Digital Signal Processing

Project Report

Aim:

To extract group delay function from various methods, and test its robustness to noise .

Introduction:

- The information in speech signals is represented in terms of features derived from short-time Fourier analysis, the features extracted from the magnitude of the Fourier transform (FT) are considered, ignoring the phase component.
- It was not exploited fully due to difficulty in computing the phase and also in processing the phase function. The information in the short-time FT phase function can be extracted by processing the derivative of the FT phase, i.e., the group delay function.
- This project will review properties of the group delay functions are reviewed, highlighting the importance of the FT phase for representing information in the speech signal.
- It capture the characteristics of the vocal-tract system in the form of formants or through a modified group delay function .
- Also, it's robustness to noise channel will be analysed.

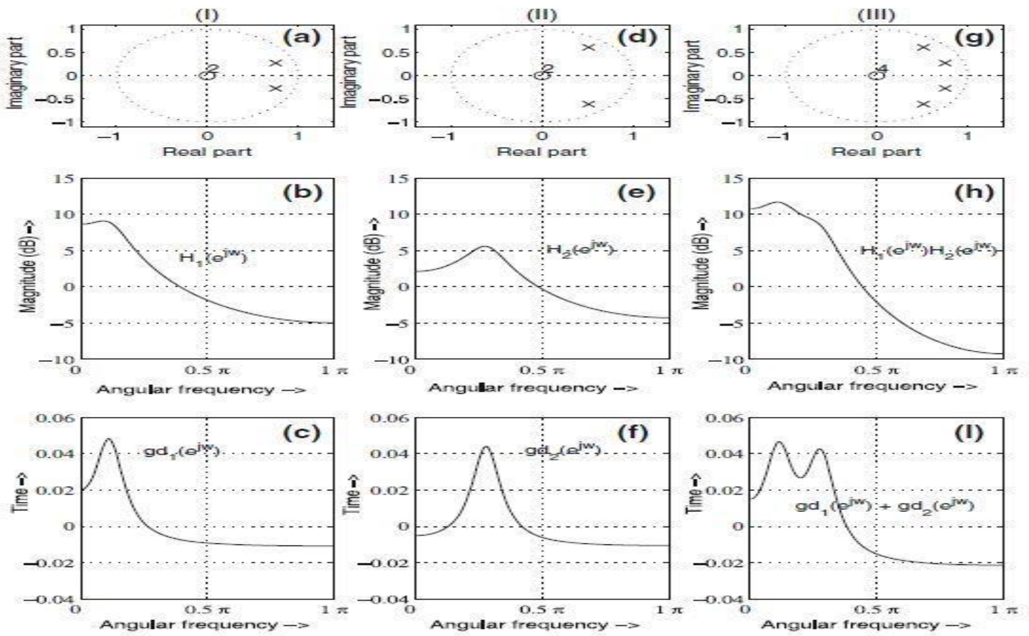
Properties of Group Delay：

We assume that H1 & H2 are minimum-phase signals, with zeros & poles inside the unit circle.

We can see that ,in Magnitude spectrum the envelope is not clearly seen as they are getting multiplied ,where as in Phase spectrum (GD),they are added.

- Note:

  1.This is useful because, we can then consider the speech as cascade of M resonators, and perform analysis, and the group delay will be additive in nature .



2. This nature,is observed only in Minimum phase signal,for Min/Mixed phase its different mechanism .Also,the derived group delay from magnitude and phase spectrum are different , in case of Min/Max signal.

(iv) For minimum phase signals

$$\tau_p(e^{j\omega}) = \tau_m(e^{j\omega}).$$

(v) For maximum phase signals

$$\tau_p(e^{j\omega}) = -\tau_m(e^{j\omega}).$$

(vi) For mixed phase signals

$$\tau_p(e^{j\omega}) \neq \tau_m(e^{j\omega}).$$

Methods for Group Delay Extraction :

1.Linear Prediction Analysis :

Introduction : Speech is a convolution of both Glottal voice box (H(z)) as well as its surrounding environment (E(z)).

So, the corresponding transfer function will be:

- All pole filter equation of H(z) -Speech.

$$H(z) = 1/(1 + \sum_{k=1}^{p} a_k z^{-k}),$$

where ak's are Linear Prediction Coefficients .

Algorithm :

