Cochlear Signal Processing: Modelling and Integration with Machine Learning

Prof. Eliathamby Ambikairajah
School of Electrical Engineering and Telecommunications
University of New South Wales (UNSW), Sydney, Australia

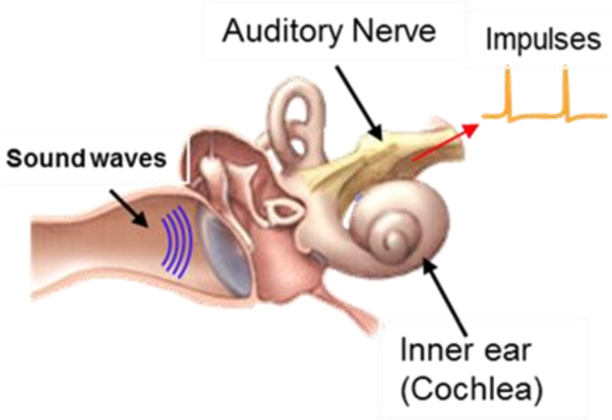
IEEE Sri Lanka Section Signal Processing Society Chapter

03 January 2022

UNSW SYDNEY

Contents

"How
are
you?"



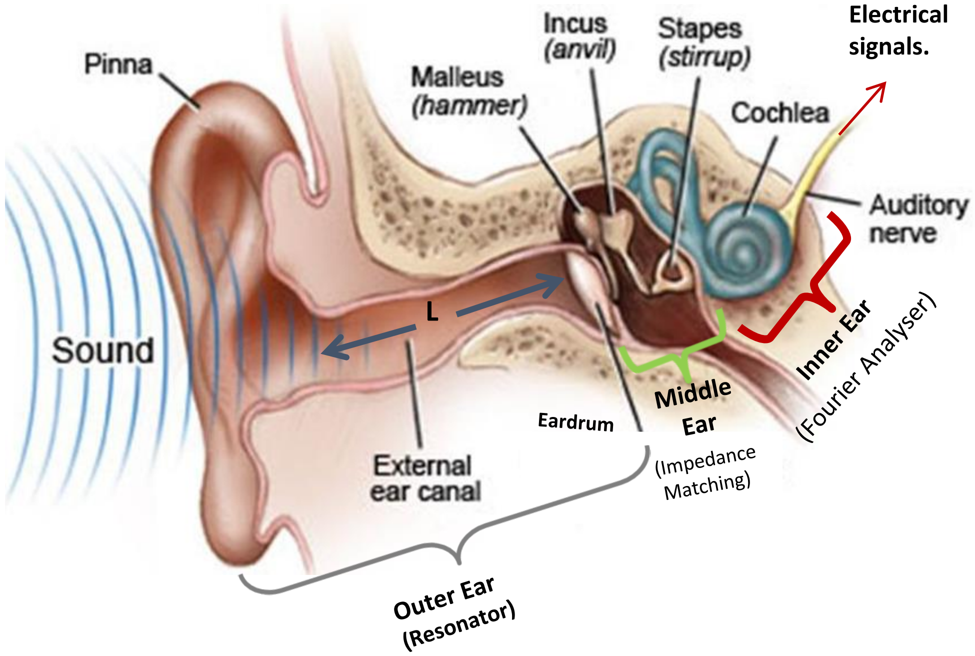
Auditory Nerve
Impulses
Sound waves
Inner ear
(Cochlea)

Introduction to the Human Auditory System

- ✓ The peripheral auditory system is divided into the Outer Ear, Middle Ear, and Inner Ear.
- ✓ The human ear can respond to minute pressure variations in the air if they are in the audible frequency range, roughly 20 Hz - 20 kHz

Sounds	Level
Faint	20dB (A faint Whisper is 30dB)
Soft (Quiet)	40dB
Moderate	60dB (normal conversation)
Loud	80dB (alarm clocks, vacuum cleaners)
Very Loud	90dB(Blenders);110dB (Concerts, car horns)
Uncomfortable	120dB (jet planes during take off)
Painful and dangerous	130dB(Jackhammers); 140dB(Gunshots) *Use hearing protection

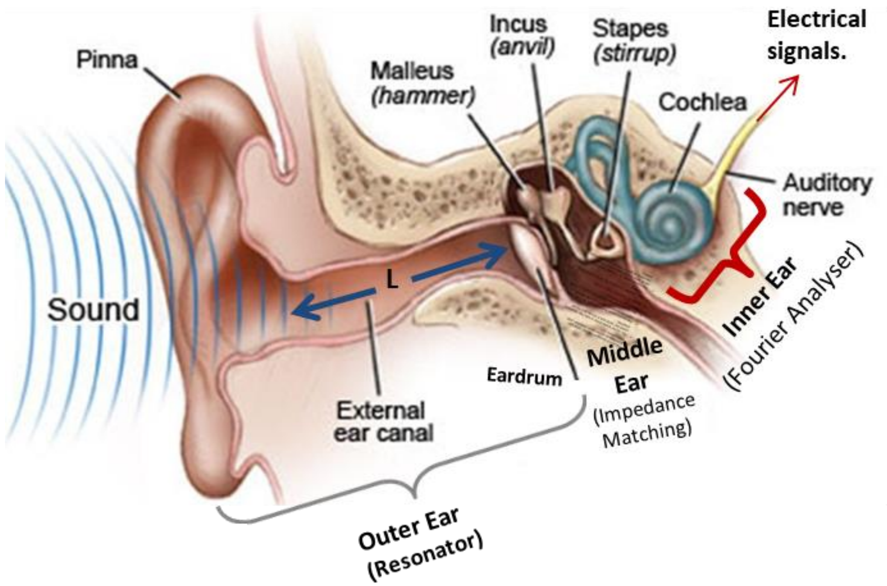
- ✓ Over 85 dB for extended periods can cause permanent hearing loss
- ✓ Zero decibels (0 dB) represent the absolute threshold of human hearing, below which we cannot hear a sound.

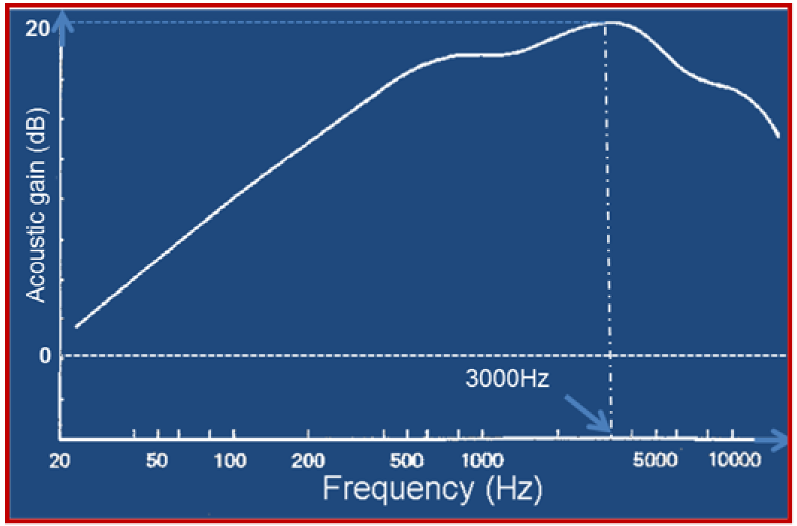


The peripheral auditory system and in particular the cochlea can be viewed as a real-time spectrum analyser.

Outer Ear (Air Vibration): A resonator

- ✓ The pinna surround the ear canal and functions as sound wave reflectors and attenuators .
- ✓ The sound waves enter a tube-like structure called ear canal and it serves as a sound amplifier.
- ✓ The sound waves travel through the canal and reach the eardrum and cause it to vibrate
- ✓ The length (L) of the human ear canal is 2.8 cm (and 7 mm in diameter)
- ✓ Speed of sound (c) = 340.3 m/sec ;
- ✓ The resonant frequency (f) of the canal is = $\frac{c}{4L}$ = 3,038Hz.
- ✓ The human outer ear is most sensitive at about 3kHz and provides about 20dB (decibels) of gain to the eardrum at around 3000Hz.

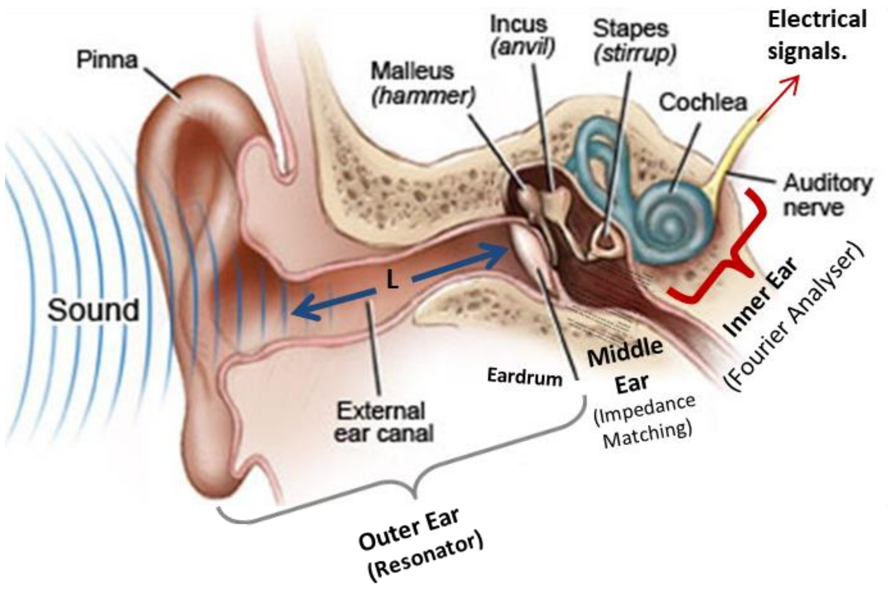


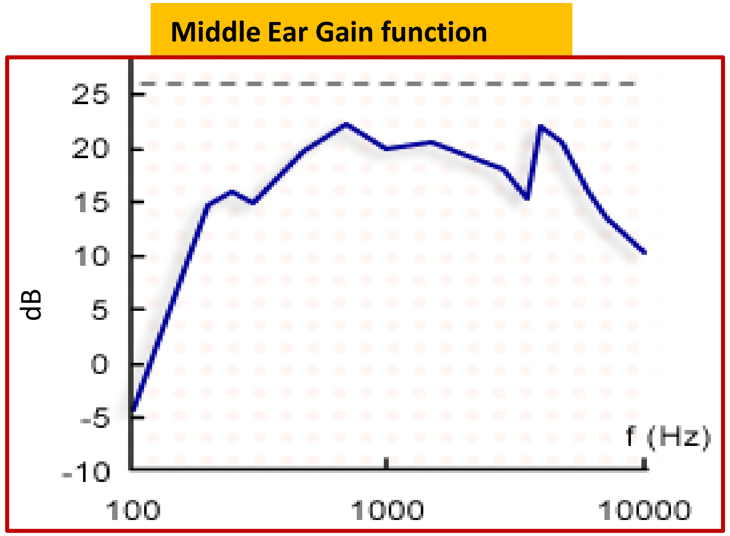


Outer ear is a low-Q bandpass filter
(Representative figure only)

Middle Ear: An Impedance Matcher & an Amplifier

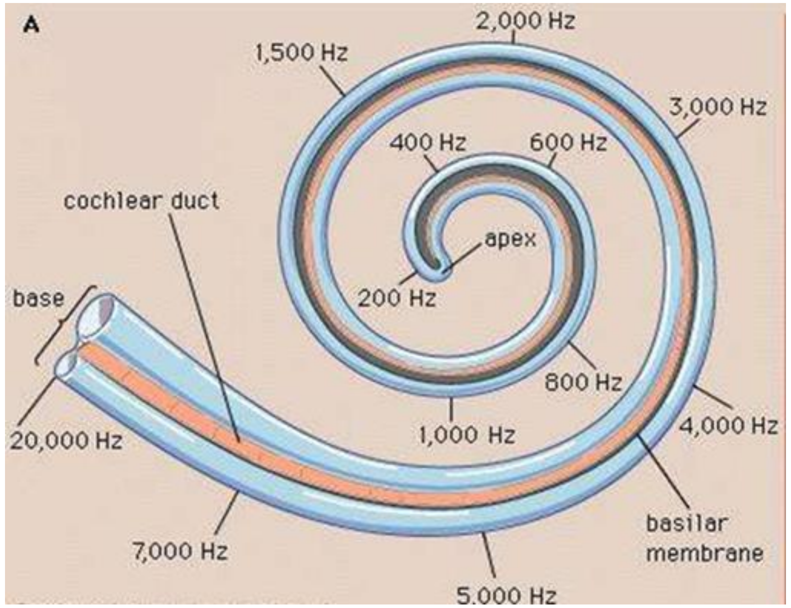
- ✓ Middle ear transforms the vibrating motion of the eardrum into motion of the stapes via the two tiny bones, the malleus and incus .
- ✓ The pressure of the sound waves on the oval window is around 25 times higher than on the eardrum.
- ✓ Since the sound Intensity (I) is proportional ($\propto$) to the square pressure (P^2) , the sound intensity increases 625 times (or 28dB)
- ✓ Middle ear converts acoustic energy to mechanical energy and mechanical energy to hydraulic energy

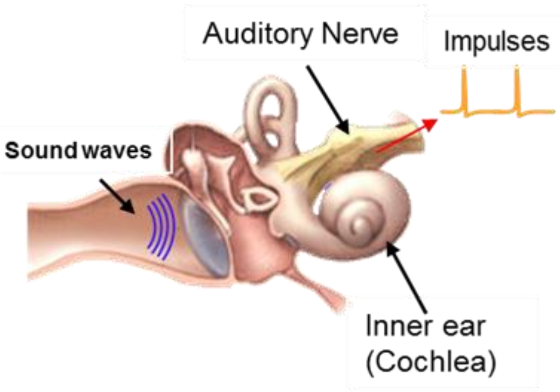


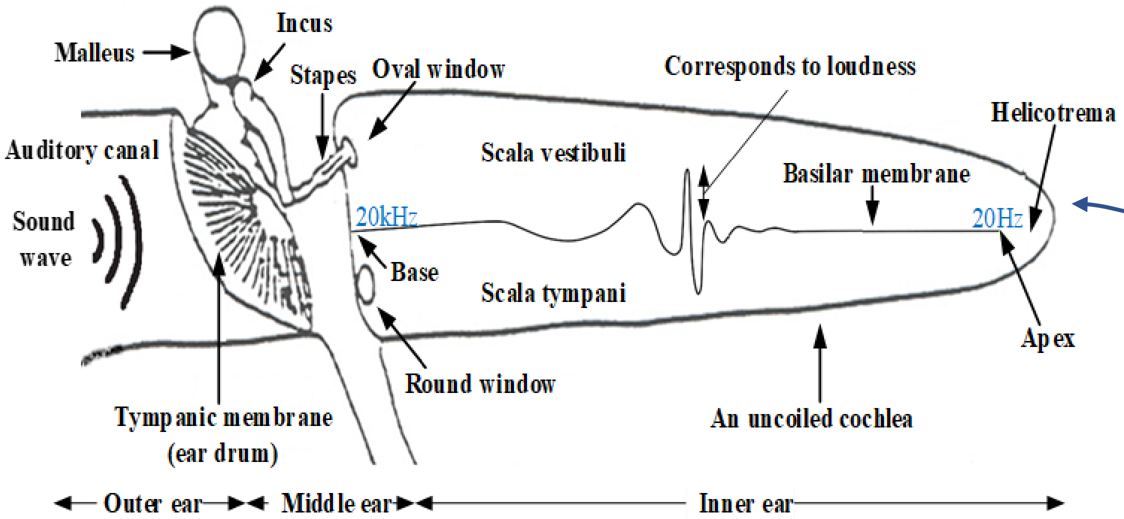


Inner Ear

- ✓ The **inner ear** consists of the cochlea responsible for converting the vibrations of sound waves into electrochemical impulses which are passed on to the brain via the auditory nerve.
- ✓ The cochlea is a spiral shaped structure which is about 3.5 cm in length if uncoiled.
- ✓ The cochlea is divided along its length by the basilar membrane (BM) which partitions the cochlear into two fluid canals (scala vestibuli and scala tympani).



