Multipitch detection

Daoud Hasnaa, Lahlali Kenza

January 28, 2019

1 Windowing and Fourier transform

- 1) The choice of the window size must be done considering the frequency of the signal. If these factors evolve, this must be taken into account. The lowest detectable frequency (F0) is determined by the size – duration – of the window.F_0= 5 $\frac{SamplingRate}{Windowsize}$
  In this case, we have : $5\frac{22400}{700} = 160$ Hz is the lowest detectable frequency, which means that we will be able to detect notes with highest frequencies.

- 2) The offset is the chosen delay from which we consider the signal, it is chosen to be about 100 ms. It's main purpose is to discard the attack of the sound.

- 3) The frequency precision is given by $dF_{min} = \frac{F_s}{Nfft}$, where $Nfft$ is the size of the FFT window. For the file "$A3_{piano}.wav$", The precision is 1.43 Hz.

2 Fundamental frequency estimation by the spectral sum method

2.1 Computation of the spectral sum

- 4) The maximum frequency that can be used for the computation of S is equal to $\frac{H*R*Fs}{Nfft}$
- 5) The maximum value of R that can be used for the computation of S is R_{max} for which $\frac{H*R*Fs}{Nfft}$ is equal to Nequist Frequency. So Rmax = $\frac{Nfft}{2H}$, we find $R_{max} = 4096$.
- 6) Below we code the function Spectral Prod and spectral sum. In this part we use spectral sum

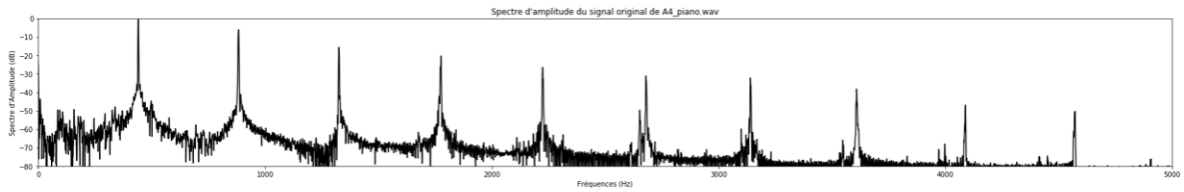
2.2 Finding the spectral sum maximum

We chose $F_{min} = 50$ and $F_{max} = 900$ and we find $N_{min} = 51$ and $N_{max} = 922$.

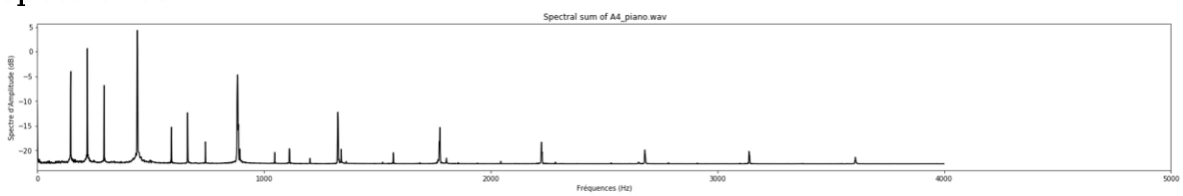
For each signal, we checked if the corresponding notes (ie A3, A4, etc.) correspond to the frequencies obtained by our algorithms. We checked the values in the website *http://www.contrabass.com/pages/frequency.html* to compare them to our values.

