Problem Set 7

Issued: 9 November 2004	Due: 30 November 2004
Legend: $\heartsuit \equiv$ extra credit (not required)	
N.B. For all problems explain your reasoning! No explanation $\Rightarrow$ No credit $\Rightarrow$ No joy	
When making plots, be sure to label all axes, provide numerical tick marks, and specify the units of measurement.	
Collaboration is encouraged so long as everyone understands and works on all problems. Please indicate the names of your collaborators.	

These problems are designed to get you thinking about quantitative approaches to issues in hearing. In problem 1 you will determine a lower bound on the size of an acoustic "particle" and estimate the number of collisions an air molecule experiences per second. In problem 2 you will analyze some of George von Békésy's (Nobel Prize, 1961) classic measurements of the traveling wave. In problem 3 you will explore the relation between tuning curves (measured versus frequency at fixed position) and traveling waves (measured versus position at fixed frequency) and show how measurements of one can be used to deduce features of the other. In problem 4 (optional) you will be challenged to derive and solve the equations for a simple cochlear model. In problems 5 & 6 you will compare your model responses to actual measurements of basilar-membrane motion made in the squirrel monkey. Does the simple model agree with experiment? In problem 7 you will investigate a classic model for the physiological basis of musical consonance based on the notion of the "critical band."

Problem 1: Molecules and Sound Particles. In their chapter on the physics of sound in *The Speech Chain*, Denes and Pinson confuse the molecules in air (i.e., the molecules of nitrogren, oxygen, etc) with the much larger "fluid particles" used to derive the equations of acoustics.

1. The average distance a molecule in air travels before colliding with another molecule is known as its *mean free path.* The mean free path sets a conservative lower bound on the size of an acoustic particle. Estimate the mean free path for a typical molecule in air at room temperature. [Hint: Estimate the mean free path as the distance the molecule needs to travel before it sweeps out a volume equal to the average empty volume surrounding each molecule of air. Use the ideal gas law and a plausible estimate of molecular dimensions.]

2. At what sound frequency would the wavelength of sound become comparable to the mean free path? Does sound propagate in air at these frequencies? Explain why or why not.

3. Determine the typical number of air molecules in a cube one mean free path on a side.

4. Estimate the typical (root-mean-square) velocity of the molecules in air at room temperature. [Hint: Use thermodynamic equipartition of energy which says that the translational kinetic energy of an air molecule moving in three dimensions is typically $\frac{3}{2}kT$.] Use your answer to determine the typical number of collisons a molecule of air experiences during one second.

Problem 2: Békésy's Traveling Waves. This problem is based on the measurements of Békésy reproduced in the course notes (Fig. 11-58 of Békésy 1960, or Fig. 7.14 of Yost 1994, or Fig. 3.8 of Pickles, 1988).

1. According to Békésy's measurements, what is the *instantaneous* speed of the traveling wave that results from a 100 Hz stimulus as it passes a point 30 mm from the human stapes? What is the speed of a 200 Hz stimulus at the same point? Compare your results with the speed of sound in water.

2. The waves Békésy describes are very different from plane waves. Explain why, illustrating your remarks by using Békésy's measurements to deduce the time waveform of the partition displacement that results from a stimulus consisting of two pure tones of equal amplitude with frequencies of 100 Hz and 200 Hz. (Assume, following Békésy's procedure, that the stimulus is applied directly to the stapes; filtering by the middle ear can then be neglected.) Plot a snapshot of the time-domain waveform as it would appear (1) at the stapes, (2) at 28 mm from the stapes, and (3) at 30 mm from the stapes.

Problem 3: Width of the Excitation Pattern. In this problem you will estimate the number of inner hair cells stimulated by a low-level pure tone. Figure 1 below shows a neural tuning curve measured in the cat. Imagine presenting a threshold-level pure tone at the neuron's characteristic frequency. The tone sets up a traveling wave whose envelope—and the corresponding pattern of hair-cell excitation—has a certain width, Δx. Count as "stimulated" any hair cell whose stereociliary motion is at least $1/10$ as large as the motion at the characteristic place.[1] Derive a relation between the bandwidth, Δf, of the tuning curve and the spatial width, Δx, of the excitation pattern. For simplicity, assume that the relation between ear-canal sound pressure and stereociliary deflection is linear and that the middle-ear transfer function does not significantly affect the shape of the tip of the tuning curve. Justify these assumptions. Explain your procedure for estimating the spatial "spread of excitation" from the neural tuning curve (Hint: Use local scaling symmetry). Potentially useful facts: (1) in this region the cat cochlear position-frequency map is approximately exponential:

$$f_{\mathrm{CF}}(x) \approx f_{\max} e^{-x/l} ,$$

where $f_{\max} \approx 57$ kHz and $l \approx 5$ mm; (2) the width of a hair cell is roughly 10 μm. What fraction of the total number of inner hair cells is stimulated by a threshold-level pure tone? Comment on the canonical characterization of neural tuning as "sharp."

