FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Multimedia and New Services

Lab 1 - Sampling and Quantisation of Media Signals

Mário Santos
up201503406

Mestrado Integrado em Engenharia Informática e Computação

February 1, 2021

Multimedia and New Services
Lab 1 - Sampling and Quantisation of Media Signals

Mário Santos
up201503406

Mestrado Integrado em Engenharia Informática e Computação

February 1, 2021

Contents

List of Figures

List of Tables

Chapter 1

1. Optional/Introductory Part

1.1 Methodology

An .mp3 file was converted into 2 .wav files, one was converted with a sample rate of 11025 Hz and the other with a sample rate of 44100 Hz. A perceptual comparison was made between the 2 generated files. Figure 1.1 displays the media information regarding the file.

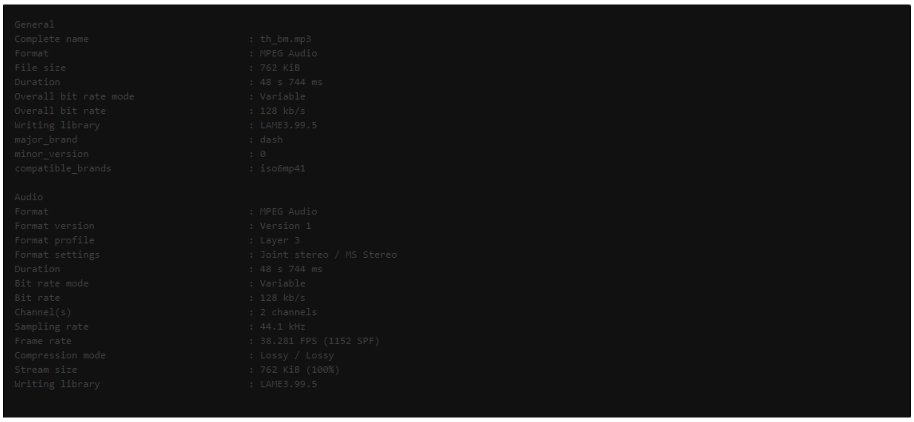

Figure 1.1: Media information of used file

1.2 Assessment

The sample rate determines how many samples, or measurements, of the sound are taken each second. The more samples that are taken, the more detail about the signals' structure is recorded, and the shape of the sound wave is captured more accurately. This results in an increase in audio quality. A lower sound quality can be perceived for the file converted with a sample rate of 11025Hz, this is due to the aliasing effect, caused by the sample rate being smaller than the double of it's maximum frequency (Nyquist). We could not perceive any differences in the file recorded with a 44100Hz sample rate when compared to the original one, this is because the original file already had a 44100Hz sample rate.

Chapter 2

2. Variation of the sampling frequency with/without filters

For this experiment the Matlab scripts "amostragemInterp_semFiltro.m" and "amostragemInterp_comFiltro.m" were used. An analysis was performed on the analytical and perceptual effects that subsampling and interpolation had on a sound file, using different sampling factors. To analytically compare the generated sound files the, Maximum Error, PSNR and Mean Squared Error (MSE) were used. The selected sound file was 'bugsbunny1.wav', with a sample rate of 11025Hz, a total length of 2.2916 seconds, 8 bits per sample, and a total of 47315 samples. The file's waveform and frequency spectrum are displayed in Fig. 2.1.

