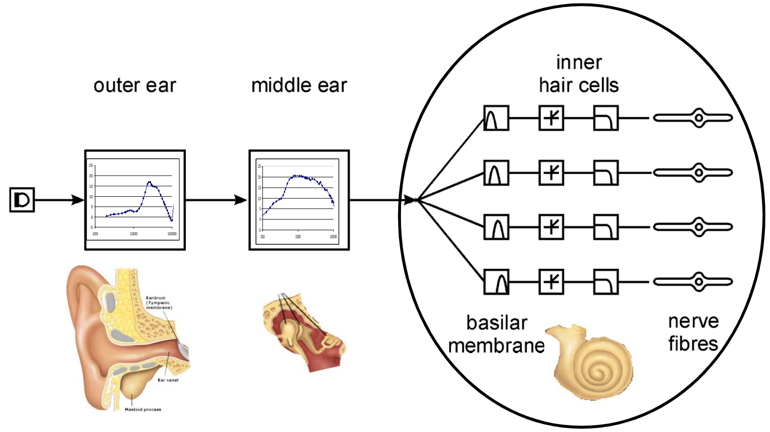
# AUDL 4007
# Auditory Perception

## Week 1

## The cochlea & auditory nerve: Obligatory stages of auditory processing



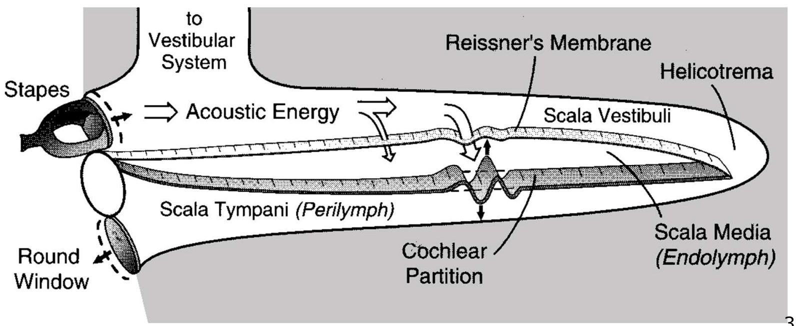
## Think of the ear as a collection of ‘systems’, transforming sounds to be sent to the brain





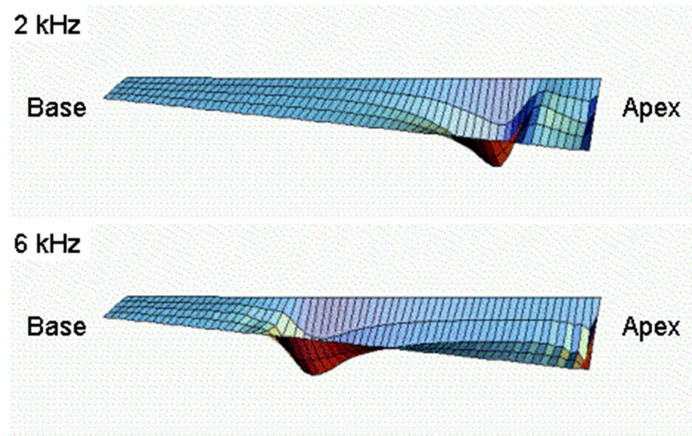

## Neural firing depends upon basilar membrane vibration

## Imagine the cochlea unrolled

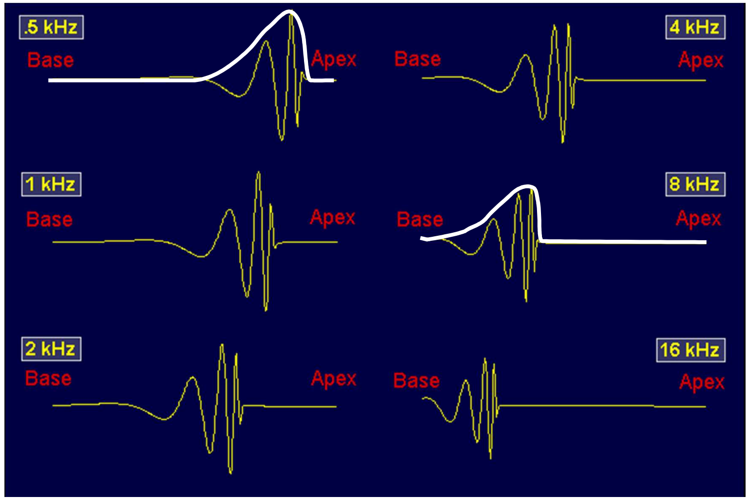





## Basilar membrane motion to two sinusoids of different frequency



# Defining the envelope of the travelling wave

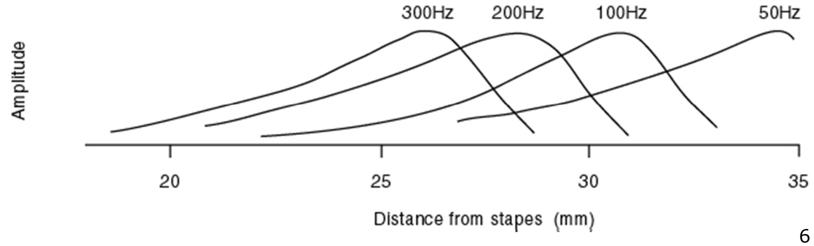



allkhalf.mov



# A crucial distinction
## *excitation pattern* vs. *frequency response*

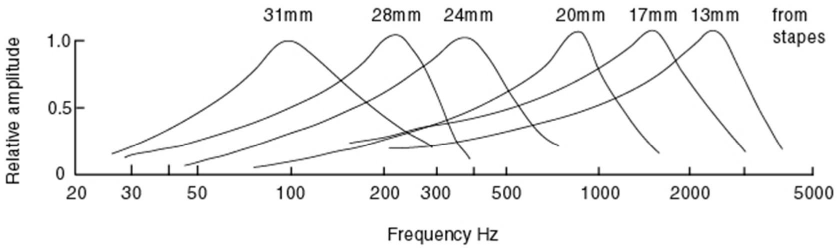
- Excitation pattern — the vibration pattern across the basilar membrane to a single sound.
  - Input = 1 sound.
  - Measure at many places along the BM.
- Essentially the envelope of the travelling wave
- Related to a *spectrum* (amplitude by frequency).





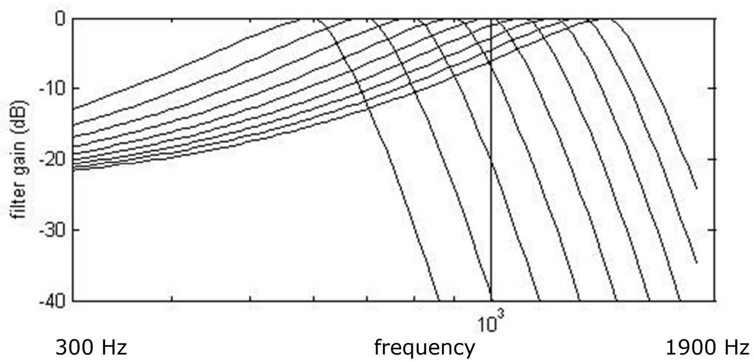

# A crucial distinction
## *excitation pattern* vs. *frequency response*

- Frequency response — the amount of vibration shown by a particular place on the BM to sinusoids of varying frequency.
  - Input = many sinusoids.
  - Measure at a single place on the BM.
  - Band-pass filters at each position along the basilar membrane.



