TTT4295 - Acoustic Signal Processing

Assigment 1

Music Box

performed by

Kristian Gøystdal
Njål Svensen

Report by

Kristian Gøystdal

Department of Electronic Systems
Norwegian University of Science and Technology

Summary

This is a set of guidelines for writing the report. This document may also be used as a template for the reports.
The Summary is the report's "commercial". After reading the summary, the reader should understand quite well what has been done. The most important results and conclusions should be reported, hence it should be very similar to the Conclusions chapter.

Contents

Contents

Chapter 1

Introduction

This report presents the analysis of a music box recording using Fast Fourier Transform (FFT) to identify the fundamental frequencies and their harmonics. The goal is to extract the melody and bass notes played by the music box, and to evaluate the tuning accuracy of the instrument by calculating the deviation of the harmonic peaks from their ideal frequencies in cents.

Chapter 2

Theory

2.1 FFT

The Fast Fourier Transform (FFT) is an algorithm that can be used compute the Discrete Fourier Transform (DFT) of a time signal. The DFT converts a sequence of complex numbers in the time domain into another sequence of complex numbers in the frequency domain.

To increase the efficiency of the FFT, zero-padding can be used to increase the length of a signal by appending zeros to its end. This will improve the frequency resolution of the FFT, as the frequency bins will be more closely spaced without changing the actual frequency content of the signal.

2.2 Note Frequency Calculation

The frequency of a musical note can be calculated using the formula:

$$f = f_0 \cdot 2^{\frac{n-n_0}{12}} \tag{2.1}$$

where f is the frequency of the note, f_0 is the reference frequency (440 Hz for A4), n is the number of semitones away from the reference note, and n_0 is the semitone index of the reference note (A4=9).

2.3 Musical Notes and Frequencies

Deviation in cents can be calculated using the formula:

$$\text{Cents} = 1200 \cdot \log_2 \left(\frac{f_{\text{actual}}}{f_{\text{ideal}}} \right) \tag{2.2}$$

where f_{actual} is the frequency of the detected peak, and f_{ideal} is the ideal frequency of the corresponding note. Deviation in cents provides a measure of how much a note is out of tune relative to the ideal frequency, where 100 cents is equal to one semitone.

Chapter 3

Method and Equipment

3.1 Recording setup

To achieve the recording used for this assignment, a smartphone was used to record the sound of a music box in a quiet room. The phone was placed on a table next to the music box, approximately 10 cm away from it. The recording was done using a application that allows for lossless audio recordings in the WAV format.

When playing the music box, we made sure that each note from the music box was played separately, allowing for a clear split between the notes in the recording later. Some notes were played at the same time as an accord, which we couldnt avoid due to the setup of the music box.

3.2 Post-processing

To process the recording, Python was used with the libraries NumPy, Matplotlib and SoundDevice. The recording was loaded into Python using the SoundDevice library, and then split into individual notes by manually selecting the start and end points of each note in the waveform. The starting point of a segment was chosen to be just before the note started, and the end point was chosen to be the startpoint of the next note. This was to ensure that the full note was captured and that it had decayed fully before the next note started.

3.2.1 FFT

To analyze the frequency content of each note, Fast Fourier Transform (FFT) was used. The FFT was computed using the NumPy library, which provides an efficient implementation of the algorithm.

To ensure that the FFT has enough frequency resolution, the length of each note segment was adjusted to be a power of two. This was achieved by zero-padding, which means finding the next power of two greater than the longest note segment, and padding the shorter segments with zeros to match this length.

The FFT was then computed for each segment, and the peak frequencies were identified using the NumPy function `np.argmax()`. The peak frequencies were then plotted using Matplotlib, with the highest peak set to 0 dB.

Harmonic peaks were identified by looking for peaks in the FFT magnitude spectrum that were integer multiples of the fundamental frequency. This was done by looking at the spectrum and identifying peaks that were evenly spaced apart. These were then marked on the plot using vertical lines.

Using the peak frequencies, the perfect harmonic peaks were calculated by multiplying the fundamental frequency by integers (2, 3, 4, etc.). These were used to compare with the actual harmonic peaks identified in the FFT and calculate the deviation in cents.

3.2.2 Note identification

To identify the musical notes corresponding to the detected frequencies, the standard tuning frequency of A4 (440 Hz) was used as a reference. The frequency-to-note mapping was done using Equation 2.1, which calculates the note number from the frequency. The resulting n value was rounded to the nearest integer to determine the closest musical note.

The note names were then determined using a predefined list of note names corresponding to MIDI note numbers. The octave was calculated by dividing the MIDI note number by 12 and adjusting for the octave offset, which in this case was 4.

The identified notes were then categorized into melody and bass based on their frequencies. Notes with frequencies above middle C (approximately 261.63 Hz) were classified as melody, while those below were classified as bass.