♡ **Problem 4: Transmission-line Model of the Cochlea.** In this problem you will analyze the simple, one-dimensional model of cochlear mechanics illustrated in Fig. 2. Represent the "unrolled" [illegible] to two fluid filled chambers, representing the scala vestibuli and scala tympani, separated by an elastic membrane, representing the cochlear partition. Assume that the wavelength of the traveling wave is long compared to the height of the scalae. The pressures in the two scalae are then approximately uniform in any cross section and depend only on the longitudinal distance from the stapes, x.

1. Assume that fluid viscosity is negligible and use Newton's 2nd law to derive an expression between the fluid pressure, p_{v}, and the longitudinal fluid velocity, u_{v}, in the scala vestibuli. In particular, show that
$$\frac{\partial p_{\mathrm{v}}}{\partial x} = -\rho \frac{\partial u_{\mathrm{v}}}{\partial t} . \tag{1}$$
where ρ is the density of the fluid. (Hint: Consider the forces on the fluid element in a cross sectional slice through the scala vestibuli of length Δx.) Derive the analogous equation relating p_{t} and u_{t} in the scala tympani. Assume that both u_{v} and u_{t} are positive when the fluid particle flows in the direction of increasing x (i.e., towards the helicotrema).

2. Assume that the fluid is incompressible and derive an expression between u_{v} and the membrane velocity using conservation of mass. Assume that the membrane displacement, $d(x,t)$, is

[1] Extra credit: Justify this criterion based on your knowledge of the difference between rate and synchrony measures of neural threshold.

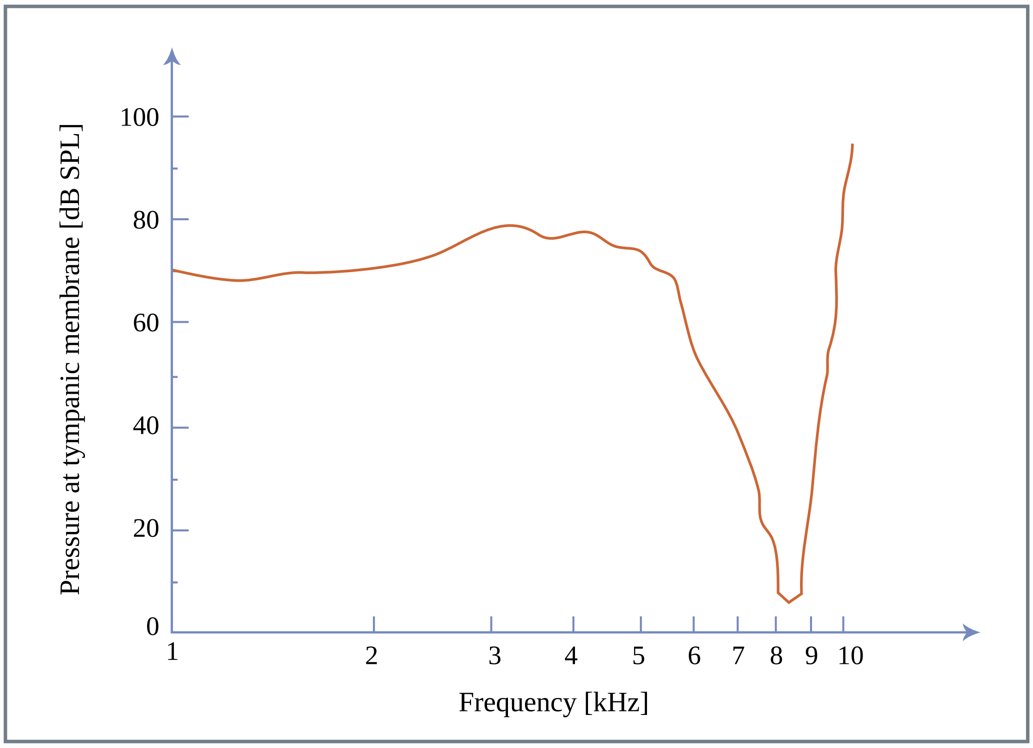

Figure by MIT OCW

Figure 1: Neural tuning measured in the cat by Liberman and Kiang (1978, Fig. 1; see also Fig. 4.3 of Pickles (1988)).

orthogonal to its surface and let d be positive when the membrane is displaced downwards (i.e., into the scala tympani). In particular, show that

$$S\frac{\partial u_{\mathrm{v}}}{\partial x} = -b\frac{\partial d}{\partial t}\,, \tag{2}$$

where S is the constant cross-sectional area of each scala and b is the width of the membrane. (Hint: Consider fluid flow into and out of the cross-sectional slice of length Δx.) Likewise, show that

$$S\frac{\partial u_{\mathrm{t}}}{\partial x} = b\frac{\partial d}{\partial t}\,. \tag{3}$$

3. Consider now the motion of the cochlear partition, which moves in response to the pressure difference across its surface. Represent a small section Δx of the partition as a simple harmonic oscillator with an effective mass $\mu\Delta x$, damping $\gamma\Delta x$, and stiffness $\kappa\Delta x$. (Thus, μ, γ, and κ are the mass, damping, and stiffness per unit length.) Derive the equation of motion for this section using Newton's 2nd law. In particular, show that

$$b(p_{\mathrm{v}} - p_{\mathrm{t}}) = \mu\frac{\partial^2 d}{\partial t^2} + \gamma\frac{\partial d}{\partial t} + \kappa d\,. \tag{4}$$

4. Show that the quantity $S(v_{\mathrm{v}} + v_{\mathrm{t}})$ is constant, independent of position. Explain why this constant must be zero. Use the results to show that $p_{\mathrm{v}} + p_{\mathrm{t}}$ is also constant, independent of position.

