Harvard-MIT Division of Health Sciences and Technology
HST.723: Neural Coding and Perception of Sound
Instructor: John Guinan

HST 723. Neural Coding and Perception of Sound: Functional cochlear mechanics.
by John Guinan

Sources for reviewing Cochlear Mechanics:

Geisler, C. D. (**1998**). *From sound to synapse. Physiology of the mammalian ear.* (Oxford Univ. Press, New York.).

Patuzzi, R. (**1996**). "Cochlear Micomechanics and Macromechanics," in *The Cochlea*, edited by P.J. Dallos, A.N. Popper and R.R. Fay (Springer-Verlag, New York), pp. 186-257.

Pickles, J. O. (**1988**). *An Introduction to the Physiology of Hearing* (Academic Press, New York).

Hearing Mechanics in brief

Sound is collected by the outer ear and led to the Middle ear. The middle ear acts as a transformer to maximize power transfer from sound in air (a low impedance medium) to sound in water (a high impedance medium). The inner ear: (1) amplifies the sound, (2) separates it into frequency components, and (3) transduces it into time patterns of nerve spikes across auditory-nerve fibers that innervate different frequency regions of the cochlea. These auditory-nerve fibers then carry the encoded sound signal to the central nervous system.

Figures removed due to copyright reasons.

Sound gathering by the external ear is directional and this directionality is important for sound localization, particularly in elevation.

Figures removed due to copyright reasons.

The mammalian cochlea has three fluid compartments, Scala Vestibuli, Scala Media and Scala Tympani (SV, SM, ST). SV & ST are filled with "perilymph" which is high in Sodium like other extracellular fluids. SM is filled with "endolymph", which is high in Potassium like intracellular fluids. The cells surrounding SM have insulating borders and SM is 80 to 100 mV positive relative to the surrounding tissue. This "Endocochlear Potential" (EP) provides a battery that powers hair-cell transduction.

The stiff Basilar Membrane (BM) divides the cochlea into two mechanical fluid compartments (SV+SM, and ST). Reissner's membrane, which divides SV and SM, has a high mechanical compliance so SV and SM act as one compartment.

On top of the BM is the "Organ of Corti" which contains several kinds of supporting cells and two kinds of sensory cells, outer and inner hair cells (OHCs and IHCs). IHCs are the main sensory cells and are synapsed on by the bulk of the auditory-nerve fibers that carry information to the central nervous system. OHCs have mainly a mechanical function and are involved in amplifying sound energy traveling along the cochlea. Hair cells are activated by bending their stereocilia (hairs).

Above the hair cells is a gelatinous structure, the Tectorial Membrane (TM). The tops of OHC stereocilia are attached to the TM. IHC stereocilia come close to the TM but are not usually connected to it. Relative motion between the TM and the top of the organ of Corti (the Reticular Lamina) bends the hair-cell stereocilia, thereby causing hair-cell excitation.

It is easiest to understand what is happening in the cochlea by viewing it as uncoiled (Fig. 5.1). The coiling of the cochlea mainly acts to make it compact.

The bony-walls of the cochlea have two main holes, the "Oval Window" in which the Stapes footplate sits, and the "Round Window" which has a mechanically compliant membrane that separates the Scala Tympani fluid from the air-filled middle-ear cavity. There is also a fluid connection to the vestibular apparatus, but since this is enclosed in bone with no major holes, there is little sound-energy that goes to the vestibular apparatus.

A simplified overview of cochlear mechanics

Sound at the tympanic membrane produces motion of the middle ear ossicles. This leads to stapes motion in-and-out of the cochlea which displaces fluid in Scala Vestibuli. Since the cochlear walls are rigid, displacement *in* at the stapes is matched by Scala Tympani displacement *out* at the round window. In most species there is a connection between Scala Vestibuli and Scala Tympani at the end of the cochlea, the "Helicotrema". For very low frequency sounds, stapes displacements lead to fluid flow through the Helicotrema. However, higher frequency sounds (i.e. for most sounds of interest) the acoustic impedance of the fluid mass in the cochlea is higher than the effective acoustic impedance of the stiffness of the cochlear partition (which is mostly due to the stiffness of the basilar membrane). The result is that when the stapes moves in, part of the cochlear partition moves down and the round window bulges out. The stiffness of the basilar membrane is about two orders of magnitude lower in the apex than in the base, this makes displacement of the cochlear partition easier at the apex than at the base. In contrast, if the displacement is at the apex, fluid along a longer part of the cochlea must be moved which is equivalent to more mass, so movements near the base are easier from a mass point of view. The resultant balance between fluid mass and BM stiffness causes the effective resonant point to move along the length of the cochlea with the high-frequency resonant point near the stapes and the low-frequency resonant point near the helicotrema. This is an oversimplification, but it is a good way to remember that the place near the stapes responds best to high frequencies. In reality, the fluid movement is a surface wave along the cochlear partition, not a bulk flow, and BM stiffness may be the dominant factor in setting the resonance position.