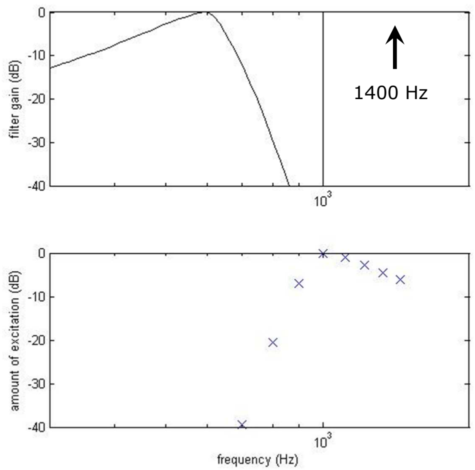



# Two sides of the same coin: Deriving excitation patterns for a 1 kHz sinusoid from frequency responses




Note shallower slope to lower frequencies (left) for frequency responses

## Frequency responses with centre frequencies running from 1400 – 600 Hz

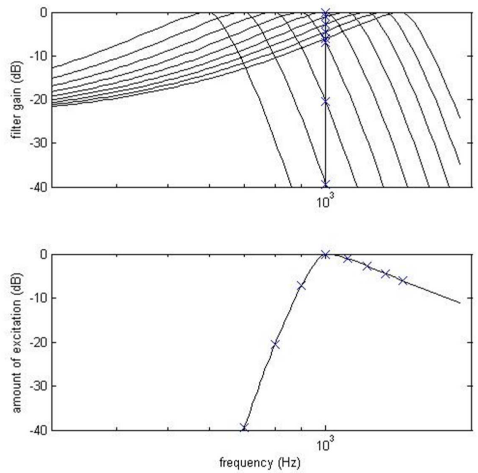





## Deriving excitation pattern from auditory filters

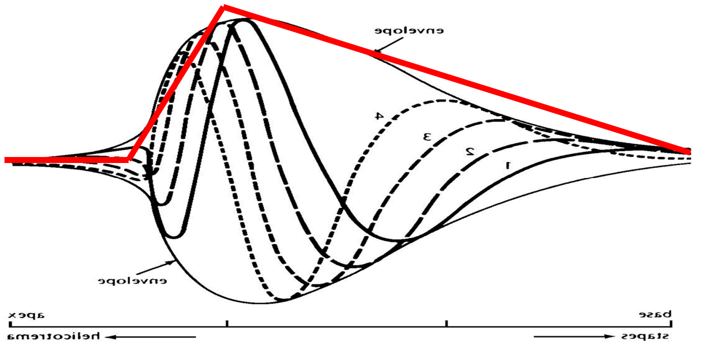
Note shallower slope to left

Note shallower slope to right






Now the other way around:
filter shapes from excitation patterns
Flip the orientation of the axis and schematise



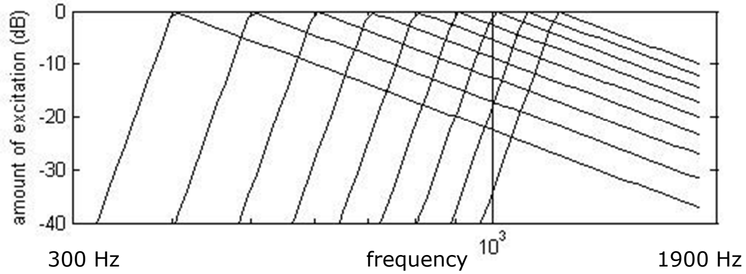
apex ← → base

low frequencies high

Note shallower slope to right

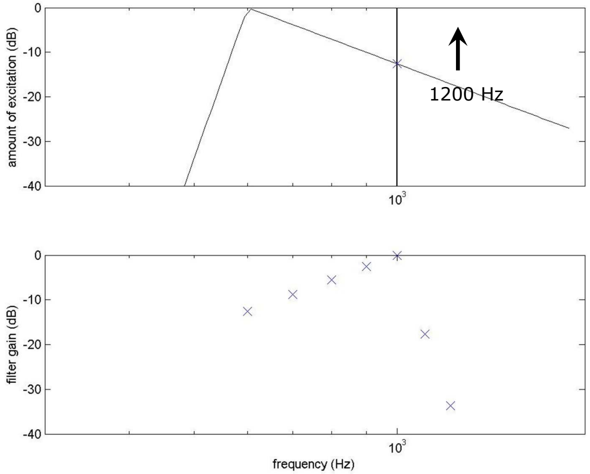


## The other side of the coin: Deriving a frequency response at 1 kHz from excitation patterns




Note shallower slope to higher frequencies (right) for excitation patterns

# Excitation patterns with centre frequencies running from 1200 – 400 Hz

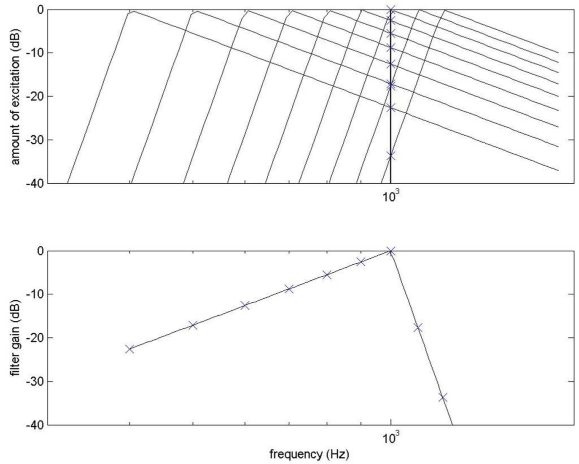





# Deriving frequency responses from excitation patterns

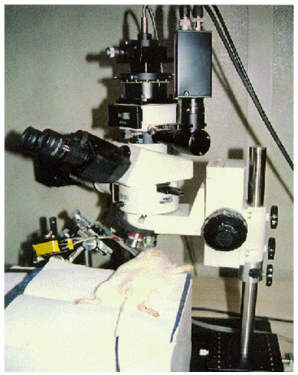
Note shallower slope to right

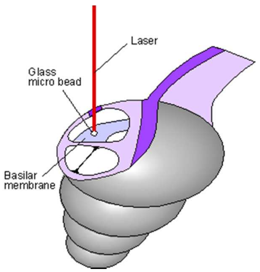
Note shallower slope to left

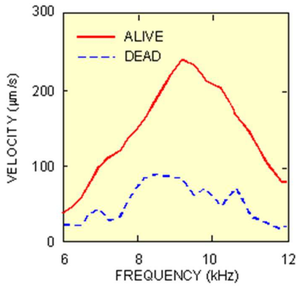





# Laser Doppler Velocimetry






http://www.wadalab.mech.tohoku.ac.jp/bmldv-e.html



Modern measure-ments of the frequency response of the basilar membrane

Consider the frequency response of a *single place* on the BM

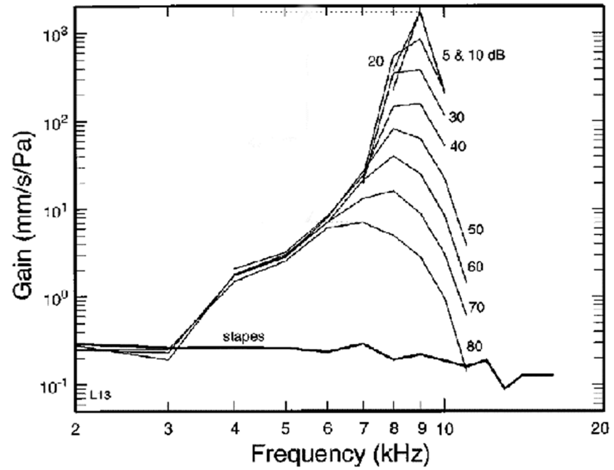



FIG. 10. A family of isointensity curves representing the gain (velocity divided by stimulus pressure) of basilar-membrane responses to tone pips as a function of frequency (abscissa) and intensity (parameter, in dB SPL). The thick line at bottom indicates the average motion of the stapes (Ruggero *et al.*, 1990). Data recorded in cochlea L13.