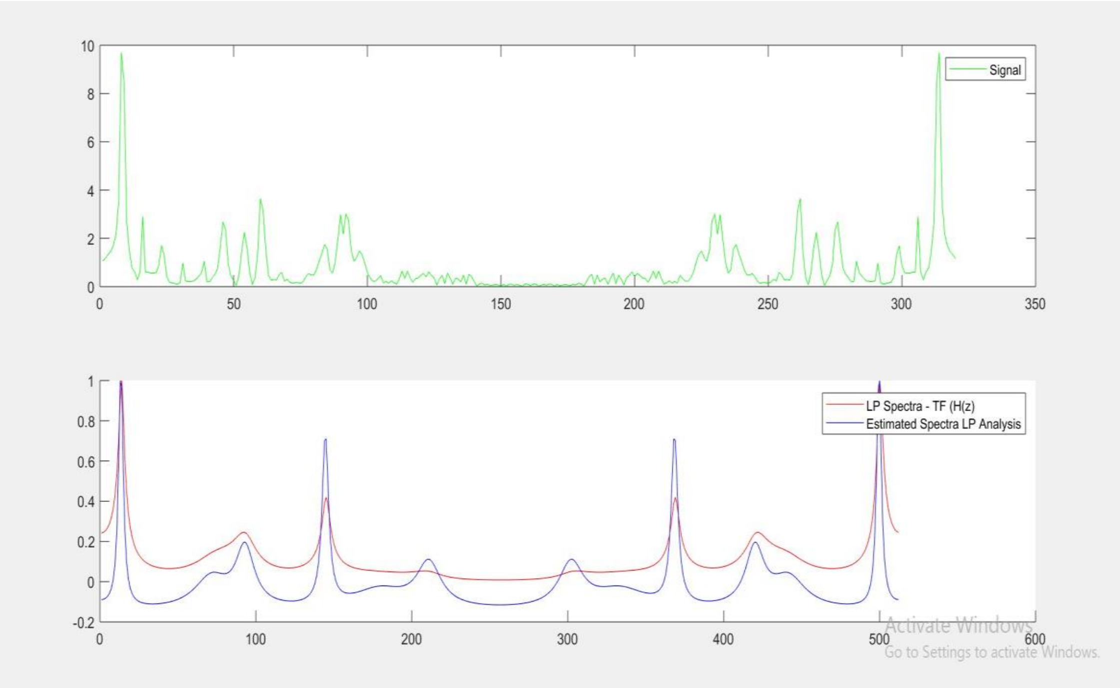
- Denoise the speech signal, and split into frames for analysis.
- Obtain LPC coefficients, N=12 or 14 in number - {a0,a1...ap}
- A 512-point DFT to the sequence &appended with appropriate number of zeros is computed using FFT pruning algorithm.
- The phases spectrum is obtained from the complex DFT-coefficients using the standard arctan function subroutine.
- The derivative of the phase spectrum is then obtained by computing the differences between successive values.
- Plot the derivative signal, which is the group delay and compare it with the Fourier Transform of the speech signal ,to identify the peaks .

Results:

Filename : lpcanalysis.m

Speech Filename : S1.wav

Note: S1.wav is a vowel speech sample, whose time scale is in milliseconds.

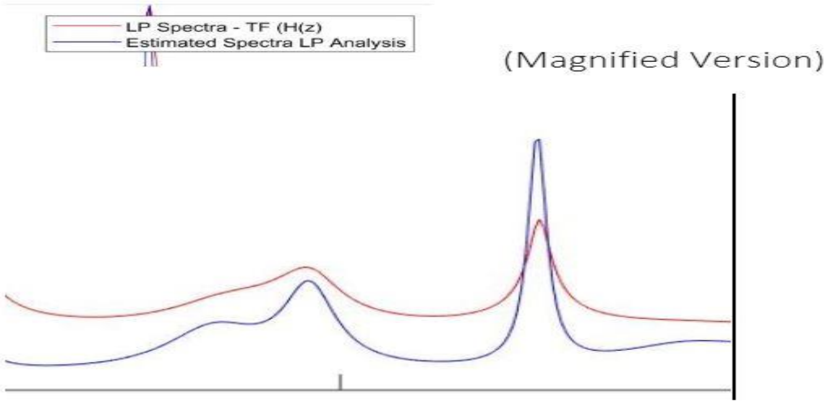


Remarks :

The first plot corresponds to FFT of the speech signal ,the second plot there is

LP Spectra plot (only the H(z) functions) & estimated LP Analysis, and the order of linear prediction coefficient was set to 12.

The following part of the graph, describes how LP Spectra does not highlight's the overlapped formats whereas LP analysis , clearly amplifies and differentiates between two formats.



2.Direct Group Delay

So, the group delay function is derived both from x(n) and n*x(n) .Here,

X =fft(x(n)) (Speech Signal) , and Y=fft(n*x(n)).

$$\begin{aligned}\tau_x(e^{j\omega}) &= -\{\mathrm{Im}\}[\frac{d(\log(X(e^{j\omega})))}{d\omega}] \\ &= \frac{X_R(e^{j\omega})Y_R(e^{j\omega}) + Y_I(e^{j\omega})X_I(e^{j\omega})}{|X(e^{j\omega})|^2},\end{aligned}$$

Algorithm :