Cochlear Macromechanics

A pure-tone sound sets up a "Traveling Wave" motion along the basilar membrane. The peak of the traveling wave is near the stapes (the base of the cochlea) for high frequencies and moves toward the apex for low frequencies. Throughout most of the frequency range, the location of the peak is logarythmically related to the frequency of the tone so peaks an octave apart are separated by the same distance along the length of the cochlea.

The measurements shown on this page were obtained by von Bekesy using human cadavers. In comparison to the basilar-membrane motions in intact, live cochleas, cadaver frequency patterns are similar but cadaver responses are linear and don't show sharp peaks or evidence for traveling-wave amplification within the cochlea.

Figures removed due
to copyright reasons.

Cochlear Macromechanics

The figure below shows responses from a live chinchilla cochlea illustrating important properties of basilar-membrane response: (1) near the peak of the response to a tone, the amplitude of basilar membrane motion grows far slower than linear (i.e. is "compressive"), while far from the peak, the response is linear (panels a, b), (2) at low sound levels the peak response shows a large gain in motion relative to stapes velocity (panel b), and (3) at low sound levels the response is strongly frequency selective and is similar to the tuning of auditory-nerve fiber responses (which are shown by the dotted line labeled "Neural" in panel d).

Scaling: The response pattern seen here applies across the cochlea, at least for frequencies > 1 kHz.

Figure removed due
to copyright reasons.

Models of Cochlear Macromechanics

A fluid-mechanical model can be made of the cochlea by considering the mechanical properties in many thin cross sections through the uncoiled cochlea. In such a model, energy is coupled along the length of the cochlea by the motion of the fluids. At any point along the length of the cochlea, the response depends on the fluid-pressure differential across the cochlear partition, and the mechanical impedance of the cochlear partition at that point. The resulting model is a transmission-line with a characteristic impedance that changes along the length of the cochlea. This cochlear transmission line is strongly frequency dispersive, which is different from the typical transmission line found in engineering. However, it is similar to typical transmission lines in allowing energy propagation in both directions.

The transmission-line properties of the cochlea lead to a set of phenomena called "Otoacoustic Emissions" (OAEs) in which energy that originated in the cochlea is carried backward along the cochlear transmission line, goes backward through the middle ear, and appears in the ear canal. OAEs can be measured with a sensitive microphone in almost all normal ears. OAEs give important clues to cochlear function but are outside the scope of this course.

Models usually assume that cochlear mass and stiffness lead to a single resonance at each place along the length of the cochlea. However, data from a variety of sources suggest that there are multiple resonances that shape the response at each cochlear place and that these resonances may be due to different vibrational modes in the motion of the basilar membrane and organ of Corti.

Figure removed due
to copyright reasons.

The vibration pattern of the cochlea.

In the simplest view, when the cochlear partition moves up, there is a sheering between the tectorial membrane and the reticular lamina (the black line at the top of the hair cells in Fig. 3.3) that leads to bending of the hair-cell stereocilia. OHC stereocilia tips are attached to the tectorial membrane and are directly bent. IHC stereocilia are freestanding and are bent by the fluid motion. However, measurements of basilar-membrane and organ of Corti motion (in exised cochleas) shows that the motion is much more complex and changes as a function of frequency.

Hair-Cell Mechano-Electric Transduction: Tip links.

Hair-cell excitation is produced when the stereocilia are bent in the direction of the tallest stereocilia. The leading theory to explain this is that thin wisps of material between the tops of the stereocilia (the "tip links") are stretched by bending the stereocilia. When the tip links are stretched, they increase the probability of opening channels at the tops of the stereocilia (the little "trap-door" shown at the base of the Gating Spring in Fig. 6.6). Opening these channels allows current to flow into the hair cell.

