

Alla Borisyuk
Mathematical Biosciences Institute
Ohio State University

Physiology and mathematical modeling of the auditory system

Contents

1	Introduction	3
1.1	Auditory system at a glance	3
1.2	Sound characteristics	7
2	Peripheral auditory system	9
2.1	Outer ear	9
2.2	Middle ear	10
2.3	Inner ear. Cochlea. Hair cells.	10
2.4	Mathematical modeling of the peripheral auditory system	13
3	Auditory Nerve (AN)	20
3.1	AN structure	20
3.2	Response properties	21
3.3	How is AN activity used by brain?	25
3.4	Modeling of the auditory nerve	27
4	Cochlear nuclei	27
4.1	Basic features of the CN structure	28
4.2	Innervation by the auditory nerve fibers	29
4.3	Main CN output targets	30
4.4	Classifications of cells in the CN	31
4.5	Properties of main cell types	32
4.6	Modeling of the cochlear nuclei	38
5	Superior olive. Sound localization, Jeffress model	39
5.1	Medial nucleus of the trapezoid body (MNTB)	40
5.2	Lateral superior olivary nucleus (LSO)	40
5.3	Medial superior olivary nucleus (MSO)	41
5.4	Sound localization. Coincidence detector model	41

6	Midbrain	48
6.1	Cellular organization and physiology of mammalian IC	48
6.2	Modeling of the IPD sensitivity in the inferior colliculus	49
7	Thalamus and cortex	58
	References	60

1 Introduction

It is hard to imagine what the world would be like if it was put on “mute”. For most people the vast and diverse stream of auditory information is an indispensable part of the environment perception. Our auditory system, among other things, allows us to understand speech, appreciate music, and locate sound sources in space. To be able to perform all perceptual functions the auditory system employs complex multi-stage processing of auditory information. For example, the sound waves are separated into frequency bands, which later in the system are often fused to create perception of pitch, small amplitude and frequency modulations are detected and analyzed, timing and amplitude differences between the ears are computed, and everything is combined with information from other sensory systems.

In this chapter we will discuss how the nervous system processes auditory information. We will need to combine some knowledge from anatomy (to know which structures participate), electrophysiology (that tells us the properties of these structures), and, finally, mathematical modeling (as a powerful tool for studying the mechanisms of the observed phenomena). In fact, one of the goals of this chapter is to show examples of types of mathematical modeling that have been used in the auditory research.

The material in this chapter is based on lectures given by Catherine Carr and Michael Reed at the Mathematical Biosciences Institute, Ohio State University in April 2003. It also uses material from the following books: “The Mammalian auditory pathway: neurophysiology” (A.N. Popper, R.R. Fay, eds. [77]), “Fundamentals of hearing: an introduction” by W.A. Yost [116], “An introduction to the psychology of hearing” by B.C.J. Moore [67], “From sound to synapse” by C. Daniel Geisler [26], chapter “Cochlear Nucleus” by E.D. Young and D. Oertel in [90]; and websites: <http://serous.med.buffalo.edu/hearing/>, <http://www.neurophys.wisc.edu/aud/training.html>.

1.1 Auditory system at a glance

Let us start by looking at the overall path that the auditory information travels through the brain. The auditory system differs significantly from the visual and somatosensory pathways in that there is no large direct pathway from peripheral receptors to the cortex. Rather, there is significant reorganization and processing of information at the intermediate stages. We will describe these stages in more details in the subsequent sections, but for now just outline the overall structure.

Sound waves traveling through the air reach the ears — the outer part of the peripheral auditory system. Then they are transmitted, mechanically, through the middle ear to the auditory part of the inner ear — the cochlea. In the cochlea

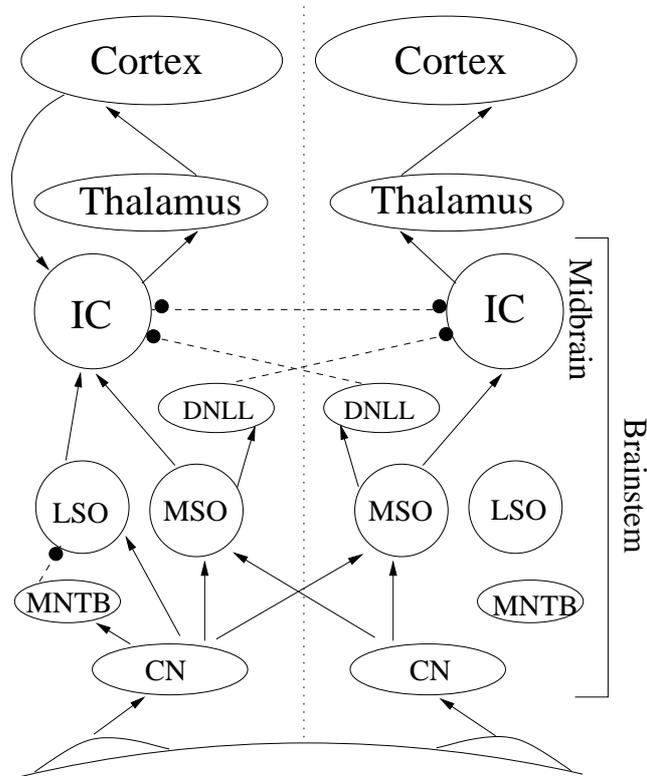


Figure 1: Schematic of some of the auditory pathways from the ears (at the bottom) to the cortex. Most of the auditory nuclei appear on both sides of the brain and are shown twice: to the right and to the left of the (dotted) midline. Abbreviations are as follows: CN – cochlear nucleus; MNTB – medial nucleus of the trapezoidal body; MSO – medial superior olive; LSO – lateral superior olive; DNLL – dorsal nucleus of the lateral lemniscus; IC – inferior colliculus. Excitatory connections are shown with solid lines; inhibitory connections — with dashed.

the mechanical signal is converted into the electrical one by auditory receptors, namely, hair cells. This transduction process is fundamental in all sensory systems as external signals of light, chemical compounds or sounds must be “translated” into the language of the central nervous system, the language of electrical signals.

Next, the signal originating from each cochlea is carried by the auditory nerve into the brainstem (see Figs. 1, 2). Figure 3 shows a 3-dimensional reconstruction of the human brainstem. It consists of three main parts: pons, medulla, and midbrain. The auditory nerve first synapses in the cochlear nuclei inside the medulla. From the cochlear nuclei, auditory information is split into at least two streams. One stream projects directly to the auditory part of the midbrain, inferior colliculus (Fig. 3); while the other goes through another set of nuclei in the medulla, called

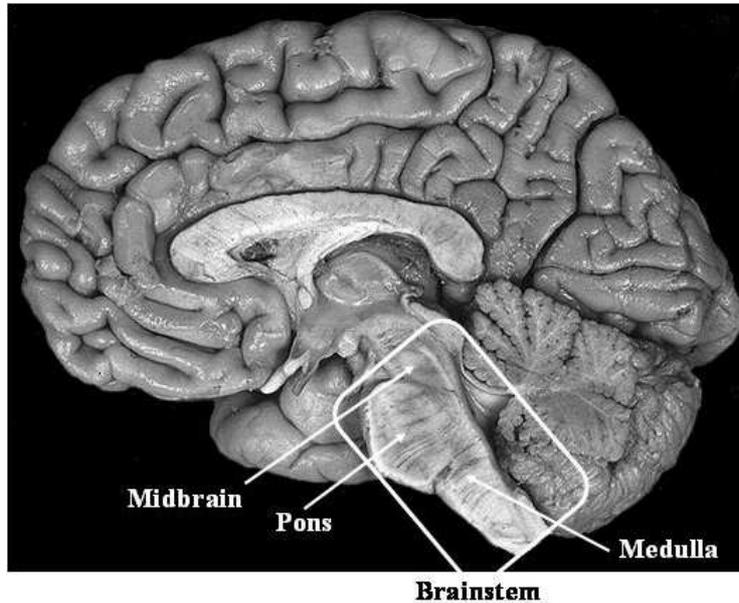


Figure 2: Location of brainstem in the human brain. Three main parts of the brainstem are indicated with arrows. (Picture of the brain is from Digital Anatomist Project [20], Department of Biological Structure, University of Washington, with permission)

superior olivary complex. The first of these pathways is thought to pick up the tiny differences in acoustic signals, that you need, for example, to differentiate between similar words. The indirect stream, which is the one that we are mostly going to consider, originates in the ventral cochlear nucleus. Auditory nerve fibers that bring information to this stream end with giant hand-like synapses. This tight connection allows the timing of the signal to be preserved up to a microsecond (which is very surprising, because the width of action potential, the “unit signal”, is on the order of milliseconds). This precisely timed information is carried on to the superior olive, both on the same side of the brain and across midline. In the superior olive the small differences in the timing and loudness of the sound in each ear are compared, and from this you can determine the direction the sound is coming from. The superior olive then projects up to the inferior colliculus.

From the inferior colliculus, both streams of information proceed to the sensory thalamus (Fig. 1-2). The auditory nucleus of the thalamus is the medial geniculate nucleus. The medial geniculate projects to the auditory cortex, located in the temporal lobes inside one of the folds of the cortex (Fig. 4).

It is important to notice that many cells at each level of auditory system are *tonotopically organized*. This means that in each of the auditory structures it is

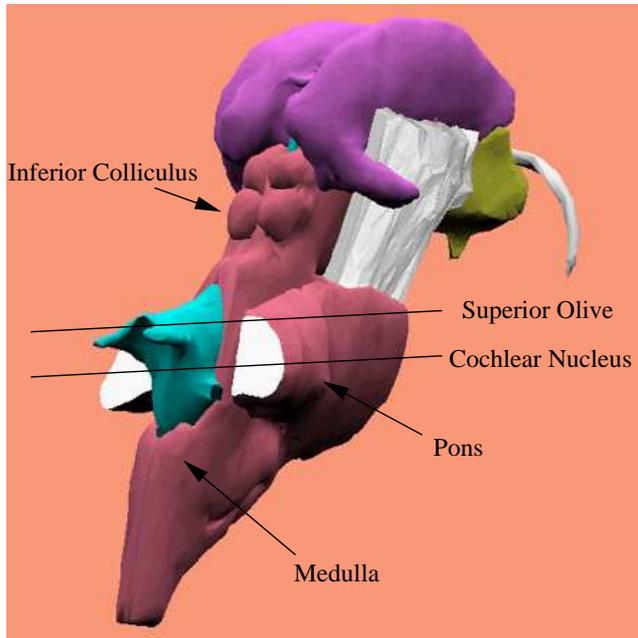


Figure 3: Reconstructed view of human brainstem (red), together with thalamus (purple) and hypothalamus (yellow) (From Digital Anatomist Project [20], Department of Biological Structure, University of Washington, with permission). Arrows point different parts of the brainstem, while lines indicate planes within which cochlear nuclei and superior olives are located.

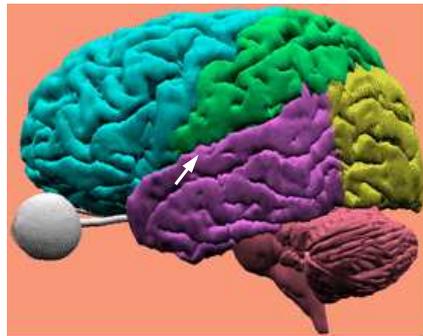


Figure 4: Lobes of the brain (From Digital Anatomist Project [20], Department of Biological Structure, University of Washington, with permission): frontal lobe (blue), parietal lobe (green), temporal lobe (purple), occipital lobe (yellow). White arrow points at the fold within which most of the auditory cortex is located.

possible to identify areas in which cells preferentially respond to sounds of certain frequencies (tones) and the preferred frequencies vary more-or-less continuously with position of the cell.

Notice also that both pathways starting at the cochlear nuclei are bilateral. The consequence of this is that localized brain lesions anywhere along the auditory pathway usually have no obvious effect on hearing. Deafness is usually caused by damage to the middle ear, cochlea, or auditory nerve. In infants the auditory system continues to develop throughout the first year of life. Neurons in the brainstem mature and many connections are just beginning to form (e.g. between brainstem nuclei, thalamic nuclei, and auditory cortex). As for other sensory systems, stimulation during this time is essential for normal development. When hearing is impaired in early life, the morphology and function of auditory neurons may be affected.

The overall structure of the sub-cortical auditory system and even some anatomical details are similar across many species. A lot of information about the auditory system has been obtained over the years from experiments, especially anatomical and electrophysiological, with cats, bats, gerbils, owls, monkeys and many other animals. Also, a large number of behavioral (psychophysical) studies have been conducted with animals as well as humans.

1.2 Sound characteristics

Here we describe some of the physical characteristics of the auditory stimuli and some of the perceptual characteristics that we ascribe to them.

Sounds are pressure waves that transfer energy and momentum from the source to places around it. One of their natural characteristics is the amount of energy transmitted by a wave in unit time, called the *power of the wave*. The power of a wave is proportional to the square of the frequency, the square of the amplitude and the wave speed. Often it is convenient to consider the power per unit area of the wavefront, i.e. the amount of energy transmitted by the wave perpendicularly to its wavefront in unit time per unit area. This is called the *intensity of a wave*. The sound waves can also be characterized by specifying their time-dependent spectra. The way these physical characteristics of sounds are usually described and quantified relates to our perception of the sounds.

The spectral properties of many sounds evoke a sensation of *pitch*. Pitch is defined as the auditory attribute on the basis of which tones may be ordered on a musical scale. The pitch of a tone is related to its frequency or periodicity. The pitch of a periodic sound wave (simple tone) is usually indicated by specifying its frequency in Hz. The pitch of a complex tone is usually indicated by the frequency of a simple tone whose pitch would be perceived to be the same. A simple tone evokes a sensation of pitch only if its frequency is between 20 Hz and 5 kHz, while

the spectrum of audible frequencies extends from 20 Hz to 20 kHz.

As a side note: in musical and psychophysical literature there are two different aspects of the notion of pitch. One is related to the frequency of a sound and is called *pitch height*; the other is related to the place in a musical scale and is called *pitch chroma*. Pitch height corresponds to the sensation of 'high' and 'low'. Pitch chroma, on the other hand, describes the perceptual phenomenon of octave equivalence, by which two sounds separated by an octave (and thus relatively distant in terms of pitch height) are nonetheless perceived as being somehow equivalent. Thus pitch chroma is organized in a circular fashion. Chroma perception is limited to the frequency range of 50-4000 Hz [122].

The intensity of the sound is commonly quantified on the logarithmic scale in Bels, or decibels. In this logarithmic scale the intensity I of a sound is quantified relative to the intensity I_0 of a reference sound, and is measured in Bels:

$$1 \text{ Bel} = \log_{10} \left(\frac{I}{I_0} \right),$$

or in decibels (dB):

$$1 \text{ dB} = \frac{1}{10} \log_{10} \left(\frac{I}{I_0} \right).$$

These expressions give the difference between the intensities in Bels or dB. For example, if I is hundred times I_0 , then the level of I is 2 Bel or 20 dB greater than that of I_0 . If we want to express an absolute intensity I using decibels then we have to use some standard reference intensity I_0 . The reference intensity most commonly used is 10^{-12} (Watt/m²), which corresponds to a pressure of 20 μ Pa or about $2 \cdot 10^{-10}$ atmosphere. The sound level of a sound relative to this standard intensity is called the *sound pressure level* or SPL of the sound. This standard reference intensity was chosen because it is very close to the faintest sound of 1 kHz that a human can detect. This lower limit to the detectable sounds is set by the Brownian motion of the molecules in the ear. They produce noise input to the auditory receptors, so the signal has to be strong enough to be detected on top of the noise. Notice that the sound pressure level can be negative, and that 0 dB sound does not mean the absence of a sound wave.

The reason for using the logarithmic scale is that when the sound intensity increases linearly from low to high, it creates the perception that the "loudness" (i.e. the perceived intensity) increases fast at first and then the increase slows down as the sound gets louder. Logarithmic scale is also useful because of the great range of intensities that the auditory system deals with. Dynamic range of human hearing is approximately 140 dB, ranging from a whisper to a jet engine. To give some examples, the sound level at an average home is about 40 dB, an average conversation is about 60 dB, and a loud rock band is about 110 dB.

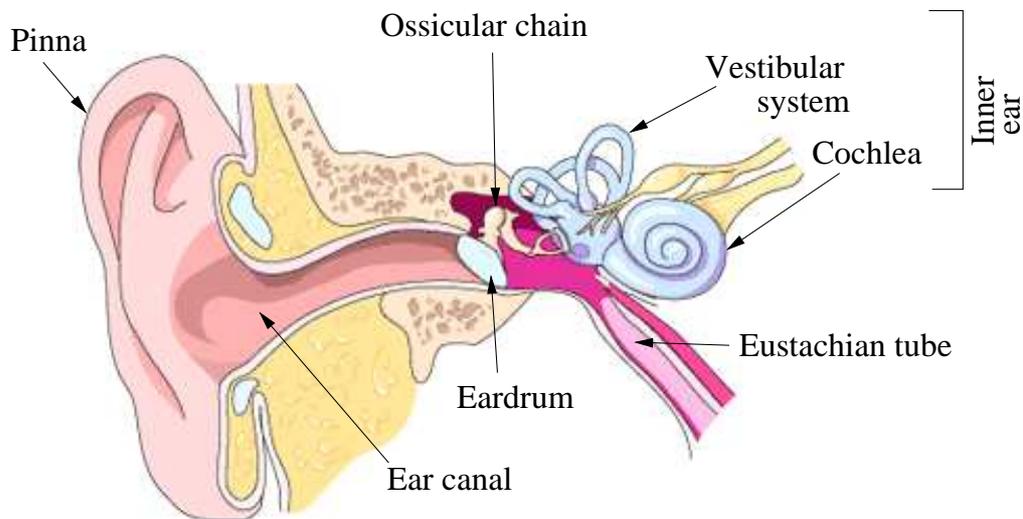


Figure 5: Peripheral auditory system. From <<http://www.workplacegroup.net/article-ear-anatomy.htm>>, permission requested.

2 Peripheral auditory system

Peripheral auditory system serves to transfer sounds from the environment to the brain. It is composed of three main components – outer, middle, and inner ear (Fig. 5).

2.1 Outer ear

The outer ear is the part that we see — *the pinna* (a structure made of cartilage that we call “the ear” in common speech) and *the ear canal*. A sound wave travelling in the air, before it reaches the entrance of the ear canal, must interact with the head, the torso, and the intrinsic shapes of the pinna. As a result, the waves can be amplified or attenuated, or their spectrum may be modified. These changes depend on the specific frequencies of the waves (e.g., they are most significant for waves of frequency above 1.5 kHz), but also on the three-dimensional position of the sound source. Our brain can use the difference of spectra between the two ears to determine the position of the original sound source. A typical way to characterize the pressure that an arbitrary sound produces at the eardrum is by using the, so called, Head-Related Transfer Functions (HRTFs). These are constructed by placing microphones at the entrance of the subject’s ear canal, and recording the generated pressure as sounds originate from point sources at various locations. The resulting functions, or their Fourier transforms – HRTF, are

surprisingly complicated functions of four variables: three space coordinates and frequency. Much research has been devoted to measurements, and computations of the external ear transformations, both for human and animal subjects (see, for example, [66],[89],[109] for humans, [68] for cats), and databases can be found on the web, for example at <http://www.ircam.fr/equipes/salles/listen/>, or http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm. The HRTFs are used in many engineering applications, most popularly in creation of realistic acoustic environments in virtual reality settings: games, virtual music halls, etc. Unfortunately, because of the diversity of the pinnae shapes and head sizes, HRTFs vary significantly from one person to another. For this reason, each of the existing models of acoustic fields, based on some average HRTF, is found adequate by only a relatively small number of listeners.

Having been collected and transformed by the pinnae, the sound waves are directed into the ear canal, where they travel towards *the eardrum* (also called *the tympanic membrane*). The ear canal maintains the proper temperature and humidity for the elasticity of the eardrum. It also contains tiny hairs that filter dust particles, and special glands that produce earwax for additional protection. The ear canal can also resonate sound waves and amplify tones in the 3000-4000 Hz range.

2.2 Middle ear

The eardrum (tympanic membrane), and the neighboring cavity with three tiny bones (*ossicular chain*) comprise the middle ear (Fig. 5). The sound waves travelling through the ear canal cause the vibrations of the tympanic membrane, and these vibrations are further transmitted through the chain of bones towards the inner ear. For the eardrum to function optimally, it should not be bent either inwards or outwards in the absence of the sound waves. This means that the air pressure should be the same in the outer and middle ear. The pressure equalization is achieved through the *eustachian tube* — a tube which connects the middle ear cavity with the back of the throat. Normally this tube is closed, but it opens with swallowing or shewing, thus working as a pressure equalizer. The middle ear bones (malleus, incus and stapes) are the smallest bones in the body. They work together as a lever system, to amplify the force of the vibrations. The malleus is attached to the tympanic membrane, the stapes enters the oval window of the inner ear, and the incus lies in between.