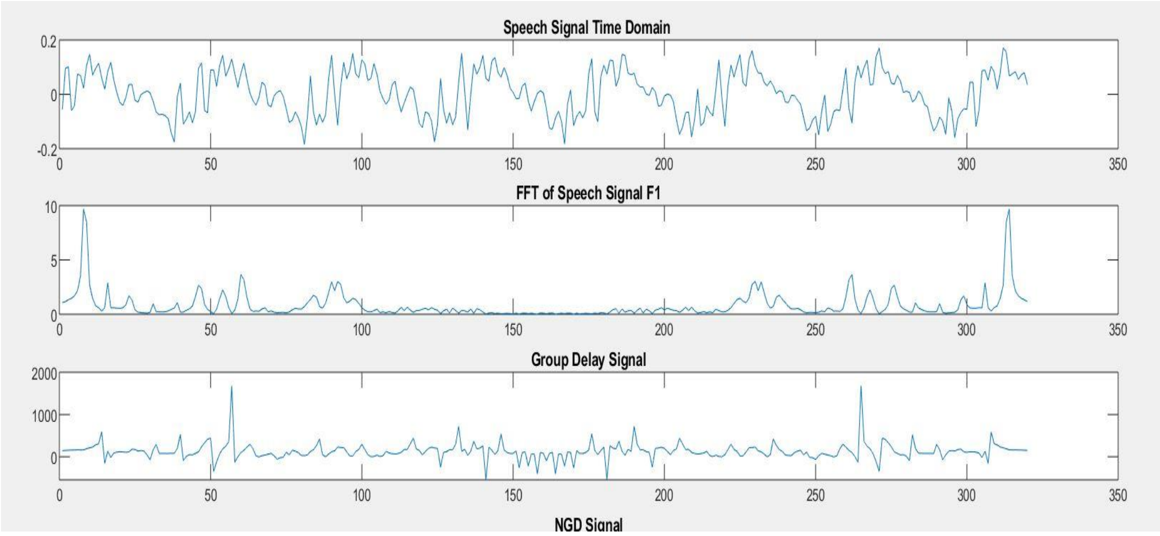
1.Compute n*x(n),for the particular frame.

2.Compute the FFTs of the signal .

3.Extraction of real and imaginary part of X & Y ,& the magnitude square of X(exp(jw)).

4.Substitute in the above mentioned formula .

Results: There are abrupt changes, group delay it is due the zeros lying in the unit circle, which will make the Group Delay function go to infinity or large value.



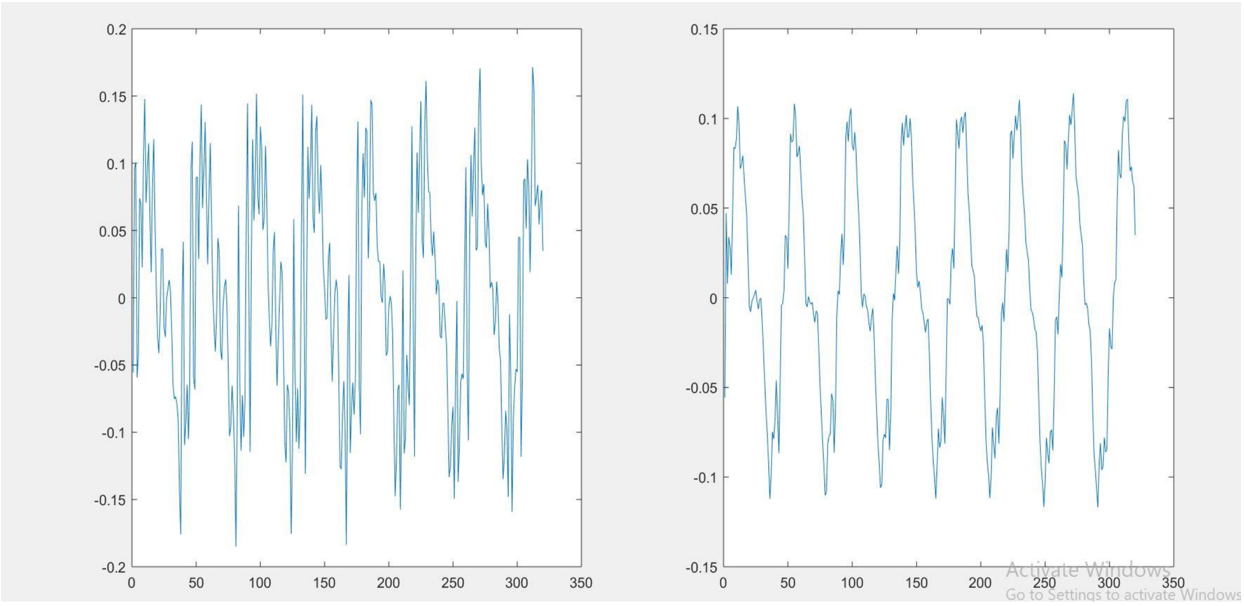
3.Modified Group Delay - Cepstral Smoothing .

The explanation for the peaky nature of above group delay ,is due to the zeros lying very close to the origin ,inside the unit circle. So, to overcome that we replace magnitude spectra with cepstrally smooth spectra (Sc(exp(jw)). This averages out the function , and will not have function values as zero.