2.2.1 A4-piano

Original signal

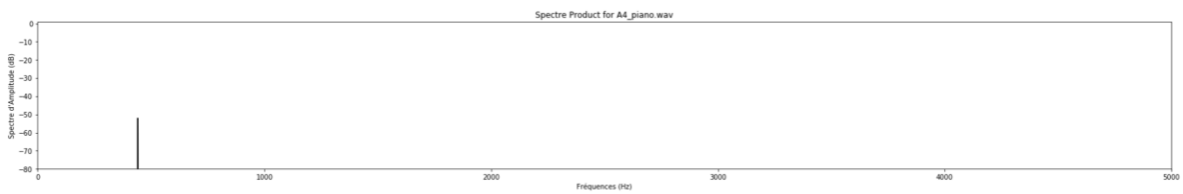


Spectral sum



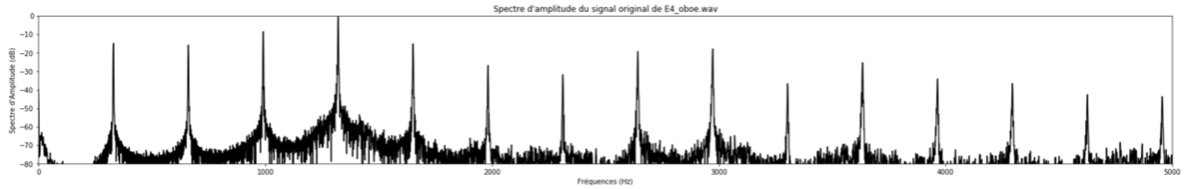
The maximal amplitude is obtained with frequency of 441.40625 Hz (fundamental frequency) and the amplitude is 1.639

Spectral Product

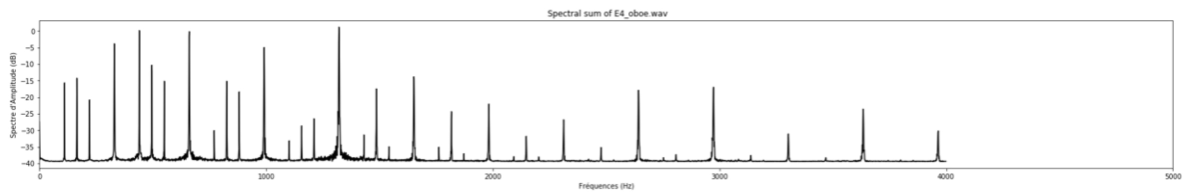


2.2.2 E4-oboe

Original signal
The figures below present the results of the method tested on *E4_oboe.wav* file.
This figure presents the signal's spectrum. We can notice that it is noisy.



Below is the spectral sum amplitudes **Spectral sum**



We can notice that the fundamental frequency is between 300 Hz and 400 Hz. Yet, the sum product detected a maximum amplitude with frequency of 440.4296875. Hence, we can conclude that the spectral sum is not very discriminating because of the noise.
We tried with the spectral product and we could detect the right frequency, the maximal amplitude is obtained with frequency of 330.078125 Hz. Thus, the Spectral Product is more accurate.

Remark : The Spectral Sum is less accurate than the Spectral Product. The Spectral Sum is less discriminative mainly because of the noise that makes the computation less accurate.

3 Subtraction of the sound corresponding to the detected fundamental frequency

3.1 Harmonic detection and suppression and stopping criterion

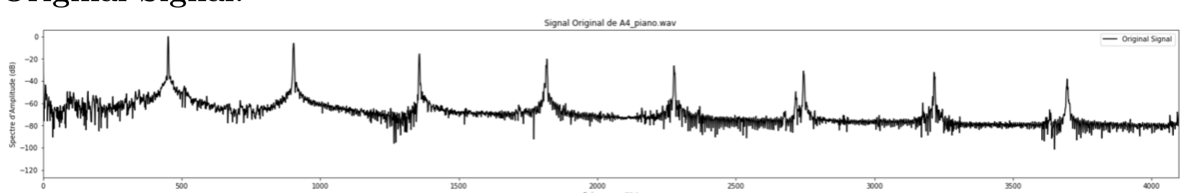
Empirically, we chose α equal to 0.02, but this α depend on the signal that we are studying. Indeed, the harmonics may be very close to each others.
Stopping Criterion Principle: As a stopping criterion, we chose to stop the deletion of the notes when the amplitude of the spectral product for a given fundamental frequency becomes negligible compared to the first detected maximum amplitude.
For this, it was necessary to set a threshold.
We chose a threshold equal to 0.0001 . the results were satisfactory for the signal *A3C4E4G4_piano* and *A4_piano*. This method didn't work well with sum product.

