EE 627 - Speech Signal Processing

Assignment # 2

1. What is the solution to the difference equation y [n] = K y[n-1] with initial condition y[0] = 1 and y[-1] = 0?Next, find the ROC of z-transform of your solution for y[n]. $K \leq 1$.

   y[n] = K^n

   $$Y(Z) = \frac{1}{1 - Kz^{-1}}$$

   The z-transform converges outside of a circle of radius K.

2. Consider a discrete-time signal x[n] passed through a bank of filters $h_k[n]$ where each filter is given by a modulated version of a baseband prototype filter h [n], i.e.,

   $h_k[n] = h[n]\exp[j(2n/N)kn]$

   where h[n], a Hamming window, is assumed causal and lies over a duration $0 \leq n < N_w$, and 2n / N is the frequency sampling factor. In this problem, you are asked to time-scale expand some simple input signals by time-scale expanding the filter bank outputs.

   a) State the constraint (with respect to the values N_w and N) such that the input x [n] is recovered when the filter bank outputs are summed.

   b) If the input to the filter bank is the unit sample δ [n], then the output of each filter is a complex exponential with "envelope" $a_k[n] = h[n]$ and phase $\theta_k[n] = (2n / N)kn$. Suppose each complex exponential output is time-expanded by two by interpolation of its envelope and phase. Derive a new constraint (with respect to values N_w and N), so that the summed filter bank outputs equal δ [n].

   c) Suppose now that the filter bank input equals

      $x[n] = \delta[n] + \delta[n - n_0]$,

      and that the filter bank outputs are time-expanded as in part (b). Derive a sufficient condition on N_w, N, and n_0 so that the summed filter bank output is given by

      $y[n] = \delta[n] + \delta[n - 2n_0]$,

      i,e., the unit samples are separated by $2n_0$ samples rather than n_0 samples.

3. Recall the “voiced” excitation to the digital model of speech production,

   $$e[n] = \sum_{q=-\infty}^{\infty} \delta[n - qP].$$

   Find the expression for the long term temporal autocorrelation, $r_e(\eta)$.

4. (MATLAB) In this MATLAB exercise,

   a) Perform the following operations with the vowel utterance of your choice:

      i. Compute the short term autocorrelation function for a hamming window of N=512 for $\eta = 0,1,2,\ldots,256$.
      ii. Compute the N=512 point magnitude spectrum of the waveform based on hamming window and stDFT.

(Note: The stDFT and conventional DFT are equivalent here because only the magnitude spectrum is required.)

iii. Repeat step (i) and (ii) after centre clipping the waveform.

b) Comment the changes in both the autocorrelation and the spectrum. What do these changes indicate about the effects of the clipping operations on the waveform?

c) Estimate the pitch using the two autocorrelation results. Which result would provide better performance in an autocorrelation procedure?