ÉCOLE CENTRALE LYON

UE ELC F-4
TP - FREQUENCY ANALYSIS
FINAL REPORT

Frequency analysis of musical sounds

Students :
Aurélien VU NGOC
Julien GUINOT

Prof. :
Sebastien OLLIVIER

October 21, 2022

Contents

List of Figures

Introduction

This report presents the results obtained by this group during their work on the practicum of the musical acoustics course. In a first part, we will focus on additive synthesis of simple harmonic sounds, both decaying and non-decaying, as well as the influence of several parameters on the signal-processing and reverse-engineering aspects of this generation. A second part with contain study results regarding time-frequency analysis of various musical instruments, as well as characterzing them through various lenses. In the last section, we will attempt, based on the reverse-engineering methods set up in the first section, and the analysis of the musical instruments done in the second one, to re-synthesize these sounds by additive synthesis.

1 Frequency analysis of a synthetic sound

The objective of this section is to synthesize simple harmonic sounds based on the additive synthesis method, which will be explained in the following paragraph. One main obejctive is to be able to reverse-engineer the frequencies of a given, known synthesized signal. This will be useful in re-synthesizing recorded sounds afterwards.
A second paragraph will address the synthesis of a damped signal, as well as its time-frequency analysis, we will also see if the previous method for reverse-engineering the signal still works with time-damped signals.

1.1 Non-damped signal

1.1.1 Synthesis and Analysis

The first objective of this section is to synthethize a non-damped signal by additive synthesis, as seen in lectures. Assume a simple harmonic signal of n frequencies F_i and of amplitudes A_i. Then, the signal can be expressed as:

$$s(t) = \sum_{i=1}^{n} A_i sin(2\pi F_i t) \tag{1}$$

To avoid signal clipping during generation, we normalize harmonic amplitudes to a maximum of 0.5 prior to synthesis. This can be seen in the matlab code. We start by synthesizing a harmonic signal of the A note ($la440Hz$), with frequency composition as read from bibliographies. The list of frequencies used as input is $(218.9Hz, 882.7Hz, 2195Hz, 3797Hz, 7898Hz$ and amplitudes are $(65dB, 30dB, 50dB, 35dB, 50dB)$. A quick reminder for the conversion of dB amplitudes to abolute amplitudes is shown below, as it will be important for evaluation of the reverse-engineering precision:

$$A_{dB} = 20log_{10}(A) \tag{2}$$

Firstly, we analyse the synthesized signal using a standard sampling rate (N_{fft}) of 44100, no window, and a signal of length $1s$. We will later analyse the influence of these parameters on spectrum analysis. The time signal as well as spectrum of the obtained signal are shown below (*Figure 1*)

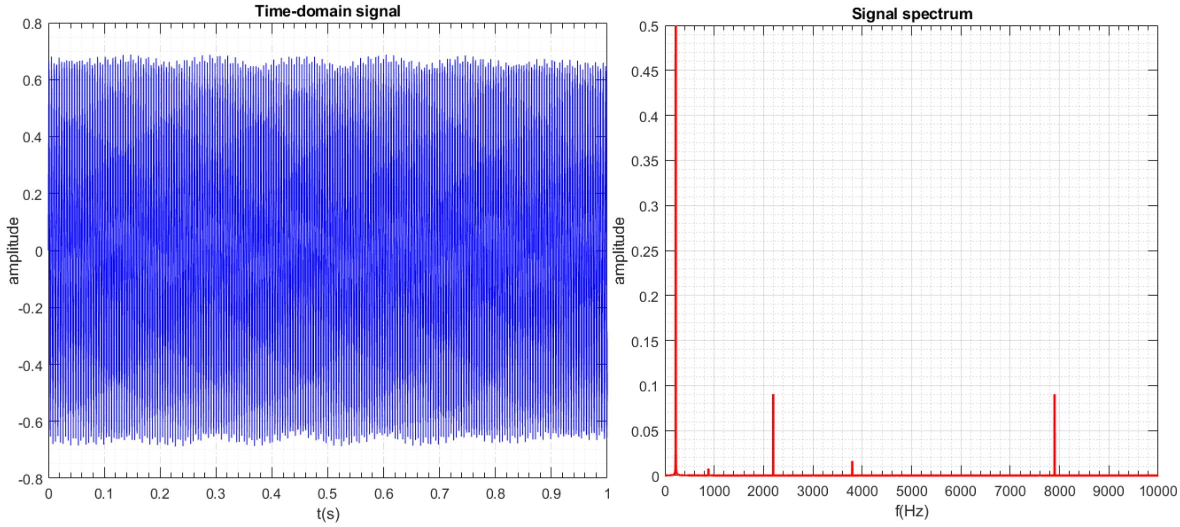

Figure 1: Spectrum and time-signal of non-damped synthesized signal

The frequencies found in the spectrum of the non-damped signal seem to correspond to the frequencies used as input for its generation. we will see later via reverse-engineering to which extent these frequencies are precise.

It is also interesting to visualize the evolution of the amplitude over time, which can be done via a spectrogram. The time-frequency spectrogram for the synthesized signal is shown below (*Figure 2*). it uses a hanning window (2048 samples), as well as a 1024-sample overlap, and $N_{fft} = 441000$. The influence of these parameters will be discussed later, but they were chosen in this case to provide the best compromise of time-frequency resolution.

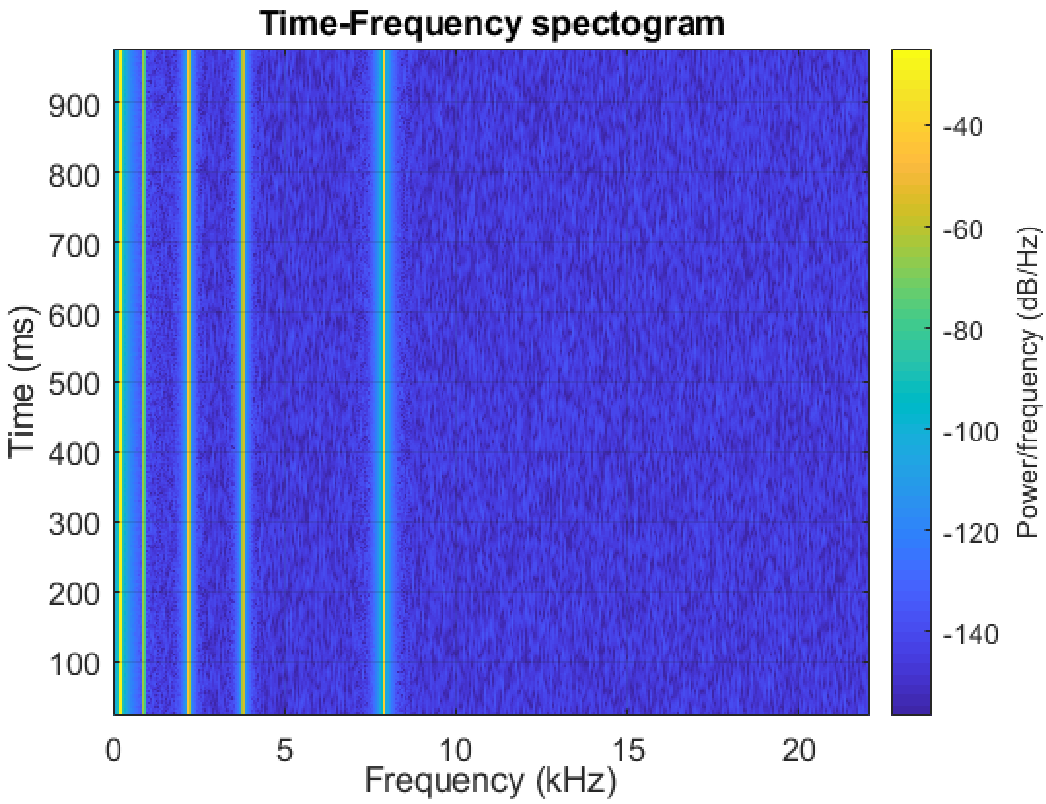

Figure 2: Time-frequency spectrogram of non-damped signal

It can be seen that the frequency contribution of every harmonic is constant over time, which is coherent with regards to the synthesis method of the signal. We will now attempt to reverse-engineer the signal and find the frequencies and amplitudes of each harmonic contribution. We will also analyze the influence of parameters of the FFT.