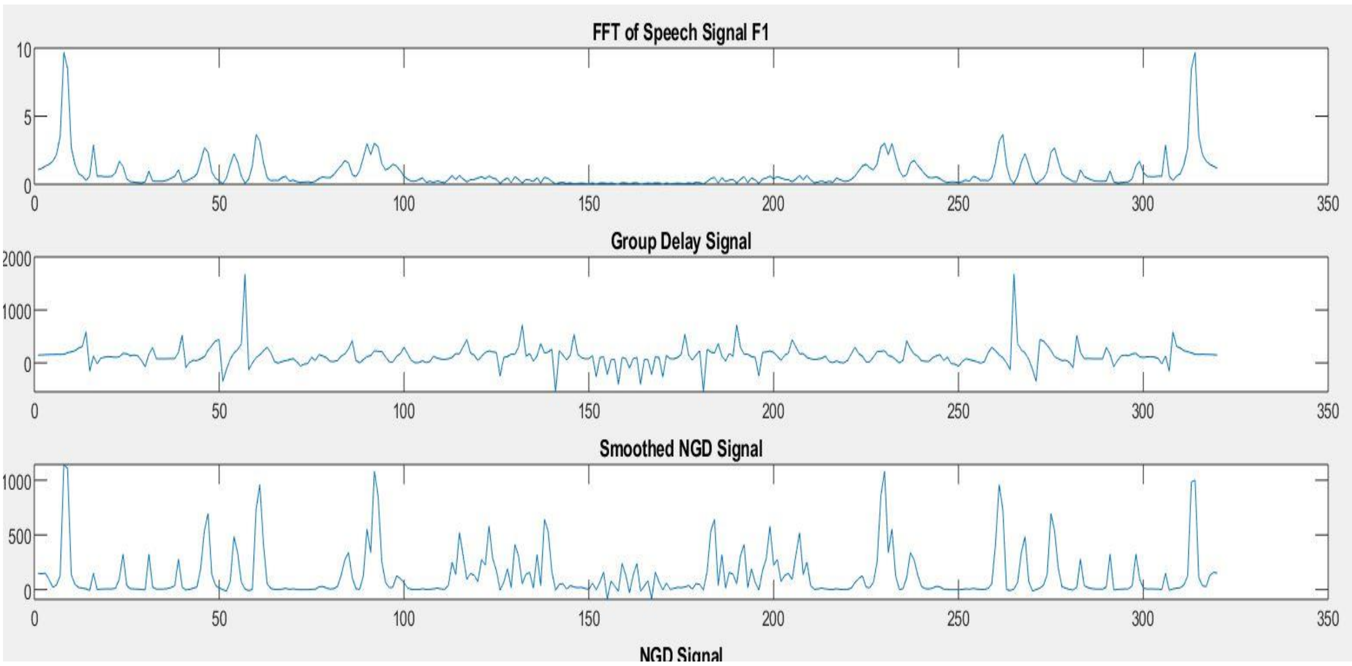
Note: The top plot is of the signal, and the bottom part is smoothed, we can see by the amplitude scale that values aren't low in latter

$$\tau_c(e^{j\omega}) = \frac{X_R(e^{j\omega})Y_R(e^{j\omega}) + Y_I(e^{j\omega})X_I(e^{j\omega})}{S_c^2(e^{j\omega})},$$

graph .



Results :



Remarks :

Comparing the first and third plot ,they are quite similar in nature .Also, the graph is less peaky in nature due to log based smoothening . But, there are few false peaks, and troughs so we create a more accurate function by introducing certain approximations and parameters.

4.Modified Group Delay - Numerator Group Delay .

Here, the denominator of the Group delay function that is square mag(X(e^jw)),is completely removed so that peaky nature is lost and the whole function is only changed by magnitude .

Algorithm

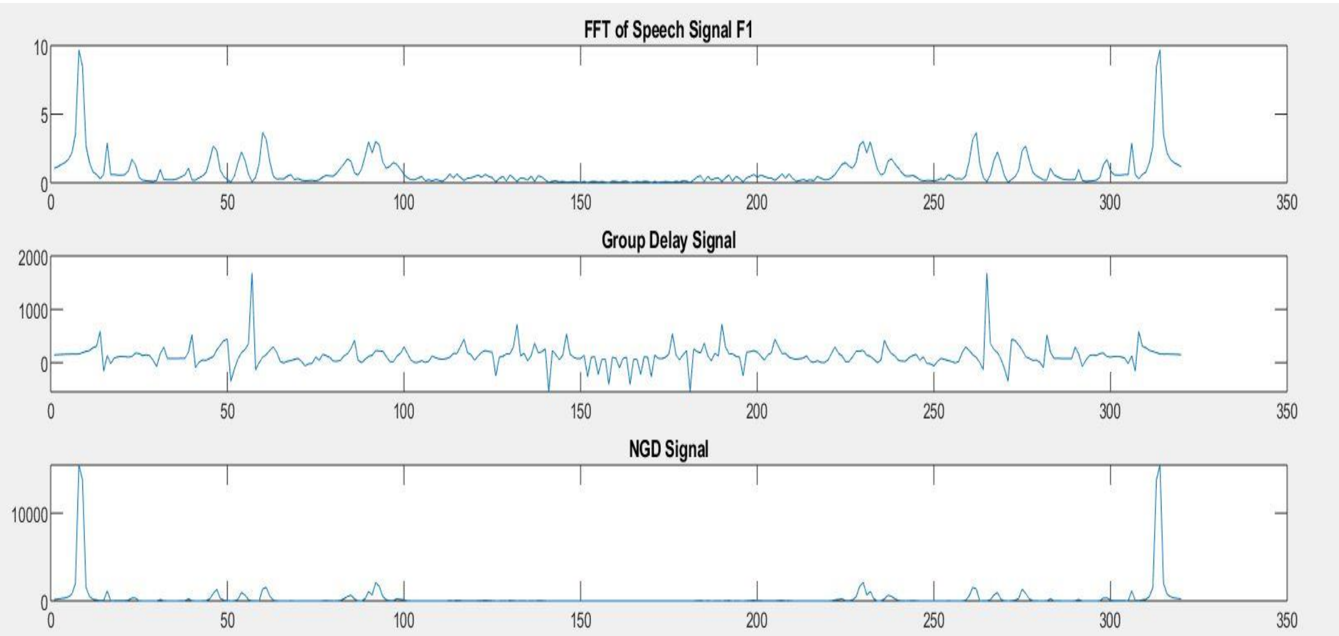
1.Compute n*x(n),for the particular frame.

