UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Electrical and Computer Engineering

ECE 537 Speech Processing
Fall 2009

EXAM 1

Thursday, October 8, 2009

- This is a CLOSED BOOK exam, but one $8\frac{1}{2}'' \times 11''$ sheet of hand-written notes (both sides) is allowed.
- There are a total of 100 points in the exam. Each problem specifies its point total. Plan your work accordingly.
- You must SHOW YOUR WORK to get full credit.
- Please be neat—we can't grade what we can't decipher.

Problem	Score	Problem	Score	Total
Name		4		
1		5		
2		6		
3		7		
Total		Total		

Name (1 point): ____

LPC Parameter Sets Containing Equivalent Information

Direct-Form Coefficients (a_k)

$$R_n(i) = \sum_{k=1}^{p} a_k R_n(|i-k|) \quad (1)$$

$$A(z) = 1 - \sum_{k=1}^{p} a_k z^{-k} \quad (2)$$

LPC Roots (r_i)

$$r_i = \text{roots}(A(z)) \quad (3)$$

$$A(z) = \prod_{i=1}^{p} (1 - r_i z^{-1}) \quad (4)$$

PARCOR Coefficients (k_i)

$$\begin{aligned} k_i &= a_i^{(i)} \\ a_j^{(i-1)} &= \frac{a_j^{(i)} + k_i a_{i-j}^{(i)}}{1-k_i^2}, \quad 1 \le j \le i-1 \end{aligned}$$

$$a_j^{(i)} = \begin{cases} k_i & j = i \\ a_j^{(i-1)} - k_i a_{i-j}^{(i-1)} & 1 \le j \le i-1 \end{cases} \quad (5)$$

Log Area Ratios (g_i)

$$g_i = \log\left(\frac{1-k_i}{1+k_i}\right) \quad (6)$$

$$k_i = \frac{1-e^{g_i}}{1+e^{g_i}} \quad (7)$$

Line Spectral Frequencies (p_n, q_n)

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1}) \quad (8)$$

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1}) \quad (9)$$

$$p_n = \arg(\text{roots}(P(z))) \quad \text{s.t.} \quad 0 < p_n < \pi$$

$$q_n = \arg(\text{roots}(Q(z))) \quad \text{s.t.} \quad 0 < q_n < \pi$$

$$P(z) = (1+z^{-1}) \prod_{n=1}^{p/2} (1-e^{jp_n} z^{-1})(1-e^{-jp_n} z^{-1}) \quad (10)$$

$$Q(z) = (1-z^{-1}) \prod_{n=1}^{p/2} (1-e^{jq_n} z^{-1})(1-e^{-jq_n} z^{-1}) \quad (11)$$

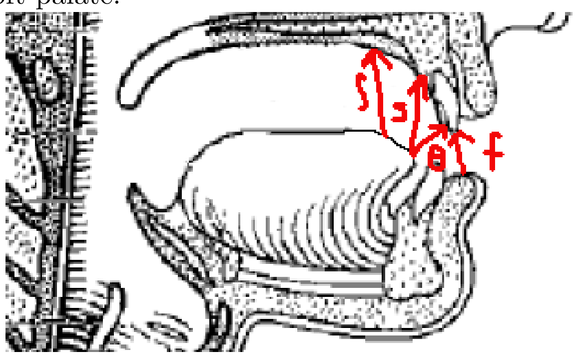
LPC Cepstrum (c_m)

$$\begin{aligned} c_0 &= \log G^2 \\ c_m &= a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k}, \quad 1 \le m \le p \\ c_m &= \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k}, \quad m > p \end{aligned}$$

$$\begin{aligned} G &= e^{c_0/2} \\ a_m &= c_m - \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k}, \quad 1 \le m \le p \end{aligned}$$

Problem 1 (12 points)

(a) Here's a midsagittal section of the head, showing the articulators of speech production. Draw four arrows on this picture, one for each of the four unvoiced fricatives of English (they are [f,θ,s,ʃ]). Each arrow should be labeled with the IPA symbol for one fricative. It should start at the articulator that produces the named fricative, and it should end at the place touched by that articulator in order to produce the fricative. For example, if [k] were a fricative, it would be portrayed by an arrow that starts on the tongue body, and ends on the soft palate.



