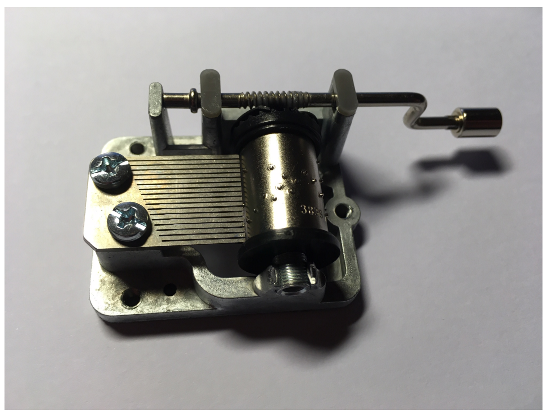
TTT4295 - Acoustic Signal Processing
Assignment 1 - Analyze the music box as a sound source

Peter Svensson

September 2, 2025



All the tasks are mandatory and should be part of your report. The deadline is specified online.

Introduction

You should make a recording of the sound from a music box which you can borrow in the acoustics lab. The music box might be mounted in a little box, or on a metal sheet. The kind of mounting can strongly affect how loud the music box plays, but it doesn't change what we are interested here: the frequencies of the generated tones.

The music box has a small "steel comb", with a number of thin short teeth, as its sound generation device. Each tooth vibrates with its own fundamental frequency and overtones, which should generate one musical tone. **Your task is to identify those tones, and inspect their spectra. Through this, you should learn about frequency resolution, harmonic relationships, and musical scales.** You can use Matlab or any other programming language of your preference for analyzing the signal in the sound recording. *You can not use any audio signal analyzing software - you must implement the spectrum calculations yourself.*

The music box as a sound source

As you can see when you use the music box, little pegs on the rotating drum push one or several teeth and then let them go. This induces quite strong vibrations in the tooth of the comb, with a number of resonance frequencies which will be present in the vibration signal. Mechanically, the tooth, or arm, functions like a *cantilever beam* which is fixed at one end and free at the other end. When vibrations are generated, so-called bending waves will be induced in the beam, and travel back and forth. However, if we look up what the resonance frequencies are for a homogeneous cantilever beam with a rectangular cross-section, we find that they do not have a harmonic relationship. This is typical for bending waves, since the wave speed is not constant for all frequencies (that is, bending waves are dispersive whereas transversal waves and longitudinal waves in air, fluids and solid media are non-dispersive). It turns out that the two first overtones are around 6.26 times, and 17.54 times, the fundamental frequency. This is not very attractive for a music instrument, but the instrument makers know how to improve this. By shaping the arms/teeth so that they are not homogeneous beams, then the resonance frequencies can be modified!

Figure 1 shows the music box seen from underneath. You might be able to see that at the right end of the teeth it looks like a tilted line. This is an extra thick part of the tooth towards the end. Furthermore, there are little modifications of each tooth at the left end as well.

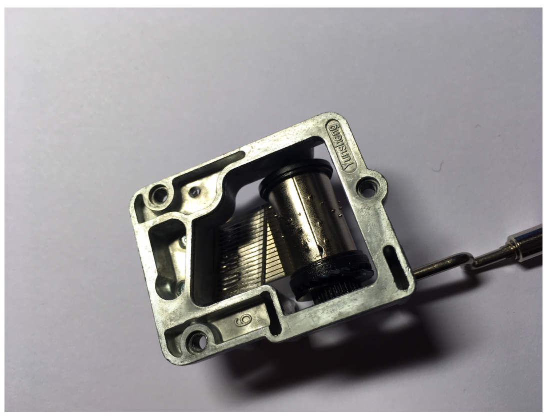


Figure 1: The music box seen from underneath.

Tasks

1. **Make a recording of the sound from the music box**
   You can use your smartphone as recording device, the quality of the microphone is not crucial. It is better if you can record it in an uncompressed format, but even a compressed-format recording will work. You should record all of the tones of the music piece.

   Hints:

   - Play the music box very slowly so that each tone dies out before the next. You can even play one tone and stop while the tone dies out. Figure 2 shows the recorded waveform, indicating two tone onsets and decays.

- You can use the Matlab function `audioread` to read several different sound file formats.

