Multimedia Services and Applications (SAM)
2021/2022

First Lab Assignment Report
Sampling and Quantisation of media signals

João Carlos Machado Rocha Pires (UP201806079)

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Contents

1 11025 Hz VS 44100 Hz **1**

2 Variation of the sampling frequency with or without filters **1**

2.1 Without filters 1

2.1.1 Sub-sampling with a factor 4 1

2.1.2 Sub-sampling with a factor 2 3

2.1.3 Comparison of the two experiences 5

2.2 With filters 5

2.2.1 Sub-sampling with a factor 4 5

2.2.2 Sub-sampling with a factor 2 8

2.2.3 Conclusions 10

3 Quantisation experiences **11**

3.0.1 Quantisation with 256 levels 11

3.0.2 Quantisation with 16 levels 12

1. 11025 Hz VS 44100 Hz

After selecting an Ed Sheeran song called *UNI*, which I have in .mp3 format since the Vinyl of the corresponding album comes with a code to redeem these songs in .mp3 format, I converted it to two files with sampling rate 11025 Hz and 44100Hz, respectively. When I listened to the first one, I noticed a distortion/noise in the audio that reminded me of old radios. Then, I heard the one with the sampling rate of an audio CD and I didn't notice any major difference when comparing to the original file.

2. Variation of the sampling frequency with or without filters

2.1 Without filters

2.1.1 Sub-sampling with a factor 4

Starting with the sub-sampling with a factor 4, in terms of perceptual quality, after the sub-sampling I noticed a "distant" sound, while after the interpolation I noticed a "slightly-distorted" sound, better than the one obtained after the sub-sampling, but (slightly) worse than the original one.

The pictures on the bottom show both the waveform and the spectrum for the original sound, for the sub-sampled sound and for the interpolated sound, respectively.

