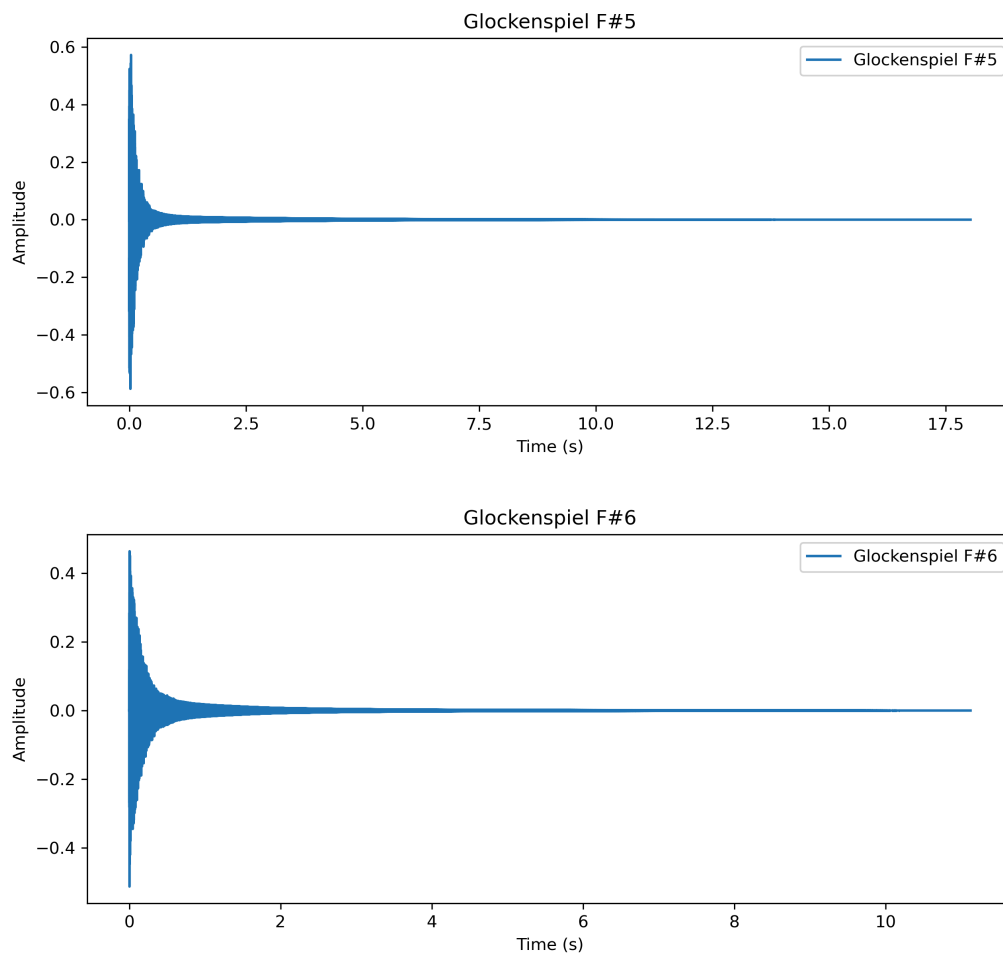


Assignment 2: Modal synthesis

Yilai Chen

1. Use python, plot the time domain waveforms of each pitches of a Glockenspiel. Label the axes amplitude and time in seconds.