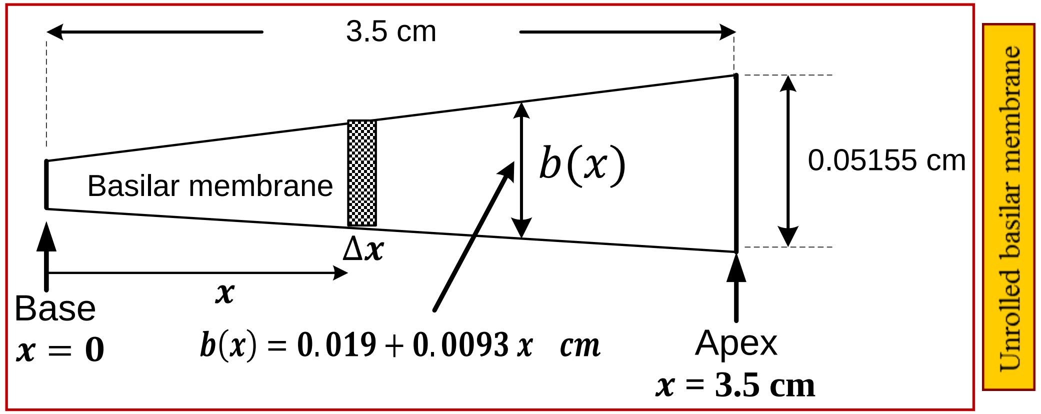




A longitudinal section of an uncoiled cochlea

Basilar Membrane (Hydro Dynamical process)

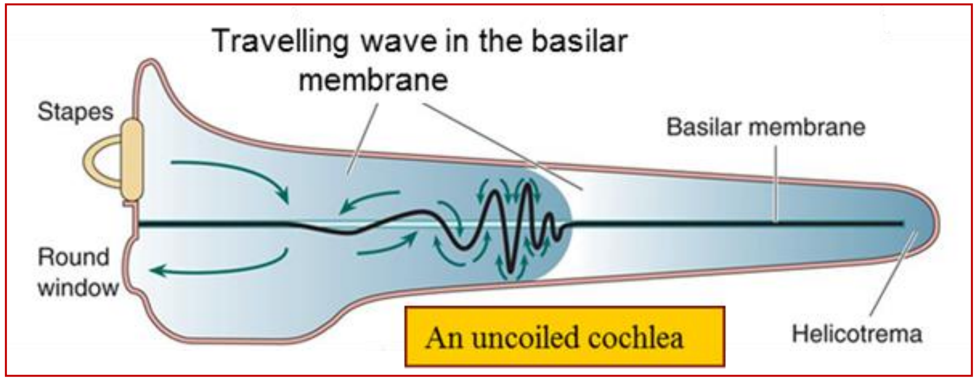
- ✓ The Basilar Membrane varies in width and stiffness along its length.
- ✓ At basal end it is narrow and stiff where as towards the apex it is wider and more flexible.
- ✓ Each point along the basilar membrane has a characteristic frequency, $f_p(x)$, to which it is most responsive.
- ✓ The maximum membrane displacement occurring at the basal end for high frequencies (20 kHz) and at the apical end for low frequencies (70Hz) .



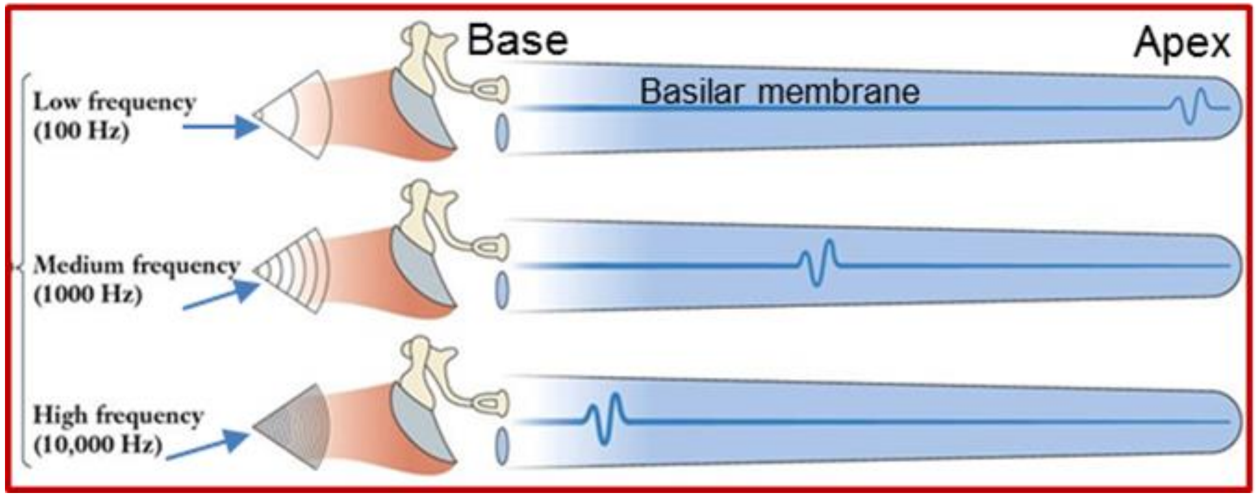
If x is the distance of a point on the basilar membrane from the stapes, then the frequency, $f_p(x)$, that produces a peak at this point is given by:

$$f_p(x) = (20000.0)\ 10^{-0.667\,x}\ Hz \qquad 0 \leq x \leq 3.5\ cm$$

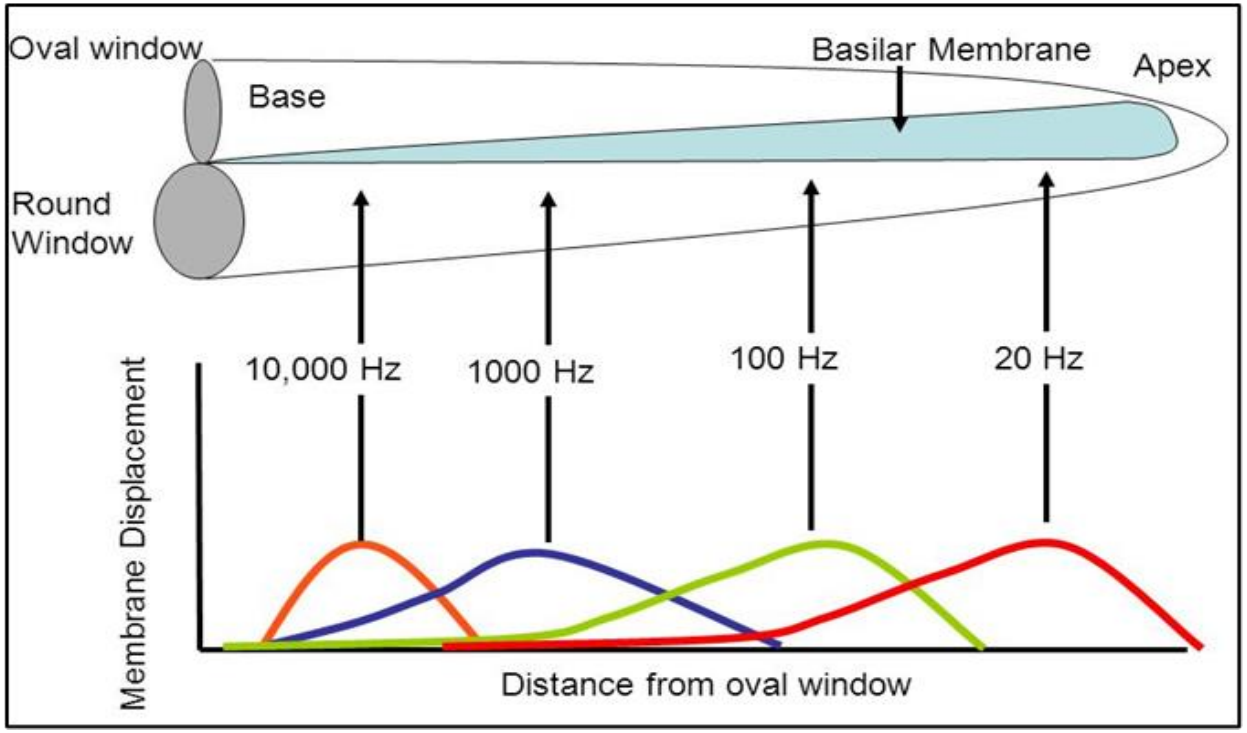
- It is evident that a 20 kHz tone at the stapes will cause the BM to vibrate at a point $x = 0$.
- A 70 Hz tone will excite the BM at a point x = 3.5 cm (i.e. at the apex)



Basilar Membrane as a Filterbank



- ✓ Different frequencies stimulate different areas of the basilar membrane
- ✓ There will be one place where the resonant frequency of the membrane matches the stimulus frequency and this place will show the maximum amount of vibration



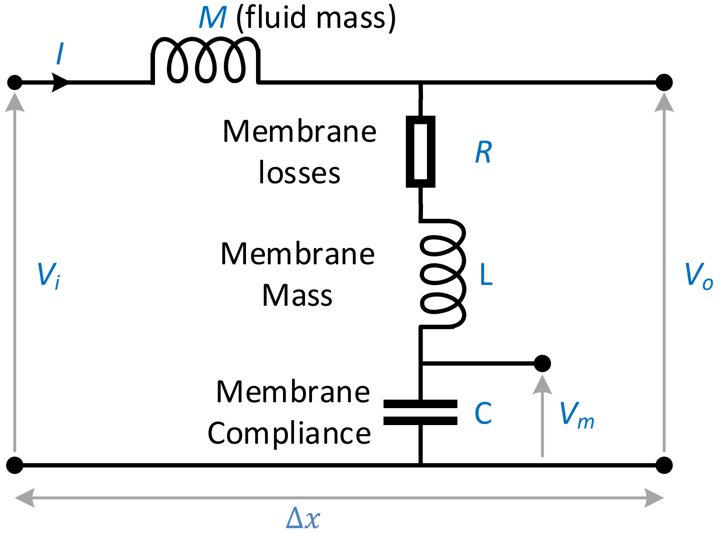
- ✓ The essential function of the basilar membrane is to act as a frequency analyser (a set of band-pass filters each responding to a different frequency region) resolving an input sound at the eardrum into its constituent frequencies

