Take Home Quiz 2

Issued 2:30 PM on Tuesday 10-Dec-2013
Due back in Room 36-791 at 1 PM Wednesday 18-Dec-2013

Each student is expected to work independently. Texts and libraries may be consulted but the course notes and homework solutions should be sufficient. You’ll do best if you look over the full exam and then do the problems you are most confident about first. If you have problems understanding a question, make an assumption and state that assumption. Look over the exam before recitation tomorrow (11-Dec-2013); questions regarding exam wording maybe answered at that time.

Please place your name at the top of each page that you hand in.
1 Problem 1.

a) 6 points

Figure 1: Adapted from Figure 45 of Fall 2013, Notes on Acoustic Sensitivity or from Fig. 5.1 of Houtgast (1974).

As part of an experiment demonstrating suppression in a psychoacoustic experiment Houtgast compared the detection threshold in direct (simultaneous) masking with the pulsation threshold for the same masker and signal.

The results shown in the second panel from the top and bottom panel of Fig. 1, which apply to experiments performed without the suppressor tone. Both the signal and the masker were derived from 1000 Hz tones: the masker was in sine phase while the maskee was in cosine phase. Results indicate that in both cases the maskee threshold ($L_T$) grows roughly 1 dB for each 1 dB increase in the masker tone ($L_1$). However for a fixed value of $L_1$, $L_T$ is roughly 20 dB greater in the pulsation threshold experiment than in the direct masking experiment.

Show that both of these results could have been predicted accurately on the basis of data published more than 20 years before Houtgast's experiment.
b) 6 points

Leibold et al.\textsuperscript{1} (2007) reported on the interaction between 5 grouped tones. They assessed the change in the threshold of each of the components (the maskee) produced by the combination of the other four tones (the maskers). We concentrate on two observations:

1) When the total bandwidth of the five components was 46 Hz (no two frequencies were different by more than 46 Hz), the threshold of each maskee component was very nearly the same (roughly 55.5 dB SPL) when the other four components (the maskers) were each held constant at 53 dB SPL.

2) When the total bandwidth of the five components was 456 Hz (no two frequencies were different by more than 456 Hz), and the levels of the four masker components were each held constant at 53 dB SPL, the threshold of the maskee components was found to differ from component to component (Table 1).

<table>
<thead>
<tr>
<th>Component</th>
<th>Frequency</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hz</td>
<td>dB SPL</td>
</tr>
<tr>
<td>$F_1$</td>
<td>799</td>
<td>49.3</td>
</tr>
<tr>
<td>$F_2$</td>
<td>895</td>
<td>53.5</td>
</tr>
<tr>
<td>$F_3$</td>
<td>1000</td>
<td>55.1</td>
</tr>
<tr>
<td>$F_4$</td>
<td>1115</td>
<td>53.2</td>
</tr>
<tr>
<td>$F_5$</td>
<td>1255</td>
<td>46.6</td>
</tr>
</tbody>
</table>

Table 1: Mean masked threshold for the individual (maskee) components of the five tone complex with a total bandwidth of 456 Hz. In each case, the other (masker) components were presented at 53 dB SPL.

\textsuperscript{1}Lori J. Leibold, Hongyang Tan, Samar Khaddam, and Walt Jesteadt. "Contributions of individual components to the overall loudness of a multitone complex", \textit{Journal of the Acoustical Society of America}, 121(5):2822-2831
b-i) Assume that $F_1$, $F_3$, and $F_5$ are fixed. Eight possible alterations of the frequencies of components $F_2$ and $F_4$ of the complex are listed below. Which alteration should reduce the variation in threshold across the five tones if the masker levels are kept at 53 dB SPL. Provide a few sentences to support your reasoning.

