}{M} \quad (11)$$

M specifies the number of output points of the DTFT. From the command, the corresponding results is

```
Xmax = 254.2287
Imax = 297
-1.3254
```

The sound file is filtered to attenuate the sinusoidal noise and its time domain signal of indices from 100:200 and its DTFT samples of time indices from 100 to 1100 is shown below:



It is obvious that the peaks in the frequency domain at $\omega = -1.3254$ and $\omega = 1.3254$ have been removed almost. And the frequency content becomes more centralized in the frequency domain after filtering. The filter is a band-stop filter from the three filter above and it stop the frequency at $\omega = -1.3254$ and $\omega = 1.3254$. By comparing the auditory perception of the signal before filtering

and that of the signal after filtering, it is obvious that the harsh part of the background sound in the filtered signal is obviously removed and the auditory effect is better.

4.Design of a Simple IIR Filter

While zeros attenuate a filtered signal, poles amplify signals that are near their frequency. In this part, how poles can be used to separate a narrow-band signal from adjacent noise will be illustrate. Such filter are commonly used to separate modulated signals from background noise in applications such as radio-frequency demodulation.

The transfer function for a second order IIR filter with complex-conjugate poles is :

$$H_i(z) = \frac{1 - r}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})} = \frac{1 - r}{1 - 2r \cos \theta z^{-1} + r^2 z^{-2}} \quad (12)$$

It has poles:

$$p_1 = re^{j\theta}, p_2 = re^{-j\theta} \quad (13)$$

To make the filter stable, the absolute value of r is smaller than 1.

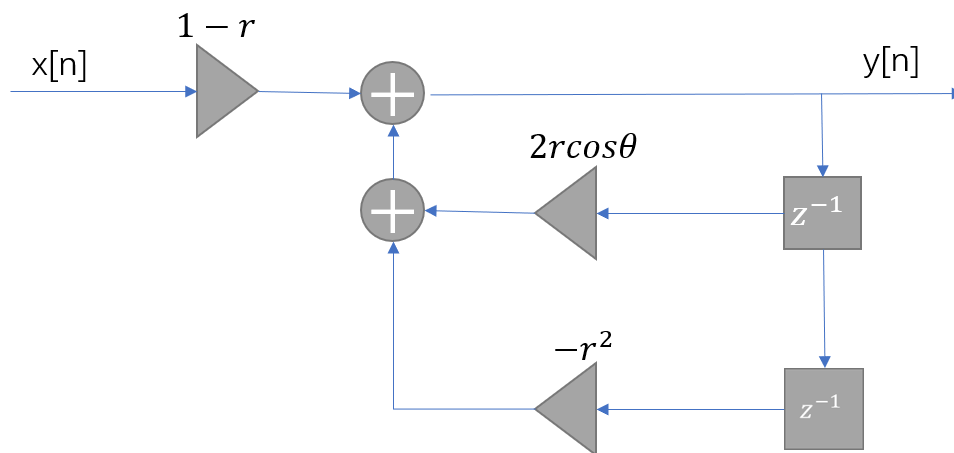
According to the transfer function, the difference equation can be derived as:

$$y[n] = (1 - r)x[n] + 2r \cos \theta y[n - 1] - r^2 y[n - 2] \quad (14)$$

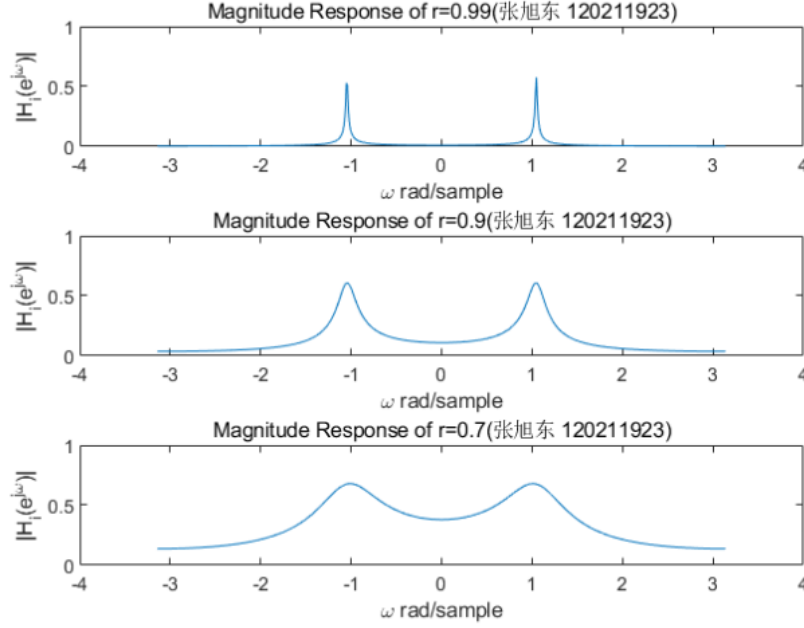
The impulse response:

$$h[n] = (1 - r)\delta[n] + 2r \cos \theta h[n - 1] - r^2 h[n - 2] \quad (15)$$

The system diagram:



According to the expression of impulse response of $h[n]$, it is obvious that $h[n]$ has infinite duration. The magnitude of the filter's frequency response $|H_i(e^{j\omega})|$ for $\theta = \frac{\pi}{3}$ in three different r values, $r = 0.99, r = 0.9, r = 0.7$ are shown as following.



It is obvious that with r becoming smaller, the maximum value magnitude response of the filter increases and the value at $\omega = 0$ increases. What's more, the absolute value of ω where the maximum value are at increases with r becoming smaller. The result can be derived from the following expression. The square magnitude of the filter's frequency response is

$$\begin{aligned}
 |H_f(e^{j\omega})|^2 &= H_f(z)H_f(z^{-1})|_{z=e^{j\omega}} \\
 &= \frac{1-r}{1-2r\cos\theta z^{-1}+r^2 z^{-2}} \frac{1-r}{1-2r\cos\theta z+r^2 z^2} \Big|_{z=e^{j\omega}} \\
 &= \frac{(1-r)^2}{1+4r^2\cos^2\theta+r^4-2r\cos\theta(z^{-1}+z)+r^2(z^{-2}+z^2)-2r^3\cos\theta(z^{-1}+\cos z)} \Big|_{z=e^{j\omega}} \\
 &= \frac{(1-r)^2}{4r^2(\cos\omega - \frac{1+r^2}{4r})^2 - \frac{(1+r^2)^2}{4} + (r^2-1)^2} \\
 &= \frac{1}{\frac{4r^2}{(1-r)^2}(\cos\omega - \frac{1+r^2}{4r})^2 + (r+1)^2 - \frac{(1+r^2)^2}{4(1-r)^2}}
 \end{aligned}$$

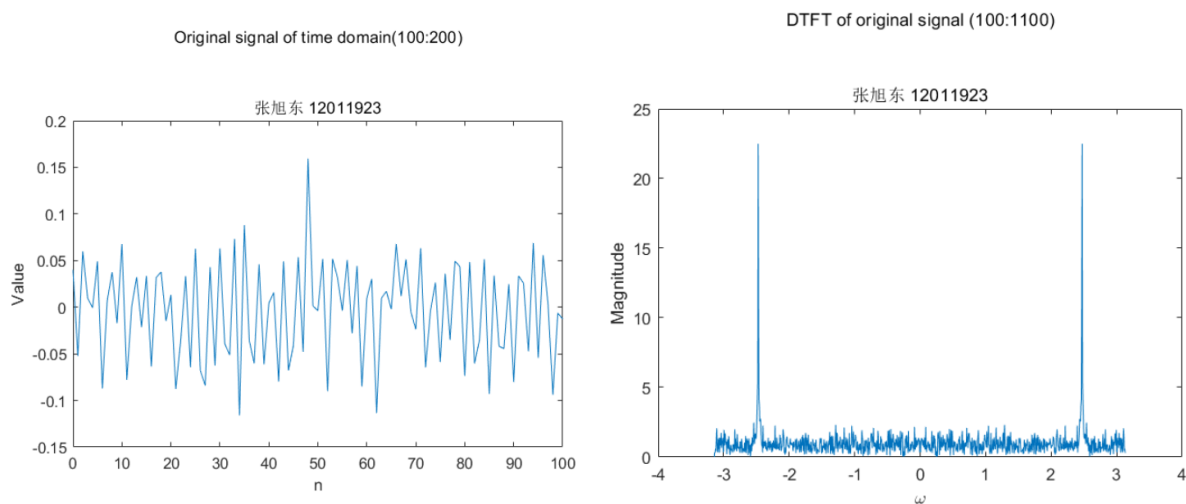
Let $h(r) = \frac{1+r^2}{4r}$ and $f(r) = (r+1)^2 - \frac{(1+r^2)^2}{4(1-r)^2}$. Their derivatives are:

$$\begin{aligned}
 h'(r) &= \left(\frac{1+r^2}{4r}\right)' = \frac{r^2-1}{4r^2} \\
 f'(r) &= \left[(r+1)^2 - \frac{(1+r^2)^2}{4(1-r)^2}\right]' = 2(r+1) + \frac{8r^5 - 24r^4 + 34r^3 - 48r^2 + 40r - 24}{16(1-r)^4} \quad (16)
 \end{aligned}$$

Because $|r| < 1$, $h'(r) < 0$. So, when r decreases in the interval $[0, 1)$, $h(r)$ increases. When $\omega \in [0, \pi]$, the maximum value of the magnitude is taken at $\omega_{max} = \cos^{-1}(\frac{1+r^2}{4r})$. When r decreases in the interval $[0, 1)$, $\frac{1+r^2}{4r}$ increases, ω_{max} decreases. When r is in the interval $[0, 1)$, $f'(r) > 0$. So when r decreases in the interval $[0, 1)$, $f(r)$ decreases. When $\omega \in [0, \pi]$, the maximum value of the magnitude increases with r decreasing in the interval $[0, 1)$. The situation is similar when $\omega \in [-\pi, 0]$. So, the above figures are consistent with the analysis of theory. What's more, with r becoming larger, the filter's bandwidth is made extremely narrow around θ .

Separate a modulated sinusoid from background noise

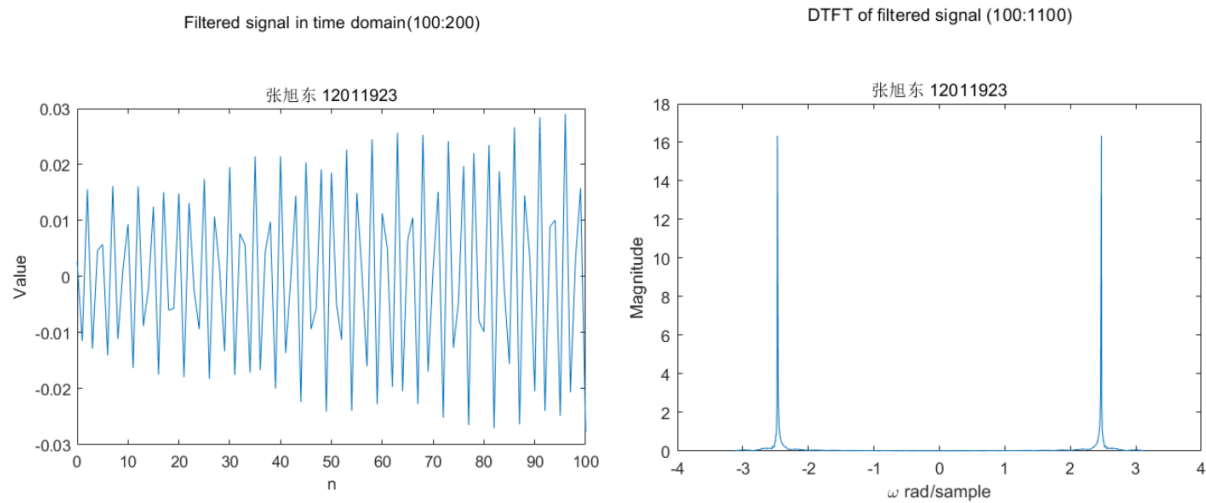
In the next experiment, the filter $H_i(z)$ is used to Separate a modulated sinusoid from background noise. The audio signal *pcm.nat* is loaded into MATLAB. Its time domain signal of indices from 100:200 and its DTFT samples of time indices from 100 to 1100 is shown below:



