The frequency analysis of the cochlea – a review of Nobili et al (1998) and Ruggero et al (1992)

by Pedro da Fonseca (pedrofon@mail.telepac.pt) – Neuroscience course
Presented in 17.12.99 to professor STEPHEN MCADAMS

Introduction

Hearing processes seems to divide into three main categories: the caption and transformation of physical sound waves, the transduction of the physical waves into neural patterns and the transformation of these neural patterns into an useful mirror of the physical world. It was thought that the processes of distortion of the physical wave sound, either to enhance the detection of sound origin, either to amplify (or reduce) certain frequencies, was restricted to the pinna (and the head) and middle ear. By the contrary, the inner ear was conceived as a passive transduction mechanism that accurately transduced the physical vibrations produced by the stapes into neural patterns. It is now apparent that the inner ear also interacts with physical waves through a mechanism that is not well known in its details but that seems to involve the motility of the outer hair cells (OHC). This physical capability of the inner ear, besides enhancing frequency discrimination,[1] also explains the existence of otoacoustic emissions, and some auditory illusions (such as the existence of combination tones, or two tone suppressing).

The two articles that are here under analysis deal primarily with the defence of this "active" model of the cochlea. In more recent literature it seems consensual that the cochlea has indeed an active participation in the displacement of internal waves and the problem is now to find a working model that correctly predict experimental results and identifies the mechanism.

Critical analysis

"How well do we understand the cochlea?"

The objective of this article is to introduce the active model of the cochlea. It starts by briefly presenting Helmholtz model of the cochlea as a set of independent strings of varying stiffness, then goes on to Békésy model of the passive cochlea, and then presents the experiments that lead to the active model of the cochlea and presents several difficulties related to this model.

The first model of the cochlea was explicitly presented by Helmholtz and according to it the cochlea has a big set of different tensors, each responding to a particular (range of) frequency(s). They were disposed inside the cochlea "as a dense set of elastic fibres tensioned across the fluid-filled cochlear duct with fibre stiffness decreasing exponentially along the

[1] See Rosenzweig (1999, 226), ??

coiled axis of the cochlea." (160) Sensors were therefore supposed to be analogous to strings, each had a particular vibration period that depended on its stiffness.

So, submitted to the presence of a complex wave – sometimes made by thousands of different waves – each sensor would respond (resonate) if and only if a (component) wave with appropriate frequency was present. Although Helmholtz' basic idea is still considered correct today, and we now even know better which particular parts of the cochlear duct correspond to the detection of specific frequencies (see our fig. 1), the specific mechanism which he devised for the cochlea seems actually quite different.