2.3 Inner ear. Cochlea. Hair cells.

The inner ear is a series of fluid-filled structures. It has two main parts: a part that is involved in balance and is called *vestibular system*, and a second part that is

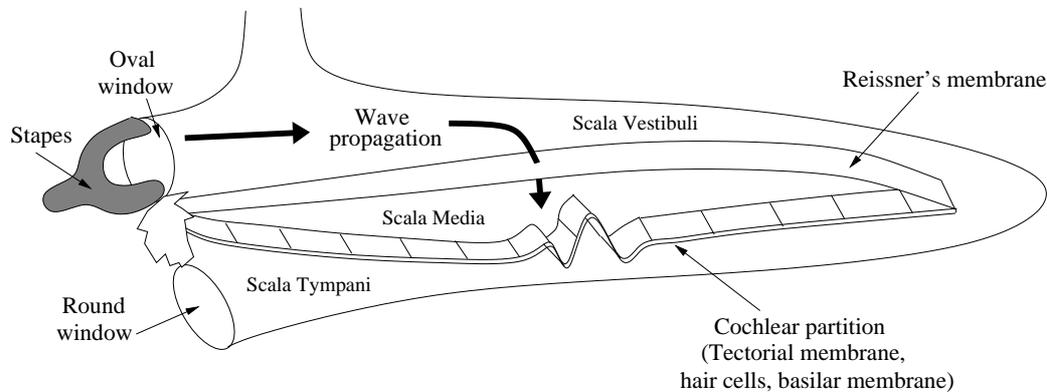


Figure 6: Schematic of a mammalian cochlea, straightened (see text).

responsible for converting sounds from mechanical vibrations into electrical signals, and is called *cochlea*. Cochlea, in mammals, is coiled as a snail's shell (Fig. 5). It has two membrane-covered openings into the middle ear: the *oval window* and the *round window*, and a membranous sac (containing receptor cells) that separates them (Fig. 6). The size of the oval window is 15 to 30 times smaller than that of the eardrum. This size difference produces amplification needed to match impedances between sound waves in the air and in the cochlear fluid. The principal function of the membranous cochlear sac is to act as a hydromechanical frequency analyzer.

Here is a rough outline of how it works (see Figure 6). The input from the middle ear arrives at the oval window and creates pressure fluctuations in the fluid. These fluctuations travel along the cochlea and eventually are dissipated by the movements of the large round window, which serves as a pressure release for incompressible fluids. As the pressure waves travel along the membranous sac, they permeate one of the walls (*Reissner's membrane*) and cause vibrations in the other (*cochlear partition*). The platform of the cochlear partition (*basilar membrane*) changes in its mechanical properties along its length from being narrow and stiff at the base of the cochlea, to being wide and compliant at the apex. Therefore, the lower the frequency of the tone the further from the oval window the vibration pattern is located.

Hair cells

Looking at higher resolution inside the cochlear partition (Fig. 7A), there are auditory sensory cells (*inner and outer hair cells*) that sit on the basilar membrane and have their stereocilia sticking out into the fluid and attached to a floppy *tectorial*

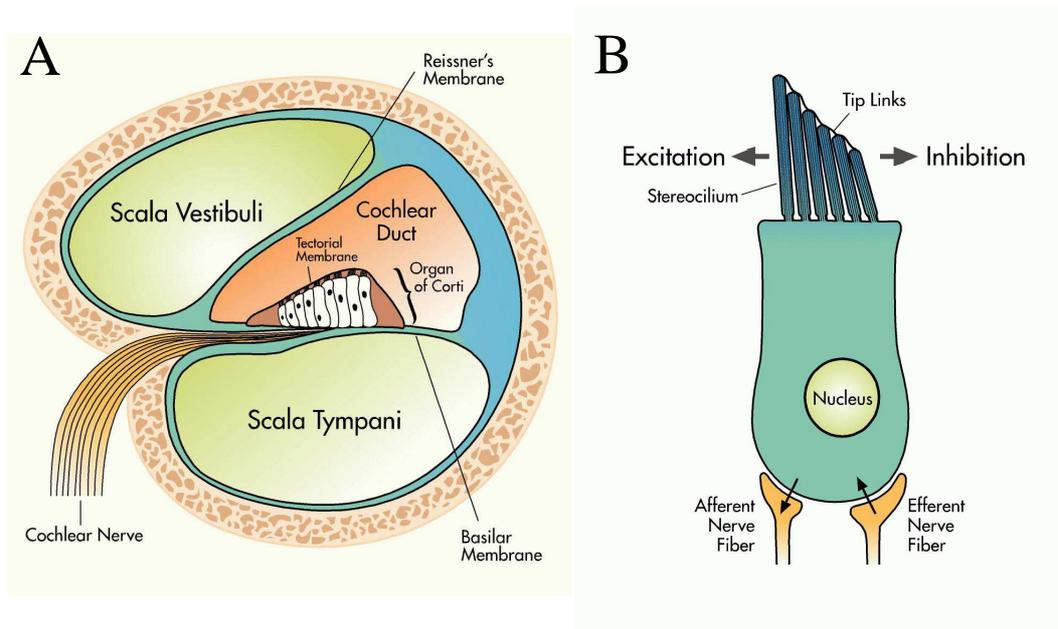


Figure 7: Location and basic structure of hair cells. *A*: Radial segment of cochlea showing main components of the cochlear partition. *B*: Schematic of a hair cell. Bold arrows indicate that excitation and inhibition of the cell voltage can be produced by deflections of the hair bundle (see text). Illustration is made by Brook L. Johnson.

membrane. Because the basilar and tectorial membranes attach to the bone at different points, their motion causes the tectorial membrane to slide across the basilar membrane, tilting the hair bundle (Fig. 7B). Using electrophysiological recordings and mechanical stimulation, it has been demonstrated that deflection of the stereocilia leads to a change of the membrane potential of the cell. Deflection towards the highest stereocilia leads to a depolarization of the cell. It is also shown that both the initial change in charge and calcium concentration occur near the tips of the stereocilia, suggesting that the transduction channels are located at the tips. Moreover, the calcium entry and charge change happens relatively fast ($10 \mu\text{sec}$), which suggests that the transduction channels are gated mechanically rather than through a second messenger system.

There are two types of auditory sensory cells: one row of inner hair cells (IHCs; they are called “inner” because they are closer to the bony central core of the twisted cochlea), and three to five rows of outer hair cells (OHCs). Inner and outer cells have different functions, and different sets of incoming and outgoing connections. Each inner hair cell’s output is read out by twenty or so *cochlear afferent neurons*

of *type I* (for characteristics of types of afferent neurons see below), and each type I afferent neuron only contacts a single inner hair cell. The output of each outer hair cell, on the other hand, is read together with many other outer hair cell outputs, by several *type II afferent neurons*. All afferent neurons send their axons to the cochlear nuclei of the brainstem and terminate there. Outer hair cells also receive descending (efferent) signals from neurons with cell bodies in the brainstem. The afferent and efferent axons together form the auditory nerve — the communication channel between the peripheral auditory system and the auditory regions of the brain.

Inner hair cells convey frequency, intensity and phase of signal. As explained above, to a first approximation, the frequency is encoded by the identity of the activated inner hair cells, i.e. those inner hair cells that are located at the appropriate place along the cochlea. The intensity of the signal is encoded by the DC component of the receptor potential; and the timing — by the AC component (see *Phase locking* below).

The function of the outer hair cells is not presently known. Two main theories are that they act as cochlear amplifiers, and that they act to affect the movement of the tectorial membrane. Another theory is that they function as motors, which alter the local micromechanics to amplify the sensitivity and frequency selectivity of the cochlea. It seems that the outer hair cells participate in the mechanical response of the basilar membrane, because loss of outer hair cells leads to a loss of sensitivity to soft sounds and decrease in the sharpness of tuning.

2.4 Mathematical modeling of the peripheral auditory system

Mathematical and computational modeling of the peripheral auditory system has a long history. For mathematical analysis and models of the external and middle ear see, e.g. review by Rosowski [84]. More recently, there has been a number of three-dimensional models of the middle ear that use the finite-element method (e.g. [44, 27, 18, 22, 25, 41, 103]). Some of these models have been successfully used for clinical applications (e.g. [41, 18]).

An even larger set of studies concentrated on the modeling of the cochlea (for reviews see, for example, [40, 28]). Early models represented cochlea as one- or two-dimensional structure and incorporated only a few elements of the presently known cochlear mechanics. Early one-dimensional models of the cochlea [23, 76] have assumed that the fluid pressure is constant over a cross-section of the cochlear channel. The fluid is assumed to be incompressible and inviscid, and the basilar membrane is modeled as a damped, forced harmonic oscillator with no elastic coupling along its length. Qualitatively, this model has been shown to capture the basic features of the

basilar membrane response. Quantitatively, however, it yields large discrepancies with measurement results [120]. Two-dimensional models by Ranke [78] and Zwislocki [121] make similar assumptions on the cochlear fluid and the basilar membrane. Ranke's model uses a deep water approximation, while Zwislocki used the shallow water theory in his model. These models were further developed in [3, 4, 53, 92] and in other works. Other two-dimensional models incorporate more sophisticated representations of the basilar membrane using, for example, elastic beam and plate theory [10, 17, 36, 37, 43, 99]. Three-dimensional models were considered by Steele and Taber [100] and de Boer [19], who used asymptotic methods and computations and obtained an improved fit of the experimental data. Their work seems to indicate that geometry may play a significant role in the problem. In particular, the effect of the spiral coiling of the cochlea on the wave dynamics remains unresolved; see [101, 105, 56, 60].

With the development of more powerful computers it became possible to construct more detailed computational models of the cochlea. A two-dimensional computational model of the cochlea was constructed by Beyer [7]. In this model the cochlea is a flat rectangular strip divided into two equal halves by a line which represents the basilar membrane. The fluid is modelled by the full Navier-Stokes equations with a viscosity term, but elastic coupling along the basilar membrane is not incorporated. Beyer has used a modification of Peskin's immersed boundary method, originally developed for modeling the fluid dynamics of the heart [75]. Several three-dimensional computational models have been reported, such as Kolston's model [45], intended to simulate the micro-mechanics of the cochlear partition in the linear regime (i.e., near the threshold of hearing), Parthasarati, Grosh and Nuttall's hybrid analytical-computational model using WKB approximations and finite-element methods, and Givelberg and Bunn's model [28] using the immersed boundary method in a three-dimensional setting.

We will consider as an example the analysis of a two-dimensional model with fluid viscosity, by Peskin [73, 74]. In this model cochlea is represented by a plane and the basilar membrane by an infinite line dividing the plane into two halves. The fluid in this model satisfies the Navier-Stokes equations with the non-linearities dropped. The distinctive feature of this model is that the location of wave peak for a particular sound frequency is strongly influenced by fluid viscosity and the negative friction of the basilar membrane. This model was studied with asymptotic and numerical methods in [54, 55]. In the example that we present here, the zeroth order approximation of the solution is found by WKB method. (The WKB method was first used for this cochlear model by Neu and Keller [69] in the case of zero membrane friction).

Problem setup: Basilar membrane at rest is situated along the x -axis. The deviation of the basilar membrane from the axis will be denoted by $h(x, t)$. Vector

(u, v) is the velocity of the fluid, and p is the pressure of the fluid (these are functions of (x, y, t)).

Assumption 1: the fluid satisfies the Navier-Stokes equations with the nonlinearities left out, i.e.

for $y \neq 0$:

$$\rho \frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} = \mu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), \quad (1)$$

$$\rho \frac{\partial v}{\partial t} + \frac{\partial p}{\partial y} = \mu \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right), \quad (2)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0. \quad (3)$$

In these equations, ρ is the density and μ is the viscosity of the fluid. The last equation represents the incompressibility of the fluid.

Notice that there is no explicit forcing term in this model. Yet, there are bounded wave solutions that move in the direction of the increasing x as if there is a source of vibration at $x = -\infty$.

Assumption 2:

- a) the basilar membrane (located at $y = 0$) has zero mass;
- b) there is no coupling along the membrane;
- c) each point of the membrane feels a restoring force that is proportional to the displacement and to the compliance (flexibility) of the membrane, the latter given by $e^{\lambda x}$ (to represent the fact that the actual membrane is more narrow and more flexible at the far end);
- d) the membrane possesses an active mechanism (mechanical amplifier, parameter β below is negative), i.e.

for $y = 0$:

$$u(x, 0, t) = 0, \quad (4)$$

$$v(x, 0, t) = \frac{\partial h}{\partial t}(x, t), \quad (5)$$

$$p(x, 0^-, t) - p(x, 0^+, t) = s_0 e^{-\lambda x} \left(h + \beta \frac{\partial h}{\partial t} \right) (x, t). \quad (6)$$

Note that the boundary conditions are applied at the rest position of the membrane $y = 0$, not at its instantaneous position $h(x, t)$. This, as well as the above assumptions is justified by the small displacements of the membrane and the fluid particles in the cochlea. Notice also that there is an unknown function $h(x, t)$ in the boundary conditions. Finding this function such that the rest of the system has a bounded

solution (u, v, p) is part of the problem. In addition, the problem is complicated by presence of the stiffness term. The parameter λ was measured (for dead cochlea) by von Békésy: $\lambda^{-1} \approx 0.7$ cm, i.e. over the length of the human cochlea (length ≈ 3.5 cm) the stiffness more than doubles.

Because of the symmetry of the system, we look for solutions that satisfy

$$p(x, y, t) = -p(x, -y, t),$$

$$u(x, y, t) = -u(x, -y, t),$$

$$v(x, y, t) = v(x, -y, t).$$

Then we can restrict the system of equations (1-3) to $y < 0$ and, using notation $y = 0$ instead of $y = 0^-$, re-write the boundary conditions (4-6) as

$$u(x, 0, t) = 0, \tag{7}$$

$$v(x, 0, t) = \frac{\partial h}{\partial t}(x, t), \tag{8}$$

$$2p(x, 0, t) = s_0 e^{-\lambda x} \left(h + \beta \frac{\partial h}{\partial t} \right) (x, t). \tag{9}$$

We also impose the condition

$$(u, v, p) \rightarrow 0 \text{ as } y \rightarrow -\infty.$$

Solution plan: We want to determine the solution that represents the steady state response of the cochlea to a pure tone. We will look for this solution as a small perturbation from the time-periodic function of the same frequency as the tone (and thus introduce a small parameter in the system). Further, we will use asymptotic expansion in the small parameter and by solving the zeroth order system of equations we will find an approximation of the solution.

We use a change of variables

$$\begin{pmatrix} u \\ v \\ p \end{pmatrix} (x, y, t, \epsilon) = \begin{pmatrix} U \\ V \\ P \end{pmatrix} (x - x_\epsilon, y/\epsilon, \epsilon) e^{i(\omega t + \frac{\Phi(x-x_\epsilon)}{\epsilon})},$$

$$h(x, t, \epsilon) = H(x - x_\epsilon, \epsilon) e^{i(\omega t + \frac{\Phi(x-x_\epsilon)}{\epsilon})},$$

where Φ is the local spatial frequency, ω is the given frequency of the external pure tone stimulus (radians/second), functions U, V, P, H, Φ are complex-valued, and the parameters ϵ and x_ϵ will be chosen later on. We set

$$X = x - x_\epsilon,$$

$$Y = y/\varepsilon,$$

and

$$\xi(X) = \Phi'(X) = \frac{\partial \Phi}{\partial X}(X).$$

In terms of the new variables:

$$\frac{\partial u}{\partial t} = i\omega \cdot U \cdot e^{i(\omega t + \frac{\Phi(x-x_\varepsilon)}{\varepsilon})},$$

$$\frac{\partial u}{\partial x} = \left[\frac{\partial U}{\partial X} + \frac{i}{\varepsilon} \xi(X) U \right] e^{i(\omega t + \frac{\Phi(x-x_\varepsilon)}{\varepsilon})},$$

$$\frac{\partial u}{\partial y} = \frac{1}{\varepsilon} \frac{\partial U}{\partial Y} e^{i(\omega t + \frac{\Phi(x-x_\varepsilon)}{\varepsilon})},$$

$$\begin{aligned} \Delta u = \Delta U = & \\ = & \left[\frac{\partial^2 U}{\partial X^2} + \frac{i}{\varepsilon} (\xi' U + 2\xi \frac{\partial U}{\partial X}) + \frac{1}{\varepsilon^2} \left(\frac{\partial^2 U}{\partial Y^2} - \xi^2 U \right) \right] e^{i(\omega t + \frac{\Phi(x-x_\varepsilon)}{\varepsilon})}. \end{aligned}$$

Then the equations (1-3) can be rewritten as:

for $Y < 0$:

$$(i\omega\rho - \mu\Delta)U + \left(i\frac{\xi}{\varepsilon} + \frac{\partial}{\partial x} \right) P = 0, \quad (10)$$

$$(i\omega\rho - \mu\Delta)V + \frac{1}{\varepsilon} \frac{\partial P}{\partial Y} = 0, \quad (11)$$

$$\left(i\frac{\xi}{\varepsilon} + \frac{\partial}{\partial x} \right) U + \frac{1}{\varepsilon} \frac{\partial V}{\partial Y} = 0, \quad (12)$$

with boundary conditions

for $Y = 0$:

$$\begin{aligned} U(X, 0, \varepsilon) &= 0, \\ V(X, 0, \varepsilon) &= i\omega H(X, 0, \varepsilon), \\ 2P(X, 0, \varepsilon) &= s_0(1 + i\omega\beta)e^{-\lambda x_\varepsilon} e^{-\lambda x} H(X, 0, \varepsilon), \\ (U, V, P) &\rightarrow 0 \quad \text{as } Y \rightarrow -\infty. \end{aligned}$$

Assumption 3. The functions U, V, P, H have the expansion

$$U = U_0 + \varepsilon U_1 + \dots,$$

$$\begin{aligned}
V &= V_0 + \varepsilon V_1 + \dots, \\
P &= \varepsilon(P_0 + \varepsilon P_1 + \dots), \\
H &= H_0 + \varepsilon H_1 + \dots.
\end{aligned}$$

We now choose

$$\varepsilon^2 = \frac{\mu\lambda^2}{\rho\omega}, \quad e^{-\lambda x_\varepsilon} = \varepsilon.$$

This choice of ε makes it indeed a small parameter for realistic values of other quantities. For example, if the fluid has characteristics of water ($\rho = 1 \text{ g/cm}^3$, $\mu = .02 \text{ g/(cm}\cdot\text{s)}$), for a 600 Hz tone ($\omega = 2\pi \cdot 600/\text{s}$) and $1/\lambda = .7 \text{ cm}$, $\varepsilon \approx .003$.

If we substitute the expansions of U, V, P, H into the equations (10-12) and collect terms with matching powers of ε , we find at the zeroth order:

for $Y < 0$:

$$\begin{aligned}
\rho\omega \left(i - \frac{1}{\lambda^2} \left(-\xi^2 + \frac{\partial^2}{\partial Y^2} \right) \right) U_0 + i\xi P_0 &= 0, \\
\rho\omega \left(i - \frac{1}{\lambda^2} \left(-\xi^2 + \frac{\partial^2}{\partial Y^2} \right) \right) V_0 + \frac{\partial P_0}{\partial Y} &= 0, \\
i\xi U_0 + \frac{\partial V_0}{\partial Y} &= 0,
\end{aligned}$$

and for $Y = 0$:

$$\begin{aligned}
U_0(X, 0) &= 0, \\
V_0(X, 0) &= i\omega H_0(X), \\
2P_0(X, 0) &= s_0(1 + i\omega\beta)e^{-\lambda x} H_0(X), \\
(U_0, V_0, P_0) &\rightarrow 0 \text{ as } Y \rightarrow -\infty.
\end{aligned}$$

Next, at the first order

for $Y < 0$:

$$\begin{aligned}
\rho\omega \left(i - \frac{1}{\lambda^2} \left(-\xi^2 + \frac{\partial^2}{\partial Y^2} \right) \right) U_1 + i\xi P_1 &= \frac{\rho\omega}{\lambda^2} i \left(\xi' + 2\xi \frac{\partial}{\partial X} \right) U_0 - \frac{\partial}{\partial X} P_0, \\
\rho\omega \left(i - \frac{1}{\lambda^2} \left(-\xi^2 + \frac{\partial^2}{\partial Y^2} \right) \right) V_1 + \frac{\partial P_1}{\partial Y} &= \frac{\rho\omega}{\lambda^2} i \left(\xi' + 2\xi \frac{\partial}{\partial X} \right) V_0, \\
i\xi U_1 + \frac{\partial V_1}{\partial Y} &= -\frac{\partial}{\partial X} U_0,
\end{aligned}$$

and for $Y = 0$:

$$\begin{aligned} U_1(X, 0) &= 0, \\ V_1(X, 0) &= i\omega H_1(X), \\ 2P_1(X, 0) &= s_0(1 + i\omega\beta)e^{-\lambda X}H_1(X), \\ (U_1, V_1, P_1) &\rightarrow 0 \text{ as } Y \rightarrow -\infty. \end{aligned}$$

Consider the zeroth order equations. For each fixed x , the functions P_0, U_0, V_0 satisfy an ODE system whose solution is given by

$$\begin{aligned} P_0(X, Y) &= P_0(X, 0)e^{\sqrt{\xi^2}Y}, \\ U_0(X, Y) &= -P_0(X, 0)\frac{i\xi}{i\omega\rho} \left(e^{\sqrt{\xi^2}Y} - e^{\sqrt{\xi^2+i\lambda^2}Y} \right), \\ V_0(X, Y) &= -P_0(X, 0)\frac{\xi^2}{i\omega\rho} \left(\frac{e^{\sqrt{\xi^2}Y}}{\sqrt{\xi^2}} - \frac{e^{\sqrt{\xi^2+i\lambda^2}Y}}{\sqrt{\xi^2+i\lambda^2}} \right), \end{aligned}$$

where $\sqrt{}$ denotes the root with positive real part.