3.2 Stopping criterion

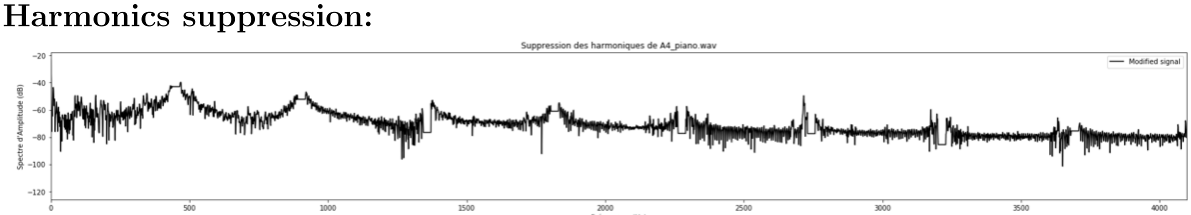
With the criterion above, we could stop the suppression from the first iteration for the signal *A4_piano*

3.2.1 A4piano

Original Signal:

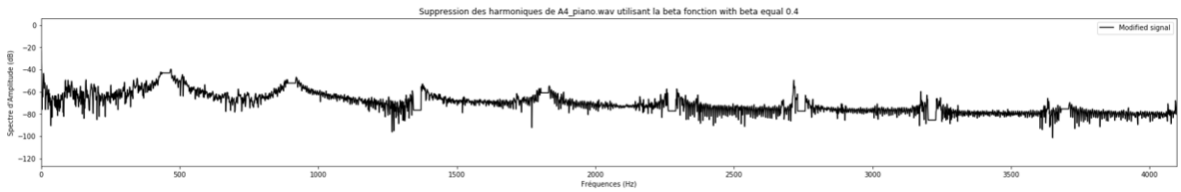


Harmonics suppression:



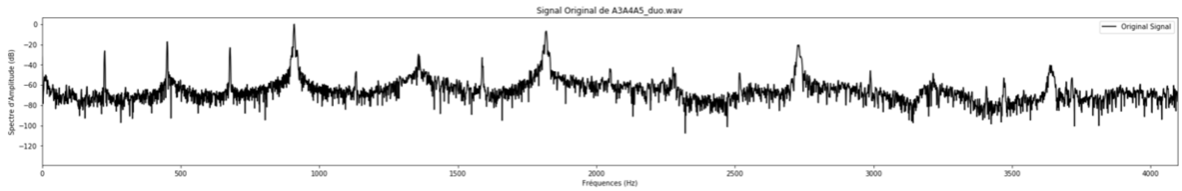
Stopping criterion:

With these below methods, we got a fundamental frequency of 441.4 Hz, and indeed, A4 corresponds to 432.00 Hz.
Taking Account of Inharmonicity with beta equal 0.4



3.2.2 A3A4A5duo

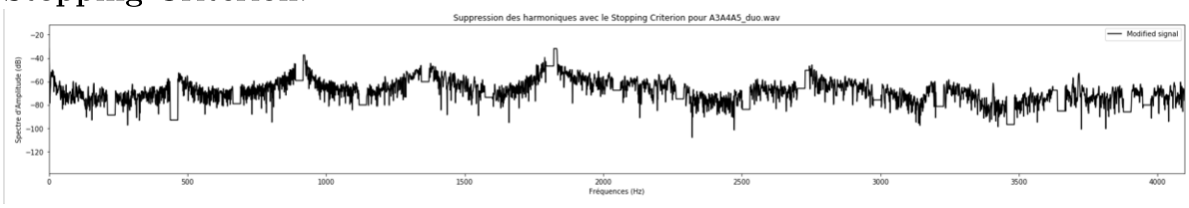
Original Signal:



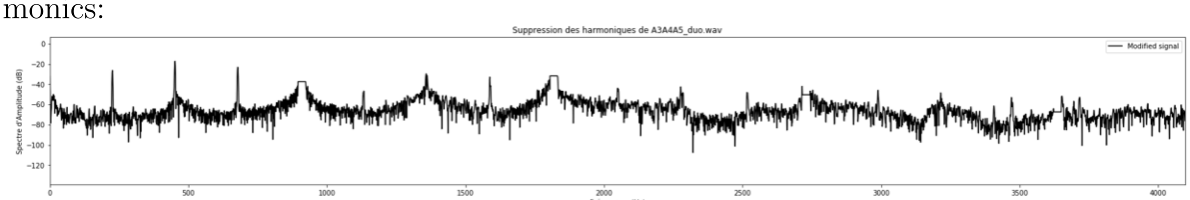
Harmonics suppression:

```
suppression of the frequency 888.671875 Hz and its harmonics
suppression of the frequency 220.703125 Hz and its harmonics
```

Stopping Criterion:



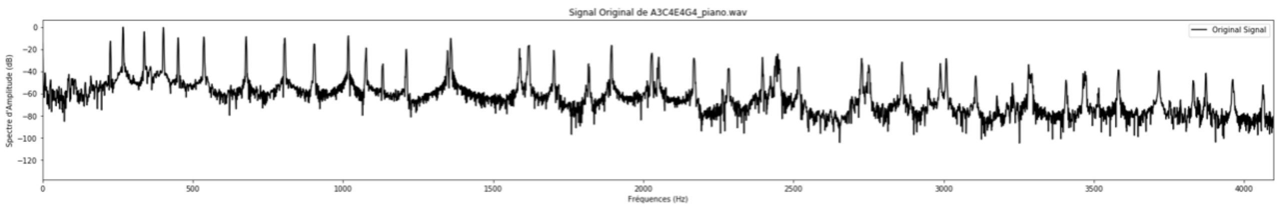
The algorithm suppressed the frequencies 888.67 and 220.70 Hz and their harmonics:



We can notice that: A3 corresponds to 216.00 HZ, A4 to 432.00 HZ and A5 to 864.00 Hz. Which is in line with the suppressed frequencies as their are multiples.

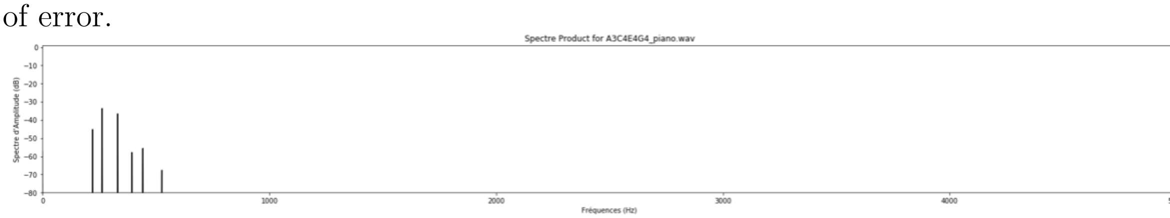
3.2.3 A3E4C4G4

The original signal:



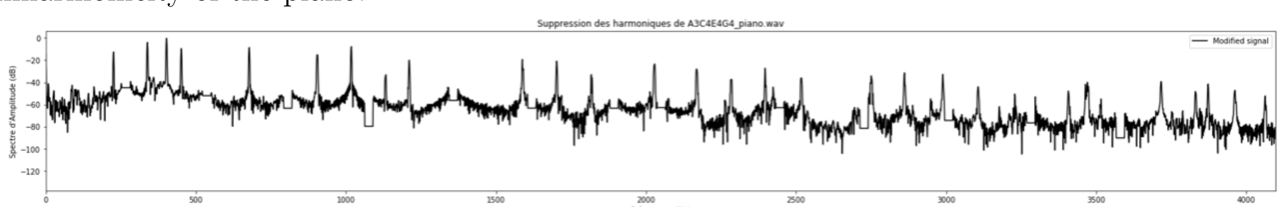
The Spectral Product

We can see that with the spectral product we could have several peaks, but not as many as those with the spectral sum, which decreases the possibility of error.