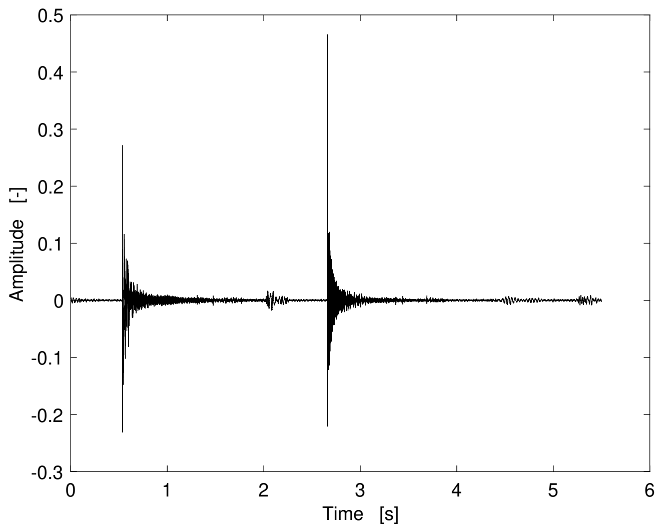


Figure 2: The recorded time signal when the music box is played very slowly, showing two tone onsets and decays.

2. **Cut the recording into individual tones**
   Cut the recording into individual tones and store them such that you can analyze them easily. Inspect the music box and check which tongue number is played for each individual tone. **Do this for all the tones of the entire music piece of the music box.** Please note that sometimes, chords are played, i.e., several tones are played at the same time. **Again: play the music box so slowly that the sound of each note has died out.**.

3. **Analyze the spectra for each individual tone**
   Take an FFT of each individual tone and identify a number of peaks. **Use the same FFT size for all notes. This might require zero-padding for most notes.** Then carry out the following analysis:

   - List the spectral peaks that you have identified.
   - What is the uncertainty in your frequency estimates?

     Hint 1: The frequency resolution of the FFT determines this.

     Hint 2: You can get a higher frequency resolution by using zero-padding when you take an FFT of the time signal.

Hint 3: These signals are of transient nature, and then you should *not* apply a window, at least not one which suppresses the start of the signal.

- Divide the found peaks into two groups: (A) (close to) harmonically related peaks, and (B) the rest of the peaks.
- How much do the spectral peaks in group A deviate from a perfect harmonic relationship?
- What are the relative levels of the harmonically related spectral peaks (group A), expressed in dB relative to the fundamental?
- **In your report you should include at least one example which shows the time signal, and the part marked which is used for the FFT/DFT.**

4. **Which musical notes are played?**
Using the fundamental frequencies from task 3 (in the group A for each tone), identify which musical notes are being played? Use the scale for equal temperament, described in the Appendix.

Also calculate the difference between the detected fundamental frequencies and those for an equal-temperament scale. **Express the differences in cents**.

5. **To remember in the report**

Your report should follow the template from blackboard. Remember to include the following in your report:

- A description of how you did the recording, including which hardware and file format you used (smartphone, laptop, something else).
- An example plot of a recorded waveform which convinces the reader that each tone had died out before the next tone started sounding. The plot should also show which part was used for the FFT.
- An example plot of a spectrum, with a frequency axis, which convinces the reader that the signals you analyzed were reasonable. The detected peaks should be marked.
- The source code (Matlab, Python etc) for your analysis, with the name of the programmer in the source code, and a date.
- *Remember to give references for text or images that you are taking from elsewhere!*

Appendix - The well-tempered musical scale

The well-tempered musical scale divides the frequency range of one octave into 12 semitones, *and makes the 12 intervals equal on a logarithmic axis.* Using a reference pitch of 440 Hz for the so-called A_4 semitone, the fundamental frequencies of the 12 semitones in octave number 4 are given as

$$f_n = 440 \cdot 2^{(n-9)/12} \text{ Hz} \tag{1}$$

The 12 semitones are given in Table 1, and by multiplying the frequencies by 2 or 1/2, 4 or 1/4, 8 or 1/8 etc, the fundamental frequencies result for one octave above or below, two octaves above or below, three octaves above or below, respectively.

The interval between two frequencies can be expressed in cents, c, defined as

$$c = 1200 \cdot \log_2 \frac{f_2}{f_1} = \frac{1200}{\log_{10}(2)} \log_{10} \frac{f_2}{f_1} \approx 3986,31 \log_{10} \frac{f_2}{f_1} \text{ cents} \quad (2)$$

This definition makes the interval between semitones in the well-tempered scale exactly 100 cents.

n	Note	$f_{\text{fundamental},n}$ [Hz]
0	C_4	261,63
1	$C\sharp_4$	277,18
2	D_4	293,66
3	$D\sharp_4$	311,13
4	E_4	329,63
5	F_4	349,23
6	$F\sharp_4$	369,99
7	G_4	392,00
8	$G\sharp_4$	415,30
9	A_4	440,00
10	$A\sharp_4$	466,16
11	H_4	493,88

Table 1: The fundamental frequencies of the 12 semitones in the 4th octave, for the well-tempered musical scale.