Next, the zeroth order equations that contain H_0 give two different formulae for H_0 in terms of $P_0(X, 0)$:

$$H_0(X) = \frac{V_0(X, 0)}{i\omega} = P_0(X, 0)\frac{\xi^2}{\omega^2\rho} \left(\frac{1}{\sqrt{\xi^2}} - \frac{1}{\sqrt{\xi^2+i\lambda^2}} \right) \quad (13)$$

and

$$H_0(X) = P_0(X, 0)\frac{2e^{\lambda X}}{s_0(1+i\omega\beta)}. \quad (14)$$

We have just expressed our zeroth order solution $U_0(X, Y), V_0(X, Y), P_0(X, Y), H_0(X)$ in terms of $P_0(X, 0)$ and $\xi(X)$. Now there are two steps remaining. Step 1 is to find a function $\xi(X)$ such that both equations for $H_0(X)$ are satisfied. Step 2 is to find $P_0(X, 0)$.

Step 1. For ξ to satisfy both (13) and (14) we need to have

$$Q(X) = \xi^2 \left(\frac{1}{\sqrt{\xi^2}} - \frac{1}{\sqrt{\xi^2+i\lambda^2}} \right), \quad (15)$$

where

$$Q(X) = \frac{2\rho\omega^2 e^{\lambda X}}{s_0(1+i\omega\beta)}.$$

The condition (15) is called a dispersion relation. It implicitly defines the local spatial frequency ξ in terms of X and ω . Equation (15) may have multiple solutions, for example, if ξ is a solution, then $-\xi$ is also a solution (a wave going in the opposite direction). Using notation $\eta = \sqrt{\xi^2}$ we can transform (15) into

$$\eta^3 - \frac{1}{2}\left(Q + \frac{i\lambda^2}{Q}\right)\eta^2 + i\lambda^2\eta - \frac{1}{2}i\lambda^2Q = 0.$$

But not every zero of this cubic satisfies the dispersion relation. First of all the real part of η has to be non-negative. Second, by definition of Q : $Q < \eta$, therefore only the zeros of the cubic such that

$$Q = \eta - \frac{\eta^2}{\sqrt{\eta^2 + i\lambda^2}}$$

satisfy the original equation. For every root of cubic that satisfies these two conditions we can then take $\xi(X) = \pm\eta$.

Step 2. Now we need to find $P_0(X, 0)$. If we multiply the first order equations by functions $-U_0, V_0, P_0$, respectively (which represent solution to the zeroth order equations with ξ replaced by $-\xi$), and integrate each equation over $Y \in (-\infty, 0)$, then we obtain after integration by parts and combining non-vanishing terms

$$\frac{\partial}{\partial X} \int_{-\infty}^0 \left[\frac{i\omega\rho\xi}{\lambda^2} (V_0^2 - U_0^2) + U_0P_0 \right] dY = 0,$$

i.e.

$$\int_{-\infty}^0 \left[\frac{i\omega\rho\xi}{\lambda^2} (V_0^2 - U_0^2) + U_0P_0 \right] dY = C_0,$$

where C_0 is a constant independent of X . Notice that all terms containing U_1, V_1 and P_1 have disappeared because of the vanishing factors in front of them. Now substitute U_0, V_0 and P_0 by their expressions in terms of $P(X, 0)$, and integrate. We get

$$(P_0(X, 0)^2) = \frac{2\omega\rho C_0 (\sqrt{\xi^2 + i\lambda^2})^3 \sqrt{\xi^2}}{\xi (\sqrt{\xi^2 + i\lambda^2} - \sqrt{\xi^2}) (\sqrt{\xi^2 + i\lambda^2} \sqrt{\xi^2 - i\lambda^2})},$$

which gives us the solution.

3 Auditory Nerve (AN)

3.1 AN structure

The auditory nerve is a collection of axons connecting the peripheral auditory system and the auditory areas of the brain. It is made up of approximately 30,000 to 55,000

nerve fibers, depending on species. About 95% of them are afferent, projecting from cell bodies in the cochlea to cochlear nuclei in the brainstem, and the rest are efferent, coming to cochlea from cells in the olivary complex (also part of the brainstem, see below). The afferent neurons are divided into Type I and Type II, based on their morphology: type I cells are large, have bipolar shape, their axons are large and myelinated, and, therefore, fast; type II cells are smaller, have different shape and non-myelinated axons. In addition, type I fibers innervate inner hair cells in many-to-one fashion, and type II fibers innervate many-to-many outer hair cells. Very little is known about the functional properties of type II afferents, partially because type I fibers are much easier for physiologist to record from, due to their large size and number. The role of efferent neurons in modifying the auditory input is also not yet clear. So, the rest of this text will focus on type I afferents.

3.2 Response properties

Spontaneous rates

In mammals, afferents can be divided into low, medium and high spontaneous rate fibers. The spontaneous rate (firing rate in the absence of stimuli) is determined by pattern of hair cell innervation, although the mechanisms are unclear. It could be that either smaller size of low spontaneous rate fibers makes them less excitable, or less transmitter is released from the hair cells to low spontaneous rate fibers. The variety of available fiber sensitivity provides a way of encoding wide range of intensities in auditory nerve (see *Intensity sensitivity* below).

Thresholds

The threshold is defined as minimal intensity of the stimulus that increases firing rate above the spontaneous level. Note that this is not a real threshold because, for example, for very low frequency fibers, sound first synchronizes spikes before there is ever an increase in rate. Usually the threshold is about 1 dB in mammals.

Latencies

Auditory fibers are also characterized by latencies in their responses. The latencies can be measured using brief stimuli, for example, clicks. Fibers tuned to different frequencies respond to a click at different latencies. Most of the delay originates from travel time of wave along basilar membrane. High frequency areas on the membrane are stimulated first, i.e. high frequency cells have shorter latencies.

Frequency tuning

If we fix the intensity of the stimulus at any particular level and look at the firing rate of the given fiber at different frequencies, we find that the response function is

usually single-peaked. This means that the cell has a preferred range of frequencies, inherited from its innervation of cochlea.

A typical way to characterize the basic properties of an auditory cell, is to probe its responses to a variety of *pure tones* (stimuli with only one frequency component and constant intensity). These responses are often summarized by marking the areas of intensity and frequency that produce a change in the firing rate of the cell (*response areas*). The border of this area is called a *tuning curve* of the cell. For auditory nerve fibers (Fig. 8) these areas usually have triangular shape, pointing down, and most of the area corresponds to an increase in firing rate (excitatory). The frequency at which the tip of the response area is located, i.e. the frequency at which the cell is most sensitive, is called the cell's characteristic frequency (CF) or best frequency (BF). The CF of each cochlear afferent is well-defined and it provides accurate information about the position on the cochlea of the hair cell that the afferent innervates. Approximately, the linear distance on the cochlea is proportional to the logarithm of CF. The shapes of the response areas change systematically with CF. The response areas of high-CF fibers have very steep high-frequency slope, and elongated tip (Fig. 8). Response areas of lower-CF afferents are relatively broader and more symmetrical. Notice that often relative bandwidth (normalized to CF), rather than absolute bandwidth, is used to characterize the sharpness of tuning; and relative bandwidth increases with CF. A common measure of frequency tuning is the Q_{10} , defined as the CF divided by the bandwidth at 10 dB above CF threshold. Q indexes also get higher with increasing CF.

Intensity tuning

Next property of the auditory nerve fibers, that we will be discussing, is the intensity tuning. Let us keep the frequency constant and consider responses, in terms of firing rate, at different intensity levels (rate-intensity function). This function is generally monotonically increasing over 40-50 dB above threshold, and then saturates. The range over which this function is increasing is termed the *neuron's dynamic range*. Maximum dynamic range is usually at CF. Notice that the dynamic range of individual cells (40 dB) is 10^7 times smaller than the total range of human hearing (110 dB). To build such a wide range out of small-range elements, the auditory system makes use of the diversity in sensitivity of individual cells. The threshold levels of less sensitive fibers are situated at the upper end of the range of the more sensitive ones. In this way one set of neurons is just beginning to fire above their spontaneous rate when the other group is beginning to fire at their maximal rate. Additionally, at very high sound levels, the frequency tuning properties of the cochlea break down. Thus, we are able to tell that a 102 dB sound is louder than a 100 dB sound, but it is hard for us to determine whether it has a different frequency.

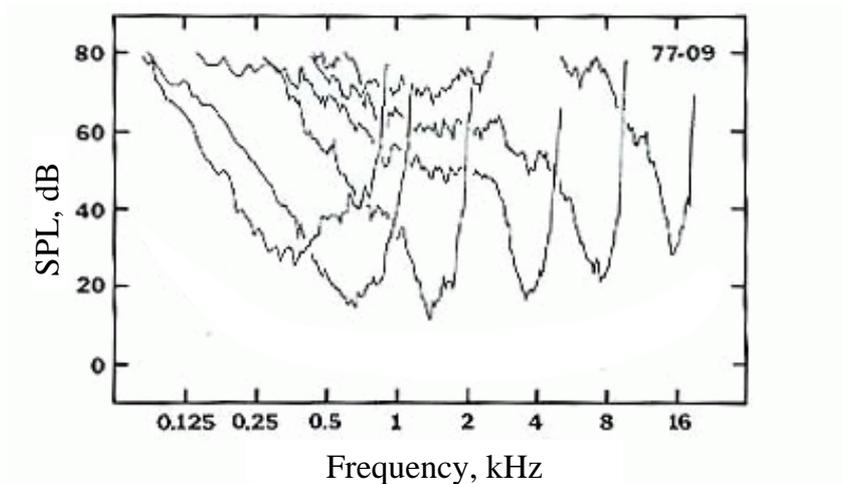


Figure 8: Examples of frequency-threshold tuning curves for chinchilla auditory nerve fibers. Area above each curve is the response area of the given fiber. Reprinted by permission of VCH Publishers, Inc. from [86].

Phase Locking (transmission of timing information)

Phase-locking of one (potentially stochastic) process with respect to another (often periodic) means that the events of the former preferentially occur at certain phases of the latter; in other words, that there is a constant phase shift between the two.

In the presence of a pure tone (periodic) stimulus, due to cochlear structure, inner hair cells are stimulated periodically by the pressure waves. If the frequency is low enough, then the depolarizations of the hair cell are also periodic and occur at certain phases of the stimulus cycle. This, in turn, generates a phase-locked release of neurotransmitter (that carries the neuronal signal to other cells) and, ultimately, leads to spikes in the auditory nerve that are also phase-locked to the stimulus cycle.

To observe or measure the phase-locking, people often build a histogram of phases of the recorded spikes within a period of the stimulus (Fig. 9). To have a meaningful resolution, given that stimuli frequencies can be quite high, bin widths have to be in the microsecond range. Phase-locking is equivalent to spikes accumulating at some parts of the periodogram, i.e. forming a peak.

A typical way to quantify phase-locking is by computing vector strength of the period-histogram [30]. Vector strength of a periodic function is the norm of the first complex Fourier component, normalized by the norm of the zeroth Fourier component. Normalization is included to remove dependence on the mean firing rate. This quantity can vary between 0 and 1, with 0 indicating low phase locking (for example,

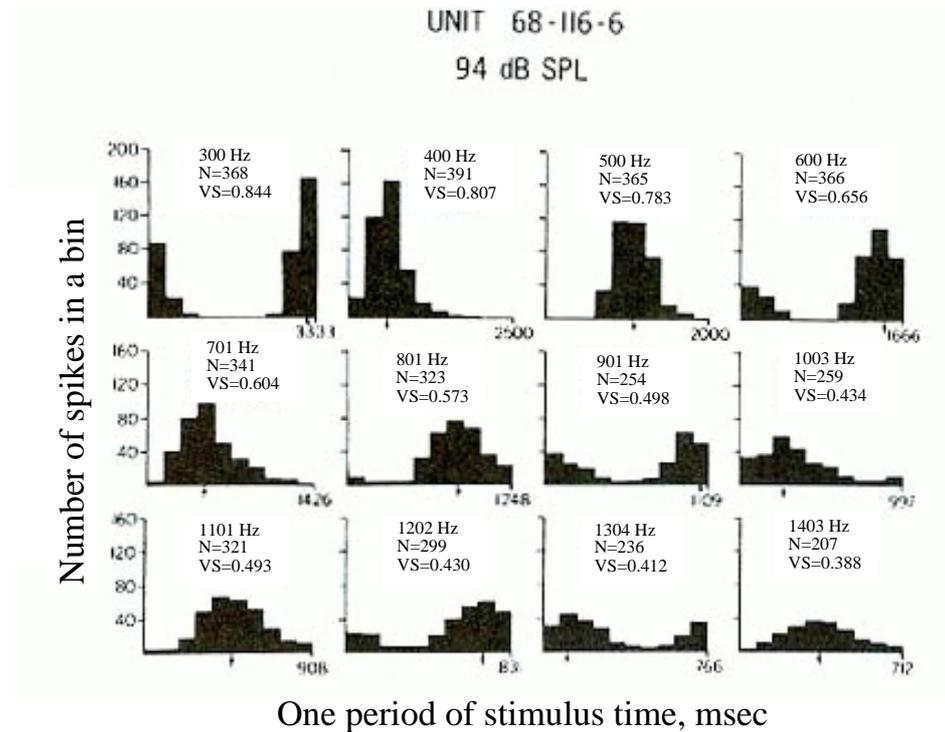


Figure 9: Periodograms of responses of an auditory nerve fiber. Each panel shows response to one stimulus frequency, and horizontal axis spans one stimulus period. Above the histogram is the stimulus frequency, total number of spikes (N) and the vector strength (VS). Reprinted by permission of Acoustical Society of America from [5].

in case of uniform distribution of spikes), and 1 indicating high phase locking (when all spikes fall in the same bin). Figure 9 shows an example of periodograms and vector strengths of an auditory fiber in response to stimuli of different frequencies.

Phase-locking generally occurs for stimuli with frequencies up to 4-6 kHz. Even high frequency fibers, when stimulated with low frequencies at intensities 10-20 dB below their rate threshold, will show phase locking of spikes, without increase in firing rate [39]. Decline of phase-locking with increase in frequency originates in the parallel reduction of the AC component of the inner hair cell response.

Phase-locking plays an important role in our ability to localize sounds, particularly at low frequencies, at which interaction of sound waves with the body and the pinna is reduced. We are able to tell which direction a sound is coming from (its azimuth), based on time delay between when it reaches our right and left ears. The idea is that the sound, originating on one side of the head, must follow a longer path and consequently takes a longer time traveling to one ear than to the other. Then,

due to the phase locking, the time delay is translated into the phase shift between spike trains, originating in the different sides of the brain. These delays are detected in some parts of the auditory pathway superior olivary complex, with precision as small as 20 microseconds. We will come back to the issue of phase-locking and sound localization when we talk about superior olivary complex below.

3.3 How is AN activity used by brain?

One of the basic questions one can ask about how the AN signal is decoded is the following: how does a neuron differentiate between change in rate due to change in intensity vs change in rate due to change in frequency? While the complete answer is not yet clear, it seems that the key is to look not at the responses of the single cell, but at the distribution of the responses across a population. The basic view is that the loudness can be carried by the total level of the population activity, and the frequency — by the position of the activity peak within the population (place principle) and/or by the temporal structure of the spike trains (using phase-locking — volley principle).

In this simple view, one considers complex sounds as made up of sinusoids (Fourier components). This linear approximation works well (at normal acoustic pressures) for the outer ear, less well for middle ear and only partially for the inner ear. The inner ear, anatomically, breaks sound into bands of frequencies, that persist for many stages of processing. How the information from these different bands is put back together is unknown.

In reality things are not linear. We will give two examples of non-linearities. The first example shows that the responses of AN fibers to tones depend on the spectral and temporal context. We illustrate this by describing a two-tone suppression: response of AN fiber to a given tone can be suppressed by a presence of another tone of a given frequency. The second example illustrates the non-linearity in a phenomenon of “hearing the missing fundamental”: tone combinations can create perception of frequencies that are not really present in the signal. Now we describe these examples in more detail.

Example 1. *Two-tone suppression* occurs in recordings from cochlea or afferents. It is defined as reduction in response to one tone in presence of a second tone. It depends upon levels and frequencies of the two tones. To illustrate this phenomenon, one can map response area of a cell (Fig. 10), then choose a test tone near CF (triangle in Fig. 10), and mark by shading the tonal stimuli that suppress responses to the test tone. Regions of overlap of the shading with the response area show that some tones that are excitatory by themselves, can suppress responses to CF tones. Latency for suppression is as short as initial response, i.e. it is not caused by efferents. It is shown that major component of the suppression originates from

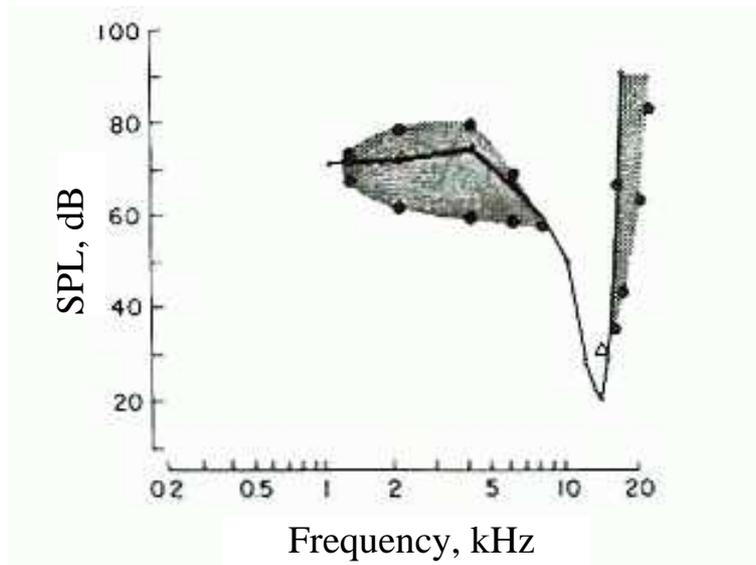


Figure 10: **Example of two-tone suppression.** Tuning curve (black line), showing the lower border of the response area of an auditory nerve fiber. Response to test tone (triangle) can be suppressed by any of the tones in the shaded area. Recording from chinchilla auditory nerve fiber, reprinted from [33] with permission from Elsevier.

basilar membrane motion.

Example 2. The second example of non-linearity is that in some specific types of experiments it is possible to hear a sound of certain frequency even though the ear is being stimulated by several other frequencies. This phenomenon is sometimes referred to as *“hearing the missing fundamental”*. For example, three frequencies are played simultaneously $F_1=4000$ Hz, $F_2=6000$ Hz, $F_3=8000$ Hz. Rather than hearing these three frequencies, the listener actually hears a 2000 Hz sound. It has been shown that the region of the cochlea with its threshold tuned to 2000 Hz do not fire any faster, while the neurons “tuned” to 4000 Hz, 6000 Hz and 8000 Hz are firing well above their spontaneous rate. To account for this paradox some scientist have proposed that frequency discrimination is not done based on basilar membrane resonance, but on timing information. Due to phase locking and because all frequencies present are multiples of 2000 Hz, multiple auditory nerve fibers will fire simultaneously at every cycle of the 2000 Hz oscillation. It is hypothesized that if only this timing information is used to interpolate the frequency content of the sound, the brain can be tricked into hearing the missing fundamental.

3.4 Modeling of the auditory nerve

Mathematical modeling of the responses of the auditory nerve fibers has mostly been phenomenological. The models aimed to describe as accurately as possible the experimentally known characteristics of fibers' responses. For example, to match the responses to single simple tones or clicks and to temporal combinations of different stimuli [16, 104, 119], or to determine the lower thresholds of the single-pulse responses [111]. Prevalence of this type of modeling is due to the fact that the biophysics of the hair cells, where auditory nerve fibers originate, has not been worked out in sufficient detail to allow construction of more biophysically realistic models. However, the auditory nerve fiber modeling has recently become one of the examples of how just how much theoretical studies can contribute to the field of neuroscience. Recent models of Heinz et al. [34, 35] provide a link between auditory fiber physiology and human psychophysics. They find how results equivalent to human psychophysical performance can be extracted from accurate models of the fiber responses. In essence, this demonstrates the “code” in auditory nerve fibers via which the auditory system extracts all necessary information.