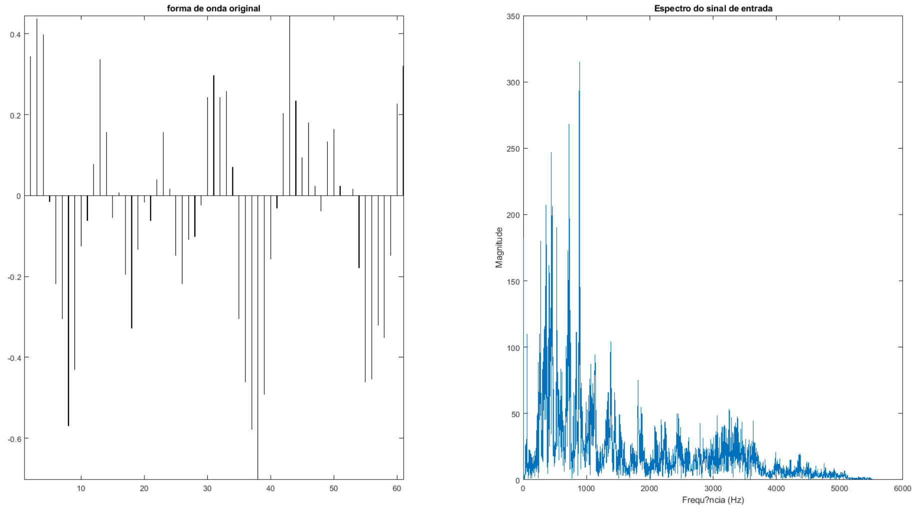

Figure 2.1: Original waveform and spectrum

Subsampling with a factor of i means that for each i samples in the original sound only 1 sample will be recorded.

2.1 Sampling without Filter (2. ii))

Figures 2.2 and 2.3 display the waveform and spectrum of the sound signal after being subsampled by a factor of 4 and 2 respectively. Given that the file has a sample rate of 11025Hz, the Nyquist frequency is 5512.5Hz. When subsampling 4 to 1 the maximum frequency is double that of Nyquist frequency, this results in frequencies that are not present in the original signal (distortions). When subsampling 2 to 1, the maximum frequency is also exceeds the Nyquist frequency, but not by as much as when subsampling 4 to 1, thus we can still observe some distortions, but the signal resembles the original signal more, when compared to the 2 to 1 subsampling. Perceptually, when subsampling 2 to 1 the sound is much clearer compared with the 4 to 1 subsampling, where the distortions are more noticeable and result in a 'static-like' sound.

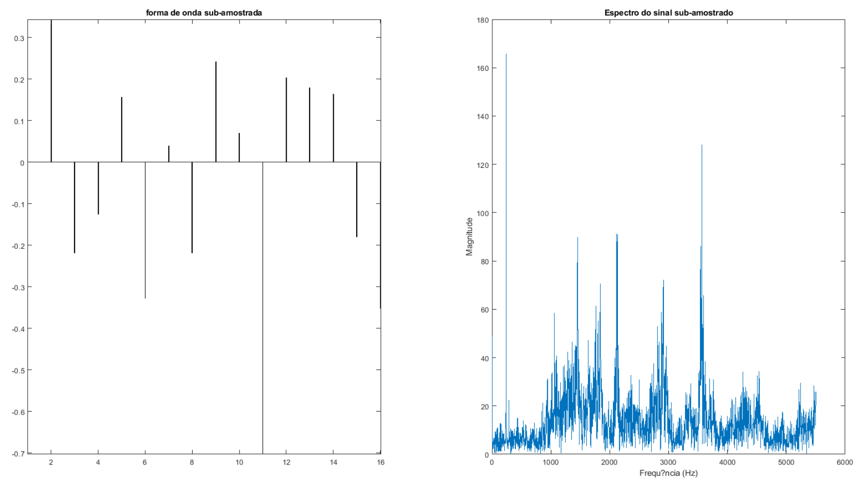

Figure 2.2: Waveform and spectrum of sound after being subsampled by a factor of 4