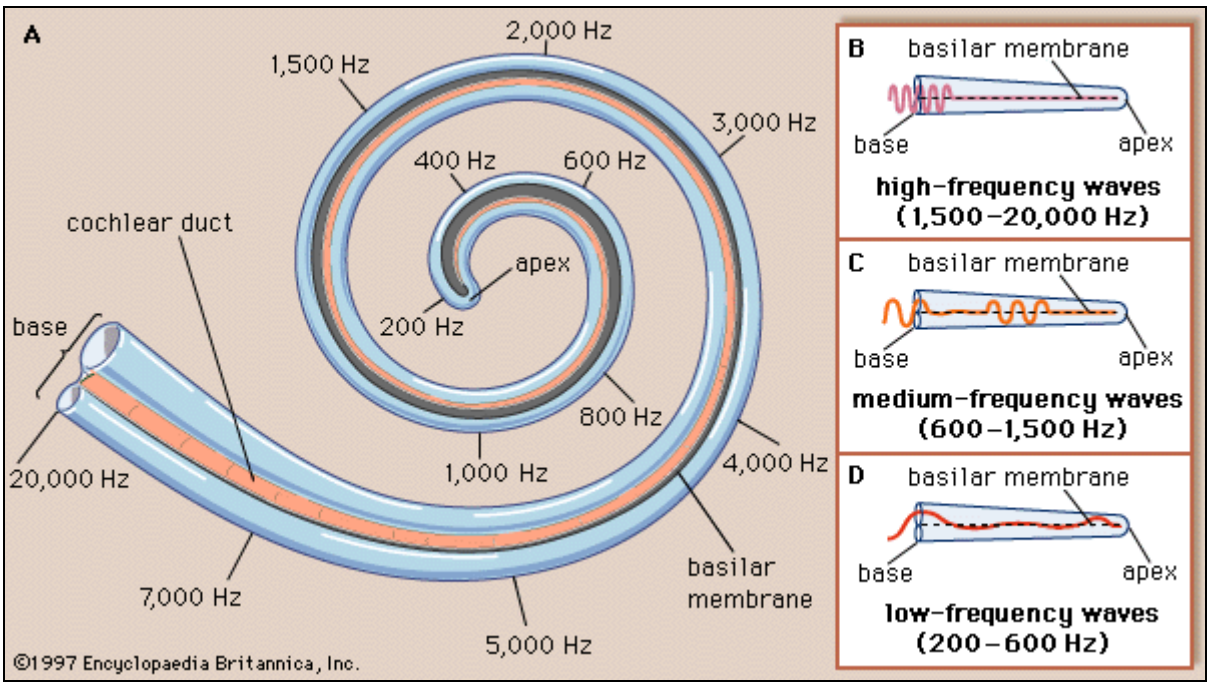

Figure showing how the cochlea makes a certain kind of Fourier analysis of complex ways by dividing them into their component parts along its axis. This model is common to both Helmholtz and Békésy models. On the right we can see the wave generated inside the cochlea, as predicted by Békésy's model and responsible for the decomposition (a better picture is given in our fig. 2, *infra*). Taken from site:
http://www.britannica.com/bcom/eb/article/8/0,5716,117508+21,00.html

---

A much more realistic model of the cochlea was developed by Georg von Békésy in the 1930s. Békésy showed that the presence of a fluid in the cochlea provoked the interdependence of movements throughout the basilar membrane. The main reason is that the fluid creates a kind of counter reaction to the initial force (made by the wave originated at the stapes) in the adjacent areas of the basilar membrane (BM). This drastically changes the wave pattern imprinted on the BM. Without the fluid the "local-force impulse would produce an instantaneous membrane acceleration" being the only resistance the "local mass". (see p.160, box 1)[2]. This would generate a movement similar to that depicted in the article (see fig. C, box 1, p. 160).

Békésy understood that the presence of a fluid would generate a counter resistance on the immediate sides of the original pressure (see fig. D) so that "any local basilar membrane

[2] Numbers and other references in brackets refer to the article under analysis except if otherwise noted.

oscillation generates forces that tend to drive flaking modules to swing with opposite phases." (160) Taking also into account the growing flexibility and more volume of the BM towards the apex, and the viscosity of the fluid we get a very different wave imparted in the BM (see fig. E, and also *infra*, fig. 2). This wave grows in amplitude and frequency until a peak is produced and then rapidly collapses. The length of this wave through the OC is inversely proportional to its frequency.