4 Cochlear nuclei

As fibers of the auditory nerve enter the cochlear nuclei (CN), they branch to form multiple parallel representations of the environment. This allows parallel computations of different sound features. For example, sound localization and identification of a sound are performed in parallel. This separation of processing pathways exists both in the CN and further upstream in the brainstem.

Some information carried by auditory nerve fibers is relayed with great fidelity by specific neurons over specific pathways to higher centers of the brain. Other CN neurons modify the incoming spike trains substantially, with only certain elements of the input signal extracted prior to transmission to the next station in the auditory pathway.

One of the features that is inherited in CN from the auditory nerve is tonotopy (continuous variation of the characteristic frequency with the position of the cell, as on basilar membrane). But, interestingly, tonotopic projections are not point-to-point. Each characteristic frequency point on basilar membrane projects to an iso-frequency *plane* across the extent of the cochlear nucleus. Thus cochlear place representation is expanded into a second dimension in brain. These tonotopic sheets are preserved in projections all the way to cortex. We contrast this with the visual and somatosensory systems, where maps reflect a location of stimulus in space and the representations are point to point.

In this section we will first describe general structure of the CN and outline its

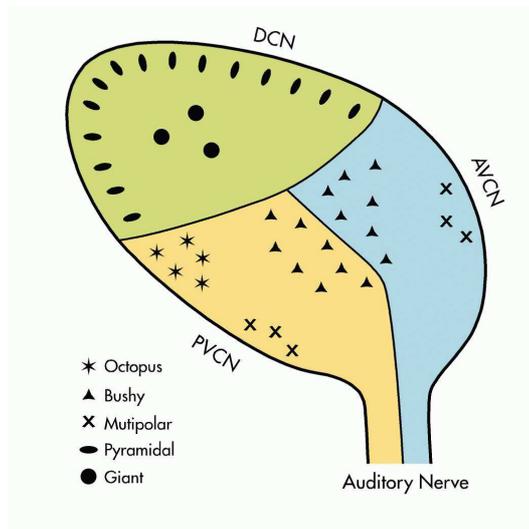


Figure 11: Schematic of cell type location in CN. Main parts of the CN are marked by color (green is dorsal cochlear nucleus (DCN), yellow is posteroventral cochlear nucleus (PVCN) and blue is anteroventral cochlear nucleus (AVCN)). Cell types are marked by symbols. Illustration is made by Brook L. Johnson.

input and output streams. Then we will describe in more details the properties of cells that make information processing in the CN possible.

4.1 Basic features of the CN structure

As we mentioned above, there are parallel streams of information processing in the cochlear nuclei. Thus, the CN is naturally subdivided into areas that house cells with different specialized properties, receive different patterns of inputs and project to different targets. The main subdivision structures are the ventral cochlear nucleus (VCN) and the dorsal cochlear nucleus (DCN) (Fig. 11).

The VCN contains four types of principal cells: globular bushy cells, spherical bushy cells, multipolar cells and octopus cells. We will describe their properties below, and just mention now that they play a very important role in processing different sorts of timing information about the stimuli.

Unlike VCN, DCN is layered and has interneurons. The outermost layer, called the superficial or molecular layer, contains cell bodies and axons of several types of small interneurons. The second layer, called the pyramidal cell layer, has the cell bodies of pyramidal cells, the most numerous of the DCN cell type, and cartwheel and granule cells. The deep layer contains the axons of auditory nerve fibers as well as giant cells and vertical cells.

4.2 Innervation by the auditory nerve fibers

The main external input to the CN comes from the auditory nerve fibers. There is also input to granule cells that brings multimodal information from widespread regions of the brain, including somatosensory, vestibular and motor regions, but we will not consider it here.

As a reminder, in mammals there are two types of the auditory nerve fibers. Type I fibers innervate one inner hair cell each, the fibers are thick and myelinated, and they constitute 90-95% of all fibers. Type II fibers innervate outer hairs cells, they are thinner and unmyelinated.

Type I fibers in the cochlear nuclei form two branches: the ascending branch goes to the anteroventral region of the cochlear nucleus (AVCN) and the descending branch to the posteroventral region (PVCN) and parts of DCN. Type II fibers project to DCN, but because it is not clear if they carry auditory information, type II fibers are, once again, not considered here.

Auditory nerve fibers make synapses to all cell types in the cochlear nuclei, except in the molecular layer of the DCN and in the granule cells region. The innervation is tonotopic within each principal cell type in the VCN and in the deep layer of DCN.

Nerves form different types of terminals onto different cell types and/or different cochlear nucleus divisions. The terminals range from small to large endbulbs. The largest are the endbulbs of Held (*calices of Held*) onto the bushy cells. Each such terminal contains hundreds of synapses. This allows to inject a lot of current into postsynaptic cell every time the pre-synaptic signal arrives. Perhaps only one endbulb is needed to fire a cell, and the transmission through these connections is very fast and reliable. Other CN cell types receive input from auditory nerves through more varicose or bouton-like terminals, located at their dendrites. In these cases more integration is required to fire a cell.

All auditory nerve synapses use glutamate as the neurotransmitter, and are depolarizing (excitatory). Often postsynaptic cells have specialized “fast” receptors (of AMPA type), important in mediating precise coding. Interestingly, in young animals glutamate receptors in the VCN are dominated by very slow NMDA receptors and are replaced by AMPA receptors in the course of development. The reason for this change is not presently clear.

Another important feature that is modified with age is synaptic plasticity. In many neuronal systems the strength of synapses are modified, both on the short (during a single stimulus response) and on the long time scale. In contrast, synaptic transmission by adult auditory nerve fibers shows little plasticity. Synaptic depression (decrease in synaptic efficacy on time scale of 100 msec) is prominent in young animals and decreases with age.

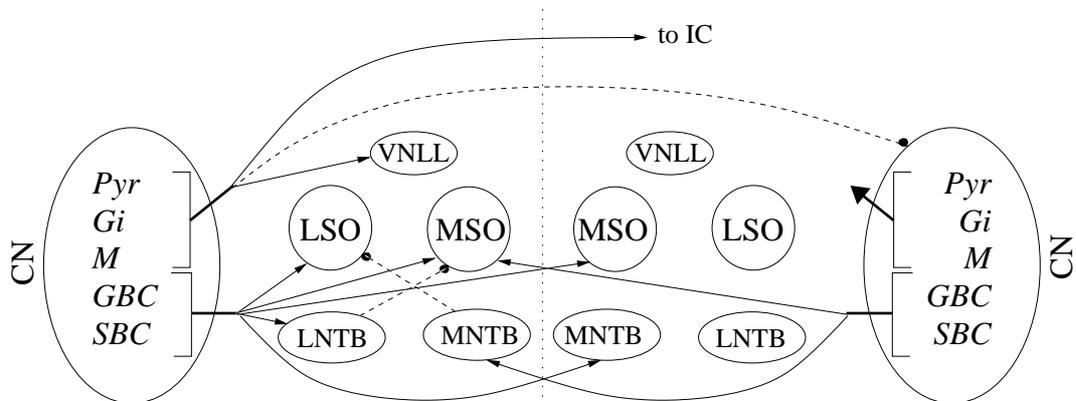


Figure 12: Main ascending connections from cochlear nucleus (CN). Main CN cell types: globular bushy cells (GBC), spherical bushy cells (SBC), multipolar (M), giant (Gi), pyramidal (Pyr) make connections to medial nucleus of the trapezoidal body (MNTB), lateral nucleus of the trapezoidal body (LNTB), medial superior olive (MSO), lateral superior olive (LSO), ventral nucleus of lateral lemniscus (VNLL) and inferior colliculus (IC). All connections are symmetric with respect to midline (dotted line), but not all are shown. Excitatory connections are shown with solid lines and inhibitory with dashed.

4.3 Main CN output targets

Figure 12 shows the two major fiber bundles that leave the cochlear nucleus in relation to other structures in the brainstem. It also shows some of the auditory neuronal circuits of the brainstem. At the output of the cochlear nucleus, spherical and globular bushy cells (SBC and GBC) project to the medial (MSO) and lateral (LSO) superior olivary nuclei and the trapezoid body (MNTB and LNTB). The MSO and LSO compare input from the two ears and perform the initial computations necessary for sound localization in the horizontal plane (see next section). Excitatory inputs to the superior olive come mostly from the spherical bushy cells and inhibitory inputs come from the globular bushy cells via the inhibitory interneurons in the trapezoid body.

Multipolar, giant, and pyramidal cells project directly to the inferior colliculus, where all ascending auditory pathways converge (see section 6). The octopus cells project to the superior paraolivary nucleus and to the ventral nucleus of the lateral lemniscus (VNLL). The granule cells axons project to the molecular layer of the DCN, where they form parallel fibers. These parallel fibers run orthogonal to the auditory nerve fibers and cross isofrequency layers. The input from these two sources is combined by the principal cells of the DCN – the pyramidal cells.

4.4 Classifications of cells in the CN

In the current prevalent view, each cochlear nucleus cell type corresponds to a unique pattern of response to sound; this is consistent with the idea that each type is involved in a different aspect of the analysis of the information in the auditory nerve. The diversity of these patterns can be accounted for by three features that vary among the principal cell types: (1) the pattern of the innervation of the cell by ANFs, (2) the electrical properties of the cells, and (3) the interneuronal circuitry associated with the cell.

To study the correlation between physiological and anatomical properties of the cells experimentally, it is necessary to make intracellular recordings of responses to various current injections and then fill the cells with a dye to image their shape. These experiments are hard to perform and while there seems to be evidence that the correlation between structure and function in these cells is strong, this evidence is not conclusive.

Physiological types

Three major types of VCN cells (bushy, multipolar and octopus) probably correspond to different physiological types: primary-like, chopper and onset.

Primary-like (spherical bushy) and primary-like with notch (globular bushy) responses are very much like auditory nerve. Initial high burst of spikes (>1,000 Hz) followed by decline to maximum rate of no more than 250 Hz. Pause (notch) is mostly due to refractory period of cell.

Choppers are major response type in PVCN, also found in other divisions. In response to high frequency tones, they discharge regularly, independent of stimulus frequency and phase with firing rates sustained at up to 200-500 Hz. These appear to be multipolar cells.

Onset responses have spikes at the onset of stimulus, and then fewer or none. Standard deviation of first spike latency is very small, about 100 μ sec. Those few spikes that are produced during the ongoing stimulus phase lock to low frequency tones. These cells are located mostly in PVCN. Some of them are octopus cells and others are large multipolars. The properties of these responses (precise onset and brief transient activation) allow them to play an important role in temporal coding.

DCN cells exhibit wide variety of response types. Their functions probably are related to pinna movement and echo suppression (which allows you to not hear yourself many times when you speak in a small room). In addition, there is also somatosensory input through granule cells that has to be combined with the auditory one.

In vitro physiology

To complicate things further, slice physiology in VCN provides additional view of types. In these studies intracellular recordings are made from isolated cells. It is possible to either control the current flowing across the cell membrane and measure the induced voltage (current clamp), or to fix the membrane voltage and to record the resulting trans-membrane current (voltage clamp). The recordings reveal two major physiological response types. Type 1 fires regular train of action potentials in response to depolarization. Also, it has relatively linear current-voltage curve. Type 2 fires just one, or a few, action potentials in response to ongoing depolarization. In addition it has very nonlinear current-voltage curve, with zero (reversal potential) at about -70 mV and much steeper slope in depolarizing than in hyperpolarizing direction.

Differences between types 1 and 2 are largely due to change in potassium currents. Type 2 cells have low threshold potassium current that is partially activates near rest values of voltage and strongly activates when voltage rises, repolarizing the membrane and halting the response. Many type 1 cells are multipolar and many type 2 cells are bushy (see more on this below, under Bushy and Multipolar cells).

4.5 Properties of main cell types

Bushy cells

Bushy cells have short ($<200 \mu\text{m}$), bushy dendritic trees. Their synaptic input is located mainly on the soma, with few synapses on their dendrites. Two subtypes of bushy cells are recognized as spherical and globular.

Bushy cells receive inputs from auditory nerve through large synaptic terminals (calyx of Held), have primary-like responses and accurate temporal coding. As mentioned earlier, spherical bushy cells and globular bushy cells differ in their locations and their projections. Spherical cells are located in the more anterior region, and project to MSO, while globular cells project to LSO and trapezoid body. Spherical cells have one or a few short dendrites that terminate in a dense bush-like structure near the soma, and globulars have more ovoid somata and larger, more diffuse dendritic trees. Spherical cells also have lower characteristic frequencies, better phase-locking and are specialized for accurate encoding of AN signal. Globular bushy cells sometimes chop or have onset responses, receive more and smaller endbulbs, and have higher characteristic frequencies. Probably these cells started out as one type and have diverged with evolution of sensitivity to higher frequencies.

Bushy cells respond to sound with well-timed action potentials. The characteristics of both the synaptic current and the postsynaptic cell properties are made so that they shape the timing of the response.

Responses to sounds

Figure 13A shows typical responses to tones of bushy cells. Large unitary synaptic events (excitatory post-synaptic potentials; EPSPs) from between one and three endbulbs cause spherical bushy cells to fire whenever the auditory nerve fibers do (except when the cell is refractory). This one-spike-in, one-spike-out mode of processing means that the responses to sound of spherical bushy cells resemble those of auditory nerve fibers, and for this reason they are called “primary-like”. Evidence that primary-like responses reflect a “one-spike-in, one-spike-out” is provided by their action potential (AP) shapes. The AP is preceded by a pre-potential (reflecting the build up of the endbulb activity) which are almost always followed by the postsynaptic component of the spike, demonstrating security of the synapse.

Globular bushy cells give similar response, primary-like-with-notch. They differ from primary in that in the beginning of the response there is a precisely timed peak followed by a notch. Precise timing is helped by the convergence of large number of fibers. If a globular cell needs one input to fire, then there is a very high probability that one of the incoming fibers will activate early on and cause a spike. This initial peak will be followed by recovery period (refractoriness), generating notch.

Electrical characteristics

As a response to injection of constant depolarizing current bushy cells produce one to three spikes at the onset and then settle to a slightly depolarized constant voltage value. This is due to the low-threshold potassium current. It is partially activated at rest, then during spiking it strongly activates increasing membrane conductance and thereby shortening the membrane time constant and repolarizing the cell. To induce firing, the input must be very strong and fast. The short time constant blocks temporal integration of inputs. Rapid temporal processing permits bushy cells to preserve information about the stimulus waveform information that is necessary for sound localization.

Multipolar cells

Multipolar cells have multiple, long dendrites that extend away from the soma in several directions. These cells are also sometimes referred to as “stellate”. Two major classes of multipolar cells have been described. The cells of one group, T-multipolars (planar), have a stellate morphology with dendrites aligned with auditory nerve fibers, suggesting that these cells receive input from a restricted range of best frequencies. Their axons project through the trapezoid body (hence the “T”) to the contralateral IC. Cells of the second group, D-multipolars (radiate), have dendritic fields that are not aligned with auditory nerve fibers. Their axons project to the contralateral cochlear nucleus.

Both T- and D-multipolar cells in the VCN serve the role of interneurons through their axon collaterals. These terminate locally within VCN as well as projecting to

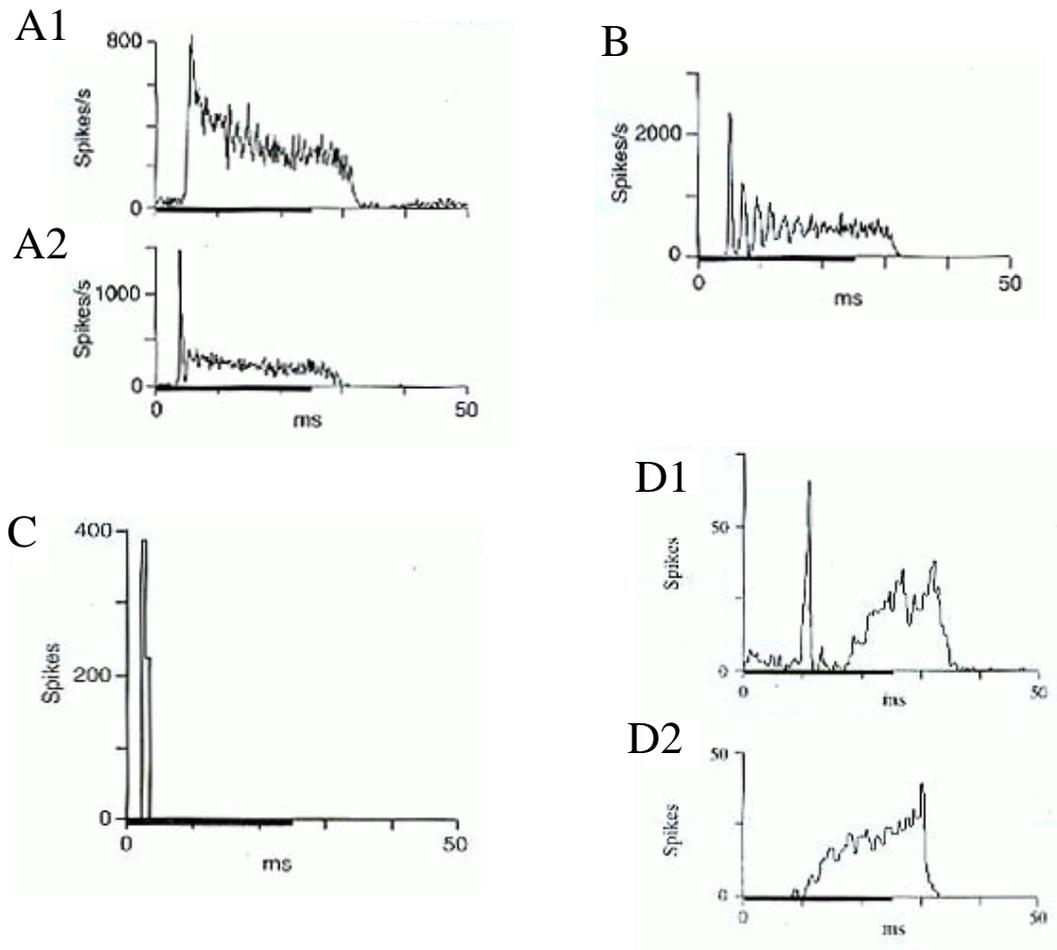


Figure 13: Different response types. *A1*: Primary-like response (bushy cell). *A2*: Primary-like with notch response (bushy cell). *B*: Chopper response from T-multipolar cell. *C*: Single-spike onset response. *D1*: Pauser response. *D2*: builder response. In all panels dark black line under the axes is the tone presentation. Panels A and B are used by permission of The American Physiological Society from [8]. Panel C is reprinted from [42] with permission from Elsevier. Panel D is reprinted by permission of VCH Publishers, Inc. from [29].

the deep DCN. T-Multipolars are excitatory and D-multipolars are inhibitory and glycin-ergic.

Responses to sounds

Figure 13B shows responses characteristic of T-multipolar cells. T-multipolar cells respond to tones by firing at regular intervals independent of frequency of the

tone, a pattern called chopping. The reproducibility of firing gives histograms of responses to sound a series of characteristic modes that is independent of the fine structure of the sound. The intervals between nodes are of equal duration and correspond to the intervals between spikes, with one spike per node. The peaks are large at the onset of the response because the latency to the first spike is quite reproducible in chopper neurons; the peaks fade away over the first 20 msec of the response, as small variations in interspike interval accumulate and spike times in successive stimulus repetitions diverge. The chopping firing pattern must arise from the intrinsic properties of the cells themselves, since it does not reflect properties of the inputs. Moreover, the same pattern is elicited by depolarization with steady currents (Fig. 14).

D-multipolar cells are broadly tuned and respond best to stimuli like noise but only weakly to tones. This is due to the fact that D-multipolars receive input from many auditory nerve fibers on their somata and on dendrites that spread across the frequencies. D-multipolars also respond with a precisely timed onset spike to tones, but (unlike octopus cells) they give some steady discharge after the onset spike. Also, unlike octopus cells, the firing pattern in response to tones is shaped by inhibition along with the intrinsic electrical properties.

Electrical characteristics

Unlike bushy cells, multipolars are capable of temporal summation of successive EPSPs. Recall from above, that bushy and multipolar cells have AMPA receptors with similar rapid kinetics. The differences between unitary voltage responses arise because the decay of the unitary event, and therefore the degree of temporal integration, is determined by the membrane time constant of the cell. For bushy cells this is short (2-4 msec), whereas for multipolar cells it is longer (5-10 msec).