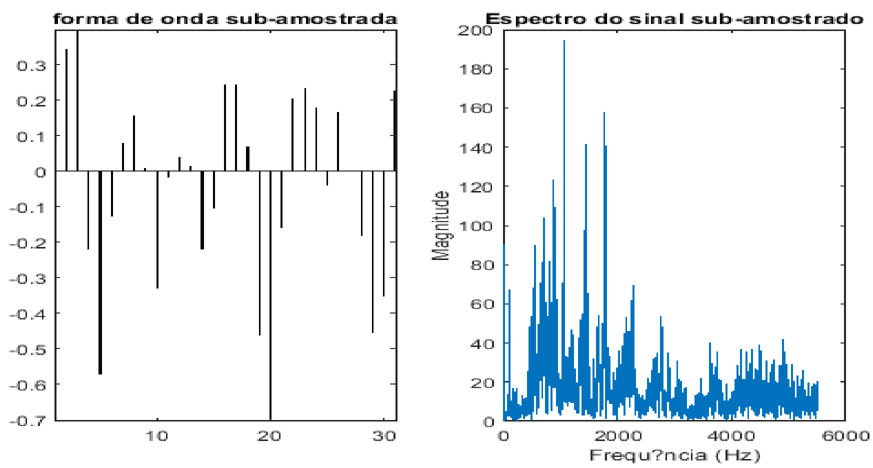

Figure 2.3: Waveform and spectrum of sound after being subsampled by a factor of 2

Figures 2.4 and 2.5 display the waveform and spectrum of the sound signal after being subsampled by a factor of 4 and 2, respectively, and then interpolated. No pre-filters or filters were used and once more we can observe frequencies that aren't present in the original signal, particularly for the higher range of frequencies. However in both cases the signal more closely resembles the original signal when compared to the subsampled signals. When subsampling 2 to 1 and interpolating, the perceived sound quality is close to the original and in the case of subsampling 4 to 1 and interpolating, the perceived sound quality is noticeably worse.

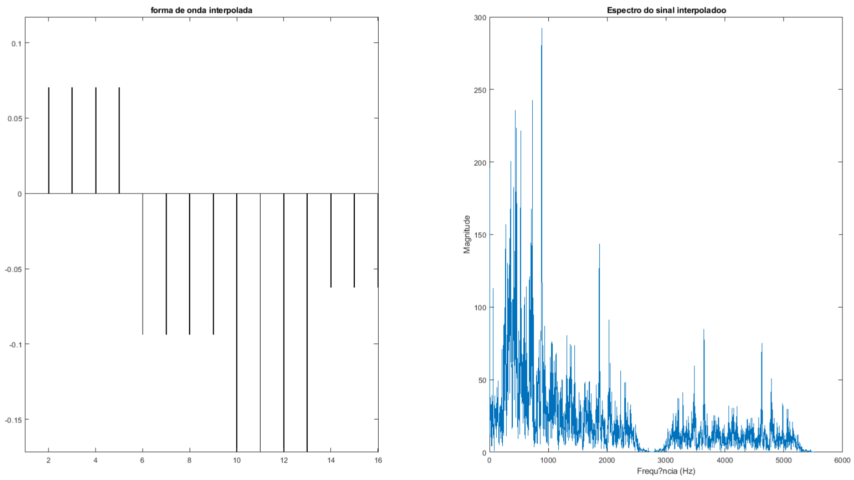

Figure 2.4: Waveform and spectrum of sound after being subsampled by a factor of 4 and then interpolated

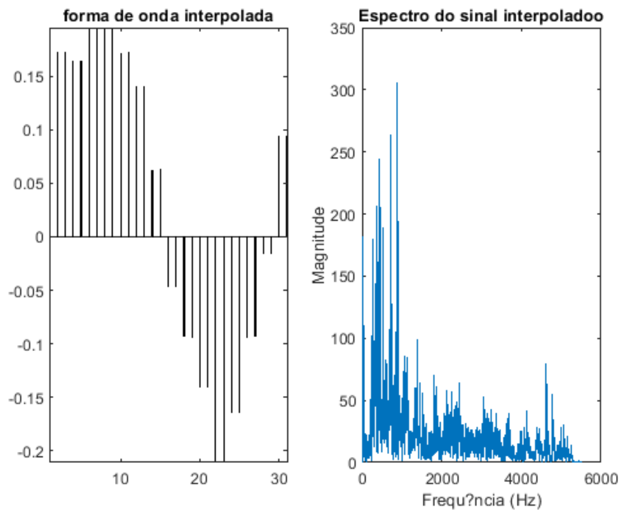

Figure 2.5: Waveform and spectrum of sound after being subsampled by a factor of 2 and then interpolated

Table 2.1 displays the Maximum Error, the PNRS and MSE obtained from each experiment.