Here I selected two sound files: `Gsp_ME_f_L-sus_F#5` and `Gsp_ME_f_L-sus_F#6`



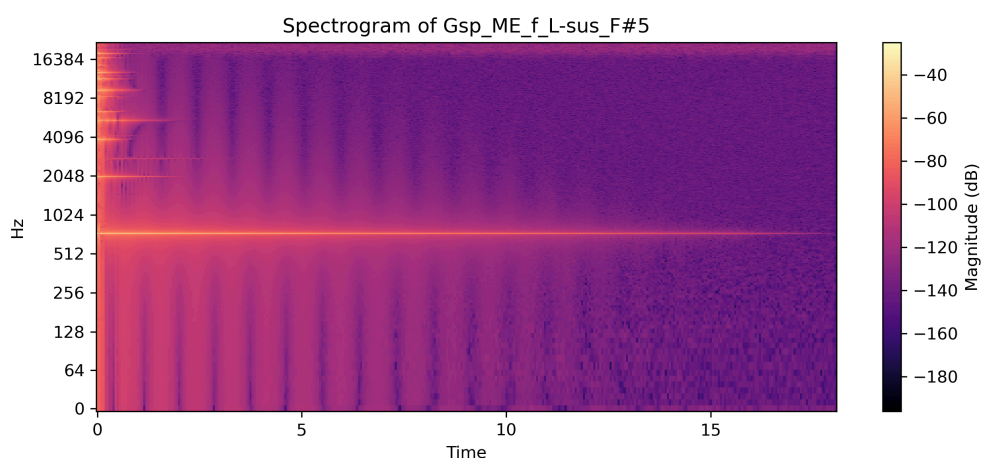
The image shows the waveforms of two Glockenspiel notes, F#5 and F#6. Both exhibit a sharp attack followed by an exponential decay, characteristic of percussive instruments.

2. Compute and plot the magnitude spectra at two different times of each sound using a rectangular window of size 100 ms. Locate the partials below 15kHz having the most energy, and estimate their frequencies and amplitudes at these time positions. Do the first spectral analysis as close as possible to the attack (onset of the note) and the second analysis about 500-1000 ms later.

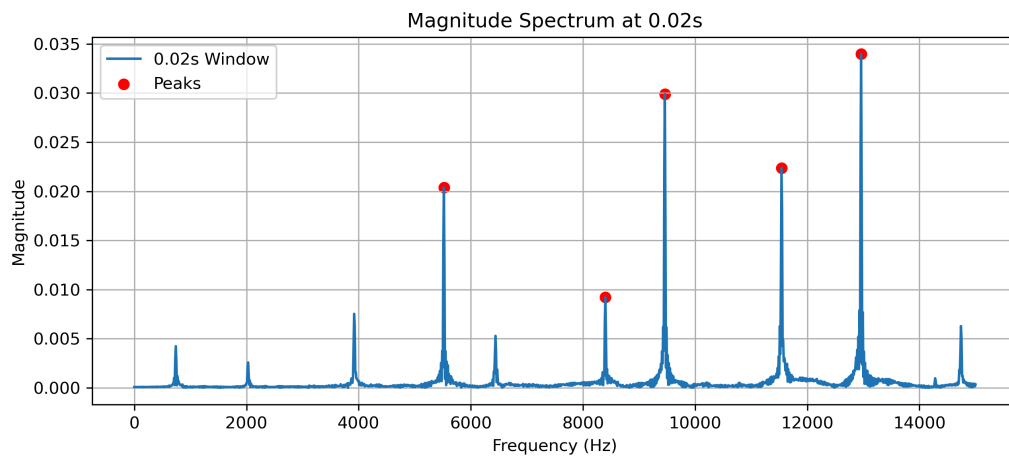
Here I chose two times: 0.02 seconds and 0.6 seconds

The output of codes are listed as follows:

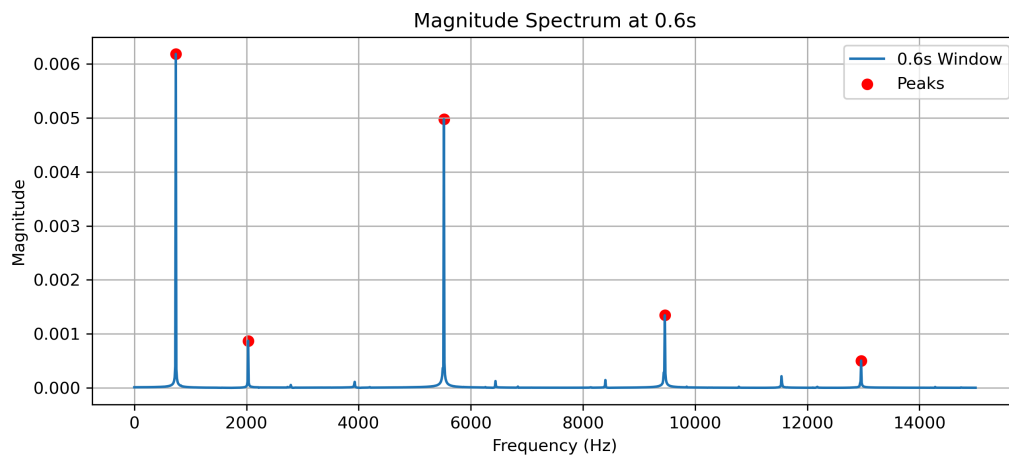
```
Time 0.020s for Gsp_ME_f_L-sus_F#5: Top Partial (Freq, Amplitude)
12960.0 Hz, 0.03398
9460.0 Hz, 0.02990
11540.0 Hz, 0.02238
5520.0 Hz, 0.02039
8400.0 Hz, 0.00921
Time 0.600s for Gsp_ME_f_L-sus_F#5: Top Partial (Freq, Amplitude)
740.0 Hz, 0.00619
5520.0 Hz, 0.00498
9460.0 Hz, 0.00134
2030.0 Hz, 0.00087
12960.0 Hz, 0.00050
Time 0.020s for Gsp_ME_f_L-sus_F#6: Top Partial (Freq, Amplitude)
7710.0 Hz, 0.03265
7910.0 Hz, 0.01081
1480.0 Hz, 0.00911
12510.0 Hz, 0.00680
7680.0 Hz, 0.00488
Time 0.600s for Gsp_ME_f_L-sus_F#6: Top Partial (Freq, Amplitude)
1480.0 Hz, 0.01159
7910.0 Hz, 0.00109
5300.0 Hz, 0.00032
12510.0 Hz, 0.00025
7720.0 Hz, 0.00012
```



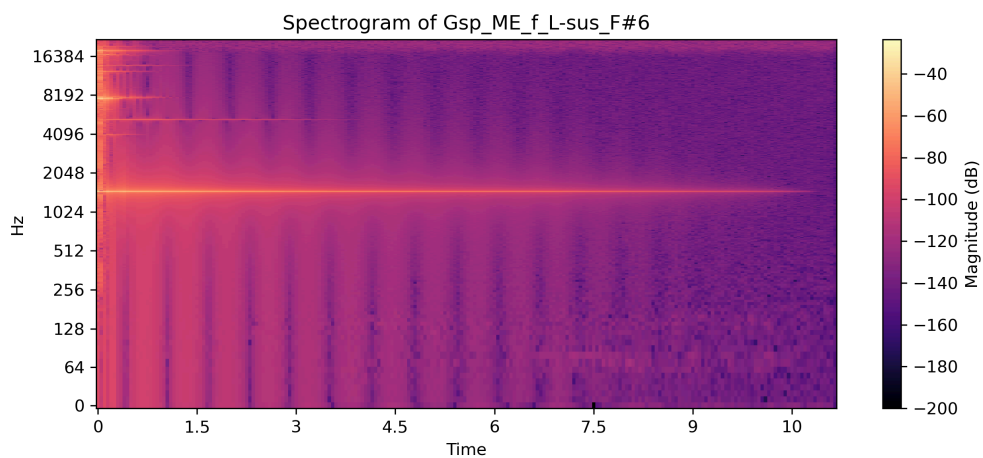
The spectrogram shows the frequency content of the Glockenspiel F#5 note, with strong harmonic components and a long decay, characteristic of its resonant nature.