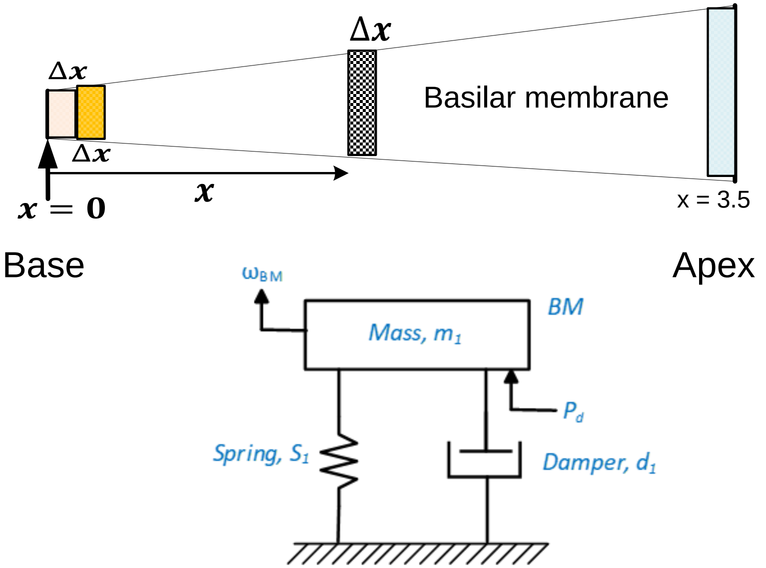
Cochlear Animation

Models of the Cochlea

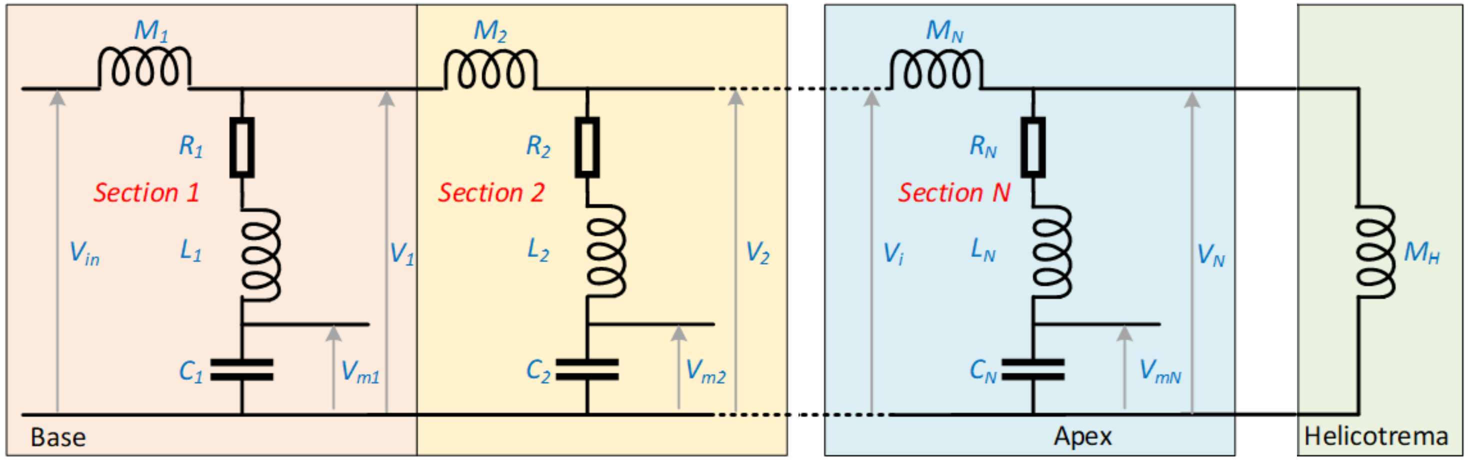
Cochlear Modelling

- ✓ A simple electrical model of a section (Δx) of the Basilar Membrane

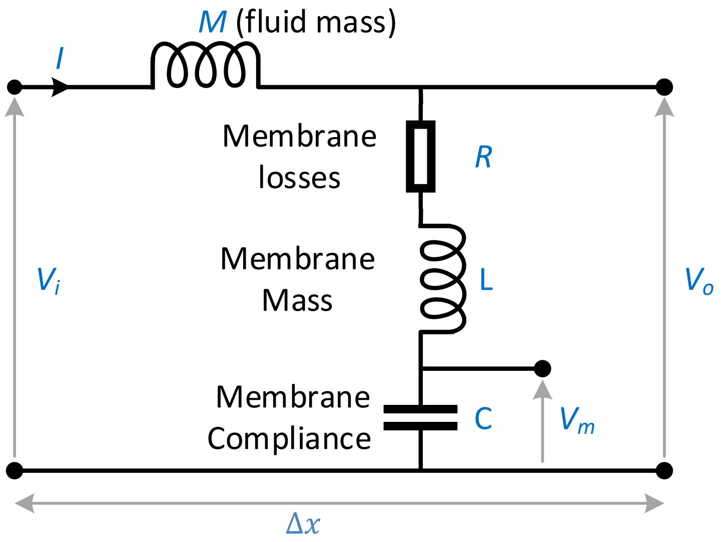




- ✓ A Transmission Line Model of the Basilar Membrane



Transmission Line Model of the Cochlea – Cascade Model



Pressure Transfer Function:

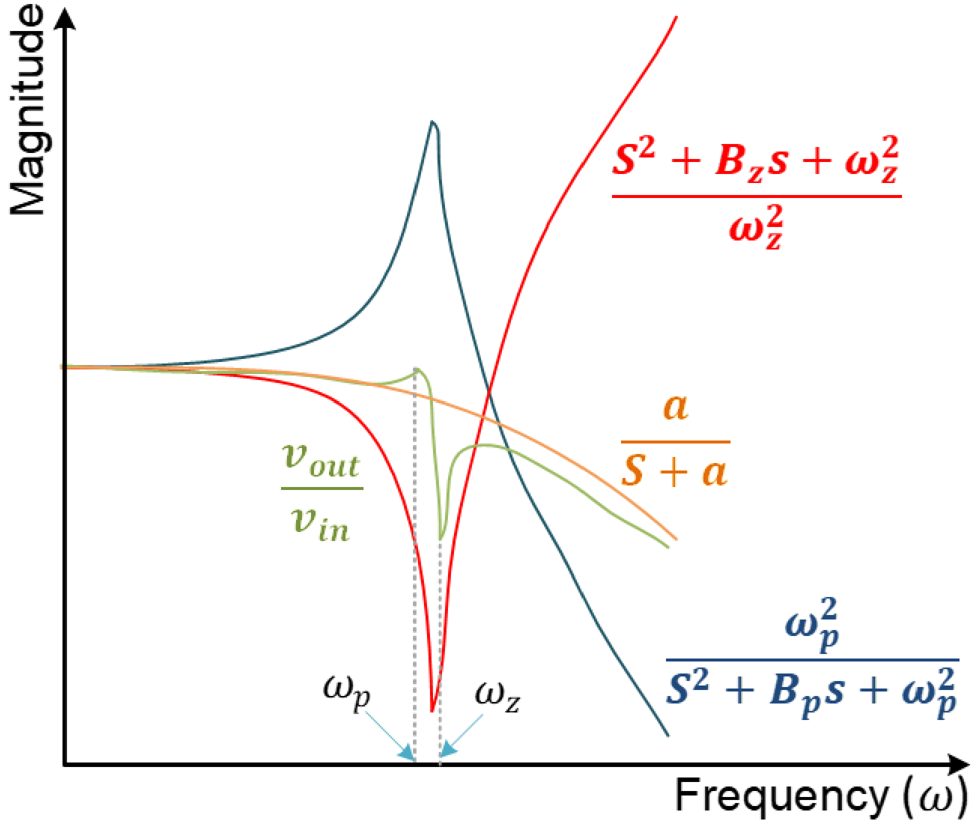
$$\frac{V_o(s)}{V_i(s)} = K\,\frac{a}{s+a}\,\frac{\omega_p^2}{s^2+B_p s+\omega_p^2}\,\frac{s^2+B_z s+\omega_z^2}{\omega_z^2}$$

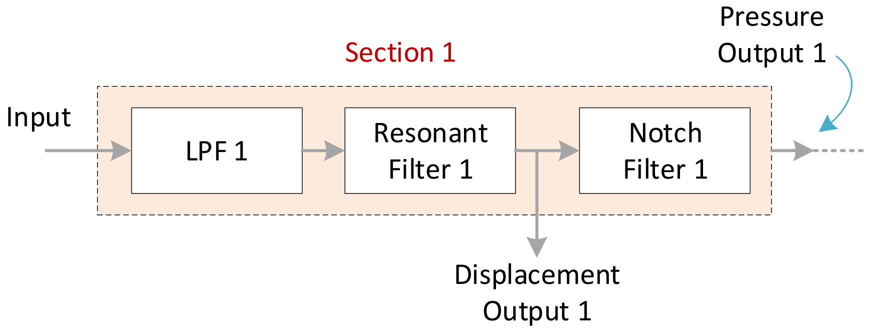
Low pass filter **Resonant pole** **Resonant zero**

Displacement Transfer Function:

$$\frac{V_m(s)}{V_i(s)} = K\,\frac{a}{s+a}\,\frac{\omega_p^2}{s^2+B_p s+\omega_p^2}$$

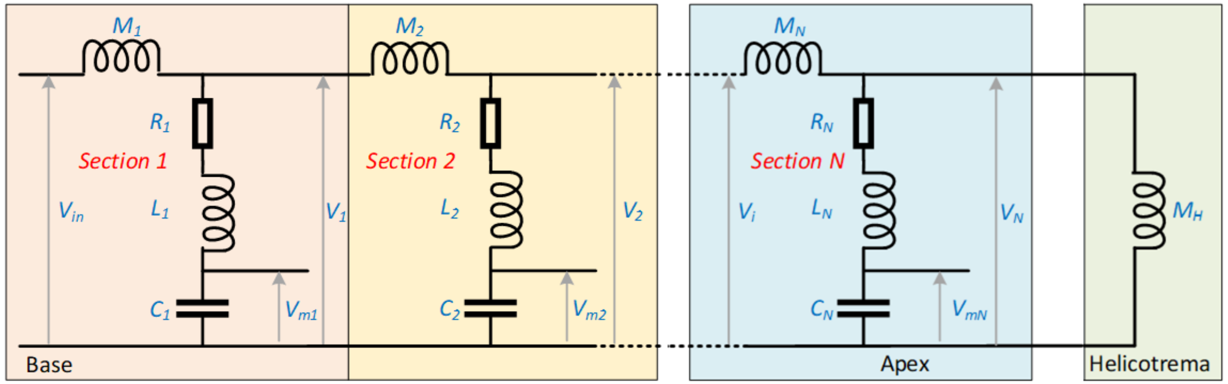
Low pass filter **Resonant pole**

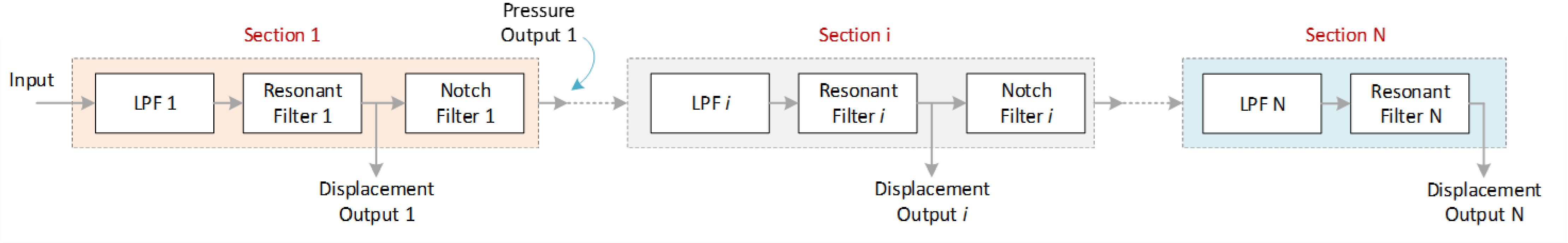




Transmission Line Model of the Cochlea – Cascade Model

- ✓ The basic model of the cochlea is a transmission line model in which each section of the basilar membrane is modelled as a cascade of low pass, notch filters and resonators.



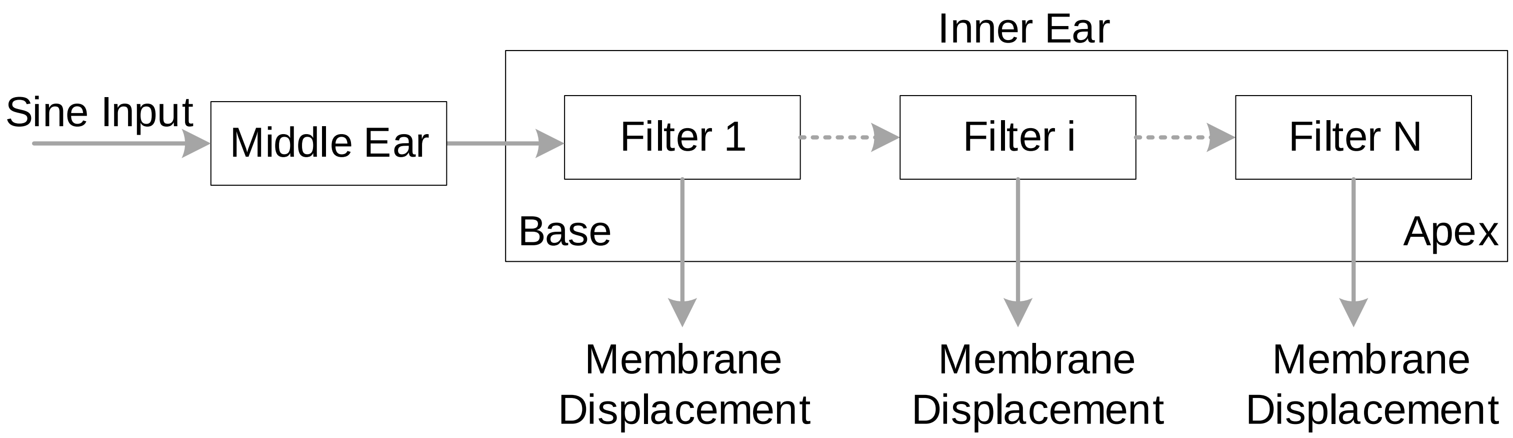


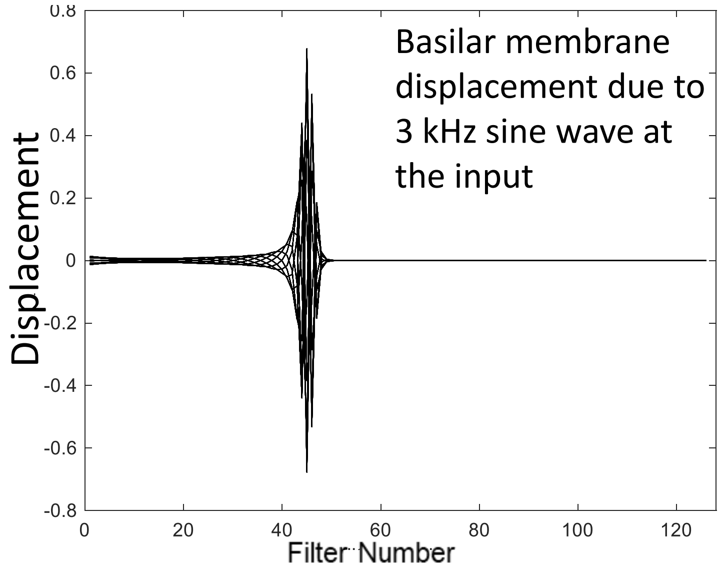
$$\frac{v_o(z)}{v_i(z)} = K \; \frac{1-a_0}{1-a_0 z^{-1}} \; \frac{1-b_1+b_2}{1-b_1 z^{-1}+b_2 z^{-2}} \; \frac{1-a_1 z^{-1}+a_2 z^{-2}}{1-a_1+a_2}.$$

