The Auditory System and Human Sound-Localization Behavior Short Answers Exercises Chapter 7: Acoustic Localization Cues

Exercise 7.1:

This exercise refers to the double-pole coordinates shown in Figure 1.7A, where the iso-elevation and iso-azimuth lines are indicated as parallel small circles on the unit sphere. For a target in the frontal hemifield at fixed azimuth α_0 deg, the iso-azimuth circle on the sphere has radius $\cos(\alpha_0)$. For straight ahead, $\alpha_0 = 0$, the radius is 1.0, and the target's elevation angles can run over the full range from $[-\pi/2, +\pi/2]$. At the far-lateral positions, $\alpha_0 = \pm\pi/2$, the radius is zero, and hence all elevation angles are confined to zero deg. At the intermediate azimuths, the allowed elevations will run from $[-\pi/2 + \alpha_0, +\pi/2 - \alpha_0]$. Indeed, this behavior is described by

$$\sin\varepsilon \in [-\cos\alpha_0, +\cos\alpha_0]$$

Exercise 7.2:

When the maximum Δx equals precisely one full wavelength of the tone, λ m, the azimuth angles at straight ahead and at ± 90 deg will all yield the same interuaral phase difference (IPD) of zero deg, and hence these three locations will be ambiguous for sound localization.
The maximum L-R phase difference is determined as

$$\Delta\Phi_{\text{MAX}} = 2\pi \cdot r \cdot (\pi/2 + 1) \cdot f \approx 0.0046 \cdot f \;\; \text{rad}$$

For $f < 1335 = f_{\text{MAX}}$ Hz there is never an ambiguity as the IPD remains below 2π. However, for higher frequencies, the IPD may (or may not) exceed 2π, depending on the actual (unknown) azimuth angle. For a general azimuth angle in the frontal hemifield it is

$$\Delta\Phi(\alpha) = 2\pi \cdot r \cdot (\alpha + \sin(\alpha)) \cdot f \quad \text{with} \;\; \alpha \in [-\pi/2, +\pi/2]$$

When the phase difference is 3π (path-length difference $3\lambda/2$ there will be 4 ambiguous azimuth angles between $[-\pi/2, +\pi/2]$, at 4π there are 5, and so on:

$$\Delta x = n\pi \quad \text{for} \;\; n \geq 2 \quad \Rightarrow \quad N(\alpha) = (n+1) \;\; \text{at} \;\; [-\pi/2, -\pi/2 + \frac{\pi}{n+1}, \cdots, \pi/2 - \frac{\pi}{n+1}, +\pi/2]$$

Each increment in the number of ambiguous azimuths occurs at frequencies

$$f(n) = n \cdot \frac{c}{2 \cdot r \cdot (1 + \pi/2)} = n \cdot 667 \;\; \text{Hz} \quad \text{with} \;\; n \geq 2$$

Exercise 7.3:

(a) The cross-correlation function between $x_L(t)$ and $x_R(t) = x_L(t-\Delta T)$ is (suppose a periodic signal with period T):

$$\Phi_{LR}(\tau) = \frac{1}{T}\int_0^T x_L(t)\cdot x_R(t+\tau)dt = \frac{1}{T}\int_0^T x_L(t)\cdot x_L(t-\Delta T+\tau)dt = \Phi_{LL}(\tau-\Delta T)$$

which is the delayed auto-correlation function.

(b) We take $x_L(t) = A\sin(\omega_0 t)$, with period $T = 2\pi/\omega_0$:

$$\Phi_{LR}(\tau) = \frac{A^2\omega_0}{2\pi}\int_0^{2\pi/\omega_0} \sin(\omega_0 t)\cdot\sin(\omega_0(t-\Delta T+\tau)dt$$

We find:

$$\Phi_{LR}(\tau) = \frac{A^2}{2}\cdot\cos(\omega_0(\tau-\Delta T))$$

(c) The half-wave rectified tone is described in the following way

$$x_L(t) = \begin{cases} \sin(\omega_0 t) & 0 \le t \le \pi/\omega_0 \\ 0 & \pi/\omega_0 \ \le t \le 2\pi/\omega_0 \end{cases}$$

