UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Electrical and Computer Engineering
Instructor: Mark Hasegawa-Johnson

ECE 537 Speech Processing

Problem Set 3

Fall 2009

Issued: Wed Sep. 16, 2009 **Due:** Wed Sept. 23, 2009

Reading for problem set 3: Flanagan, Allen & Hasegawa-Johnson 3.1-5

Problem 3.1

Fig. 3.3 in the textbook shows two different lumped-element approximations for the wave equation in a short section of tube. These two approximations are valid when the tube is so short that $\tan kl \approx sL_a l$ and $\cot kl \approx 1/sC_a l$ (where $s = j\omega$ is frequency in radians/second), yielding Eq. (3.34):

$$z_a \approx \frac{1}{2}(R_a + sL_a)l$$
$$y_b \approx (G_a + sC_a)l$$

where l is the length of the tube section.

(a) Find the ABCD-matrix for the T-network shown in Fig. 3.3(a). That is, find the frequency responses $A(s)$, $B(s)$, $C(s)$ and $D(s)$ so that

$$\begin{bmatrix} p_2 \\ U_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} p_1 \\ U_1 \end{bmatrix}$$

Write your answer in terms of resonant frequency of the T-section, $\omega_0 = 1/\sqrt{L_a C_a l^2}$, and the admittance and impedance loss factors, $\alpha_y = G_a/C_a$ and $\alpha_z = R_a/L_a$. Your answers for $A(s)$ and $D(s)$ should contain no variables other than s, ω_0, α_y, and α_z; your answer for $B(s)$ may also include L_a and R_a and l, and your answer for $C(s)$ may also include G_a, C_a, and l.

Notice that I am asking you for equations in terms of the acoustic quantities p and U, instead of the electrical quantities E and I that are shown in the figure. This change is inconsequential; the textbook switches to acoustic quantities beginning in Sec. 3.21.

Please approximate $2 + y_b z_a \approx 2$. The lumped-element approximations $\tan kl \approx sL_a l$ and $\cot kl \approx 1/sC_a l$ are inaccurate at high frequencies, therefore $2 + y_b z_a \approx 2$ is a pretty good approximation. This approximation does *not* apply to $1 + y_b z_a$; that one, you should expand into a complete quadratic function of s. As a result, your complete ABCD matrix will be quadratic in s.

(b) Prove that the ABCD matrix for the Π-network in Fig. 3.3(b) is identical to that of the T-network in 3.3(a).

Notice that the Figure swaps notation between Fig. 3.3(a) and Fig. 3.3(b): y_a in Fig. 3.3(b) refers to the same quantity as $y_b/2$ in Fig. 3.3(a); y_b in Fig. 3.3(b) refers to the same quantity as $y_a/2$ in Fig. 3.3(a).

(c) In discrete-time implementations of the ABCD-matrix method, it is important to avoid frequency-domain aliasing. One way to avoid aliasing is to use the bilinear transform,

$$s = \Omega_c \frac{1 - z^{-1}}{1 + z^{-1}}$$

where Ω_c should be chosen to be above the frequencies of interest, but below the Nyquist rate; 2-3 kHz is usually reasonable.

Use the bilinear transform to convert the ABCD-matrix from the frequency variable s (radians/second) to the frequency variable z (exponentiated radians/sample). Hint: if this seems difficult, then you are thinking too hard.

(d) Parts (a) and (b) of this problem refer to modeling of a tube section in the frequency domain. When we model the tube section in the time domain, we usually model the forward-going and backward-going waves $p_+(t - x/c)$ and $p_-(t + x/c)$ instead of the pressure and volume velocity. As you know, these waves are related to pressure and volume velocity by

$$\begin{aligned} p(x,t) &= p_+(t - x/c) + p_-(t + x/c) \\ U(x,t) &= Y_0(p_+(t - x/c) - p_-(t + x/c)) \end{aligned}$$

Take the Laplace transform of these two quantities at the locations $x = 0$ (in order to find $p_1(s)$ and $U_1(s)$) and $x = l$ (in order to find $p_2(s)$ and $U_2(s)$). You will need to use the delay property of the Laplace transform, $x(t - \tau) \leftrightarrow x(s)e^{-s\tau}$. Eliminate the variables $p_+(s)$ and $p_-(s)$ in order to find the ABCD matrix that computes $p_2(s)$ and $U_2(s)$ in terms of $p_1(s)$ and $U_1(s)$.

(e) Demonstrate that the ABCD matrix you found in parts (a) and (b) is a low-frequency linearization of the matrix you found in part (d).