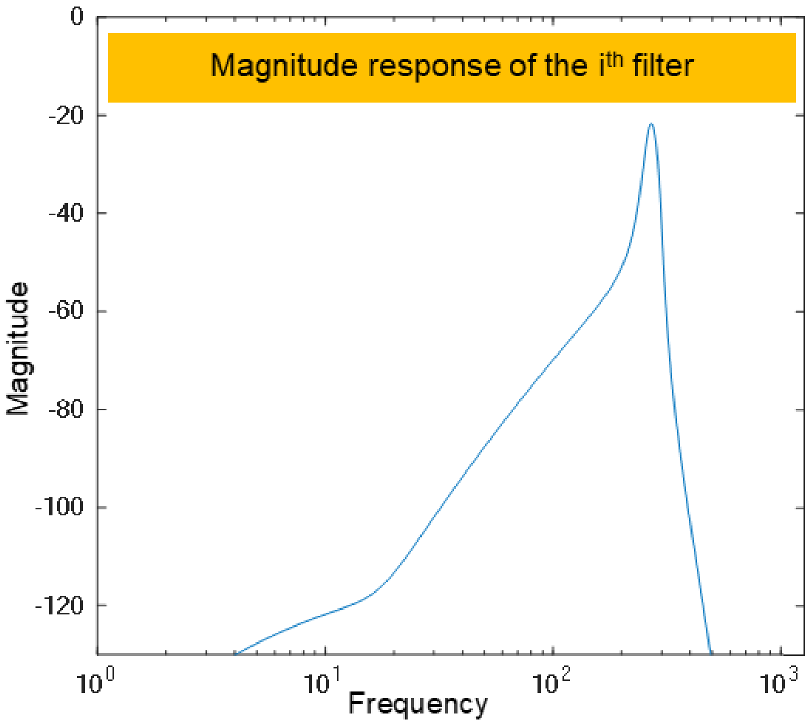
$$\frac{v_m(z)}{v_i(z)} = K \; \frac{1-a_0}{1-a_0 z^{-1}} \; \frac{(1-b_1+b_2) z^{-1}}{1-b_1 z^{-1}+b_2 z^{-2}},$$

Digital Filter Implementation

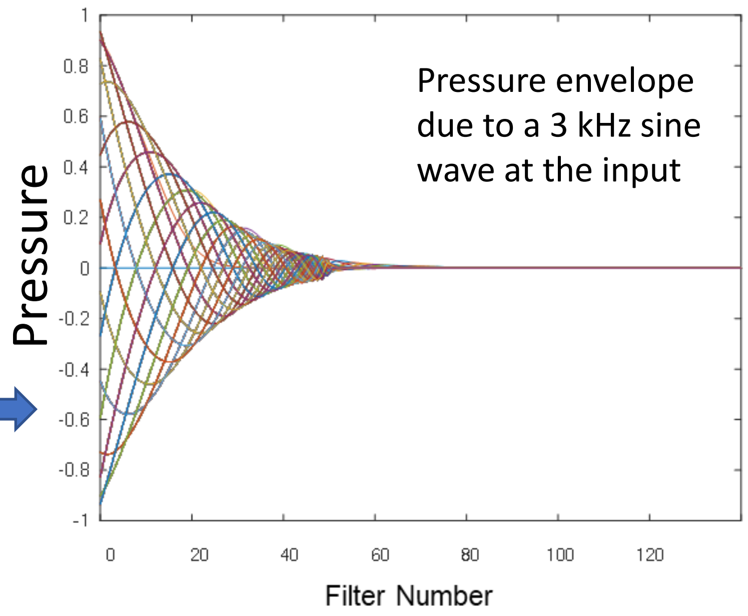
Membrane displacement and Pressure envelope for a sinusoidal input



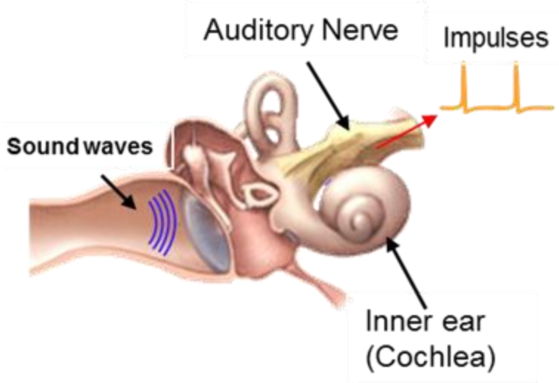




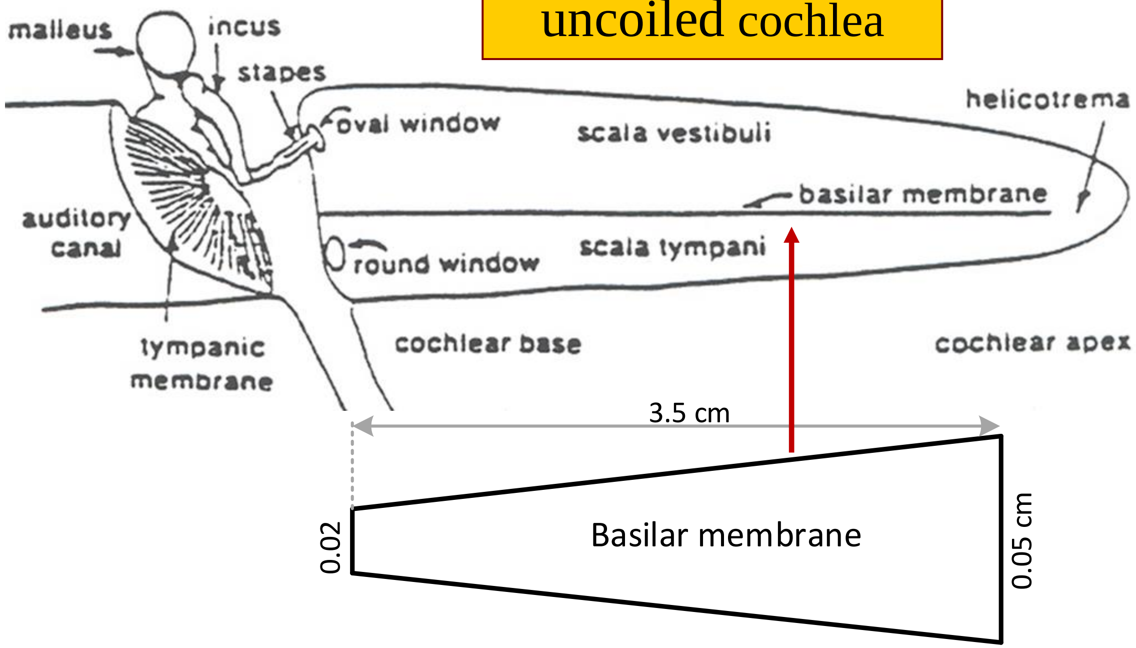
- ✓ Observing the envelope, the pressure is high at the basal end and decays down to zero at the resonant position.

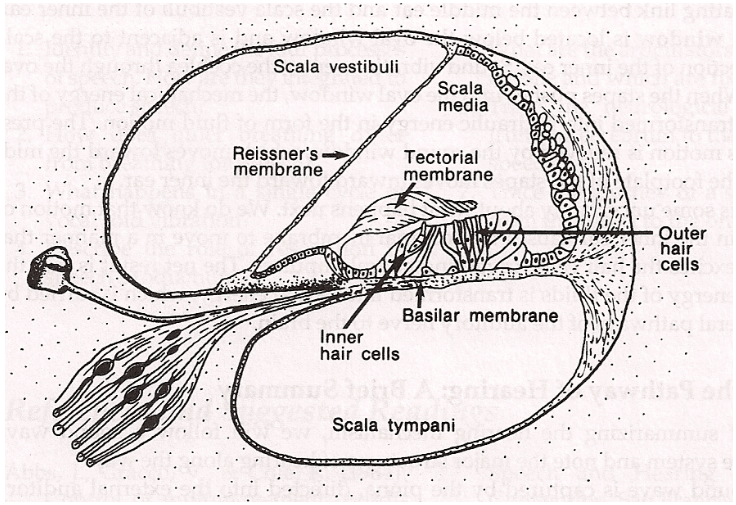


Organ of Corti

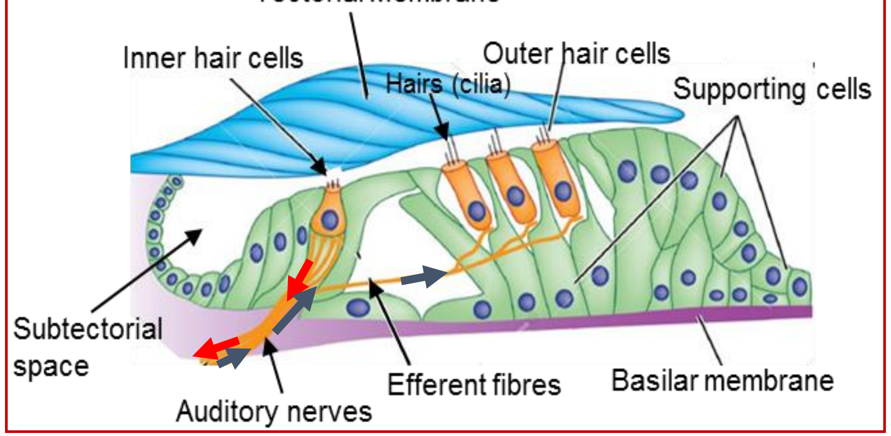


uncoiled cochlea



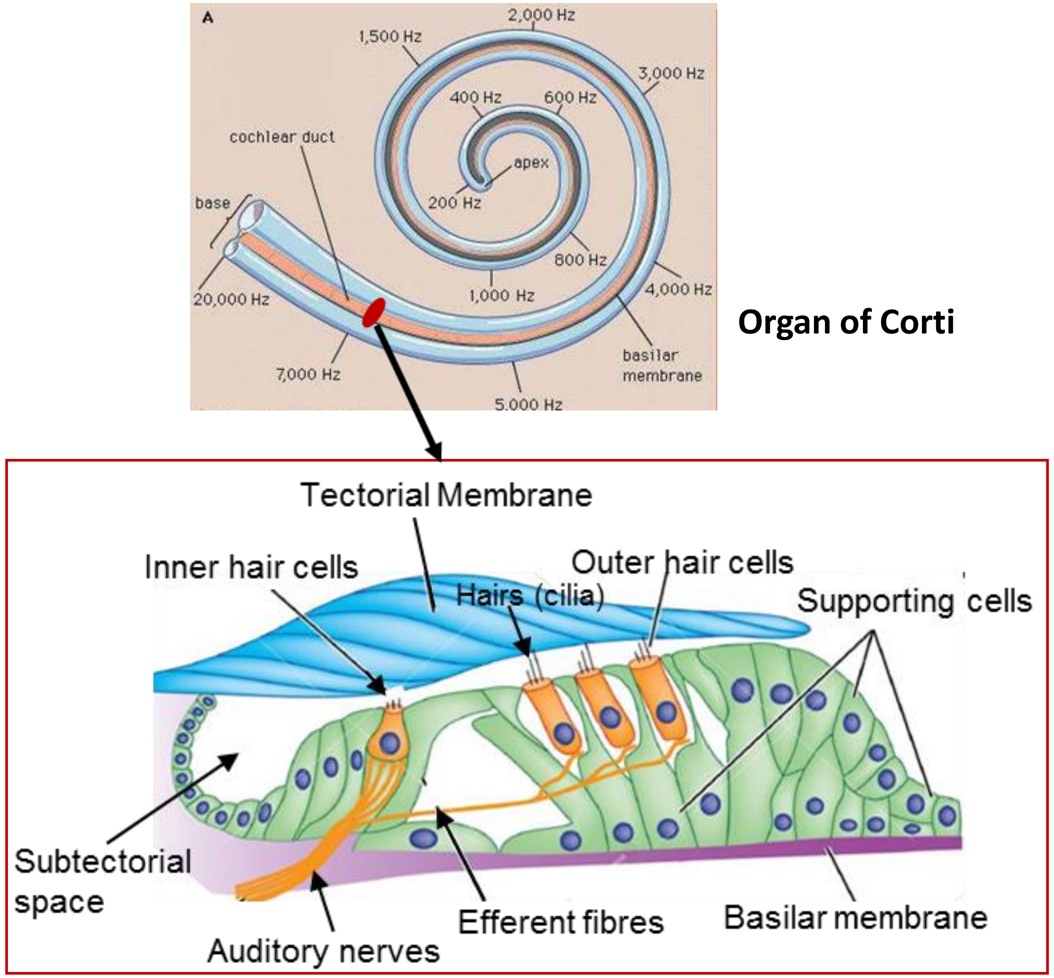


A cross section of Organ of Corti

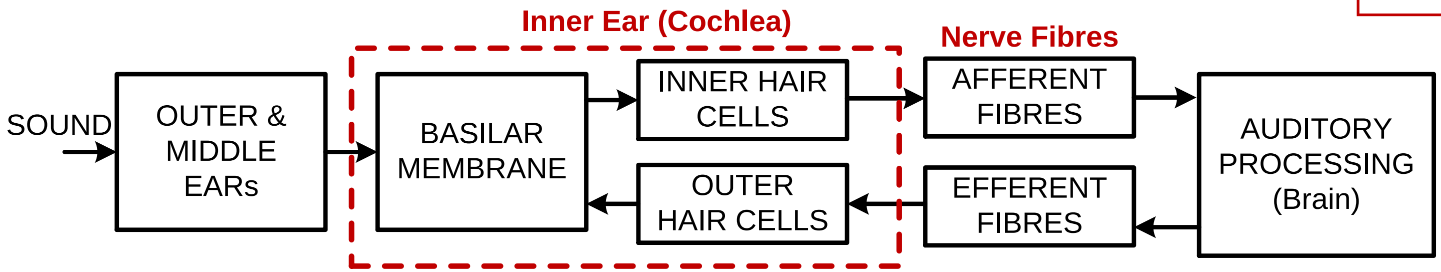


Organ of Corti

- ✓ Attached to the basilar membrane and running its entire length is the organ of corti containing some 30,000 sensory hair cells.
- ✓ There are two types of sensory hair cells:
  - One row of inner hair cells, whose cilia float freely in the fluid-filled region called subtectorial space
  - Three rows of outer hair cells whose cilia are attached to the tectorial membrane
- ✓ When the basilar membrane deflects, due to pressure wave in the cochlear fluid it triggers the inner hair cells to transmit nerve impulses to brain.