1.1.2 Reverse-engineering frequencies and amplitudes

We will first attempt to find the frequencies and amplitudes of each sinusoid for the previous spectrum, without changing any of the FFT parameters. we do this by finding the local maxima of the spectrum, and associating each frequency to its peak value. Then, we find the amplitudes by tusing *Equation 2* after un-normalizing the amplitudes. The original and reverse-engineered amplitudes and frequencies, as well as the relative error between the two is shown in the following table (*Table 1*):

	1	2	3	4	5
Given frequency (Hz)	218.9	882.7	2195	3797	7898
REd frequency (Hz)	220	884	2196	3798	7899
Error margin (%)	0.5025	0.1473	0.0456	0.0263	0.0127
Given amplitude	0.0355	0.00063	0.00632	0.0011	0.0063
REd amplitude	0.0355	0.00055	0.00643	0.0011	0.00642
Error margin (%)	0	11.96	1.69	1.8	1.65

Table 1: Reverse-engineered values for basic

As an arbitrary frequency precision criterion, we will require that reverse-engineered frequencies be at least within an 8^{th} tone range from the original frequency. Given that $\boxed{2^{1/48} = 1.0145}$, this represents a 1.45% error margin, which, as is shown in the previous table, is respected for the current analysis parameters for all frequencies. Regarding amplitudes, we attach less value to precision, only orders of magnitude must be respected. The first thing to notice is that Reverse-engineered frequency values are all integers. This is a coincidence caused by the fact that the length of the signal is 1s, sampling frequency is 44100Hz, and $N_{fft} = 44100$. This will change in the following paragraph. Thus, the reverse-engineering procedure works for this first case.

As a side note, consider the low-range spectrum, between 20 and 160Hz. Then, if we want 8th-tone precision for this range, the smallest 8th-tone interval being that of 20Hz, we need the distance between two consecutive fft points to be smaller than the 8th-tone interval. assuming that we analyze the totality of the human hearing range (20hz-20kHz), this amounts to the critical value of:

$$N_{fft} = \frac{20000 - 20}{20 * 2^{1/48} - 20} = 68602 \tag{3}$$

However, results show that at 20Hz, precision with this metric is only of 2.1%, which does not correspond to the given frequency precision metric. This is possibly due to our method of reverse-engineering frequencies in finding local maxima. However, starting at 25Hz, the precision criterion is met, so, as we continue to work on the project, we will assume this to be adequate until we come back to it later

1.1.3 Influence of analysis parameters on precision

We will now look at the precision of this method in the case of variations N_{fft} and the use of various windows. We expect that, at constant sampling frequency, for lower N_{fft}, the precision of the spectrum will be lower. Error margin results for varying N_{fft}, and windows for all other parameters constant, are shown in the following tables:

	1	2	3	4	5
Error margin (N=1024)	18.04	7.33	2.02	0.94	0.33
Error margin (N=8182)	3.417	0.75	0.18	0.07	0.045
Error margin (N=44100)	0.50	0.14	0.04	0.02	0.01

Table 2: Error margins for varying N_{fft}

As expected, as N_{fft} increases, the precision of the measure increases, and it is possible to discern frequencies with great precision (almost 16th tone precision). For the sake of analysis, we will use $N_fft = 88200$, which is a value superior to the critical value computed before, and $N_{fft} = 2048$ for parameter analysis, to be able to notice error margin variation.

For a non-decaying generated signal, neither the window used nor the length of the signal influence the precision of the reverse-engineered frequencies. It can be assumed that this is because of the simple nature of the signal. for now, we will consider that the ideal parameters are the use of a Hanning window for maximal tone separation, $N_{fft} = 44100$, and the longest signal possible.

For this part, we are able to reverse-engineer the frequencies and amplitudes of the generated signals with satisfying precision.

	1	2	3	4	5
Error margin (no window) %	8.20	2.45	1.04	0.37	0.33
Error margin (hanning window) %	8.20	2.45	1.04	0.37	0.33
Error margin (blackman window) %	8.20	2.45	1.04	0.37	0.33
Error margin (flat-top window) %	8.20	2.45	1.04	0.37	0.33

Table 3: Error margins for various windows

1.2 Damped signal

We will now study the process of generating a damped signal. We take the expression of the signal in *Equation 1* and add damping for each individual harmonic. The new signal can be expressed as:

$$s(t) = \sum_{i=1}^{n} A_i sin(2\pi F_i t)e^{-\frac{t}{\tau_i}} \tag{4}$$

In our case, we test-generate the same signal as before, with damped harmonics corresponding to a plucked-string sound. This means fast attack, low sustain, and intermediate release times. We can then choose a mean $\tau = 0.8s$ value for a 1.5s length signal, with a random seed to lengthen or shorten this attenuation time according to the harmonic, which seems to produce a realistic simulation of either a high A note on a guitar or an A note on a xylophone. The time signal, frequency spectrum and spectrogram using the same parameters as in the previous section are plotted in the following *Figure 3*:

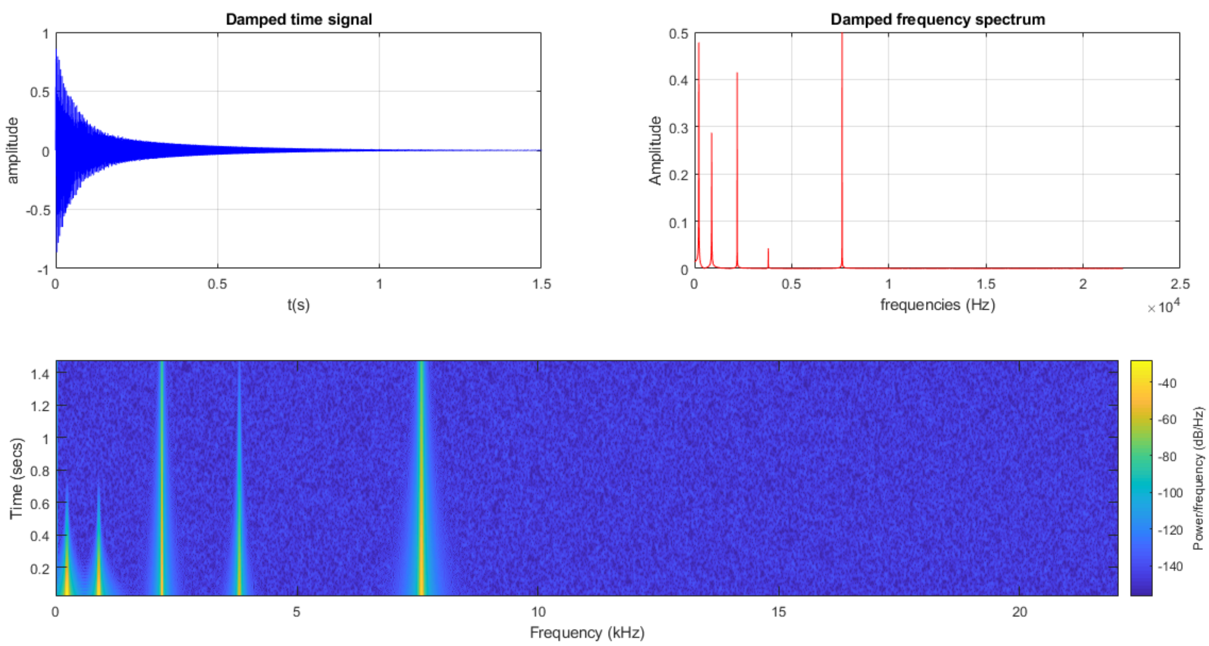

Figure 3: Synthethised damped signal characteristics

We can see that the various harmonics of the signal are damped at varying speeds, given the random aspect of decay time generation, and the auditive aspects of the signal are now much more realistic. We can also see that introducing dampening to the signal has made the harmonics of the spectrum quite a bit larger, and thus might influence our ability to reverse-engineer frequencies and amplitudes.