When the stereocilia are bent for a long time (tens of ms) the channels begin to close (i.e. "adapt). This adaptation is accomplished by moving the attachment point of the spring up and down the tallest stereocilia. In normal sound stimulation the stereocilia move back and forth so fast that there is no significant adaptation (however, adaptation is important in hair cells in the vestibular apparatus that act as gravity receptors and can see steady deflections).

There is evidence from some non-mammalian vertebrates that the hair-cell stereocilia channel + tip-link apparatus can actually supply sound-frequency mechanical energy (up to 1 kHz) which may act to amplify mechanical responses. Whether this happens in mammals is an open question that we will return to later.

Figures removed due
to copyright reasons.

When the stereocilia are bent and the tip links cause transducer channels to open current flows into the hair cell. The driving force for the current is provided by the potential difference from scala media to the hair-cell interior (both are high potassium so concentration gradients have little influence).

This "receptor current" causes a voltage drop (the "receptor potential") as it flows out across the membrane of the body of the hair cell. This cell membrane (as in all cells) acts like a resistor in parallel with a capacitor and has a cut off frequency of a few hundred Hz (in the apex) to at most a few thousand Hz in the base.

Depolarization at the base of the hair cell causes transmitter to be released and leads to activation of auditory-nerve fibers.

Figure removed due
to copyright reasons.

The ion channels that are opened in the stereocilia are nonlinear and have current vs. displacement curves in the form shown in Fig. 3.21. Displacement in one direction produces a big positive current, while an equivalent displacement in the other direction produces only a small negative current. Thus, averaged over a cycle, there is a net DC produced. This DC is particularly important in releasing transmitter at frequencies above the membrane cut-off frequency.

Figure removed due
to copyright reasons.

Hair-Cell Mechano-Electric Transduction: Shape across frequency

The combination of the non-linear stereocilia transducer channel and the hair-cell membrane low-pass filter act to change the shape of the receptor potential across different frequencies. At low frequencies the AC component of the receptor potential is largest while at high frequencies the DC component of the receptor potential is largest. Since the receptor potential produces the transmitter release, these hair-cell properties are the main reason that cycle-by-cycle timing information is preserved in auditory nerve fibers at low frequencies but not at high frequencies.

Figure removed due
to copyright reasons.

The "Cochlear Amplifier"

Compared to the velocity of the stapes for low-level sounds, the basilar membrane shows a great increase in velocity at the peak of its response. Model calculations indicate that this increase in mechanical response is largely due to the injection of sound-frequency energy into the traveling wave, a process called the "cochlear amplifier". A variety of evidence strongly points to the Outer Hair Cells as the source of this cochlear amplifier. There are two competing theories for the motor element in the Cochlear Amplifier: (1) OHC somatic motility, and (2) Stereocilia motility.

Outer Hair Cell somatic Motility

When you put a voltage across OHCs, they change their length. The length change is controlled by the OHC transmembrane voltage. Depolarizing the OHC makes it shorter: hyperpolarizing the OHC makes it longer. The OHC length change appears to be due to an unusual protein that lines the side walls of OHCs. When the voltage across the OHC membrane is changed, the protein is presumed to change its shape thereby causing a change in the length of the OHC.

Figure removed due
to copyright reasons.

OHC Somatic Motility and the "Cochlear Amplifier"

How OHC somatic motility leads to the cochlear amplifier is not clear. The mechanical structures that couple OHCs to the basilar membrane are complex (Fig 2.10). However, the presumed mechanism would work something like this: The basilar membrane moves up causing sheer between the reticular lamina and the tectorial membrane that bends OHC stereocilia. This stereocilia bending opens channels and allows current to enter the OHC. The resulting OHC receptor potential causes an OHC length change (This length change is accompanied by a much bigger stiffness change. Whether the length or the stiffness change is most important is not clear). This length change is then somehow coupled back to the basilar membrane in the right phase to amplify the traveling wave.

Figure removed due
to copyright reasons.

