UCLA — Electrical Engineering Dept.
EE114: Speech and Image Processing
Midterm Exam
Wednesday, February 6, 2013

Solutions

1. Consider a continuous-time signal consisting of a windowed cosine at F_0 Hz,

$$x_{\mathrm{a}}(t) = \begin{cases} \cos(2\pi F_0 t), & 0 \leq t < 1/F_0, \\ 0, & \text{elsewhere}, \end{cases}$$

sampled at a sampling rate of $F_s = 1/T = 4F_0$ Hz.

(a) (5 points) Find and sketch the sampled signal, $x(n) = x_{\mathrm{a}}(nT)$, n integer.

> **Solution:**
> $$x(n) = [\underline{1}, 0, -1, 0].$$

(b) (5 points) Find $X(e^{j\omega})$, the DTFT of $x(n)$.

> **Solution:**
> $$X\left(e^{j\omega}\right) = \sum_{n=0}^{3} x(n)\, e^{-jn\omega} = 1 - e^{-j2\omega}.$$

(c) (5 points) Find $X(k)$, the 4-point DFT of $x(n)$.

> **Solution:**
> $$X(k) = X\left(e^{j\omega}\right)\Big|_{\omega=2\pi k/4} = 1 - e^{-jk\pi} = [\underline{0}, 2, 0, 2].$$

(d) (5 points) Using $X(k)$, evaluate the inverse DFT, $\hat{x}(n)$. Explain why $\hat{x}(n) = x(n)$ or $\hat{x}(n) \neq x(n)$ for $-\infty < n < \infty$.

> **Solution:**
> $$\hat{x}(n) = \frac{1}{4}\sum_{k=0}^{3} X(k)\, e^{-j2\pi nk/4} = [\ldots, \underline{1}, 0, -1, 0, \ldots],$$
> periodic of period $N = 4$, therefore $\hat{x}(n) = x(n)$ only for $0 \leq n \leq 3$, but $\hat{x}(n) \neq x(n)$ otherwise.

Note that $x_{\mathrm{a}}(t) = w_{\mathrm{a}}(t)\, c_{\mathrm{a}}(t)$, where $w_{\mathrm{a}}(t) = 1$, $0 \leq t < 1/F_0$, and zero elsewhere, and $c_{\mathrm{a}}(t) = \cos(2\pi F_0 t)$, $-\infty < t < \infty$.

(e) (5 points) Find the DTFT of $w(n) = w_{\mathrm{a}}(nT)$, $W(e^{j\omega})$.

Solution: $w(n) = [\underline{1}, 1, 1, 1]$, so:

$$W\left(e^{j\omega}\right) = \sum_{n=0}^{3} w(n)\, e^{-jn\omega} = 1 + e^{-j\omega} + e^{-j2\omega} + e^{-j3\omega}.$$

(f) (5 points (bonus)) Find the DTFT of $c(n) = c_{\mathrm{a}}(nT)$ using the DTFT definition.

Solution: $c(n) = [\ldots, \underline{1}, 0, -1, 0, \ldots]$, periodic of period $N = 4$. Also, remember that:

$$\sum_{n=-\infty}^{\infty} e^{-jn\omega} = 2\pi \sum_{k=-\infty}^{\infty} \delta(\omega - 2\pi k).$$

$$\begin{aligned} C\left(e^{j\omega}\right) &= \sum_{n=-\infty}^{\infty} c(n)\, e^{-jn\omega} = \sum_{n=-\infty}^{\infty} e^{-j4n\omega} - \sum_{n=-\infty}^{\infty} e^{-j(2+4n)\omega} \\ &= 2\pi \sum_{k=-\infty}^{\infty} \delta(4\omega - 2\pi k)\left(1 - e^{-j2\omega}\right) = \frac{\pi}{2} \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{\pi}{2}k\right)\left(1 - e^{-j2\omega}\right) \\ &= \frac{\pi}{2} \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{\pi}{2}k\right)\left(1 - e^{-j\pi k}\right) = \pi \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{\pi}{2}(2k-1)\right). \end{aligned}$$