Now that we have generated a dampened sound, we wish to analyze the influence of the parameters of the spectrogram on frequency and time resolution. We will now be varying the following parameters to observe their influence : Window, window sample size, chunk size, and overlap size. The objective is to attain a satisfying compromise between time resolution and frequency resolution. We will begin by analyzing window types, with four windows: Flat-top, Hanning, Hamming, And Blackman, all with the same sample size (1024). The results are shown below

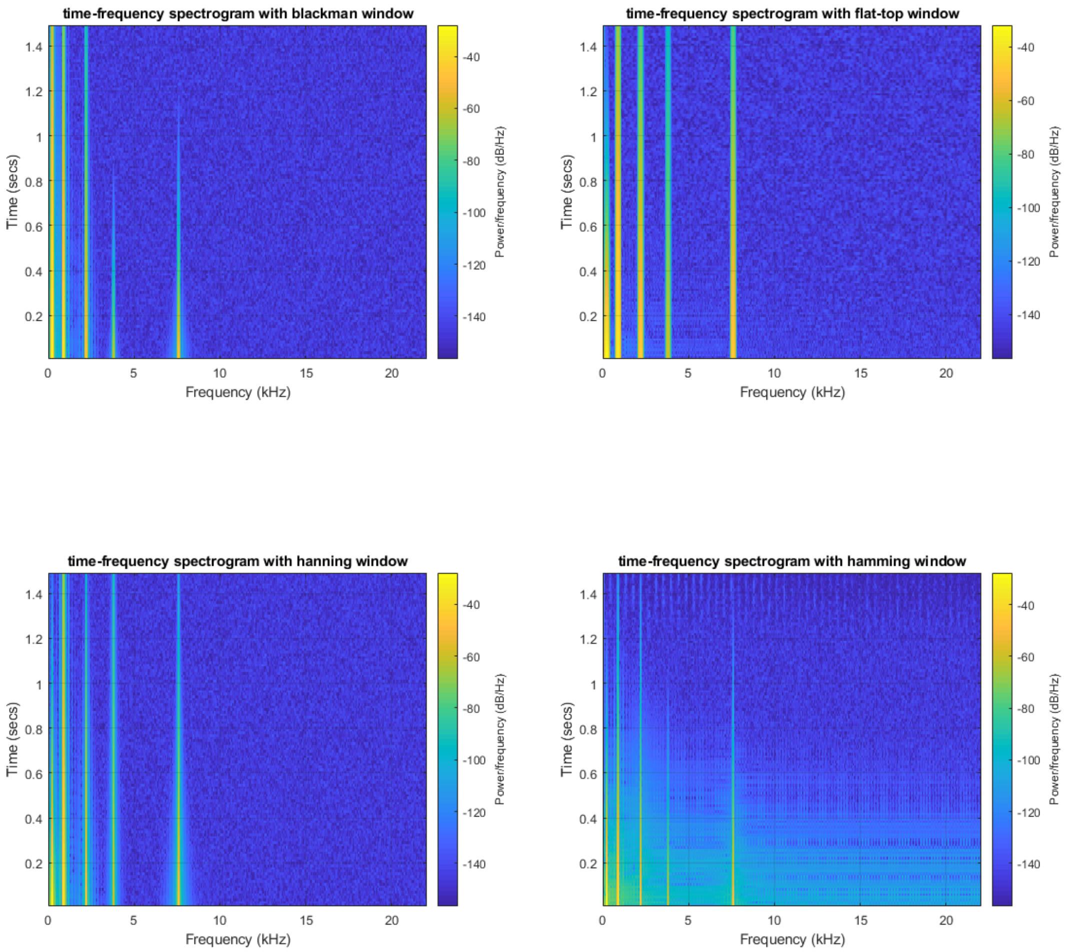

Figure 4: Blackman, Flat-top, Hanning and Hamming windows

We notice that the windows that seem to provide optimal frequency resolution for the same sample size are the Hamming window and the Blackman window. the Hanning and Flat-top window provide higher time-resolution, but in our case we are mostly interested in determining frequencies with a high degree of precision. Given the fact that the hamming window offers less contrast, and that it is thus harder to determine the exact frequency of the harmonic, we will continue our study using the blackman window. we will now study the influence of the sample size used by the blackman window on resolution, by analyzing the spectrograms of four sample-size Blackman windows (512,1024,2048,4096), at constant overlap size (256):

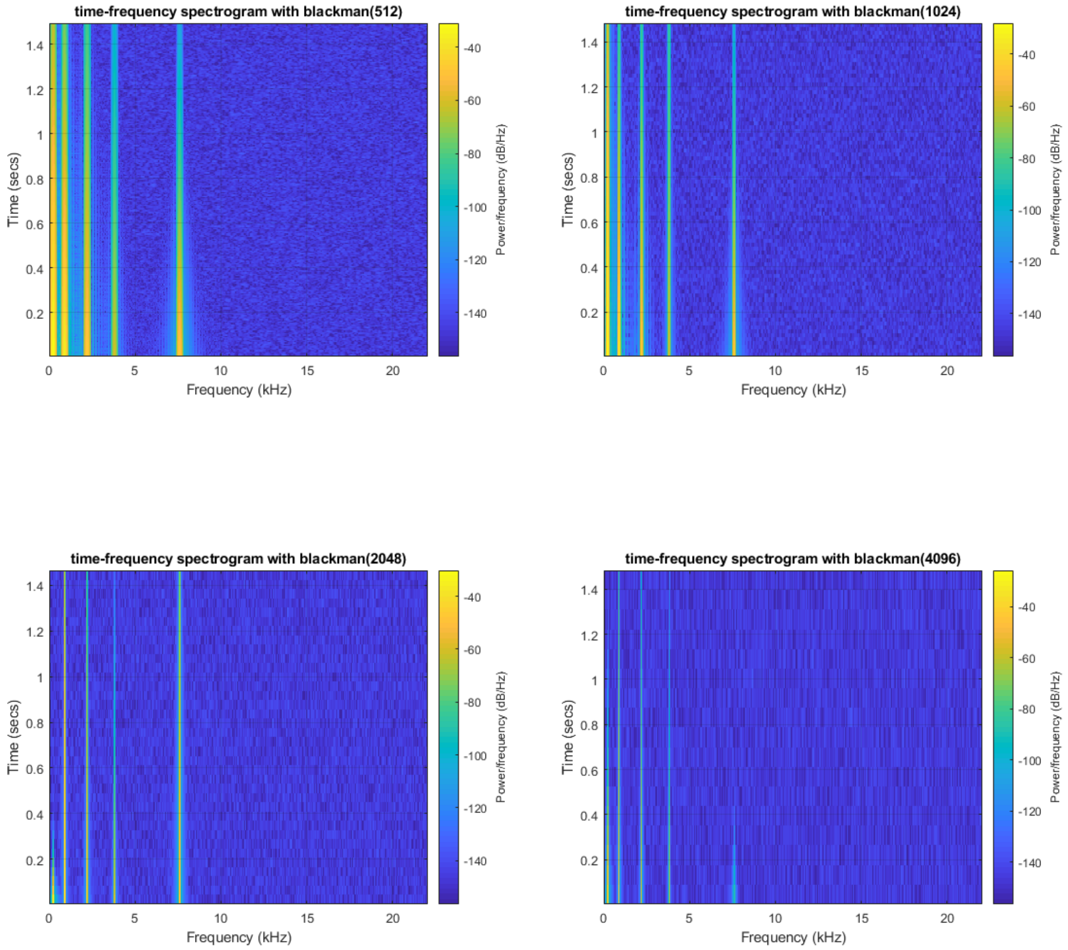

Figure 5: Blackman window with varying sample size

As window sample size goes up, frequency resolution also goes up. However, it can also be seen that for each increment of sample size, time resolution also goes down, to the point that the Blackman(4096) window becomes quite blocky. This, we will prever a higher sample size, without going too overboard. Perhaps one of the following parameters can counter-act this blockiness, allowing us to continue with a higher sample size. We will now study overlap sample size. Indeed, a better way to evaluate frequency placement is to overlap samples, which provides a way to cross-compare information bewteen samples and obtain a better estimation for eache frequency. The following spectrograms show the resulsts obtained for a blackman (2048) window, for varying overlap sizes (128,256,512,1024):

