ECE 537 Fundamentals of Speech Processing
Problem Set 1

UNIVERSITY OF ILLINOIS
Department of Electrical and Computer Engineering

Assigned: Monday, 8/22/2022; Due: Monday, 8/29/2022
Reading: Fletcher & Munson, *Loudness, Its Definition, Measurement, and Calculation*, Bell System Technical Journal 12(4):377-430, 1933

1. For this question, let's consider a sound composed of just two pure tones, at $f_1 = 200$Hz and $f_2 = 1000$Hz. Suppose that the tone at 200Hz has an RMS amplitude of $P_1 = 2$ units, and the tone at 1000Hz also has an RMS amplitude of $P_2 = 2$ units. The peak amplitude of each pure tone is therefore $2\sqrt{2}$ units; the RMS amplitude of the composite tone is $2\sqrt{2}$ units; the peak amplitude of the composite tone is $4\sqrt{2}$ units (verify these facts for yourself). In each part of this problem set, you will be asked to compute something under three different assumptions about the "unit," i.e., under three different loudnesses of this pure tone. For the quiet tone, the unit is 10^{-9} atmospheres, thus the RMS amplitude of each pure tone is $P_k = 2 \times 10^{-9}$ atmospheres. For the medium tone, the unit is 10^{-7} atmospheres, thus the RMS amplitude of each pure tone is $P_k = 2 \times 10^{-7}$ atmospheres. For the loud tone, the unit is 10^{-5} atmospheres, thus the RMS amplitude of each pure tone is $P_k = 2 \times 10^{-5}$ atmospheres.

   (a) (1 point) Look up the value of one standard atmosphere, in Pascals. Having determined that value, compute (to one significant figure) the RMS amplitudes of the pure tone components (P_k) of the quiet tone, of the medium tone, and of the loud tone.

   Solution: One standard atmosphere is $101,325 \approx 10^5$ Pascals. Thus, to one significant figure, the RMS amplitude of each component is $P_k = 2 \times 10^{-4}$Pa for the quiet tone, $P_k = 2 \times 10^{-2}$Pa for the medium tone, and $P_K = 2$Pa for the loud tone.

   (b) (1 point) The density of air and the speed of sound vary considerably as a function of pressure and temperature. At 20°C at one standard atmospheric pressure, the density of air is about $1.204kg/m^3$, and the speed of sound is about $343m/s$. At this temperature and pressure, find the intensity of each pure tone component (J_k), in Watts/m^2, to one significant figure, for the quiet tone, the medium tone, and the loud tone.

   Solution: $\rho c \approx 413kg/m^2s$, and $J_k = P^2/\rho c$. Therefore J_k is roughly 10^{-10}W/m^2 for the quiet tone, 10^{-6}W/m^2 for the medium tone, and 10^{-2}W/m^2 for the loud tone.

   (c) (1 point) To the nearest 1dB, find the intensity level of each pure tone component (β_k) for the quiet tone, the medium tone, and the loud tone. Notice that (Fletcher & Munson) give the reference intensity in Watts/cm^2; you will want to convert it into Watts/m^2.

   Solution: The reference intensity is 10^{-16}W/cm^2, which is 10^{-12}W/m^2, therefore β_k is 20dB for the quiet tone, 60dB for the medium tone, and 100dB for the loud tone.

   (d) (1 point) To one significant figure, find the total intensity $J = J_1 + J_2$ (in Watts/m^2) of the quiet tone, of the medium tone, and of the loud tone.

Solution: The total intensity is just the sum of the intensities of the two components, thus for the quiet tone it's $2 \times 10^{-10} W/m^2$, for the medium tone it's $2 \times 10^{-6} W/m^2$, and for the loud tone it's $2 \times 10^{-2} W/m^2$.

(e) (1 point) To the nearest 1dB, find the total intensity level β of the quiet tone, the total intensity level of the medium tone, and the total intensity level of the loud tone.

Solution: The total intensity is $\beta = 10 \log_{10} J$, which is 23dB for the quiet tone, 63dB for the medium tone, and 103dB for the loud tone.

(f) (1 point) To the nearest 1dB ($\pm$1dB), find the loudness levels L_1 and L_2 of the two pure tone components of the quiet tone, of the medium tone, and of the loud tone. Suggestion: use Fig. 4; if the point you want is not right on top of one of the curves shown, estimate it visually. Are the loudness levels (L_k) larger or smaller than the intensity levels (β_k)?