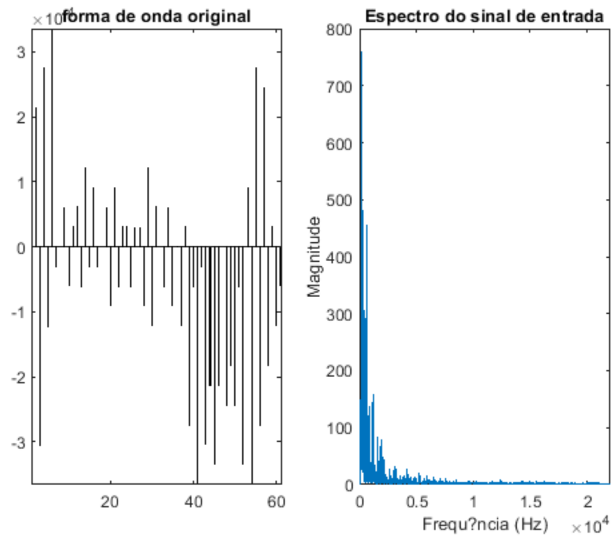

Figure 2.1: Waveform and spectrum for the original sound

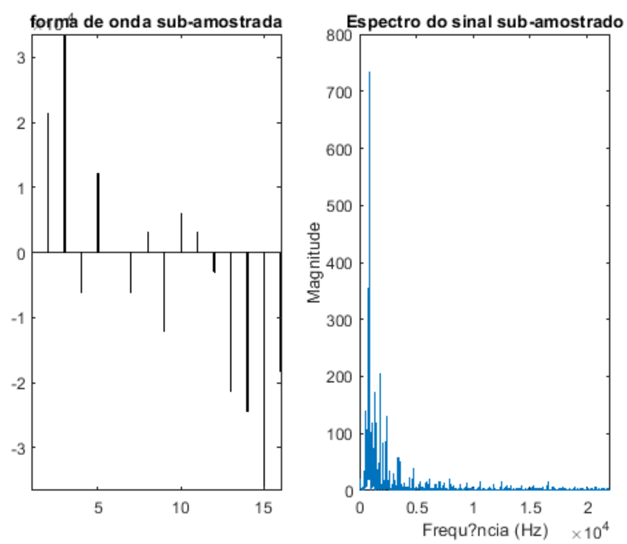

Figure 2.2: Waveform and spectrum for the sub-sampled sound

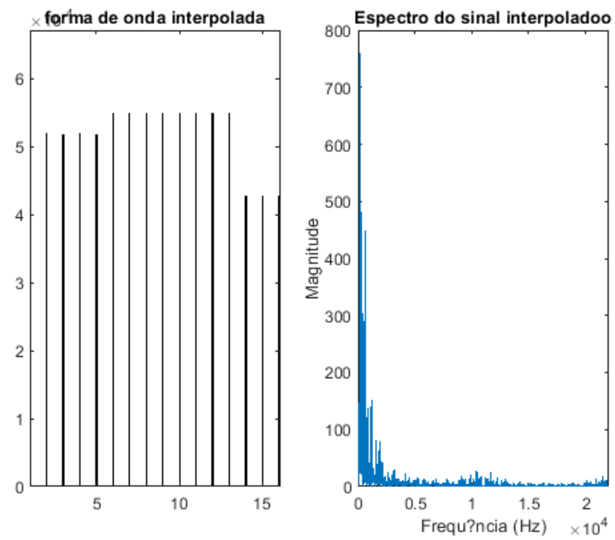

Figure 2.3: Waveform and spectrum for the interpolated sound

Comparing all the graphics, it's noticeable that the sub-samples and interpolated sounds have more representativeness in high frequencies (spectrum analysis) and less in lower frequencies (when comparing both to the original sound).

Regarding the waveform, it's notorious that the sub-sampled sound includes only some values of the original waveform, which was expected, and the interpolated sound waveform only includes positive values.

In terms of the obtained mean square error, the value was 8.251×10^{-5} (PSNR = 34.8165).

2.1.2 Sub-sampling with a factor 2

Doing the same process as in the sub-section above, but this time with a sub-sampling with a factor 2, in terms of perceptual quality, the sub-sampled sound seems more similar to the original sound that with a factor 4. I didn't noticed any major difference between both interpolated sounds.

Again, as above, the following pictures show both the waveform and the spectrum for the original sound, for the sub-sampled sound and for the interpolated sound, respectively.