Stereocilia Motility and the "Cochlear Amplifier"

As mentioned earlier, the bending of stereocilia pulls on the tip links which pulls on the gating channels and increases their open probability. Such a physical system is likely to be bilateral, i.e. if the gating channel closes for some reason it pulls on the tip link which then pulls on the stereocilia and this pull might be able to move the whole structure. The open probability of the channels has been shown to depend on the calcium concentration at the site of the channel inside the stereocilia. Furthermore, when the channel is open, calcium from the extracellular fluid flows into the cell. This could then make a feedback loop which at the right frequency could have a positive gain so that stereocilia motion at sound frequencies is actually amplified. Such an amplifying effect has been shown for certain non-mammalian hair cells. Since the structures involved are similar in mammals, this mechanism might also work in mammals.

Figure removed due
to copyright reasons.

The "Cochlear Amplifier": OHC Somatic Motility vs. Stereocilia Motility

Both theories for the motor element of the cochlear amplifier appear to be limited in the upper frequency at which they would work by inherent low-pass filters. OHC somatic motility is driven by the membrane voltage but the membrane low-pass filter limits the high-frequency voltage change produced for a given receptor current (or stereocilia bending). Stereocilia motility is limited by the low-pass produced by the time-integration needed for diffusion to build up the calcium concentration at the intra-stereociliar side of the transduction channel.

Despite the parallel high-pass limitations, OHC somatic motility seems more likely to be the cochlear amplifier motor. OHC somatic motility can produce much more force per hair cell than the stereocilia motor. More, importantly, positive energy output by stereocilia motility has never been shown above about 1 kHz, whereas OHC somatic motility has been shown to be able to drive the whole cochlear partition (with the stereocilia channels blocked) up to 70 kHz.

Figure removed due
to copyright reasons.

Figure removed due
to copyright reasons.

Cochlear Amplification can be reduced by a second tone, this produces "Suppression"

Figure removed due
to copyright reasons.

Cochlear Nonlinearity and Suppression

We have already noted that within the peak region of basilar-membrane response, the response grows nonlinearly (compressively) with sound level. The same mechanisms that produce this compressive growth also produce a phenomena known as "two-tone suppression." We will call this just "suppression" because although it is most easily understood by considering two tones, it operates under a much wider range of circumstances. Suppression is usually divided into "high-side" and "low-side" suppression according to whether the frequency of the suppressor (Fs) is above or below the frequency of the test tone (Ftt) whose response is being suppressed.

High-side suppression

High-side suppression has a low threshold, i.e. the suppressor tone begins to suppress when it is comparable in level to the test tone. High-side suppression has a low growth rate, i.e. the suppression increases slowly with increases in suppressor level (less than 1 dB of suppression for every 1 dB increase of the suppressor).

Figure removed due
to copyright reasons.

14

Low-side suppression

Low-side suppression has a high threshold, i.e. the suppressor tone begins to suppress only when it at a relatively high sound level (e.g. 60-80 dB SPL). Low-side suppression has a high growth rate, i.e. the suppression increases rapidly with increases in suppressor level (2-3 dB of suppression for every 1 dB increase in the suppressor). In addition, for suppressors of low enough frequency, the suppression increases and decreases within one cycle of the suppressor.

Figures removed due to copyright reasons.

Understanding Suppression

Cochlear amplification depends on the OHCs sensing the motion of the basilar membrane and then doing something (either by somatic motility or stereocilia motility) that amplifies the motion. OHC motion sensing comes about by the bending of OHC stereocilia which leads to a conductance change and OHC receptor current. However, if the stereocilia are bent too far they go into a saturation region where further bending has little or no effect. Suppression is produced when a suppressing tone causes the OHC stereocilia to go into their nonlinear region so that the stereocilia no longer properly sense the test tone. If the sensing of the test tone is reduced, the OHC motor output at the test-tone frequency will be reduced and amplification of the test tone will be reduced. The result is suppression of the test-tone response.

A key thing to know in understanding suppression is that the traveling wave of the test tone is amplified only in a relatively small region basal to the peak of the test tone response. A high-side suppressor can have a big response in this region because its peak response is in this region, which accounts for the low threshold of high-side suppression. However, a high-side suppressor will only affect part of the test-tone amplification region, and the growth of the suppressor response is compressive, so high-side suppression grows slowly. In contrast, low side suppression has a high threshold because high sound levels are required before the suppressor response is big in the test-tone amplification region. However, when the low-side suppressor reaches this threshold it acts with a high growth rate because its response grows linearly which is much faster than the compressive growth of the test-tone response.

Figure removed due
to copyright reasons.