2.Compute the FFTs of the signal .

3.Extraction of real and imaginary part of X & Y ,& the magnitude square of X(exp(jw)).

4.Multiply the corresponding real and imaginary parts and sum

Results:



Remarks :

The formants in 3rd plot and 1sd plot ,are positioned correctly but the magnitude of the peaks isn't high enough to display all . Only the higher peaks are significant.

5.Modified Group Delay - Numerator Group Delay with Gamma.

The NGD function is a good approximation , but it needs to enhance the low magnitude peaks too .

Algorithm

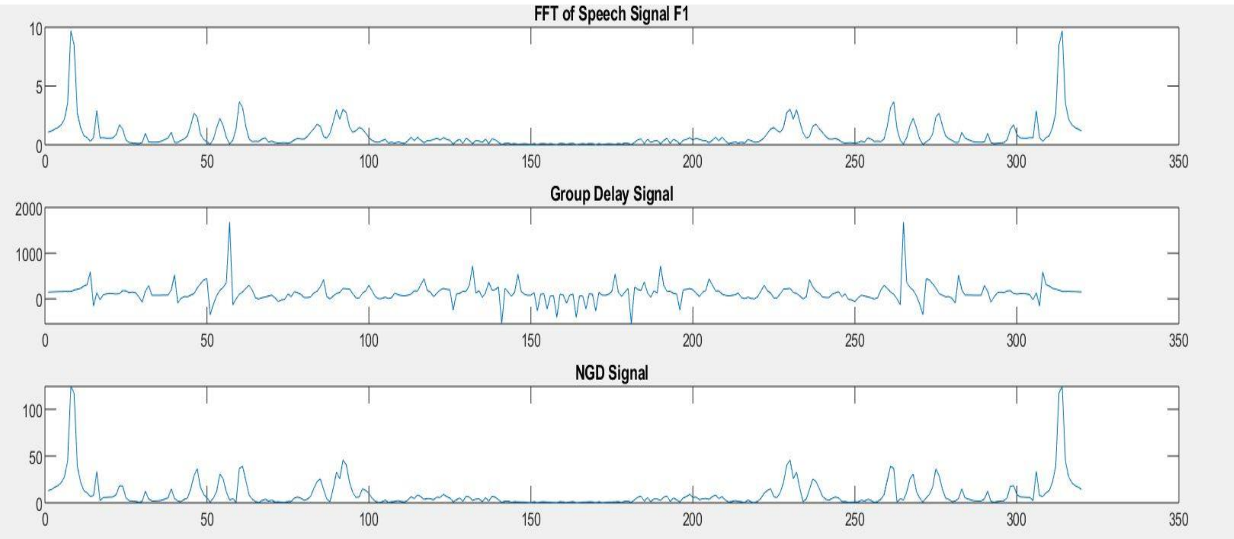
1.Compute n*x(n),for the particular frame.

2.Compute the FFTs of the signal & extraction of real and imaginary part of X & Y ,& the magnitude square of X(exp(jw)).

4.Multiply the corresponding real and imaginary parts and sum it.

5.Power of the NGD signal with gamma.

Results



Remarks :

Compared to the previous NGD plot, the powering of the plot with gamma boosts the low magnitude peaks ,a more closer version to fft of signal.

6.Modified Group Delay - Cepstral smoothed function with parameters .