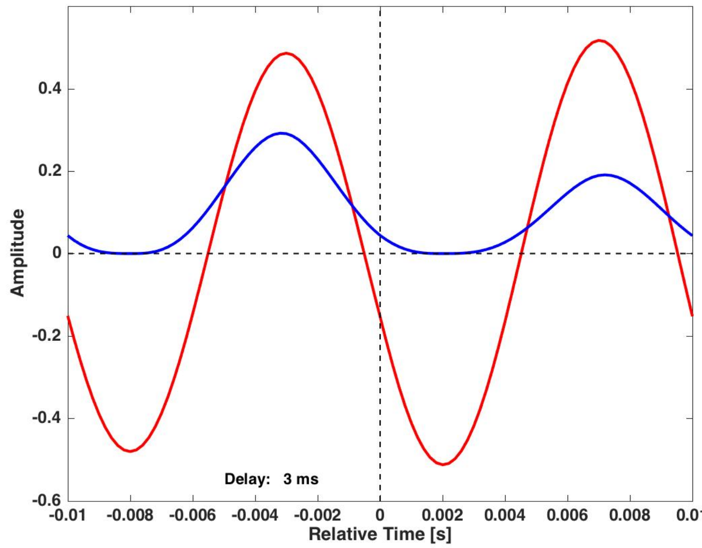
Note that

$$x_R(t) = x_L(t-\Delta T) = \begin{cases} \neq 0 & \Delta T-\tau \le t \le \pi/\omega_0+\Delta T-\tau \\ = 0 & \pi/\omega_0+\Delta T-\tau \ \le t \le 2\pi/\omega_0+\Delta T-\tau \end{cases}$$

from which we have to adapt the integration boundaries (only where the two functions overlap and are unequal to zero). For example, in case $\Delta T > \tau$:

$$\Phi_{LR}(\tau) = \frac{A^2\omega_0}{2\pi}\int_{\Delta T-\tau}^{\pi/\omega_0} \sin(\omega_0 t)\cdot\sin(\omega_0(t+\tau-\Delta T))dt$$

etc. The figure shows a calculation for a sine (blue) and the rectified sine (blue) with Matlab routine *Chapter7-Exc7-3.m*.



Cross correlation for $f_0 = 100$ *Hz and a delay of 3 ms*

Exercise 7.4:

Let's first make an estimate of the electrical conduction properties of axons. An axon is described as a cylinder of fluid, surrounded by a (partially-)insolating membrane. Current can flow inside the axon through the fluid, and across the membrane. Charge can accumulate across the membrane, which is responsible for the resting potential difference. If the axon has a radius r, membrane thickness $a \ll r$ and length L, it will have an intra-axon resistance (from the ionic fluid inside the axon) that is proportional to its length, and inversely proportional to its cross-section:

$$R_a(L) = \rho_a \cdot \frac{L}{\pi r^2}$$

with ρ_a the specific resistance of the internal fluid (in Ωm). The membrane capacitance is proportional to the membrane surface, and inversely proportional to the distance of the charged plates:

$$C_m(L) = \varepsilon_0 \varepsilon_m \cdot \frac{2\pi r L}{a} \equiv 2\pi r L \cdot c_m$$

with c_m the specific membrane capacity in F/m^2. Let's suppose that myelinated axons have their nodes of Ranvier at 1 cm distances, so that action potentials 'jump' from node to node in 1 cm steps. The velocity with which they travel is then estimated by

$$v_{\text{AP}} \approx \frac{\Delta L}{\tau_m} = \frac{0.01}{\tau_m}$$

in which the membrane time constant is determined by

$$\tau_m = R_a(\Delta L) \cdot C_m(\Delta L) = \frac{2\rho_a \Delta L^2 c_m}{r}$$