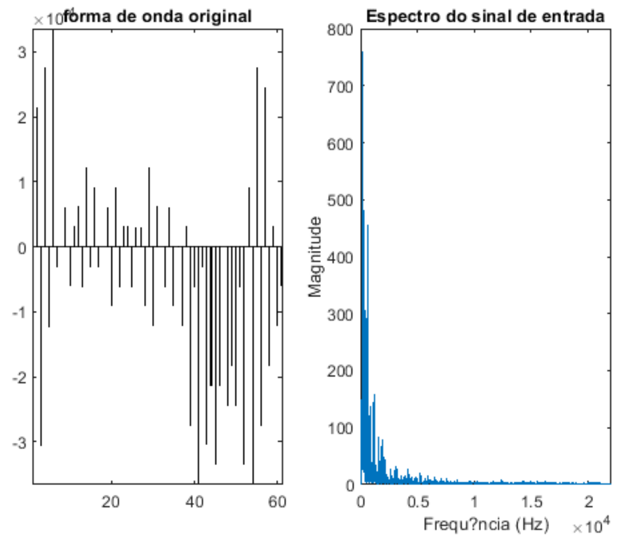

Figure 2.4: Waveform and spectrum for the original sound

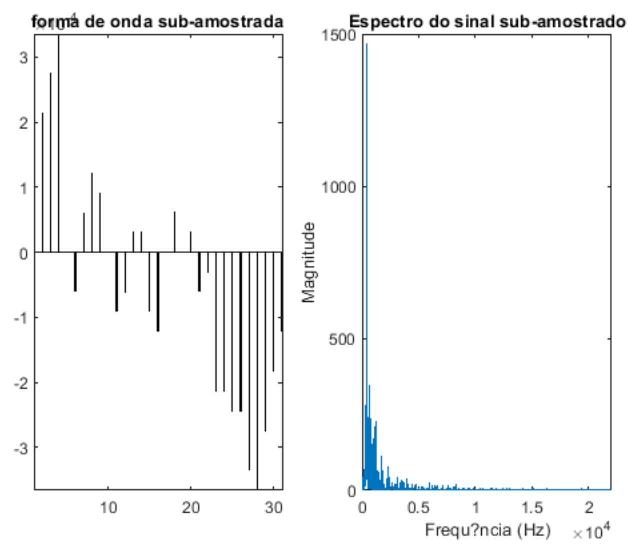

Figure 2.5: Waveform and spectrum for the sub-sampled sound

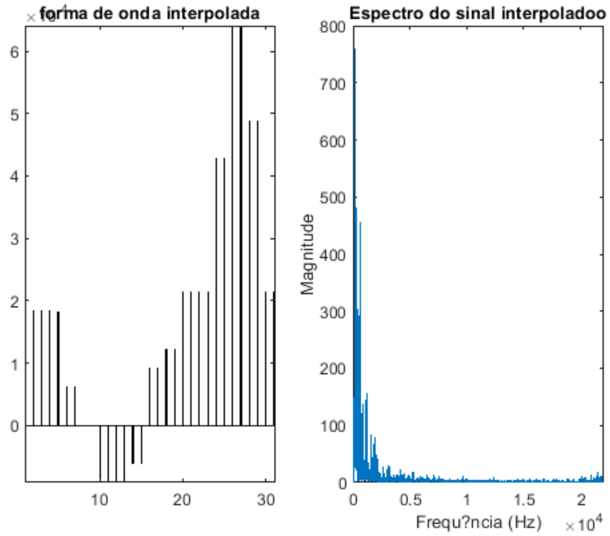

Figure 2.6: Waveform and spectrum for the interpolated sound

The first conclusion is similar, with an attenuation, to the one taken for the factor 4: the interpolated sound has higher magnitude for higher frequency values (spectrum).

However, regarding the waveform, we can easily verify that the sub-sampled sound waveform has more values than the one obtained for the factor 4 (as expected), and the interpolated sound waveform has now also negative values, being, nevertheless, still somehow different from the original one.

In terms of the obtained mean square error, the value was $2.093\,53 \times 10^{-5}$ (PSNR $= 40.7727$).

2.1.3 Comparison of the two experiences

The sub-sampling with a factor 2 produces sub-sampled and interpolated sounds closer, in terms of perceptual quality, waveform and spectrum, to the original sound.

Even when we compare the MSE, we can easily verify that its value is lower for the sub-sampling with a factor 2.

2.2 With filters

In this section, the experiences above, with the sub-sampling factors 4 and 2, were repeated, this time with the use of filters, and are reported below.

2.2.1 Sub-sampling with a factor 4

Starting again by the sub-sampling with a factor 4, let's first list the variables associated to the samples.

- Sampling frequency: 44,1KHz
- Number of samples: 447.692
- Number of samples (original sound): 447.692

Starting with the analysis in terms of perceptual quality, the sub-sampled sound seemed 'distant', but with better quality that the one from the previous experience (without the application of the pre-filter).

Regarding the interpolated sound, I didn't noticed any major difference when comparing to the original sound.

The following three plots show, as above, the waveform and the spectrum for the original, sub-sampled and interpolated sounds, respectively.