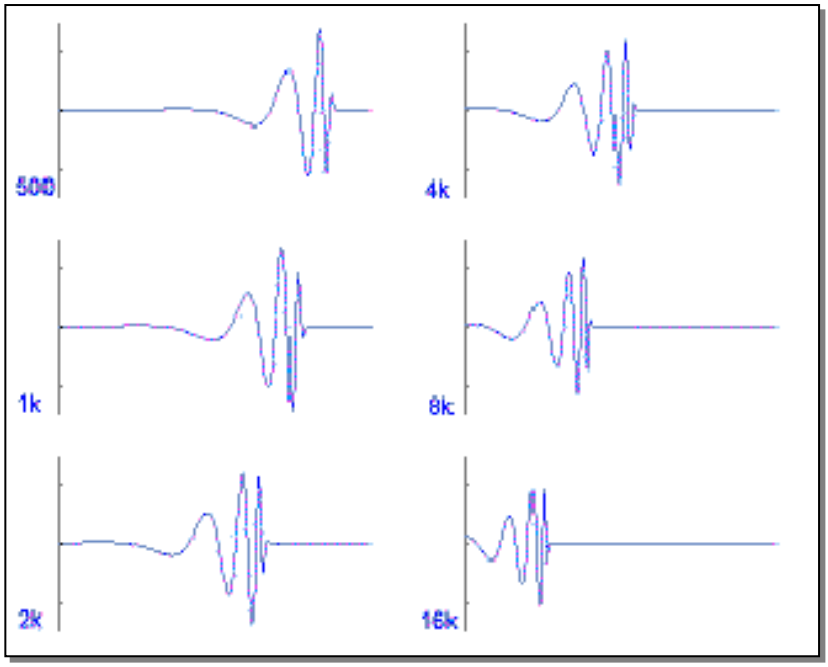

Fig. 2 The wave through the BM has its peak very close to the base in case of very high frequencies (16KHz) and close to the apex in lower frequencies (500Hz).
Taken from the Internet. At http://www.neurophys.wisc.edu/animations/

This relation between the frequency of the wave produced by the stapes and the location at which the wave at the basilar membrane has its peak permits to draw "a mechanical map that associated a 'characteristic frequency' with each site along the cochlea" (161). Since the inner hair cells of the Organ of Corti (OC) depolarise according to the movements of the BM, or the basilar membrane, it is easy to see that, according to Békésy's model, a particular inner cell will respond if and only if a particular frequency (or range of frequencies) is present above a certain threshold, which preserves Helmholtz idea of the cochlea as a frequency analyser: "If the cochlea operated linearly, it would perform like a bank of linear filters which separated the Fourier components in a sound, and travelling waves would be equivalent to linear combinations of harmonic oscillation modes as in Helmholtz's model." (164) [3]

There are several reasons that lead to the introduction of an active model of the cochlea, but the authors here focus only one: the "compressive non-linearity" or "saturation" of the cochlea response to sounds between 30 to 90 dB (see p. 162 and fig. 1-D). The reasons given in the article of Ruggero et al, and the enhancing capabilities ascribed to the active cochlea for

[3] We did not understood the following subsequent passage from this article: "Cochlear models must include the finding that instantaneous coupling among different basilar membrane portions has a long-range character that is only approximately represented by nearest-neighbour transmission-line interactions." (161) It seems that the authors are criticising a certain kind of model of the cochlea of the "transmission-line" type, but we do not know this model.

instance by Rosenzweig (1999, p. 226) or the existence of otoacoustic emissions[4], are not presented as main reasons to introduce the active model.

According to Nobili et al, this non-linearity can only be explained if, at low frequencies, "there are energy-dependent processes … that neutralize viscous damping … by positive mechanical feedback" (162). This demands a mechanically active cochlea that can amplify the wave through the basilar membrane. Of course, it is natural that at the measure that the wave present in the cochlea duct augments in amplitude, that is, with higher-volume sounds, the mechanical capabilities of the cochlea would loose some or all of their efficacy. So an active cochlea would explain the amplification phenomena but also the non-linearity that can be observed in higher sound pressures.

It is generally accepted that the only cells capable of doing this kind of mechanical 'undamping' are the outer hair cells (OHC) due to their motility, the number of efferent nerves, their seemingly lack of influence in the transmission of sound to the brain (the lack of inner cells is sufficient to cause deafness but the same is not true of OHC), and the fact that their ineffectiveness is sufficient to prevent non-linearity effects from appearing (see article by Ruggero, *infra*).

On this article the authors don't present these reasons[5] but they develop a detailed account of OHC motility and their connection with Deiter's cells. They rapidly expose the connections of OHCs with the basilar membrane (through Deiter's cells) and the tectorial membrane (162) but provide detailed functional and organic schemes of the way in which OHCs move the OC (and therefore also the BM).[6]

They continue by presenting possible ways of comprehending the motor that drives OHCs. It is know that (de)polarization of the OHC produces changes of about 4% in their length but the motor that produces this change is not well known even today.[7] The authors refer that OHC motor does not depend on ATP (unlike neuronal pumps) and cannot be an ion channel since it can "generate forces above 20kHz under the direct control of the cell transmembrane potential." (162) They also add that changes in OHC stiffness "depends primarily on the structure of the plasma membrane" (164).