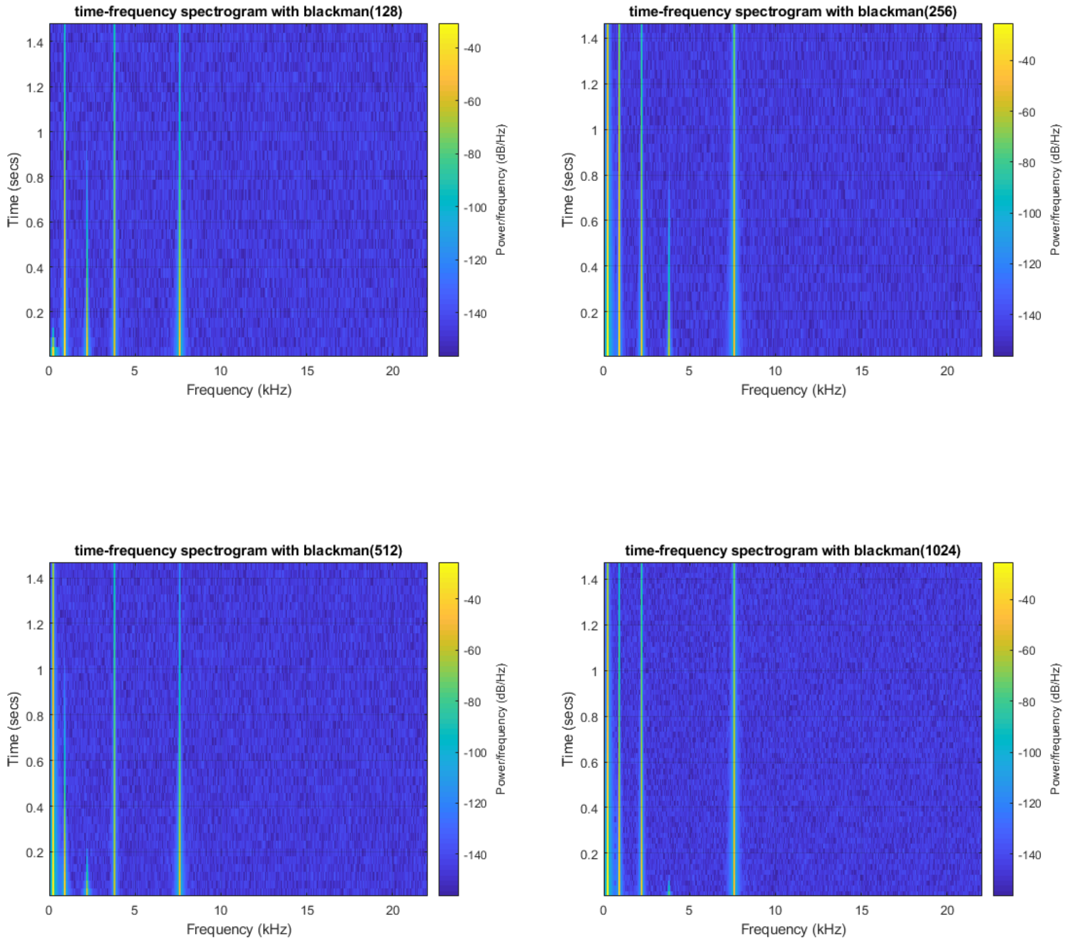

Figure 6: Blackman window with varying overlap

The main effect of overlapping samples for frequential analysis is better frequency localization, but we also notice that upping the overlap size allows for better time localization. this is a desirable effect, given that it allows for globally more resolution relative to time and frequency, which is a highly desirable quality for the application we are seeking. for the rest of this stufy, we will apply a half-chunk overlap preset, which is most desirable relative to calculation speed (as we are only using powers of 2). It is expected that the amount of DFT points for each chunk will only better the frequency resolution of each chunk, at the expense of time resolution. As we we believe these settings to be optimal for basic analysis, though we may need to adapt later on, we will now assume spectrograms calculated with Blackman(4096) windows of 2048-size overlap, and 16384-DFT points for each chunk.

2 Frequency analysis of recorded instruments

We now focus on analyzing sound recordings from various instruments, with the objective to analyze the obtained spectrograms with regard to instrument physics, acoustics, perception, timbre, frequency and signal processing. The goal being to cover a broad range of frequencies, a mix of recordings from our own personal instruments and those recorded from the musical both of us took part in this year, both winds, strings (plucked and rubbed), and percussions.

2.1 String instruments

The first sound we analyse is a recording of one of us playing "here comes the sun" by the beatles. The isolated note is supposedly a D. the following *Figure 7* shows both the spectrogram and the frequency spectrum obtained by analyzing the sound file:

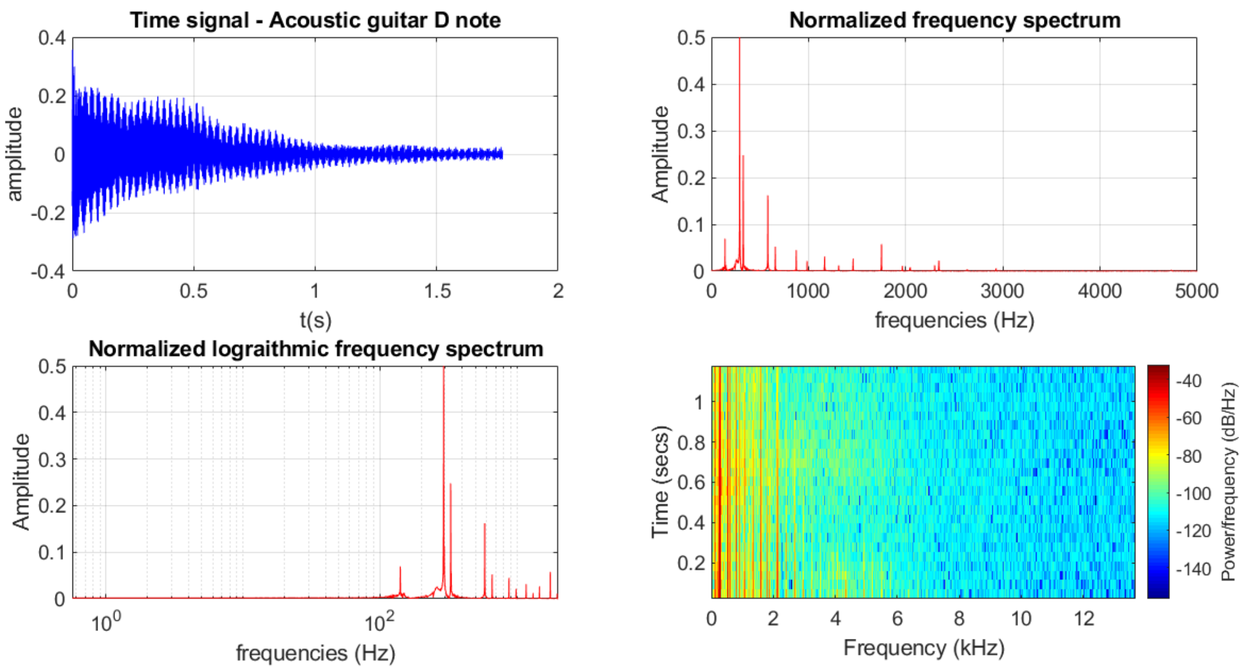

Figure 7: Characteristics of D-note played on a guitar

We will now discuss the results based on various aspects of the instruments :

- **Played note:** The first thing to notice is the strange distribution of the harmonics on the frequency spectrum. Indeed, in order, the reverse-engineered frequencies are *[140,6 291,4 331,5 581,7 657,9 874,2 1166,1 1751,2]*. Though the fundamental frequency of the played noted (D3=140Hz) can indeed be found in the spectrum, as well as the associated harmonics (291Hz, 581Hz and 874Hz), other harmonics can be observed, the fundamental being an open-string E (E4=329Hz). The slight difference in frequencies can be attributed to a badly-tuned guitar. A quick look at the score music for "here comes the sun" shows that the previously played note is indeed an open-string E. What can then be observed on the spectrum are residual harmonics from this note.

- **Instrument physics:** Firstly , the guitar is a plucked-string instrument. the theoretical model adopted for such excitations of string instruments shows that harmonics should be of the form $F_n = n * F_0$, where F_0 is the fundeamental frequency of the string *(*$F_0 = \sqrt{\frac{T}{\mu 4L^2}}$*, where* μ *is the lineic mass of the string and* T *is the applied tension).* The frequency spectrum allows to verify this, as the harmonics measured obey this law for the notes E4 and D4. Another thing to notice is the relatively long decaying times of each harmonic, which, in comparison to each other, seem to decay faster the higher the frequency. Physically, this can be explained by the fact that the main opposing force to the string's movement is air, the resistance of which is low relatively to the strings's tension. However, as the vibration frequency becomes higher, the string has to oppose that air more and more each second, which explains faster dissipation.

- **Perception and timbre:** The soundboard of the guitar, which allows for longer-decaying harmonics, as well as the low-frequency register of the played notes (this can be seen on the log-log scale of the frequency spectrum), and the openness of the played notes (even longer decaying notes, not slowed by finger contact), gives the guitar what can be called a characteristic warmth to its sound, which can be characterized by its low register and resounding openness. in terms of timbre, it can be called a 'warm' sound.

- **Signal processing**. In terms of signal processing and reverse-engineering frequencies and amplitudes, it can be noticed that the contrast relative to the previously-shown spectrograms, as well as the cleanliness of the frequency spectrums, is much less defined. This is normal, as real-life recording must come with background noise (white noise), as well as physical flaws in the instrument (a bit of frequency leakage in mid-tones at high times can be noticed, perhaps because of the slowing of the strings which results in minot pitching after a while. The issue of background noise can be somewhat resolved by applying a de-noiser in audacity or a wiener filter in matlab.

