# The Auditory System and Human Sound-Localization Behavior

## Exercises Chapter 1

**Problem 1-1** Coordinates and coordinate transformations of sound locations.
The azimuth (A) - elevation (E) system is a so-called *double-pole* coordinate system, in which the azimuth angle in the horizontal plane is specified by a rotation about a head-vertical rotation axis from the mid-sagittal plane to the source position, while the elevation angle in the vertical plane is described by a rotation about the inter-aural axis from the horizontal plane to the source (Figure 1.7A).

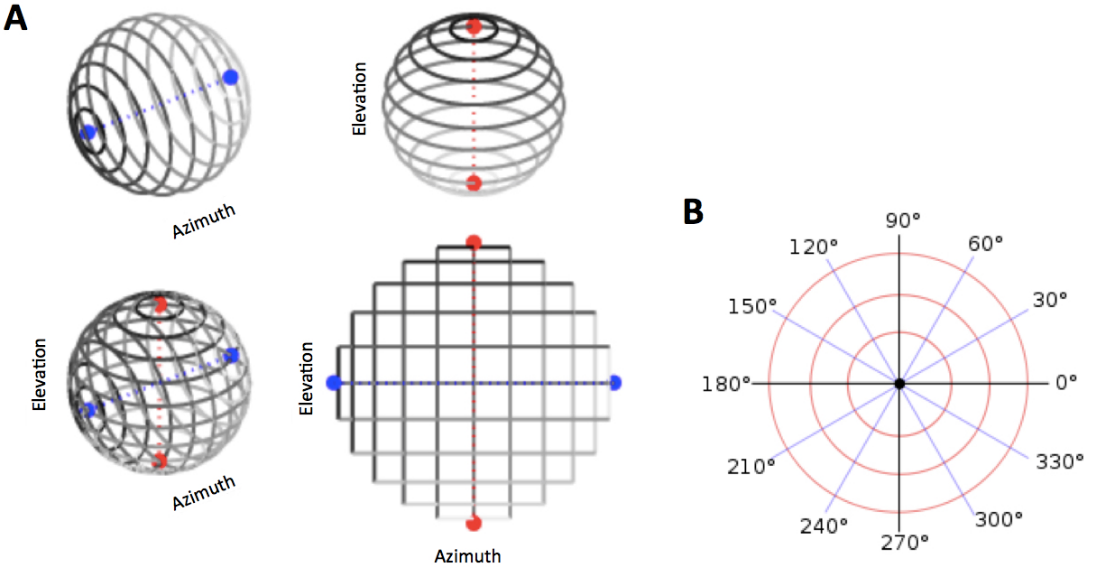


**Figure 1.7** *Left: double-pole azimuth-elevation coordinate system to specify sound locations with respect to the head of the subject. Top plots show the iso-azimuth (left) and iso-elevation (right) contour lines. Bottom plots show the full sphere (left) and the projection of (A,E) contour lines as seen from straight-ahead. Right: polar coordinates, showing iso-eccentricity lines (circles) and iso-direction lines (spokes).*

(a) Suppose a target sound is described by azimuth-elevation angles ($A_S$, $E_S$). You wish to foveate the sound with your eyes, which initially fixate at straight-ahead, i.e., (A,E)=(0,0) deg. The sound location has to be transformed into oculocentric polar coordinates, expressing the rotation amplitude, given by eccentricity angle, R, and direction, Φ (Figure 1.7B). Do the transformation, that is, calculate $R(A_S,E_S)$ and $\Phi(A_S,E_S)$.

(b) Show that for all sound locations in the frontal hemifield: $A+E \leq \pi/2$

**Problem 1-2** Audiovisual integration is useful only when sound S and visual target V are both at the *same* spatial location (see Fig. 1.6)! Often, sensory coordinates may differ substantially, but when they emerged from a single object they should be integrated. From Eqn. 1.2, determine the coefficients [*a,b,c,d*] for adequate audiovisual integration.

Often, however, the *sensory* coordinates of S and V may be identical $(A_S,E_S)=(R_V,\Phi_V)$, yet originate from *different* objects. In such cases the stimuli should not be integrated.

Draw that situation. What happens if the coefficients that you just determined for integration are applied by the sensorimotor system? Give an argument as to why this may or may not help in the identification and localization tasks.