input/
output
functions
on the
basilar
membrane

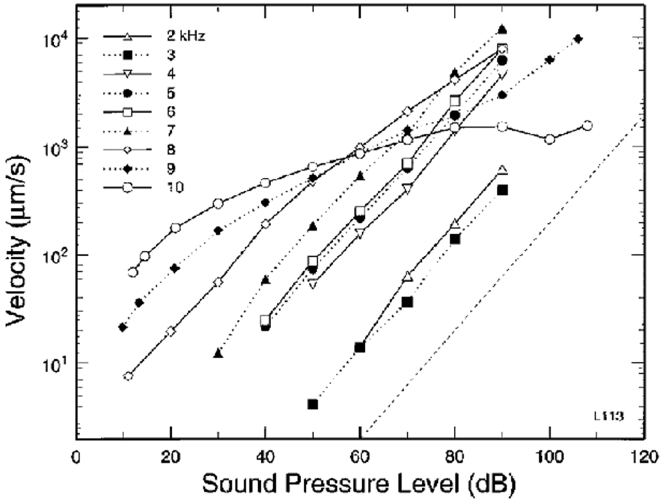
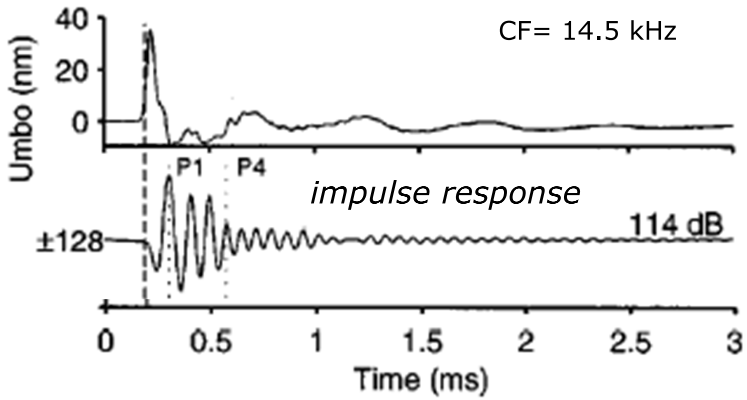



FIG. 7. Velocity-intensity functions of basilar-membrane responses to tones with frequency equal to and lower than CF (10 kHz). The straight dashed line at right has a linear slope (1 dB/dB).

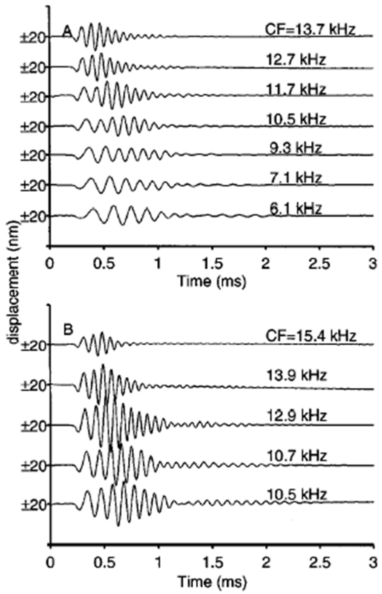


# Waveform of response to clicks on the basilar membrane (a.k.a. ?)



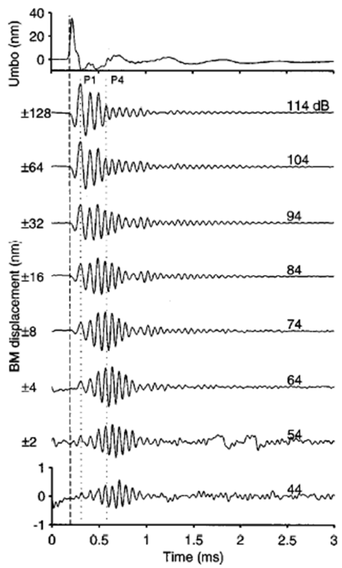



Click
responses
at various
BM places

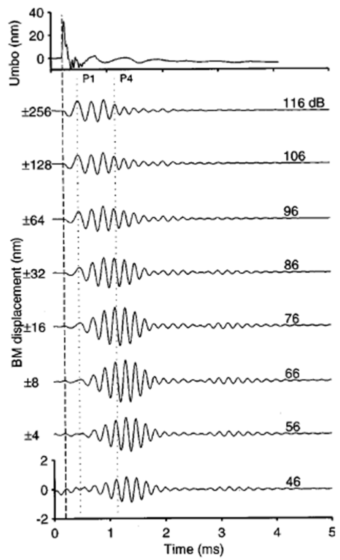




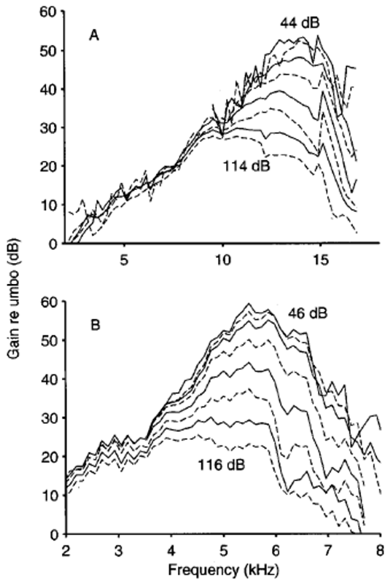

CF = 14.5 kHz




CF = 5.5 kHz

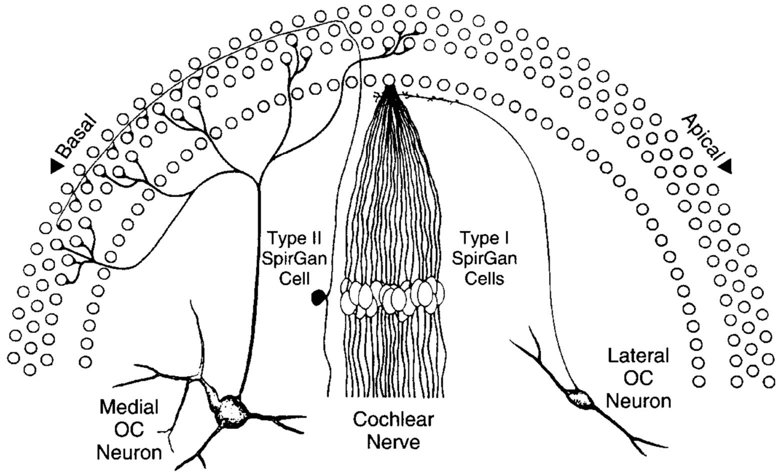


Fourier transform the response to clicks to obtain the frequency response (valid when?)