Solution: The loudness levels are $L_1 \approx -3, L_2 = 20$dB for the quiet tone, $L_1 = 50, L_2 = 60$dB for the medium tone, and $L_1 = 100, L_2 = 100$dB for the loud tone. At these two frequencies, the loudness levels are less than or equal to the intensity levels of each pure tone.

(g) (1 point) Use Table III in the article to estimate the loudness of each pure tone component (G_k) for the quiet tone, the medium tone, and the loud tone. Do not round; just report results from the table.

Solution: The loudnesses are ($G_1 \approx 0.32, G_2 = 97.5$) for the quiet tone, ($G_1 = 2200, G_2 = 4350$) for the medium tone, and ($G_1 = G_2 = 88000$) for the loud tone.

(h) (1 point) Find the total loudness (G) of the quiet tone, the total loudness of the medium tone, and the total loudness of the loud tone.

Solution: The total loudness are the sum of the component loudnesses, thus $G \approx 97.8$ for the quiet tone, $G = 6550$ for the medium tone, and $G = 176000$ for the loud tone.

(i) (1 point) Use Table III to find the total loudness level (L), in decibels, of the quiet tone, of the medium tone, and of the loud tone. Express your answer to the nearest 1dB, i.e., find the entry in the table that is nearest to what you want. Are the total loudness levels (L) larger or smaller than the intensity levels (β) of each composite tone?

Solution: The total loudness level is computed by inverting Table V, thus for the quiet tone we have $L = 20$dB, which is smaller than its intensity level. For the medium tone we have $L = 67$dB, and for the loud tone we have $L = 108$dB; both of these loudness levels are larger than the corresponding intensity levels!

2. (1 point) The data for two tones, in Fig. 6, defines the relationship $G(L) = 2G(L_k)$; notice that this curve is reasonably well approximated by the straight line $L \approx L_k + 9$, i.e., $G(L_k + 9) \approx 2G(L_k)$. The data for three tones, similarly, defines the relationship $G(L) = 10G(L_k)$; notice that this curve is reasonably well approximated by the straight line $L \approx L_k + 30$, i.e., $G(L_k + 30) \approx 10G(L_k)$.

   Recall that, at the reference frequency $f = 1000$Hz, the relationship between loudness level and intensity is

$$L_k = 10 \log_{10} \left(\frac{J_k}{J_r} \right), \tag{1}$$

where $J_r = 10^{-12}\text{W/m}^2$ is the reference intensity. From Eq. 1 it's possible to compute that $L = L_k + 9$ corresponds to $J = 8J_k$, similarly $L = L_k + 30$ corresponds to $J = 1000J_k$. Since these should correspond to $G = 2G_k$ and $G = 10G_k$, respectively, we get that

$$G(10\log_{10}(8J_k)) \approx 8G(10\log_{10} J_k) \tag{2}$$

$$G(10\log_{10}(1000J_k)) \approx 10G(10\log_{10} J_k), \tag{3}$$

or, in other words,

$$G \approx C\left(\frac{J_k}{J_r}\right)^{1/3}, \tag{4}$$

where the constant C is unknown, because it appears on both sides of Eqs. 2 and 3.

Use Eq. 4 to approximate the loudness (G_2) of the $f_2 = 1000$Hz components of the quiet tone, the medium tone, and the loud tone from Problem 1. Find a value of C that makes these approximate loudness measures reasonably close to the true loudness measures, especially for the medium and loud tones.

Solution: For the quiet tone, we get $G_2 \approx C\sqrt[3]{100} = 4.64C$, compared to a true value of $G_2 = 97.5$. For the medium tone, $G_2 \approx C\sqrt[3]{10^6} = 100C$, compared to $G_2 = 4350$. For the loud tone, $G_2 \approx C\sqrt[3]{10^{10}} = 2154C$, compared to $G_2 = 88000$. The medium and loud tones are pretty well approximated if we set C to be somewhere between 40.9 and 43.5.

3. The "critical bandwidth" is the bandwidth within which the intensities of two tones add, instead of their loudnesses adding. On p. 409 of the article, the critical bandwidth is approximated as a piece-wise constant function of center frequency. Far more detailed measurements of the critical bandwidth have been made since 1933.

