Audio Processing

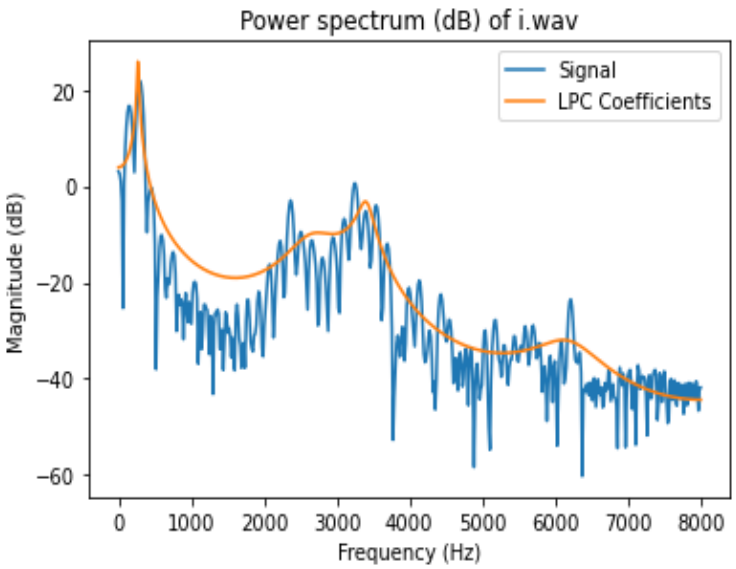
Ex04 – Report

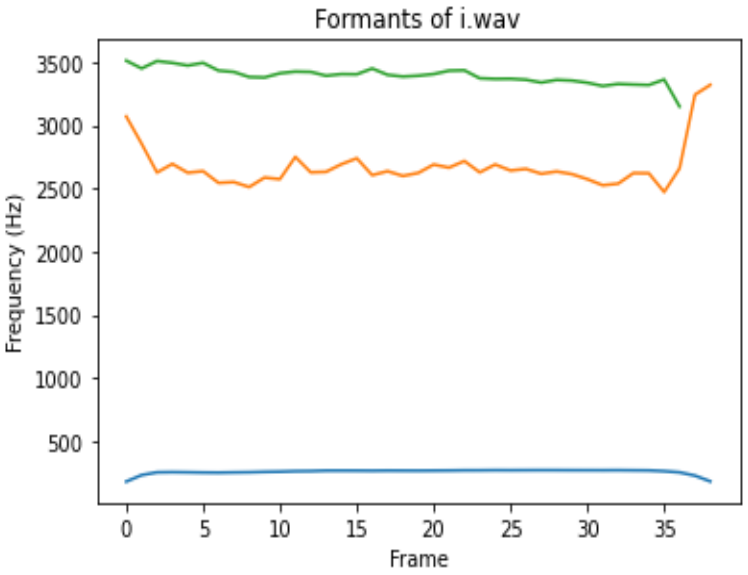
Student: Anh Huy Bui

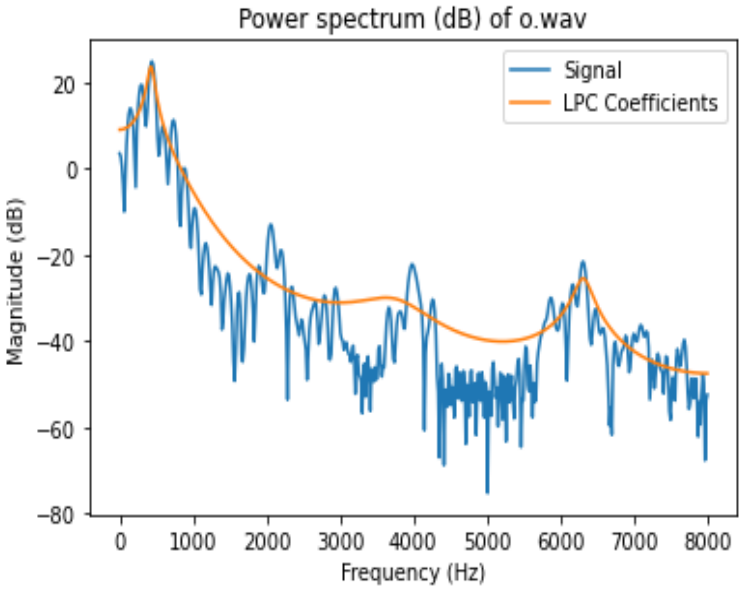
ID: 293257

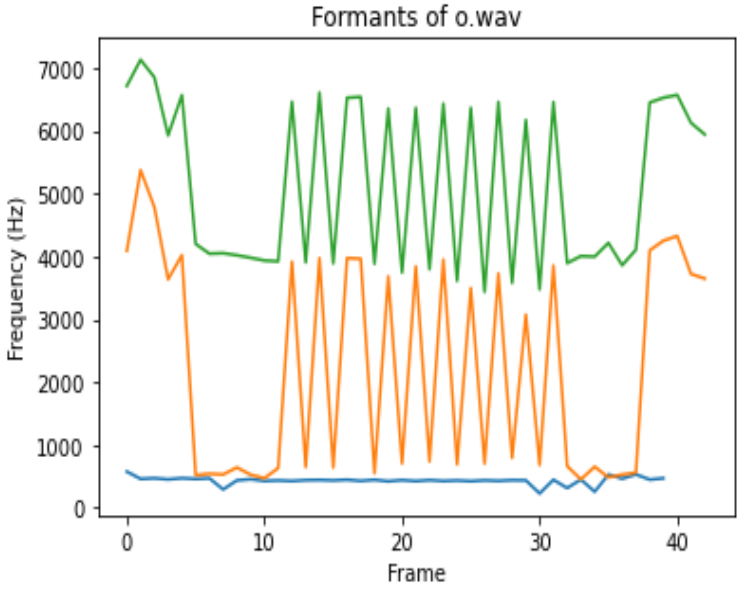
Problem 1:

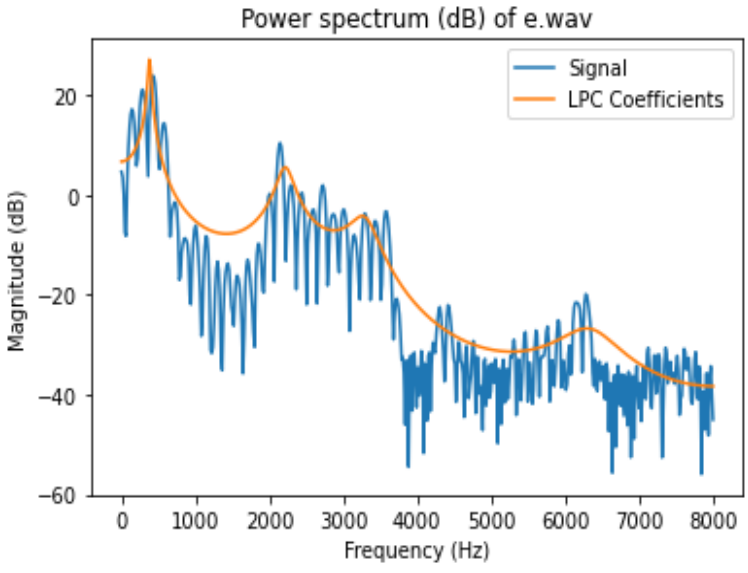
- **Filter order = 8**

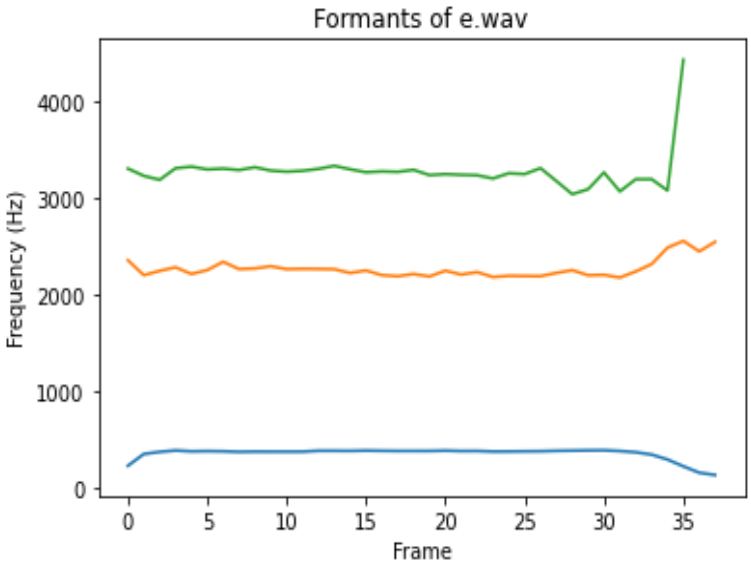


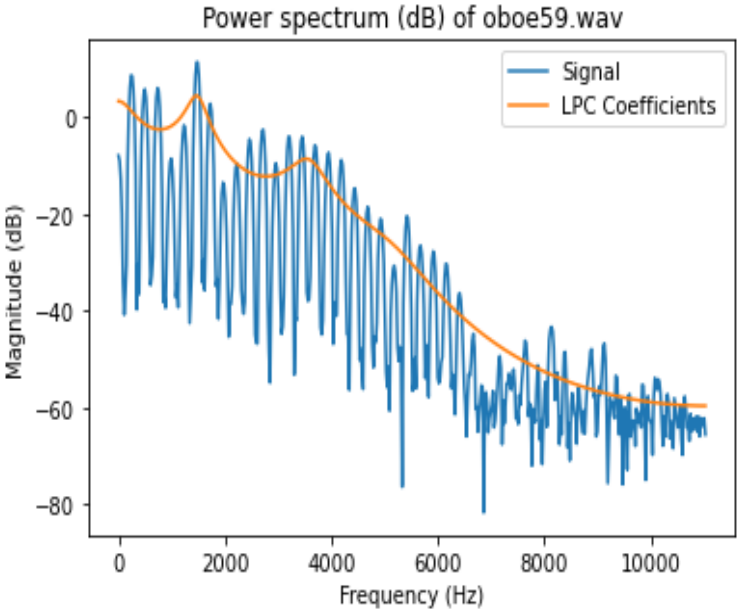


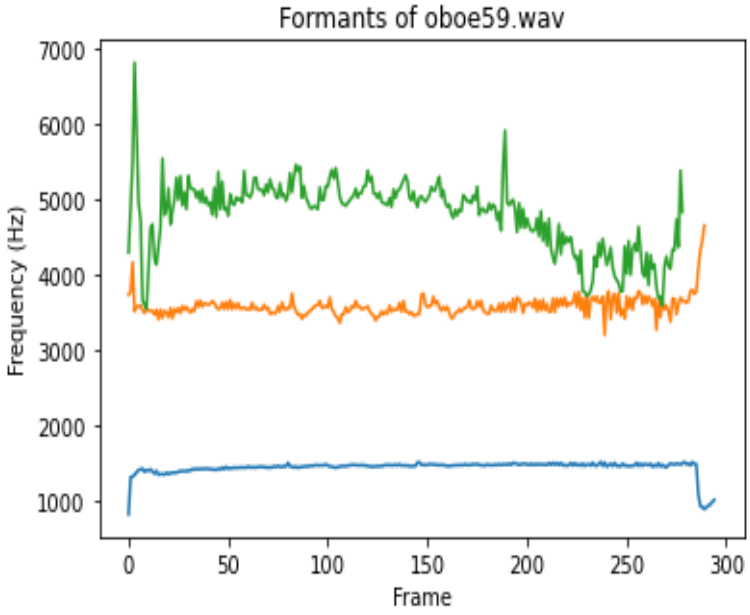




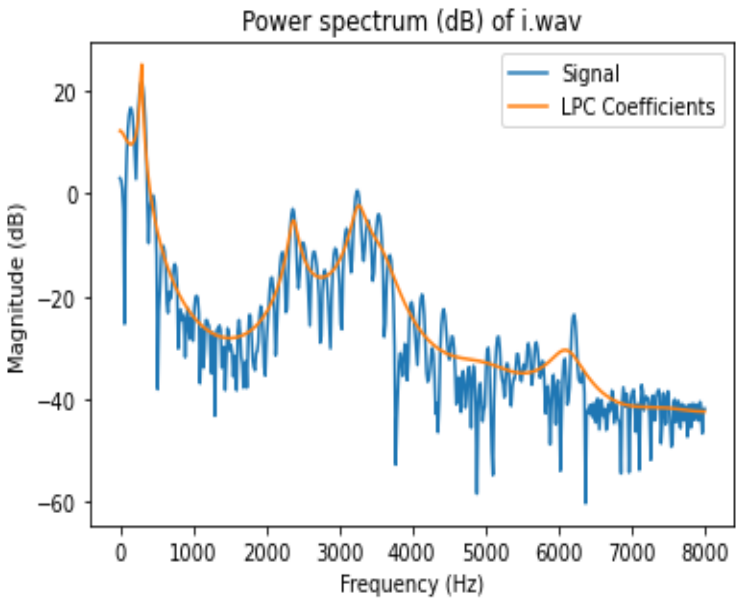


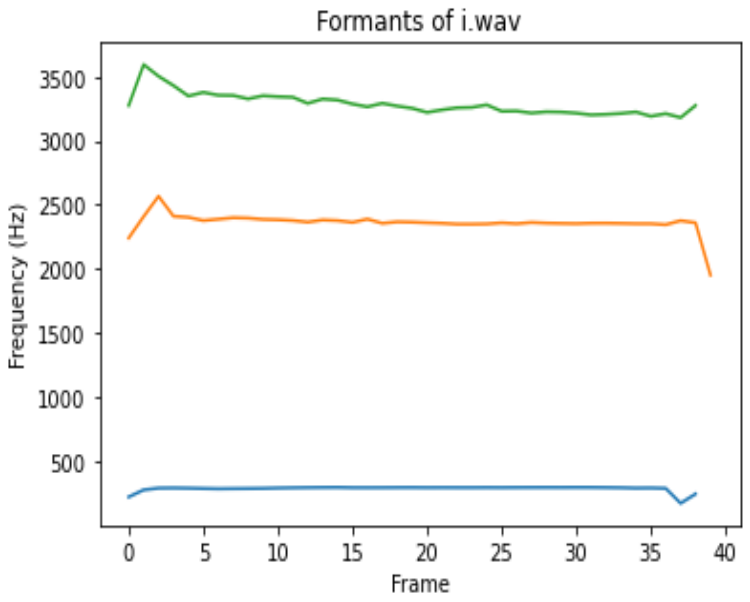