For $-\pi \leq \omega < \pi$:

$$C\left(e^{j\omega}\right) = \pi\delta\left(\omega + \frac{\pi}{2}\right) + \pi\delta\left(\omega - \frac{\pi}{2}\right).$$

(g) (5 points (bonus)) Find the DTFT of $x(n) = x_{\mathrm{a}}(nT)$ using the answers to parts (e) and (f) and the properties of the DTFT, and compare with your answer to part (b).

Solution:

$$\begin{aligned} X\left(e^{j\omega}\right) &= \frac{1}{2\pi} C\left(e^{j\omega}\right) * W\left(e^{j\omega}\right) \\ &= \frac{1}{2}\left(W\left(e^{j(\omega+\pi/2)}\right) + W\left(e^{j(\omega-\pi/2)}\right)\right) \\ &= 1 - e^{-j2\omega}. \end{aligned}$$

2. Consider the vocal tract in Fig. 1. Suppose the sound source (glottal excitation) is white noise introduced at the point indicated by an asterisk in the model. The lips are at the far right side of the model. Given are the cross-sectional area of the back cavity, $A_b = 24$ cm^2, and its length, $l_b = 12$ cm, and the cross-sectional area of the constriction, $A_c = 2$ cm^2, and its length, $l_c = 4$ cm. (Note: $A_c \ll A_b$. Also, remember that $c = 35000$ cm/s.)

   (a) (5 points) Calculate the two lowest resonance frequencies of the back cavity, $F_1^{(b)}$ and $F_2^{(b)}$.

   Solution: Both tubes are half-wavelength tubes (the back tube is closed at both ends and the front tube is open at both ends), so we use the formula: $F_n = \frac{nc}{2l}$ with $n = 1, 2$ and $l = l_b$:

$$F_1^{(b)} = \frac{c}{2l_b} = \frac{35000}{2 \times 12} \approx 1458.3 \text{ Hz}, \quad F_2^{(b)} = 2F_1 \approx 2916.7 \text{ Hz}.$$

   (b) (5 points) Calculate the two lowest resonance frequencies of the constriction, $F_1^{(c)}$ and $F_2^{(c)}$.

   Solution: Using the same formula as before:

$$F_1^{(c)} = \frac{c}{2l_c} = \frac{35000}{2 \times 4} = 4375 \text{ Hz}, \quad F_2^{(c)} = 2F_1 = 8750 \text{ Hz}.$$

   (c) (5 points) Calculate the Helmholtz resonance, F_{H}.

   Solution:

$$F_{\mathrm{H}} = \frac{c}{2\pi}\sqrt{\frac{A_c}{A_b l_b l_c}} = \frac{35000}{2\pi}\sqrt{\frac{2}{24 \times 12 \times 4}} = \frac{35000}{48\pi} \approx 232.1 \text{ Hz}.$$

   (d) (5 points) What is the first formant of the vowel modeled by this tube configuration?

   Solution: The first formant is the lowest of the resonance frequencies, i.e., F_{H}.

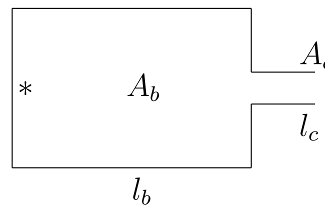


Figure 1: Vocal tract

Table 1: First three formants of vowels AA, IY, and UW.

	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
AA	570	930	2600
IY	180	1960	3320
UW	190	530	2240

3. The three vowels AA, IY, and UW, are spoken sequentially in a specific order by one speaker. Table 1 shows the formants of the three vowels for this specific speaker. Two spectrograms of the utterance of the three vowels are given in Fig. 2.