HST 723. Neural Coding and Perception of Sound: Otoacoustic Emissions.
by John Guinan

Otoacoustic Emissions (OAEs) are sounds measured in the ear canal that originate in the cochlea. Figure 1 shows a typical measurement arrangement: a sound source (in this case to generate a click) and a sensitive microphone, with the microphone output averaged over many repetitions. Figure 2 shows examples of three kinds of OAEs:

Spontaneous OAEs (SOAEs) are emissions present when no stimulus is put in the ear (found in about half of human ears with normal hearing).

Transient Evoked OAEs (TEOAEs) are emissions evoked by a brief sound (either a click or a tone burst). The TEOAE is recorded after waiting for the sound from the source to die away (as determined from measurements in a cavity).

Stimulus Frequency OAEs (SFOAEs) are emissions at the same frequency as the tone from the source. The effect of SFOAEs adding and subtracting with the source tone can be seen by sweeping the tone frequency.

More OAE Examples

Distortion Product OAEs (DPOAEs) are emissions created by distortion processes within the ear. Distortion products are created at all of the sum and difference frequencies of f_1 and f_2 but the largest distortion product is at $2f_1-f_2$, where f_1 and f_2 are pure tones of frequency f_1 and f_2 with $f_1<f_2$.

All OAEs make use of energy created by the cochlear amplifier and their presence at normal amplitudes is an indication that the cochlear amplifier is working properly. This is the basis of the use of OAEs as clinical tests.

As shown in Table 1, otoacoustic emissions have been classified by the stimulus that creates them, without consideration of the mechanisms by which they are generated.

Figures removed due
to copyright reasons.

OAEs can be classified by the processes by which they are generated: (1) linear coherent reflection mechanisms, or (2) distortion mechanisms

Some emissions, particularly DPOAEs, normally contain significant components from both distortion and linear coherent reflection mechanisms (Fig. 12, bottom).

Figures removed due to copyright considerations. Please see:
Shera, C. A., and J. J. Guinan, Jr. "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs." *J Acoust Soc Am* 105, no. 2 Pt 1 (Feb 1999): 782-98.

SFOAE properties.

SFOAEs are the basis of a new test for determining the frequency selectivity of the cochlea so we will look at them more closely. SFOAEs grow compressively. If a tone source is swept in frequency at several levels the result changes with level. At high levels, the ear-canal SPL varies slowly with frequency due to the slowly varying properties of the sound source and the input impedance of the ear. As sound level is lowered, the SFOAE becomes an increasing fraction of the total pressure and it adds or subtracts from the pressure directly produced by the source, depending on the phase relationship between them (see Fig. 1). At very low sound levels, SFOAE production is almost linear. If we scale the result at the highest level and subtract it from the lowest level, we get an approximate measure of the SFOAE. Such measures of SFOAEs show a slowly varying amplitude and a fast varying phase (Fig. 3). This fast varying phase is what one would expect if the SFOAE was an "echo" with a delay. Although this "echo" idea provides a convenient way to think about reflection emissions, it obscures the origin of this as being due to a wave group delay produced by linear coherent reflection (see next page).

Linear Coherent Reflection

Reflection emissions are produced by a process called "coherent reflection" which by itself is linear, although it takes place within a nonlinear system in which amplitudes grow compressively.

When energy travels along a transmission line (such as along the cochlea) any discontinuity in the impedance of the transmission line creates a partial reflection of the energy. We presume that there are numerous small random irregularities throughout the length of the cochlea each which reflects a little bit of energy. For most frequencies, this reflected energy cancels in the ear-canal because the phases of the various components cover 360 degrees. However, just at the peak of the traveling wave, where the amplitude of the wave is big and the phase varies linearly with frequency, the reflected waves can add and produce a substantial amplitude in the ear canal (the effect is like a narrow-band filter on a broad band noise). This process produces an reflected wave that has a long group delay which depends on the phase characteristic of the traveling wave at its peak. If we presume that traveling waves all have about 2-3 large peaks within their bandwidth (as shown in Fig. 3), then the narrower the tuning, the faster the phase vs. frequency plot and the longer the group delay. Using this in reverse, the group delay tells you about the bandwidth.

Figure removed due
to copyright reasons.

Figure removed due to copyright reasons.
Please see: Zweig G., and C. A. Shera. "The origin of
periodicity in the spectrum of evoked otoacoustic emissions."
J Acoust Soc Am 98, no. 4 (Oct 1995): 2018-47.