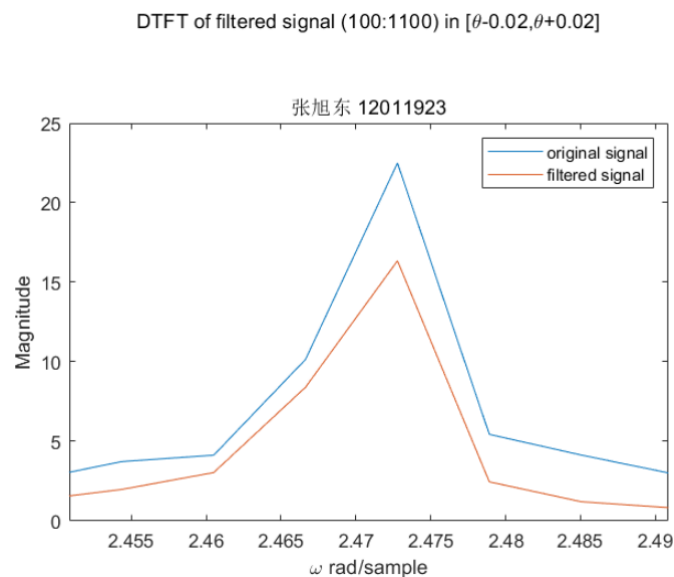
The two peaks in the spectrum correspond to the center frequency of the modulated signal while the low amplitude wideband content is the background noise. The *pcm* signal is modulated at 3146Hz and sampled at 8kHz, so $\theta = 2\pi \frac{3146}{8k} = 2.47086$. The function *IIRfilter* that implement the filter with $r = 0.995$ and $\theta = 2.47086$ is shown as below. In the code, the recursive difference equation is implemented and $y[n]$ is assumed to be equal to 0 for negative values of n .

```
function y=IIRfilter(x,theta,r)
    y(1)=(1-r)*x(1);
    y(2)=(1-r)*x(2)+2*r*cos(theta)*y(1);
    for i=3:length(x)
        y(i)=(1-r)*x(i)+2*r*cos(theta)*y(i-1)-r*r*y(i-2);
    end
end
```

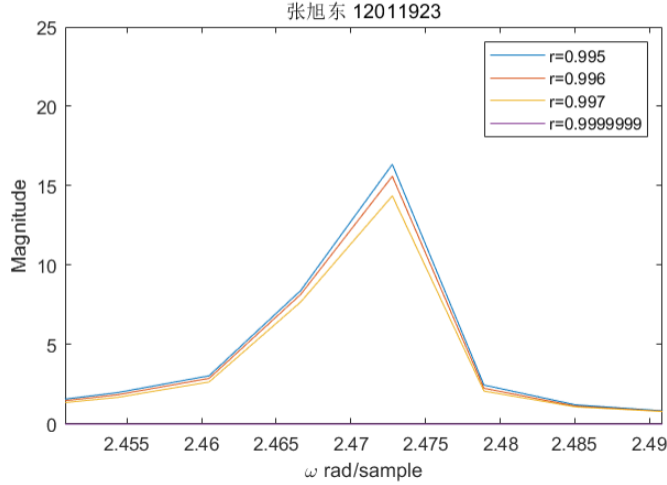

The filtered signal's time domain signal of indices from 100:200 and its DTFT samples of time indices from 100 to 1100 is shown below:



From the two pictures above, the background noise has been removed and the sound is more clear. Before filtering, there are many small fluctuations which represents the noise. After filtering, these fluctuation are removed. In order to see the DTFT around $\omega = \theta$ more clearly, the portion of this DTFT for the values of ω in the range $[\theta - 0.02, \theta + 0.02]$ is shown as below:



From the picture, what can be found is that the magnitude of the signal in frequency domain decrease through the filter, which can be gotten from the above three filter. When r becoming larger from 0.95, the peak value become smaller and smaller. When $r = 0.9999999$, the peak value is 0. So, $r = 0.9999999$ is not effective for this application.