   (a) (1 point) A reasonable estimate of critical bandwidth, valid at frequencies above about 400Hz, is one musical tone, a.k.a. one sixth of an octave. Frequency f_2 is said to be one sixth of an octave above f_1 if

   $$\log_2\left(\frac{f_2}{f_1}\right) = \frac{1}{6} \tag{5}$$

   Use Eq. 5 to estimate the critical bandwidth, $\Delta f = f_2 - f_1$, for f_1 values of 2000Hz, 4000Hz, 8000Hz, and 16,000Hz. Compare your values of Δf to the rough critical bandwidths specified on p. 409 of the article.

   Solution: The values are

   $$f_1(2^{1/6} - 1) = \begin{cases} 245 & f_1 = 2000 \\ 490 & f_1 = 4000 \\ 980 & f_1 = 8000 \\ 1960 & f_1 = 16,000 \end{cases}$$

   These values are slightly larger than the values specified in the article.

   (b) (1 point) An empirical estimate of critical bandwidth is the "equivalent rectangular bandwidth" or ERB, defined to be the bandwidth of a rectangular bandpass filter that has the same energy as the auditory filter at any given center frequency. Moore and Glasberg (*Hearing Research* 47(1-2):103-138, 1990) published the following formula for ERB:

   $$\text{ERB}(f) = 24.7\,(4.37f + 1)\,, \tag{6}$$

   where in this formula, f is measured in kilohertz. Find ERB(f) for values of $f \in \{2, 4, 8, 16\}$kHz, and compare these to the values specified in the article.

Solution: This formula gives

$$\text{ERB}(f) = \begin{cases} 241 & f = 2 \\ 456 & f = 4 \\ 888 & f = 8 \\ 1750 & f = 16 \end{cases}$$

These values are larger than the critical bandwidths specified by Fletcher & Munson, but a little bit smaller than 1/6 of an octave.

4. When two tones are within one critical bandwidth of each other, their intensities add ($J = J_1 + J_2$), and loudness is therefore sub-additive ($G < G_1 + G_2$). When the tones are separated by more than about 1000Hz, their loudnesses add ($G = G_1 + G_2$). When $f_2 - f_1$ is more than a critical bandwidth, but less than 1000Hz, the loudness of the composite tone is somewhere between these two results.

   In every part of this problem, assume that $J_2 = J_1$, $L_2 = L_1$, and $G_2 = G_1$. The only things we will vary are the frequencies f_2 and f_1, and the level $L_k = L_1 = L_2$.

   (a) (1 point) Consider, first of all, the loudness that results if the tones are within one critical band of each other, therefore the intensities add ($J = J_1 + J_2 = 2J_1$). Assume that loudness is given by the cube-root approximation in Problem 2. Suppose that we define the loudness of the composite tone to be $G = G_1 + b_2 G_2 = G_1(1 + b_2)$. What is the value of b_2?

Solution: If $f_2 = f_1$, then the intensities add, $J = J_2 + J_1 = 2J_1$, so

$$\begin{aligned} G &= C\left(\frac{J}{J_r}\right)^{1/3} \\ &= C\left(\frac{2J_1}{J_r}\right)^{1/3} \\ &= 2^{1/3} C\left(\frac{J_1}{J_r}\right)^{1/3} \\ &= 2^{1/3} G_1 \end{aligned}$$

where $2^{1/3} \approx 1.26$, therefore in this case $b_2 \approx 0.26$.

   (b) (1 point) The Fletcher & Munson model predicts that b_k is given by their Eq. (17) unless Eq. (17) is greater than 1.0, in which case $b_k = 1.0$. Fletcher & Munson specify that Eq. (17) should only be used if Δf is larger than a critical band; the smallest possible critical band, in their model, is 100Hz. Sketch b_k as a function of Δf, over the range $100 \le \Delta f \le 1000$Hz. Draw two different plots on the same axis: one plot showing b_k if $L_k = 20$dB, and one plot showing b_k if $L_k = 60$dB.

Solution: If $L_k = 20$ dB, $b_k = \min\left(1.0, 2.6\left(\frac{250+\Delta f}{1000}\right)\right)$. This should be a straight line rising from 0.91 when $\Delta f = 100$ to 1.0 when $\Delta f = 130$Hz, then it should be 1.0 for all higher values of Δf.

If $L_k = 60$ dB, $b_k = \min\left(1.0, 0.8\left(\frac{250+\Delta f}{1000}\right)\right)$. This should be a straight line rising from 0.28 when $\Delta f = 100$, up to a value of 1.0 when $\Delta f = 1000$Hz.