5. Simplify the equations by introducing the variables $p \equiv p_{\mathrm{v}} - p_{\mathrm{t}}$ and $u \equiv S(u_{\mathrm{v}} - u_{\mathrm{t}})/2$. Assume that p and u have sinusoidal time dependence and denote their Fourier transforms by upper-case letters (so that, e.g., $p(x,t) = \mathrm{Re}\{P(x,\omega)e^{i\omega t}\}$, where ω is 2π times the frequency of

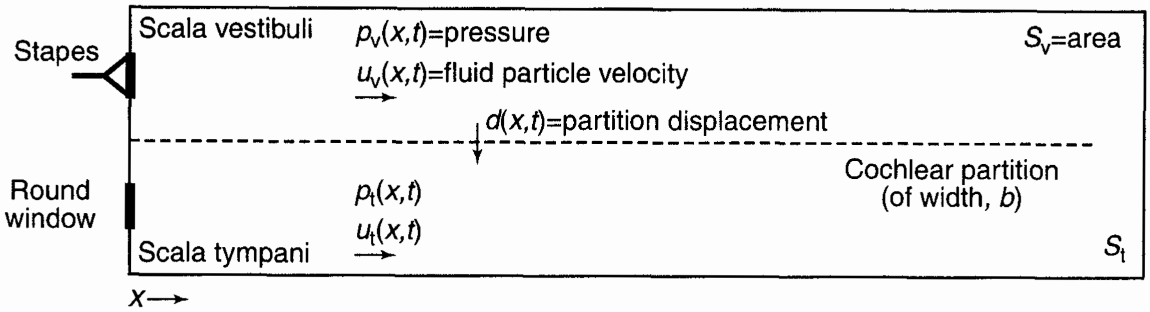

Figure 2: Schematic diagram.

stimulation). Show that one obtains a pair of first-order ordinary differential equations for P and U:

$$\frac{dP}{dx} = -ZU ; \tag{5}$$

$$\frac{dU}{dx} = -YP , \tag{6}$$

where

$$Z(x, \omega) \equiv i\omega M , \tag{7}$$

and

$$Y(x, \omega) \equiv \frac{1}{i\omega L(x) + R(x) + 1/i\omega C(x)} . \tag{8}$$

What are the values of M, L, R, and C in terms of the mechanical properties of the cochlea (e.g, ρ, b, μ, etc)?

Equations (5) and (6) have the same form as the equations describing an electrical transmission line with series impedance Z and shunt admittance Y per unit length. The long-wave model is thus often referred to as a *one-dimensional transmission-line model.*

6. Decouple the transmission-line equations to obtain a *wave equation* for $P(x, \omega)$ at frequency ω:

$$\frac{d^2P}{dx^2} + \frac{1}{\lambda\!\!\!^{-2}} P = 0 . \tag{9}$$

[Hint: Differentiate Eq. (5) for dP/dx and substitute Eq. (6) for dU/dx into the result.] What is $\lambda\!\!\!^{-}(x, \omega)$ in terms of Z and Y?

7. Solve the wave equation for $P(x, \omega)$ [Eq. (9)] assuming that Z and Y are constant, independent of position. Discuss the character of the solution $Pe^{i\omega t}$ when $\lambda \equiv 2\pi\lambda\!\!\!^{-}(x)$ is real. Provide a physical interpretation of λ. How is the solution modified if λ has an imaginary part?

8. The mechanical properties of the cochlea (e.g., the mass and stiffness of the partition) vary with position. But if they change gradually enough, the cochlea might be expected to act locally much as a uniform transmission line. This assumption allows one to obtain an approximate solution for the forward-traveling pressure wave using the so-called "WKB approximation."[2] Here, you will show that when the mechanical properties of the medium vary with position,

[2] "WKB" stands for Wentzel, Kramers, and Brillion, who applied this approximation technique to Schrödinger's equation and the problem of a wave packet moving in a potential.

the amplitude of the wave changes even when λ is entirely real. First, justify assuming a trial solution of the form

$$P(x) \approx A(x)e^{-i\int_0^x dx'/\lambda(x')} , \qquad (10)$$

where $A(x)$ is some function to be determined (note that the dependence on ω has been omitted for clarity). Substitute this expression into the wave equation and assume that the second spatial derivative of $A(x)$ is "small" and can be neglected. Solve the resulting equation and determine the function $A(x)$. Discuss the qualitative differences between this solution for $P(x,\omega)$ and the result you obtained assuming that $\lambda = \text{constant}$.