	Max Error	**PSNR**	**MSE**
Sample Factor - 4	0.0187	14.3214	0.0036
Sample Factor - 2	0.0056	19.5319	0.0006

Table 2.1: Overview of measurements after unfiltered sampling

We can see the audio file subsampled by a factor of 4 presents the lowest PSNR and the highest Maximum and Mean Squared Error, an indication of worse quality. These measurements indicate that the audio subsampled by a factor of 2 has a higher quality (lower Maximum and Medium Squared Error, higher PSNR), this is expected given that it has double the sample rate of the file subsampled with a factor of 2.

2.2 Sampling with Filter (2. iii)

Given that the audio file has a sample rate of 11025Hz, the Nyquist frequency is 5512.5Hz. By subsampling with a factor of 4, the sample rate is now 2756Hz, meaning that max frequency to avoid the aliasing effect is 1378Hz and an anti-aliasing filter should be used ahead of the sampler to attenuate the frequencies above that limit. By subsampling with a factor of 2, the sample rate is now 5512.5Hz, meaning that max frequency to avoid the aliasing effect is 2756Hz and an anti-aliasing filter should be used ahead of the sampler to attenuate the frequencies above that limit.

In figures 2.6, 2.7, we can see the pre-filters impulse and frequency response when subsampling by a factor of 4 and 2 respectively. In the case of subsampling by a factor of 4, the filter begins to cut off frequencies higher than approximately 2000Hz, and in the case of subsampling by a factor of 2, it begins to cut off frequencies higher than roughly 4000Hz.

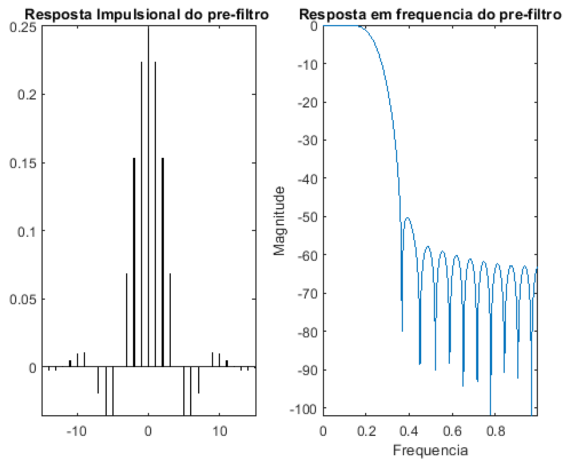

Figure 2.6: Pre-filter impulse and frequency response with a subsampling factor of 4

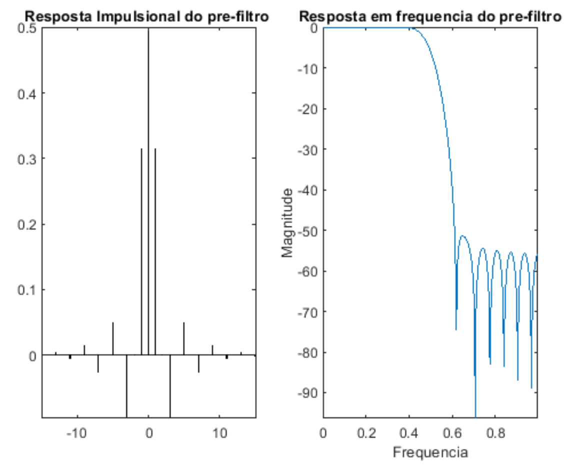

Figure 2.7: Pre-filter Impulse and frequency response with a subsampling factor of 2

In figures 2.8, 2.9, we can see the filters impulse and frequency response when subsampled by a factor of 4 and 2 respectively. In the case of subsampling by a factor of 4, the filter begins to cuts off between approximately 3000Hz and 7000Hz, and frequencies over 8000Hz. When subsampling by a factor of 2, the filter begins to cuts off frequencies over 5000Hz, roughly.

