Homework 5 (ECE6255 Spring 2010) Grade=4/100

1. Consider an all-pole vocal tract transfer function of the form (make sure you understand the materials in Section 6.4 of Quatieri, this important topic will come back again in linear prediction, all-pole. analysis of speech signals):

$$V(z) = \frac{1}{\prod_{k=1}^{q}\left(1-c_k z^{-1}\right)\left(1-c_k^* z^{-1}\right)}$$

where $c_k = r_k e^{j\theta_k}$; $\quad |c_k| = r_k < 1$ for stability.

Show that the correponding cepstrum is

$$\hat{v}(n) = \begin{cases} 2\sum_{k=1}^{q} \frac{(r_k)^n}{n}\cos(\theta_k n) & n > 0 \\ 0 & \text{otherwise} \end{cases}.$$

2. Show that the following are equivalent: real cepstrum defined in (b) is equivalent to computing the even part of the complex cepstrum defined in (a), i.e. the real cepstrum discards the phase part of the complex cepstrum (make sure you understand the discussion in Section 6.3 of Quatieri).

(a) $\quad c(n) = \dfrac{\hat{x}(n)+\hat{x}(-n)}{2}$, with $\hat{x}(n)$ the complex cepstrum;

(b) $\quad$ equivalently, $c(n) = \dfrac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Re}[\hat{X}(\omega)]e^{j\omega n}$,

or $c(n) = \dfrac{1}{2\pi}\int_{-\pi}^{\pi} \log|X(\omega)|\, e^{j\omega n} d\omega$.

3. To smooth the log magnitude spectrum $\log|X(e^{j\omega})|$ of a signal $x(n)$, its cepstrum $c(n)$ is often windowed by a lifter $l(n)$, usually a rectangular window, in the cesptral domain to generate a modified cepstrum $c_l(n) = c(n)\bullet l(n)$ and then Fourier transformed to produced $\tilde{X}(e^{j\omega})$

(a) Since multiplication in the cepstral domain results in convolution of the Fourier transforms, if $l(n)$ has a Fourier Transform $L(e^{j\omega})$, write $\tilde{X}(e^{j\omega})$ in an expression relating the log magnitude spectrum and $L(e^{j\omega})$;

(b) To smooth $\log\left|X(e^{j\omega})\right|$, what type of cepstral window, or lifter $l(n)$, should be used?

(c) Compare the use of the rectangular cepstral window and a Hamming window in the time domain;

(d) How long should the cesptral window be? Why?

4. Note in Section 2.8.2 on Page 34 of Quatieri, when all the poles and zeros of a rational transfer function are inside the unit circle, the transfer function is called minimum phase. An example was given in Problem 1 above. The corresponding impulse response is referred to as a minimum phase sequence. On the other hand, we have a maximum phase transfer function is all the poles and zeros are outside the unit circle. In general, we have mixed phase rational transfer functions for the z-transform defined in Eq. (6.9) of Quatieri.

   If $x(n)$ is a minimum phase sequence of finite length N with complex cepstrum $\hat{x}(n)$, and $y(n) = \alpha^n x(n)$ with complex cepstrum $\hat{y}(n)$.

   (a) If $0 < \alpha < 1$; how will $\hat{y}(n)$ be related to $\hat{x}(n)$?

   (b) Let $\max_i |z_i|$ denote the magnitude of the maximum magnitude zero of $X(z)$, how should α be chosen so that $y(n)$ is no longer minimum phase?

   (c) Let $\min_i |z_i|$ denote the magnitude of the minimum magnitude zero of $X(z)$, how should α be chosen so that $y(n)$ is maximum phase?

5. Write a MATLAB program to compute the complex and real cepstra of the following signals:

   1) $x_1(n) = \delta(n) - .85\delta(n-99)$

   2) $x_2(n) = \sin(\frac{2\pi n}{100}) \quad 0 \le n \le 99$

   3) $x_3(n) = 0.95^n \quad 0 \le n \le 99$

   4) a section of voiced speech

   5) a section of unvoiced speech

   For each of the first 3 signals, plot the waveform, the complex cepstrum, and the real cepstrum. For the speech signals (the last two signals), plot the signal, the log magnitude spectrum, the real cepstrum, and the lowpass liftered log magnitude spectrum (on top of the log magnitude spectrum to show the smoothing effect). For the real speech signals, take segments of duration with more than 400 samples each. Use a Hamming window before computing the cepstra of the speech files. Use the file 1.wav to test your program. Do the same speech analysis on 5.wav, state your observations.

6. **Extra Credit** (+25%, 1/100 in total): For your signals in parts (4) and (5) above compute log magnitude and the lowpass liftered log magnitude spectra, and real cepstra for a long section (or even for the entire utterance) containing multiple sounds. Use the displays shown on Slide 60 in Lecture 16-18. Explain the result.