9. At its basal end, the cochlear "transmission line" is driven by the motion of the stapes. To compare model predictions with experiment we need to normalize basilar-membrane (BM) velocity by the velocity of the stapes and thereby obtain a "transfer function,"

$$T(x,\omega) \equiv \frac{\text{BM velocity}}{\text{stapes velocity}} , \qquad (11)$$

that depends only on the mechanics of the cochlea. Assume continuity of volume velocity and obtain an expression for the ratio, T, of membrane velocity at point x to the velocity of the stapes. Your expression should involve b, Z, Y, P, and the area of the stapes footplate (or oval window), S_{ow}.

10. Based on the forms for Z and Y obtained above (i.e., Eqs. 7 and 8) show that λ has the form

$$\lambda(x,\omega) = \frac{l}{4N}\frac{\left(1-\beta^2+i\delta\beta\right)^{1/2}}{\beta} , \qquad (12)$$

where $\beta(x,\omega) \equiv \omega/\omega_{\text{r}}(x)$, $N \equiv (l/4)\sqrt{M/L}$, $\omega_{\text{r}}(x) \equiv 1/\sqrt{LC}$ is 2π times the "resonant" frequency of a section of membrane, and $\delta \equiv \omega_{\text{r}}RC$ is the dimensionless damping parameter. Comment on the significance of the fact that $\lambda(x,\omega)$ is a function of the ratio $\omega/\omega_{\text{r}}(x)$.

Problem 5: Comparing Theory and Experiment. In this problem you will compare the one-dimensional transmission-line model with actual measurements of basilar-membrane motion. To solve for the model response, one substitutes the equation for the wavelength of the traveling wave $2\pi\lambda(x,\omega)$ [from Eq. 12 from Problem 4, part 10] into the expression you obtained for the basilar-membrane transfer function $T(x,\omega)$ [from Problem 4, part 9], uses the WKB approximation to solve for the pressure $P(x,\omega)$ [from Problem 4, part 8], and evaluates the necessary integrals. When the smoke clears, the transfer function, T, becomes[3]

$$T(x,\omega) \approx T_0 i\beta(x,\omega)\left[\frac{\omega_{\text{max}}}{\omega_{\text{r}}(x)}\right]^{1/2}\frac{e^{-i4N\{\sin^{-1}[\beta(x,\omega)-i\delta/2]-\sin^{-1}[\beta(0,\omega)-i\delta/2]\}}}{[1-\beta^2(x,\omega)+i\delta\beta(x,\omega)]^{3/4}} , \qquad (13)$$

where T_0 is a real, dimensionless constant and the dimensionless constants N and δ (defined along with $\beta(x,\omega)$ in Problem 4, part 10) have been assumed independent of position.

1. Rhode's (1971) measurements of the amplitude and phase of $T(x_0,\omega)$ in the squirrel monkey (made as a function of angular frequency ω at some point x_0) are shown in Fig. 3. Compute $T(x_0,\omega)$ from Eq. (13)—using, for example, Matlab or some similar program—and vary the free parameters (T_0, x_0, N, and δ) to try to obtain a decent fit to the data.[4] A decent fit should

[3] Much extra credit: Derive Eq. (13) for $T(x,\omega)$.

[4] Compute your model responses over the range 1–10 kHz using at least 256 points/decade resolution. Plot your amplitude results on a logarithmic frequency scale and your phase results on a linear frequency scale, as in Fig. 3. You may find the Matlab functions angle() and unwrap() helpful for computing your model phase responses. (When trying to match the phase data, the function unwrap() will help you remove discontinuities from the model phase response by adding integer multiples of 2π when appropriate.) And when converting your magnitudes to dB, be careful to take the common and not the natural logarithm!

Rhode's Measurements of T (Animal 69-473)			
Amplitude		Phase	
f [kHz]	$\|T\|$ [dB]	f [kHz]	$\angle T$ [rad]
1.0	−6.0	1.0	−0.6
1.5	−2.9	1.5	−2.0
2.0	−1.2	2.0	−3.4
2.5	4.2	2.5	−4.5
3.0	3.0	3.0	−6.3
3.5	9.2	3.5	−7.1
4.0	5.7	4.0	−8.5
4.5	10.7	4.5	−10.0
5.0	12.3	5.0	−11.1
5.5	15.6	5.5	−13.0
6.1	19.8	6.0	−14.2
6.4	21.6	6.2	−14.9
6.7	23.6	6.4	−15.6
7.0	25.6	6.6	−16.6
7.2	25.4	6.8	−17.3
7.4	27.4	7.0	−18.4
7.6	27.2	7.2	−19.1
7.8	29.3	7.4	−19.6
8.0	24.1	7.6	−21.5
8.2	19.4	7.8	−22.9
8.3	15.2	8.0	−23.9
8.4	8.5	8.2	−24.8
8.7	−0.8	8.4	−25.3
8.9	−10.2	8.6	−26.1
9.0	−13.1	8.8	−27.5
9.3	−21.2	9.0	−28.0
9.5	−28.0	9.3	−27.8
10.0	−24.9	10.0	−28.5