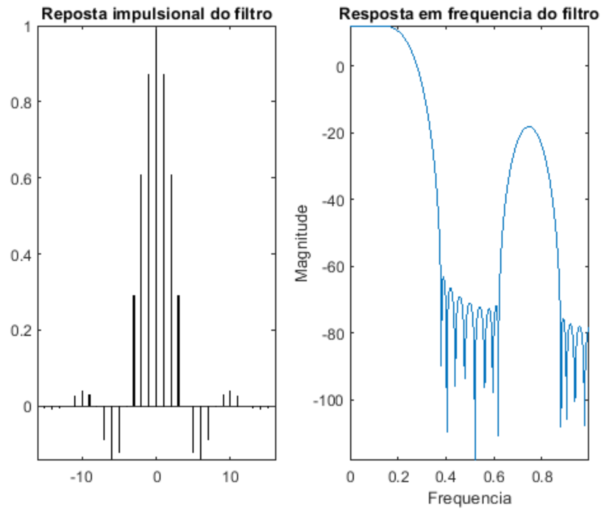

Figure 2.8: Filter impulse and frequency response with a subsampling factor of 4 and interpolated

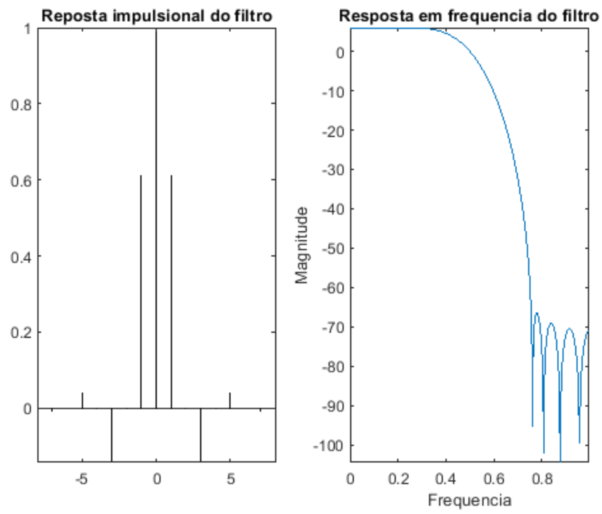

Figure 2.9: Filter Impulse and frequency response with a subsampling factor of 2 and interpolated

Figures 2.10 and 2.11 display the waveform and spectrum of the sound signal after being subsampled by a factor of 4 and 2 using the pre-filters. In both cases we can see there is less distortion when compared to the unfiltered subsampled signals, particularly in the higher frequencies.

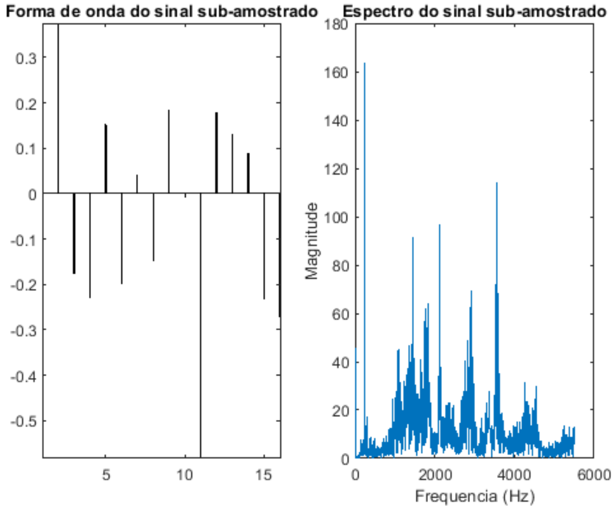

Figure 2.10: Waveform and spectrum of sound after being subsampled by a factor of 4