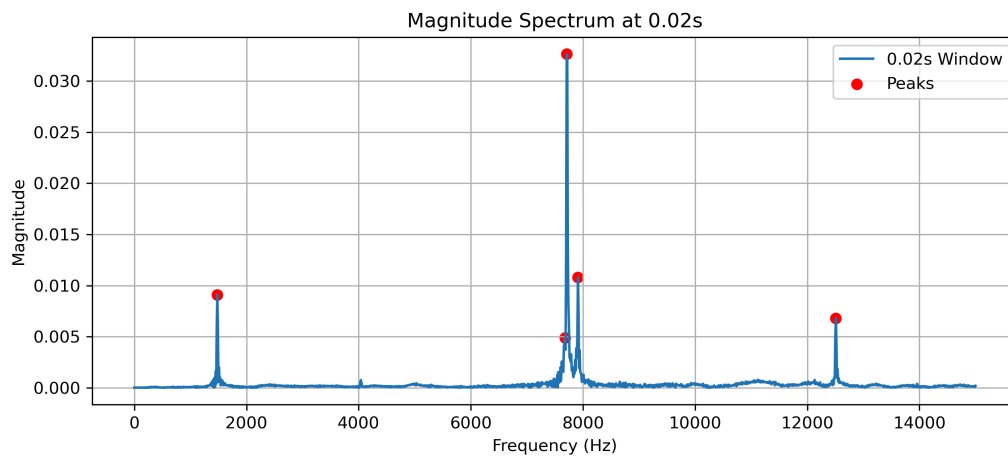
The magnitude spectrum at 0.02s highlights the dominant frequency components of the sound, with red markers indicating the peaks corresponding to prominent harmonics.



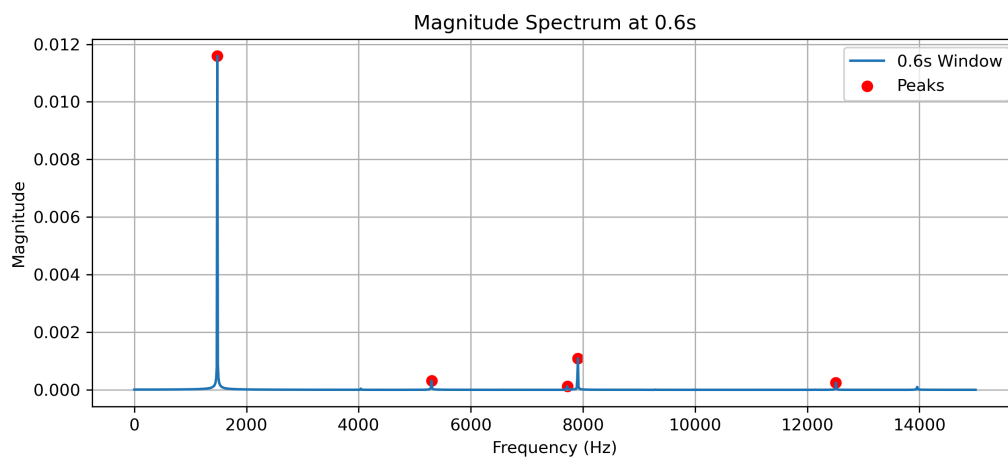
The magnitude spectrum at 0.6s shows a decay in harmonic amplitudes over time, with the fundamental and a few higher harmonics still present but significantly reduced. (Note that the range of the vertical axis is different, and the actual value decays greatly over time.)



The spectrogram of Glockenspiel F#6 reveals strong harmonic content with a clear fundamental frequency and overtones, exhibiting a long decay characteristic of its resonant nature.



The magnitude spectrum at 0.02s highlights dominant frequency peaks, with strong harmonics indicating the resonant nature of the sound.