# Innervation of the cochlea




90-95% of afferents are myelinated, synapsing with a single inner hair cell (IHC).



# Four aspects of firing patterns on the auditory nerve

- The coding of intensity.
- The representation of the place code.
- The representation of temporal fine structure (for intervals ranging up to ≈20 ms).
- The representation of gross temporal structure.

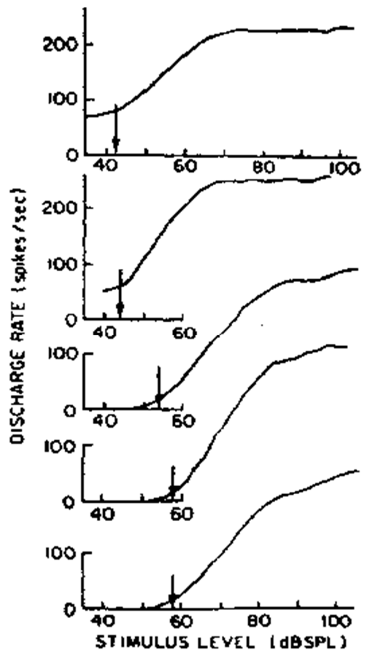


*Intensity*

Rate-level functions for auditory nerve fibres

Observe!

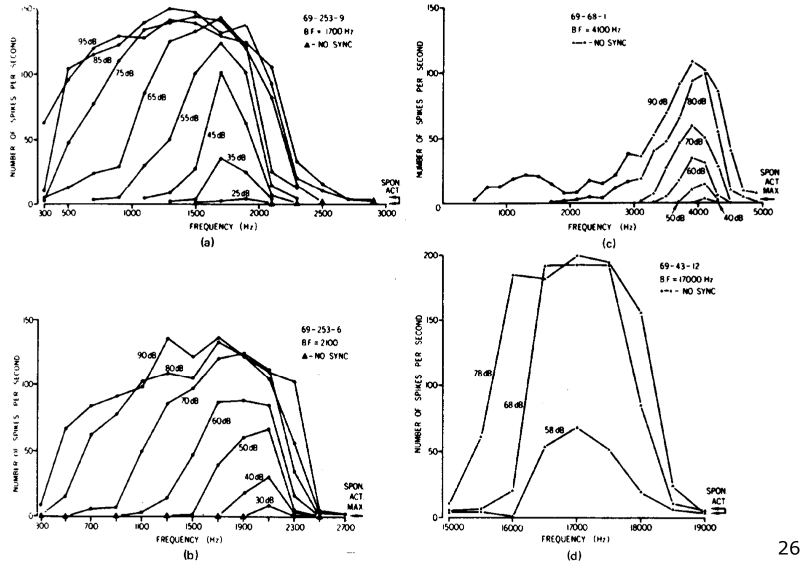
- Threshold
- Saturation
- Limited dynamic range



# However, firing rates depend not only on sinusoidal sound intensity but also on sound ...

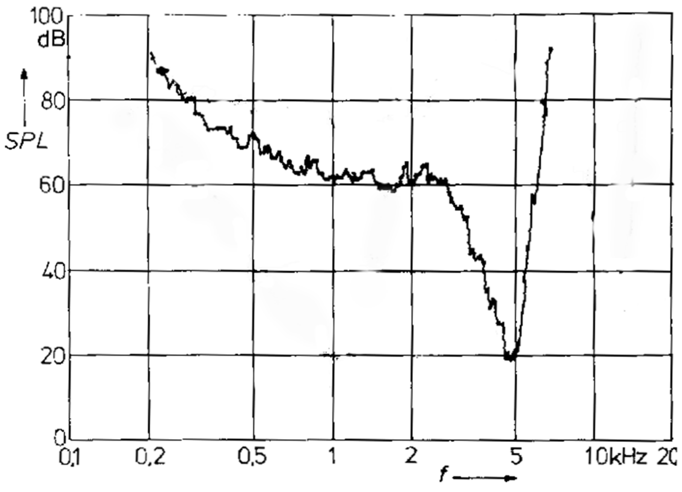


# Firing rate across frequency and level



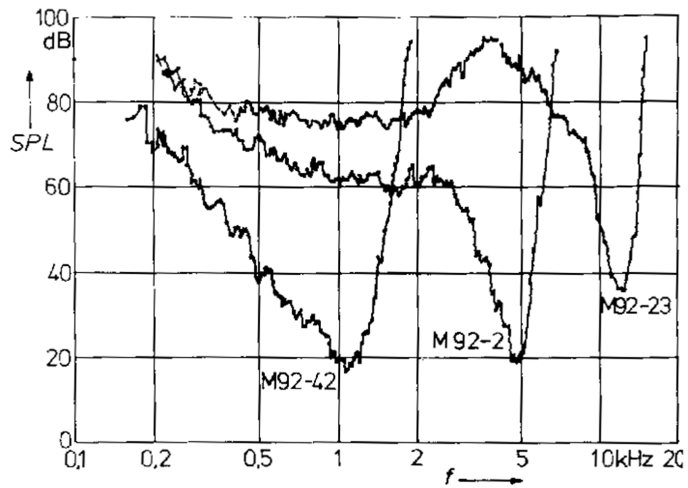



# 'Audiograms' of single auditory nerve fibres reflect BM tuning

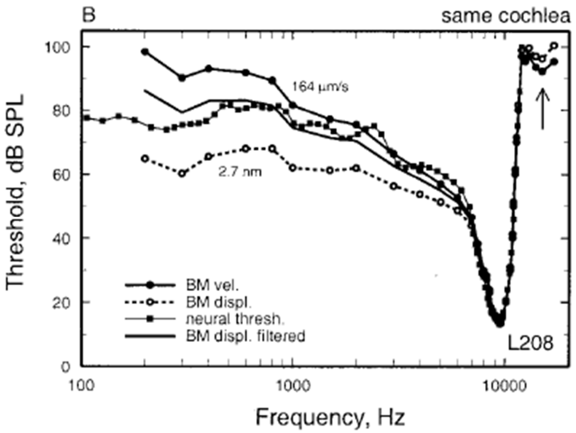





# The 'best' frequency of a particular tuning curve depends upon the BM position of the IHC to which the afferent neuron is synapsing



# BM and neural tuning compared




'filtered' is high-pass filter at 3.8 dB/octave. From Ruggero *et al*. 2000

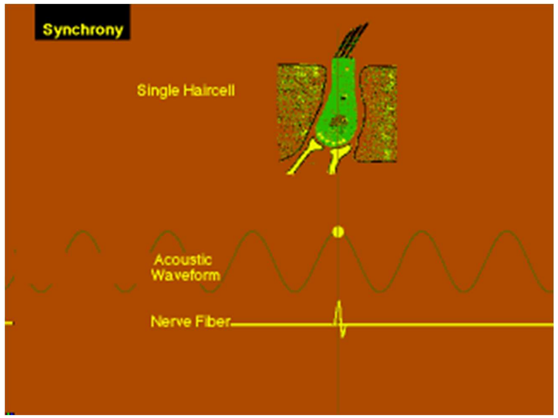


# *Temporal coding (up to ≈ 5 kHz)*

Information about stimulus frequency is not only coded by *which* nerve fibres are active (the *place* code) but also by *when* the fibres fire (the *time* code).