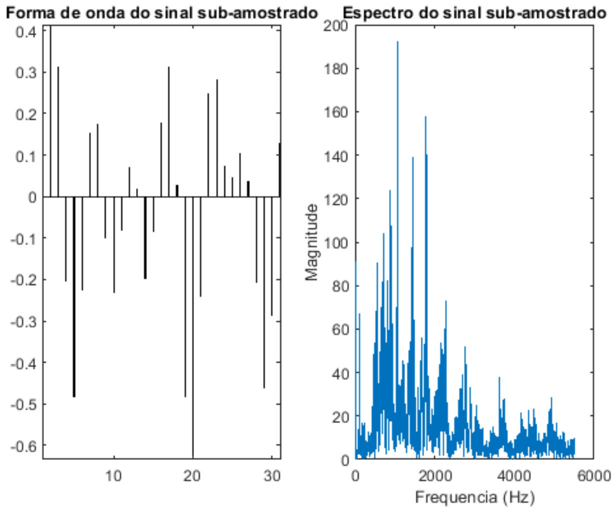

Figure 2.11: Waveform and spectrum of sound after being subsampled by a factor of 2

Figures 2.12 and 2.13 display the waveform and spectrum of the sound signal after being subsampled by a factor of 4 and 2 using pre-filters and then interpolated using filters. In both cases we can see there is less distortion when compared to the unfiltered interpolated signals.

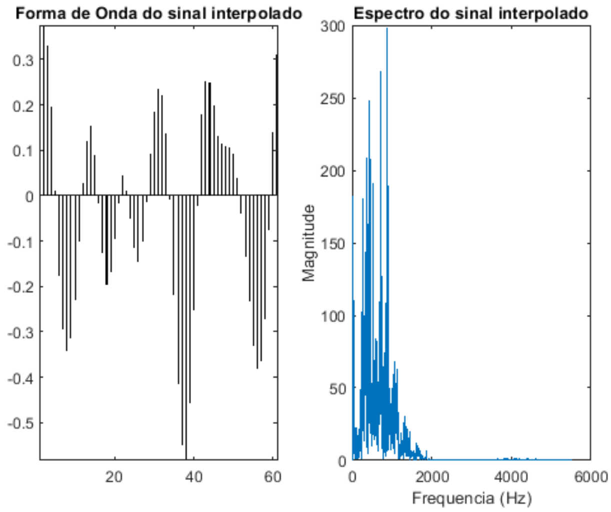

Figure 2.12: Waveform and spectrum of signal after being subsampled by a factor of 4 and then interpolated, using pre-filter and filter

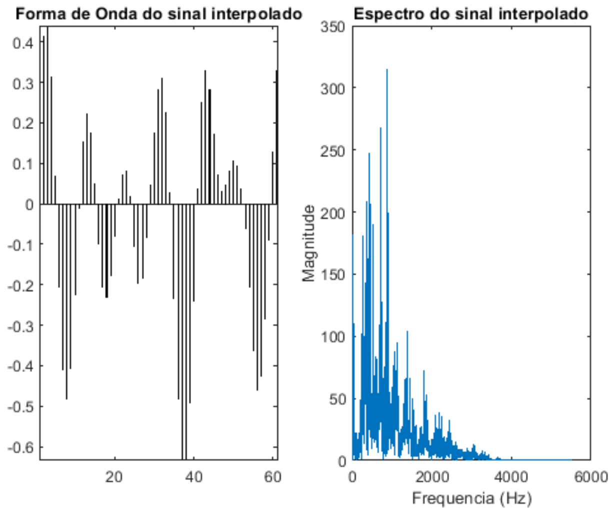

Figure 2.13: Waveform and spectrum of sound after being subsampled by a factor of 2 and then interpolated, using pre-filter and filter

	Max Error	**PSNR**	**MSE**
Sample Factor - 4	0.0049	20.1449	3.50e-23
Sample Factor - 2	0.0020	23.9276	2.42e-27

Table 2.2: Overview of measurements after filtered sampling

These measurements indicate that the audio subsampled by a factor of 2 is more accurate (lower MSE), this is expected given that it has double the sample rate of the file subsampled with a factor of 2. Both errors are also multiple magnitudes smaller when compared to the unfiltered versions, indicating a much higher accuracy to the original sound, and therefore a higher quality. For the same value of k (sampling factor), the script "amostragemInterp_comFiltr.m" produced the best results, in terms of analytical performance measures and perceptual quality. Based on these experiments we can see that the pre-filters and filters are a useful method to improve the quality of a subsampled audio, and the ones used did successfully cut off the desired frequencies.