Table 1: Data for the measurements shown in Fig. 3. Note that the frequencies at which the amplitude and phase are measured are not always the same.

do a reasonable job reproducing both the amplitude and the phase simultaneously. (Hint: Try $N \sim 3$ and $\delta \sim 1/10$ as reasonable first guesses for N and δ.) List the parameter values of your fit and plot the resulting function T together with the data. Note that measurements in the squirrel monkey suggest that

$$\omega_{\rm r}(x) = \omega_{\rm max} e^{-x/l} , \tag{14}$$

where $\omega_{\rm max}/2\pi \approx 50$ kHz and $l \approx 5$ mm. For your reference the values of the data points are listed in Table 1.

2. Based on your experience varying the parameters values, explain the physical significance of the parameters T_0, x_0, N, and δ. What features of T depend on the value of each? How accurately are the parameter values determined by the data (i.e., how sensitively does the fit depend on the value of each)? Assess the overall quality of your fit. What features of the model response are similar to the data? What features are not?

♡ 3. Using your best-fit parameters, plot the real and imaginary parts of $\lambda(x,\omega)$ [i.e., $2\pi\lambda(x,\omega)$ from Eq. 12 in Problem 4, part 10]. Make two plots: (1) the first at x_0 as a function of f and

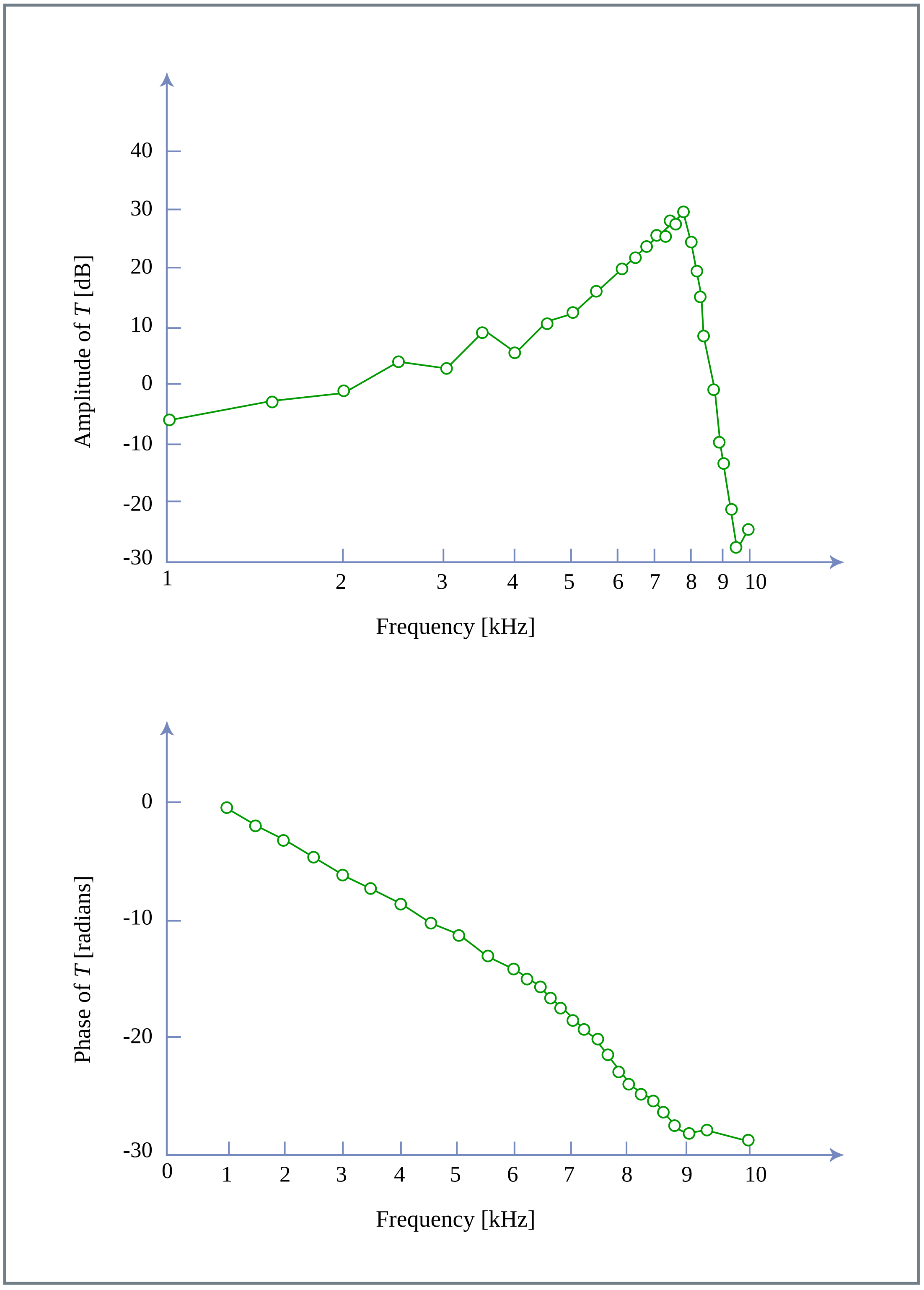

Figure by MIT OCW

Figure 3: Basilar-membrane transfer function measured in the squirrel monkey (Rhode, 1971).