so that the speed is

$$v_{\mathsf{AP}} \approx \frac{1}{2\rho_a c_m \Delta L} \cdot r \propto r$$

So now substitute some numbers (extracted from Hobbie; take myelinated axon):

$$\begin{aligned}
\Delta L &\approx 0.01 \text{ m} \\
r &\approx 3\mu \text{ m} \\
\rho_a &\approx 40 \ \Omega\text{m} \\
c_m &\approx 0.13\mu \text{ F/m}^2
\end{aligned}$$

from which we estimate

$$v_{\mathsf{AP}} \approx 30 \text{ m/s}$$

For an azimuth angle at α_0 deg the interaural time difference is given by

$$\Delta T = \frac{\Delta x_{Head}}{c} = \frac{r_{Head} \cdot (\alpha_0 + \sin(\alpha_0))}{c}$$

with $r_{Head} \sim 0.1$ m and $c \approx 343$ m/s this is

$$\Delta T = 2.9 \cdot 10^{-4} \cdot (\alpha_0 + \sin(\alpha_0))$$

E.g., for an azimuth angle of 10 deg this time difference becomes a length difference of

$$\Delta L_{Axon} = 8.7 \cdot 10^{-3} \cdot (0.17 + 0.17) = 3 \text{ mm}$$

Exercise 7.5:

This requires some geometry calculations to determine the difference in distance from the two ears to the sound source, and the fact that sound intensity is inversely proportional to the squared distance to the source (as the head is acoustically transparent, this is the only factor for a reduction in intensity).

The coordinates of the source in the horizontal plane (say, forward/lateral) re. center of the head are

$$[x_S, y_S] = [R\cos(\alpha_0), R\sin(\alpha_0)]$$

As the right ear and left ear have coordinates

$$[x_{re}, y_{re}] = [x_{re}, 0] = \quad \text{and } [x_{le}, y_{le}] = [-x_{re}, 0]$$

the coordinates of the source re. the right ear are

$$[x_{Sr}, y_{Sr}] = [R\cos(\alpha_0) - x_{re}, R\sin(\alpha_0)]$$

and to the left ear they are

$$[x_{Sl}, y_{Sl}] = [R\cos(\alpha_0) + x_{re}, R\sin(\alpha_0)]$$

So the acoustic distances of the source to the left and right ear are given by

$$D_{Sr} = \sqrt{R \cdot (R - 2 \cdot x_{re}\cos(\alpha_0))} \text{ and } D_{Sl} = \sqrt{R \cdot (R + 2 \cdot x_{re}\cos(\alpha_0))}$$

The intensity of the source at the right/left ear is then

$$I_r = \frac{I_0}{D_{Sr}^2} \text{ and } I_l = \frac{I_0}{D_{Sl}^2}$$

and hence, the intensity difference is written as

$$\Delta I = \frac{I_0}{R} \cdot \frac{4x_{re}\cos(\alpha_0)}{R^2 - 4x_{re}^2\cos^2(\alpha_0)}$$

The numbers are $R = 0.7$ m, $\alpha_0 = 45$ deg, and $x_{re} = 0.075$ m, leading to

$$\Delta I = \frac{I_0}{0.7} \cdot \frac{4 \cdot 0.075 \cdot 0.71}{0.49 - 4 \cdot 0.075^2 \cdot 0.5} \approx 0.63 \cdot I_0 \approx I_0 - 2 \text{ dB}$$

(the intensity difference scales with the absolute source intensity).