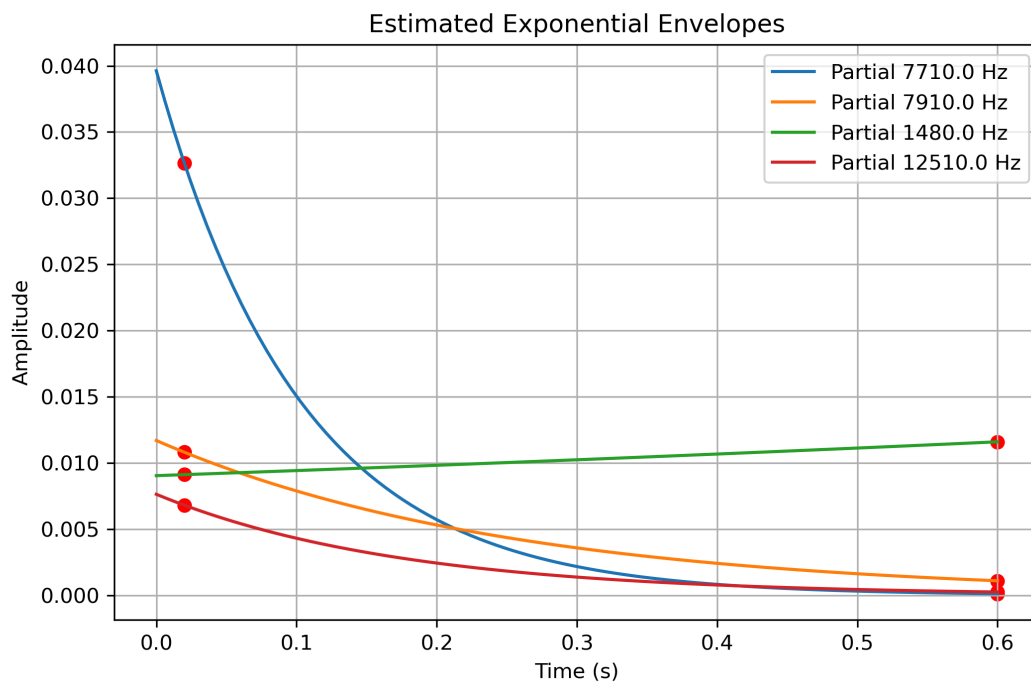
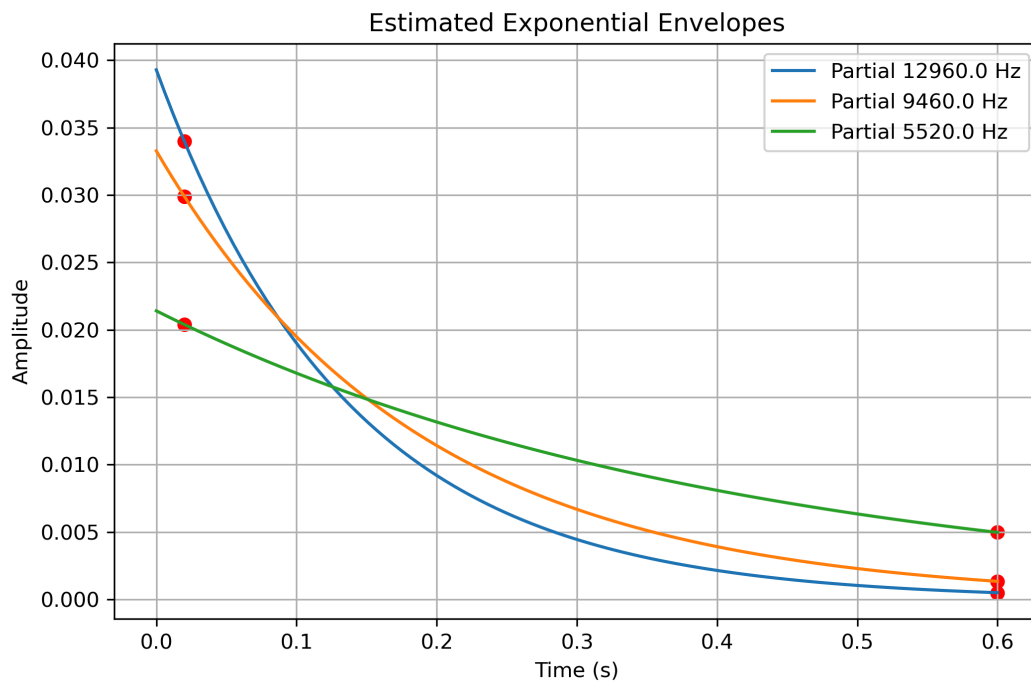
The magnitude spectrum at 0.6s shows a significant decay in harmonic amplitudes, with the fundamental frequency still dominant while higher harmonics diminish.

3. Compute the parameters of an exponential amplitude envelope for each of the partials you found from your spectral analysis at the two time positions. Plot the amplitude envelopes of these partials.