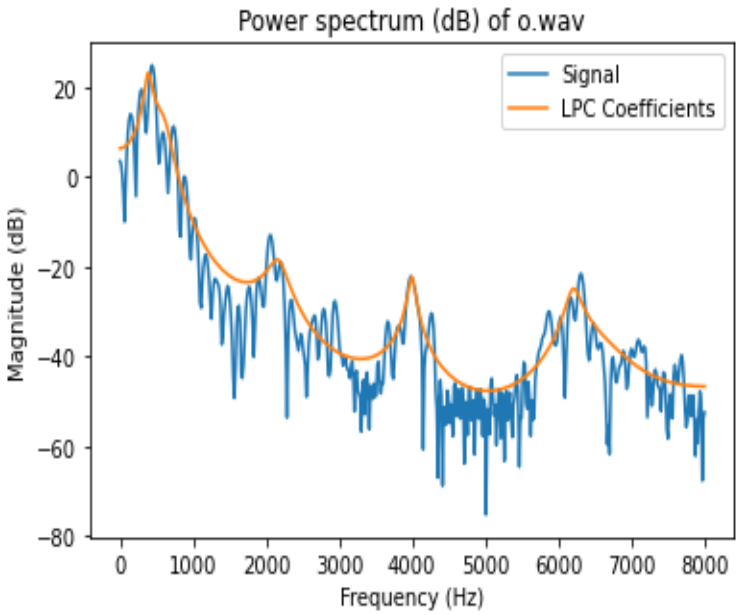

Power spectrum (dB) of oboe59.wav
Signal
LPC Coefficients
Magnitude (dB)
0
−20
−40
−60
−80
0
2000
4000
6000
8000
10000
Frequency (Hz)

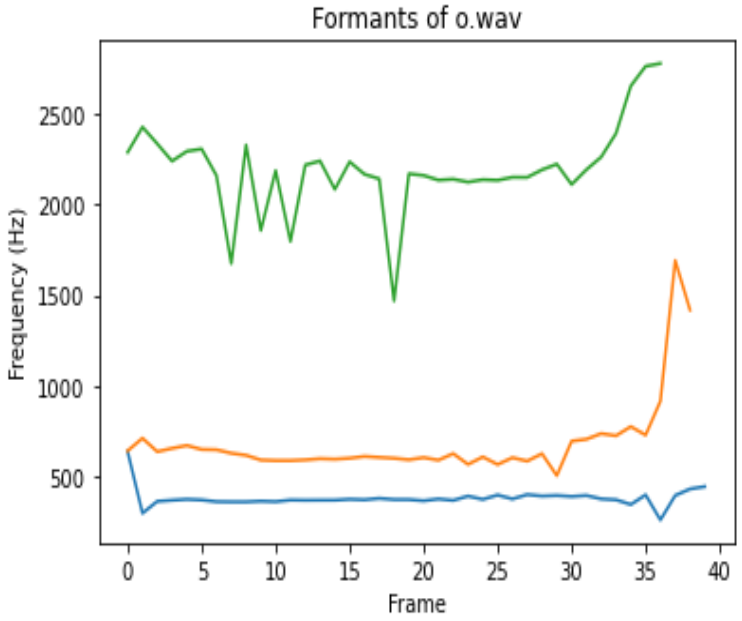

Formants of oboe59.wav
Frequency (Hz)
7000
6000
5000
4000
3000
2000
1000
0
50
100
150
200
250
300
Frame