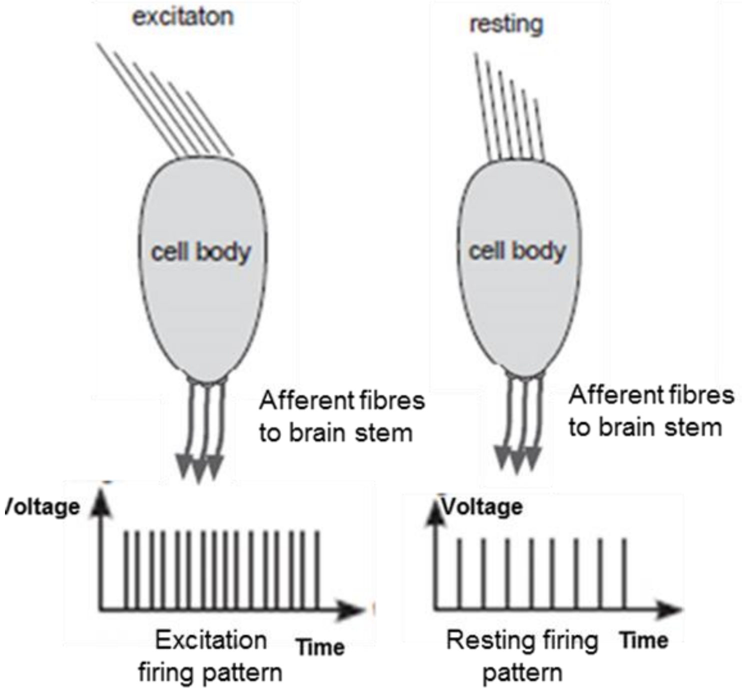


Organ of Corti

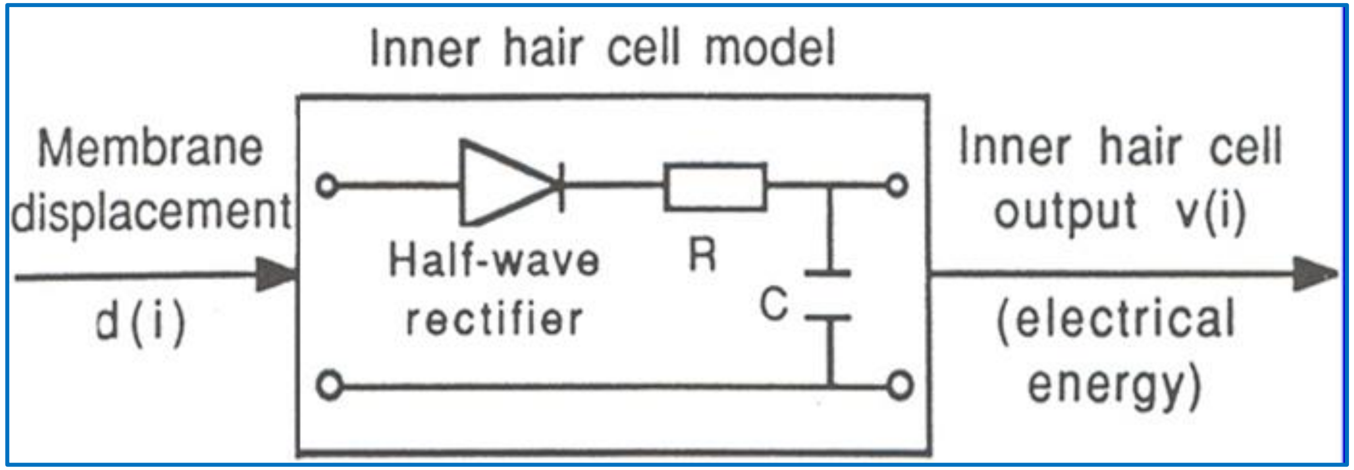


A simplified Diagram of a Human Auditory System

Mechanical to Neural Transduction (Electro Chemical)



- ✓ The mechanical displacement to electrical energy transduction process takes place in the inner hair cells
- ✓ Bending of the inner hair cell cilia due to basilar membrane displacement produces a change in the overall resistance (reduces it) of the inner hair cell, thus modulating current flow through the hair cell.

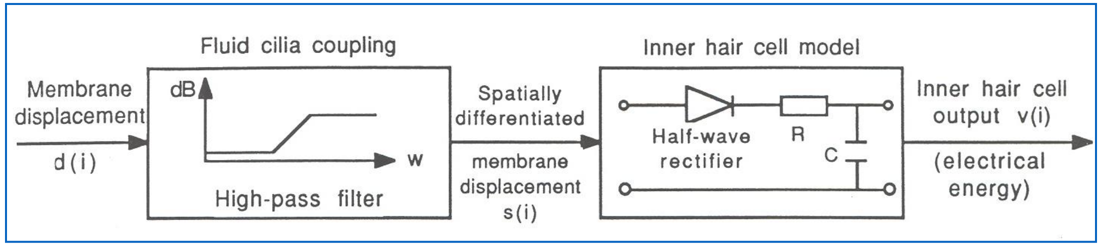


Inner Hair Cell model

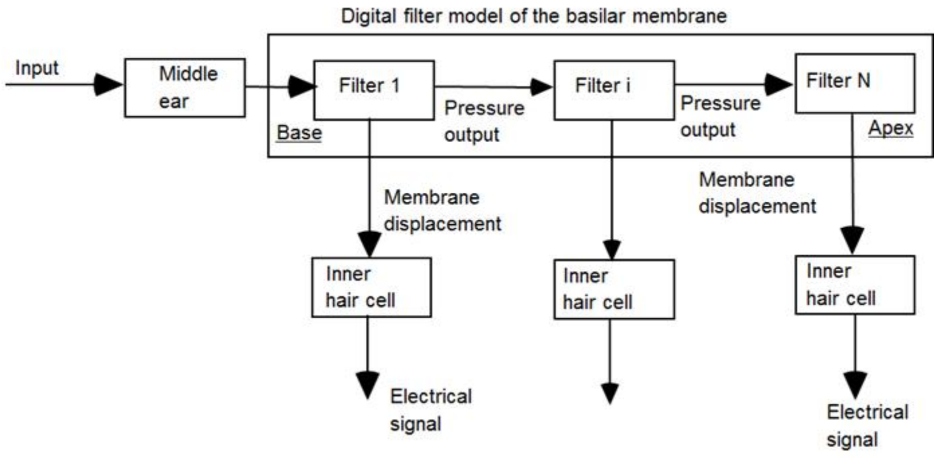
Here bending the inner hair cell cilia is simulated by charging of the capacitor and returning to the initial position of the cilia is equivalent to discharging the capacitor.

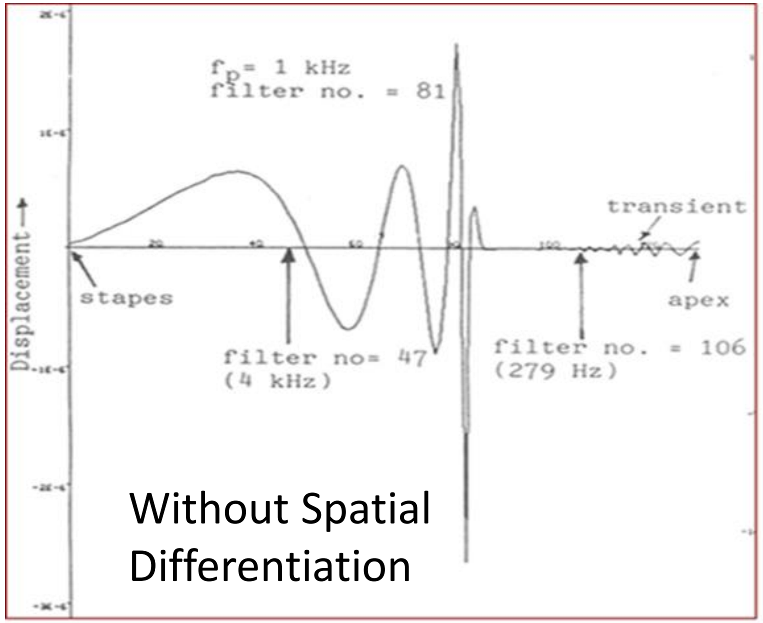
Mechanical to Neural Transduction (Electro Chemical)

- ✓ There is coupling between the cilia of the inner hair cells , through the fluid in the subtectorial space.
- ✓ Spatial differentiation of the membrane displacement represents this coupling

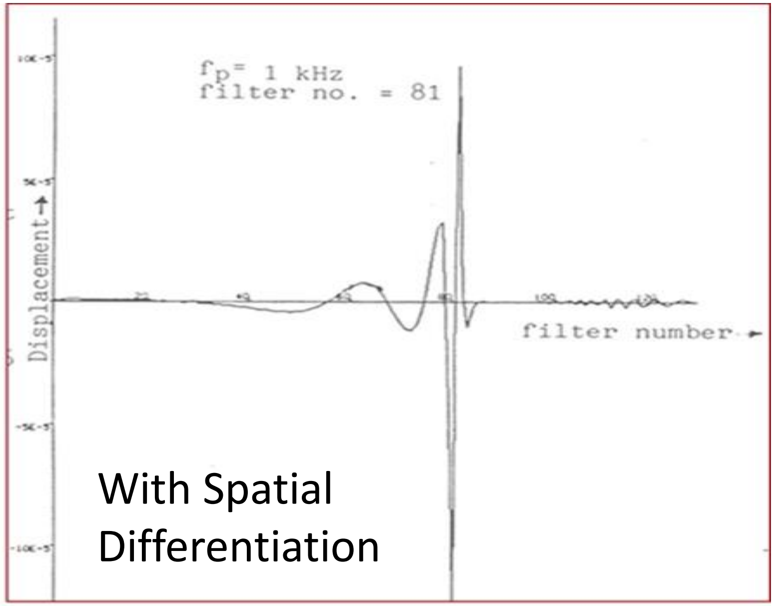


Transmission Line Model





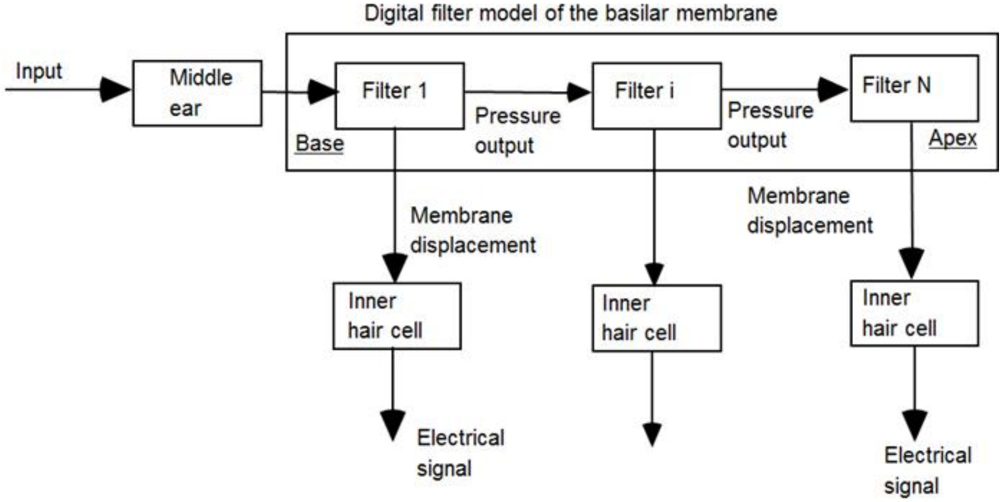
Without Spatial Differentiation

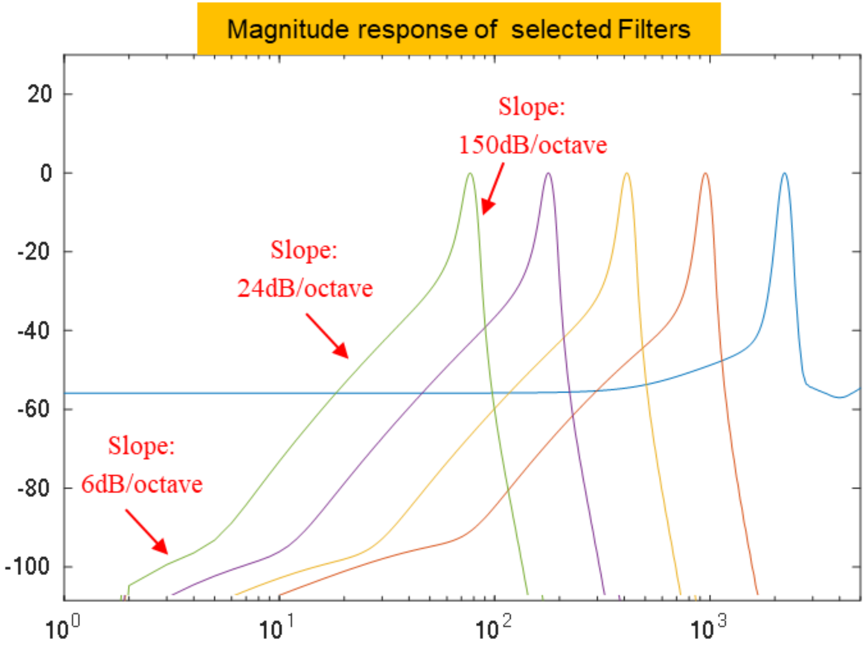


With Spatial Differentiation

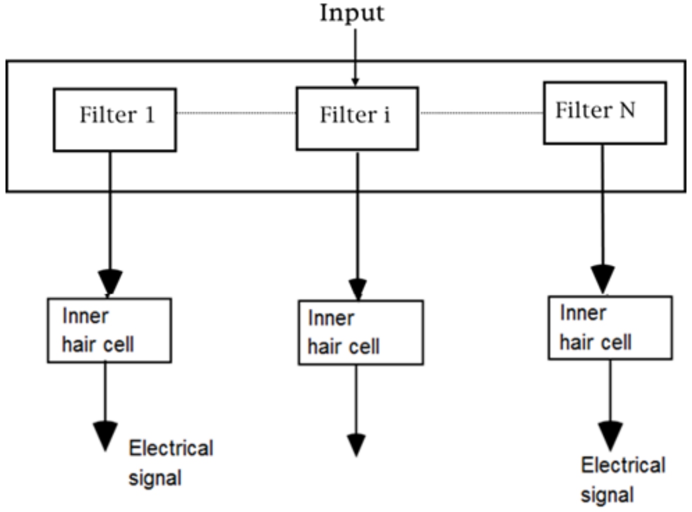
Cochlear Modelling: Cascade and Parallel Models