Extrapolated Measurements of T (Animal 73-104)		
f [kHz]	$\|T\|$ [dB]	$\angle T$ [rad]
2.0	1.7	−2.8
4.6	15.4	−10.6
5.4	19.6	−12.9
6.0	24.6	−15.0
6.3	29.5	−16.5
6.4	33.4	−17.2
6.6	40.3	−18.3
6.8	41.0	−19.5
7.0	50.7	−20.8
7.1	61.5	−22.2
7.4	80.9	−25.8
7.6	80.3	−29.3
7.8	51.3	−34.8
8.1	5.6	−34.0

Table 2: Data for the measurements shown in Fig. 4.

(2) the second at the best frequency (i.e., approximately 7.8 kHz) as a function of x. Discuss how the real and imaginary parts of $\lambda(x,\omega)$ determine the behavior of the wave (see part 7 of Problem 4).

Problem 6: Theory and Experiment, Revisited. Cochlear mechanics is now known to be extremely labile. Healthy preparations show nonlinear responses at all but the lowest (and highest) sound levels. Rhode's (1971) measurements were made on what is now thought to be a compromised preparation at sound levels of 70–90 dB SPL. Figure 4 shows a more modern estimate of T obtained by extrapolating Rhode's later measurements to sound-levels near threshold (Zweig, 1991).

1. Compute $T(x,\omega)$ as in Problem 5 and attempt to obtain an approximate fit to the data by varying the free parameters (i.e., T_0, x_0, N, and δ). (Hint: Use your earlier experience to obtain initial estimates.) List the parameter values of your fit and plot the resulting function T together with the data. For your reference the values of the data points are listed in Table 2.

2. Assess the overall quality of your fit. What features of the data are captured by the model response? What features are not? A number of simplifying assumptions were made in deriving the model. Which assumptions do you believe to be the most significant (i.e., most likely to underlie any discrepancy between the model predictions and the data)? Explain your reasoning.

Problem 7: A Model for the Physiological Basis of Harmony. Certain musical intervals are called *consonant* because listeners consider pitches separated by these intervals to "sound good" when played together, e.g., on a piano. The intervals traditionally considered consonant are listed in Table 3. Note how the relative frequencies of the tones in a consonant interval are given by ratios of small integers. In this problem you will investigate a model (originally due to Helmholtz and extended and modernized by Plomp) for the physiological basis for this striking "numerological" fact, a fact that underlies the rules of harmony governing much of Western music.

