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

ECE 537 Speech Processing

Problem Set 2
Fall 2009

Issued: Wed Sep. 2, 2009 **Due:** Wed Sept. 16, 2009

Reading for problem set 2: Flanagan, Allen & Hasegawa-Johnson 3.1-3

Problem 2.1

(a) What is the RMS average displacement of air particles for a pure-tone plane wave having a pressure of 0 dB-SPL at 1 kHz?

Solution: The RMS pressure at 0 dB SPL is 20μPa, which is 2×10^{-5} Pa. Since the characteristic impedance of air is $\rho_0 c = 407$ kg/m^{2}s, the RMS particle velocity of this wave must be $v_{RMS} = 2 \times 10^{-5}/407$m/s. This being a sinusoid, the RMS particle displacement is $x_{RMS} = v_{RMS}/\omega = 2 \times 10^{-5}/407/2000\pi$.

(b) Compare this to the thermal velocity of a nitrogen molecule. The thermal energy of a free air molecule is $E_T = (3/2)kT$, where $k = 1.38 \times 10^{-23}$ is Boltzmann's constant. Thermal energy is a form of spread-spectrum kinetic energy, i.e., the molecule has an RMS thermal velocity v_T (spread across all frequencies) of $E_T = (1/2)m|v_T|^2$, where m is the mass of the nitrogen molecule. What is v_T?

Solution: $v_T = \sqrt{3kT/m}$. There are 28g=0.028kg of Nitrogen in one mole, so $m = 0.028/(6.02\times 10^{23})$. At freezing ($T = 273K$), $v_T = \sqrt{3 \times 1.38 \times 273 \times 0.028/6.02} = 2.29$m/s.

(c) Why is the thermal vibration of air molecules not audible?

Solution: Thermal movements are all in different directions, changing all the time, with energy spread over all frequencies; the average velocity of any block of air is zero. A very small average velocity is enough to set up an acoustic wave.

Problem 2.2

A person is speaking at an intensity of 66 dB-SPL, as measured with a sound level meter at 1 meter.

(a) Find the total power in the voice assuming that the level is uniform around the head.

Solution: First, 66 dB SPL is equal to a pressure of $0.04 = 20 \times 10^{-6} \times 10^{(66/20)}$. This corresponds to an intensity (power/area) of $I_0 \equiv |P^2|/\rho_0 c = 0.04^2/407 = 3.93 \times 10^{-6}$ [Watts/m^2]. The area of the sphere at 1 meter is $4\pi r^2 = 4\pi$. Thus the total speech power is about $16\pi \approx 50$ [μW].

(b) Find the total power assuming that the intensity varies as

$$I(\theta,\phi) = I_0 \cos(\theta/2)\cos(\phi/2) \tag{1}$$

where θ is the angle in the horizontal plane, and ϕ in the vertical plane, relative to the "straight ahead" direction $\theta = 0, \phi = 0$.

Solution: The total power $\mathcal{P}$ is given by the integral of the intensity over the area:

$$\mathcal{P} = I_0 \times \int_{-\pi/2}^{\pi/2}\int_{-\pi}^{\pi} \cos(\theta/2)\cos(\phi/2)(r^2\cos\phi d\theta d\phi)$$

$$= \frac{1}{2} I_0 r^2 \times \int_{-\pi/2}^{\pi/2}\int_{-\pi}^{\pi} \cos(\theta/2)\left[\cos(3\phi/2) + \cos(\phi/2)\right] d\theta d\phi$$

$$= 2 I_0 r^2 \times \int_{-\pi/2}^{\pi/2} \left[\cos(3\phi/2) + \cos(\phi/2)\right] d\phi$$

Thus

$$= 2r^2 I_0 \times \left[\frac{2}{3}\sin(3\phi/2) + 2\sin(\phi/2)\right]_{-\pi/2}^{\pi/2} = \frac{16\sqrt{2}}{3} r^2 I_0$$

Total power is about 60% of part a.

Problem 2.3

(a) How many millibels [mB] in 1 bel [B]? Solution: 1000 mB = 1 B

(b) Give the formula for the intensity in mB units. Solution: The intensity in mB is $1000\log_{10}(I/I_{ref})$ where $I_{ref} = 10^{-12}$ W/m^2.