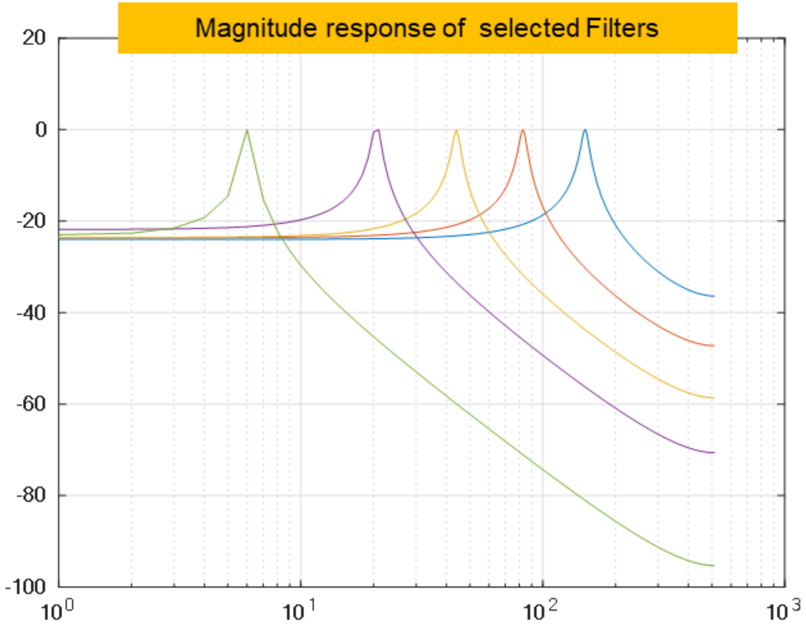
Transmission Line Model



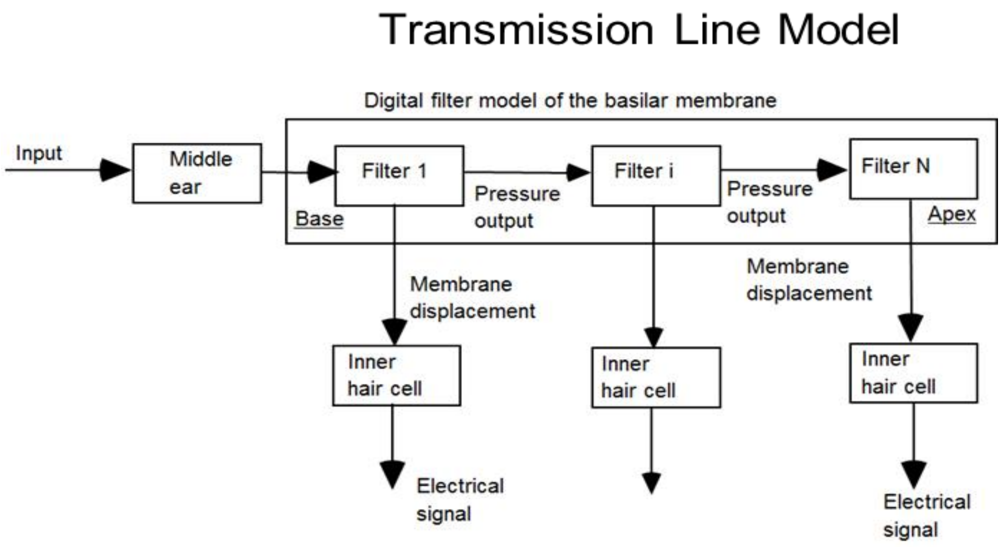


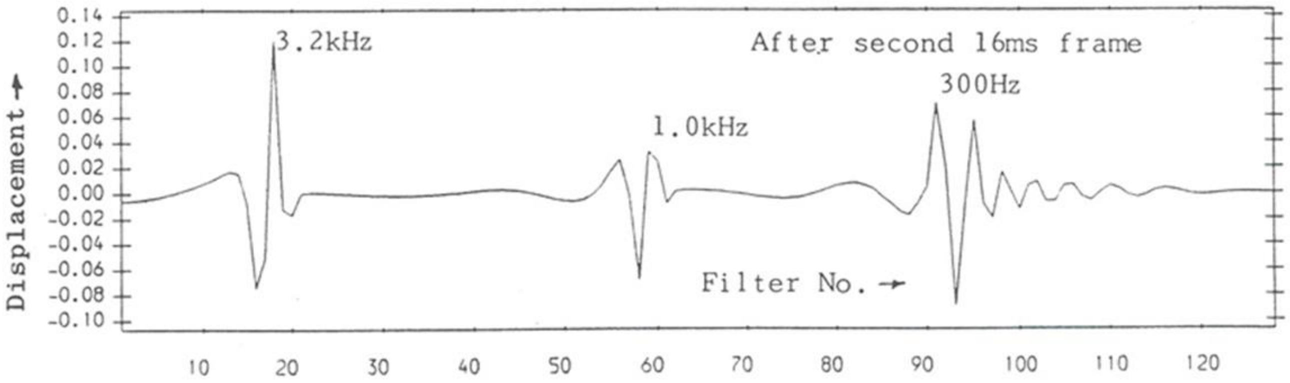
Parallel Filter Bank Model

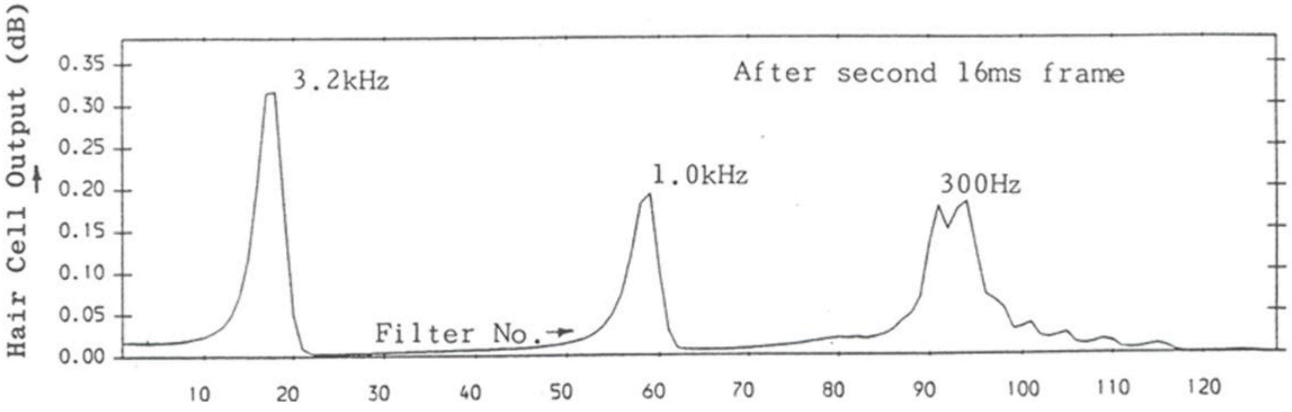




Sinusoidal components at the input





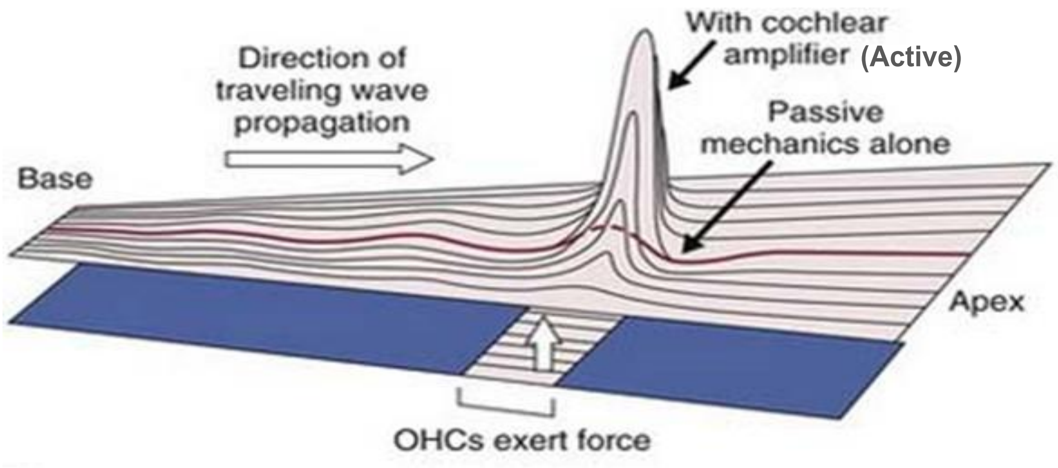


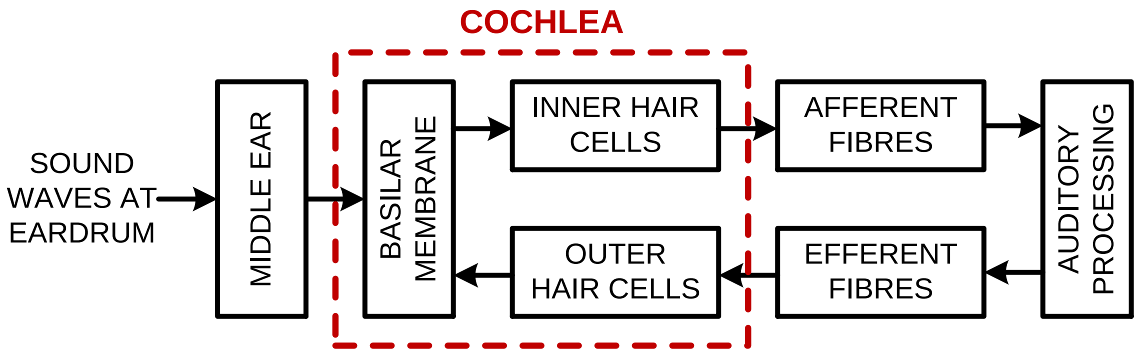
- ✓ The basilar membrane displacement and the corresponding inner hair cell output in response to a sum of three sinusoidal components applied at the input (3.2kHz, 1kHz and 300 Hz).
- ✓ The inner hair cell model output shows the spectral components in the input signal.

Adaptive Models of the Cochlea

Cochlear Response with Active Gain

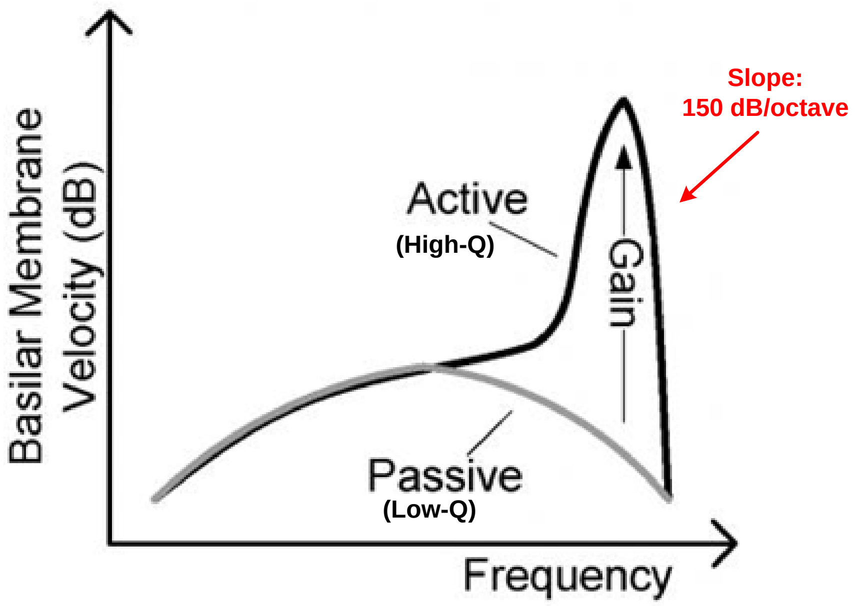
- ✓ Human auditory system can process a vast range of sounds spanning some twelve orders of magnitude of input pressure intensity
- ✓ In order to achieve this, the cochlea makes use of both passive and active systems
- ✓ The outer hair cells (OHC) provide this active mechanism - they amplify the motion picked up by the IHC at low input sound levels at that frequency