Chapter 4

Results

4.1 Segmented Audio Signal

The audio signal was successfully segmented into individual notes based on the manually selected time splits. Each segment corresponds to a distinct note played by the music box. As seen in Figure 4.1, the waveform of Segment 1 clearly shows the onset and decay of the note, and shows that there note is allowed to die out before another note occurs. This segmentation allows for isolated analysis of each note's frequency content.

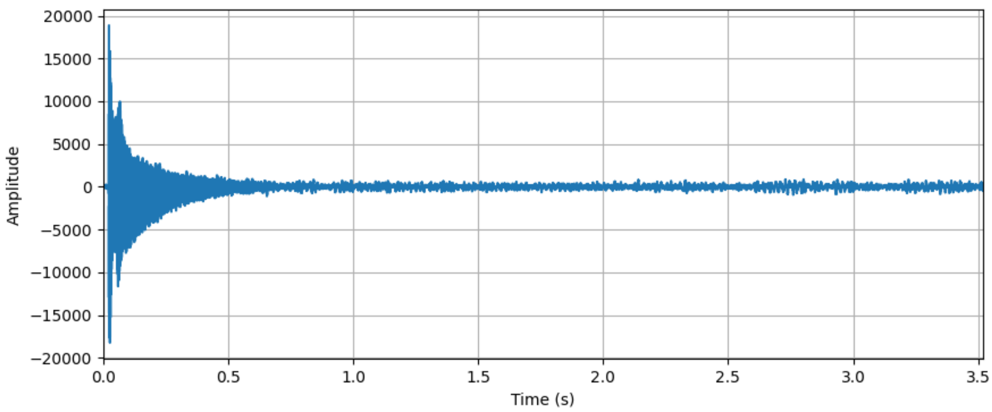

Figure 4.1: Waveform of Segment 1

4.2 FFT Analysis of Music Box Recording

4.2.1 Segmented FFT Analysis

If we look at the FFT of Segment 1 in Figure 4.2, we can see that the fundamental frequency is approximately 1106.56 Hz.

4.2.2 Overharmonic Peaks Identification

Also as seen in Figure 4.2, the harmonic peaks are identified by looking for peaks in the FFT magnitude spectrum that are integer multiples of the fundamental frequency. The deviation of the actual harmonic peaks from the perfect harmonic frequencies is calculated in cents and is summarized in Table 4.1.

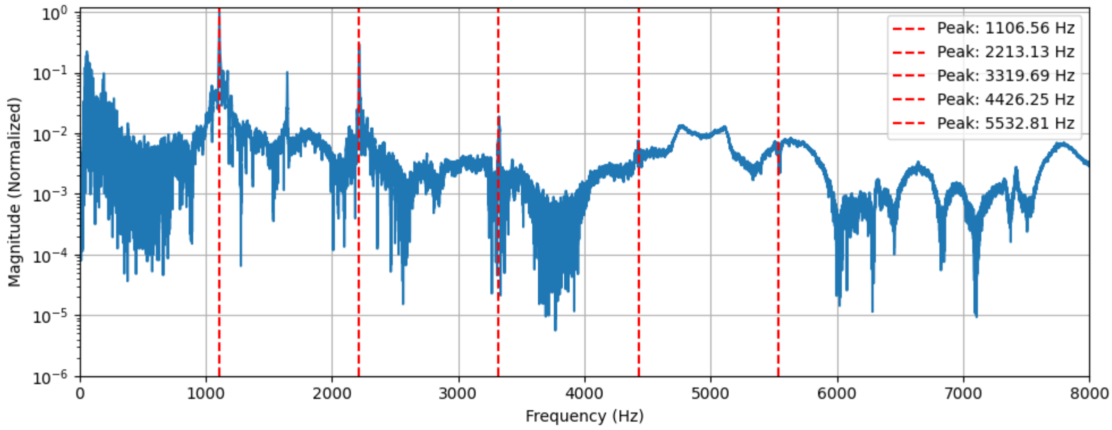

Figure 4.2: FFT of Segment 1 with identified peaks

Table 4.1: Deviation of Harmonic Peaks from Perfect Harmonics for Segment 1

Harmonic	Actual Frequency (Hz)	Deviation (cents)
1st (Fundamental)	1106.56	0
2nd	2213.12	0
3rd	3319.68	-5.85
4th	4426.24	-11.73
5th	5532.80	-17.61

4.3 Note Identification and Melody Extraction

Using Equation 2.1, the identified peak frequencies were mapped to their corresponding musical notes. The notes were then categorized into melody and bass based on their frequencies. Notes above middle C (approximately 261.63 Hz) were classified as melody, while those below were classified as bass.

The melody frequencies and their corresponding notes are summarized in Table 4.2. Also the deviation in cents from the ideal frequency is included in this table.

The bass frequencies and their corresponding notes are summarized in Table 4.3.