(b) The following table lists five pairs of phonemes, and five different speech articulators. If the two phonemes in the ith phone pair both use the same position of the jth articulator, put an "**S**" in box i, j; if they use a different position of the jth articulator, put a "**D**" in box i, j.

Articulator	[d] vs. [n]	[d] vs. [t]	[d] vs. [r]	[d] vs. [b]	[i] vs. [ɑ]
Lips	S	S	S	D	S
Tongue Tip	S	S	D	D	S
Tongue Body	S	S	S	S	D
Velum	D	S	S	S	S
Glottis	S	D	S	S	S

Problem 2 (12 points)

(a) The low-frequency linearization of the relationship between pressure, p, and volume velocity, U, for a tube of length l and area A open at both ends is

$$p = j\omega LU, \quad L = \frac{\rho_0 l}{A}$$

Give the MKS units for each of the following variables:

- p: Solution: [Pa] $= [N/m^2] = [kg/m \cdot s^2]$
- ω: Solution: [radians/s] $= [1/s]$
- ρ_0: Solution: $[kg/m^3]$
- l: Solution: $[m]$
- A: Solution: $[m^2]$
- L: Solution: $[kg/m^4]$
- U: Solution: $[m^3/s]$

(b) The low-frequency linearization of the relationship between pressure and volume velocity at the open end of a tank (a tube closed at the opposite end) is

$$p = U/sC, \quad C = \frac{lA}{\rho_0 c^2}$$

Demonstrate that U/sC is measured in units of Pascals.

Solution: $U = [m^3/s]$, $s = [1/s]$, $C = [m^3/(kg/m^3)(m^2/s^2)] = [m^4 s^2/kg]$, so U/sC has units of $[(m^3/s)/(1/s)(m^4 s^2/kg)] = [kg/m \cdot s^2] =$ [Pa].

Problem 3 (15 points)

A tube of length l and cross-sectional area A is terminated, at $x = 0$, by a radiation impedance $Z_R(\omega)$.

(a) Assume that $Z_R(\omega)$ is the complex-valued radiation impedance of an ideal hemispherical source of radius a and area A. Write $Z_R(\omega)$ in terms of a, A, ω, ρ_0, and c.

Solution:

$$Z_R(\omega) = \frac{\rho_0 c}{A}\left(\frac{j\omega a/c}{1 + j\omega a/c}\right)$$

(b) Pressure and volume velocity at position x within the tube can be written as

$$\begin{aligned} P(x,\omega) &= Z_0(U_+(\omega)e^{-j\omega x/c} + U_-(\omega)e^{j\omega x/c}) \\ U(x,\omega) &= U_+(\omega)e^{-j\omega x/c} - U_-(\omega)e^{j\omega x/c} \end{aligned}$$

Find $\Gamma(\omega) = U_-(\omega)/U_+(\omega)$ in terms of $Z_R(\omega)$ and $Z_0 = \rho_0 c/A$.

Solution:

$$\Gamma = \frac{Z_r - Z_0}{Z_r + Z_0}$$

(c) The specific impedance looking into the tube from any position, x, is given by

$$Z(x,\omega) = \frac{P(x,\omega)}{U(x,\omega)}$$

Write $Z(x,\omega)$ in terms of $\Gamma(\omega)$, x, ω, and c.

Solution:

$$Z(x,\omega) = Z_0 \frac{1 + \Gamma(\omega)e^{2j\omega x/c}}{1 - \Gamma(\omega)e^{2j\omega x/c}}$$

Problem 4 (15 points)

A particular Helmholtz resonator is made up of a back cavity closed at one end, with area A_b and length l_b, and a front cavity open at both ends, with area A_f and length l_f, $A_b \gg A_f$. Assume that the radiation impedance is $Z_R = 0$.