(c) Give the formula for the sound pressure level in cB (centibel) units. Solution: The sound pressure level in cB is $200\log_{10}(P/P_{ref})$ where $P_{ref} = 20 \times 10^{-6}$ [Pa].

Problem 2.4

Demonstrate that $P_{ref} \equiv 20$ μPa is the same as $I_{ref} \equiv 10^{-12}$ [W/m^2].

Solution: The first is given by the formula $|P_{ref}|^2/\rho c$ and the second is given by I_{ref}. Thus we need to show that numerically these two quantities are similar (almost identical). Since $(20 \times 10^{-6})^2/407 \approx 10^{-12}$, the two references are nearly the same.

Problem 2.5

A bottle has a neck diameter of $d = 1$cm and is $l = 1$cm long. It is connected to the body of the bottle "barrel" which is $D = 5$cm in diameter and $L = 10$cm long. Treat the barrel as a short piece of transmission line, closed at one end, which looks like a compliance $C = V_{barrel}/\rho_0 c^2$, and the neck which look like a mass $M = \rho_0 l/A_{neck}$. These two impedances are in series, since they both see the same volume velocity (flow).

Find the resonant frequency of the bottle. Solution: $F_{res} = 1/2\pi\sqrt{LC} = 342$Hz.

Problem 2.6

Sketch a cross-section of the vocal tract. Locate the lips, tongue tip, tongue body, epiglottis, glottis, pharynx, alveolar ridge, hard palate, velum (soft palate), and uvula.

Problem 2.7

Suppose that a tube with characteristic impedance Z_0 [kg/m^{4}s] is terminated in a cap whose acoustic impedance is $Z_L(s)$. Find the formula for the reflectance $R(s)$ in terms of the load impedance $Z_L(s)$ and the characteristic impedance z_0 if:

(a) $Z_L(s) = R$ [kg/m^{4}s] Solution: $\Gamma = (R - Z_0)/(R + Z_0)$

(b) $Z_L(s) = 1/sC$ [kg/m^{4}s]Solution: $\Gamma = (1 - sCZ_0)/(1 + sCZ_0)$

(c) $Z_L(s) = r + sM$ [kg/m^{4}s]Solution: $\Gamma = (sM + r - Z_0)/(sM + r + Z_0)$

Problem 2.8

Consider a two-tube model of the vocal tract, with two tubes of lengths $L_1 = 8$cm and $L_2 = 9$cm, with cross-sectional areas of $A_1 = 1\text{cm}^2$ and $A_2 = 5\text{cm}^2$. Assume that the terminating impedance at the lips is $Z_L = 0$, and the terminating impedance at the glottis is $Z_G = \infty$.

(a) The resonant frequencies of this system are those for which the sum of front and back cavity impedance is zero. Write an equation that, if solved, would tell you exactly the resonant frequencies of the front and back cavities.

Solution: Pressure at the junction is given by $p = -UZ_b = UZ_f$. This equation has two solutions: either p and U are both zero, or $Z_b + Z_f = 0$ (in which case U can be anything, and $p = UZ_f$). The resonant frequency of the back cavity is $Z_b = -j(\rho_0 c/A_1)\cot(kL_1)$, that of the front cavity is $Z_f = j(\rho_0 c/A_2)\tan(kL_2)$, so $k = \omega/c$ must satisfy

$$\cot(kL_1)/A_1 - \tan(kL_2)/A_2 = 0$$

(b) Approximate the front cavity impedance as zero; find the resonant frequencies of the back cavity. Solution: If $\rho_0 c/A_2 \approx 0$, the resonances are given by $\cot(kL_1) = 0$, thus $F_{res} = c/4L_1 + nc/2L_1$ for any integer n. If $c = 354m/s$ (at body temperature) and $L_1 = 9cm$, front cavity resonances are 983Hz, 2950Hz, etc.