2.3 Measurements Overview

In this section an overview of the metrics measured for the sampling experiment is presented in table 2.3.

	Max Error	PSNR	MSE
Sample Factor - 4 No Filter	0.0187	14.3214	0.0036
Sample Factor - 2 No Filter	0.0056	19.5319	0.0006
Sample Factor - 4 Pre-Filtered	0.0049	20.1449	3.5016e-23
Sample Factor - 2 Pre-Filtered	0.0020	23.9276	2.4172e-27

Table 2.3: Overview of measurements after sampling

Chapter 3

Quantisation Experiments

In the following experiments an audio signal was quantised using 2 different number of quantisation levels, 256 and 16. Originally, the audio file used 2 channels, using Audacity we converted the file to Mono (1 channel). The signal was verified to have been digitised using the PCM format (16 bits per sample and 44,1kHz sampling frequency), shown in Fig. 3.1.


```
General
Complete name                  : th_bm_short_1channel.wav
Format                         : Wave
File size                      : 564 KiB
Duration                       : 6 s 550 ms
Overall bit rate mode          : Constant
Overall bit rate               : 706 kb/s

Audio
Format                         : PCM
Format settings                : Little / Signed
Codec ID                       : 1
Duration                       : 6 s 550 ms
Bit rate mode                  : Constant
Bit rate                       : 705.6 kb/s
Channel(s)                     : 1 channel
Sampling rate                  : 44.1 kHz
Bit depth                      : 16 bits
Stream size                    : 564 KiB (100%)
```

Figure 3.1: Audio file information

3.1 256 Quantisation Levels

Figures 3.2 and 3.3 display the waveform and spectrum of the sound signal after being subsampled by a factor of 4 and after interpolation, across the whole length of the audio, and for a particular section (sample index in range [2000,2200]), respectively.

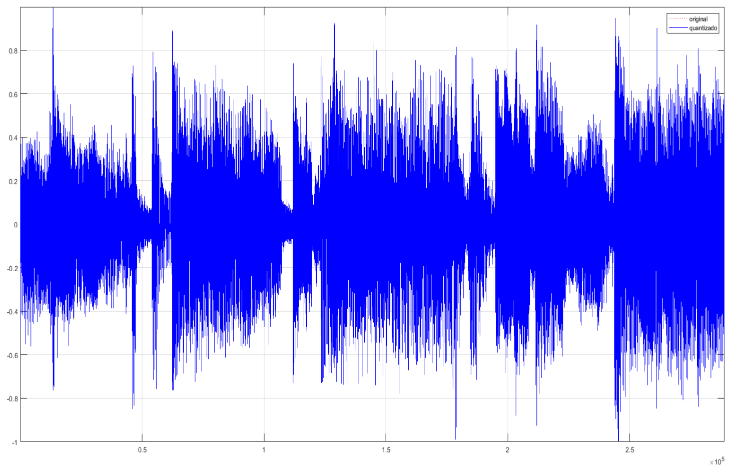

Figure 3.2: Full audio signal waveform after 256 level quantisation
Red: Original signal
Blue: Quantised Signal

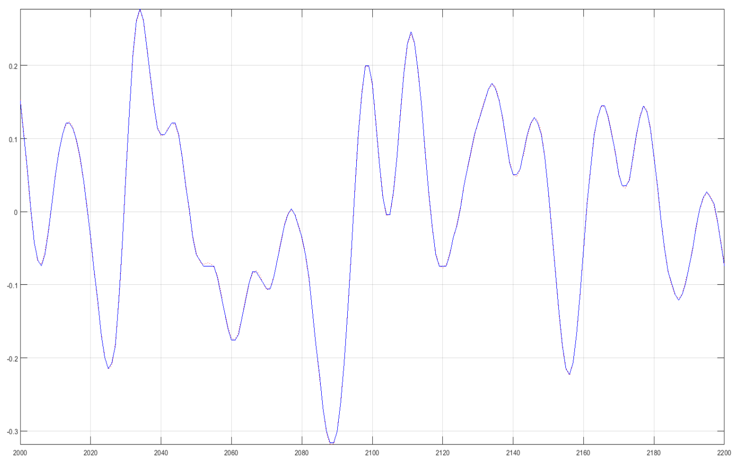

Figure 3.3: Section of audio signal waveform after 256 level quantisation
Red: Original signal
Blue: Quantised Signal

3.2 16 Quantisation Levels

Figures 3.2 and 3.3 display the waveform and spectrum of the sound signal after being subsampled by a factor of 2 and after interpolation, across the whole length of the audio, and for a particular section (sample index in range [2000,2200]), respectively.