We now analyze a piano playing a C note. contrary to guitars, which are plucked, the strings in a piano are hammered, which is a much more punctual excitation, which might lead to less harmonics:

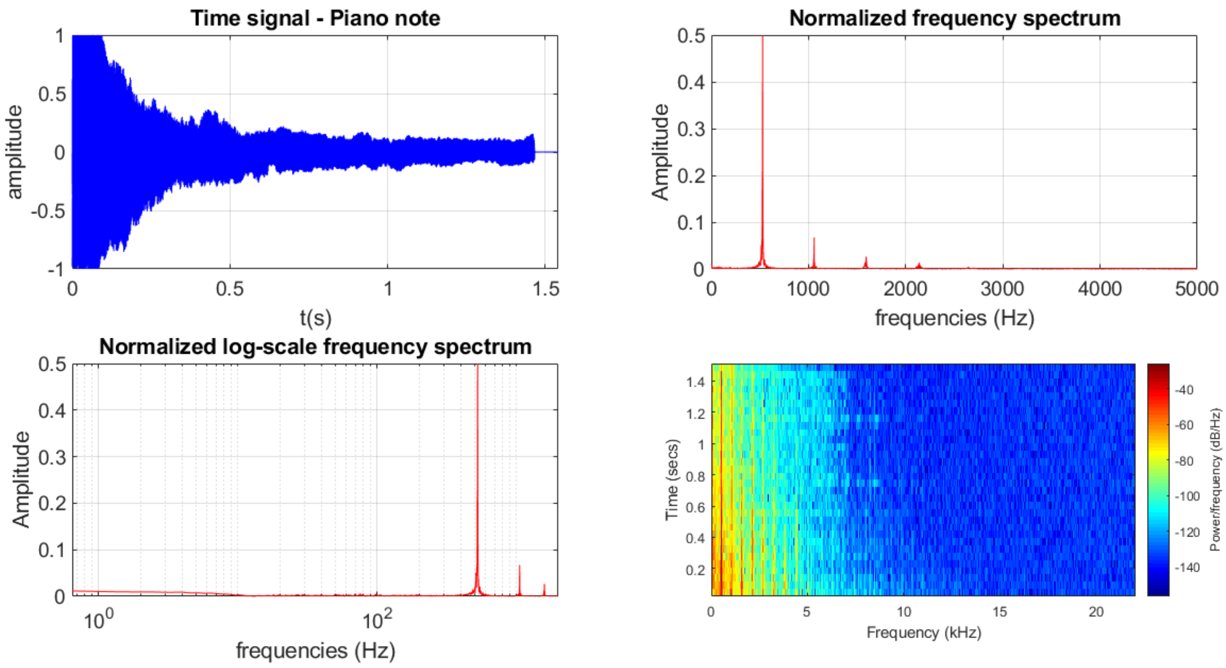

Figure 8: Characteristics of C-note played on a piano

- **Played note:** The note played here is a C (C4=523.25Hz). Indeed, the resulting harmonics as well as the fundemental (reverse-engineered value: 520.6Hz) can all be observed on the spectrum, with no residual note pollution this time.

- **Instrument physics:** Contrary to the guitar, piano notes are struck, not plucked, which makes a difference in terms of harmonics participation. Indeed, it can be seen that higher-order harmonics have a negligible participation compared to the fundamental C harmonic, du to the punctuality of the excitation..

- **Perception and timbre:** The fact that there is much less harmonic participation in higher-order harmonics means that the piano sound is closer to a tuning fork sound that that of the guitar. the almost sole participiation of the fundamental harmonic makes it a very bright and pure sound, almost without pollution by other frequencies.

- **Signal processing:** in this case, due to better recording conditions, the contribution of background noise was much less detrimental to frequency reading. However, it can be notices that in the lower-frequency range, where this particular note holds the most harmonic contibution, there is what can be considered as "frequency radiation", in which structure and string vibration contributes to other frequency contributions at higher times. it is also interesting to notice that the small background noise contribution is within the very-low frequency range, as shown by the log-scale frequency spectrum.

2.2 Wind instruments

Saxophone
The first wind instrument sound we analyse is a recording of a saxophone. This is an sample from the Commuz' 2020 recording. The note, according to the sheet is supposedly a G4, and we will try to find this using the frequency analysis below.

- **Played note:** As said above, the note is supposedly a G4 (392.00 Hz). After a fft, we are able to reverse-engineer a fundamental frequency : $393.7Hz$, which is noticebly the same frequency.

- **Instrument physics:** Thes axophone is usually said to be equivalent to a conical duct with closed-end and open-end conditions. However, this is obtained with a very specific design of the mouthpiece. Here we have a saxohpone alto, which is recurved, and the physics behind it are more complex. But if we restrain our study to the case of a soprano saxophone : the mouthpiece is considered a close-end condition and the pavilion is of exponential section, which enables the open-end condition on that side of the bore. Then the frequencies are such that : $f_n = n\frac{c_0}{2L}$.

- **Perception and timbre:** The first thing to notice is that the 2^{nd} peak ($786.6Hz$) is higher than the first one. It corresponds exactly to the second harmonics ($2*393.7 = 787.4$). This peak is one of the main characteristics of the saxophone special sound. It is responsible for the "body" part of the saxophone sound with a bit of "tinny" sound to give the saxophone its warm round sound. Then when zooming in on both the time signal and the spectrogram, we can see a light vibrato. Harmonics

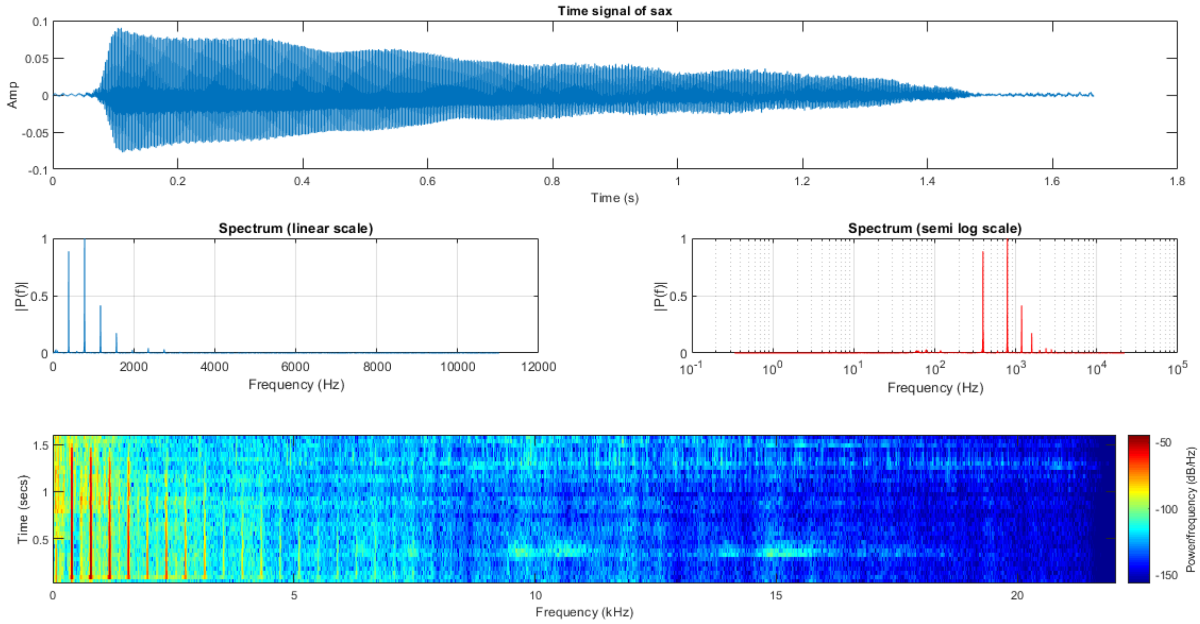

Figure 9: Characteristics of G4-note played on a saxophone alto

are oscillating around their mean value, but not the fundamental frequency. Is this why we still hear the note correctly pitched? Or is it because our ears sensitivity is not sufficient to hear such small and quick variation in frequency (though we still hear the vibrato, which means we have to some extent this sensitivity).

About the technique of vibrato, using a string instrument, it is extremely simple to achieve it since the only requirement is for the player to press the string harder in order to change the tension of the string. Again, the frequency is linked to the applied tension, varying this tension thus brings the sound high/low to create the vibrato. Here with the saxophone, the technique is different because there is no direct correlation between the frequency and the flow of air. Indeed, there is so many parameters to take into account in the mouthpiece, that studies have failed to create an artificial saxophone player. Only humans and their "sensitivity" to the mouthpiece manage to create a sound.