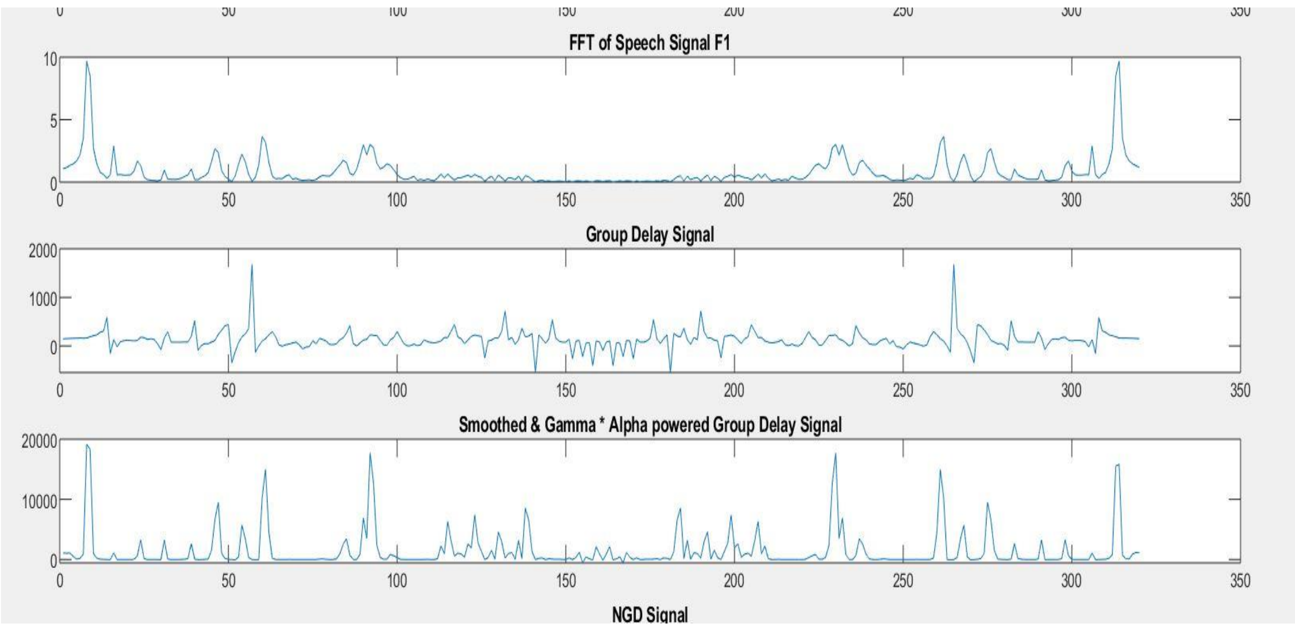
In certain functions, just replacing the denominator with (mag(Sc(jw))^2) isn’t sufficient. Hence, new parameters gamma and alpha are included.

$$\tau_c(e^{j\omega}) = \left(\frac{\tau(e^{j\omega})}{|\tau(e^{j\omega})|}\right)(|\tau(e^{j\omega})|)^{\alpha},$$

$$\tau(e^{j\omega}) = \left(\frac{X_R(e^{j\omega})Y_R(e^{j\omega}) + Y_I(e^{j\omega})X_I(e^{j\omega})}{S_c(e^{j\omega})^{2\gamma}}\right).$$

vo parameters α and γ can vary such that $(0< \alpha \leq 1.0)$ and $(0< \gamma \leq 1.0)$.

Results:



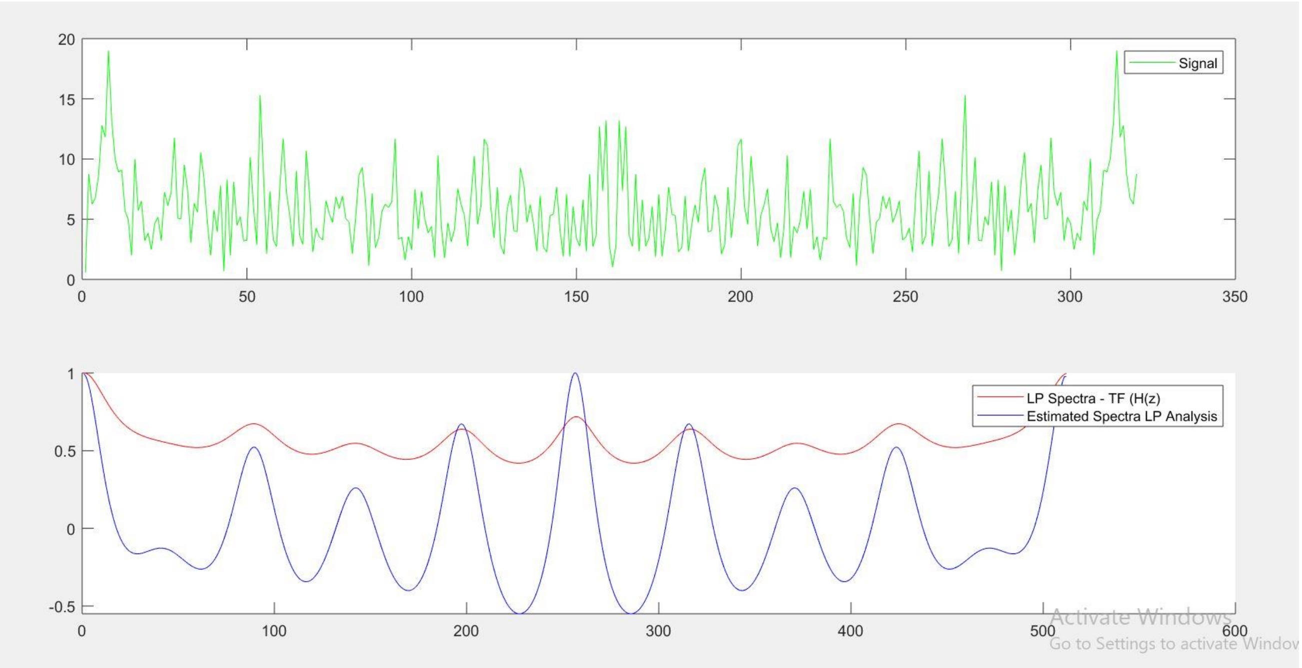
Remarks :

Here, there are few amplifications at some peaks. This method is more effective when the signal is affected too much with zeros.