1. Increase $F_2$ and increase $F_4$.
2. Do not change $F_2$ and increase $F_4$.
3. Decrease $F_2$ and increase $F_4$.
4. Increase $F_2$ and do not change $F_4$.
5. Decrease $F_2$ and do not change $F_4$.
6. Increase $F_2$ and decrease $F_4$.
7. Do not change $F_2$ and decrease $F_4$.
8. Decrease $F_2$ and decrease $F_4$.

b-ii) Would the masked threshold of the 1000 Hz component ($F_3$) increase, decrease, or remain the same as the result of the alterations you suggest in part (c-i)? Provide a few sentences to support your reasoning.
c) 13 points

![Graph](image)

Figure 2: Spectrum of notched noise. The relative width of the notch is \( D \).

A listener detects an \( f_0 = 2000 \) Hz tone masked by flat spectrum noise with a "spectrum level" \( N_0 = 30 \) dB SIL/Hz that is filtered externally to produce notched noise (Fig. 2). Note that in air, dB SIL = dB SPL. A graph of the person's masked threshold as a function of the relative notch width \( D \) is presented in Fig. 3. You should assume the shape of this listener's auditory filter \( (H(f)) \) is symmetric about the tone frequency \( f_0 \): \( H(f_0 - \epsilon) = H(f_0 + \epsilon) \).

![Graph](image)

Figure 3: Masked thresholds as a function of relative notch width \( D \).

c-i) Describe how the Power Spectrum Model of Masking would account for the result shown in Fig. 3.

c-ii) Estimate the values of \( A \) and \( B \). Explain your reasoning.

c-iii) Describe how you would use the results in Fig. 3 to estimate the width of this listener's auditory filter used to detect a tone of frequency \( f_0 \). The width may be defined as the notch width \( (Df_0, \) in Hz) at which the auditory filter attenuates by 10 dB more than it attenuates at \( f_0 \):

\[
H \left( f_0 \left( 1 \pm \frac{D}{2} \right) \right) = \frac{1}{10} H(f_0).
\]
2 Problem 2.

In all parts of this problem, make the following assumptions. The tube segments are rigid and of uniform circular cross-sectional area. The sound pressure is zero where the tube system opens to the outside. The speed of sound is 340 m/s.

Figure 4: Tube structure for Problem 1-a.

a) (6 points) In this part you are to consider a tube (Fig. 4) with

\[
\begin{align*}
L_B &= 6.0 \text{ cm} & L_F &= 14.0 \text{ cm} \\
A_B &= 0.5 \text{ sq cm} & A_F &= 2.0 \text{ sq cm}
\end{align*}
\]

Determine the three lowest natural frequencies of this structure.

b) (13 points)

A uniform acoustic tube with cross section 3 sq cm and length \( L = 20 \text{ cm} \) is closed at one end is driven by a cosinusoidal pressure source at the other end (Fig. 6):

\[
p(t) = P_S \cos(2\pi F_s t) = \mathcal{R} \left[ P_S e^{\text{i}\omega_s t} \right]
\]

where \( P_S = P_S, \omega_s = 2\pi F_s \), and if \( z = x + jy \), \( \mathcal{R}[z] = x \). The pressure and volume velocity within the tube are given by:

\[
\begin{align*}
p(x, t) &= \mathcal{R} \left[ P(x) e^{\text{i}\omega_s t} \right] \\
u(x, t) &= \mathcal{R} \left[ U(x) e^{\text{i}\omega_s t} \right]
\end{align*}
\]
Figure 5: Problem 1-b.

Figure 6: Problem 1-b. The top panel shows the tube structure (uniform acoustic tube with cross section 3 sq cm and length $L = 20$ cm). The middle panel shows the distribution of pressure $|P(x)|$ and volume velocity $|U(x)|$ (assuming $z_0 = 1$ acoustic ohm) when the tube is driven by a sinusoidal pressure source of pressure amplitude $P_S$ located at $x = -20$ cm. The bottom panel is provided for you to sketch the phases of $P(x)$ and $U(x)$ relative to the pressure of the source.
b-i) (4 points)

What “boundary conditions” must the following four quantities satisfy:

\[ P(-L) \quad P(0) \]
\[ U(-L) \quad U(0) \]

b-ii) (5 points)

The middle panel of Fig. 6 shows the amplitude of pressure \( |P(x)| \) and volume velocity \( |U(x)| \) as a function of position within the tube.