(a) The resonant frequencies of this system are denoted ω_n, for $n = 1, 2, 3, \ldots$. Write an equation in terms of $\omega, l_f, l_b, A_f, A_b, \rho_0$, and c that is satisfied at the resonant frequencies ($\omega = \omega_n$, $n \in \{1, 2, 3, \ldots\}$) and not at any other frequency.

Solution:

$$j\frac{\rho_0 c}{A_b}\cot\left(\frac{\omega l_b}{c}\right) - j\frac{\rho_0 c}{A_f}\tan\left(\frac{\omega l_f}{c}\right) = 0$$

(b) The equation you wrote in part (a) can be linearized, at low frequencies, as

$$\frac{1}{\omega C} - \omega L = 0$$

What are L and C?

Solution:

$$L = \frac{\rho_0 l_f}{A_f}, \quad C = \frac{l_b A_b}{\rho_0 c^2}$$

(c) Suppose the back cavity volume is $l_b A_b = 256\text{cm}^2$ (about a quarter of a liter). For what values of l_f and A_f does this system have a resonant frequency of $\frac{\omega_1}{2\pi} \leq 250\text{Hz}$? Assume $c \approx 350 m/s$ at body temperature.

Solution:

$$\frac{A_f}{l_f} \leq A_b l_b \left(\frac{500\pi}{c}\right)^2 = 5.1\text{mm}$$

Problem 5 (15 points)

Based on $P = Z_0(U_+ + U_-)$ and $U = U_+ - U_-$ it is possible to show that the pressure and volume velocity at the right end of a pipe, p_2 and U_2, are related to the pressure and volume velocity at its left end, p_1 and U_1, by

$$\left[\begin{array}{c} p_2 \\ U_2 \end{array}\right] = \left[\begin{array}{cc} \cos(kl) & -jZ_0\sin(kl) \\ -jY_0\sin(kl) & \cos(kl) \end{array}\right]\left[\begin{array}{c} p_1 \\ U_1 \end{array}\right], \tag{12}$$

where $k = \omega/c$, $Z_0 = 1/Y_0 = \rho_0 c/A$, A is the cross-sectional area of the pipe, and l is its length.

(a) Find a second-order Taylor series approximation of Eq. 12, valid for sufficiently small ω.

Solution:

$$\left[\begin{array}{c} p_2 \\ U_2 \end{array}\right] = \left[\begin{array}{cc} 1-\frac{1}{2}(kl)^2 & -jZ_0 kl \\ -jY_0 kl & 1-\frac{1}{2}(kl)^2 \end{array}\right]\left[\begin{array}{c} p_1 \\ U_1 \end{array}\right]$$

(b) For what frequencies is the second-order Taylor series approximation (part a) valid?

Solution: The second-order approximation for a cosine is valid for $|kl| \ll \frac{\pi}{2}$, or $|\omega| \ll \frac{\pi c}{2l}$, or $|f| \ll \frac{c}{4l}$. In fact, the approximation is not too bad even at $kl = \pi/2$: $1-\frac{1}{2}(\frac{\pi}{2})^2 = -0.23$. At $kl = \pi/4$, the error is very small: $1-\frac{1}{2}(\frac{\pi}{4})^2 = 0.6917$, whereas $\cos(\pi/4) = 1/\sqrt{2} = 0.707$.

Problem 6 (15 points)

The cepstrum of speech is a rapidly decaying sequence (it can be shown that $|c_x[m]| < \frac{1}{m}e^{-am}$), therefore it is hard to build good probabilistic models of $c_x[m]$. Automatic speech recognition combats cepstral decay using anti-tapered windows, i.e., windows that are small in the center and large at the edges—the opposite of a Hamming window. The easiest such window to analyze (though not the most common) is the anti-Hamming window,

$$c_y[m] = \begin{cases} c_x[m]\left(0.54 - 0.46\cos(\frac{2\pi m}{N-1})\right) & -(\frac{N-1}{2}) \le m \le (\frac{N-1}{2}) \\ 0 & \text{else} \end{cases} \tag{13}$$

(a) Make the following definitions:

$$C_x(\omega) \equiv \ln|X(e^{j\omega})| \tag{14}$$

$$C_y(\omega) \equiv \sum_{m=-(N-1)/2}^{(N-1)/2} c_y[m]e^{-j\omega m} \tag{15}$$

Under these definitions, $C_y(\omega) = \frac{1}{2\pi}W(\omega) * C_x(\omega)$. Find $W(\omega)$. You may write your answer in terms of $W_R(\omega) = \sin(\omega N/2)/\sin(\omega/2)$.