Harmonics Suppression:

Since all of fundamental frequencies are multiples of the fundamental frequency of A3, we can predict that we will detect one frequency and delete all of its harmonics. This way we can delete all of the signal harmonics. However, we can see in the figure below that the algorithm detected 2 frequencies (the fundamental frequency of A3 220 Hz, and the frequency 888 Hz) . The frequency 888 Hz corresponds to the note A5 (880 Hz), which is due to the inharmonicity of the piano.



Stopping Criterion:

```
suppression of the frequency 261.71875 Hz and its harmonics
suppression of the frequency 331.0546875 Hz and its harmonics
suppression of the frequency 441.40625 Hz and its harmonics
suppression of the frequency 392.578125 Hz and its harmonics
```

When we compare the previous frequencies with the notes A3 (ie 220 Hz), E4 (ie 329.63Hz), C4 (ie 261.63 Hz), G4 (392 Hz). We notice 3 of our frequencies correspond to C4, E4 and G4, nonetheless our algorithm detect 441.40 Hz instead of 220 Hz, which is a Harmonic of A3.

4 Some refinements or options

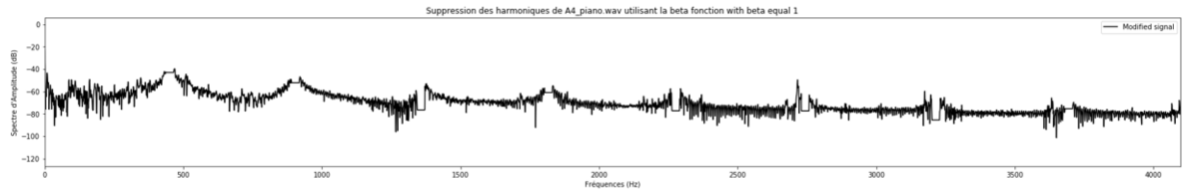
4.1 A different F0 detector

We adapted the previous previous **Harmonic Supression** to use the Spectral "Product" instead of the Spectral "Sum". The results are much better with the Spectral Product then with the Spectral Sum. The graphs are in the previous section they correspond to the graph "Harmonic suppression".

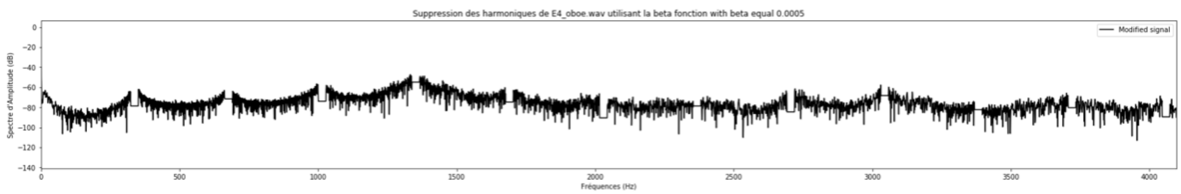
4.2 Taking account of inharmonicity

We modified "Detect harmonic" algorithm adding a term $\sqrt{1 + k^2 * \beta}$. We used $\beta = 1$ for $A4 - piano$ and for $E4 - eboe$ we get a β equal to 0.0005.

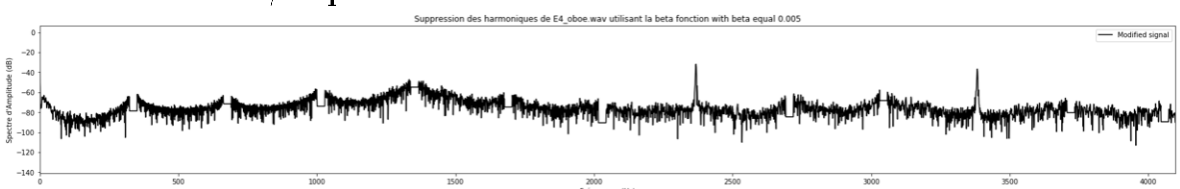
For A4piano



For E4oboe with β equal 0.0005



For E4oboe with β equal 0.005



We notice that a slide difference in β value for E4oboe has a big impact on the recognition of the frequencies which is not the case for the A4 piano. We realised that taking into consideration a variable treshold in our previous algorithms enabled us to counter the problem of inharmonicity.