Note that this will be a lossless approximation, i.e., it works for $R_a = G_a = 0$.

Problem 3.2

This problem considers the glottal model shown in Fig. 3.14 of the textbook.

(a) In time-domain simulations of the vocal tract, it is common to neglect the glottal mass (L_g), and model the glottis as a pure resistance ($R_g(t)$). Fig. 3.14 shows one set of relationships between the glottal pressure and volume velocity, namely $p_g(t) = R_g(t)(U_S(t) - U_g(t))$ (where I have used the convenient notation $U_S(t) \equiv P_S/R_G(t)$). A second relationship is dictated by propagation within the first tube section of the vocal tract, namely (if we assume that $x = 0$ at the glottis):

$$\begin{aligned} p_g(t) &= p_+(t) + p_-(t) \\ U_g(t) &= Y_0(p_+(t) - p_-(t)) \end{aligned}$$

Impose continuity at the junction in order to find an equation for $p_+(t)$ as a function of $p_-(t)$, $U_S(t)$, and $R_g(t)$. The result should look similar to the Kelly-Lochbaum equations, but with the addition of a term related to $U_S(t)$.

(b) In frequency-domain simulations, it is common to use both R_g and L_g. Because frequency-domain simulations don't deal very well with time-varying circuit elements, however, it is common to fix these elements at some static value, e.g., the value corresponding to a 1cm$\times$1mm glottal aperture.

A frequency-domain simulation might start with either a Thevenin equivalent circuit, as shown in Fig. 3.14 (lungs are a fixed pressure source independent of volume velocity) or a Norton equivalent circuit (lungs are a fixed volume velocity source $U_S \equiv P_S/R_g$, independent of pressure).

The Norton equivalent circuit is shown in Fig. 3.18.

Find two ABCD matrices: one $[p_g, U_g]$ to $[P_S, U_S]$ for the Thevenin equivalent circuit of Fig. 3.14, one relating these two vectors for the Norton equivalent circuit of Fig. 3.18. Argue that the ABCD matrix for the Norton equivalent is closer to being an identity matrix than that for the Thevenin equivalent. In general, identity matrices are a good thing.

Problem 3.3

As demonstrated in class, the radiation impedance shown in Eq. (3.37) (that of a hemispherically symmetric source) can be written as a resistance in parallel with an acoustic mass.

(a) Suppose that you have the following two equations for $p_L(s)$ and $U_L(s)$, the pressure and volume velocity at the lips:

$$\begin{bmatrix} p_L \\ U_L \end{bmatrix} = \begin{bmatrix} A(s) & B(s) \\ C(s) & D(s) \end{bmatrix} \begin{bmatrix} P_S \\ U_S \end{bmatrix} \tag{1}$$

$$p_L(s) = Z_R(s) U_L(s) \tag{2}$$

where the ABCD-matrix is the product of the Norton equivalent matrix in problem 3.2(b), times several consecutive tube matrices as shown in problem 3.1, and $Z_R(s)$ is the radiation impedance $Z_R = z_s(\rho c/A)$, z_s as shown in Eq. (3.37). Combine Eqs. 1 and 2 in order to find $T(s)$ as a function of C, D, and Z_R, where $T(s)$ is defined as

$$T(s) \equiv \frac{U_L(s)}{U_S(s)} \tag{3}$$

Notice that, in the Norton equivalent version of the glottis, U_S is a specified function, but P_S is unknown! Therefore, you will need to solve for P_S first (in terms of U_S, A, B, C, and D) before you solve for $U_L(s)$.

(b) Time-domain simulations of the vocal tract must cope with the frequency dependence of the radiation impedance. Fortunately, it's pretty easy to deal with. Remember that the reflection coefficient $\Gamma = (Z_R - Z_0)/(Z_R + Z_0)$ is the ratio p_-/p_+ at the lips. When Z_R is frequency-dependent, this relationship no longer works in the time domain, but it still works in the frequency domain:

$$p_-(s) = \Gamma(s)p_+(s)$$

therefore $p_-(t)$ can be computed as a convolution:

$$p_-(t) = \Gamma(t) * p_+(t)$$

Assume that $Z_R(s)$ is defined by a resistor in parallel with an acoustic mass, as shown in Eq. (3.37). Find the resulting radiation reflection response, $\Gamma(t)$.

(c) Convert $\Gamma(s)$ to $\Gamma(z)$ using the bilinear transform, $s = \Omega_c(1-z^{-1})/(1+z^{-1})$. Design a digital filter structure that will implement the convolution $p_-[n] = \Gamma[n] * p_+[n]$.