Octopus cells

Octopus cells in VCN (Fig. 11) are contacted by short collaterals of large numbers of auditory nerve fibers through small terminal boutons. Octopus cell dendrites are oriented in one direction, inspiring their name. The orientation is perpendicular to the auditory nerve fibers so that the cell bodies encounter fibers with the lowest best frequencies, and the long dendrites extend toward fibers that encode higher frequencies. Each cell receives input from roughly one-third of the tonotopic range, i.e. the frequency tuning is broad and contribution of each fiber to the octopus cell response is very small. Also EPSPs that an octopus cell receives are very brief, between 1 and 2 msec in duration.

Responses to sounds

Octopus cells behave like bushy cells in that the low input resistance prevents temporal summation of inputs. But instead of receiving a few large inputs, octopus cells receive small inputs from many fibers. Thus, many fibers have to fire simultaneously to drive an octopus cell.

Such synchronous firing is not achieved with ongoing pure tone stimuli. Rather, it happens at stimulus transients. For example, at the onset of a pure tone (thus the term onset response, Fig. 13C), or at a rapid fluctuation of a broadband stimulus such as at the onset of a syllable or during a train of clicks. In fact an octopus cell can faithfully follow a train of clicks with rates up to 500 Hz.

Electrical characteristics

In the case of octopus cells, the fast membrane time constant or high membrane conductance is generated by presence of two opposing voltage-sensitive currents. One of them activates with hyperpolarization and has a reversal potential near -40 mV, the other is a potassium current (reversal potential near -80 mV) that activates with depolarization. At rest both currents are partially activated. In fact, they are quite large, but they compensate one another. In addition, the voltage-gated conductances of these currents are very sensitive to voltage near resting potential. This means that any voltage fluctuation activates these currents which counter the voltage change. As a result, EPSPs are always small and brief and the cell is only sensitive to synchronous inputs. Moreover, any depolarizing current that rises too slowly is bound to activate the potassium conductance and to preclude the cell from responding. This mechanism makes octopus cells sensitive to the rate of rise of their input, not its amplitude. In a sense they are sensors of the derivative of the input.

Pyramidal cells

Pyramidal (also called “fusiform”) neurons in the DCN are bipolar, with a spiny dendritic tree in the molecular layer and a smooth dendritic tree in the deep layer. The cell bodies of pyramidal cells form a band in the pyramidal cell layer. The smooth dendrites are flattened in the plane of the isofrequency sheets, where they receive input from the auditory nerve. The spiny dendrites span the molecular layer and are contacted by parallel fibers at the spines.

Responses to sounds

Pyramidal neurons in the DCN show pauser and buildup responses to sound. The examples shown in Fig. 13D are typical of anesthetized animals, where the inhibitory circuits of the DCN are weakened. The response shows a poorly timed, long latency onset spike followed by a prominent pause or a slow buildup in response with a long latency.

Electrical properties

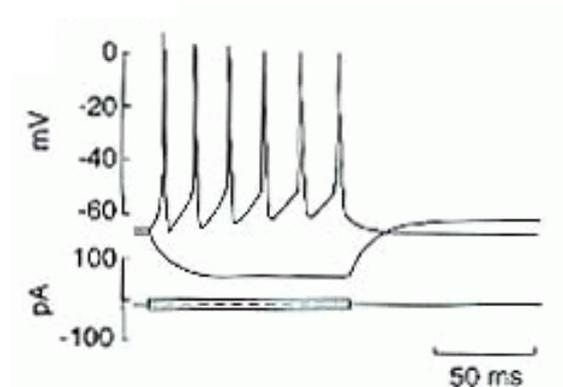


Figure 14: Responses of T-multipolar cell to intracellular injections of depolarizing and hyperpolarizing currents. Membrane potential (top) and current time courses (bottom). Spiking is in response to the depolarizing current. Reproduced with permission from [59], Copyright 1991 by the Society for Neuroscience.

The pauser and buildup characteristics seem to derive from a transient potassium conductance. This conductance is inactivated at rest. If a sufficient depolarizing stimulus arrives, the cell simply responds with a spike. However, if the cell is initially hyperpolarized, this removes the potassium current inactivation. Any subsequent depolarization will activate the potassium current transiently, producing the long latency of a pauser or buildup response. *In vivo* a sufficient hyperpolarization is provided by inhibitory synaptic inputs as an aftereffect of a strong response to an acoustic stimulus.

Some of the other cell types

Giant cells are large multipolar cells located in the deep layers of the DCN. They have large, sparsely branching, dendritic trees that cross isofrequency sheets. Giant-cell axons project to the contralateral inferior colliculus.

Granule cells are microneurons whose axons, the parallel fibers, provide a major excitatory input to DCN through the molecular layer. Granule-cell axons terminate on spines of the dendrites of pyramidal cells, on spines of cartwheel cells, and on the stellate cells.

The vertical cells are inhibitory interneurons that project to their isofrequency sheets in both DCN and VCN; they inhibit all of the principal cells in the cochlear nucleus, except the octopus cells. Vertical cells are narrowly tuned, and they respond most strongly to tones at a frequency near their characteristic frequency.

Cartwheel cells are inhibitory interneurons whose numerous cell bodies lie in the pyramidal cell layer of the DCN. Their dendrites span the molecular layer and are

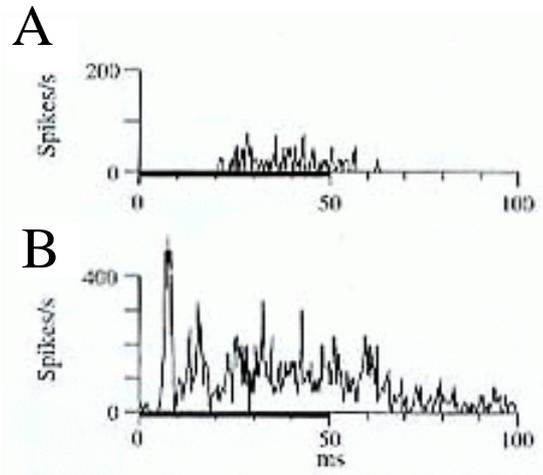


Figure 15: Weak (A) and strong (B) responses of cartwheel cells. Used by permission of The American Physiological Society from [71].

densely covered with spines that are contacted by parallel fibers. They contact pyramidal, giant, and other cartwheel cells through glycinergic synapses. The reversal potential of the cartwheel-to-cartwheel cell synapses lies a few millivolts above the resting potential and below the threshold for firing so that its effect is depolarizing for the cell at rest but hyperpolarizing when the cell has been depolarized by other inputs. Thus, the effects of the cartwheel cell network are likely to be context-dependent: excitatory when the cells are at rest but stabilizing when the cells are excited.

Figure 15 shows responses to sound of cartwheel cells. Cartwheel cells are the only cells in the cochlear nucleus with complex action potentials, which reflect a combined calcium and sodium spike. Many of the cartwheel cells respond weakly to sounds and no particular pattern of response is consistently observed. They probably mainly transmit non-auditory information to the principal cells of the DCN.

As a side note: cartwheel cells share many features with cerebellar Purkinje cells. For example, both of these fire complex action potentials; and genetic mutations affect Purkinje cells and cartwheel cells similarly.

4.6 Modeling of the cochlear nuclei

As data on the membrane properties of cochlear nucleus neurons have accumulated, it has become possible to use biophysical models of Hodgkin-Huxley type to explore the different behaviors described in the previous section. These models give researchers ability to test whether the biophysical mechanisms discussed above can

indeed underlie the observed cellular responses. In particular, when many ionic currents have been identified in a given cell type, a model helps identify which of them are the key ingredients in shaping each feature of the cell's behavior.

For example, a model of bushy cells [85] demonstrates that the presence of a low-threshold potassium current in the soma of the cell can account for responses of both spherical and globular bushy cells depending on number and strength of inputs. When inputs are few and large the response matches characteristics of the primary-like response, but with larger number of inputs it becomes primary-like with notch. Another computational study of globular bushy cells [47] uses a much simpler underlying mathematical model (integrate-and-fire rather than Hodgkin-Huxley type). This allows to study in generality what increases or decreases synchrony in the output vs. the input, using globular bushy cell as an example.

Some of the other models have dealt in detail with responses of various cells in the dorsal cochlear nucleus [9, 80].

Another example of mathematical modeling in cochlear nucleus is a model of multipolar cell [6]. This is a more complex multi-compartmental model, which demonstrates which combination of known somatic and dendritic currents can accurately reproduce properties of a chopper response (regular firing with irregular or constant inputs).

5 Superior olive. Sound localization, Jeffress model

The superior olive is a cellular complex in the brainstem of about 4 mm long. The cytoarchitecture of this complex defines three parts: the medial superior olive, the lateral superior olive, and the nucleus of the trapezoid body. It should be emphasized that the trapezoid body itself is not a nucleus; it is a bundle of fibers. The entire superior olive complex is surrounded by small cellular groups known as the preolivary or periolivary nuclei. The olivary nuclear complex is the first level in the auditory system where binaural integration of auditory signals occur. It is the key station to performing many of the binaural computations, including, in large part, localization of sounds.

Two nuclei within superior olivary complex that have been most extensively studied are the medial and lateral superior olivary nuclei (MSO and LSO). Other nuclei, such as medial and lateral nuclei of the trapezoid body (MNTB and LNTB) are named according to their position with respect to the MSO and LSO. There is also a group of cells called olivacochlear neurons. They project to the cochlea. These neurons can be activated by sound, and cause suppression of spontaneous and tone evoked activity in auditory nerves.

5.1 Medial nucleus of the trapezoid body (MNTB)

The cell body of neurons in the MNTB receives input from axonal projection of cells in the contralateral VCN, primarily from the globular bushy cells (GBC). The GBC-to-MNTB synapses are large, endbulb type, i.e. they also provide reliable synaptic connection. For that reason the responses of many MNTB cells are similar to their primary excitatory input, the globular bushy cells. Some cells with chopper type responses are also found. Because convergence of the bilateral auditory information has not occurred yet at the level of MNTB, the neurons there are responsive exclusively to sounds presented to the ear contralateral to the nucleus itself. As usual, each cell responds best to a characteristic frequency. They then send a short axon to the LSO where it forms inhibitory glycine-ergic synapses.

5.2 Lateral superior olivary nucleus (LSO)

Principal cells receive inputs onto soma and proximal dendrites. Excitatory input is received directly from the ipsilateral VCN. Inhibitory input is from the contralateral VCN, transmitted through the MNTB. Even though there is an extra synapse in the path of the inhibitory signal, inputs from each ear arrive at the LSO simultaneously. This is possible due to the large reliable synapses associated with MNTB. The cells that receive excitatory input from ipsilateral ear and inhibitory input from the other ear are often called IE cells.

Because of the organization of their inputs, LSO cells are excited by the sound which is louder in the ipsilateral ear and softer in the contralateral ear. There is almost no response when a sound is louder in the contralateral ear than in the ipsilateral ear. As a result the sensitivity functions of LSO neurons to the interaural level difference (ILD) are sigmoidal. There are successful computational models by Michael Reed and his colleagues (e.g. [81]) that demonstrate how a spectrum of sound ILDs can be coded by a population of the sigmoidally-tuned LSO cells.

The characteristic frequencies of LSO cells are predominantly in the high frequency range complimenting the range of frequencies over which the MSO is responsive. The frequency tuning is sharp, as narrow as in the cochlear nucleus and the auditory nerve, because ipsilateral and contralateral inputs are well matched in frequency. However, ipsilateral (excitatory) inputs are slightly broader tuned. Therefore, to maintain a response level of an LSO neuron, as contralateral sound moves away from characteristic frequency it must become louder to continue counteracting the ipsilateral signal.

5.3 Medial superior olivary nucleus (MSO)

The cell bodies of neurons in the MSO receive input from two sets of dendrites. One projects laterally from the cell and gets its input from the ipsilateral VCN. The other projects medially from the cell and receives its input from the contralateral VCN. The cells in the MSO, as in LSO, are classified according to their response to these two sets of dendritic inputs. If a cell is excited by both contralateral and ipsilateral input, it is classified as an EE cell. If a cell is excited contralaterally, but inhibited by ipsilateral input, it is classified as an EI cell. If a cell is excited by ipsilateral, but inhibited by contralateral input, it is classified as an IE cell. If a cell is inhibited by both contralateral and ipsilateral input, that is it fires below its spontaneous rate when either ear is presented with an appropriate stimulus, it is classified as an II cell.

MSO neurons respond to both binaural and monaural stimuli. The monaural response however will always be submaximal when compared to similar binaural stimuli. The response tends to be frequency selective and the best frequency is typically the same for both ipsilateral and contralateral input in a given cell. The characteristic frequencies in the MSO tend to be in the low end of the hearing spectrum. In addition to a characteristic frequency, many neurons in the MSO respond best to a specific delay between ipsilateral and contralateral stimulus termed the “characteristic delay”. They also have a “least characteristic delay” to which they respond worse than to monaural input. If the characteristic delay is preserved across frequencies or modulation frequencies, the cell is termed to be of “peak-type”. If the least characteristic delay is preserved across frequencies the cell is termed “trough-type”. Most of the cells in the MSO are of peak-type, and they form the basis for a famous Jeffress theory, which is described below.

5.4 Sound localization. Coincidence detector model

Many animals, especially those that function in darkness, rely on sound localization for finding their prey, avoiding predators, and locating members of their own species. Therefore, it is not surprising that many mechanisms for sound localization, well-suited for different circumstances, have been created in the course of evolution.

Primary cues for sound localization are binaural. They include interaural differences in sound pressure level, time of arrival and frequency spectrum of the sound. It is also possible to localize sounds monaurally, but it is much less efficient (see, e.g., [106]). Historically, the most studied binaural cues have been interaural time and sound pressure level differences. Interaural time difference (ITD) arises because the sound, originating on one side of the head, must follow a longer path and consequently takes a longer time traveling to the distal than proximal ear (Fig. 16).

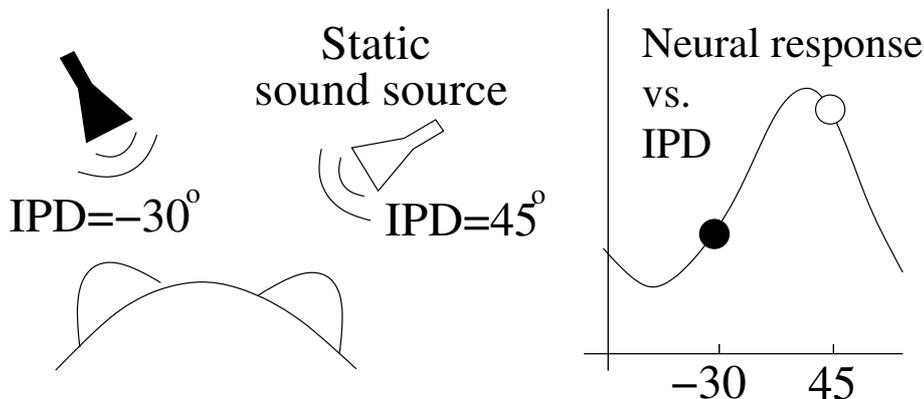


Figure 16: When a sound source is placed not directly in front of the animal, but off-center, it generates ITD (and IPD). Two schematic sound source locations are shown in the left panel. If the response of a neuron is recorded at the same time, then it can be plotted vs. corresponding IPD (schematically shown in the right panel). This plot of response vs. IPD is the static tuning curve of the neuron.

The interaural level difference (ILD) results from the interference of the head and ears of the listener with the propagation of the sound so that a “sound shadow” may develop, i.e., the sound may be attenuated on the far side relative to the near one. In natural settings an acoustic signal usually gives rise to both time and amplitude differences. Under certain conditions only one of the cues retains sufficient magnitude to serve the listener. In particular, according to the duplex theory [79], interaural level differences are the primary cue for localizing high-frequency tones, whereas interaural time differences are responsible for the localization of the low frequency tones. The basis for this distinction is that for high-frequency sounds the wavelength is small compared with the size of the head and the head becomes an obstacle. At lower frequency sounds the wave length is longer and the head is acoustically transparent.

In recent years many departures from the duplex theory have been found. For example, for complex high-frequency signals ITD between low-frequency envelopes can serve as a localization cue [115]; direct physical measurement of ILD at the ears of the listener (e.g., in cats [108] and humans [21] show that ILD does not vary in a systematic fashion with azimuth; further, the interaural differences can be altered by factors such as head and pinnae shape; finally, interaural spectral differences, that develop for complex acoustic signals (particularly with high frequency components), provide considerable directional information that can enhance the localization task.

Nevertheless, time and level differences continue to be considered as primary

sound localization cues. Many species are very acutely tuned to them. Humans, for example, can readily sense interaural disparities of only a few decibels or only tens of microseconds [79, 102]. The sensitivity is even higher, for example, in some avians [46]. Besides psychophysical evidence, there is a great amount of physiological data, showing that in various structures along the auditory pathway (medial and lateral superior olive (MSO and LSO), dorsal nucleus of lateral lemniscus (DNLL), inferior colliculus (IC), auditory cortex) low-frequency neurons are sensitive to ITDs and high-frequency neurons to ILDs. What frequency is “low” or “high” varies with species. Humans are sensitive to ITDs for tone frequencies up to about 1500 Hz. Neural sensitivity to ITD in cats, for example, goes up to 3kHz [50], and in owls, up to 6-8kHz [65].

Interaural phase difference (IPD)

Rose et al. [83] found that for some cells in the inferior colliculus of the cat the discharge rate was a periodic function of interaural time delay, with a period equal to the stimulus wavelength. They argued that the periodic nature of the response showed that the cells were sensitive to the interaural phase difference (IPD) rather than ITD. Similar cells were found in other parts of the mammalian auditory system (e.g., [30]). Of course, the issue of distinction between IPD and ITD arises only when the sound carrier frequency is changed, because, for example, if the cell is sensitive to ITD, then its preferred IPD (the one that elicits the largest response) would change linearly with carrier frequency; if, on the other hand, the cell is sensitive to IPD, its preferred IPD would stay the same. For pure tones of a fixed frequency, talking about IPD and ITD is redundant and we will sometimes use them interchangeably, assuming that the stimulus is a pure tone at the characteristic frequency of the neuron.

Schematic of the experiments

In experimental studies concerning the role of binaural cues and corresponding neuronal sensitivity, the sound is usually played from speakers located at one of the positions around the listener (free-field), or via headphones (dichotic). Headphone stimulation is, of course, an idealized sound representation (for example, the stimuli in the headphones are perceived as located within the head, [118]). On the other hand, it may be hard to control the parameters of the stimuli in the free-field studies. For example, the simplest geometrical model of IPD induction, due to Woodworth [110], that assumes constant speed and circular head shape, does not match the direct measurements of IPD [48] (Gourevitch [31] notes that for a 400 Hz signal at 15° off midline, the observed IPD is about three times greater than IPD based on the Woodworth’s model). A more detailed model by Kuhn [49] takes into account physics of sound wave encountering the head as an obstacle. It shows

the dependence of IPD on the head size and the sound carrier frequency for humans and provides better fit for the data, but also demonstrates that extracting the information available to the auditory system in a free-field study is not a trivial task.

Neuronal mechanisms of IPD-sensitivity: Superior olive

Many contemporary models of binaural interaction are based on the idea due to Jeffress [38]. He suggested the following model: an array of neurons receives signals from the two ears through delay lines (axons that run along the array and provide conduction delay). The key assumption is that each neuron fires maximally when the signals from both ears arrive to that neuron simultaneously, in other words — when the conductance delay exactly compensates the IPD. These neurons are referred to as coincidence detectors. Mathematically, we can also think about their action as computing the temporal correlation of signals. The Jeffress-like arrangement was anatomically found in the avian homolog of medial superior olive (the first site of binaural interaction), the nucleus laminaris [117]. In mammals, recordings from single cells in MSO confirm that many cells do have the properties of coincidence detectors [30, 97, 112]. By recording the responses of these neurons to binaurally presented tones with different IPDs (simulating sounds at different locations around the head of the animal), one produces the (static) tuning curve (schematic in Fig. 16). It is usually a relatively smooth curve that has a clearly distinguishable peak. The peak of the binaural response appears when the ITD is approximately equal (with a different sign) to the difference between times of maximal responses to monaural signals [30], as expected for coincidence detectors.

In natural environments the relative positions of a sound source and the head are free to change and, therefore, the binaural sound localization cues are likely to be changing (dynamic) too. In experiments (e.g., [113]), the binaural cue produced by a moving sound source may be simulated by temporal variation of IPD. (In humans dynamic IPD modulation creates the percept of sound motion.) If the MSO cells act like coincidence detectors, they are expected to signal the instantaneous value of the interaural delay and show no evidence for motion sensitivity. This is confirmed experimentally (e.g., [98]).

Coincidence detector model

Cell body

We describe here a coincidence detector point neuron model of the Morris-Lecar type (after Agmon-Snir et al. [2]). The main dynamic variables are the voltage V and a gating variable w .

$$C\dot{V} = -g_{Na}m_{\infty}(V)(V - V_{Na}) - g_K \cdot w \cdot (V - V_K) - g_L(V - V_L) - G_{syn}(V - V_{syn})$$

$$\tau_w(V)\dot{w} = w_\infty - w$$

Parameters are chosen in a realistic range to tune the model neuron into the phasic firing regime. Namely, this cell will fire only one spike in response to a depolarizing current injection (Fig. 17), as observed experimentally (e.g. [82]). Most of the parameters are the same as in Agmon-Snir et al. [2], unless noted otherwise.