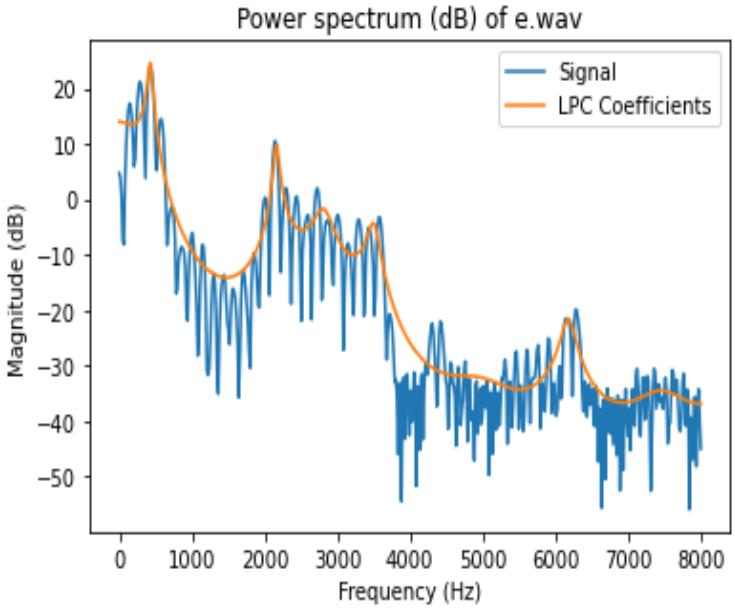
- **Filter order = 16**

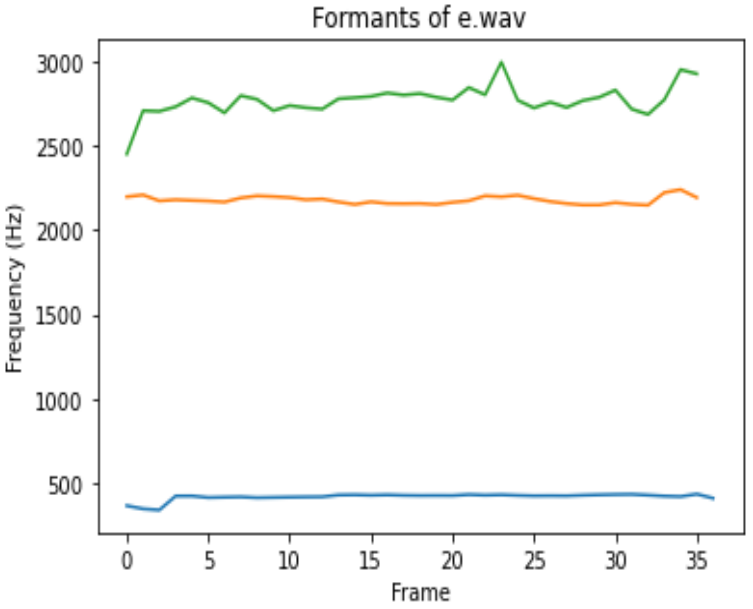


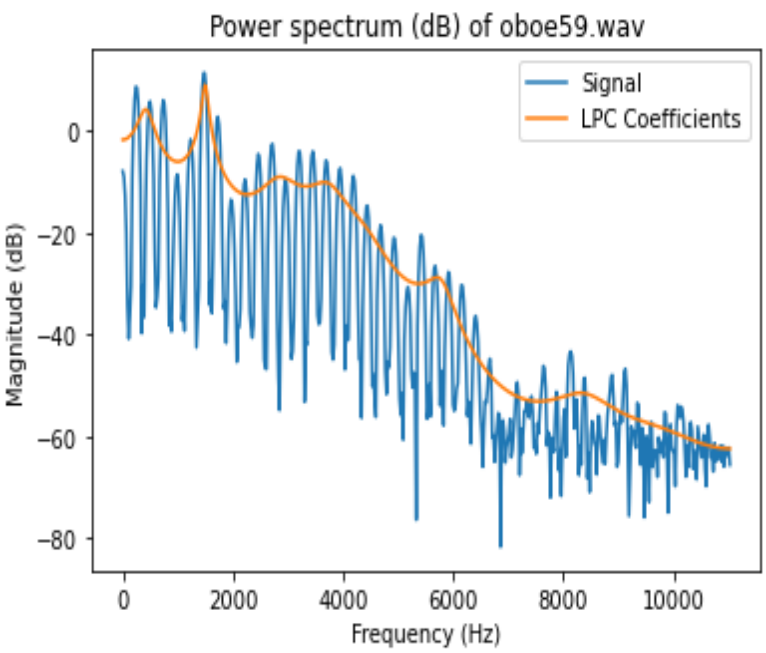


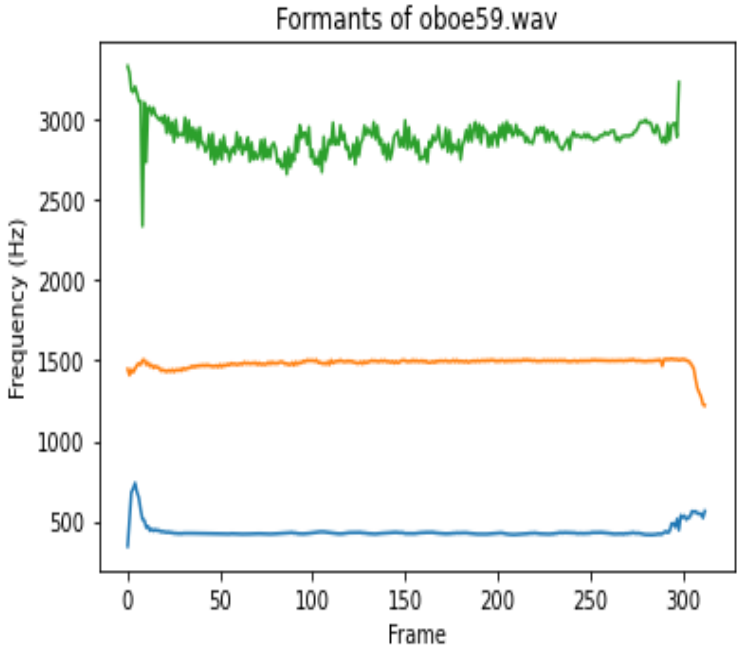




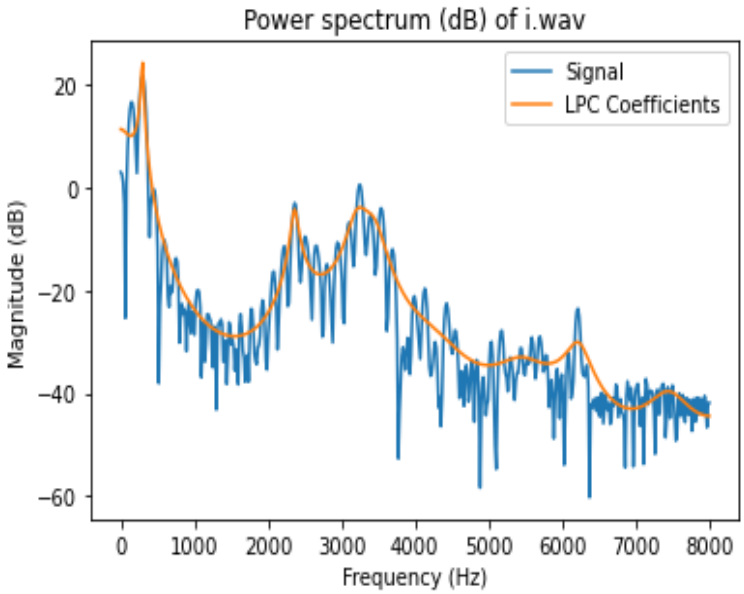