Specify as closely as possible the value of \( F_3 \). If \( F_1 < F_3 < F_2 \) provide expressions for the values of \( F_1 \) and \( F_2 \).

b-iii) (4 points)

On the axes provided in the third panel of Fig. 6, sketch and label the phase of the pressure \( P(x) \) and volume velocity \( U(x) \) relative to the source as functions of position \( x \).

c) (6 points)

A uniform acoustic tube closed at one end and open at the other has natural frequencies given by

\[ F_n = \frac{2n - 1}{4} \frac{c}{L} \quad n = 1, 2, \ldots \]

where \( c \) is the speed of sound and \( L \) is the length of the tube.

Assume the length of the tube \( L \) is 20 cm. You are given the task of perturbing the cross sectional area over small lengths of this tube to produce a tube-system with the first three natural frequencies:

- \( F_1 = 450 \) Hz
- \( F_2 = 1250 \) Hz
- \( F_3 = 2100 \) Hz

At what location(s) and in what manner(s) (increase or decrease) might you perturb the cross sectional area of the tube to achieve such a result, at least approximately. Sketch the resulting tube and explain your reasoning.
Question 3

1. (10 points) As part of your intergalactic travels you come across a set of recordings. You identify the vowels and plot the F1-F2 vowel space and observe that the front vowels have lower F2 and decreased amplitude of high frequency spectra relative to American English vowels.

As a scientific emissary of SHBT@Earth describe:
  a. What factors may have caused this difference.
  b. What other adaptations you might expect in their perceptual system and/or their environment.
  c. How you identified the vowels.

2. (10 points) A spectrogram is shown on the following page.

  a. Identify the 5 vowels.
  b. Estimate the median vocal tract length of the speaker from all vowels.
  c. Identify where the following waveform is located in the spectrogram (the nearest time tick is sufficient).

![Figure. A waveform segment belonging to one of the two spectrograms.]

3. (5 points) State True or False (if False, rectify to be a True statement)

  a. The harmonics in a spectrum will shift to the left as the vocal tract is increased in length.
  b. The area of the vowel quadrilateral reduces when a speaker increases speaking rate.
  c. Assuming that the amplitude and fundamental frequency of voicing remains the same, the vowel /u/ in shoot will be louder and longer than the vowel /a/ in shot.
Figure: A sentence comprising 5 vowels
Figure 1 shows a possible acoustic model of the auditory periphery. A middle ear connects an air-filled cylindrical ear canal (on the left) with a semi-infinite fluid-filled tube (the cochlea, on the right). Both tubes have rigid walls. The tympanic membrane and the stapes are modeled as rigid pistons with areas that match the cross-sectional areas of their respective tubes ($A_{tm}$ and $A_s$). The pistons are assumed massless. The malleus and incus span an air-filled tympanic cavity and are represented as rigid colinear massless rods (of length $l_m$ and $l_i$) connected at the malleoincudal joint, which is also assumed rigid. Both the tympanic membrane and the stapes are held in place by perfectly elastic (i.e., infinite compliance, zero stiffness) rings too small to be seen in the diagram. The volume of the tympanic cavity is large enough to be assumed infinite, and viscous losses in the air, fluid, and elastic rings are assumed negligible and should be ignored.

![Figure 1: An acoustic model of the auditory periphery.](image)

(a) You observe that a 2-kHz plane wave traveling down the model ear canal is perfectly transmitted to the cochlea (i.e., no acoustic power is reflected at the tympanic membrane). Use your observation to solve for the area of the stapes, $A_s$, in terms of other geometrical and acoustical parameters.

(b) Use your formula for $A_s$ to estimate the area of the human stapes footplate, using the approximate anatomical values $A_{tm} = 60 \text{ mm}^2$, $l_m = 5 \text{ mm}$, and $l_i = 4 \text{ mm}$. Compare your answer with the actual, anatomical area of the stapes footplate. Your estimated stapes area will not, of course, be exactly correct. Explain what you think is the most important reason for the discrepancy and why. Assuming perfect transmission at 2 kHz, how might you modify the model to yield a more accurate estimate? Explain your reasoning.