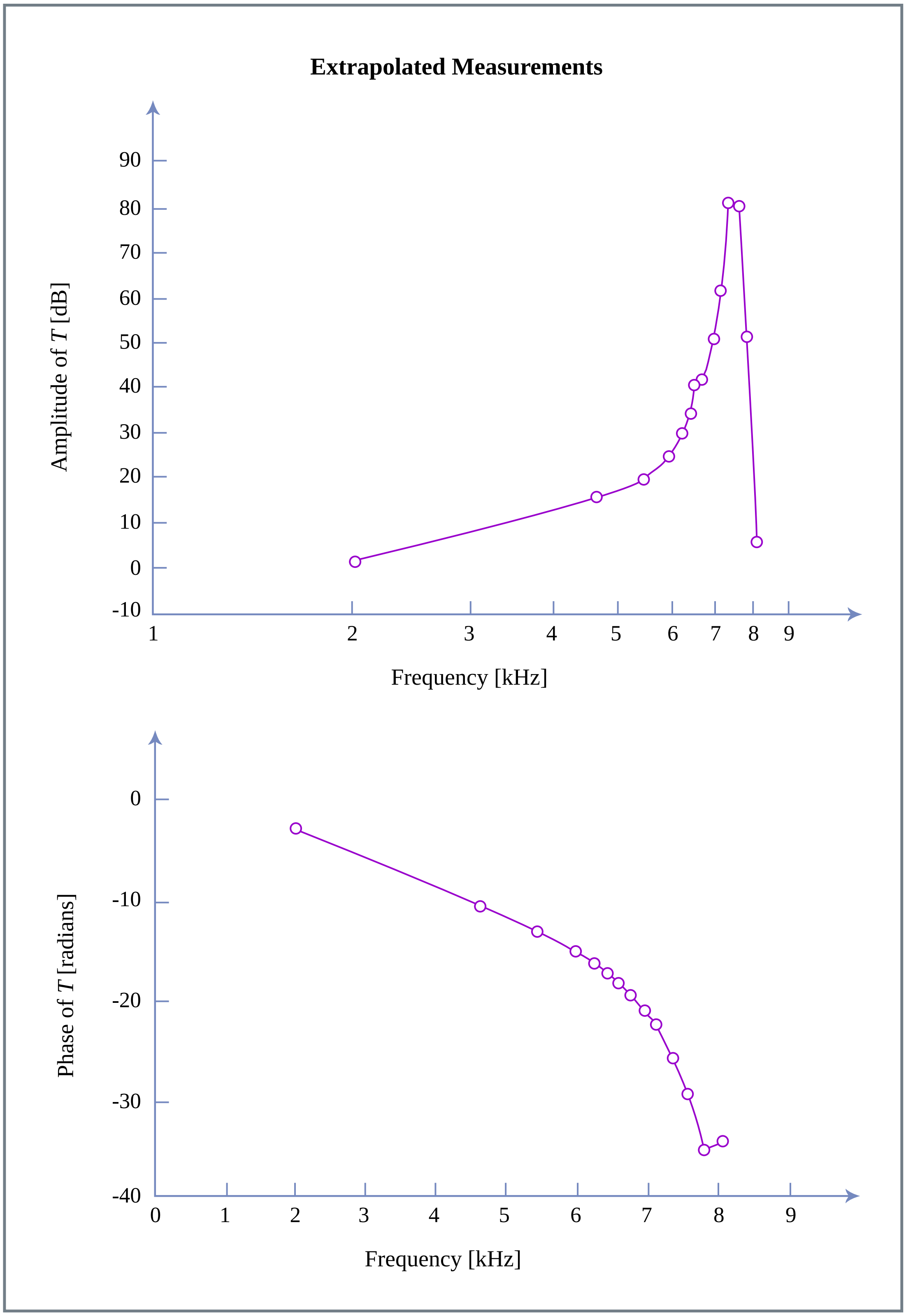

Figure by MIT OCW

Figure 4: Basilar-membrane transfer function obtained by extrapolating Rhode's measurements to low sound levels (Zweig, 1991).

Consonant Intervals		
Name of interval	**Notes (in the key of C-major)**	**Ideal frequency ratio**
Octave	C-C	2/1
Fifth	C-G	3/2
Fourth	C-F	4/3
Major Third	C-E	5/4
Minor Third	E-G	6/5
Major Sixth	C-A	5/3
Minor Sixth	E-C	8/5

Table 3: The consonant intervals and their ideal frequency ratios.

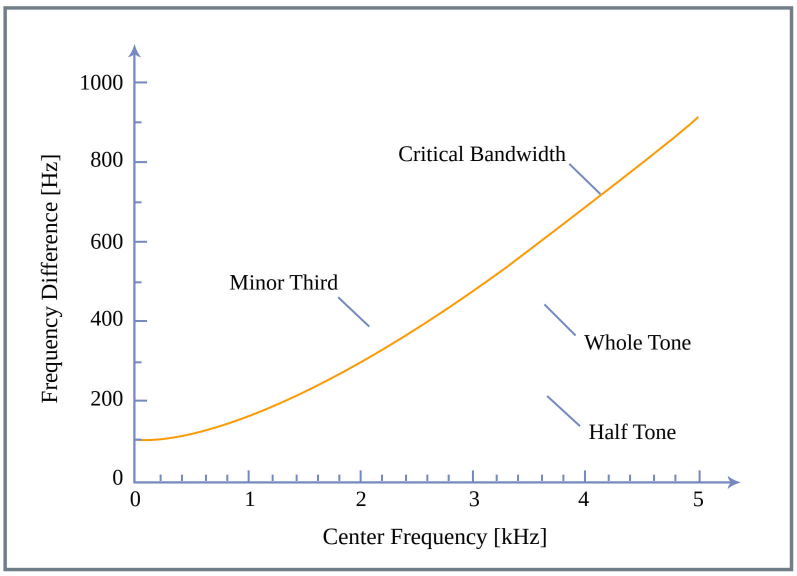

Figure by MIT OCW

Figure 5: Dependence of the critical bandwidth on frequency, as approximated by Zwicker and Terhardt's (1980) formula given in the text. The frequency differences corresponding to three musical intervals are shown for comparison. Note that the data imply that a given musical interval, e.g., a minor third, sounds more consonant at high frequencies than it does at low.

Central to the model is the psychophysical idea of the "critical bandwidth," which arises from the frequency analysis performed by the cochlea. At low and moderate sound levels, frequency components in a complex sound that lie farther apart than a critical bandwidth excite separate groups of nerve fibers (cf. Problem #3), and such components can be "heard out" as individual tones. But frequency components lying within a critical bandwidth of one another excite an overlapping group of fibers, producing an interaction that results in sensations of beats and/or "roughness." Figure 5 illustrates how the critical bandwidth varies with frequency. The critical bandwidth is roughly constant at low frequencies and roughly proportional to frequency at high frequencies.