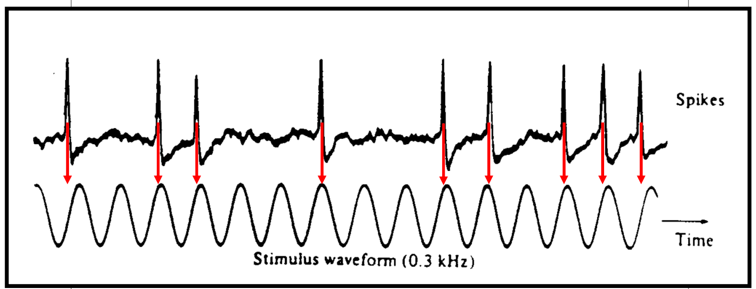
The firing of auditory nerve fibres is synchronized to movements of the hair cell cilia (at low enough frequencies)




Play transdct.mov



Auditory nerves tend to fire to low-frequency sounds at particular waveform times (*phase locking*).




*Not* the same as firing *rate*!

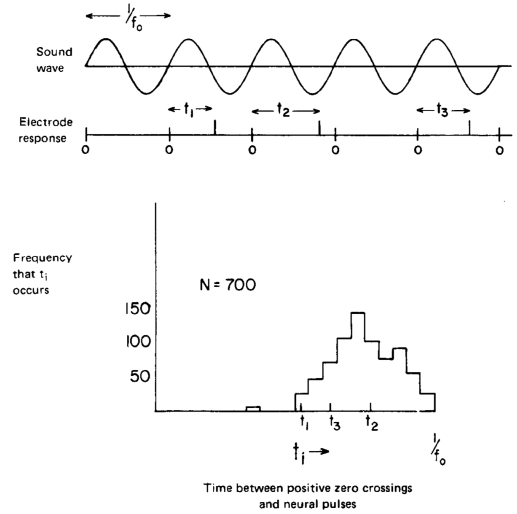
*Evans (1975)*

# But phase-locking is limited to lower frequencies ...

- Synchrony of neural firing is strong up to about 1-2 kHz.
- No evidence of synchrony above 5 kHz.
- The degree of synchrony decreases steadily over the mid-frequency range.