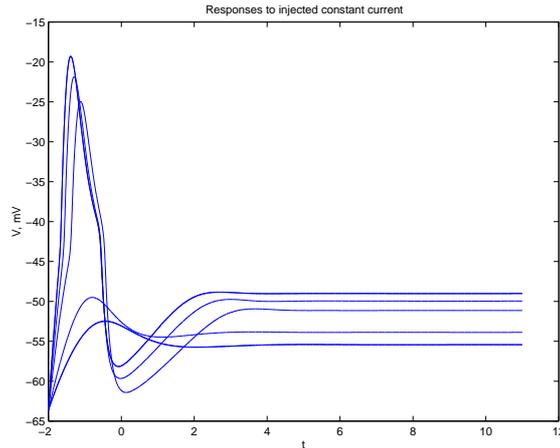


Figure 17: Responses to increasing level of depolarizing current. The cell fires only once in response to a constant level injection of depolarizing superthreshold current.

Input

Input is modeled as 2 delay lines (one from each ear). Each line makes 12 synapses on the cell (Fig. 18A). On each stimulus cycle every synapse can produce a synaptic event with probability 0.7 (at 500 Hz stimulus frequency this corresponds to average firing rate of the afferents equal to 350 Hz) (Fig. 18B). In case a synaptic event is produced, it results in the alpha-function change of postsynaptic conductance, namely

$$G_{syn,unit} = A \frac{t - \phi}{t_p} \exp\left(1 - \frac{t - \phi}{t_p}\right).$$

Here, $t_p = .1$ msec, amplitude A is chosen to produce reasonable firing rates at the given number of inputs and ϕ is chosen from a Gaussian distribution $N(\phi_0, \sigma^2)$ for one side and $N(IPD, \sigma^2)$ for the other side (IPD – interaural phase disparity, $\sigma = 0.4$ – measure of phase-locking, ϕ_0 – preferred phase) (see Fig. 19A).

Figures 19 and 20 show sample voltage and synaptic conductance time courses and a tuning curve of the model neuron. The tuning curve was obtained by computing number of spike elicited during a 1 sec stimulus presentation. This tuning curve shows that the neuron is a pretty good coincidence detector.

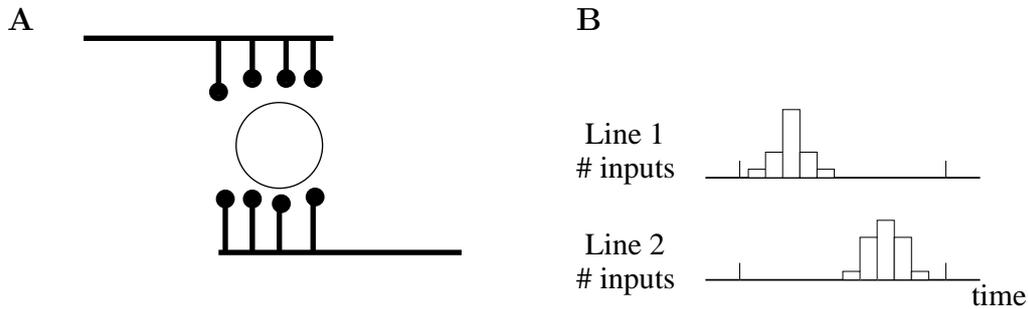


Figure 18: Schematic of the input to the model cell. *A*: Delay lines. *B*: Example of distribution of input events times within stimulus cycle. Total number of events is also random, see text

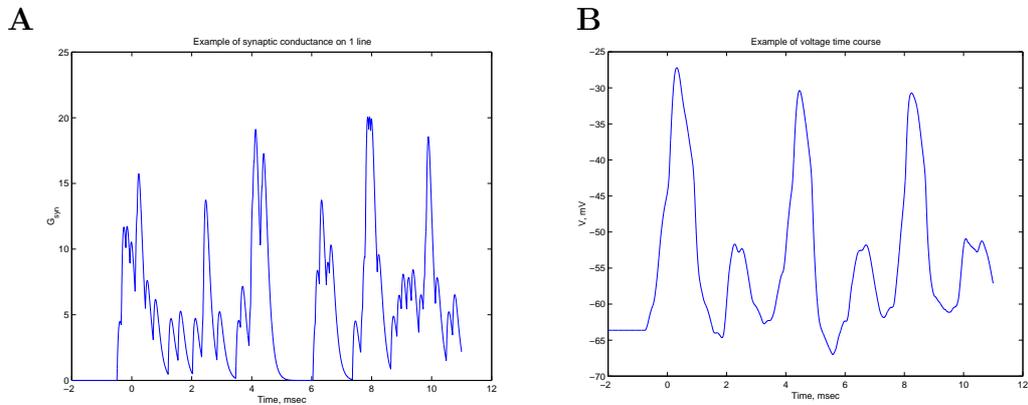


Figure 19: Sample conductance and voltage traces. *A*: Synaptic conductance of one line of inputs over 6 stimulus cycles. *B*: Corresponding voltage trace.

However, the known coincident-detector avian auditory neurons have a very specific structure. They have bipolar dendrites whose length varies continuously with characteristic frequency of the neuron (neurons with higher characteristic frequencies have shorter dendrites). The function of the dendrites was hard to unravel experimentally, until a mathematical model [2] suggested that propagation of signal along the dendrites resulted in nonlinear summation of conductances at the soma, improving coincidence detection. The argument goes as follows. Let us measure the quality of coincidence detection as the difference between maximum and minimum of the tuning curve. Let us also assume that the unitary synaptic conductance is adjusted in such a way that when signals from both sides arrive simultaneously the probability of firing is the same for either a point neuron or a bipolar neuron (this

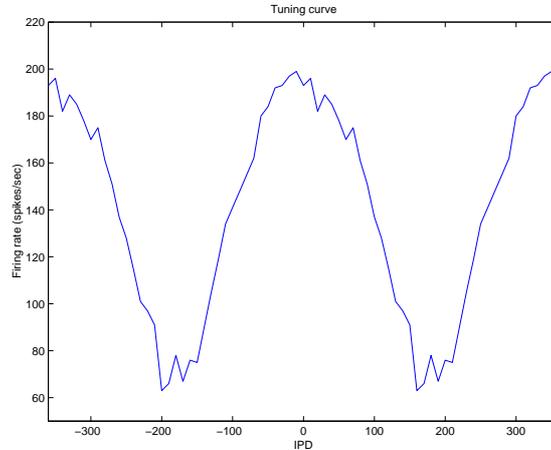


Figure 20: Tuning curve of the model neuron. It is computed with small amount of averaging.

ensures that the maximum of the tuning curve is the same), i.e. let us say that the cell needs n simultaneous input spikes to fire ($n/2$ from each side). First, we need to show that the quality of coincidence detection is better for a neuron with bipolar dendrite than for a point neuron (i.e. minimum of the tuning curve is lower with the dendrite included). The key observation is that the synaptic current produced from a number of input spikes arriving at the same location is smaller than the current produced by the same number of inputs arriving at several different locations, because the driving force for excitatory current decreases with increase in local voltage. When the signals from different sides arrive half-cycle apart (worst ITD) and by chance one side produces n input spikes, it is still enough to fire a point neuron (which does not care where the inputs came from), but in the bipolar dendrite neuron all n spikes will now be concentrated on one side, producing smaller (insufficient) synaptic current, i.e. probability of firing will be smaller with dendrite included. Second, we need to show that longer dendrites improve coincidence detection at lower rather than higher input frequencies. This happens because of the jitter present in the input trains. Higher frequencies mean larger overlap between inputs from both sides even when the phase shift is maximal. These occasional “false” coincidences, will only add a small fraction to the conductance. For the point neuron it means only a small change in synaptic current. However, in case with dendrite the small conductance arrives at the different dendrite, producing relatively large current, i.e. higher probability of the “false” response. As a result, at a given level of jitter, the minimum of the tuning curve increases with frequency, i.e. the dendrites become less advantageous at higher frequencies.

In addition, more recent computational models of MSO argue that Jeffress model

may need to be seriously revised. In fact, it has been shown experimentally that MSO cells that are ITD-sensitive may be receiving substantial location-tuned inhibitory input [13]. A computational model [13] confirms that if that inhibition arrives at the MSO cell with a fixed time delay relative to excitation from the same side of the brain, it may allow to explain a larger range of MSO response properties than traditional excitation-only coincidence detector Jeffress-type model.

6 Midbrain

The inferior colliculus (IC) is the midbrain target of all ascending auditory information. It has two major divisions, the central nucleus and dorsal cortex, and both are tonotopically organized. The inputs from brainstem auditory nuclei create multiple tonotopic maps to form what are believed to be locally segregated functional zones for processing of different aspects of the auditory stimuli. The central nucleus receives both direct monaural input from cochlear nuclei and indirect binaural input from the superior olive.

6.1 Cellular organization and physiology of mammalian IC

The IC appears to be an auditory integrative station as well as a switchboard. It is responsive to interaural delay and amplitude differences and may provide a spatial map of the auditory environment, although this has not been directly shown, except in some birds, e.g., barn owls. Changes in the spectrum of sounds such as amplitude and frequency modulation appear to have a separate representation in the IC. This sensitivity to spectral changes possibly provides the building blocks for neurons responsive to specific phonemes and intonations necessary to recognize speech. Finally, the IC is involved in integration and routing of multi-modal sensory perception. In particular, it sends projections which are involved in ocular reflexes and coordination between auditory and visual systems. It also modifies activity in regions of the brain responsible for attention and learning.

One of the very interesting properties of the inferior colliculus responses is its sensitivity to interaural time (or phase) differences. As we explained above, ITD sensitivity first arises in the MSO, due to precise preservation of timing information (phase locking) and coincidence detection. Notably, preservation of timing information starts to deteriorate in the IC. In particular even its low frequency cells do not phase lock as well as cells at the previous levels. Yet, in spite of apparent lack of timing information, new properties of ITD sensitivity, such as non-linear responses to dynamic changes in ITD arise in the inferior colliculus. This has recently attracted attention of modelers to the IC [14, 15, 88, 11, 12].

6.2 Modeling of the IPD sensitivity in the inferior colliculus

The inferior colliculus (IC), in the midbrain, is a crucial structure in the central auditory pathway (Fig. 1). Ascending projections from many auditory brain stem nuclei converge there. Nevertheless, its role in the processing of auditory information is not well understood. Recent studies have revealed that many neurons in the IC respond differently to auditory stimuli presented under static or dynamic conditions and the response also depends on the history of stimulation. These include responses to dynamically varying binaural cues: interaural phase [62, 95, 96] and interaural level differences [87], and monaural stimuli such as slowly modulated frequency [57]. At least some of the sensitivity to dynamic temporal features is shown to originate in IC or in the projections to IC, because, as we mentioned above, responses of cells in lower level structures (MSO and dorsal nuclei of lateral lemniscus (DNLL) for the low-frequency pathway, Fig. 1) follow the instantaneous value of the stimulus [98]. A history-dependence of the responses, similar to that in IC, but of even greater magnitude, has been observed in the auditory cortex [58]. These findings suggest that there might be a hierarchy of representations of auditory signals at subsequent levels of the auditory system.

Sensitivity to simulated motion in IC has been demonstrated with binaural beats (periodic change of IPD through 360°), [95, 114] and partial range sweeps (periodic modulation of interaural phase in a limited range) [96, 62]. For example, it was shown that the binaural beat response is sometimes shifted with respect to the static tuning curve (Fig. 22, 23), responses to IPD sweeps form hysteresis loops (Fig. 24), and the response to the dynamic stimulus can be strong outside of the statically determined excitatory receptive field (example in Fig. 24A, arrow). Thus, there is a transformation of response properties between MSO and IC, but there is no agreement on how this transformation is achieved. Both intrinsic and synaptic mechanisms have been implicated.

One hypothesis is that the observed history-dependence of responses is reflective of “adaptation of excitation” (firing rate adaptation) [62, 98]. The firing rate adaptation is a slow negative feedback mechanism that can have various underlying cellular or network mechanisms. It can manifest itself in extracellular recordings of responses to a constant stimulus as the decrease of probability of firing with time, and in *in vitro* recordings as the increase of the interspike interval in response to the injection of a constant depolarizing current. Both extracellular [62, 98], and intracellular [72, 93] recordings have registered firing rate adaptation in IC. It has also been shown by Cai et al., in a modeling study [15], that addition of a negative feedback, in the form of a calcium-activated hyperpolarizing membrane conductance (one of the adaptation mechanisms), to a motion-insensitive model of an IC neuron, induced sensitivity to dynamic stimuli.

Another hypothesis [98] is that the motion is the result of interaction between excitatory and inhibitory inputs, aided by adaptation and post-inhibitory rebound. Post-inhibitory rebound (PIR) is observed when the neuron fires upon its release from hyperpolarization; this has been demonstrated in IC slice recordings [72, 93]. McAlpine and Palmer [64] argued that leaving the key role in generation of motion sensitivity to inhibition contradicts their data. They showed that sensitivity to the apparent motion cues is decreased in the presence of the inhibitory transmitter, GABA, and increased in the presence of inhibitory transmission blocker, bicuculline.

Unification of the hypotheses. To examine the role of various mechanisms in shaping IC response properties, mathematical models of an IC cell were developed in which the cell receives IPD-tuned excitatory and inhibitory inputs and possesses the properties of adaptation and post-inhibitory rebound.

In earlier modeling studies of IC, Cai et al. [14, 15] developed detailed cell-based spike-generating models, involving multiple stages of auditory processing. One of their goals was to describe some of the data that require a hierarchy of binaural neurons. Recent models [11, 12] take a very different approach. These models are minimal in that they only involve the components whose role we want to test and do not explicitly include spikes (firing-rate-type models). Exclusion of spikes from the model is based on the following consideration. First, IC cells, unlike neurons in MSO, do not phase lock to the carrier frequency of the stimulus. In other words, spike timing in IC is less precise. Second, it is assumed that the convergent afferents from MSO represent a distribution of preferred phases and phase-locking properties. As a result, either through filtering by the IC cell or dispersion among the inputs, spike-timing information is not of critical importance to these IC cells. This minimalistic approach greatly facilitates the examination of computations that can be performed with rate-coded inputs (spike-timing precision on the order of 10 ms), consistent with the type of input information available to neurons at higher levels of processing. In addition, these models are computationally efficient, can be easily implemented, and require minimal assumptions about the underlying neural system.

It is shown in the modeling studies [11, 12] that intrinsic cellular mechanisms, such as adaptation and post-inhibitory rebound, together with the interaction of excitatory and inhibitory inputs, contribute to the shaping of dynamic responses. It is also shown that each of these mechanisms has a specific influence on response features. Therefore, if any of them is left out of consideration, the remaining two cannot adequately account for all experimental data.

Model

Cellular properties

The model cell [12] represents a low-frequency-tuned IC neuron that responds

to interaural phase cues. Since phase-locking to the carrier frequency among such cells is infrequent and not tight [51], this model assumes that binaural input is rate-encoded. In this model the spikes are eliminated by implicitly averaging over a short time scale (say, 10 ms). Here, we take as the cell's observable state variable its averaged voltage relative to rest (V). This formulation is in the spirit of early rate models (e.g., [32]), but extended to include intrinsic cellular mechanisms such as adaptation and rebound. Parameter values are not optimized, most of them are chosen to have representative values within physiological constraints, as specified in this section.

The current-balance equation for the IC cell is:

$$C\dot{V} = -g_L(V - V_L) - \bar{g}_a a(V - V_a) - I_{syn}(t) - I_{PIR} + I. \quad (16)$$

The terms on the right-hand side of the equation represent (in this order) the leakage, adaptation, synaptic, post inhibitory rebound, and applied currents. The applied current I is zero for the examples considered here and all other terms are defined below.

Leakage current has reversal potential $V_L = 0$ mV and conductance $g_L = 0.2$ mS/cm². Thus, with $C = 1$ μ F/cm², the membrane time constant $\tau_V = C/g_L$ is 5 ms.

The adaptation is modeled by a slowly-activating voltage-gated potassium current $\bar{g}_a a(V - V_a)$. Its gating variable a satisfies

$$\tau_a \dot{a} = a_\infty(V) - a. \quad (17)$$

Here $a_\infty(V) = 1/(1 + \exp(-(V - \theta_a)/k_a))$, time constant $\tau_a = 150$ ms, $k_a = 5$ mV, $\theta_a = 30$ mV, maximal conductance $\bar{g}_a = 0.4$ mS/cm², reversal potential $V_a = -30$ mV. The parameters k_a , θ_a are chosen so that little adaptation occurs below firing threshold; and \bar{g}_a so that when fully adapted for large inputs, the firing rate is reduced by 67%. The value of τ_a , assumed to be voltage-independent, matches the adaptation time scale range seen in spike trains (Semple, unpublished) and it leads to dynamic effects that provide good comparison with results over the range of stimulus rates used in experiments.

The model's readout variable, r , represents firing rate, normalized by an arbitrary constant. It is an instantaneous threshold-linear function of V : $r = 0$ for $V < V_{th}$ and $r = K \cdot (V - V_{th})$ for $V \geq V_{th}$, where $V_{th} = 10$ mV. We set (as in [12]) $K = 0.04075$ mV⁻¹. Its value does not influence the cell's responsiveness, only setting the scale for r .

Synaptic inputs

As proposed by Spitzer and Semple [98] and implemented in the spike-based model of Cai et al. [14, 15], this model IC cell receives binaural excitatory input

from neurons in ipsilateral MSO (e.g., [1, 70]) and indirect (via DNLL [91]) bin-aural inhibitory input from the contralateral MSO. Thus DNLL, although it is not specifically included in the model, is assumed (as in Cai et al. [14]) to serve as an instantaneous converter of excitation to inhibition.

The afferent input from MSO is assumed to be tonotopically organized and IPD-tuned. Typical tuning curves of MSO neurons are approximately sinusoidal with preferred (maximum) phases in contralateral space [97, 112]. We use a sign convention that the IPD is positive if the phase is leading in the ear contralateral to the IC cell, to choose the preferred phases $\phi_E = 40^\circ$ for excitation and $\phi_I = -100^\circ$ for inhibition. Preferred phases of MSO cells are broadly distributed [63, 58]. Similar ranges work in the model, and in that sense, the particular values picked are arbitrary.

While individual MSO cells are phase-locked to the carrier frequency [97, 112], we assume that the effective input to an IC neuron is rate-encoded, as explained above. Under this assumption, the synaptic conductance transients from incoming MSO spikes are smeared into a smooth time course that traces the tuning curve of the respective presynaptic population (Fig. 21). Therefore, the smoothed synaptic conductances, averaged over input lines and short time scales, g_E and g_I , are proportional to the firing rates of the respective MSO populations and they depend on t if IPD is dynamic. Then

$$I_{syn} = -g_E(V - V_E) - g_I(V - V_I), \quad (18)$$

$$g_{E,I} = \bar{g}_{E,I} [0.55 + 0.45 \cos(IPD(t) - \phi_{E,I})]. \quad (19)$$

Reversal potentials are $V_E = 100$ mV, $V_I = -30$ mV, maximum conductance values $\bar{g}_E = 0.3$ mS/cm², $\bar{g}_I = 0.43$ mS/cm².

Rebound mechanism

The post-inhibitory rebound (PIR) mechanism is implemented as a transient inward current (I_{PIR} in the equation (16)):

$$I_{PIR} = g_{PIR} \cdot m \cdot h \cdot (V - V_{PIR}). \quad (20)$$

The current's gating variables, fast (instantaneous) activation, m , and slow inactivation, h , satisfy:

$$m = m_\infty(V), \quad (21)$$

$$\tau_h \dot{h} = h_\infty(V) - h, \quad (22)$$

where $m_\infty(V) = 1/(1 + e^{-(V - \theta_m)/k_m})$, $h_\infty(V) = 1 - 1/(1 + e^{-(V - \theta_h)/k_h})$, $\tau_h = 150$ ms and $V_{PIR} = 100$ mV. The parameter values: $k_m = 4.55$ mV, $\theta_m = 9$ mV, $k_h = .11$

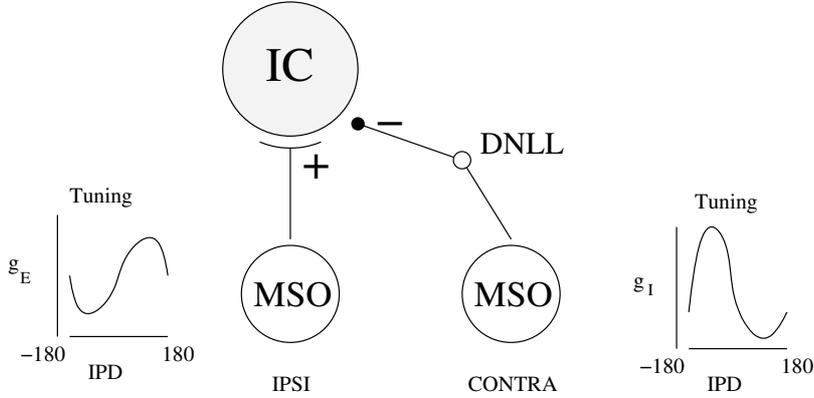


Figure 21: Schematic of the model. A neuron in the inferior colliculus (marked IC), receiving excitatory input from ipsilateral medial superior olive (MSO); contralateral input is inhibitory. Each of the MSO-labelled units represents an average over several MSO cells with similar response properties (see text). DNLL (dorsal nucleus of lateral lemniscus) is a relay nucleus, not included in the model. Interaural phase difference (IPD) is the parameter whose value at any time defines the stimulus. The tuning curves show the stimulus (IPD) tuning of the MSO populations. They also represent dependence of synaptic conductances (g_E and g_I) on IPD (see text). Used by permission of The American Physiological Society from [12].

mV, $\theta_h = -11$ mV, are chosen to provide only small current at steady state for any constant V and maximum responsiveness for voltages near rest. Maximum conductance g_{PIR} equals 0.35 mS/cm² in Fig. 24 and zero elsewhere.