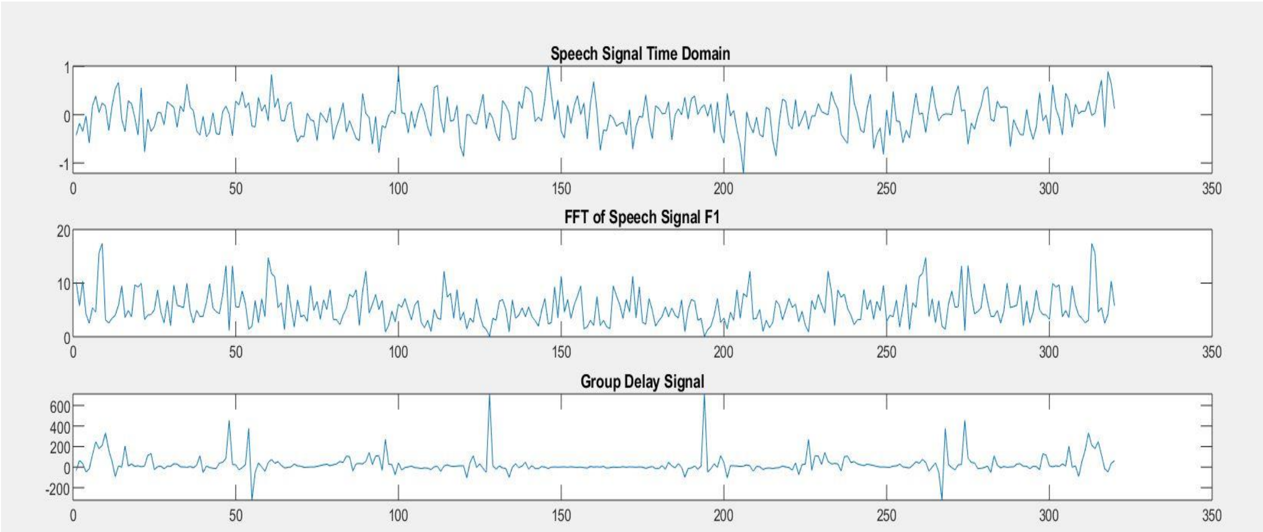
PART 2 – Robustness to Noise

So, all the above process are now tested with addition of AWGN noise of SNR - 10dB noise .