5. Filter Design Using Truncation

Ideally, a lowpass filter with cutoff frequency ω_c has a frequency response of

$$h_{ideal}(e^{j\omega}) = \begin{cases} 1 & |\omega| \leq \omega_c \\ 0 & \omega_c < |\omega| \leq \pi \end{cases} \quad (17)$$

and a corresponding impulse response of

$$h_{ideal}[n] = \frac{\omega_c}{\pi} \text{sinc}\left(\frac{\omega_c n}{\pi}\right), -\infty < n < \infty \quad (18)$$

However, no real filter can have this frequency response because h_{ideal} is infinite in duration. One method for creating a realizable approximation to an ideal filter is to truncate this impulse response outside of $n \in [-M, M]$.

$$h_{trunc}[n] = \begin{cases} \frac{\omega_c}{\pi} \text{sinc}\left(\frac{\omega_c n}{\pi}\right) & n = -M, \dots, 0, 1, \dots, M \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

Due to the modulation property of the DTFT, the frequency response of the truncated filtered is the result of convolving the magnitude response of the ideal filter with the DTFT of the truncating window. A truncated impulse response is of finite duration, yet the filter is still non causal. In order to make the FIR filter causal, it must be shifted to the right by M units. For a filter of size $N = 2M + 1$, this shifted and truncated filter is given by

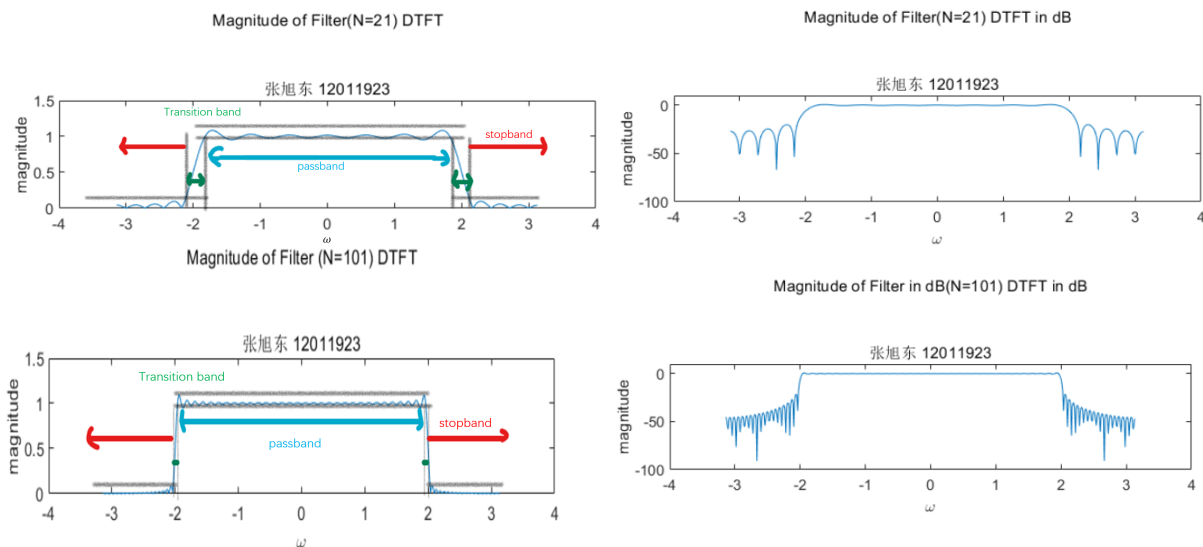
$$h[n] = \begin{cases} \frac{\omega_c}{\pi} \text{sinc}\left(\frac{\omega_c}{\pi}\left(n - \frac{N-1}{2}\right)\right) & n = 0, 1, \dots, N-1 \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

This time shift of $\frac{N-1}{2}$ units to the right corresponds to multiplying the frequency response by $e^{-j\omega\frac{N-1}{2}}$, which not affects the magnitude response of the filter but adds a factor of $-j\omega\frac{N-1}{2}$ to the phase response. Such a filter is called linear phase because the phase is a linear function of ω .