Exercise 7.6:

The two pinna reflections give rise to two different delays, $\Delta T_1 = 2\Delta x_1/c$ and $\Delta T_2 = 2\Delta x_2/c$, and the total signal in the ear canal, with the two (unattenuated) reflections is:

$$p'(t) = p(t) + \int_0^\infty p(t - \tau - \Delta T_1)d\tau + \int_0^\infty p(t - \tau - \Delta T_2)d\tau$$

In the frequency domain this becomes

$$P'(\omega) = P(\omega) \cdot (1 + \exp(-i\omega\Delta T_1) + \exp(-i\omega\Delta T_2))$$

and the transfer characteristic becomes

$$H(\omega) = \frac{P'(\omega)}{P(\omega)} = (1 + \exp(-i\omega\Delta T_1) + \exp(-i\omega\Delta T_2))$$

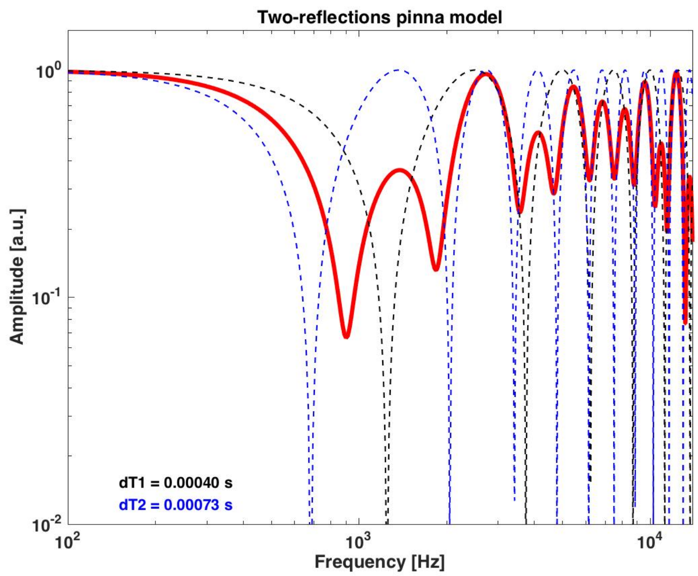
The HRTF corresponds to the amplitude characteristic, which is

$$G(\omega) = \sqrt{Re(H)^2 + Im(H)^2}$$

and this yields

$$\Rightarrow G(\omega) = \sqrt{3 + 2\cos(\omega\Delta T_1) + 2\cos(\omega\Delta T_2) + 2\cos(\omega(\Delta T_1 - \Delta T_2))}$$

An example of this amplitude characteristic is shown in the figure:



Red: two-delay model of the pinna for $\Delta T_1 = 0.0004$ s and $\Delta T_2 = 0.00073$ s. Black: one-delay model for ΔT_1 (Eqn. 7.33); blue: one-delay model for ΔT_2. Note that the first notch of the blue curve is shifted to a lower frequency than the black curve ($f_1 = 1/(2\Delta T_2)$ = 685 Hz, vs. $f_1 = 1/(2\Delta T_1)$ = 1250 Hz), and that the first notch of the two-delay model lies between these two single-reflection notches.

The figure was generated with Matlab function *Chapter7-Exc7-6.m*
(Note that the two example delays are quite long: for a total path-length difference of 3 cm the delay would be 0.00009 s, leading to a first notch at 5.7 kHz).

Exercise 7.7:

Suppose that the reflection is modified by a particular linear filter, given by $r(\tau)$. We then rewrite Eqn. 7.31 as follows

$$p'(t) = p(t) + \int_0^\infty p(t - \tau - \Delta T) \cdot r(\tau) d\tau$$

which in the frequency domain becomes (and take $K(\omega) \equiv FT[r(\tau)] = Re[K(\omega)] + i \cdot Im[K(\omega)] \equiv R(K) + i \cdot I(K)$) reads:

$$P'(\omega) = P(\omega) \cdot (1 + K(\omega) \cdot \exp(-i\omega\Delta T))$$

and the amplitude characteristic becomes (see also the previous exercise)

$$G(\omega) = \sqrt{1 + ||K(\omega)||^2 + 2R(K(\omega)) \cdot \cos(\omega\Delta T) + 2I(K(\omega)) \cdot \sin(\omega\Delta T)}$$

(note that for $K(\omega) = 1$ this reduces indeed to Eqn. 7.33).