(c) Approximate the back cavity impedance as infinity; find the resonant frequencies of the front cavity. Solution: If $\rho_0 c/A_1 \approx \infty$, the resonances are given by $\tan(kL_2) = -\infty$, equivalent to $\tan(kL_2) = 0$, thus $F_{res} = c/4L_2 + nc/2L_2$ for any integer n. If $c = 354m/s$ (at body temperature) and $L_2 = 8cm$, back cavity resonances are 1106Hz, 3319Hz, etc.

(d) Now use matlab to find the first three resonant frequencies of the whole system, using the equation you wrote in part (a). Do this as follows: plot $Z_f(f)$, the front-cavity impedance, and $-Z_b(f)$, the negative back cavity impedance, as functions of frequency f; zoom in on the places where they cross in order to learn their frequencies.

Solution: Cross-over frequencies are at 760Hz, 1328Hz, and 2810Hz.

(e) Of the resonant frequencies you uncovered in part (d), which resonances are "front-cavity resonances"? Which are "back-cavity resonances"?

Solution: The formant frequencies at 760Hz and 2810Hz are close to the resonances of the decoupled back cavity, so they are often called "back cavity resonances." The formant frequency at 1328Hz is close to the resonance of the decoupled front cavity, so it is often called a "front cavity resonance." Notice that the decoupled resonances at 983Hz and 1106Hz have been pushed apart, by coupling, to 783Hz and 1328Hz. Coupling always pushes apart resonances that would otherwise be very close together, though not always by quite this much!

Problem 2.9

Working in matlab, create an M-tube time-domain simulation of the vocal tract. Choose M, the number of tubes, so that the propagation delay from one end of a tube to the other is equal to one sample time (choose a convenient sampling frqeuency like 8kHz or 10kHz or 11.025kHz). Define the variables $p_+(m,n)$ and $p_-(m,n)$, the forward-going and backward-going waves, respectively, in the mth tube section at sample time $t = n/F_s$.

Write a function, `UL=tubemodel(A,UG,ZG,ZL)`, that computes the volume velocity at the lips $u_L(n)$ given the volume velocity at the glottis $u_G(n)$, given the areas $A = [a_1, \ldots, a_M]$ of the M tubes in the vocal tract, and given the terminating impedances Z_G of the glottis and Z_L of the lips. You may assume for now that the terminating impedances are real-valued constants, independent of frequency; we will correct that assumption in the next homework.

Solution: Here is some example code:

```
function UL=tubemodel(A,UG,ZG,ZL)
% UL=tubemodel(A,UG,ZG,ZL)
  M=length(A);
  ZRatio = [407./A, ZL]./[ZG, 407./A];
  Gamma=(ZRatio-1)./(ZRatio+1);
  Pplus = zeros(1,M);
```

```
Pminus = zeros(1,M);
UL = zeros(size(UG));
for t=1:length(UL),
  tmp = (1+Gamma).*[UG(t),Pplus]-Gamma.*[Pminus,0];
  Pminus = (1-Gamma(2:(M+1))).*[Pminus(2:M),0]+Gamma(2:(M+1)).*Pplus;
  Pplus = tmp(1:M);
  UL(t) = tmp(M+1);
end
```

(a) Test your model by finding the impulse response of a uniform tube, i.e., $a_m = 5\text{cm}^2$ for all m. Set Z_L to be roughly equal to the characteristic impedance of a tube with a 1m^2 cross-sectional area; set Z_G to be roughly equal to the characteristic impedance of a 1mm×1cm rectangular slit. Use a vocal tract length somewhere between 15cm (typical of female heads) and 17cm (typical of male heads). Plot the frequency response, $H(\omega) = U_L(\omega)/U_G(\omega)$, the Fourier transform of the impulse response.

Solution: The sampling frequency is given by $F_s = c/D$ where D is the length of a tube section. $Z_L = 407$ (or 425 if you assume $c = 354 m/s$), $Z_G = 407 \times 10^5$. You should see resonances at 500Hz, 1500Hz, etc.

uniformtube-eps-converted-to.pdf

(b) Test your model by finding and plotting the frequency response of the two-tube model you solved for in problem 8.

Solution: You should see resonances at the formant frequencies of the vowel [a], that is, 750Hz, 1250Hz, 2750Hz, etc.

aamodel-eps-converted-to.pdf