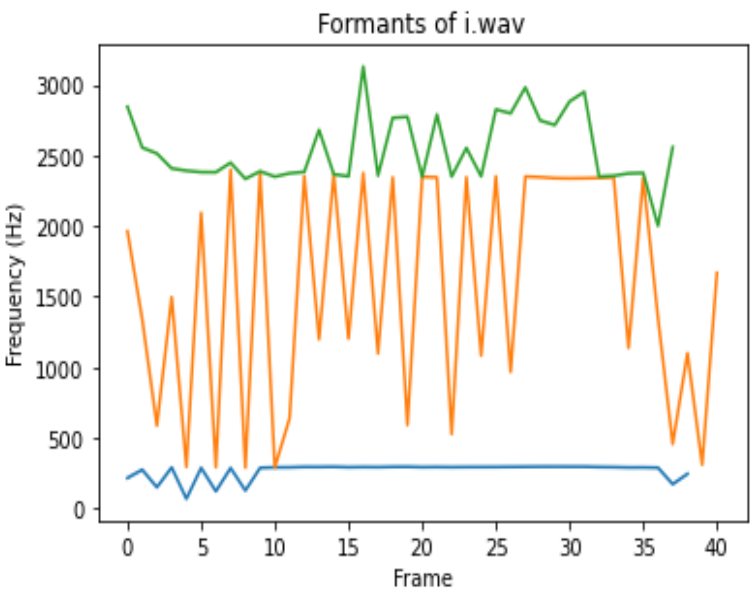


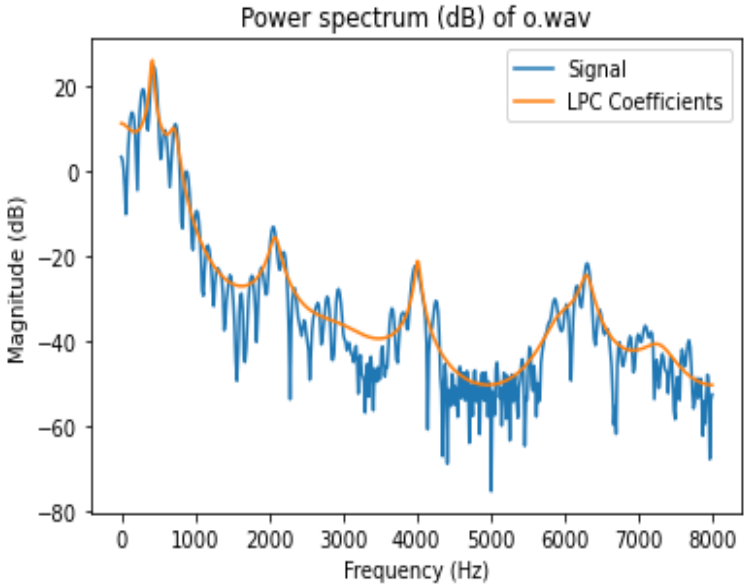

Power spectrum (dB) of oboe59.wav
Signal
LPC Coefficients
Magnitude (dB)
0
−20
−40
−60
−80
0
2000
4000
6000
8000
10000
Frequency (Hz)