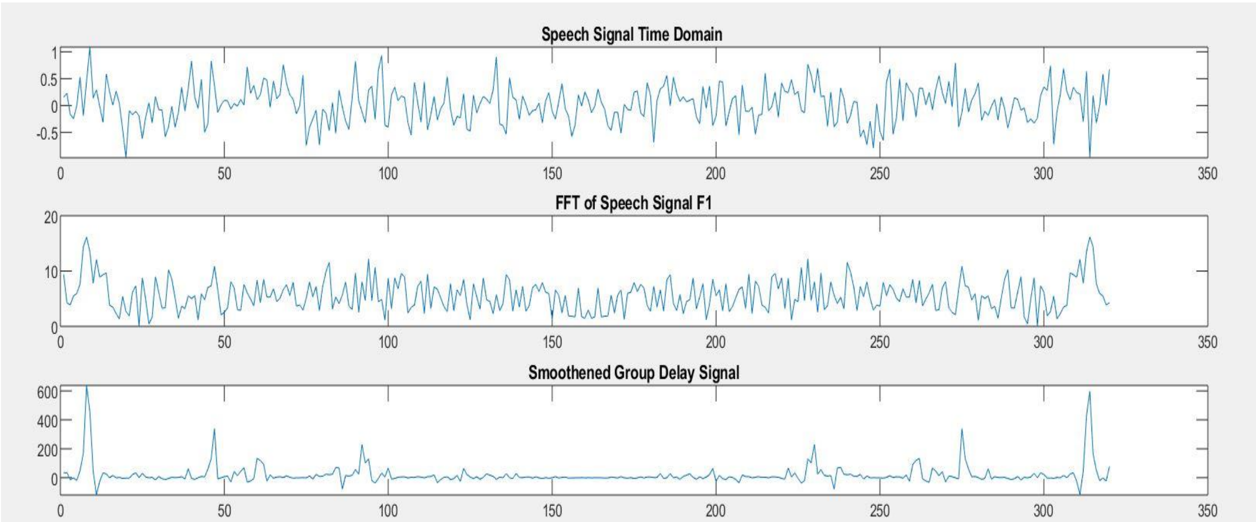
1.LP Analysis



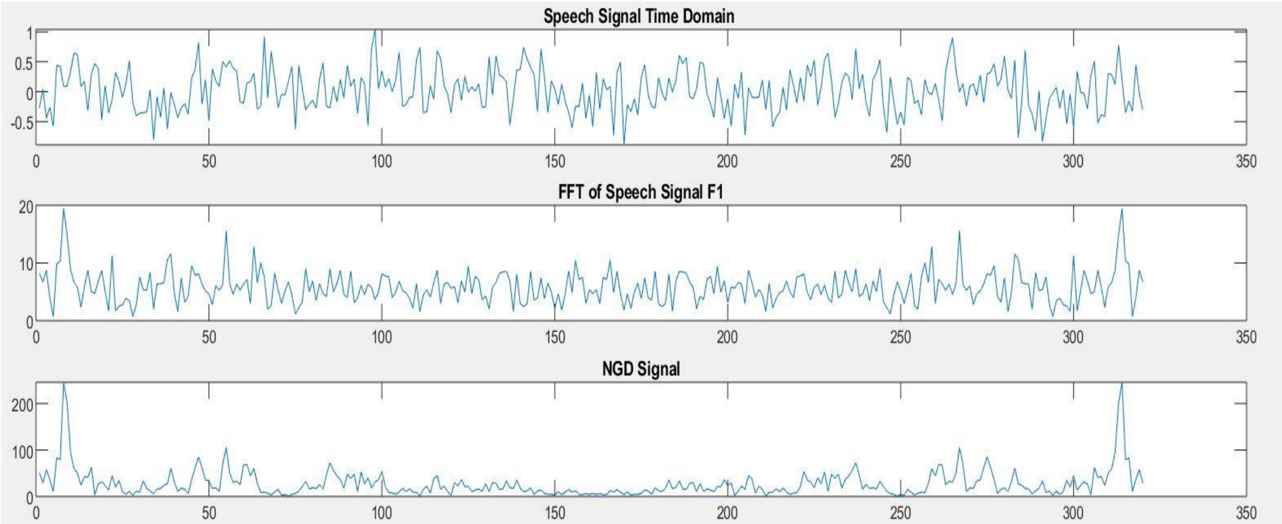
2.Group Delay



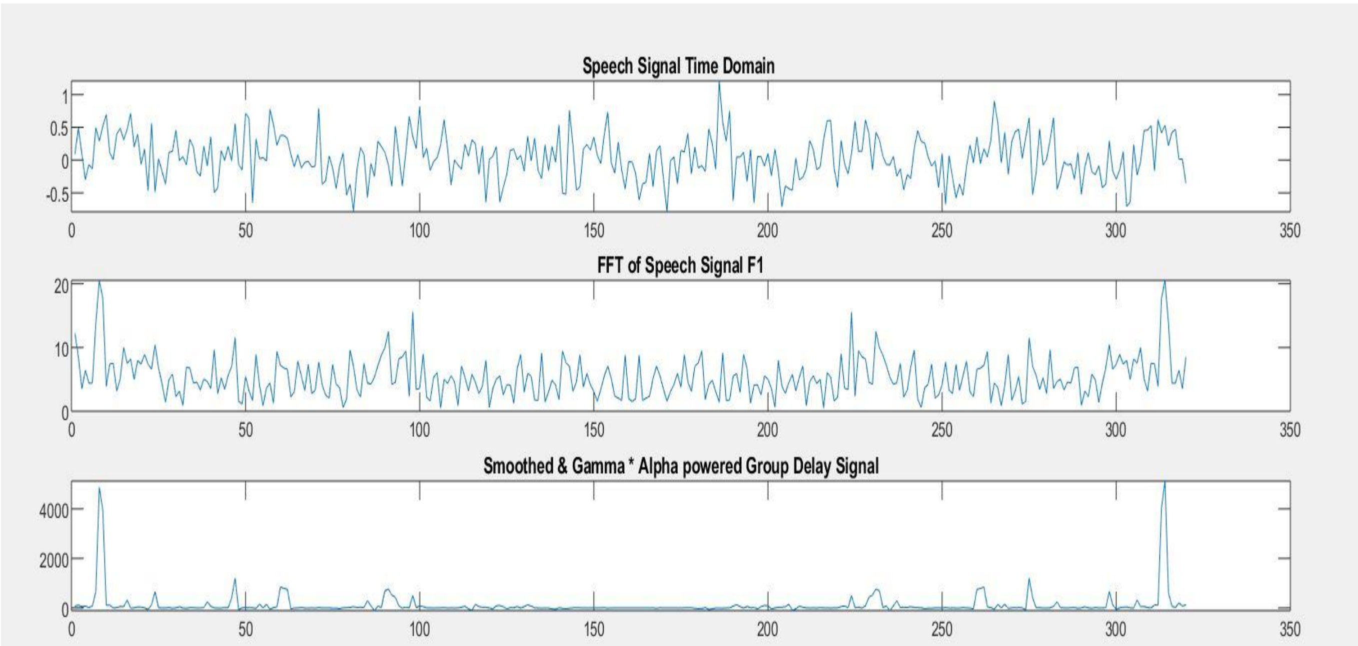
3.Modified Group Delay –Spectral Smoothing.



4. Numerator Group Delay (Powered)



5.Final Modified Group Delay



Conclusion:

1.The best version of the modified group delay cannot exist, as it largely depends on the speech data set ,and also the choice of parameters. A very good approximation is Numerator Group Delay , with gamma as parameter.

2.Robustness to Noise: Some of the proposed methods aren't that noise efficient in nature. Here NGD's performance is slightly better than the rest.

Name:Niharika Vadlmudi

Roll Number :2018122008

TA Mentor :Krishna.G