Plomp has measured the relative roughness (or dissonance) of two pure tones as a function of their frequency separation and obtained a curve qualitatively similar to that shown in Fig. 6. We will now use this curve and the assumption of "dissonance additivity" to compute the "total dissonance" of two complex stimuli with harmonic spectra similar to that produced by a musical instrument. Let each of the two stimuli consist of 6 harmonic partials with fundamental frequencies f_1 and f_2, respectively. (In other words, let stimulus #1 consist of six tones at frequencies $\{f_1, 2f_1, 3f_1, 4f_1, 5f_1, 6f_1\}$ and stimulus #2 of six tones at frequencies $\{f_2, 2f_2, 3f_2, 4f_2, 5f_2, 6f_2\}$.) Define the total dissonance of a complex harmonic sound to be the sum of the dissonance values, D, computed for all pair-wise combinations of harmonic partials.

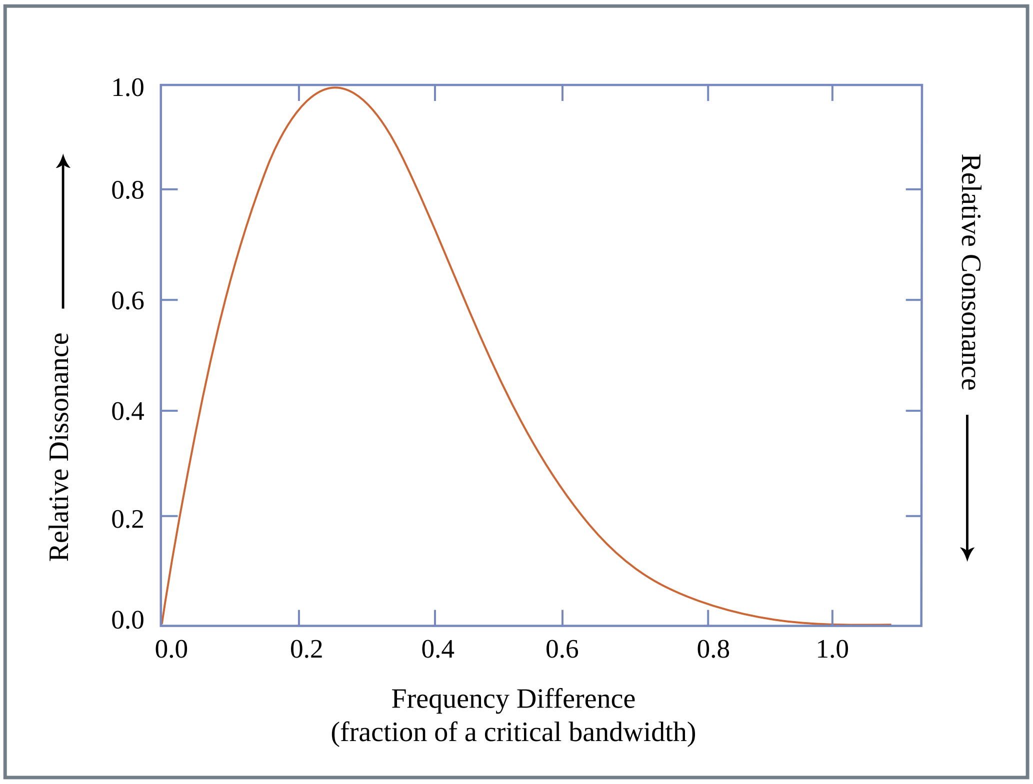

Figure by MIT OCW

Figure 6: Qualitative form of Plomp's empirical curve representing the relative dissonance of two pure tones as a function of their frequency separation expressed as a fraction, r, of a critical bandwidth. Empirically, two tones are consonant if their frequency separation is either greater than a critical bandwidth ($r > 1$) or is sufficiently small that the listener perceives slow beats (i.e., a single, amplitude-modulated tone). Maximum dissonance occurs at a frequency separation of approximately one quarter of a critical band.

1. For $f_1 = 250$ Hz, compute the total dissonance of stimulus #1, played by itself.[5]

2. Take $f_1 = 250$ Hz and $f_2 = 265$ Hz (so that f_2 is approximately one half-step higher than f_1) and compute the total dissonance when both stimuli are played together.

3. With the fundamental of stimulus #1 fixed at $f_1 = 250$ Hz, compute and plot the total dissonance of both stimuli played together as a function of f_2. Vary f_2 over slightly more than an octave, from a half-step below 250 Hz to a half-step above 500 Hz. Use a frequency resolution of at least 128 pts/octave. Plot the total dissonance curve and show that it has a number of fairly sharp local minima, corresponding (in this model) to local consonance maxima. Identify the frequencies f_2 at which these consonance maxima occur and express them as ratios relative to f_1. How do your results compare with ratios identified in Table 3?

[5] For computational convenience, approximate Plomp's curve of dissonance values, $D(r)$, with the formula

$$D(r) = \frac{g(r/2r_0)}{g(1/\sqrt{2})},$$

where $g(x) = xe^{-x^2}$ and $r_0 \equiv \sqrt{2}/8$. Use Zwicker and Terhardt's approximate analytic expression for the critical bandwidth, CB(f), as a function of frequency:

$$\mathrm{CB}(f)/\mathrm{Hz} = 25 + 75\left[1 + 1.4(f/\mathrm{kHz})^2\right]^{0.69}.$$