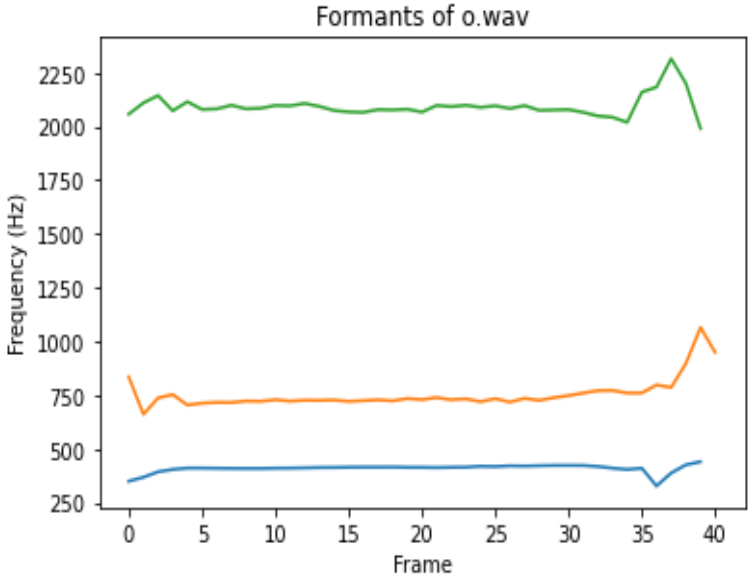

Formants of oboe59.wav
Frequency (Hz)
3000
2500
2000
1500
1000
500
0
50
100
150
200
250
300
Frame