A difficulty is that it seems that OHCs should not be able to respond to very high frequencies (like the '20kHz and above' the authors referred). On the one hand the "internal viscous forces" would increase leading to spending increasingly more energy to overcome the increasing inner resistance to movement. There is another reason that we did not understand very well, the authors say that "the OHC transducer current is shunted by the cell membrane capacitance" (164) this seems to mean that the time that an outer signal needs to pass the cell membrane to activate the motor of the OHC is longer than the time it would take to imprint movements of 20kHz, but we are not very sure that it is this that was meant by the authors. In any case they say afterwards, in the same page, that "the potential across the basolateral membrane of the OHCs might be sufficient to drive motility".

---

[4] The existence of otoacoustic emissions is of course referred by the authors in the end of the article but it is not presented as an argument to sustain the active character of the cochlea.

[5] With the exception of otoacoustic emissions that appear latter on but not as main reasons to accept OHC preponderance in the physical response of the cochlea.

[6] See their fig. 2-A, B & C. An animation is also present in the Internet that shows the movement of the OC caused by the OHC at the site http://www.neurophys.wisc.edu/animations/

[7] We have made a rapid search through PubMed and found out some recent articles that provide possible explanations for OHC motor. But it seems no final model was yet achieved.

Another point that seemed obscure to us is the way in which "viscous coupling [between OHCs and Deiter's cells] makes the force applied to the basilar membrane by the OHCs increase with frequency". (164). The same difficulty occurred regarding the undamping effect of viscous forces described in connection with the proportional displacement between tectorial membrane (radial displacement) and basilar membrane. The reason given by the authors is that "because viscous force, which affects the motion of the cochlear partition is proportional to velocity, but with opposite sign, it produces undamping" (164). It was not clear to us why should the viscous force augment with frequency in some cases (inside OHCs) and diminish in others.[8]

The authors also emphasize the fact that neurons (and inner cells at a particular location) are only fine tuned to a specific frequency when loudness is nearer the threshold level. At higher volumes a certain neuron responds to a wider set of frequencies. This was already discovered by Kiang in 1965 in an experience with cats. But the objective of the authors is to show that, because of this imprecision, to "filter the perceptually relevant features of sound, the cochlea exploits the saturation property of the OHC amplifier" (165). But we didn't understood how this was accomplished. The main problem of the tuning curve of inner cells is that, on normal hearing levels, it is difficult to decide if, for instance "the stimulus was a weak tone on 1200Hz or a stronger tone of 500 or 1800Hz or any frequency in between" (Rosenzweig 1999, 227). If this is the problem then equalization and tone-on-tone suppression do not seem to help. Although they show exemplary the uses in which the active OHC can be put (see also *infra* my final discussion at the end of the paper).

The last three topics – otoacoustic emissions, feedback control, and cochlear evolution – express somewhat lateral topics of which knowledge is somewhat scarce. The authors focus our ignorance on the way otoacoustic emissions are produced and suggest that its spectrum might be "dominated" by the middle ear. Regarding feedback control several pathways from the brain which control the OHC are identified. "Forward transduction" and the role of calcium ions is also treated. Regarding the last topic it is noted that in other vertebrates cochleas do not have OHC. Although the authors do not say it explicitly, these kinds of cochlea also show signs of active amplification and otoacoustical emissions. Because it is possible in other vertebrates that an active response can be achieved even without OHC, this opens at least the possibility that there may be other mechanisms involved in the physical manipulation of the stimulus at the cochlear level.

"Basilar membrane responses to two-tone and broadband stimuli"

In this article the authors present forceful reasons to think that some nonlinearities that occur in two-tone suppression or intermodulation (that results in a combination tone) are present in the movement of the basilar membrane of the cochlea. They also showed that the use of furosemide eliminates the nonlinearity effects, with all probability due to its impairment on the functioning of the OHCs.

Two techniques were used to capture the basilar membrane velocity, the emission of gamma photons in a source attached to the membrane and captured by a immobile detector, and, the second method, a laser was reflected on microbeads (10-30μm) also attached to the membrane. Because velocities in the emission of electromagnetic frequency alter the frequency of the radiation it is possible to know, studying the reflected radiation, which was the relative velocity of the emission (or reflector) of radiation relatively to the detector. This

[8] We should add that our 'maîtrise' was in philosophy. Mathematics and physics are not our strong points.

very complex apparatus was applied on the study of several anaesthetised chinchillas, in which the inner ear was revealed by cutting aside all its protector membranes. Expectedly, the results were not the same for all the chinchillas since some of them had better inner ears than others.

