

Speech Processing Lab - Week 2

Name: S U Swakath

Roll number: 180020036

Google Colab Link : <https://colab.research.google.com/drive/153hIFDhg5gRcPzGAx3ki8oiu639V72uso?sharing>

Aim

- To understand the difference between stationary and non-stationary signals.
- To get a feel about the non stationary nature of speech signals.
- To understand the limitations of Fourier transform in case of non-stationary signals.

Theory

Introduction

System is a physical entity that exists and produces signal. Signal is a physical quantity that is measurable. Signals are categorized into several classes based on the nature of the signal. Stationary vs/ Non-Stationary is one such classification.

What are Stationary and Non-Stationary signals ?

A signal is said to be stationary if its frequency or spectral contents are not changing with respect to time. One such example of stationary wave would be the sine wave that we generate using a function generator by selecting a fixed frequency. In this case the frequency content of the generated sine wave will not change with time, hence it is a stationary signal.

Further, the definition of stationarity should not be confused with the time varying amplitude in the time domain as in the case of sine wave. Stationarity is linked only to the behavior of the frequency contents of the signal with respect to time and nothing else.

Any signal that does not satisfy the stationary signal criteria is classified as a Non - Stationary signal.

Motivation

Most of the signals, systems and signal processing concepts are taught based on the implied assumption of stationarity of the signal under consideration. Speech signal processing deviates in this aspect. This is because, speech is an example for non-stationary signal where as conventional synthetic signals like sine wave, triangular wave and square wave are stationary in nature. Hence different approaches and tools are needed to process the speech signal. To appreciate the new approaches and also tools developed exclusively for speech processing, we should get a feel about stationary vs non-stationary classification of signal and to understand the limitation of Fourier Representation in analysing an non-stationary signal.

Problem A

Generation of a singletone sine wave and its spectrum.

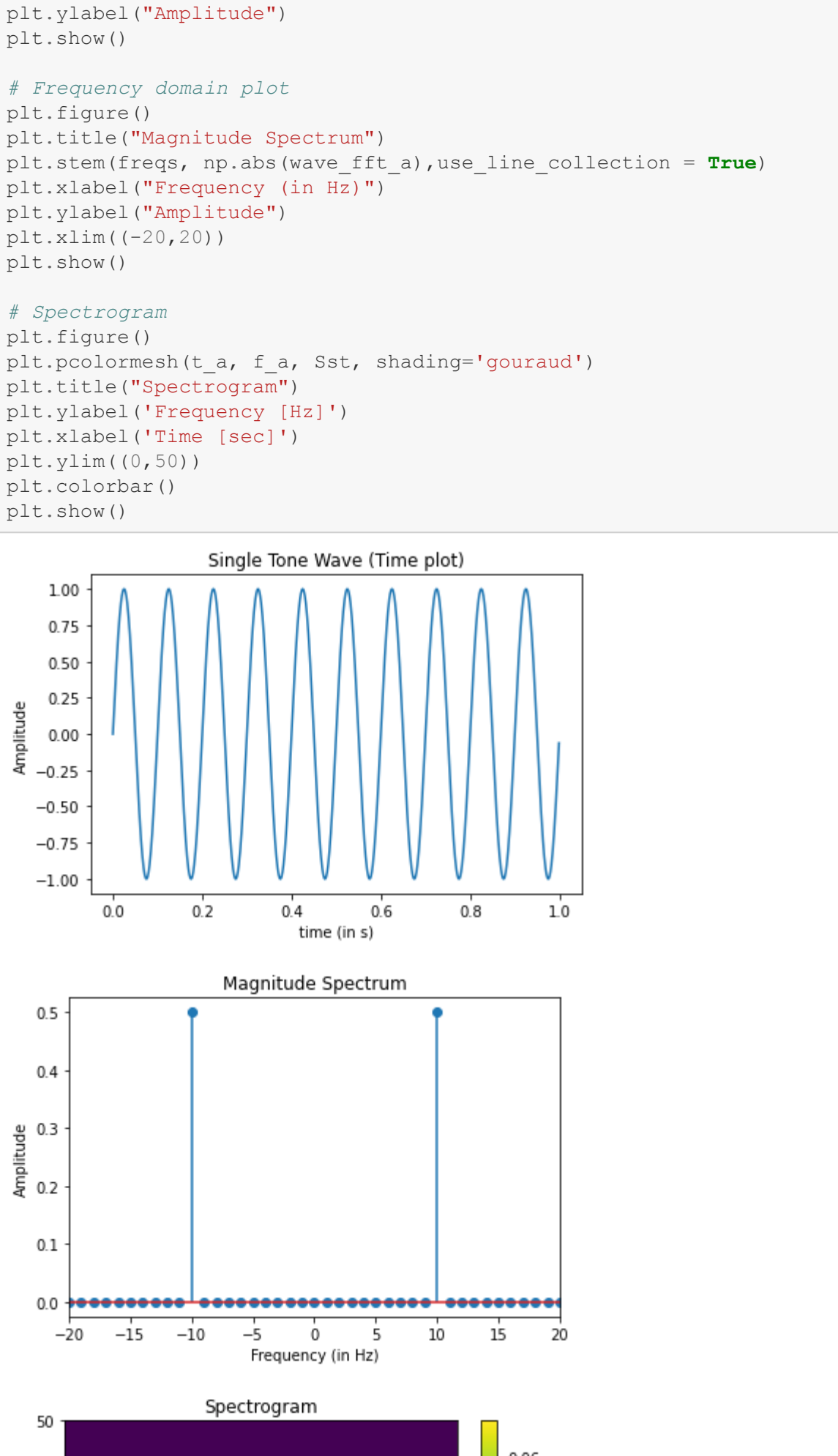
a. Generate a 10 Hz sine wave sampled at 1000 Hz sampling frequency and for a duration of 1 sec for this problem.