Table 4.2: Identified Melody Frequencies, Corresponding Notes, and Deviation in Cents

Segment	Frequency (Hz)	Note	Deviation (cents)
1	1106.56	C#6	-3.39
3	1174.57	D6	-0.13
5	1319.50	E6	1.30
6	1401.42	F6	5.58
7	591.87	D5	13.33
8	1105.89	C#6	-4.44
10	1171.75	D6	-4.29
11	1319.41	E6	1.19
12	1396.42	F6	-0.61
13	1863.25	A#6	-1.30
15	591.91	D5	13.45
16	1767.40	A6	7.27
17	1174.57	D6	-0.13
18	1401.38	F6	5.53
19	1751.76	A6	-8.12
20	1171.62	D6	-4.48
21	1656.67	G#6	-4.75
22	5145.13	E8	-42.84
25	1562.33	G6	-6.25
26	1396.63	F6	-0.35
27	1173.52	D6	-1.68
28	1047.30	C6	1.33
29	1171.67	D6	-4.42
30	470.87	A#4	17.39
31	591.91	D5	13.45
33	1106.44	C#6	-3.59
34	1174.57	D6	-0.13
36	1319.46	E6	1.24
37	1770.05	A6	9.86
38	1401.30	F6	5.43
39	591.91	D5	13.45
40	1106.14	C#6	-4.05
41	1171.79	D6	-4.23
42	699.70	F5	3.09
43	1319.41	E6	1.19
44	1396.34	F6	-0.72
45	1863.55	A#6	-1.03
46	591.87	D5	13.33
47	1757.56	A6	-2.40
49	1401.26	F6	5.37
50	1767.24	A6	7.10
51	1767.53	A6	7.39
52	2337.53	D7	-8.71
53	470.83	A#4	17.24
54	1106.60	C#6	-3.32
55	1562.38	G6	-6.20
56	2211.57	C#7	-4.61

Table 4.3: Identified Bass Frequencies, Corresponding Notes, and Deviation in Cents

Segment	Frequency (Hz)	Note	Deviation (cents)
2	53.20	G#1	+42.46
4	86.85	F2	-9.13
9	53.83	A1	-37.13
14	56.31	A1	+40.89
23	57.11	A#1	-34.72
24	59.26	A#1	+29.10
32	76.80	D#2	-22.08
35	42.86	F1	-31.92
48	87.86	F2	+10.87

Chapter 5

Discussions

As shown in Table 4.1, we can see that the deviation increases with higher harmonics, indicating that the music box's tuning is not perfectly harmonic.

In Table 4.2, we can see that the melody notes also have some deviation from the ideal frequencies, with deviations ranging from approximately -42.84 to +17.39 cents. This indicates that while the music box is generally well-tuned, there are some notes that are slightly out of tune. The bass notes in Table 4.3 also show deviations, with values ranging from -37.13 to +42.46 cents. As 100 cents is equal to one semitone, these deviations are relatively small but still noticeable to a trained ear.

Overall, the analysis shows that the music box is reasonably well-tuned, but there are some imperfections in the tuning of both the melody and bass notes. These deviations could be due to various factors, such as the mechanical nature of the music box or slight variations in the manufacturing of the comb teeth. Further analysis could involve comparing these results with other music boxes or exploring methods to improve the tuning accuracy.

Chapter 6

Conclusions

The analysis of the music box recording using FFT has successfully identified the fundamental frequencies and their harmonics. The melody and bass notes were extracted, and their tuning accuracy was evaluated by calculating the deviation from ideal frequencies in cents.

Overall, the results indicate that the music box is reasonably well-tuned, with most notes falling within acceptable limits of deviation. However, some outliers were identified, particularly in the higher register, which may be more susceptible to manufacturing imperfections. Further investigation into the mechanical properties of the music box could provide insights into these tuning discrepancies and inform potential improvements in design or construction.

Chapter 7

References

The References chapter should name **all** literature that has been used in the assignment. If you have used an equation from a book or a standard *or the assignment text*, make sure to cite it. To cite a reference, use square brackets, as in [1]. There are many good ways to list your sources, pick one bibliography style and stick with it for all the sources in your report.

Literature references usually look like this (see next page): [Ref.nr.] author, title, publisher, place, year, pages: [1] Nilsson, James W., Electric circuits, Fourth edition, Addison Wesley, New York, 1993, 923 p.

Bibliography

[1] Nilsson, James W., Electric circuits, Fourth edition, Addison Wesley, New York, 1993, 923 p.

Appendix A

Appendix

The appendix should contain observations, calculations and results that are too long to be included in the previous parts. Data sheets, source code, and things like that can also be included here.

The appendix often includes extra material which works as a compliment to the report for those who want deeper insights into derivations, assumptions made and for those who want to look closer at the results. Raw data might in some cases be possible to include, and might be valuable for the reader.

Appendix B

Appendix

If you have data from different measurements, you might want to use multiple appendices.