For the five main harmonics obtained at time = 0.02 in the previous step, I analyze whether they are still the main harmonics at 0.6 seconds. If so, I synthesize them as fixed harmonics. Therefore, for audio 1 and 2, I found three and four harmonics to synthesize respectively.

After calculation, the output is as follows:

Gsp_ME_f_L-sus_F#5: Top Partial: (A, gamma)
 Partial 12960.0 Hz: A = 0.0393, gamma = 7.2606
 Partial 9460.0 Hz: A = 0.0333, gamma = 5.3482
 Partial 5520.0 Hz: A = 0.0214, gamma = 2.4321
 Gsp_ME_f_L-sus_F#6: Top Partial: (A, gamma)
 Partial 7710.0 Hz: A = 0.0396, gamma = 9.6857
 Partial 7910.0 Hz: A = 0.0117, gamma = 3.9471
 Partial 1480.0 Hz: A = 0.0090, gamma = -0.4151
 Partial 12510.0 Hz: A = 0.0076, gamma = 5.7185



F#5 and F#6 for time decay, which are calculated and fitted using the decay factor and the initial amplitude.

The 1480 Hz may be a disturbing noise, but for the sake of consistency, I still included it in the calculation.

4. Make an additive synthesis of the two sound samples.

Finally, I got the simulated audio `Gsp_ME_f_L-sus_F#5_synthesized.wav` and `Gsp_ME_f_L-sus_F#6_synthesized.wav`.

5. Play a little melody

I tried to write a little melody using synthesized audio: `Gsp_ME_f_L-sus_F#6_melody` and `Gsp_ME_f_L-sus_F#6_melody`

6. Analysis and Comparison of Synthesized and Real Recordings

The synthesis results show that **F#6 was successfully reproduced**, closely resembling the original sound, while **F#5 exhibited noticeable discrepancies**. Several factors contribute to this difference in accuracy.

6.1 Key Differences in Partial Selection and Decay Rates

- At **Time 0.020s**, F#5 had a dominant partial at **11540.0 Hz with an amplitude of 0.02238**, which was among the most significant frequency components at that moment.
- However, by **Time 0.600s**, this partial had decayed completely, making it **invisible in the spectral analysis** and thus excluded from synthesis.
- In contrast, **F#6 maintained its key partials more consistently over time**. This suggests a much slower decay rate, making it easier to model accurately.

6.2 Differences in Decay Characteristics (γ values)

- F#5's partials exhibit steep decay rates

, e.g.,:

- **12960 Hz**: $\gamma=7.2606$
- **9460 Hz**: $\gamma=5.3482$
- **5520 Hz**: $\gamma=2.4321$
- **11540 Hz** was completely omitted, likely due to a high decay rate.
- In contrast, **F#6 had a more gradual decay**, especially for **1480 Hz** ($\gamma=-0.4151$), which actually showed a slight **increase in amplitude**.

6.3 Insufficient Number of Selected Partials & Insufficient time points were used to fit the decay curve.

- The **number of selected partials** for synthesis might have been too low, leading to missing spectral content.
- For **F#5**, at **Time 0.020s**, there were **strong partials at 11540 Hz and 8400 Hz**, but these were not included in the final synthesis model.
- This issue was less prominent in **F#6**, as its most dominant partials were consistently captured.
- Only two time points were used, which may be too few.

7. Possible Improvements

7.1 Better Partial Selection Strategy

- Instead of selecting partials based solely on their presence at one time point, consider **tracking their evolution over time**.
- Use a threshold-based method to **retain fast-decaying partials** that contribute significantly at early time points.

7.2 Adaptive Envelope Modeling

- Modify the decay modeling to **account for partials with rapid energy loss**.
- Instead of **fitting a single exponential function** to all partials, consider a **piecewise approach**, where different segments of the sound have different decay rates.

7.3 Increase the Number of Partial in Synthesis

- A greater number of partials should be included in synthesis, especially for **F#5**, where high-frequency content significantly affects timbre.
- Consider **dynamic partial selection**, where different sets of partials are used for different time segments.