**Problem 1-3** Inverse problems are often ill-posed. A good example is given by the problem of perception: on the basis of a limited number of measurements (sensory observations, e.g. foveation of points in the visual scene through saccades), the brain has to make an estimate (inference) about the environment and the stimuli that caused the percept. Mathematically, a problem is well-posed if there exists a unique and stable solution to the problem. A solution is stable if it resists (small) perturbations of the starting values. If solutions are not stable, or unique, the problem is ill-posed.

The latter may even happen for seemingly trivial calculations such as taking a derivative. As a numerical exercise we look at the following example: suppose that we have to determine the derivative of a function, f(x):

$$q(x) = \frac{df}{dx}$$

However, it could occur that instead of *f(x)* we have to deal with a slightly perturbed measurement of this function, say:

$$f_n(x) = f(x) + \frac{\sin(nx)}{\sqrt{n}}$$

Clearly, for $n \to \infty$ the difference between perturbed and original function, $\|f - f_n\|$, approaches zero. Show that for the derivative, however, this difference becomes arbitrarily high. This means that the operation is unstable, and hence calculating the derivative is ill-posed.

**Problem 1-4** Figure 1.8 shows a simple two-layer (input-output) neural network consisting of N linear neurons in each layer. The activity of the input layer is drawn above the neurons. Only neurons k-2, k, k+2 and k+4 receive input strengths of 1, 3, 2 and 1 units, respectively.

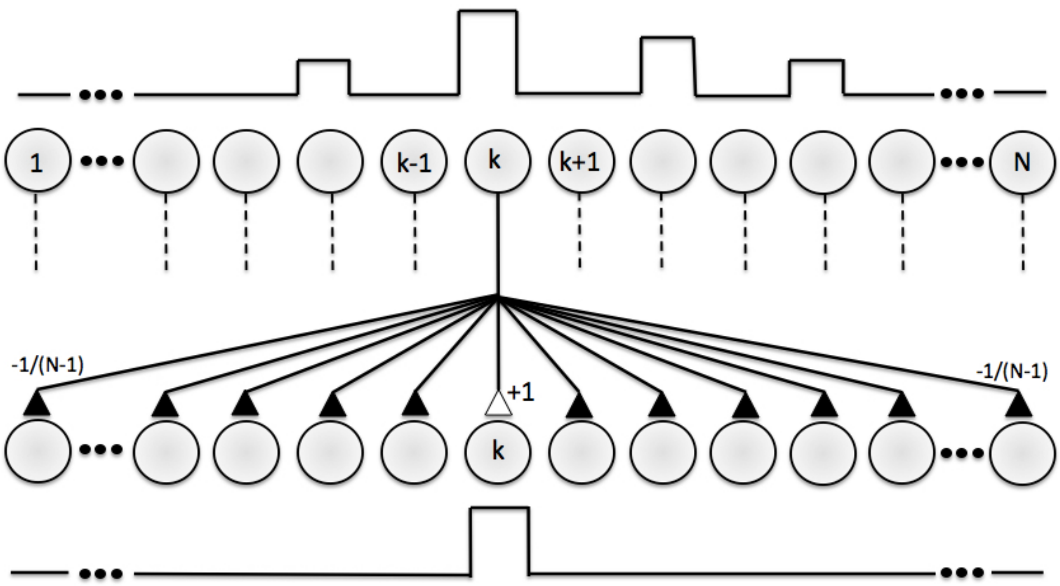


**Figure 1.8** A *Winner-Take-All (WTA) two-layer feedforward network of linear neurons. The connection scheme (here highlighted for input neuron k only) is identical for all neurons.*

The neurons of the network have repetitive connection patterns; only the connections of input neuron *k* are drawn for clarity. Each neuron in the input layer excites its corresponding output neuron with synaptic strength +1, and inhibits all the other neurons of the network with synaptic strengths -1/(N-1). In this way, the total synaptic weight from each input neuron sums to zero. The activity of an output neuron is determined by:

$$y_k = \sum_{n=1}^{N} w_{kn} \cdot x_n$$

with $w_{kn}$ the synaptic connection from input neuron *n* to output neuron *k*, and $x_n$ the activity of input neuron *n*. Take N=11 and k=5. Show that the network indeed operates as a WTA network by calculating the activities of all N output neurons. In modeling saliency maps, WTA networks play an important role, as they weed out the contributions from all competing active neurons except from the one neuron with the strongest activation.