- **Signal processing:** The saxophone was recorded using a condenser microphone (Shure Beta 98H/C) which is adapted to this kind of instrument, but it was recorded live during a concert. Because of that, we have some background noise, but the sample is clear enough to process our analysis and have conclusive results.

Flute
Now let's focus on the flute sound. The sample is also from the Commuz' 2020 and recorded live.

- **Played note:** The reverse-engineered frequency value of this particular sound is $784.6161Hz$ and corresponds to G5 ($783.99Hz$).

- **Instrument physics:** Contrary to the saxophone, the flute is considered as a cylindrical duct with open-end condition at both ends. Moreover the flute has no

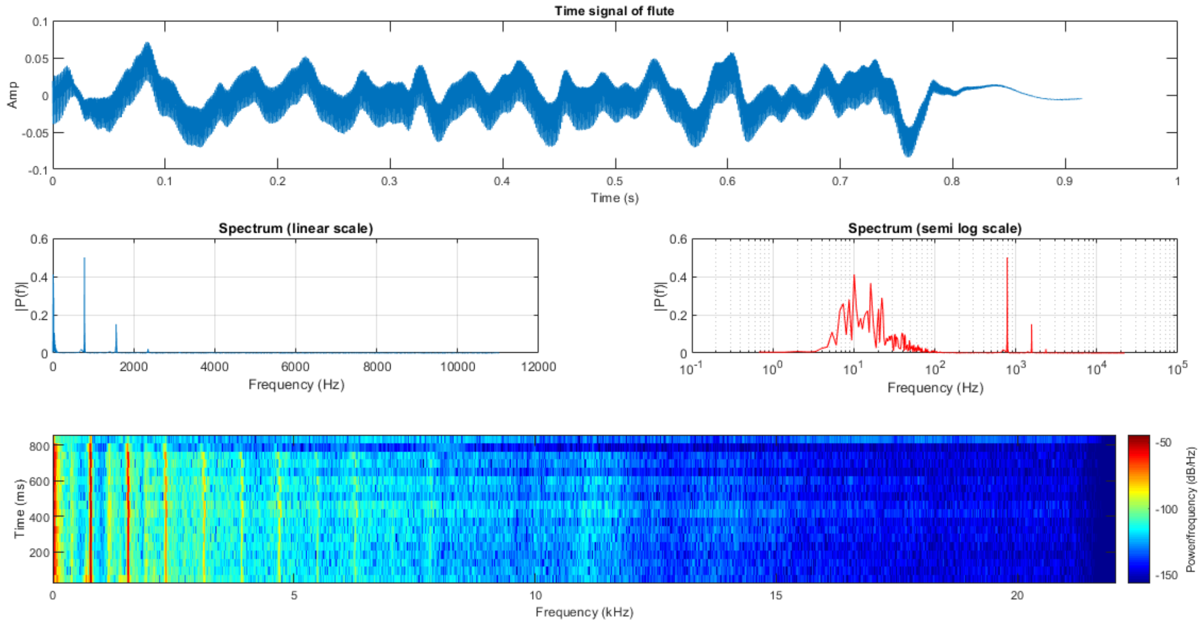

Figure 10: Characteristics of G5-note played on a flute

pavilion which changes dramatically the output impedence of the system. But the frequencies allowed inside the bore are the same as the saxophone : $f_n = n\frac{c_0}{2L}$.

- **Perception and timbre:** The spectrums (log and linear) show not much harmonics and we can expect the sound to be really "pure" (meaning, close to a sinus). Furthermore, when zooming in on the time signal and observing one period, the shape is indeed close to a sinusoid (see Figure 11) :

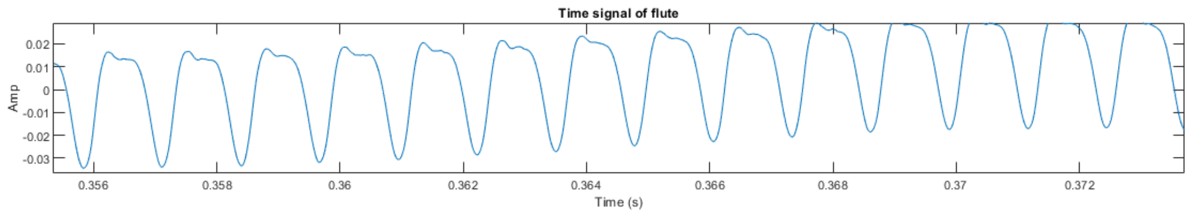

Figure 11: Zoom in the time signal of the flute : close to a perfect sinus

When listening to this particular sample, one can tell the flute sounds "hesitant". We can also see these characteristics on the time signal as it oscillates quasi-randomly in amplitude. This might participate in creating a range of frequencies in the low-tones because the analysis can see this oscillation as a modulation in amplitude of low frequency.

- **Signal processing:** The recording uses a Shure SM81. We can observe on the spectrogram that higher frequencies almost fade out at the same time as the fundamental stops. The sound must have a full spectrum, even if it is played loud or quite.

Another particular point it is interesting to notice is the presence of a large amount of low frequencies (see the log-scaled spectrum of Figure 10). One can explain this phenomenom by saying that the recording is not good because of background noise, low-pitched hum and ground noise. The signal then needs EQ-ing and this one of the reason mixing consoles and professionnal mixing engineers use a high pass filter to cut this unwanted range of frequencies. But one can also explain the phenomenon by pointing out the hesitation of the player (see above).

Trumpet
Because we did not have a trumpet player nor a trumpet within easy reach, we decided it would nice to analyze the trumpet sound from a well-known musician. As a result, here is in Figure 12 the first note of "Blue in Green" by Miles Davis. What makes his sound so special on this tune?

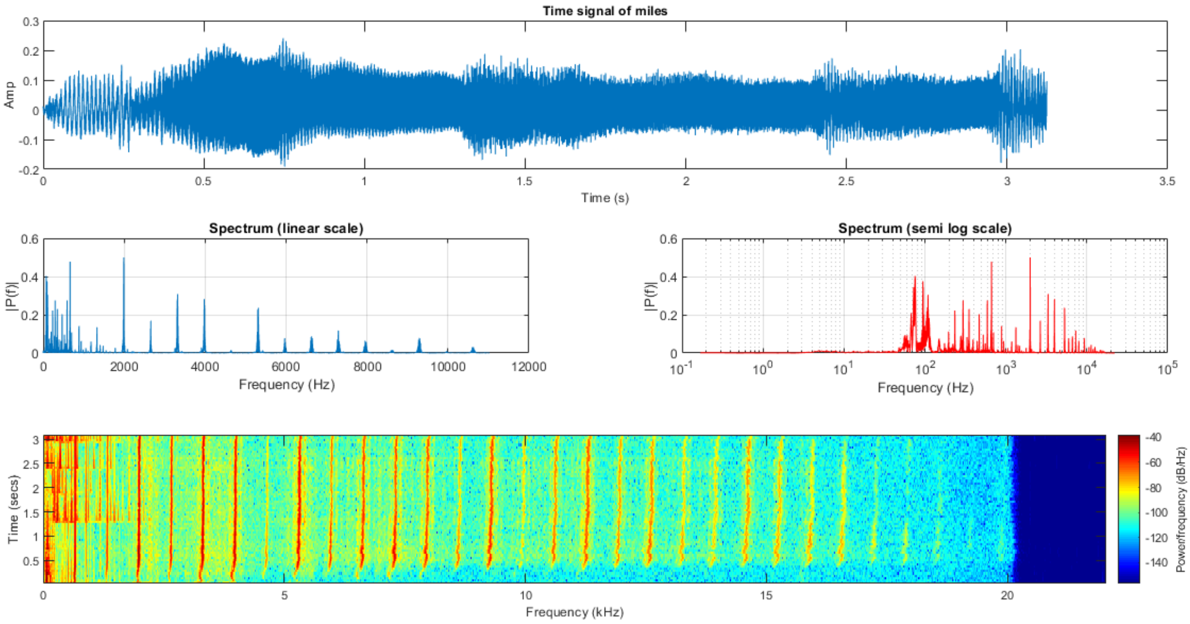

Figure 12: Characteristics of first note of "Blue in Green" played by Miles Davis on his trumpet

- **Played note:** According to the sheet, the note is a E5 ($659.25Hz$) and the reversed-engineered value of the fundamental is $664.7Hz$. Interestingly, Miles seems to play a little too high-pitched (maybe on purpose) and in fact, the difference is about eigth a tone.