To examine the effect of filter size on the frequency characteristics of the filter, the function $LPFtrunc(N)$ computes the truncated and shifted impulse response of size N for a lowpass filter with a cutoff frequency of $\omega_c = 2.0$.

```
function h=LPFtrunc(N)
    wc=2.0;
    n=0:N-1;
    h=wc/pi*sinc(wc/pi*(n-(N-1)/2));
end
```

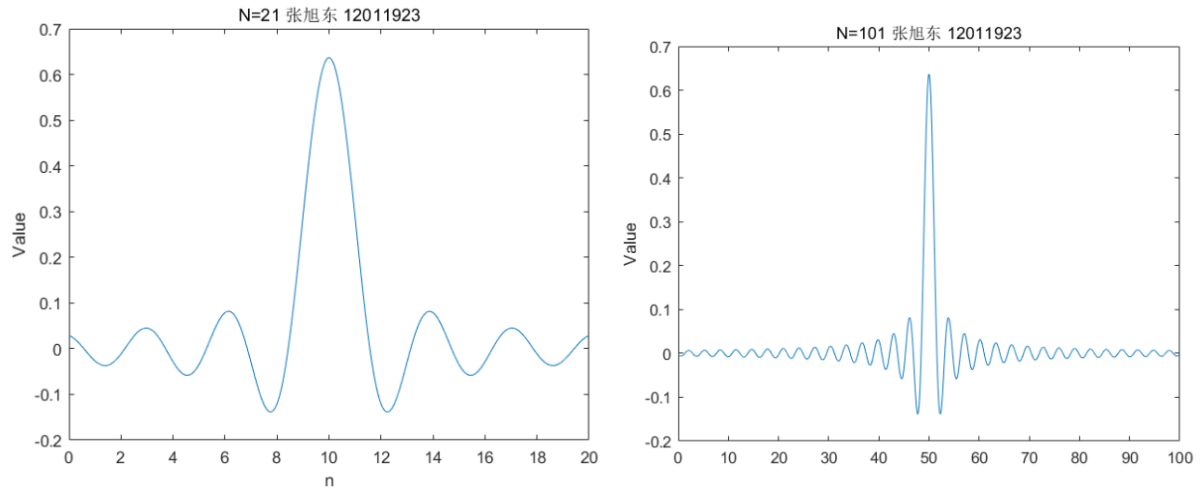
For $N = 21$ and $N = 101$, the magnitude response of the filter in original scale and in decibels are



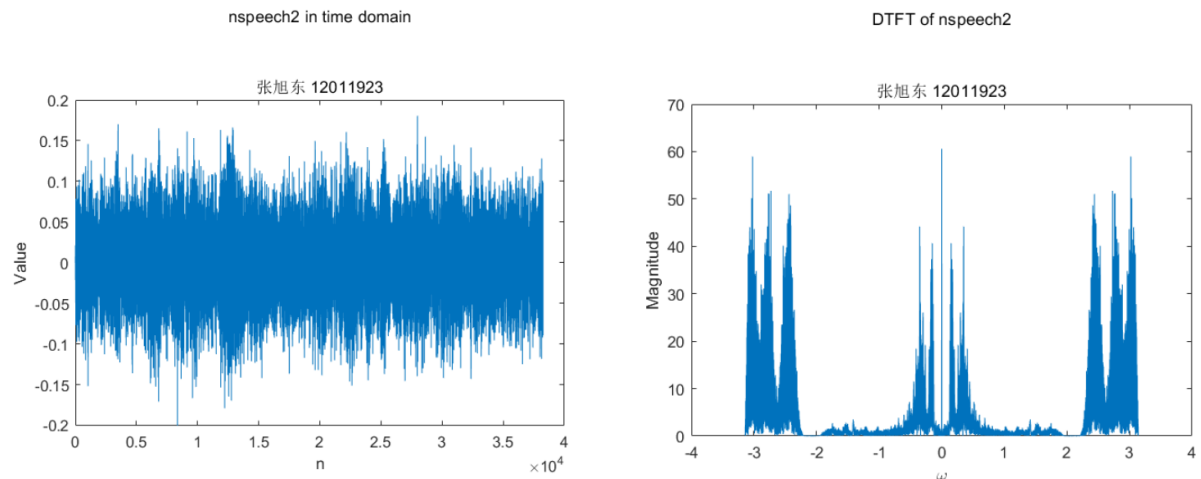
From the above pictures, it is obvious that with N becoming larger, the frequency response of the filter is more ideal, and the bandwidth of passband is longer and that of transition band is shorter, which means the performance is better. What's more, with N becoming larger, the magnitude of fluctuation in passband become smaller and smaller. However, the fluctuation on the border of passband and transition band isn't decreasing, and it will never disappear because of the finite duration of the truncated filter, which is called Gibbs phenomenon.