Stimuli

The only sound input parameter that is varied is IPD as represented in equations (18) and (19); the sound’s carrier frequency and pressure level are fixed. The static tuning curve is generated by presentation of constant-IPD stimuli. The dynamic stimuli are chosen from two classes. First, the binaural beat stimulus is generated as: $IPD(t) = 360^\circ \cdot f_b \cdot t \pmod{360^\circ}$, with beat frequency f_b that can be negative or positive. Second, the partial range sweep is generated by $IPD(t) = P_c + P_d \cdot \text{triang}(P_r t / 360)$. Here, $\text{triang}(\cdot)$ is a periodic function (period 1) defined by $\text{triang}(x) = 4x - 1$ for $0 \leq x < 1/2$ and $\text{triang}(x) = -4x + 3$ for $1/2 \leq x < 1$. The stimulus parameters are the sweep’s center P_c ; sweep depth P_d (usually 45°); and sweep rate P_r (in degrees per second, usually $360^\circ/\text{s}$). We call the sweep’s half cycle where IPD increases – the “up-sweep”, and where IPD decreases – the “down-sweep”.

Simulation data analysis

The response’s relative phase (Figs. 22 and 23) is the mean phase of the response

to a binaural beat stimulus minus the mean phase of the static tuning curve. The mean phase is the direction of the vector that determines the response’s vector strength [30]. To compute the mean phase we collect a number of response values (r_j) and corresponding stimulus values (IPD_j in radians). The average phase φ is such that $\tan \varphi = (\sum r_j \sin(IPD_j)) / (\sum r_j \cos(IPD_j))$. For the static tuning curve the responses were collected at IPDs ranging from -180° to 180° in 10° intervals. For the binaural beat the response to the last stimulus cycle is used.

Example of results: binaural beats

The binaural beat stimulus has been used widely in experimental studies (e.g., [52, 61, 96, 107]) to create interaural phase modulation: a linear sweep through a full range of phases (Fig. 22). Positive beat frequency means IPD increases during each cycle; negative beat frequency — IPD decreases (Fig. 22A). Using the linear form of $IPD(t)$ to represent a binaural beat stimulus in the model, we obtain responses whose time courses are shown in Fig. 22A, lower.

To compare responses between different beat stimuli and also with the static response, we plot the instantaneous firing rate vs. the corresponding IPD. Fig. 22B shows an example that is typical of many experimentally recorded responses (e.g., [52, 95, 96]). Replotting our computed response from 27A in 27C reveals features in accord with the experimental data (compare 27B and 27C): dynamic IPD modulation increases the sharpness of tuning; preferred phases of responses to the two directions of the beat are shifted in opposite directions from the static tuning curve; maximum dynamic responses are significantly higher than maximum static responses. These features, especially the phase shift, depend on beat frequency (Fig. 22C, D).

Dependence of phase on the beat frequency

In Fig. 22D the direction of phase shift is what one would expect from a lag in response. Fig. 22C is what would be expected in the presence of adaptation (e.g., [94]; as the firing rate rises, the response is “chopped off” by adaptation at the later part of the cycle. This shifts the average of the response to earlier phases). In particular, if we interpret the IPD change as an acoustic motion, then for a stimulus moving in a given direction there can be a phase advance or phase-lag, depending on the speed of the simulated motion (beat frequency).

To characterize the dependence of the phase shift on parameters, we define the response phase as in a vector-strength computation (see above). We take the average phase of the static response as a reference ($\varphi_0 = 0.1623$). Fig. 23A shows the relative phase of the response ($\tilde{\varphi}_{f_b} - \varphi_0$) vs. beat frequency (f_b). At the smallest beat frequencies the tuning is close to the static case and the relative phase is close to zero. As the absolute value of beat frequency increases, so does the phase

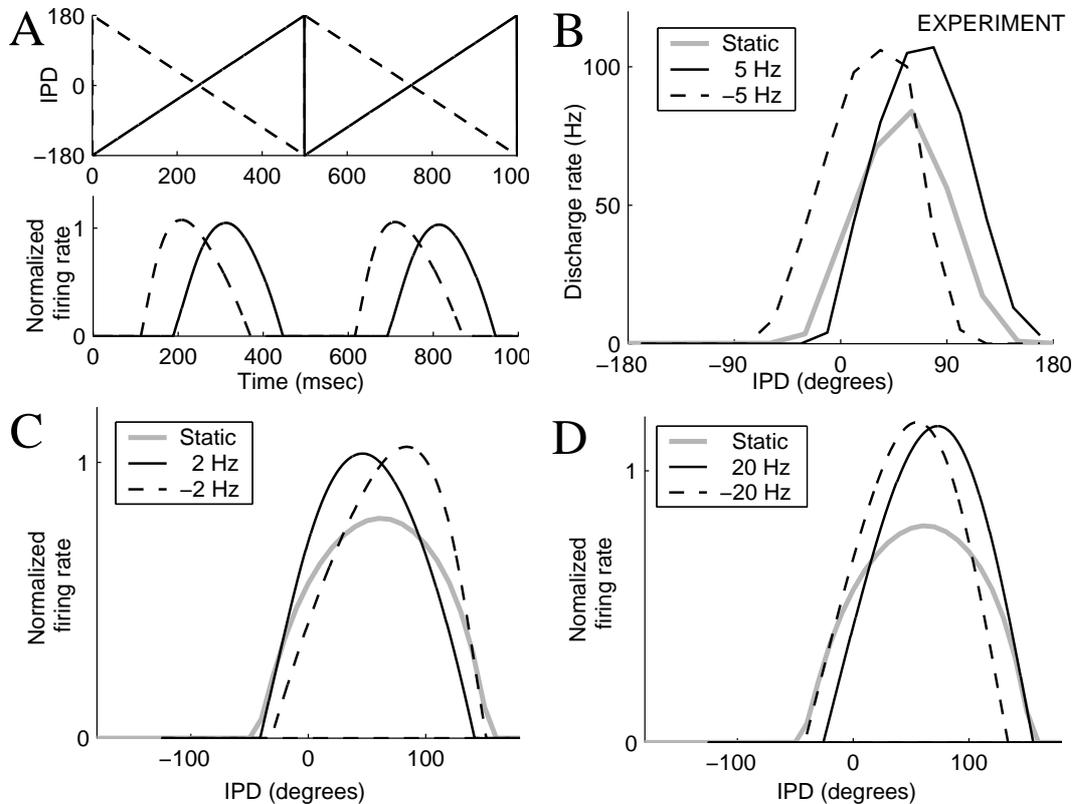


Figure 22: Responses to binaural beat stimuli. *A*: Time courses of IPD stimuli (upper) and corresponding model IC responses (lower). Beat frequencies 2 Hz (solid) and -2 Hz (dashed). *B-D*: Black lines show responses to binaural beats of positive (solid) and negative (dashed) frequencies. Grey line is the static tuning curve. *B*: Example of experimentally recorded response. Carrier frequency 400 Hz, presented at binaural SPL 70 dB. Beat frequency 5 Hz. *C, D*: Model responses over the final cycle of stimulation vs corresponding IPD. Beat frequency is equal to 2 and 20 Hz, respectively. Used by permission of The American Physiological Society from [12].

advance. The phase-advance is due to the firing rate adaptation (see above). At yet higher frequencies a phase lag develops. This reflects the RC -properties of the cell membrane — the membrane time constant prevents the cell from responding fast enough to follow the high frequency phase modulation.

Plots such as in Fig. 23A have been published for several different IC neurons [98, 12] and examples are reproduced in Fig. 23B. Whereas the present model does not have a goal of quantitative match between the modeling results and experimental data, there are two qualitative differences between Figs. 23A and 23B. First, the range of values of the phase shift is considerably larger in the experimental

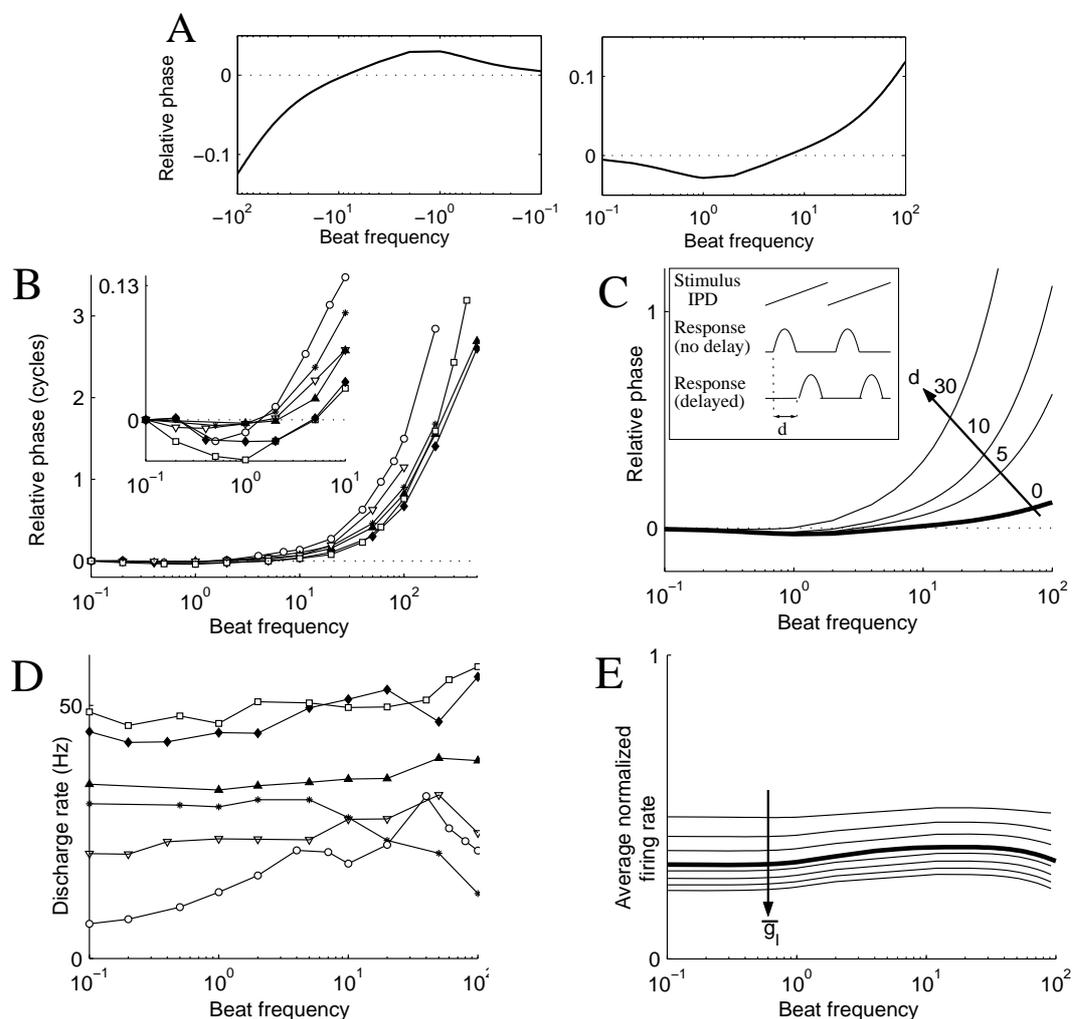


Figure 23: *A*: Dependence of phase shift on beat frequency in the model neuron. Phase shift is relative to the static tuning curve (see text). *B*: Experimentally recorded phase (relative to the phase at 0.1 Hz beat frequency) vs beat frequency. Six different cells are shown. Carrier frequencies and binaural SPLs are: 1400 Hz, 80 dB (star); 1250 Hz, 80 dB (circle); 200 Hz, 70 dB (filled triangle); 800 Hz, 80 dB (diamond); 900 Hz, 70 dB (empty triangle); 1000 Hz, 80 dB (square). Inset is a zoom-in at low beat frequencies. *C*: Phase of the model response, recomputed after accounting for transmission delay from outer ear to IC, for various values of delay (d), as marked. Inset illustrates computation of delay-adjusted phase. *D*, *E*: Average firing rate as a function of beat frequency: experimental (*D*, same cells with the same markers as in *B*) and in the model (*E*). Multiple curves in *E* – for different values of \bar{g}_I ($\bar{g}_I = 0.1, 0.2, 0.3, 0.43, 0.5, 0.6, 0.7, 0.8$ mS/cm²). Thick line — for the parameter values given in *Methods*. Used by permission of The American Physiological Society from [12].

recordings, particularly, the phase lag at higher beat frequencies. Second, it had been reported in the experimental data, that there is no appreciable phase advance at lower beat frequencies. Fig. 23B suggests that the phase remains approximately constant before shifting in the positive direction. Only in the close-up view of the phase at low beat frequencies (Fig. 23B, inset) can the phase-advance be detected. It is a surprising observation, given that the firing rate, or spike frequency, adaptation has been reported in IC cells on many occasions (e.g., [62, 96]) and that, as we noted above, the phase advance is a generic feature of cells with adapting spike trains.

Role of transmission delay

The modeling study [12] suggested that the absence of phase advance and high values of phase lag should be attributed to the transmission delay (five to tens of milliseconds) that exists in transmission of the stimulus to the IC. This delay was not included in the model, as outlined in *Model* above. In fact, we can find its contribution analytically (Fig. 23C, inset). Consider the response without transmission delay to the beat of positive frequency f_b . In the phase computation, the response at each of the chosen times (each bin of PSTH) is considered as a vector from the origin with length proportional to the recorded firing rate and pointing at the corresponding value of IPD. In polar coordinates $\tilde{\mathbf{r}}_j = (r_j, IPD_j)$, where r_j is the j th recorded data point (normalized firing rate), IPD_j is the corresponding IPD. The average phase of response ($\tilde{\varphi}_{f_b}$) can be found as a phase of the sum of these vectors. If the response is delayed, then the vectors should be $\tilde{\mathbf{r}}_j = (r_j, IPD_j + d \cdot f_b \cdot 360^\circ)$, because the same firing rates would correspond to different IPDs (see Fig. 23C, inset). The vectors are the same as without transmission delay, just rotated by $d \cdot f_b$ fractions of the cycle. Therefore, the phase with transmission delay $\varphi_{f_b} = \tilde{\varphi}_{f_b} + d \cdot f_b$. If we plot the relative phase with the transmission delay, it will be $\varphi_{f_b} - \varphi_0 = (\tilde{\varphi}_{f_b} - \varphi_0) + d \cdot f_b$. The phase of the static tuning curve (φ_0) does not depend on the transmission delay. Fig. 23C shows the graph from 23A right, modified by various transmission delays (thick curve is without delay — same as in 23A right). With transmission delay included in the model, the range of observed phase shifts increases dramatically and the phase-advance is masked. In practice neurons are expected to have various transmission delays, which would result in a variety of phase-frequency curves, even if the cellular parameters were similar.

Mean firing rate

Another interesting property of response to binaural beats is that the mean firing rate stays approximately constant over a wide range of beat frequencies (Fig. 23D and [98]). The particular value, however, is different for different cells. The same is valid for the model system (Fig. 23E). The thick curve is the firing rate for the

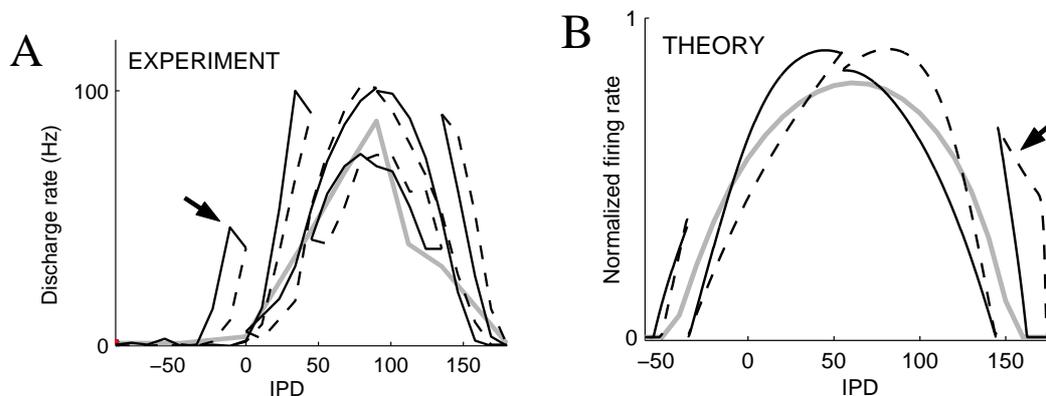


Figure 24: Responses to partial range sweeps. “Rise-from-nowhere”. *A*: Experimentally observed response arising in silent region of the static tuning curve (marked with an arrow). Carrier frequency 1050 Hz; binaural SPL 40 Hz. *B*: Model responses to various sweeps. The rightmost sweep produces “rise-from-nowhere” (arrow). Used by permission of The American Physiological Society from [12].

standard parameter values. It is nearly constant across all tested beat frequencies (0.1 to 100 Hz). The constant value can be modified by change in other parameters of the system, e.g. \bar{g}_I (shown).

Other results

The model that we have just described allows to show [12] that the presence of firing rate adaptation together with IPD-tuned excitatory and inhibitory inputs can explain sharper tuning and phase shifts in response to beats; discontinuous responses to overlapping IPD sweeps; and a decrease in motion sensitivity in the presence of added inhibition. Also, a post-inhibitory rebound mechanism (modeled as a transient membrane current activated by release from prolonged hyperpolarization) allows to explain the strong excitatory dynamic response to sweeps in the silent portion of the static tuning curve (Fig. 24A,B arrows). It also allows to make predictions for future *in vitro* and *in vivo* experiments and suggest experiments that might help to clarify what are the sources of inhibition. For further details and discussion of this model see [12].

7 Thalamus and cortex

The auditory part of the thalamus is the medial geniculate body. It is situated in the caudal part of the thalamus and is the last major relay station for ascending auditory fibers before they reach the cortex. There is a tonotopic arrangement in

the medial geniculate body in which low frequencies are represented laterally and high frequencies are located medially in the principal part. The main projection of the medial geniculate body is to the primary auditory cortex. The medial geniculate body also sends fibers to other thalamic nuclei and may play a part in a regulatory feedback system, with descending projections to the inferior colliculus, the nucleus of the lateral lemniscus, the trapezoid body, and the superior olivary nucleus.

The primary auditory cortex is a primary auditory reception area in cortex and it receives its projections from the medial geniculate body. Areas immediately adjacent to the primary auditory cortex are auditory-association areas. These association areas receive signals from the primary auditory cortex and send projections to other parts of the cortex. Tonotopic organization of the auditory cortex is particularly complex. In the simplest analysis, high frequencies are represented anteriorly and low frequencies posteriorly in the auditory cortex. Each auditory cortical area is reciprocally connected to an area of the same type area in the contralateral hemisphere. In addition, auditory-association areas connect with other sensory-association areas such as somatosensory and visual. They also send projections that converge in the parietotemporal language area. It appears that the higher level of integration in the association areas is responsible for more complex interpretation of sounds. These properties of the auditory cortex may explain why patients with lesions in one of the cortical hemispheres have little difficulty with hearing as measured by presenting simple sounds, e.g., pure tones. However, such patients may have impaired ability to discriminate the distorted or interrupted speech patterns and have difficulty focusing on one stimulus if a competing message, especially meaningful, i.e., speech is presented at the same time.

There is also a descending efferent auditory pathway that parallels the afferent pathway and is influenced by ascending fibers via multiple feedback loops. The specific function of this system in audition is not well understood, but clearly modulates central processing and regulates the input from peripheral receptors.

Because auditory cortex seems to deal primarily with higher brain functions, its modeling has been very scarce. However, recently there has been an promising set of studies by Middlebrooks and colleagues, in which they teach neural networks to interpret cortical recordings and make predictions of the psychophysical performance [24]. More details will have to be unravelled about biophysical properties of cortical auditory cells before more detailed modeling can be successfully used.

