Department of Electrical and Computer Engineering
Digital Speech Processing
Homework No. 3

Problem 1

Which of the following pairs of tones is perceived as the louder tone (and by how many phons):

1. 20 dB intensity level at 1000 Hz or 20 dB intensity level at 4000 Hz
2. 40 dB intensity level at 250 Hz or 40 dB intensity level at 1000 Hz
3. 50 dB intensity level at 500 Hz or 30 dB intensity level at 2000 Hz

Problem 2

What is the perceived pitch (in Mels) for the following tones:

1. 50 Hz
2. 250 Hz
3. 500 Hz
4. 2000 Hz
5. 5000 Hz

Problem 3

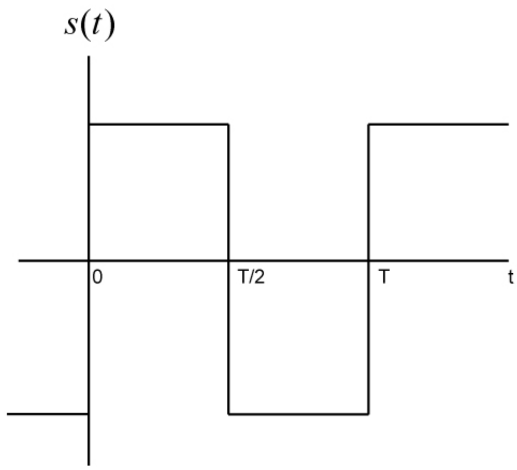


Figure 1: Periodic square wave input to ear.

If the input to the ear is the periodic square wave, $s(t)$, shown in Figure 1, with $T = 10$ msec period, what frequencies are present at the input to the basilar membrane (assuming all processing at the outer and middle ears is linear). Show the response of the basilar membrane to the periodic square wave at the stapes end and at the apical end.

Problem 4

By substitution, show that the set of equations:

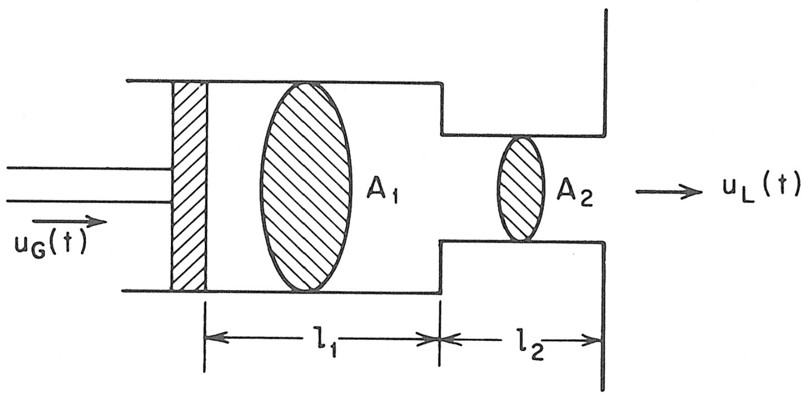
$$\begin{aligned} u(x,t) &= [u^+(t-x/c) - u^-(t+x/c)] \\ p(x,t) &= \frac{\rho c}{A}[u^+(t-x/c) + u^-(t+x/c)] \end{aligned}$$

are solutions to the partial differential equations:

$$\begin{aligned} -\frac{\partial p}{\partial x} &= \frac{\rho}{A}\frac{\partial u}{\partial t} \\ -\frac{\partial u}{\partial x} &= \frac{A}{\rho c^2}\frac{\partial p}{\partial t} \end{aligned}$$

Problem 5

Consider an ideal lossless tube model for the production of vowels consisting of 2 sections as shown in the figure below. Assume that the terminations at the glottis and lips are completely lossless.



For the above conditions the system function of the lossless tube model is obtained from the following equation:

$$V_a(s) = \frac{0.5(1 + r_G)(1 + r_L)(1 + r_1)e^{-s(\tau_1+\tau_2)}}{1 + r_1 r_G e^{-s2\tau_1} + r_1 r_L e^{-s2\tau_2} + r_L r_G e^{-s2(\tau_1+\tau_2)}}$$

by substituting $r_G = r_L = 1$ and:

$$r_1 = \frac{A_2 - A_1}{A_2 + A_1}.$$

(a) Show that the poles of the system are on the $j\Omega$ axis and are located at values of Ω satisfying the equations:

$$\cos[\Omega(\tau_1 + \tau_2)] + r_1 \cos[\Omega(\tau_2 - \tau_1)] = 0$$

or equivalently

$$\frac{A_1}{A_2} \tan(\Omega\tau_2) = \cot(\Omega\tau_1)$$

where $\tau_1 = l_1/c$, $\tau_2 = l_2/c$ and c is the velocity of sound.
(b) The values of Ω that satisfy the equations derived in (1) are the formant frequencies of the lossless tube model. By judicious choice of the parameters l_1, l_2, A_1, and A_2 we can approximate the vocal tract configurations of vowels, and by solving the above equations obtain the formant frequencies for the model. The following table gives parameters for several vowel configurations. Solve for

Vowel	l_1	A_1	l_2	A_2
/i/	9 cm	8 cm^2	6 cm	1 cm^2
/ae/	4 cm	1 cm^2	13 cm	8 cm^2
/a/	9 cm	1 cm^2	8 cm	7 cm^2
/Λ /	17 cm	6 cm^2	0 cm	6 cm^2

Table 1: Parameters of the two tube model for various vowels.

the formant frequencies for each case. (Note that the nonlinear equations must be solved graphically or iteratively.) Use $c = 35000$ cm/sec.

Problem 6

Write a MATLAB program to play out a sequence of audio files (including file separation beep tones) as specified in the text file `filelist.txt`. Use the following as the text file for your program:

```
s5_synthetic.wav
beep_fs_10000.wav
s5.wav
beep_fs_16000.wav
s5_synthetic.wav
```

You can load the speech and beep wave files from the class website.