Regarding both two tone suppression and two tone distortion (introducing of a combined note) it was found that both these effects have mechanical counterparts in the vibration of the basilar membrane. Of course, because the authors could only study the vibration of a particular spot in the basilar membrane (in this case 3.5 mm from the base) the studies were therefore restricted to studying characteristic frequency (CF) effects and each cochlea had, depending on the exact location of the frequency emitter or reflector, different CFs.

More specifically, in relation to two-note suppression, the authors found, by studying the vibration velocity of the basilar membrane with a characteristic frequency (CF) of 8kHz, that the introduction of a second frequency could decrease the velocity of vibration of the 8kHz note. Regarding the rate of suppression, it was found to depend on the intensity and frequency of the first (non-suppressing) tone, and also on the intensity and frequency of the suppressing tone.

The suppression effect is bigger if the first tone has low levels, so that "the input-output functions are not simply shifted to higher probe levels, but are also linearized." (308). Which means, we think, that suppression effects disappear at high levels of probe tone input. The suppression effect is also bigger when the frequency of the probe note equals the CF (see the authors fig. 2). They also observed that "as the intensity of the suppressor note is raised, the magnitude of the mechanical suppression effect increases" (308) with regard to the 8kHz note. They also detected that the rate of suppression varies with the frequency of the suppressing note: frequencies higher than CF produce low suppressions "in the order of 0.4 dB dB $^{-1}$" but with frequencies "well bellow CF … the rate is substantially higher (between 0.65 and 1.42 dB dB $^{-1}$)." (308) In the example shown the authors have produced the suppression of the 8kHz tone by adding a 300Hz tone. In this particular example the effects of suppression are minimum when the tones are of equal intensity (70 dB) or even if the suppressing tone is much higher than the probe-tone but does not depass the 70dB range. (For instance when the probe-tone has 40dB and the suppressing tone 70dB, the suppressing effects are minimum.) The suppressing effects, however, rise increasingly when the suppressor tone augments its pitch. For 80, 90 or 100dB, the 300Hz suppressing note is capable of generating a strong suppressing effect on the 8kHz note.

Regarding two-tone distortion, the experience was made in a way that the two primary notes had frequencies such that the resulting combining note would have a frequency of CF, in this case 10kHz. The basilar membrane was then measured at this CF site and it was found that the 10kHz combining note appeared even when the primary notes were no longer visible at CF location. (This happened when the primary notes had frequencies below 80 dB – see their fig. 3.) Fig. 4 (309) depicts an experience made by the authors in which it can be seen that the distortion effects (i.e. the volume of the combined tone when compared to the volume of the primary ones) decrease when the primary tones increase in volume. It also is shown the effects of changing the relation between the ratio of the primary frequencies. Ratios of 1.1, 1.2 and 1.3 were tried. In general the ratio of 1.2 showed the best results (although it seems to us that, because $2f_1–f_2$ is always equal to 10kHz, it is not clear that the results obtained by the authors apply to any two frequencies of 1.2 ratio but to those particular primary frequencies – 12.5 and 15 kHz. So why did they not indicate the values of the primary frequencies instead of giving us their ratios?)

Until now we have studied just what Nobili et al (1997, 165, see *supra*) called the "tone-on-tone suppression" (and creation). We are now going to see the effects of the saturation effect that these other authors described in an experience with clicks. The point of this experience made by Ruggero et al. is to show that the responses to clicks of varying intensity do not grow linearly with the intensity of the stimuli. Showing precisely the kind of compressive non-linearity that we have talked about. Since the clicks do not have just a single intensity but each click produces, along 6ms, several signals of (same frequency but) different intensity (see Ruggero fig. 5) it is possible to see the linearity grow of the basilar membrane velocity at CF up until a certain threshold is achieved, after that the non-linear response becomes apparent. These results are systematised in figure 6, and the authors conclude that "the gain functions change systematically as a function of click level. Peak gains are largest … at the lowest stimulus intensities" (311). Also, at the measure that gains become less prominent there is also a "shift to lower frequencies" (311). The authors also show that the results achieved through clicks are comparable to those achieved with tones: "the CF specific compressive nonlinearities evident in responses to tones or clicks appear to be the same and do not preclude predicting the responses to one type of stimulus from these evoked by the other." (311)