References

- [1] Adams JC. Ascending projections to the inferior colliculus, *J. Comp. Neurol.* 183: 519-538, 1979

- [2] Agmon-Snir H, Carr CE, Rinzel J. A case study for dendritic function: improving the performance of auditory coincidence detectors, *Nature* 393: 268-272, 1998
- [3] Allen JB. Two-dimensional cochlear fluid model: new results, *J. Acoust. Soc. Am.* 61: 110-119, 1977
- [4] Allen JB, Sondhi MM. Cochlear macromechanics: time domain solutions, *J. Acoust. Soc. Am.* 66: 123-132, 1979
- [5] Anderson DJ, Rose JE, Hind JE, Brugge JF. Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: frequency and intensity effects, *J. Acoust. Soc. Am.* 49: 1131-1139, 1971
- [6] Banks MI and Sachs MB. Regularity analysis in a compartmental model of chopper units in the anteroventral cochlear nucleus. *J. Neurophysiol.* 65: 606-629, 1991
- [7] Beyer RP. A computational model of the cochlea using the immersed boundary method, *J. Comp. Phys.* 98: 145-162, 1992
- [8] Blackburn CC and Sachs MB. Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis, *J. Neurophys.* 62: 1303-1329, 1989
- [9] Blum JJ, Reed MC, Davies JM. A computational model for signal processing by the dorsal cochlear nucleus. II. Responses to broadband and notch noise. *J. Acoust. Soc. Am.* 98: 181-91, 1995
- [10] Bogert BP. Determination of the effects of dissipation in the cochlear partition by means of a network representing the basilar membrane, *J. Acoust. Soc. Am.* 23: 151-154, 1951
- [11] Borisyuk A, Semple MN, Rinzel J. Computational model for the dynamic aspects of sound processing in the auditory midbrain, *Neurocomputing* 38: 1127-1134, 2001
- [12] Borisyuk A, Semple MN and Rinzel J. Adaptation and Inhibition Underlie Responses to Time-Varying Interaural Phase Cues in a Model of Inferior Colliculus Neurons, *J. Neurophysiol.* 88: 2134-2146, 2002
- [13] Brand A, Behrend O, Marquardt T, McAlpine D, Grothe B. Precise inhibition is essential for microsecond interaural time difference coding, *Nature* 417: 543-547, 2002

- [14] Cai H, Carney LH, Colburn HS. A model for binaural response properties of inferior colliculus neurons. I. A model with interaural time difference-sensitive excitatory and inhibitory inputs. *J. Acoust. Soc. Am.* 103: 475-493, 1998a
- [15] Cai H, Carney LH, Colburn HS. A model for binaural response properties of inferior colliculus neurons. II. A model with interaural time difference-sensitive excitatory and inhibitory inputs and an adaptation mechanism. *J. Acoust. Soc. Am.* 103: 494-506, 1998b
- [16] Carney LH. A model for the responses of low-frequency auditory-nerve fibers in cat. *J. Acoust. Soc. Am.* 93: 401-417, 1993
- [17] Chadwick RF. Studies in cochlear mechanics, in: M.H. Holmes, A. Rubenfeld (Eds.), *Mathematical Modeling of the Hearing Process*, Lecture Notes in Biomathematics, vol. 43, Springer, Berlin, 1980.
- [18] Daniel SJ, Funnell WR, Zeitouni AG, Schloss MD, Rappaport J. Clinical applications of a finite-element model of the human middle ear. *J. Otolaryngol.* 30: 340-346, 2001
- [19] de Boer BP. Solving cochlear mechanics problems with higher order differential equations. *J. Acoust. Soc. Am.* 72: 1427-1434, 1982
- [20] Digital Anatomist Project, Department of Biological Structure, University of Washington, <http://www9.biostr.washington.edu>
- [21] Feddersen WE, Sandal ET, Teas DC, Jeffress LA. Localization of high frequency tones. *J. Acoust. Soc. Am.* 29: 988-991, 1957
- [22] Ferris P, Prendergast PJ. Middle-ear dynamics before and after ossicular replacement. *J. Biomech.* 33: 581-590, 2000
- [23] Fletcher H. On the dynamics of the cochlea. *J. Acoust. Soc. Am.* 23: 637-645, 1951
- [24] Furukawa S, Xu L, and Middlebrooks JC. Coding of Sound-Source Location by Ensembles of Cortical Neurons. *J. Neurosci.* 20: 1216 - 1228, 2000
- [25] Gan RZ, Sun Q, Dyer RK Jr, Chang KH, Dormer KJ. Three-dimensional modeling of middle ear biomechanics and its applications. *Otol. Neurotol.* 23: 271-280, 2002
- [26] Geisler CD. *From sound to synapse: physiology of the mammalian ear.* Oxford University Press, 1998

- [27] Gil-Carcedo E, Perez Lopez B, Vallejo LA, Gil-Carcedo LM, Montoya F. Computerized 3-D model to study biomechanics of the middle ear using the finite element method. *Acta Otorrinolaringol. Esp.* 53:527-537, 2002. Spanish.
- [28] Givelberg E, Bunn J. A Comprehensive Three-Dimensional Model of the Cochlea. *J. Comput. Phys.* 191: 377-391, 2003
- [29] Godfrey DA, Kiang NYS, and Norris BE. Single unit activity in the dorsal cochlear nucleus of the cat. *J. Comp. Neurol.* 162: 269-284, 1975
- [30] Goldberg J and Brown P. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli some physiological mechanisms for sound localization. *J. Neurophysiol.* 32: 613-636, 1969
- [31] Gourevitch G. Binaural hearing in land mammals. In: Yost WA and Gourevitch G (Eds.) *Directional Hearing*, Springer-Verlag, 1987
- [32] Grossberg S. Contour enhancement, short term memory, and constancies in reverberating neural networks. *Stud Appl Math* 52: 213-257, 1973
- [33] Harris DM. Action potential suppression, tuning curves and thresholds: comparison with single fiber data. *Hear. Res.* 1: 133-154, 1979
- [34] Heinz MG, Colburn HS, and Carney LH. Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. *Neural Computation* 13: 2273-2316, 2001
- [35] Heinz MG, Colburn HS, and Carney LH. Evaluating auditory performance limits: II. One-parameter discrimination with random level variation. *Neural Computation* 13, 2317-2339, 2001
- [36] Holmes MN. A mathematical model of the dynamics of the inner ear, *J. Fluid Mech.* 116: 59-75, 1982
- [37] Inselberg A, Chadwick RF. Mathematical model of the cochlea. i: formulation and solution, *SIAM J. Appl. Math.* 30: 149-163, 1976
- [38] Jeffress LA. A place theory of sound localization. *J. Comp. Physiol. Psych.* 41: 35-39, 1948
- [39] Johnson, DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J. Acoust. Soc. Am.* 68: 1115-1122, 1980

- [40] Keener JP, Sneyd J. *Mathematical physiology*, chapter 23. Springer-Verlag, 1998.
- [41] Kelly DJ, Prendergast PJ, Blayney AW. The effect of prosthesis design on vibration of the reconstructed ossicular chain: a comparative finite element analysis of four prostheses. *Otol. Neurotol.* 24: 11-19, 2003
- [42] Kiang NYS, Morest DK, Godfrey DA, Guinan JJ, and Kane EC. Stimulus coding at coudal levels of the cat's auditory nervous system: I. Response characteristics of single units. In *Basic Mechanisms in Hearing* (Moller AR and Boston P., eds.) New York Academic, pp. 455-478, 1973
- [43] Kim DO, Milnar CE, Pfeiffer RR. A system of nonlinear differential equations modeling basilar membrane motion. *J. Acoust. Soc. Am.* 54: 1517-1529, 1973
- [44] Koike T, Wada H, Kobayashi T. Modeling of the human middle ear using the finite-element method. *J. Acoust. Soc. Am.* 111: 1306-1317, 2002
- [45] Kolston PJ. Comparing in vitro, in situ and in vivo experimental data in a three dimensional model of mammalian cochlear mechanics, *Proc. Natl. Acad. Sci. USA* 96: 3676-3681, 1999
- [46] Konishi M. Study of sound localization by owls and its relevance to humans. *Comp. Biochem. Phys. A* 126: 459-469, 2000
- [47] Kuhlmann L, Burkitt AN, Paolini A, Clark GM. Summation of spatiotemporal input patterns in leaky integrate-and-fire neurons: application to neurons in the cochlear nucleus receiving converging auditory nerve fiber input. *J. Comput. Neurosci.* 12: 55-73, 2002
- [48] Kuhn GF. Model for the interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.* 62: 157-167, 1977
- [49] Kuhn GF. Physical acoustics and measurements pertaining to directional hearing. In: Yost WA and Gourevitch G (Eds.) *Directional Hearing*, Springer-Verlag, 1987
- [50] Kuwada S and Yin TCT. Physiological studies of directional hearing. In: Yost WA and Gourevitch G (Eds.) *Directional Hearing*, Springer-Verlag, 1987
- [51] Kuwada S, Yin TCT, Syka J, Buunen TJF, Wickesberg RE. Binaural interaction in low-frequency neurons in inferior colliculus of the cat. IV. Comparison of monaural and binaural response properties. *J. Neurophysiol.* 51: 1306-1325, 1984

- [52] Kuwada S, Yin TCT, Wickesberg RE. Response of cat inferior colliculus neurons to binaural beat stimuli: possible mechanisms for sound localization. *Science* 206 (4418): 586-588, 1979
- [53] Lesser MB, Berkley DA. Fluid mechanics of the cochlea. Part i, *J. Fluid. Mech.* 51: 497-512, 1972
- [54] Leveque RJ, Peskin CS, Lax PD. Solution of a two-dimensional cochlea model using transform techniques. *SIAM J. Appl. Math.* 45: 450-464, 1985
- [55] Leveque RJ, Peskin CS, Lax PD. Solution of a two-dimensional cochlea model with fluid viscosity. *SIAM J. Appl. Math.* 48: 191-213, 1988
- [56] Loh CH. Multiple scale analysis of the spirally coiled cochlea, *J. Acoust. Soc. Am.* 74: 95-103, 1983
- [57] Malone BJ and Semple MN. Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. *J. Neurophysiol.* 86: 1113-1130, 2001
- [58] Malone BJ, Scott BH and Semple MN. Context-dependent adaptive coding of interaural phase disparity in the auditory cortex of awake macaques. *J. Neurosci.* 22, 2002
- [59] Manis PB and Marx SO. Outward currents in isolated ventral cochlear nucleus neurons, *J. Neurosci.* 11: 2865-2880, 1991
- [60] Manoussaki D, Chadwick RS. Effects of geometry on fluid loading in a coiled cochlea, *SIAM J. Appl. Math.* 61: 369-386, 2000
- [61] McAlpine D, Jiang D, Shackleton TM, Palmer AR. Convergent input from brainstem coincidence detectors onto delay-sensitive neurons in the inferior colliculus. *J. Neurosci.* 18: 6026-6039, 1998
- [62] McAlpine D, Jiang D, Shackleton TM, Palmer AR. Responses of neurons in the inferior colliculus to dynamic interaural phase cues: evidence for a mechanism of binaural adaptation. *J. Neurophysiol.* 83: 1356-1365, 2000
- [63] McAlpine D, Jiang D, Palmer AR. A neural code for low-frequency sound localization in mammals. *Nat. Neurosci.* 4: 396-401, 2001
- [64] McAlpine D and Palmer AR. Blocking GABAergic inhibition increases sensitivity to sound motion cues in the inferior colliculus. *J. Neurosci.* 22: 1443-1453, 2002

- [65] Moiseff A. and Konishi M. Neuronal and behavioral sensitivity to binaural time differences in the owl. *J. Neurosci.* 1: 40-48, 1981
- [66] Moller H, Sorensen MF, Hammershoi D and Jensen CB. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.* 43: 300-321, 1995
- [67] Moore BCJ. An introduction to the psychology of hearing, 5th ed. Academic Press, 2003
- [68] Musicant AD, Chan JCK, and Hind JE. Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. *J. Acoust. Soc. Am.* 87: 757-781, 1990
- [69] Neu JC and Keller JB. Asymptotic analysis of a viscous cochlear model. *J. Acoust. Soc. Am.* 1985
- [70] Oliver DL, Beckius GE, and Shneiderman A. Axonal projections from the lateral and medial superior olive to the inferior colliculus of the cat: a study using electron microscopic autoradiography. *J. Comp. Neurol.* 360: 17-32, 1995
- [71] Parham K and Kim DO. Spontaneous and sound-evoked discharge characteristics of complex-spiking neurons in the dorsal cochlear nucleus of the unanesthetized decerebrate cat. *J. Neurophys.* 73: 550-561, 1975
- [72] Peruzzi D, Sivaramakrishnan S, Oliver DL. Identification of cell types in brain slices of the inferior colliculus. *J. Neurosci.* 101: 403-416, 2000
- [73] Peskin CS. Partial differential equations in biology. Courant Inst. Lecture Notes, Courant Inst. of Mathematical Sciences, NYU, New York, 1975-76
- [74] Peskin CS. Lectures on mathematical aspects of physiology. *Lectures in Appl. Math.* 19: 38-69, 1981
- [75] Peskin CS. The immersed boundary method, *Acta Numerica* 11, 479-517, 2002
- [76] Peterson LC, Bogert BP. A dynamical theory of the cochlea, *J. Acoust. Soc. Am.* 22: 369-381, 1950
- [77] Popper AN, Fay RR, eds. The mammalian auditory pathway: neurophysiology. New York : Springer-Verlag, 1992
- [78] Ranke OF. Theory of operation of the cochlea: a contribution to the hydrodynamics of the cochlea. *J. Acoust. Soc. Am.* 22: 772-777, 1950

- [79] Rayleigh L. On our perception of sound direction. *Phil. Mag.* 13: 214-232, 1907
- [80] Reed MC, Blum JJ. A computational model for signal processing by the dorsal cochlear nucleus. I. Responses to pure tones. *J. Acoust. Soc. Am.* 97: 425-438, 1995
- [81] Reed MC, Blum J. A model for the computation and encoding of azimuthal information by the Lateral Superior Olive. *J. Acoust. Soc. Amer.* 88, 1442-1453, 1990
- [82] Reyes AD, Rubel EW, Spain WJ. *In vitro* analysis of optimal stimuli for phase-locking and time-delayed modulation of firing in avian nucleus laminaris neurons. *J. Neurosci.* 16: 993-1000, 1994
- [83] Rose JE, Gross NB, Geisler CD, Hind JE. Some neural mechanisms in inferior colliculus of cat which may be relevant to localization of a sound source. *J Neurophysiol* 29: 288-314, 1966
- [84] Rosowski JJ. Models of external- and middle-ear function. In: *Auditory Computation*, edited by H.L. Hawkins, T.A. McMullen, A.N. Popper, and R.R. Fay. New York: Springer-Verlag, 1996
- [85] Rothman JS, Young ED, Manis PB. Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: implications of a computational model. *J. Neurophysiol.* 70: 2562-2583, 1993
- [86] Ruggero MA and Semple MN, "Acoustics, Physiological", *Encyclopedia of Applied Physics*, vol. 1, 1991
- [87] Sanes DH, Malone BJ, Semple MN. Role of synaptic inhibition in processing of dynamic binaural level stimuli. *J. Neurosci.* 18: 794-803, 1998
- [88] Shackleton TM, McAlpine D, Palmer AR. Modelling convergent input onto interaural-delay-sensitive inferior colliculus neurones. *Hearing Res.* 149: 199-215, 2000
- [89] Shaw EAG. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J. Acoust. Soc. Am.* 56: 1848-1861, 1974
- [90] Shepherd GM, editor. *The Synaptic Organization of the Brain*, 5th ed., Oxford University Press, 2004

- [91] Shneiderman A, Oliver DL. EM autoradiographic study of the projections from the dorsal nucleus of the lateral lemniscus: a possible source of inhibitory input to the inferior colliculus. *J. Comp. Neurol.* 286: 28-47, 1989
- [92] Siebert WM. Ranke revisited a simple short-wave cochlear model, *J. Acoust. Soc. Am.* 56: 596-600, 1974
- [93] Sivaramakrishnan S, Oliver DL. Distinct K currents result in physiologically distinct cell types in the inferior colliculus of the rat. *J. Neurosci.* 21: 2861-2877, 2001
- [94] Smith GD, Cox CL, Sherman SM, Rinzel J. Spike-frequency adaptation in sinusoidally-driven thalamocortical relay neurons. *Thalamus and related systems* 1: 1-22, 2001
- [95] Spitzer MW and Semple MN. Interaural phase coding in auditory midbrain: influence of dynamic stimulus features. *Science* 254: 721-724, 1991
- [96] Spitzer MW and Semple MN. Responses of inferior colliculus neurons to time-varying interaural phase disparity: effects of shifting the locus of virtual motion. *J. Neurophysiol.* 69: 1245-1263, 1993
- [97] Spitzer MW and Semple MN. Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. *J. Neurophysiol.* 73: 1668-1690, 1995
- [98] Spitzer MW and Semple MN. Transformation of binaural response properties in the ascending pathway: influence of time-varying interaural phase disparity. *J. Neurophysiol.* 80: 3062-3076, 1998
- [99] Steele CR. Behaviour of the basilar membrane with pure-tone excitation, *J. Acoust. Soc. Am.* 55, 148-162, 1974
- [100] Steele CR, Taber LA. Comparison of wkb calculations and experimental results for three-dimensional cochlear models, *J. Acoust. Soc. Am.* 65, 1007-1018, 1979
- [101] Steele CR, Zais JG. Effect of coiling in a cochlear model, *J. Acoust. Soc. Am.* 77: 1849-1852, 1985
- [102] Stevens SS and Newman EB. The localization of actual sources of sound. *Am. J. Psychol.* 48: 297-306, 1936

- [103] Sun Q, Gan RZ, Chang KH, Dormer KJ. Computer-integrated finite element modeling of human middle ear. *Biomech. Model Mechanobiol.* 1: 109-122, 2002
- [104] Tan Q, Carney LH. A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. *J. Acoust. Soc. Am.* 114: 2007-2020, 2003
- [105] Viergever MA. Basilar membrane motion in a spiral shaped cochlea. *J. Acoust. Soc. Am.* 64: 1048-1053, 1978
- [106] Watkins AJ. The monaural perception of azimuth: a synthesis approach. In: Gatehouse RW (Ed.) *Localization of sound: theory and applications*. Amphora Press, 1982
- [107] Wernick JS, Starr A. Binaural interaction in superior olivary complex of cat – an analysis of field potentials evoked by binaural-beat stimuli. *J. Neurophysiol.* 31: 428, 1968
- [108] Wiener FM, Pfeiffer RR, Backus ASN. On the sound pressure transformation by the head and auditory meatus of the cat. *Acta oto-laryngol (Stockh.)* 61: 255-269, 1966
- [109] Wightman FL, and Kistler DJ. Sound localization. In *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R. Fay. New York: Springer-Verlag, pp 155-192, 1993
- [110] Woodworth RS. *Experimental psychology*. New York: Holt, 1938
- [111] Xu Y, Collins LM. Predicting the threshold of single-pulse electrical stimuli using a stochastic auditory nerve model: the effects of noise. *IEEE Trans Biomed Eng.* 50: 825-835, 2003
- [112] Yin TCT and Chan JCK. Interaural time sensitivity in medial superior olive of cat. *J. Neurophysiol.* 64: 465-488, 1990
- [113] Yin TCT and Kuwada S. Binaural interaction in low frequency neurons in inferior colliculus of the cat. II. Effects of changing rate and direction of interaural phase. *J. Neurophysiol.* 50: 1000-1019, 1983a
- [114] Yin TCT and Kuwada S. Binaural interaction in low frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. *J. Neurophysiol.* 50: 1020-1042, 1983b

- [115] Yin TCT, Kuwada S, Sujaku Y. Interaural time sensitivity of high frequency neurons in the inferior colliculus. *J. Acoust. Soc. Am.* 76: 1401-1410, 1984
- [116] Yost WA. *Fundamentals of hearing: an introduction*, 4th ed. Academic Press, 2000
- [117] Young SR and Rubel EW. Frequency-specific projections of individual neurons in chick brainstem auditory nuclei. *J. Neurosci.* 3: 1373-1378, 1983
- [118] Yost WA and Hafter ER. Lateralization. In: Yost WA and Gourevitch G (Eds.) *Directional Hearing*, Springer-Verlag, 1987
- [119] Zhang X, Heinz MG, Bruce IC, Carney LH. A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am.* 109: 648-670, 2001
- [120] Zweig G, Lipes R, Pierce JR. The cochlear compromise, *J. Acoust. Soc. Am.* 59: 975-982, 1976
- [121] Zwislocki JJ. Analysis of some auditory characteristics, in: R.R. Buck, R.D. Luce, E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Wiley, New York, 1965.
- [122] <http://hwr.nici.kun.nl/miami/taxonomy/node20.html>