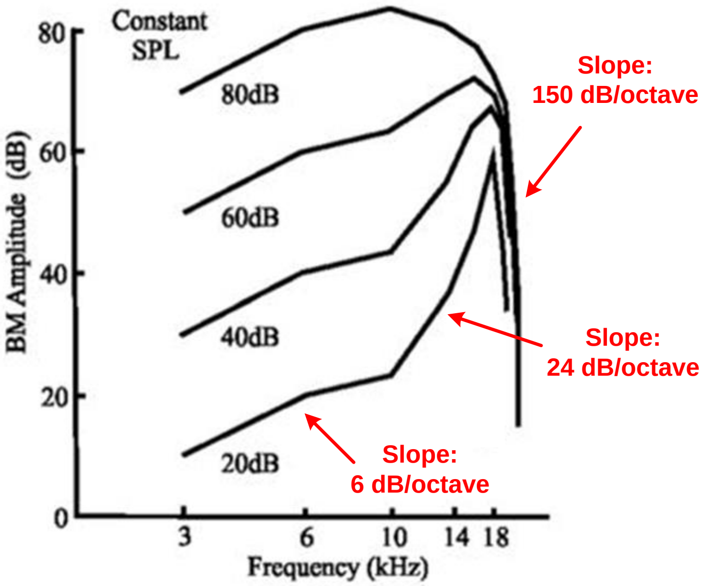




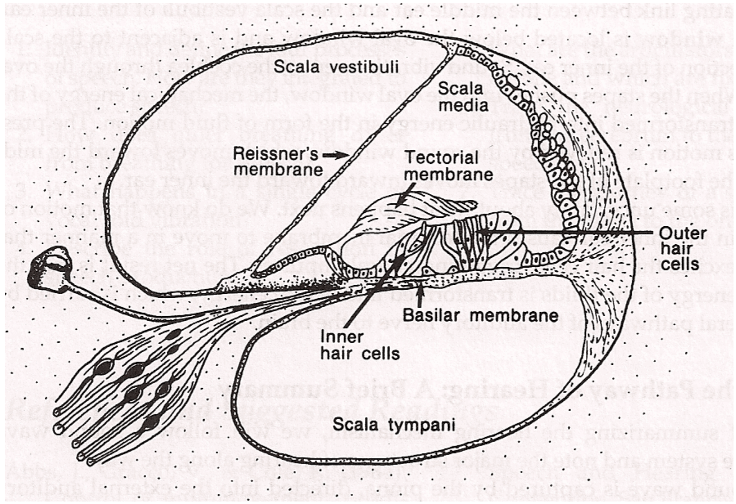
Transmission Line Model with Feedback

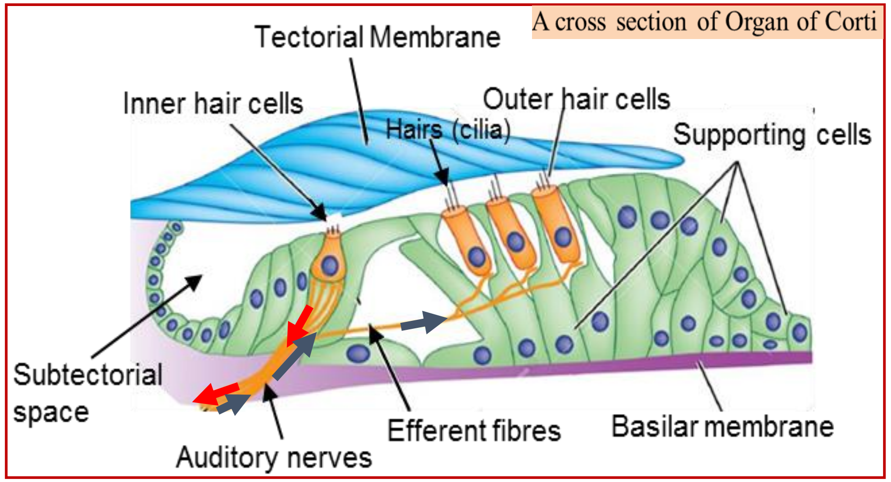
- ✓ The basilar membrane within the cochlea is normally in a passive state.
- ✓ Upon stimulation by a frequency of low amplitude, the section of the basilar membrane corresponding to that frequency transitions to an active state (adaptively higher-Q spectral decomposition).
- ✓ It is surprising how this large number of locally acting feedback loops can act together to give a large and uniform amplification of the global response of the BM.
- ✓ It would be desirable to have an active model of the cochlea that incorporates the level-dependent adaptive gain and adaptive frequency selectivity properties.