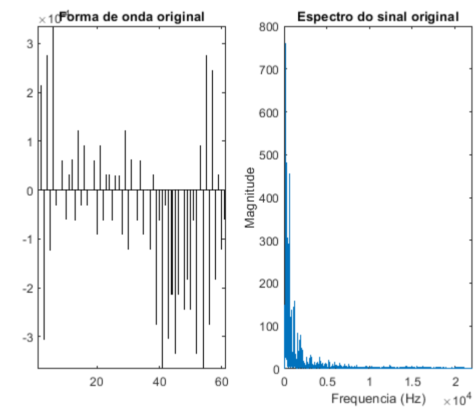

Figure 2.7: Waveform and spectrum for the original sound

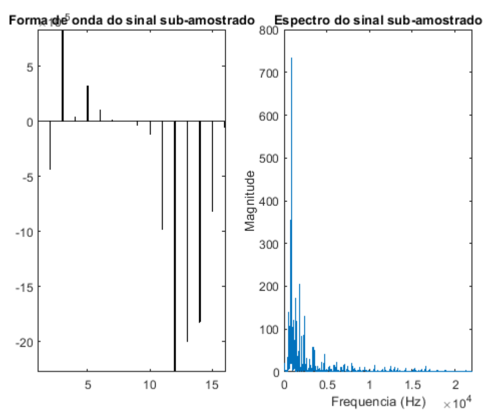

Figure 2.8: Waveform and spectrum for the sub-sampled sound

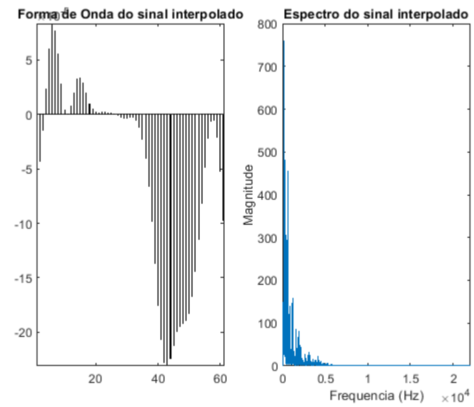

Figure 2.9: Waveform and spectrum for the interpolated sound

Comparing the above plots with the ones obtained in the previous experience, we can conclude that with the application of pre and interpolated filters, we obtain more represented values in the waveform, mostly on the negative vertical axis, which translates into a sound closer to the original one.

The following two pictures show the impact of both the pre and interpolated filters. On the left, we have the Impulse Response and, on the right, the Frequency Response.

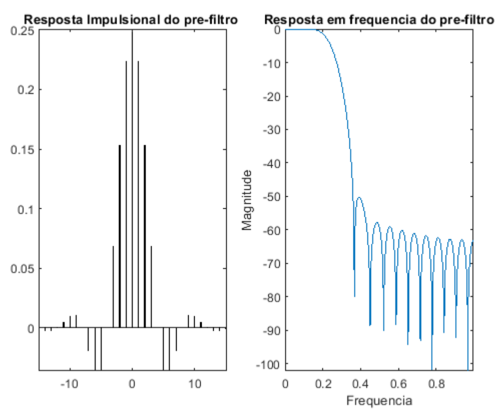

Figure 2.10: Impulse and frequency responses of the pre-filter

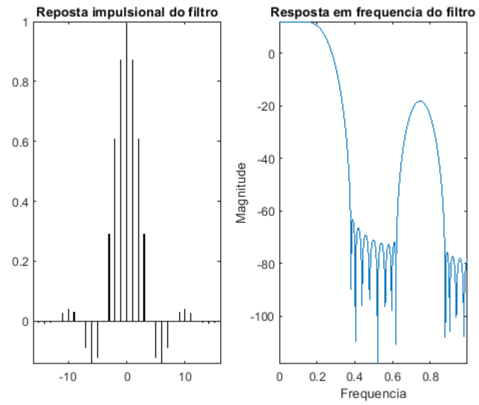

Figure 2.11: Impulse and frequency responses of the interpolated filter

For this experiment, the obtained mean squared error (from the comparison of the original sound with the interpolated one) was 1.3213×10^{-5} (PSNR of the interpolated sound $= 42.7715$), a much lower value than above.

2.2.2 Sub-sampling with a factor 2

Moving to the sub-sampling with a factor 2 (with filters), the variables associated to the samples have the same values as above, reason why they are not included here.

Starting by the perceptual quality, the sub-sampled sound is not 'distant' anymore and the interpolated sound is (at least as far as I could recognize) the same as the original one.

The following three plots show, as above, the waveform and the spectrum for the original, sub-sampled and interpolated sounds, respectively.