For the stopband ripple, the number of it becomes more and more with N becoming larger while the magnitude of fluctuation become smaller and smaller. That is because the sinc function in the time domain will have shaper main lobes with N increasing. the frequency response of the truncated filtered is the result of convolving the magnitude response of the ideal filter with the

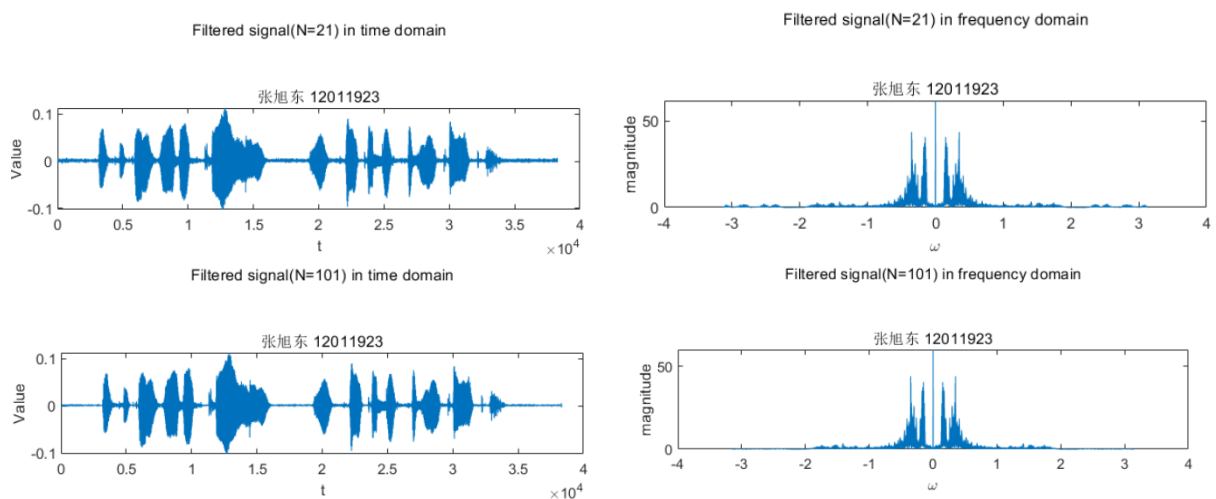
DTFT of the truncating window. So, more ripples will be generated with N increasing.



The audio signal *nspeech2.nat* is loaded into MATLAB. Its time domain signal and its DTFT samples of time indices from 100 to 1100 is shown below:



The high frequency signal is noise and it is band limited to $|\omega| > 2.2$ while the low frequency signal is speech and it is band limited to $|\omega| < 1.8$. The results after applying the two filters to this signal is shown as below:



Listening to the two result, it is obvious that the background noise in the signal filtered by the truncated filter whose size is $N = 101$ is less than that in the signal filtered by the truncated filter whose size is $N = 21$, which means the performance gets better and better with N increasing. For the frequency domain, the high frequency is removed basically. However, the rest of high frequency in the signal filtered by the truncated filter whose size is $N = 101$ is less than that in the signal filtered by the truncated filter whose size is $N = 21$, which is consistent with the hearing.

6.Conclusion

In this lab, some basic concepts of the digital filters have learned. Filters are mainly divided into two classifications: finite impulse response(FIR) and infinite impulse response(IIR). What's more, zeros attenuate a filtered signal while poles amplify signals that are near their frequency. Besides, how to design a filter directly from the transfer function and by implementing truncated filter by convolving the magnitude response of the ideal filter with the DTFT of the truncating window is more understood. The first one is easy to realize but it is hard to get the specifications. The second one is easy to set the specifications but how to choose the proper size of truncated filter need to be considered more.