   (a) (5 points) What is the sequence in which the vowels were uttered? Justify your answer on the basis of the spectrograms.

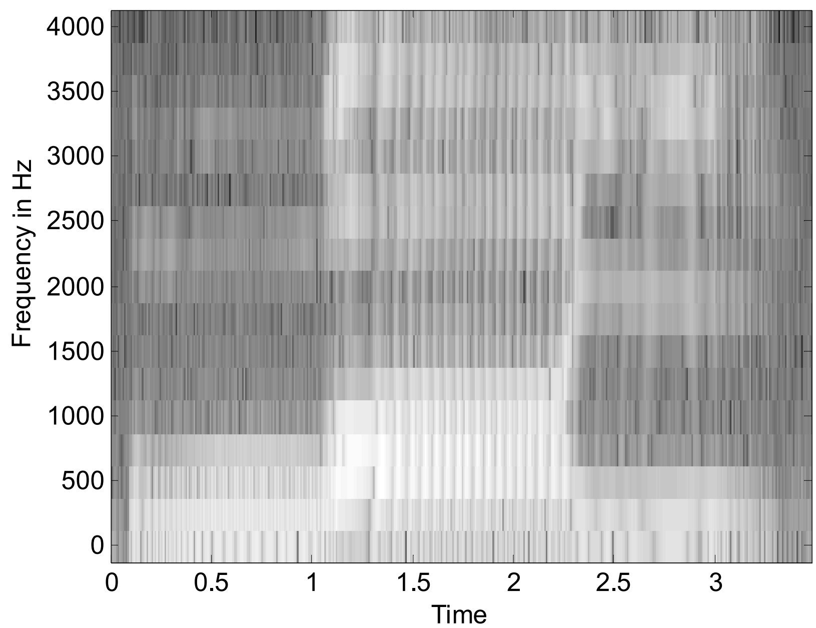
   Solution: It is clear from the wideband spectrogram (spectrogram 1) that the first vowel is UW, the second vowel is AA, and the third one is IY. The wideband spectrogram allows us to see the formants of the three vowels.

   (b) (5 points) Describe qualitatively (and briefly) the two spectrograms. Describe what you see.

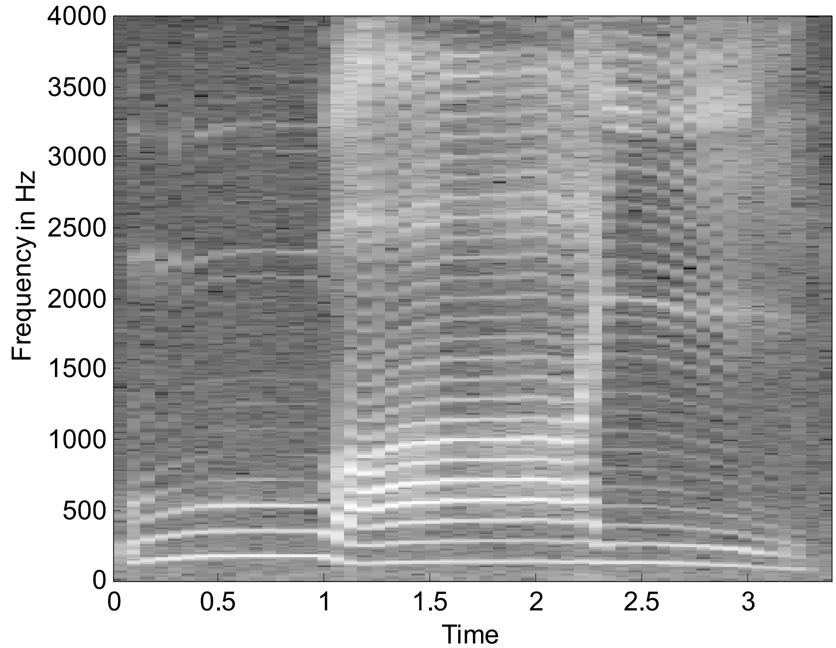
   Solution: Both spectrograms are time-frequency representations of the three-vowel sequence. In spectrogram 1 pitch and harmonics cannot be resolved, but formants can. In spectrogram 2 pitch and harmonics can be clearly detected.