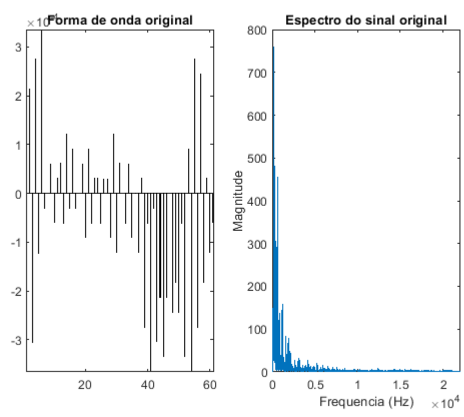

Figure 2.12: Waveform and spectrum for the original sound

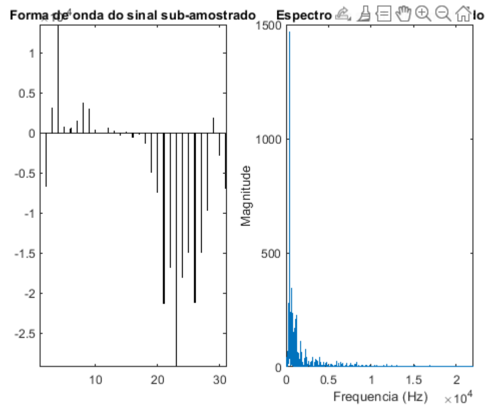

Figure 2.13: Waveform and spectrum for the sub-sampled sound

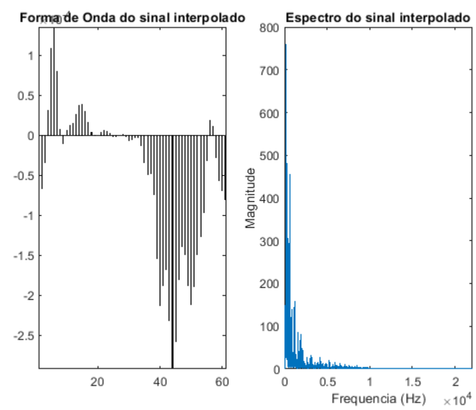

Figure 2.14: Waveform and spectrum for the interpolated sound

Again, the following two pictures show the impact of both the pre and interpolated filters. On the left, we have the Impulse Response and, on the right, the Frequency Response.

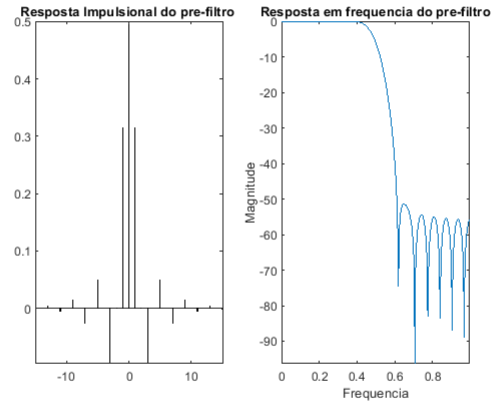

Figure 2.15: Impulse and frequency responses of the pre-filter

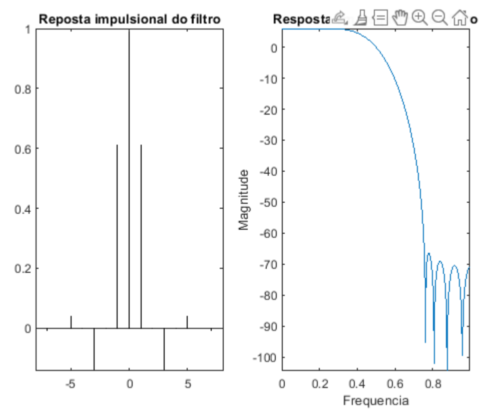

Figure 2.16: Impulse and frequency responses of the interpolated filter

For this experiment, the obtained mean squared error (from the comparison of the original sound with the interpolated one) was $5.366\,63 \times 10^{-6}$ (PSNR of the interpolated sound = 46.6845), a much lower value than above.

2.2.3 Conclusions

To start the conclusion of this two experiences, with and without the use of filters, the first major conclusion we can find is that, for the same value of k (4 and 2), the program that uses filters leads to better results.

Regarding the use of the filters, by the plots above, we can see that for k=4, the frequency response of both the pre and interpolated filters includes a higher variation in frequencies between 0.4 and 0.6 and above 0.8, while for k=2, those variations only occur for frequencies above 0.6/0.7. The reason for this is pretty intuitive: taking into consideration that k=4 leads to worse sub-sampled and interpolated sounds, the filters response are more diversified and impactful that for k=2.

The sampling frequency was 44,1KHz for both k=4 and k=2. According to the Nyquist frequency, 44.1KHz should be equal or higher than the double of the maximum frequency observed in the sound signal. That said, the ideal cut-off frequency of these filters should be equal or smaller than half of the Nyquist frequency, which means that should be equal or smaller than 22,05KHz.

3. Quantisation experiences

This third and last part of the report focus on quantisation experiences with 256 and 16 levels.