Procedure

1. Generate a singletone wave using numpy library.

```
SingleTone wave = $sin(2*pi*10*t); 0<=t<=1s
```

2. Generate time domain plot, frequency domain (amplitude) plot and spectrogram using scipy and matplotlib libraries.



Observation

From above frequency domain plot we can observe that the signal contain a single frequency component at 10 Hz. From spectrogram we can observe that the frequency property of the signal is constant throughout time. Hence the signal is a stationary signal.

Problem B

a. Generate a multitone sine wave composed of 10, 50 and 100 Hz frequency components. Use a sampling frequency of 1000Hz and a wave of duration 1 sec for the problem.

Procedure

1. Generate a multitone sine wave using numpy library.

```
Multitone wave = $sin(2*pi*10*t) + sin(2*pi*50*t) + sin(2*pi*100*t); 0<=t<=1s
```

2. Generate time domain plot, frequency domain (amplitude) plot and spectrogram using scipy and matplotlib libraries.



Problem C

Generation of a non-stationary multitone sine wave and its spectrum.

a. Generate a non-stationary multitone sine wave made of different combinations of 10, 50 and 100 Hz components. Example combination:

- $\text{Sin}(2\pi/10t) \cdot 0 < t < 0.2 \text{ sec}$
- $\text{Sin}(2\pi/50t) + \text{Sin}(2\pi/10t) \cdot 0.2 < t < 0.5 \text{ sec}$
- $\text{Sin}(2\pi/10t) + \text{Sin}(2\pi/50t) + \text{Sin}(2\pi/100t) \cdot 0.5 \leq t < 1 \text{ sec}$

b. Generate the frequency spectrum taking the whole signal at once and taking each stationary part of the signal individually. Record your observations on the spectrum obtained by using the whole signal, the stationary parts individually and the limitations of Fourier Transform when we consider the whole non-stationary speech signal.