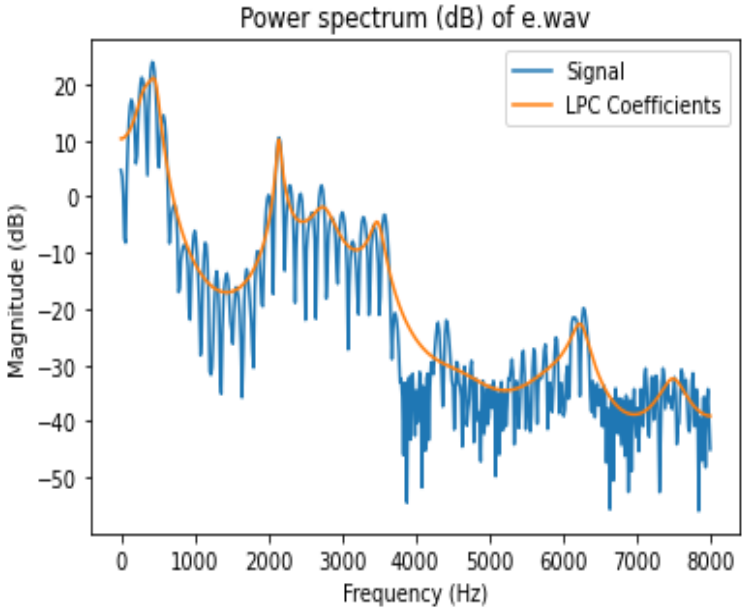
- **Filter order = 20**

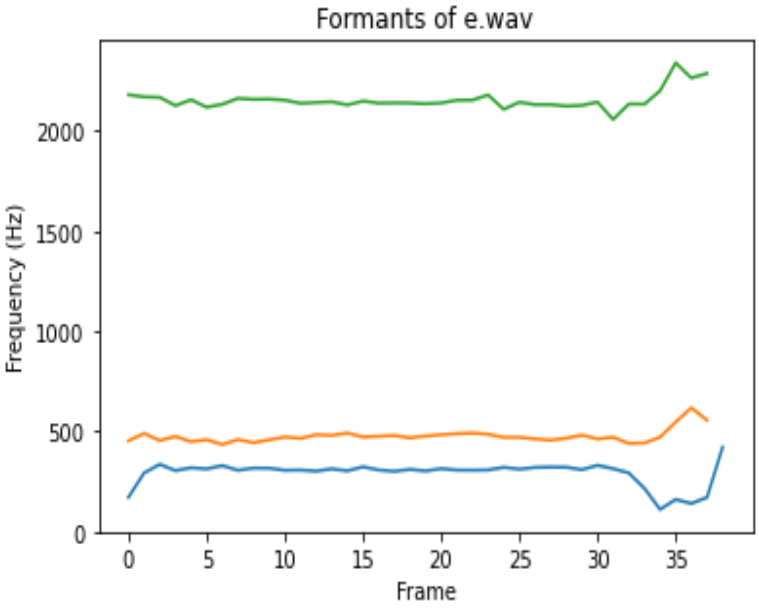


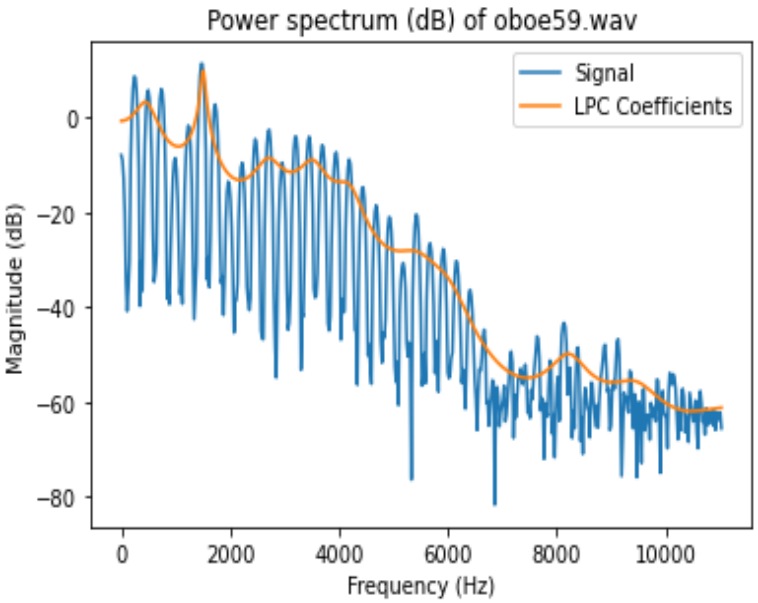


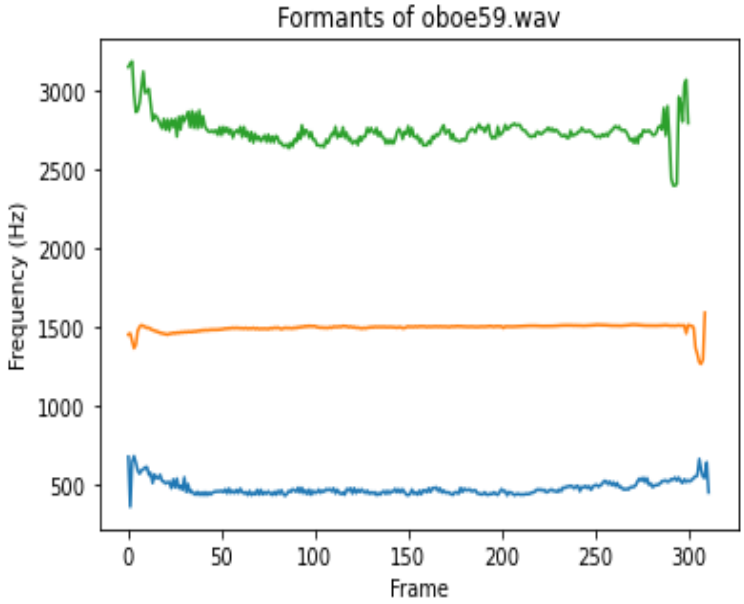












	i	o	e	oboe
Order = 8	259.53538 2674.1015 3392.5319	434.18361 2250.6473 5142.8207	352.29864 2265.0264 3275.0652	1440.0622 3581.6966 4793.3522
Order = 16	***286.0335*** ***2360.9112*** ***3287.9208***	385.22134 688.1659 2195.7403	***418.7426*** ***2177.6947*** ***2768.6828***	437.0447 1484.6430 2885.2586
Order = 20	268.5158 1568.9482 2538.6713	***409.2858*** ***754.1582*** ***2093.8191***	291.0970 475.9952 2155.8655	471.8560 1490.7389 2741.8838
Reference	294 2343 3251	406 727 2090	434 2148 2763	

(Bold indicates best suit to reference value)

Model order and peaks:

As can be seen from the left graphs, higher value of filter order results in more accurate LPC model, which stick sharply with signal DFT. In other words, lower value cause lost in information (signal peaks). However, too high value may contain too much information to process. As in this case, we only want to keep 3 first peaks, so it will be the best if those peak can represent the main frequency of the signal, yet high value of order gives some addition peaks (not the main one), so with only 3 peaks recorded, they cannot represent the whole signal. For example: Vowel “e” with order = 20 from the table is completely different from reference.

Problem 2:

Roots of the LP polynomial got from np.roots are used to calculate the angle/frequency at which the magnitude of transfer function returns maximum.

Conclusion:
I have plotted additional figures to display different effects of different orders.

In my opinion, when observing these figures, especially the “i” of order = 20, I find that peaks of “i” closest to reference is when order = 16. With an increase order, the formants is not stable throughout the whole “i” signal. Similar to “o” when it comes to too low (order = 8).

Bonus:

An oboe is a wind instrument, so it returns somewhat similar to speech processing. However, its FFT shows that the frequencies vary much faster then voice signal. It also requires a different order to get the best peaks. Formants figures always show that there is some minor variation of peaks throughout the whole signal.