3.0.1 Quantisation with 256 levels

Starting with the perceptual quality analysis, with 256 levels (quantisation step: 0.0045) I didn't noticed any difference in comparison with the original sound.

The first of the following two graphics shows the signal of the original sound in red and the quantized one in blue, for the interval [0.0453515 - 0.0498866] seconds, for one of 500 samples obtained.

The second one is a 'zoom' of a part of the first one, showing 200 samples with the same color schema.

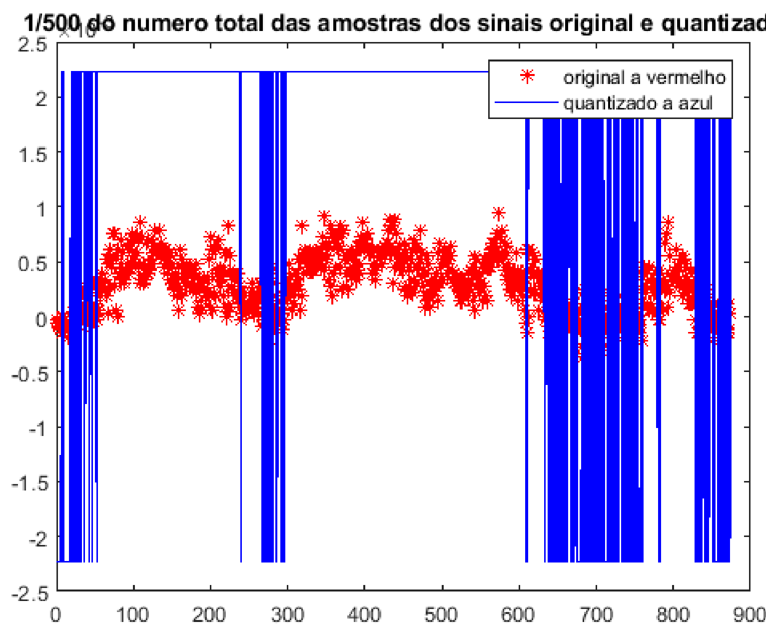

Figure 3.1: Signal in the interval [0.0453515 - 0.0498866] seconds

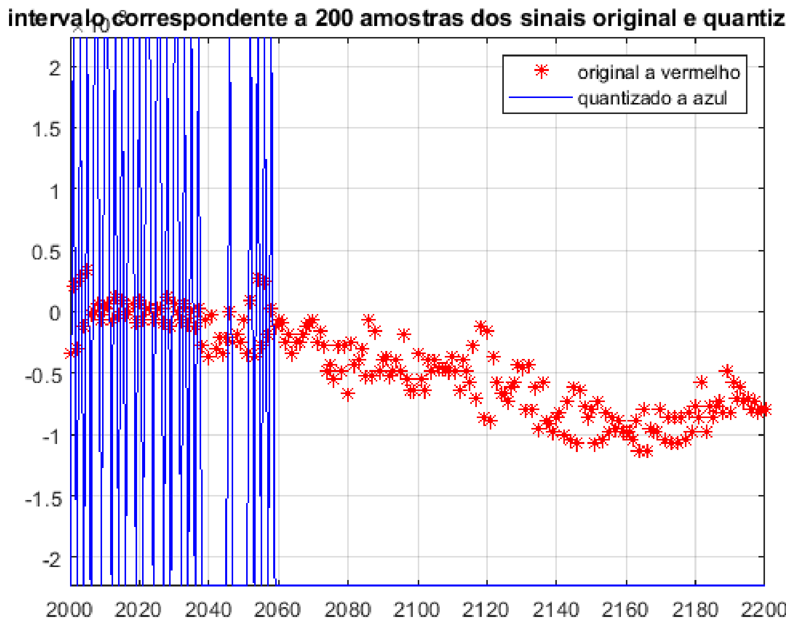

Figure 3.2: 'Zoom' of the previous plot with an interval composed by 200 samples

As expected, we can find zones where the original signal doesn't have a quantized representation.

For this experiment, the obtained mean squared error (from the comparison of the original sound with the interpolated one) was $1.676\,65 \times 10^{-6}$ (PSNR of the interpolated sound = 19.3097).