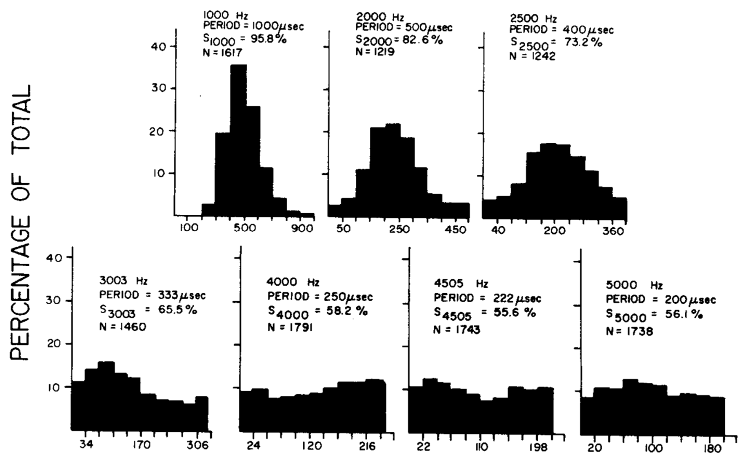


# ... as readily seen in a *period histogram*





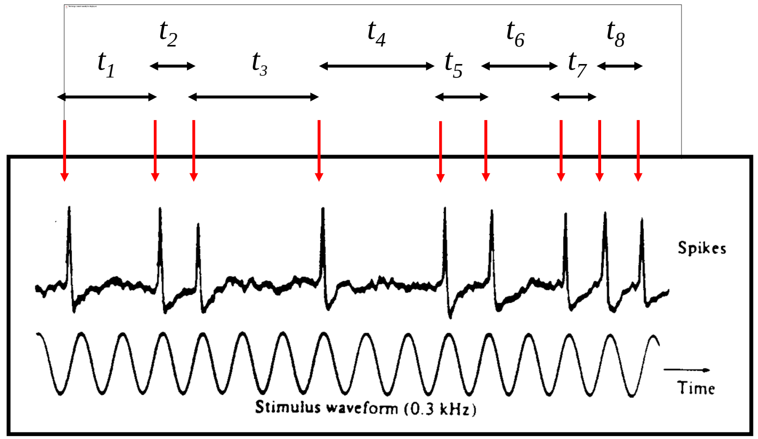

# Period histograms across frequency



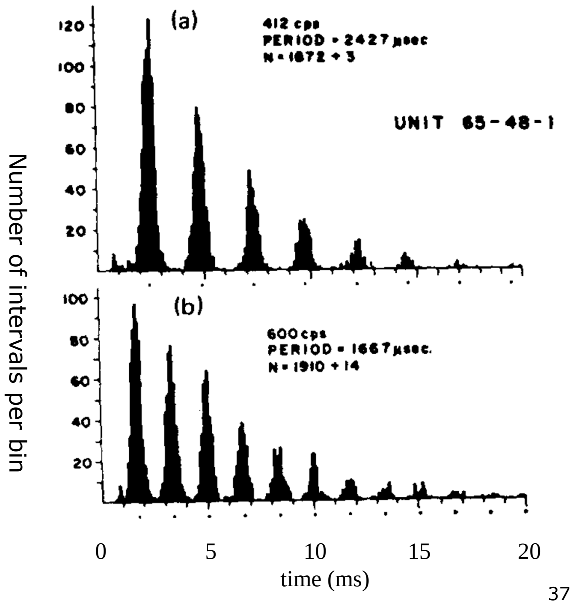

Note half-wave rectification and *synchrony index*



# Constructing an *interval histogram*

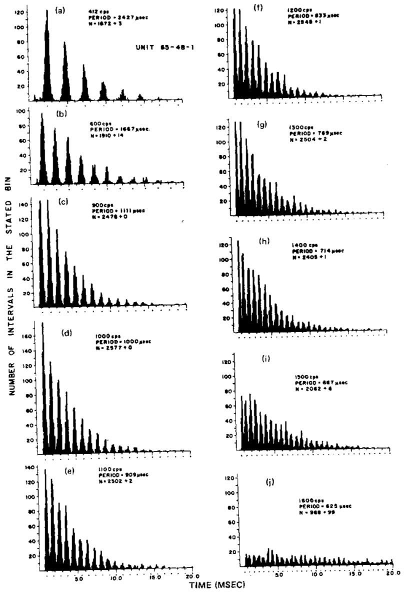


Interval histograms for a single AN fibre at two different frequencies



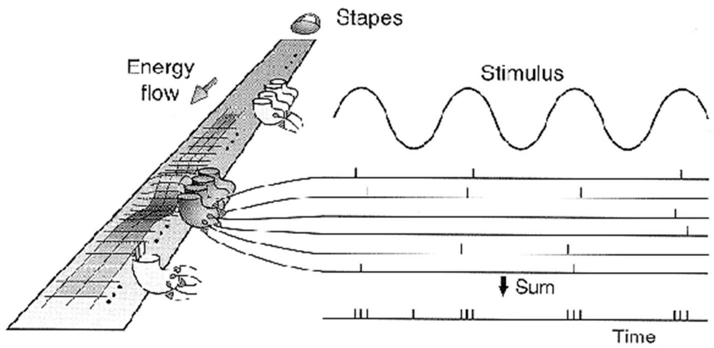



Interval histograms for a single AN fibre across frequency





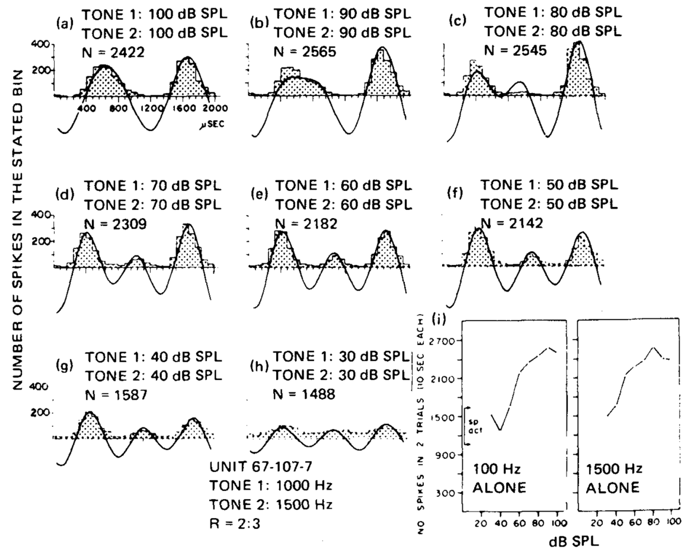

# Neural stimulation to a low frequency tone



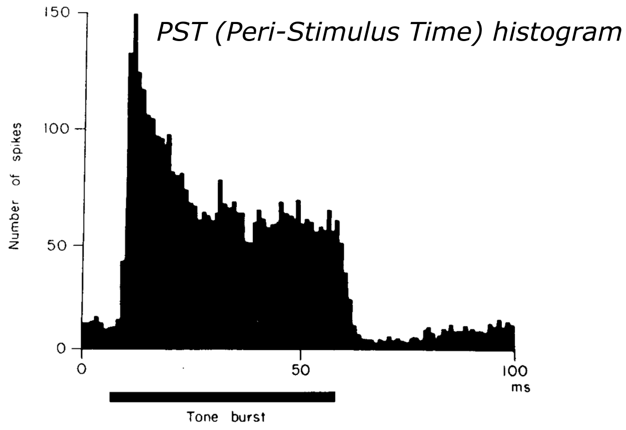

Sound energy propagates to the characteristic place of the tone where it causes deflection of the cochlear partition. Neural spikes, when they occur, are synchronized to the peaks of the local deflections. The sum of these neural spikes tends to mimic the wave shape of the local deflections.



# Period histograms to more complex sounds



## *Gross temporal structure*
## Enhanced response to sound onsets: The value of novelty



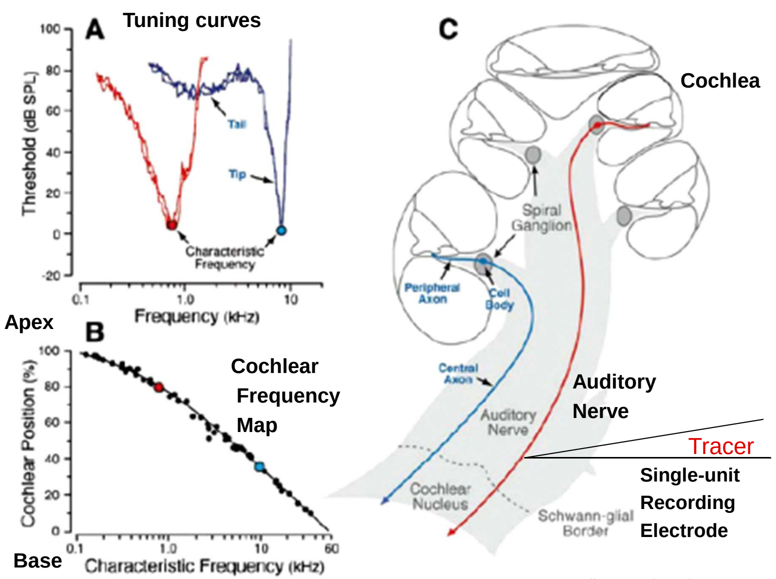



## Where we've got to …

- Outer ear channels sound to the middle ear, and can be characterized as a bandpass filter.
- Middle ear effects an efficient transfer of sound energy into the inner ear, again with the characteristics of a bandpass filter.
- Inner ear
  - Transduces basilar membrane movements into nerve firings …
  - which are synchronised to peaks in the stimulating waveform at low enough frequencies
  - Performs a mechanical frequency analysis, which can be envisioned as the result of analysis by a *filter bank.*



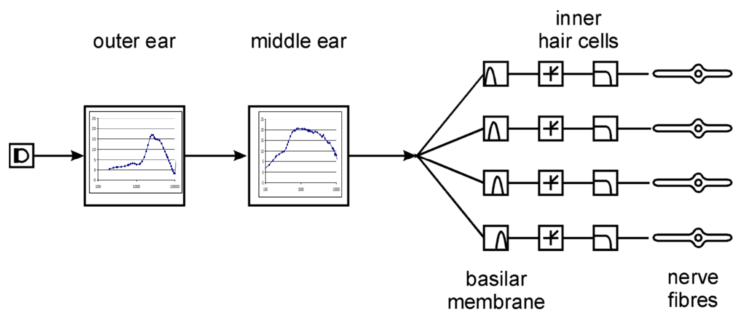
**Auditory Nerve Structure and Function**




Liberman (1982)



## A systems model of the auditory periphery



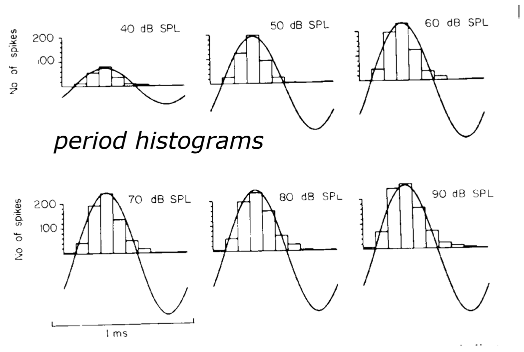
# What properties should the filter bank have?