   (c) (5 points) Which of the two spectrograms is obtained with a narrowband window and which with a wideband window? Explain your answer.

   Solution: Spectrogram 1 is a wideband spectrogram, while spectrogram 2 is a narrowband spectrogram.



(a) Spectrogram 1.



(b) Spectrogram 2.

Figure 2: Spectrograms of a sequence of the three vowels AA, IY, and UW.

(d) (5 points) What is the approximate pitch frequency?

Solution: From the narrowband spectrogram, we see that the pitch frequency is around 100 Hz.

(e) (5 points) Is the speaker most likely male or female? (Remember that male pitch is from 80 to 200 Hz, female pitch is from 150 to 350 Hz.)

Solution: Most likely the speaker is male.

4. Consider the signal:

$$x(n) = \alpha^n \cos(\pi n), \quad 0 \le n \le N-1, 0 < \alpha < 1.$$

Let $R(k)$ be the autocorrelation function of $x(n)$.

(a) (5 points) Find $R(0)$.

Solution: $x(n) = (-1)^n \alpha^n$, $0 \le n \le N-1$, so:

$$R(0) = \sum_{n=0}^{N-1} x^2(n) = \sum_{n=0}^{N-1} [(-\alpha)^n]^2 = \sum_{n=0}^{N-1} \alpha^{2n} = \frac{1-\alpha^{2N}}{1-\alpha^2}.$$

(b) (5 points) Find $R(1)$.

Solution:

$$\begin{aligned} R(1) &= \sum_{n=1}^{N-1} x(n)\, x(n-1) = \sum_{n=1}^{N-1} (-\alpha)^n (-\alpha)^{n-1} \\ &= -\frac{1}{\alpha} \sum_{n=1}^{N-1} \alpha^{2n} = -\frac{1}{\alpha}\left(\frac{1-\alpha^{2N}}{1-\alpha^2} - 1\right) = -\alpha \frac{1-\alpha^{2N-2}}{1-\alpha^2}. \end{aligned}$$

(c) (5 points) Using the autocorrelation method of linear prediction analysis, find the 1st-order prediction coefficient, a_1.

Solution:

$$a_1 = \frac{R(1)}{R(0)} = -\alpha \frac{1-\alpha^{2N-2}}{1-\alpha^2} \frac{1-\alpha^2}{1-\alpha^{2N}} = -\alpha \frac{1-\alpha^{2N-2}}{1-\alpha^{2N}}.$$

(d) (5 points) Find the corresponding error, $E_{\min}$.

Solution:

$$E_{\min} = R(0) - a_1 R(1) = \frac{1-\alpha^{2N}}{1-\alpha^2} - \alpha^2 \frac{1-\alpha^{2N-2}}{1-\alpha^{2N}} \frac{1-\alpha^{2N-2}}{1-\alpha^2}.$$

(e) (5 points) Find a_1 and $E_{\min}$ as $N \to \infty$.

Solution: From parts (c) and (d) we have:

$$a_1 \xrightarrow[N\to\infty]{} -\alpha, \quad E_{\min} \xrightarrow[N\to\infty]{} 1.$$

(f) (5 points) Assume $N \to \infty$ and let $\alpha = 0.5$. Find the variables a_1 and G for the transfer function of the linear predictive model:

$$H(z) = \frac{G}{1 - a_1 z^{-1}}.$$

Sketch the pole-zero plot.

Solution: $G = \sqrt{E_{\min}}$, so:

$$H(z) = \frac{G}{1 - a_1 z^{-1}} = \frac{1}{1 + 0.5z^{-1}},$$

with a pole at $z = -0.5$.