Solution:

$$W(\omega) = 0.54W_R(\omega) - 0.23W_R(\omega - \frac{2\pi}{N-1}) - 0.23W_R(\omega + \frac{2\pi}{N-1})$$

(b) Devise an algorithm that accepts $c_x[m]$ as input, and generates $C_y(\omega)$ as output, but requiring about half as many multiplications as the algorithm specified in Eqs. 13 and 15. Your answer should contain at least two equations: one equation to replace Eq. 13, and one equation to replace Eq. 15.

Solution: $c_x[m]$, $c_y[m]$, and $C_y(\omega)$ are all real, even sequences, therefore

$$C_y(\omega) = c_y[0] + 2\sum_{m=1}^{(N-1)/2} c_y[m]\cos(\omega m)$$

$$c_y[m] = \begin{cases} c_x[m](0.54 - 0.46\cos(\frac{2\pi m}{N-1})) & 0 \le m \le \frac{N-1}{2} \\ c_y[-m] & -\frac{N-1}{2} \le m \le -1 \\ 0 & \text{otherwise} \end{cases}$$

This algorithm requires about half as many multiplications as does Eq. 15, but the computational savings is a factor of four rather than a factor of two: this algorithm uses $N/2$ real-valued multiplications to compute each frequency sample $C_y(\omega)$, whereas Eq. 15 uses N complex multiplications to compute each frequency sample.

If the entire spectrum is desired (at least N frequency samples), then much greater computational reductions are achievable by replacing Eq. 15 with an FFT. Since $c_y[m]$ and $C_y(\omega)$ are both real, even sequences, a completely real version of the FFT can be used, whose computational complexity is $\mathcal{O}\{N \log_2 N\}$ real multiplications (as opposed to the usual $\mathcal{O}\{N \log_2 N\}$ complex multiplications).

Problem 7 (15 points)

A continuous-time filter $h(t)$ is to be modeled using the impulse-invariant transform $h[n] = h(nT)$ for sampling time $T = 1/F_s$, where

$$h(t) = e^{-\pi B_1 t} \sin(2\pi F_1 t) u(t)$$

$H(z)$ can be written as

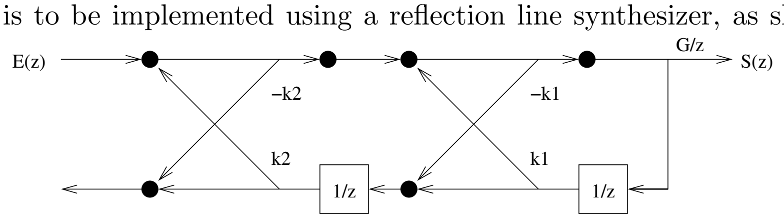
$$H(z) = \frac{Gz^{-1}}{A(z)}, \quad A(z) = 1 - a_1 z^{-1} - a_2 z^{-2}$$

(a) Find a_1, a_2, and G in terms of $\sigma_1 = \pi B_1 T$ and $\omega_1 = 2\pi F_1 T$.

Solution:

$$\begin{aligned} G &= e^{-\sigma_1} \sin(\omega_1) \\ a_1 &= 2e^{-\sigma_1} \cos(\omega_1) \\ a_2 &= -e^{-2\sigma_1} \end{aligned}$$

(b) The filter $H(z)$ is to be implemented using a reflection line synthesizer, as shown here:



Write the reflection coefficients k_1, k_2 in terms of the direct-form coefficients a_1, a_2.

Solution:

$$\begin{aligned} k_2 &= a_2 \\ k_1 &= a_1/(1 - a_2) \end{aligned}$$