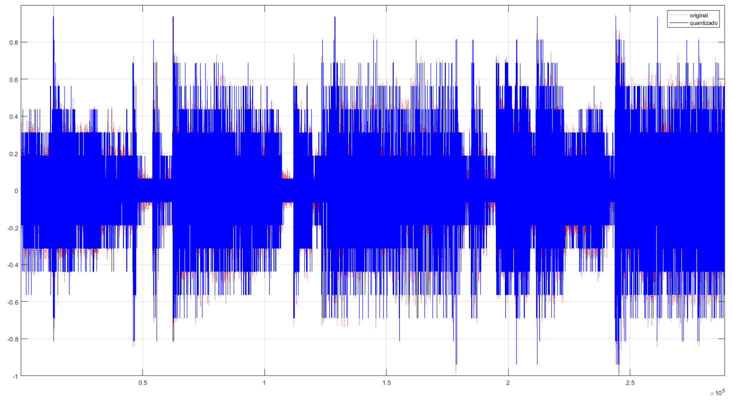

Figure 3.4: Full audio signal waveform after 16 level quantisation
Red: Original signal
Blue: Quantised Signal

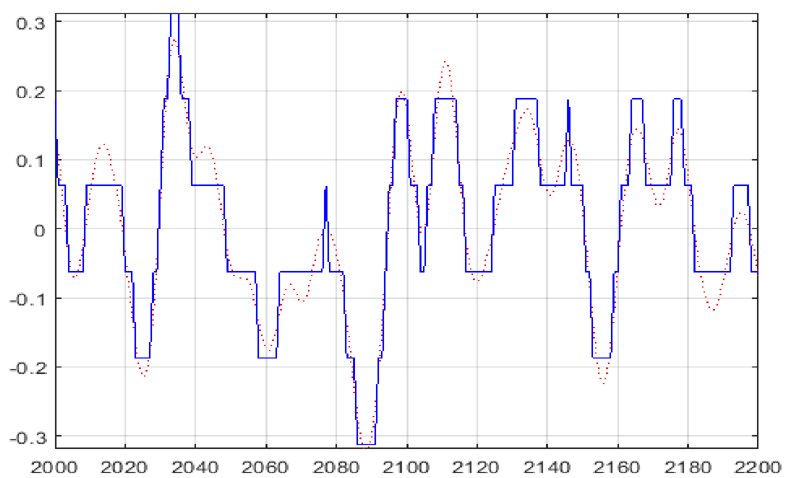

Figure 3.5: Section of audio signal waveform after 16 level quantisation
Red: Original signal
Blue: Quantised Signal

By graphical comparison it can be observed that the higher the quantisation level, the closer the signal will be to the original signal, given that more amplitude values are used. In this experiment, when using 256 quantisation levels, the generated sound is very similar to the original, however when using only 16 quantisation levels, the decrease in quality is clear, and the audio is significantly distorted.

3.3 Measurements Overview

The following table 3.1 displays the metrics for both quantisations. We can see that when quatising using 256 levels, we get a very small Max Error and MSE, and a high PSNR value. Conversely, quantising with 16 levels produced a larger Max Error and MSE, and a lower PSNR, indicating the worst precision and quality. These metrics agree with our graphical analysis, and we can conclude that the more quantisation levels are used, the higher the audio precision and quality.

	Max Error	**PSNR**	**MSE**
256 Quantisation Levels	5.0809e-06	22.4159	3.8472e-7
16 Quantisation Levels	0.0013	-13.7317	0.0063

Table 3.1: Overview of measurements after quantisation