Exercise 7.8:

The sensory spectrum is a convolution of the HRTF with the sound source spectrum, which in the frequency domain reads

$$Y_S(\omega;\varepsilon_S) = HRTF(\omega;\varepsilon_S) \cdot S(\omega)$$

Taking the logarithm (and representing log-frequency, $\Omega =^2 \log \omega$, like in the cochlea), this reads as

$$\hat{Y}_S(\Omega;\varepsilon_S) = H\hat{R}TF(\Omega;\varepsilon_S) + S(\hat{\Omega})$$

We now take the spectral correlation of the sensory signal with a particular (stored) HRTF, which corresponds to some elevation angle, ε:

$$C_{\hat{Y}}(\varepsilon;\varepsilon_S) = C[\hat{Y}_S(\Omega;\varepsilon_S), H\hat{R}TF(\Omega,\varepsilon)]$$

and by applying the definition of Eqn. 7.37, we obtain

$$C_{\hat{Y}}(\varepsilon;\varepsilon_S) = \frac{\sigma_{H_S}}{\sigma_Y} \cdot C[H\hat{R}TF(\Omega;\varepsilon_S), H\hat{R}TF(\Omega;\varepsilon)] + \frac{\sigma_S}{\sigma_Y} \cdot C[H\hat{R}TF(\Omega;\varepsilon), \hat{S}(\Omega)]$$

or, in short notation,

$$C_{\hat{Y}}(\varepsilon;\varepsilon_S) = \frac{\sigma_{H_S}}{\sigma_Y} \cdot C_H(\varepsilon;\varepsilon_S) + \frac{\sigma_S}{\sigma_Y} \cdot C_S(\varepsilon)$$

And if the second term is zero (source spectrum does not resemble an HRTF), then the correlation is maximum for that HRTF that resembles the true HRTF (provided the HRTFs are all unique).

Exercise 7.9:

Suppose that the sound source is positioned at elevation location ε_0, which is the physical sound position. Then the speaker will always generate a spectral imprint on the sensory spectrum, which is given by

$$S(\omega) = H(\omega;\varepsilon_0) \cdot X(\omega)$$

We have seen that if the source spectrum is uncorrelated with the stored HRTFs, the subject will perceive the sound at ε_0. However, if we shape the source spectrum such that it does correlate well with some other elevation angle, say ε^*, in the following way:

$$X^*(\omega) = \frac{H(\omega;\varepsilon^*)}{H(\omega;\varepsilon_0)} \cdot X(\omega)$$

then the sensory spectrum will become

$$S^*(\omega) = H(\omega;\varepsilon_0) \cdot X^*(\omega) = H(\omega;\varepsilon_0) \cdot \frac{H(\omega;\varepsilon^*)}{H(\omega;\varepsilon_0)} \cdot X(\omega) = H(\omega;\varepsilon^*) \cdot X(\omega)$$

and according to the spectral correlation model, the subject will now perceive the sound source at ε^*! In this way, a fixed single speaker can generate *all possible perceptual elevation angles*! Note that this is *not* possible for azimuth.... To simulate different azimuth angles with fixed speakers, you need at least two speakers with systematically varying relative intensities and timings.

Exercise 7.10:

We approximate this problem by simple ray-tracing: first-reflection sounds thus come from one single point on each of the four walls, floor and ceiling, that is found by equal incident-reflection angles between source-wall-ear. Eqn. 7.49 then reads

$$p_{Ear}(t) = \frac{p_0\left(t - \frac{r}{c}\right)}{r} + \sum_{n=1}^{6} \left\{ \frac{1}{r_{1n} r_{n2}} \cdot \int_0^{\infty} w_n(\tau) p_0 \left(t - \frac{r_{1n}}{c} - \frac{r_{n2}}{c} - \tau \right) d\tau \right\}$$

in which r is the shortest distance from the sound source to the center of the head, r_{1n} is the distance from the sound source to wall n and r_{n2} the distance from the wall to the center of the head. These distances are taken to a point that is the reflective midpoint of the source - wall - head triangle. Finally $w_n(\tau)$ is the filter (dampening) of wall n.