3.0.2 Quantisation with 16 levels

Moving to the quantisation with 16 levels (quantisation step: 0.0716) and stating with the perceptual quality analysis, I would say that the sound is now almost unrecognizable, having more or less 90% of noise and 10% sound, making it difficult to understand the original sound, which is imperceptible in the background.

The following two plots are the same as above, but for the quantisation with 16 levels.

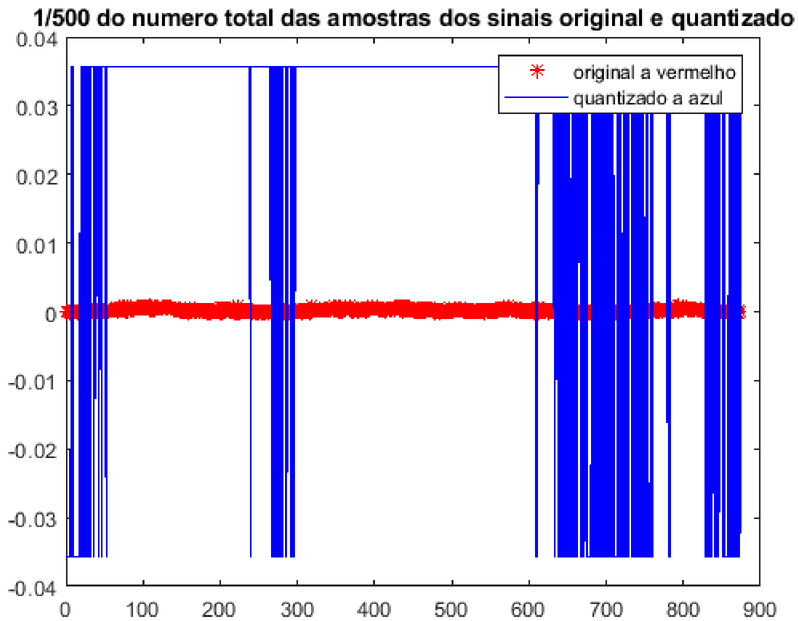

Figure 3.3: Signal in the interval [0.0453515 - 0.0498866] seconds

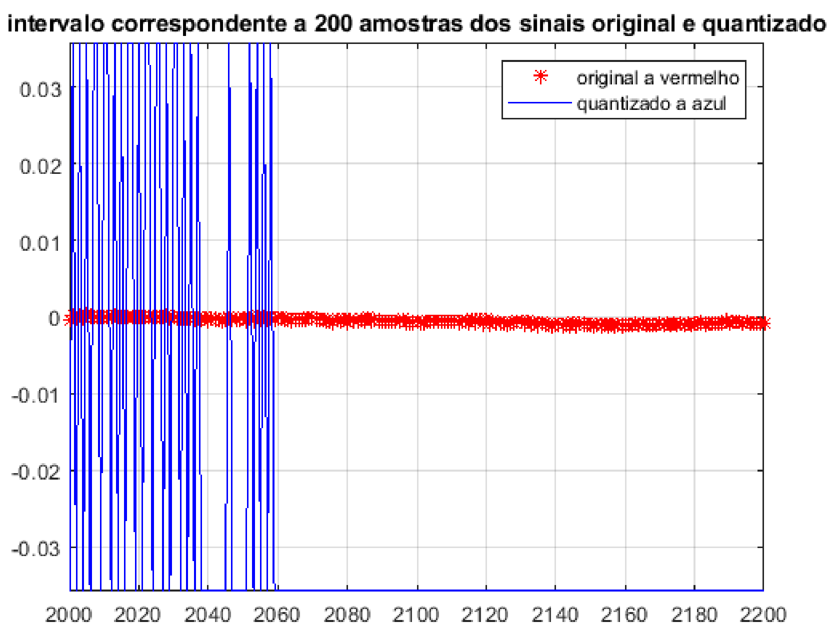

Figure 3.4: ‘Zoom’ of the previous plot with an interval composed by 200 samples

As above, we can find zones where the original signal doesn't have a quantized representation, but this time the range of those zones is higher.

For this experiment, the obtained mean squared error (from the comparison of the original sound with the interpolated one) was 0.000 434 877 (PSNR of the interpolated sound = -16.8707).