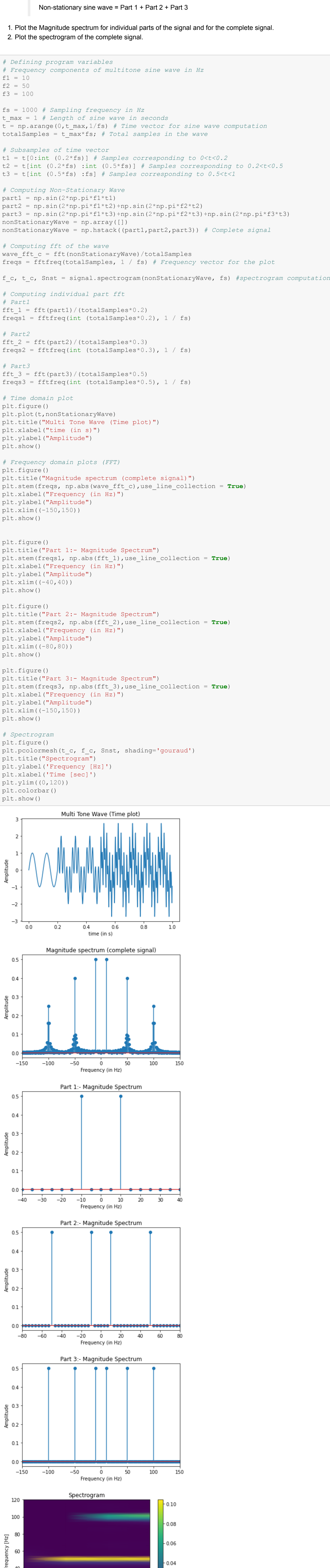
Procedure

1. Generate non-stationary sine wave using different combinations of the single tone components available.

- Part 1 = $\text{Sin}(\pi t^{10})$; $0 \leq t \leq 0.2 \text{ sec}$
- Part 2 = $\text{Sin}(\pi t^{10}) + \sin(\pi t^{50})$; $0.2 \leq t \leq 0.5 \text{ sec}$
- Part 3 = $\text{Sin}(\pi t^{10}) + \sin(\pi t^{50}) + \sin(\pi t^{100})$; $0.5 \leq t \leq 1.0 \text{ sec}$

```
Non-stationary sine wave = Part 1 + Part 2 + Part 3
```

1. Plot the Magnitude spectrum for individual parts of the signal and for the complete signal.
2. Plot the spectrogram of the complete signal.



Observations

As the sine wave generated is non-stationary the frequency property of the signal varies with time. This can be observe in the magnitude spectrum of individual parts of the entire signal, that is each part (Part 1 to Part 3) has different combinations of frequencies in it.

The magnitude spectrum of the entire signal provides the information about the different frequency components present in the signal, but not the information about when those frequency contents existed, that is, timing information is missing.

This is the limitation of Fourier Transform when we consider the whole non-stationary speech signal.

The essential information in the non-stationary signal is stored in the time varying spectrum. We cannot use the Fourier Transform to analyse an non-stationary signal as we lose the time domain information in the process. To resolve this issue we consider small segments of an non-stationary signal to be stationary and compute the spectra of each segment.

