HW1 – deep learning in speech signals

# Part 1

## Q1:

First, we will replace the using Euler’s formula:

Now, using the DTFT formula:

We know that the DTFT of is

We can plug it to the formula below and get

## Q2

Based on the previous question, setting

We’ve already seen in the previous section the inner sum, so we can apply the exact calculations from before, and get

## Q3

# Part 2

## Q1

In the picture we can see 3 entire wavelengths. In addition, on the right side of the pink segment there is almost another wavelength and on the right side there is approximately a bit more than half a wavelength. So, we calculate that there are about 4.5 wavelengths.

Since the entire frame is 0.004279 seconds, each wavelength is seconds which means the pitch is

Question 2

1. Done.
   1. We looked at the waveforms below. Using praat we see that the pitch of the EH is 240 and the pitch of AO is 193 . Both by the blue line and by the frequenct of one wavelength written in the top of the screenshots.

A screen shot of a screen

Description automatically generatedA blurry image of a black and white image

Description automatically generated

* 1. Using praat, we can automatically see the formants as points – like in the screenshots below.  
     The frequency of the red dots in “EH” is 4500, 3500, 2800, 2000, 725 and in “AO” 703, 1300, 2700, 3700, 4700 . 5 formants - like Omer Adam

A yellow rectangular object with red dots

Description automatically generatedA screenshot of a computer

Description automatically generated

### Q3

F, J = 7, 8 (written in the forum)

I = 10 – we can see in the spectrogram the part where there is speech, silence and then white noise.

D – 9 since the beginning of spectrogram seems like noise which turns slowly into some more defined word.