- Filter spacing
  - *Corresponding to tonotopic map*
- Filter bandwidth
  - *vary with frequency as on the basilar membrane*
- Filter nonlinearity
  - *vary gain and bandwidth with level as on the basilar membrane*



# Modelling the hair cell/auditory nerve synapse

- Neuro-transmitter is released when cilia are pushed in one direction only, tied to polarity of basilar membrane motion
  - *half-wave rectification*

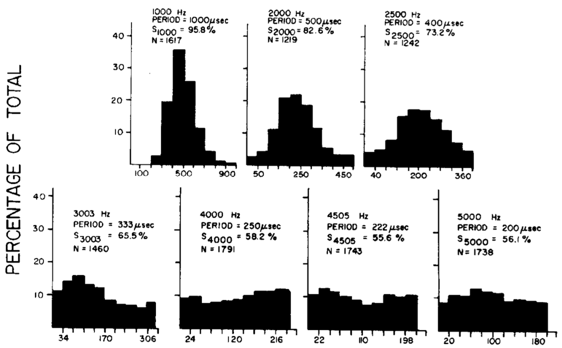





# Modelling the hair cell/auditory nerve synapse

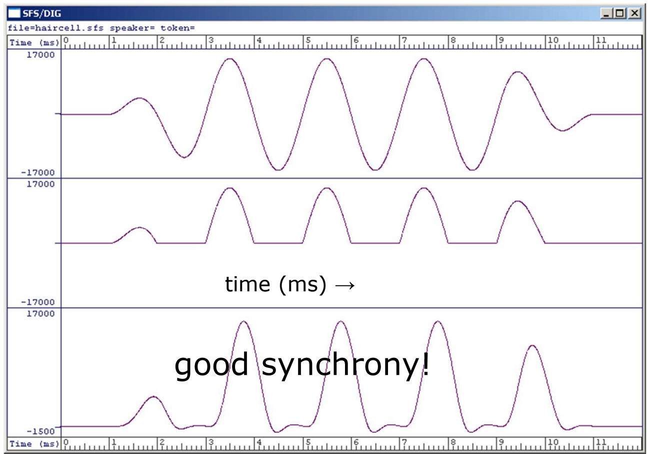
Phase-locking is limited to low frequencies
- *low-pass filtering*

*period histograms across frequency*






# Simulating hair cell transduction at 500 Hz



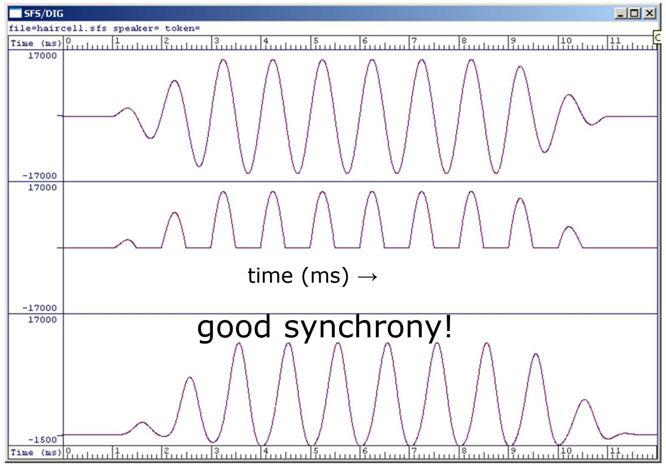

input wave

½–wave rectification

smoothing with a 1.5 kHz lowpass filter

think of this last wave as driving the auditory nerve
(e.g., as the amount of neurotransmitter in the synaptic cleft)

# Simulating hair cell transduction at 1000 Hz



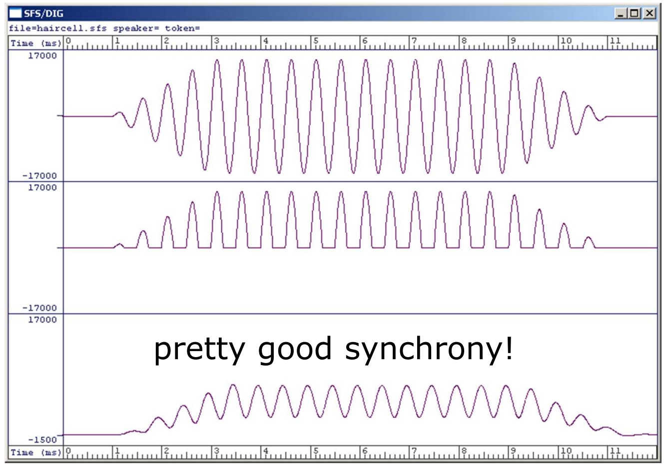

input wave

½–wave rectification

smoothing with a 1.5 kHz lowpass filter



# Simulating hair cell transduction at 2000 Hz



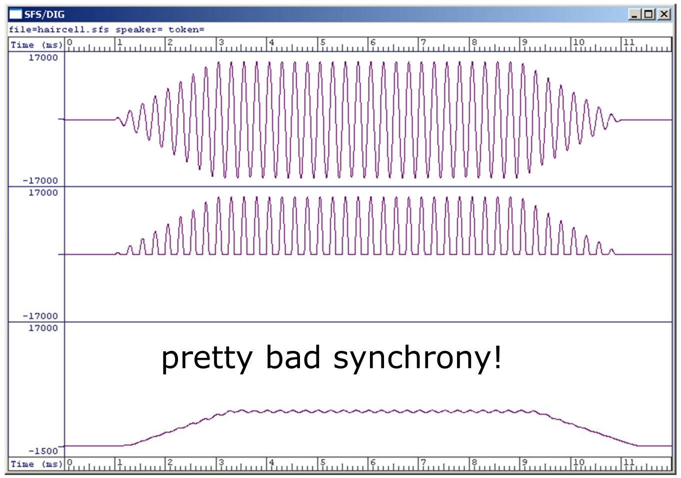

input wave

½–wave rectification

smoothing with a 1.5 kHz lowpass filter



# Simulating hair cell transduction at 4000 Hz



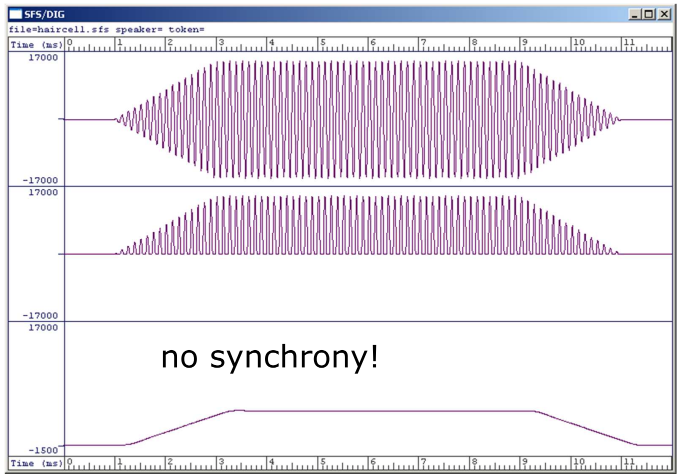

input wave

½–wave rectification

smoothing with a 1.5 kHz lowpass filter



# Simulating hair cell transduction at 8000 Hz



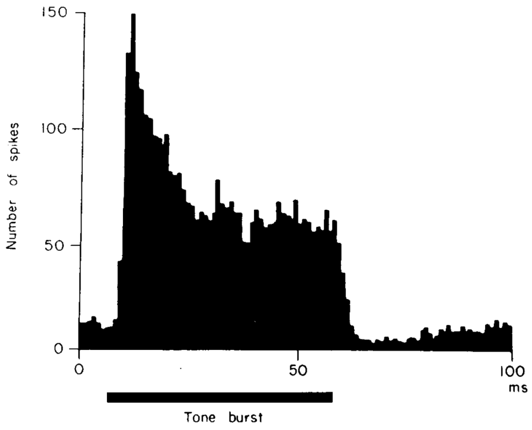

input wave

½–wave rectification

smoothing with a 1.5 kHz lowpass filter

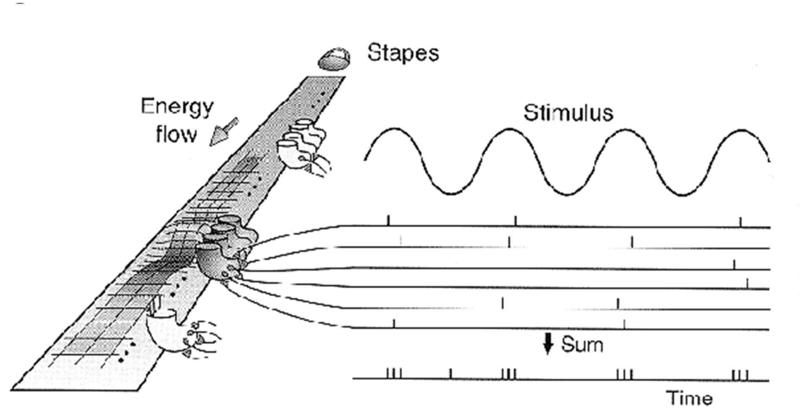
# Modelling the hair cell/auditory nerve synapse