Problem D

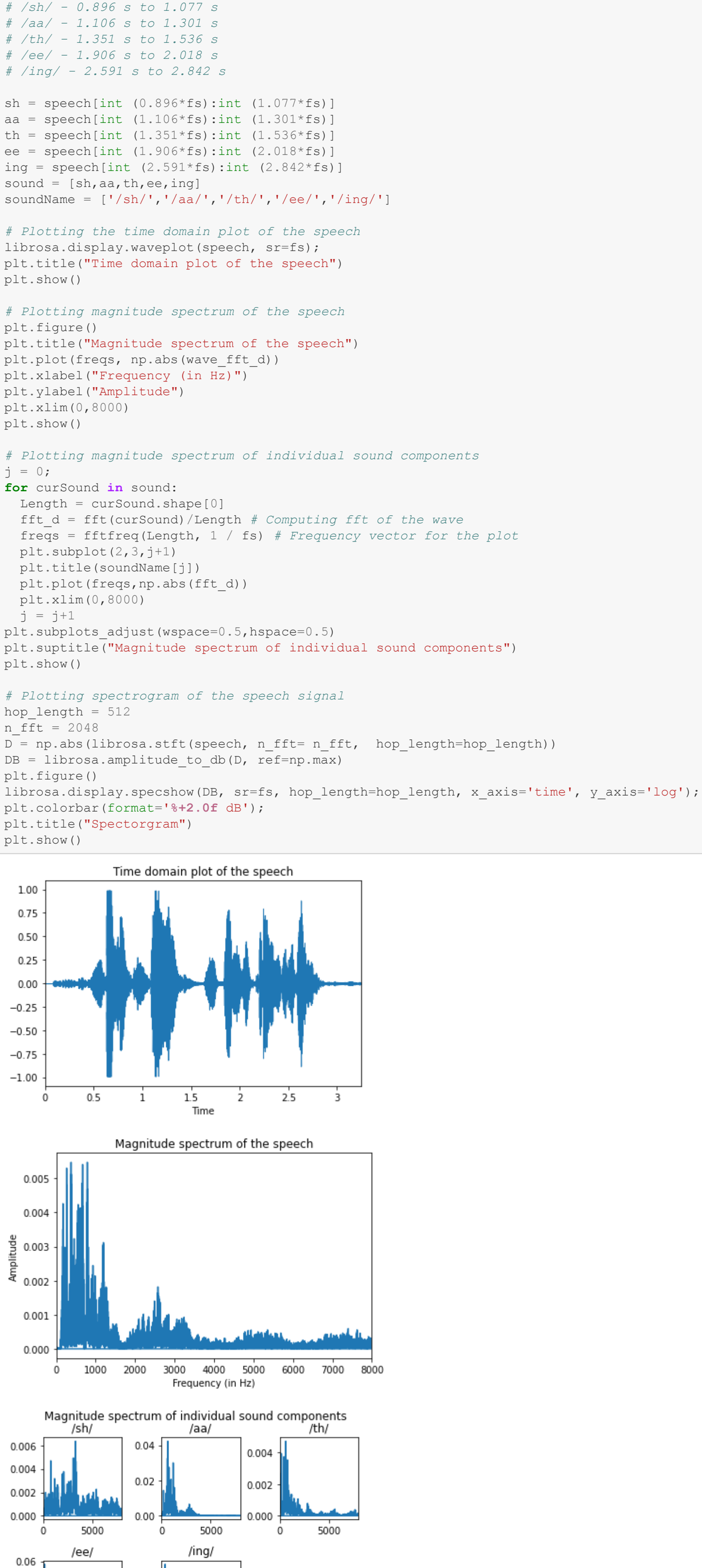
Plotting waveform and spectra of speech signal.

a. Record the phrase 'Sakshaat Speech Processing' and save it in a .wav file. Now, (re)sample the speech signal to 8KHz and plot the speech signal.

b. Compute and plot the spectrum of the different categories of sounds present in the recorded speech file. Inspect the spectral properties of sounds belonging to different categories, and comment on the nature of the whole speech signal.

Procedure

1. Record the phrase 'Sakshaat Speech Processing' using wavesurfer and save it in .wav format. Upload the file in Google drive and access it in the colob.
2. Read the .wav file using librosa library and resample it to 8K Hz
3. Use wavesurfer to identify any 5 different speech components present in the phrase 'Sakshaat Speech Processing' and record the time interval information of each component. In this lab I have identified the sounds sh,aa,th,ee,ing.
4. Plot the magnitude spectrum for the complete speech and for the individual components.
5. Plot the spectrogram of the speech signal.



Observation

From the above magnitude spectrum plots we can observe that each sound component has different spectral properties. Due to this we conclude that speech is a non-stationary signal. As discussed in the last session, a non-stationary signal cannot be analysed using Fourier Transform. Hence we use STFT to analyse speech signals

References and Tools

1. For theory concepts :- <https://yab.amrita.edu/index.php?sub=53&brch=164&sim=371&cnt=1104>
2. Basics of librosa library :- <https://librosa.org/doc/latest/mel-spectrogram.html>
3. Wavesurfer :- <https://sourceforge.net/projects/wavesurfer/>