- **Perception and timbre:** At first, there is a rich spectrum : high frequencies are almost as present as the lowest harmonics. Usually, the higher the frequency, the lower its peak amplitude. However in the present case, the amplitudes seems to lower slowly, thus creating a sound with many harmonics. When listening to Miles on this tune, his sound is particularily bright, which is the musical term to say that there is powerful high frequencies. The power spectrum is also defined essentially by a steep peak at $2kHz$ wich contributes a lot to the typical "horny" sound of the trumpet (which is one of the horns family of instruments). Miles also plays this note with a special mute that sounds almost "metallic" and very bright. We can imagine that the mute has a lot of influence over this high peak. We also have an interesting $5kHz$ peak, though smaller than the previous one, that is the core of the trumpet's sound. According to many sound engineers, the $5kHz$ peak of a trumpet represents the "brass/bell" sound.

- **Signal processing:** though the recording is of great quality, the player does not play alone (piano behind), we have then a "polluted" signal. Some of the frequencies might not be part of the trumpet spectrum, however, since Miles plays quite loud on this section, we can expect to have an accurate representation of its spectral presence. This is why the lowest part of the spectrum has many different frequencies. In other words, Miles's sound takes the high part of the spectrum, thus we can hear it very clearly among the others instruments.

3 Synthesis of replicated sounds

In this section, we will use the samples we have from different instruments to create a synthetiser. Based on the spectral analysis we have went through in section 2, we are now able to identify the frequencies that are specific to one instrument. The idea is to recreate these typical frequencies and using the additive synthesis to synthetise each instrument.

3.1 Importance of time-dampened harmonics

In the first place, we try the following method to create our synthesized sound :

1. Analysis of instruments samples: the objective is to obtain reversed-engineered frequencies from the spectrum and their corresponding amplitudes.
2. Then we use this frequencies to re-create sound using additive synthesis

Now that we have access to the frequencies that characterize a given instrument, we are able to synthesize a sound by adding those signals. The following matlab function is useful for this application (see Equation 1):

```matlab
function [s, t] = additiveSynthSimple(Freqs, Amp, Tmax,
    Fs)
    t = (0:1/Fs:Tmax)';
    s = 0*t; % vector of zeros
    for i = 1:size(Freqs)
        s = s + Amp(i)*sin(2*pi*Freqs(i)*t);
    end
    s = s/max(abs(s)); % normalization
end
```

This simple method takes one interesting parameter: *Amp*. We can choose the influence of each frequency in the time signal, so the signal shape can be modulated with those amplitudes, and so is the sound we hear. In other words, we have here the possibility to "equalize" our sound and shape it. But for the moment, let's try to make it sound like a real instrument. We have 3 options :

- Every amplitude is the same: each frequency has the same influence
- Amplitudes decrease exponentially with the frequency
- Amplitudes matches the original samples's

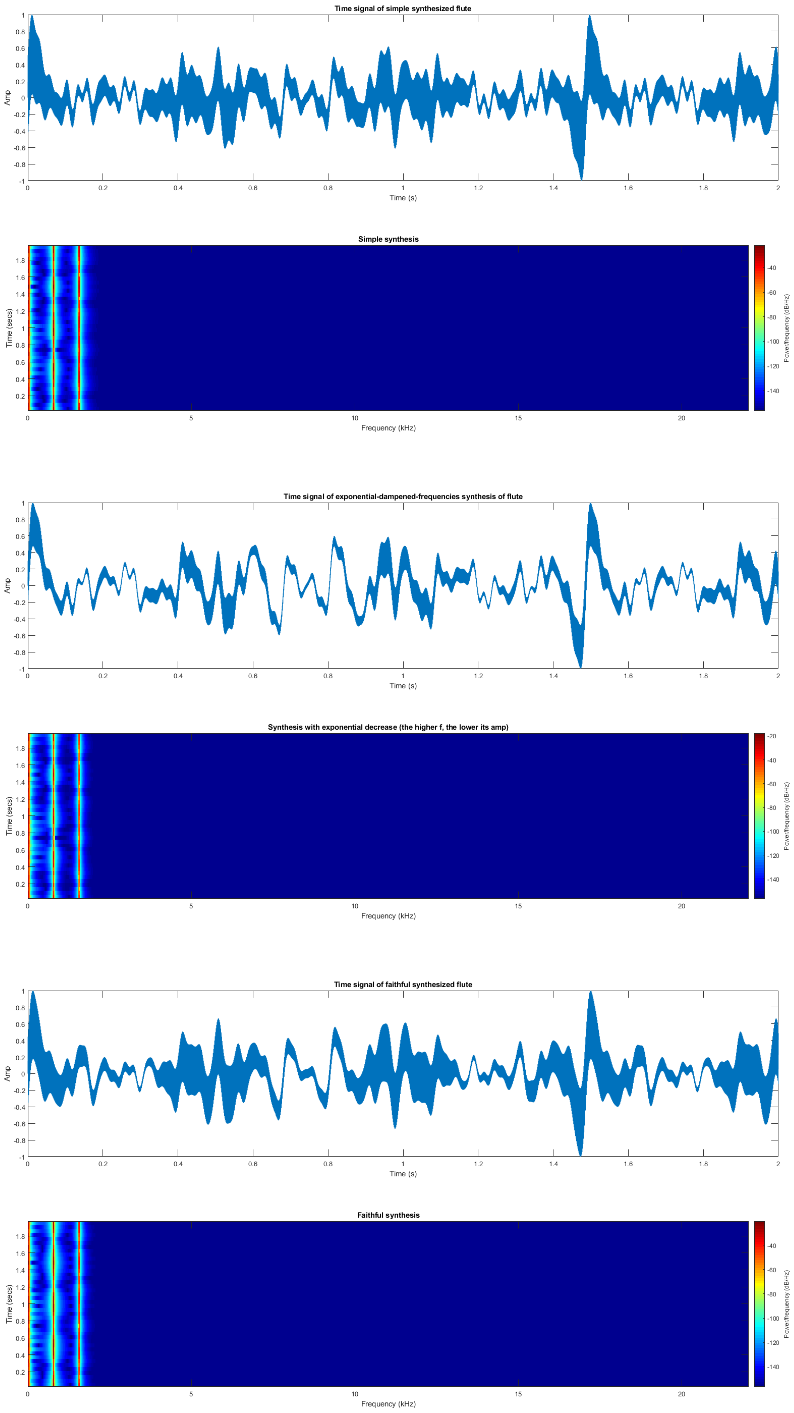

Figure 13: Synthesized flute

Even though the time signal seems quite similar to the original one (to some extent of course), the sound does not feel the same at all. You can tell very easily from the sound created that it was synthesized and not played by a real musician[1]. Moreover, it is difficult to tell which instrument is played here. In other words, our synthesizer fails to produce a good sound.

We have seen in **subsection 1.2** that real musical signals have damped frequencies. Given the current results of our synthesizer, it is clear that it lacks a lot of features, and the signal that is produced must have many more "real" features that we can simulate. We will try in the following section to implement a frequency damping to the signals.

3.2 String instruments

String instruments resonates, whereas wind instruments need a player to create sound. As a result, when the player stops blowing, a very low release time due to the small amount of space in the conduct and the naturally low-resonating space around the mouth does not allow us to replicate decaying time for wind instruments. However, in the case of resonating string instruments, quick attack, sustain and decay allows us to properly modelise release times with our damped-signal generation code.
One thing to notice is that for these resonating string instruments, decay times seem to decrease linearly with frequency. As is, higher-frequency harmonics decay faster. This is to be expected, the explication being that the strings have a higher rate of friction against air at higher vibration speeds. As an additionnal challenge to this sound-synthesis, we will attempt to replicate this behaviour by adding a linear-decreasing envelope to the decay time list. We now analyse the spectrograms and time signals produced by synthesizing guitar and piano sounds:

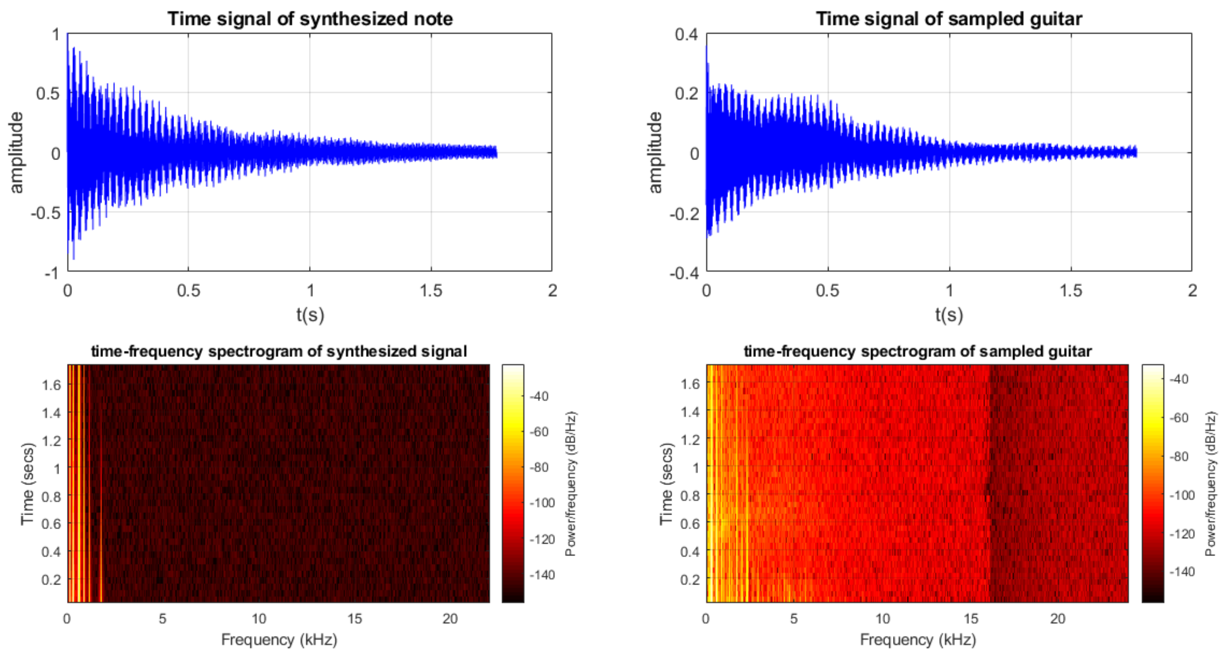

Figure 14: Synthesis of a guitar note

though there is a bit of frequency pollution due to the residual E which is explained

[1]All the samples are available in the zipped file of all functions and figures given along with this report

section 2, the guitar synthesized sound is quite close to the sample, both in terms of waveshape and time-frequency locations. Upon listening to the sampled file, though knowning that it is synthesized, it remarkably resembles the sampled sound, both in timbre and decay times. What might be missing is the slight guitar body vibration, reducing in other strings vibration, induced by the plucking excitation, which can be heard in the original sample, and the lower-amplitude harmonics which provide realistic richness to the sample. Other than that however, the result is quite satisfying in terms of signal processing and timbre. The following waveforms and spectrograms show the results for the piano.

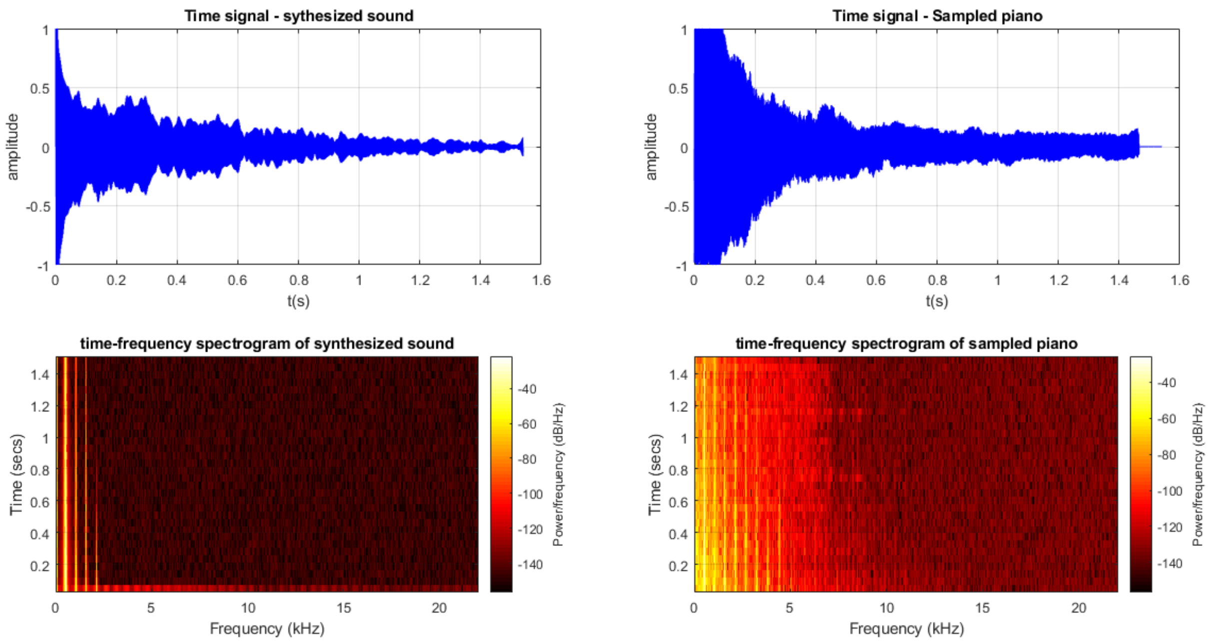

Figure 15: Synthesis of a piano note

This time, though spectrogram and waveforms differ a bit, and the richness of the piano suffers a bit from the suppression of frequency bleeding at higher frequency ranges, the sound produced very closely resembles that of the sample. This is satisfying from the standpoint of a synthesizer, as it is hard to discern either of the sound when listening to them. the is predictable, as the attack time of a struck chord is nearly instantaneous, and the high tension of strings within the piano allows for quick dissipation of higher-order harmonics, resulting in a close to pure sound. the results of the synthesizer in the case of resonant struck-string instruments are very satisfying.

3.3 Limits of our synthetizer

Obviously, the "synthesizer" coded in the above section is a very basic one. it only takes into account harmonic frequencies and amplitudes, as well as the basics of time-dampening for some string instruments as seen in the previous section. In reality, instruments and professional synthethisers present many more characteristics to a sound that are difficult for us to replicate within the scope of this project:

- The basics of a musical sound can be decomposed in four parameters : attack, decay, sustain and release. Attack dictates the necessary time for any given sound to reach its maximal amplitude (for instance, a legato violin would reach it much slower than a staccato one), Sustain the amount of time till it reaches median amplitude from maximum amplitude. Decay indicates the time it stays at median amplitude, and release the amount of time it takes to reach 0 from median amplitude. these parameters influence perception of sounds, and without such amplitude modulation, it is difficult to replicate a "real" sound. The reason guitar plucked strings are easy to replicate is because they naturally have fast attack, fast sustain, fast decay and slow release, which is what we modelize in our code.

- another simulation parameter necessary to replicate real sound would be to replicate the excitation mode of the instrument. Plucking, rubbing, striking and blowing, and even different types of blowing, all excite instruments in different manners, and it is presumptous to assume that we would be able to replicate many instruments easily with such basic parameters

- Another parameter which would have to be taken into account would be pitch bending. Tremolo, vibrato as well as bending for string instruments are all integral to best replicating said instruments,which necessitate amplitude modulation and frequency modulation over time, which are not implemented here.

Basically, though commercialized synthesizers and plugins confirm that it is possible to replicate to a better extent musical sounds, the scope of this project does not give us enough tools or data to attempt reaching such fidelity.

Conclusion

Through this report we have studied the frequencies of a synthetic sound, the peaks in the spectrum and also time damping. Additive synthesis is a simple way to synthesize sound, so we combined this method with reverse-engineered frequencies method to re-create particular sounds. In the meantime, we conducted an analysis of the characteristics of many instruments such as the saxophone, the piano, the guitar and more, as well as a quick review of the phyisics behind each of these instruments.

This report also presents a very (very) simple signal-processing theory and proved how many possibilities can be found to process and analyze a given sample (windowing, sampling rate, overlapping, ...).

Musical acoustics can appear rather simple at first glance, however the problems it tackles may be much more complex and interesting than it might seem at first. So we went through many difficulties including implementation of a good analysis of a given instrument. Because each instrument has its own physic model, it is difficult to multiply measurements as they may vary in many ways depending on the external conditions. But it was all the more interesting to focus on specific sound recordings and try to explain different phenomena we feel by listening.

Creating a synthesizer turned out to be much more difficult than expected, and this project features only a mere replica of what one can call a "synthesizer". But the very basics of the synthesis technology are tackled and implemented in this study. Facing those difficulties made us realize why the synthesizer is so recent and also why people went it first appeared.