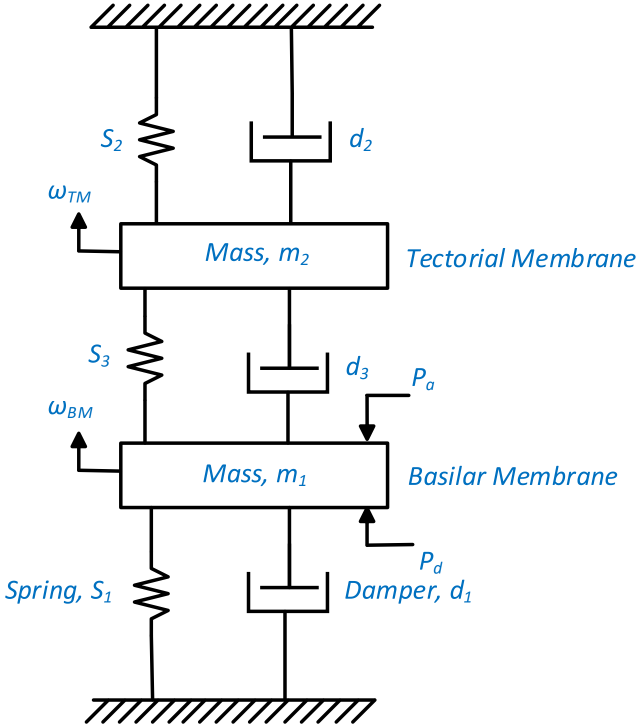


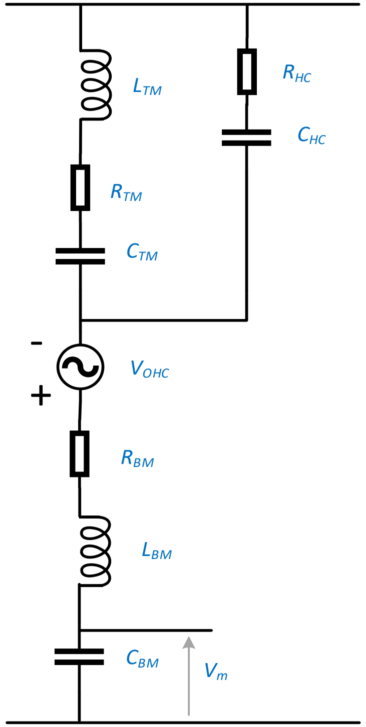


Active Cochlear Modelling

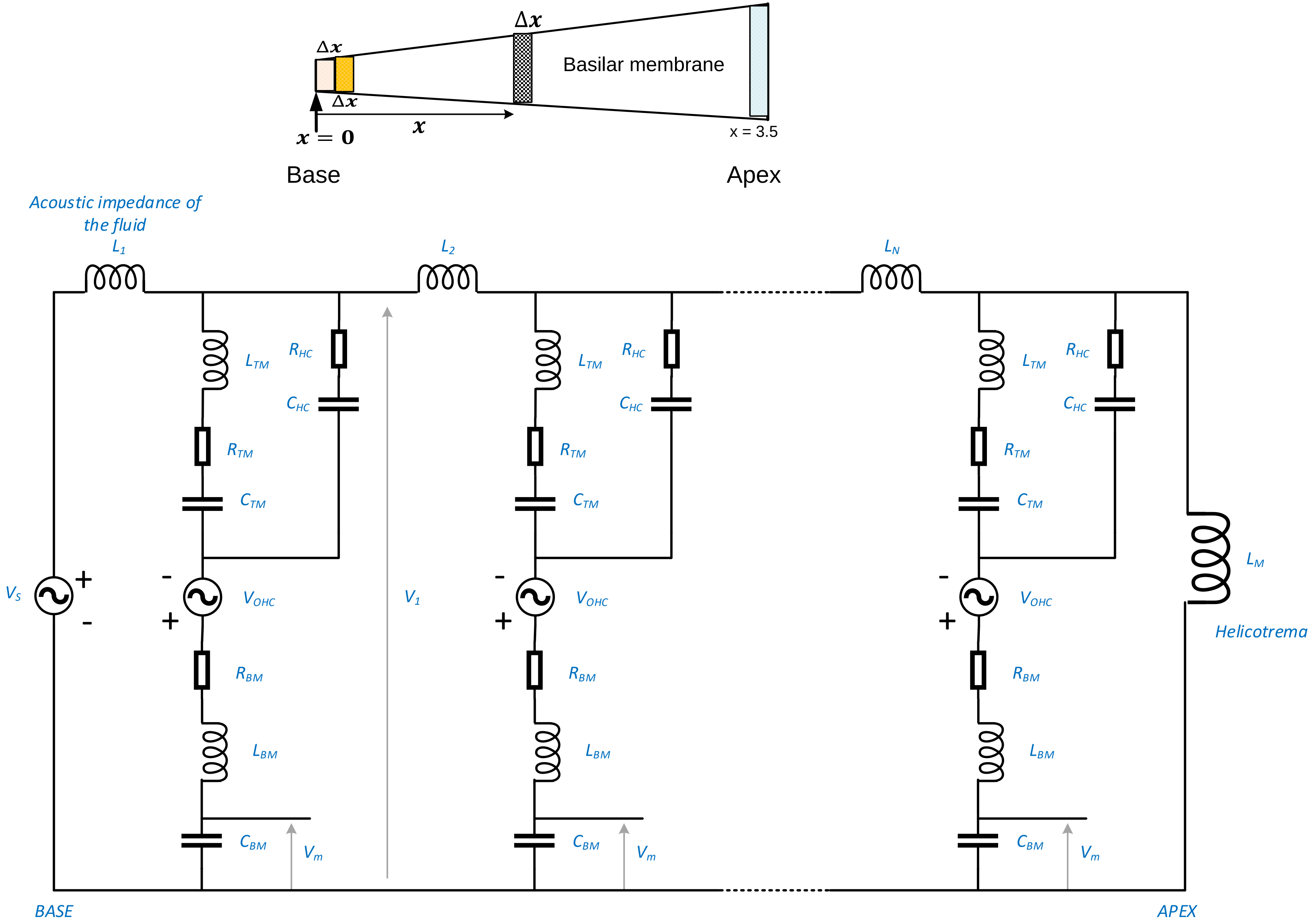




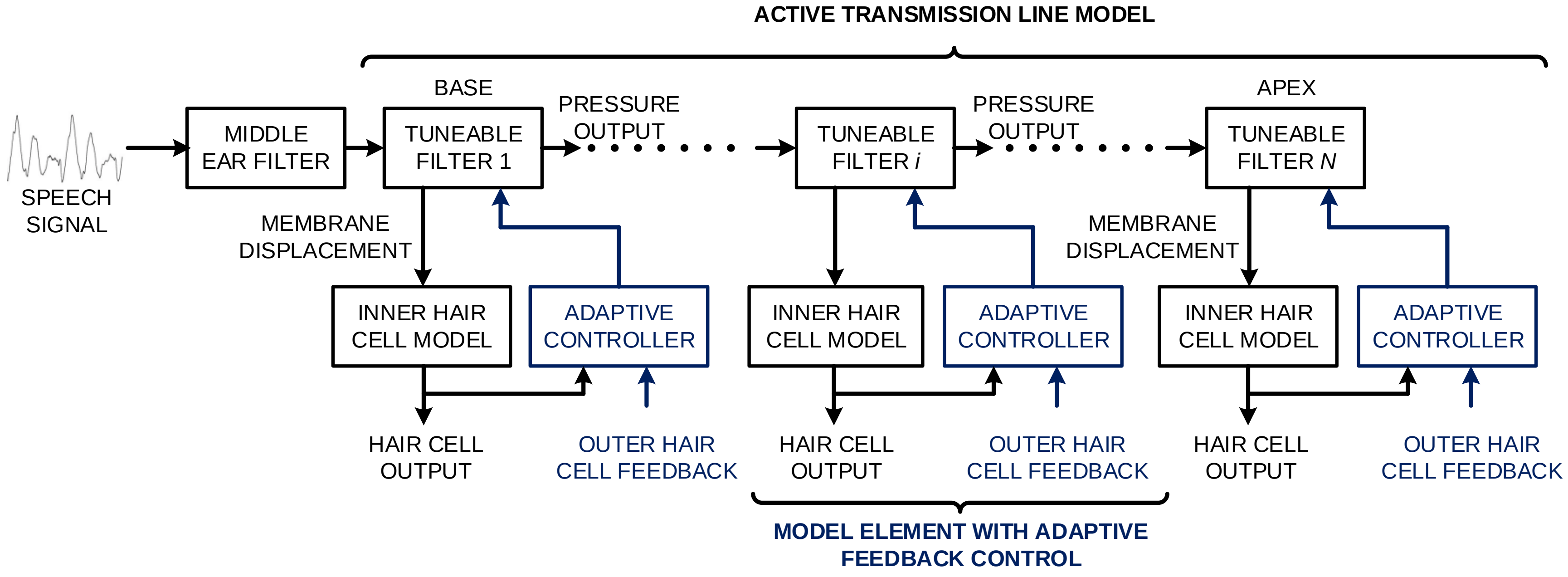


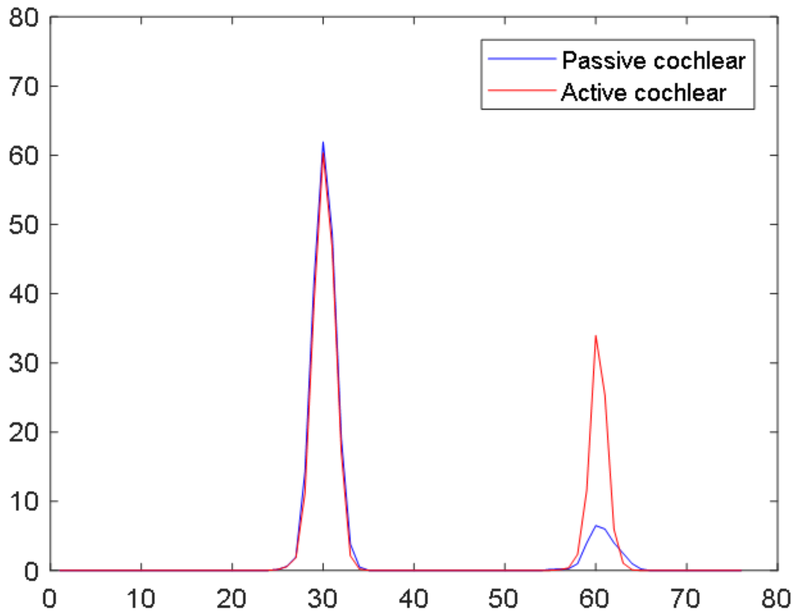


Active Cochlear Modelling



Active Transmission Line Model



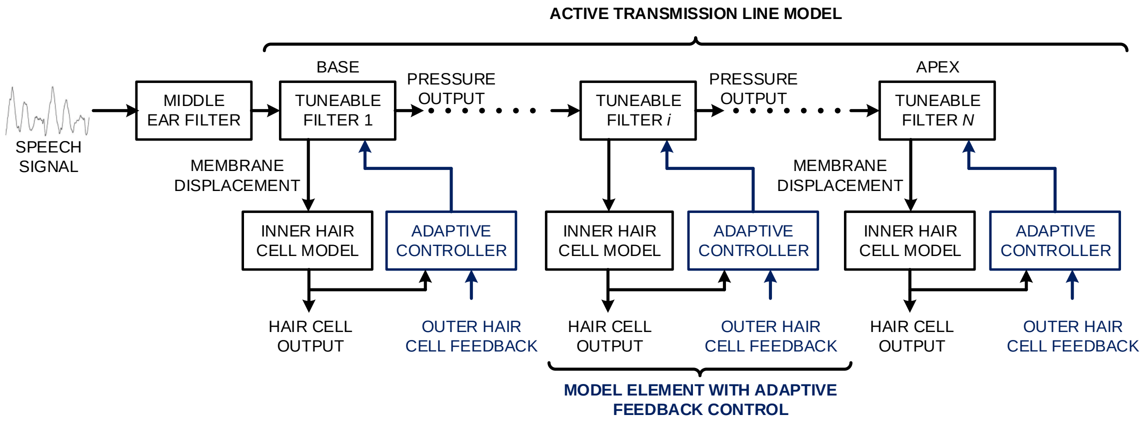


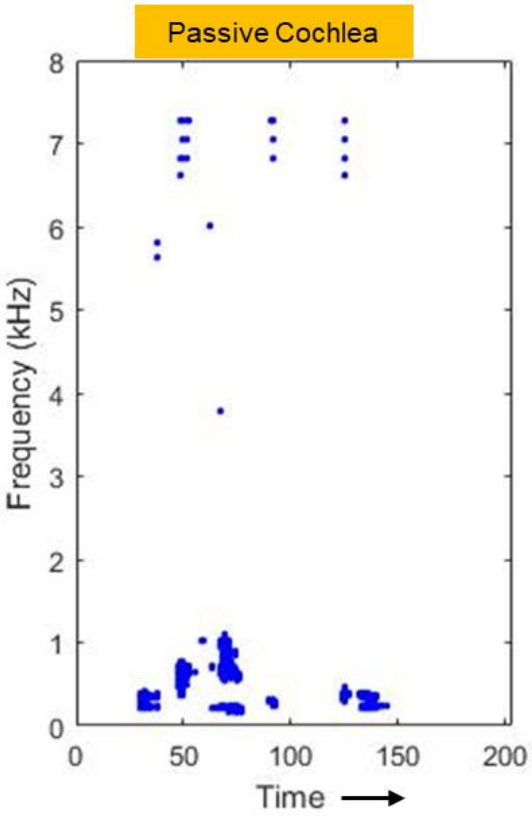
Ideally, models of the cochlea should exhibit

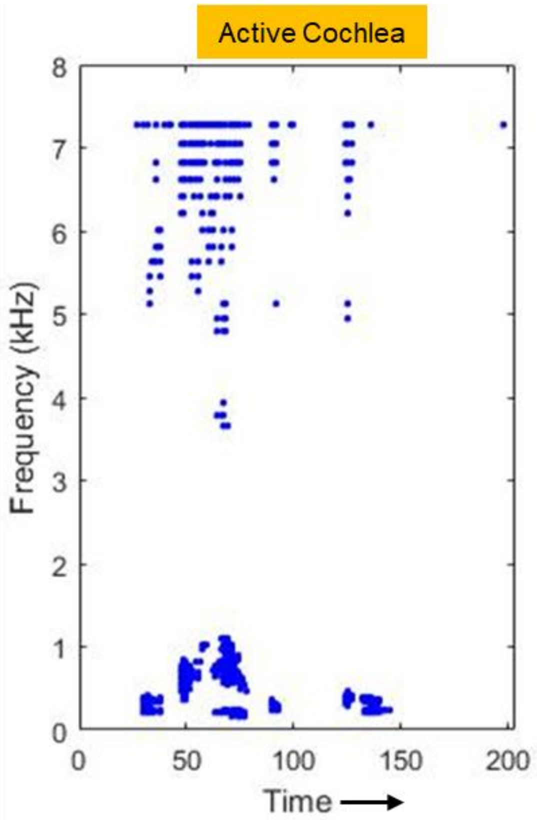
- level-dependence
- sharp auditory tuning curves
- fast adaptation

to changes in the input stimuli

Active Transmission Line Model







Integration of Cochlear Models with Machine Learning

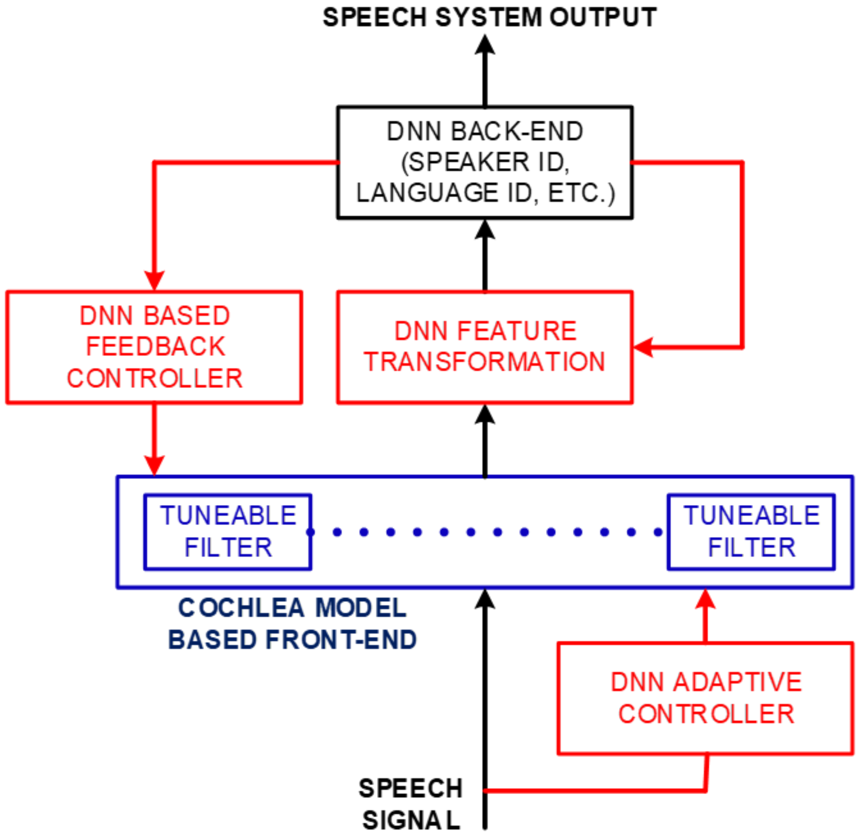
Australian Government
Australian Research Council

IEEE Sri Lanka Section Signal Processing Society Chapter

03 January 2022

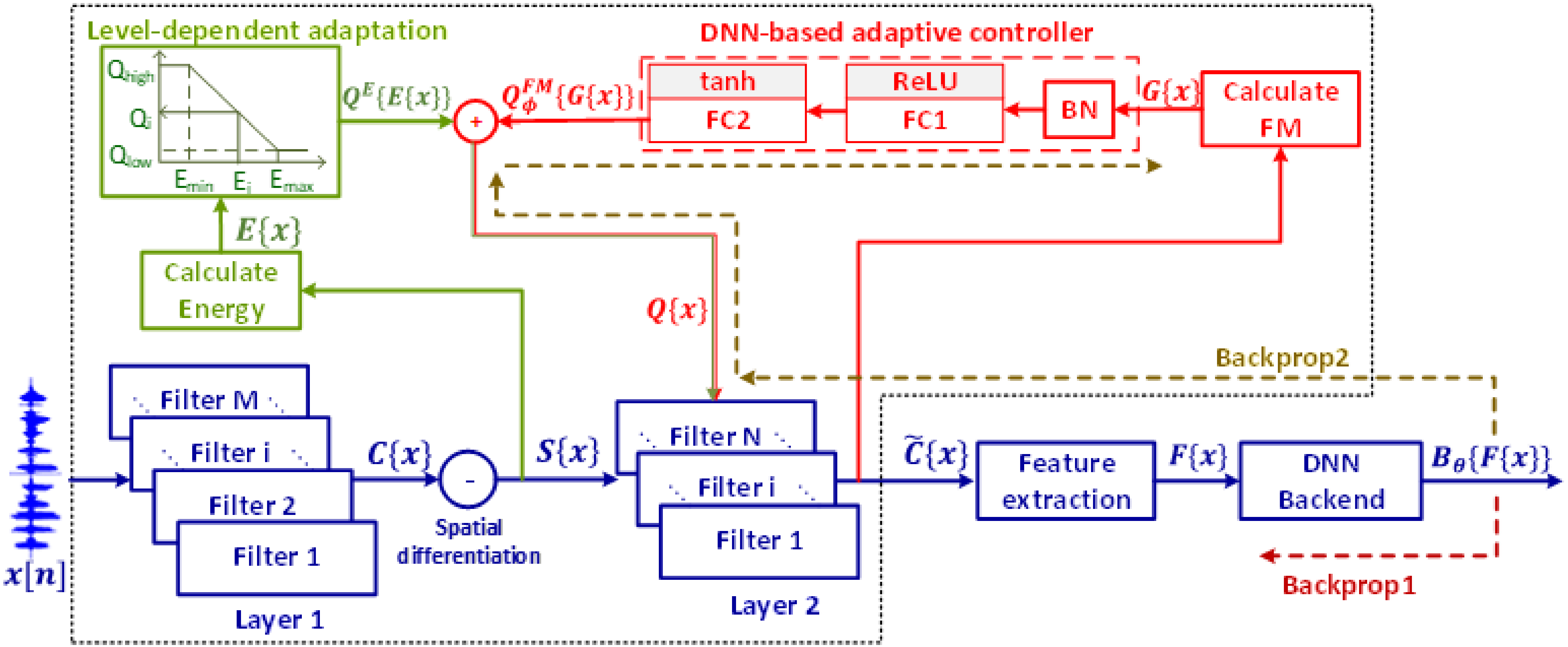
Integration Auditory Models with DNN

- ✓ A feed-forward adaptive spectral decomposition model based on the cochlea;
- ✓ A back-end dependent feedback path to improve the adaptive spectral decomposition;
- ✓ Extending the end-to-end system to learn a channel-invariant speech signal representation.

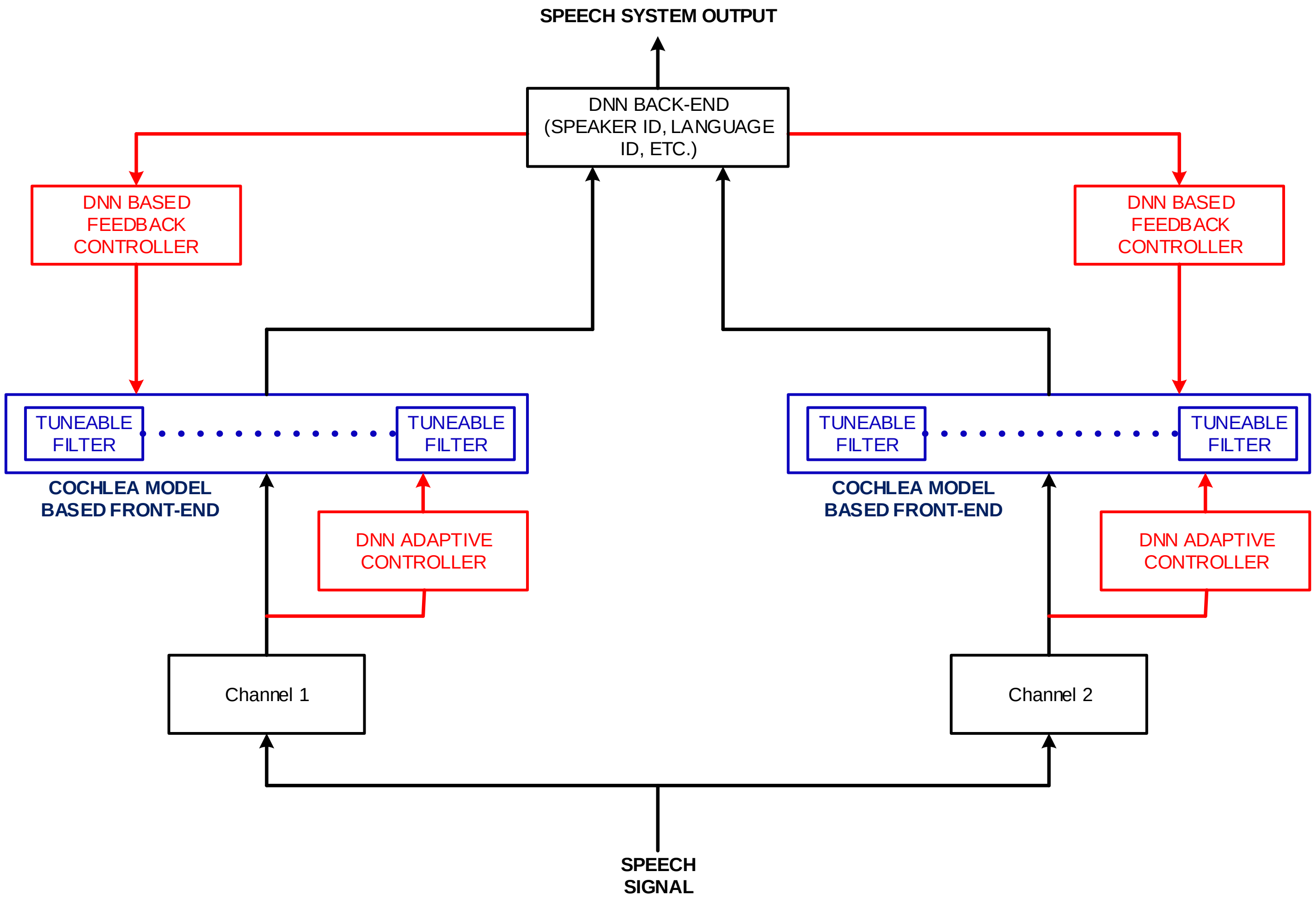


Australian Government
Australian Research Council

Integration Auditory Models (Parallel Filterbank) with DNN



Future Trends –Binaural Auditory Models for AI



✓ Exploit binaural variations to develop a robust cochlear front-end for future AI systems

Australian Government
Australian Research Council

Conclusions

- ✓ Future models of cochlea will include active feedback mechanisms to improve detection of smaller signals
- ✓ Filters in the cochlear models and adaptive feedback paths both may be implemented as deep learning models thus enabling integration with state-of-the-art speech processing systems
- ✓ This could lead to benefits for cochlear implants in terms of being able to adapt to the environment as well as learning and adapting to the individual