Homework 2 (Sample solution)

Exercise 1: Presentation in the time domain (3 Points)

Which sounds are voiced?

Periodic parts in the signal indicate voiced sounds. Each vertical line stands for one glottal stop. The intervals show the frequency of the glottal stops which is the fundamental frequency of the speaker. Vowels are generally voiced. Some consonants can also be voiced, e.g. the <d> in “dark”.

Which sounds are unvoiced?

Unvoiced sounds in a speech signal are non-periodic. In comparison to periodic sounds they almost look like random noise and usually have a smaller amplitude than periodic parts of a speech signal. E.g., the <s> in “is”, the <l> in “hill, or the <p> in “people”.

Find the sounds <a>, <o>, <u>, <sh>, <f>, and <k> in sentence.wav and compare them with the recordings of the individual sounds. Why do they look different? How is this effect called?

The articulatory process is not abrupt, in fact, it is a continuous process. Not individual sounds, but an airflow with continuous connected sounds is produced. The articulation organs are constantly moving, they don’t stay in defined positions. While pronouncing one sound they already move to the position for the following sound.
For very short vowels, the so called “duration dependent undershoot” can appear. There, the phase of the vowels is so short, that the position of the articulation organs for that vowel is not reached because the organs already move to the position of the next sound. That result in an overlap of the individual positions of the articulation vowels.
While articulating a sound, the sound always also shows features of the next or the preceding sound. In sum, every sound shows – more or less – attributes of the surrounding sounds. This effect is called **coarticulation**.
E.g., the word “shortcut” has a <sh> (non-priodic), a <o> (periodic) which overlap.

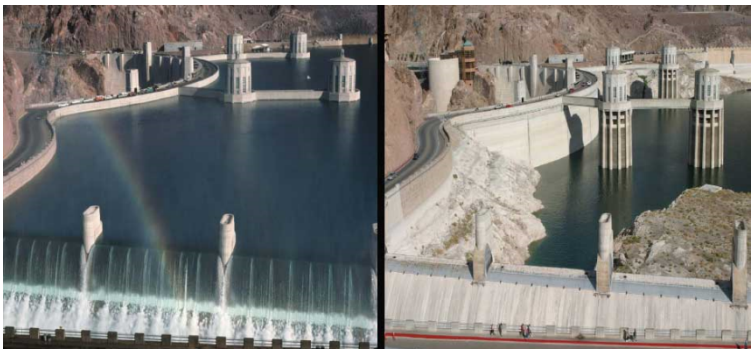
Exercice 2: LTI-Systems (8 Points)

a) **LTI-Systems:**

- **Air:** The speech signal of two persons speaking at the same time is a superposition of both speech signals. If the air would have non-linear characteristics, superposed speech of two speakers would sound distorted.
- **High pass filter:** The properties of a filter are independent from the input signal and the output is free from distortion. If the input signal is changed by the factor A, the output signal is changed by A as well.

Non LTI-Systems:

- **Dam:** The water reservoir behind a dam can keep water until the water line reaches the edge of the dam. From then on every drop that gets into the lake pours over the edge of the dam. This contradicts the linearity principle.



- **Loud speaker:** Loud speakers have non-linear characteristics which cause the total harmonic distortion effect ("Klirrfaktor").
- The human **vocal tract** is time-variant. Depending on the shape, e.g. different tongue position, of the vocal tract you produce different speech signals.
- The **position of the sun** is time invariant since it differs depending on the current day of the year.

b) **Practical Practice in Python**

Is the previous normalization process after the convolution an LTI system? Why or why not?

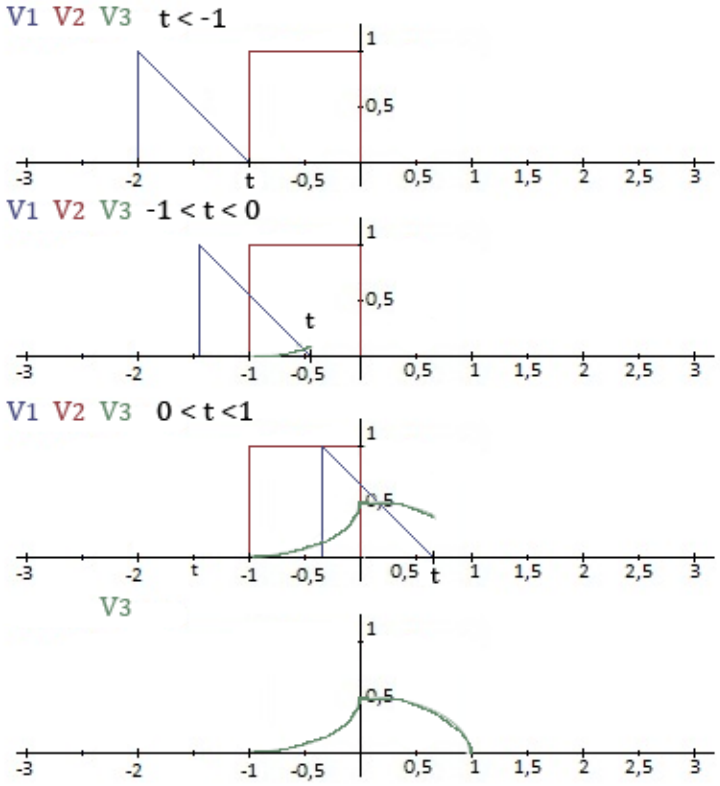
No, the function is not conform with the superposition principle. If we first summate two signals and then normalize them, we get a different result as if we would first normalize two signals and then summate them.

We are normalizing the result to avoid sample values above 1 or below -1. What is the reason that the result of a convolution can be above 1 or below -1 if both convolved signals are already bounded between -1 and 1?

In a convolution, two signals are multiplied and summated. Due to the summation values above 1 and below -1 can occur. Therefore, the result has to be normalized.

Exercise 3: Convolution and Fourier transformation (4 Points)

a)



b) **What does the Fourier transform achieve?**

With the Fourier transformation, a continuous signal can be transformed from the time- to the frequency-domain. The invers Fourier transformation can be used for the way back.

What is the difference between the Fourier transformation and the Fourier series?

A Fourier series can only be used for periodic signals. The frequency spectrum of a Fourier series is always discreet. The Fourier transformation can be used for non-periodic signals and the frequency spectrum can be continuous.

Fourier transformation

Shift theorem: $x(t - t_0) \multimap X(j\omega) \cdot e^{-i\omega t_0}$

So:

$$s(t - t_1) = a \cdot rect\left(\frac{t - t_1}{T}\right) \multimap S(j\omega) \cdot e^{-i\omega t_1} = aT \cdot si\left(\omega\frac{T}{2}\right) * e^{-i\omega t_1}$$