After describing the types of nonlinearities that we have been discussing the authors analyse now their possible causes. The notorious fact is that nonlinearities disappear after death and also in "poor state experimental cochlea" (311). And this "suggests that their origin does not reside in the largely acellular … basilar membrane" (311). Although it is widely believed that the OHC are responsible for these effects the authors claim that almost all evidence for this is 'indirect' and the best to date is the otoacoustic emissions produced "by electrical stimulation of the medial efferent system" (311). The authors have thus made a further experience that can provide stronger evidence – "almost inescable imply" (312) – that OHCs are in fact the responsible for the non-passive character of the cochlea. Their experience consisted of injecting furosemide "that rapidly and reversibly reduces the hair cells receptor potentials" (312). After the injection of furosemide it was verified that nonlinearity CF specific effects disappeared. Since inner cells (which are also affected by furosemide) cannot account for physical movements in the OC, it follows that it must be OHC that produces the alterations in the OC responsible for the enhanced movements of the basilar membrane.

Discussion.

We did not understand several points relating to the mathematical formulations (Box 3 of first article), the role of viscosity in the undamping of very high frequencies and the latency of sound described in p.310 of the second article. We also found hard to understand several concepts and discussions because they are intricately mixed with past and large literature. So it seems highly likely that most difficulties we see in these articles are due to our own difficulties in apprehending fully what is at stake. Nevertheless we highlight the following points: Nobili et al have said that "To filter the perceptually relevant features of sound, the cochlea exploits the saturation property of the OHC amplifier, which produces two main effects: equalization of the response and tone-on-tone suppression." (165) We do not understand how these effects can help to filter the relevant features of sound. Linearization, by the contrary seems to blur the distinction between levels of sound. Although it can be useful at low levels we were unable to imagine how it can be used to filter sounds in normal volume conditions (in speech for instance). It seems better, at least in a first approach, to imagine the amplification feature as something design to improve the sensibility at low level sounds. Secondly, the article from Ruggero et al. has showed that, regarding two tone

suppression, the difference of the frequencies and the volumes that are necessary to create the suppression effect, are just to high to be of perceptual interest. In the retina for instance, the suppression of flanking signals allows for the easy recognisance of lines. But how does a 300Hz suppressing a 10kHz tone, with the 300Hz showing 80dB at least volume, enhances sound perception. It seems easier to see this particular effect as an involuntary effect with no perceptual advantages, although it might be a very powerful proof of the active nature of the cochlea.

A parallel case happens in two-tone distortion. Although Tartini has appreciated its effect, and it has been widely used in music, it is hard to see its (evolutionary) advantages for perception. On the other hand, if we see it like an unwanted effect of the (active) mechanics of the cochlea, it is difficult to understand why it is present in so different species. We should also had that, if these distortions produced by the active cochlea are not as useful to frequency detection as it is necessary to explain perceptual distinction of frequency, this might reinforce the role attributed to the time structure regularities of sounds.[9]

We also would like to hear more about the hair cells presented in the vestibular channel. It is clear that the active effects of the cochlea show at the basilar membrane, but that does not necessarily imply that they must be provoked by the OC. Although it seems very clear that the OHCs play an all important role in the movements of the basilar membrane, we would like at least to know that the hair cells present in the vestibular channel do not possess motility (we were unable to find this information). If even cochleae devoid of OHCs can be active, it seems at least possible that other mechanisms might induce physical interactions inside the cochlea.

Regarding possible developments in this theme it seems clear that there are two difficulties that need to be explained. How does OHC motor works, and what are the principles that provoke contraction or dilation of the OHCs. We have read an abstract of an article in which it was proposed that the OHCs made an effort inversely proportional to the direction of the tectorial membrane. In this model this could account for the specific distortion effects found in the cochlea.

[9] For a forceful defence of the so-called volley theory see for instance Griffiths, 1998, "Analysis of temporal structure in sound by the human brain", in *Nature Neuroscience*, 1(5): 422-427. See specially fig. I which shows how very different representations at the cochlea level can be perceived as the same pitch.