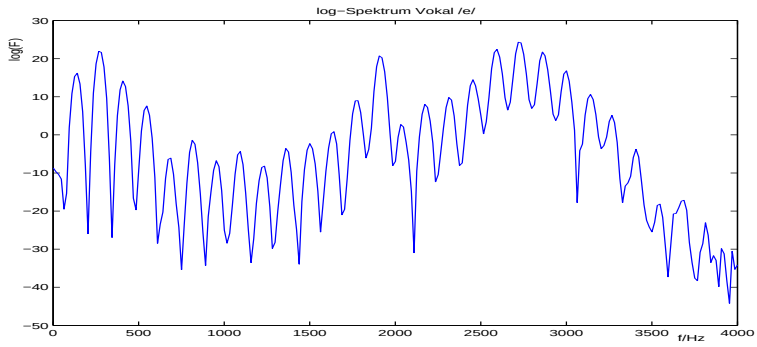
Speech Signal Processing

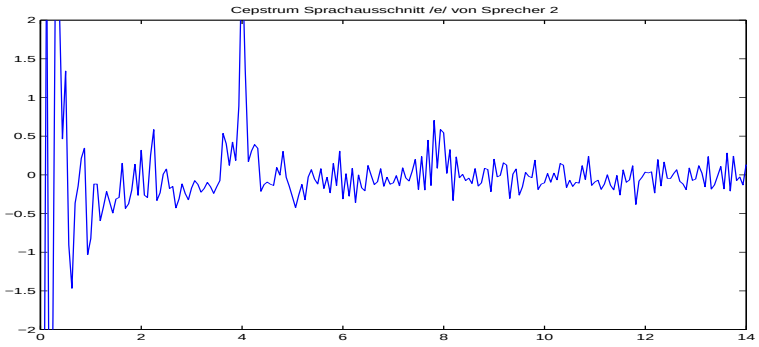
Exercise 3

1. Sampling , Windowing, DFT
   Given is a band limited speech signal $x(t)$ with upper cut-off frequency $f_g = 6kHz$.
   (a) What is the minimum sampling frequency f_a, with which perfect reconstruction is still possible?
   (b) What is the resolution in time domain, if we sample with f_a respectively $2f_a$?
   (c) Draft the spectrum of the sampled signal for the sampling frequencies f_a, $2f_a$ and $\frac{1}{2}f_a$.
   (d) What is the resolution in frequency domain, if we use a sampling frequency of f_a and a DFT-length of N=512?

2. Cepstrum
   (a) How can we make use of the *cepstrum* within speech processing?
   (b) The plot below shows the logarithm of the short-time spectral density of the vowel /e/ pronounced by speaker 1. Determine the fundamental frequency of the excitation.



   (c) In the next plot the *cepstrum* of the vowel /e/ of a second speaker is shown. Again, determine the fundamental frequency of the excitation.
   (d) Compare the fundamental frequencies of the two speakers.
   (e) Which range within the cepstrum is interesting for further signal processing?



3. Linear prediction

(a) For which purpose is *linear prediction* used in speech signal processing?

(b) Give the general predictor equation for the estimation of the signal $s[n]$?

(c) How does the number of predictor coefficients influence the LPC-spectrum?

(d) Over which time interval are the predictor coefficients estimated? (short-time stationarity of the speech signal)