- Rapid adaptation
  - *need some kind of automatic gain control (agc)*



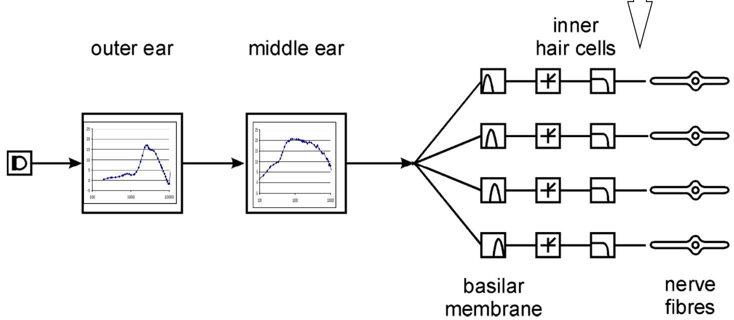



# Neural stimulation to a low frequency tone






# We're done!
# (but need agc here)



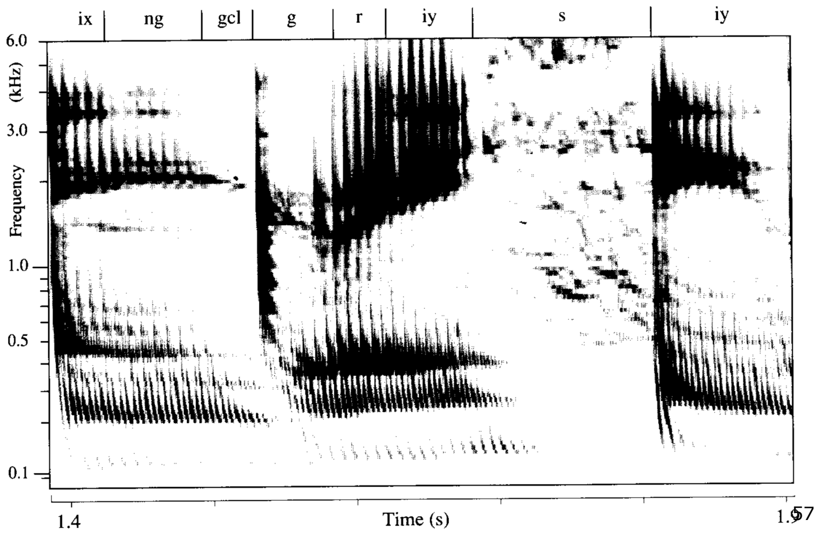



# A spectrogram with 'ear-like' processing (Giguere & Woodland, 1993)

(*typical spectrogram properties in italics)*

- A first-stage broad band-pass linear filter to mimic outer and middle ear effects *(pre-emphasis filter).*
- A filterbank whose centre frequencies are arranged in the same way as the human tonotopic (frequency to place) map ... *(equal spacing of filters in Hz).*
- with non-linear filters whose bandwidths increase as level increases (*linear filters with a fixed bandwoidth)*.
- Smearing of temporal information so as to mimic the frequency limitation of phase locking in the auditory nerve *(smearing by choice of temporal window/filter bandwidth — no extra processing )*.

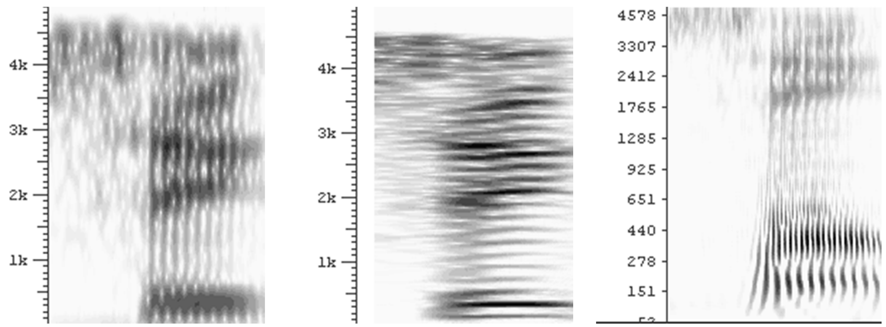
# An auditory spectrogram






# Types of Spectrogram

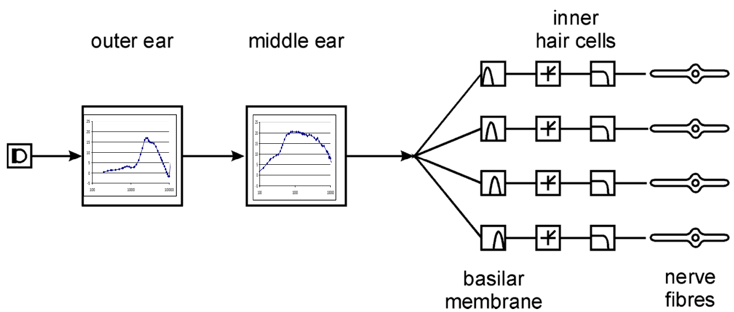
Wide-band | Narrow-band | Auditory




*An auditory spectrogram looks like a wide-band spectrogram at high frequencies and a narrow-band spectrogram at low frequencies (but with more temporal structure).*

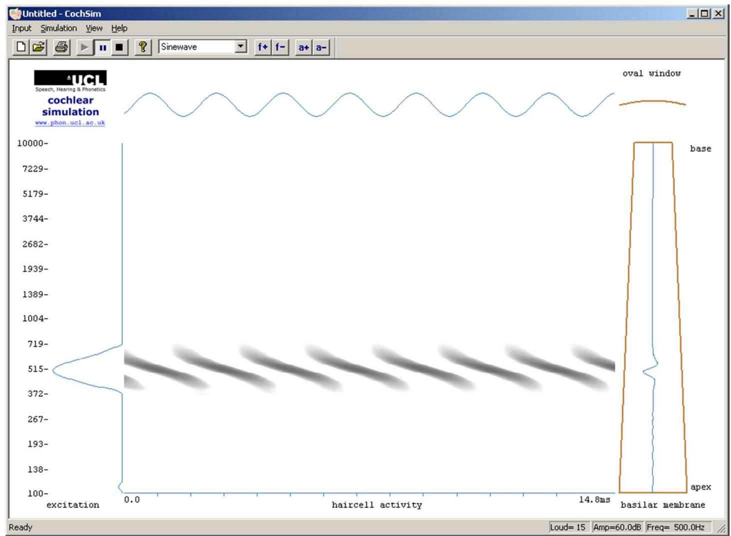


# Laboratory session: A computer implementation of essentially this model

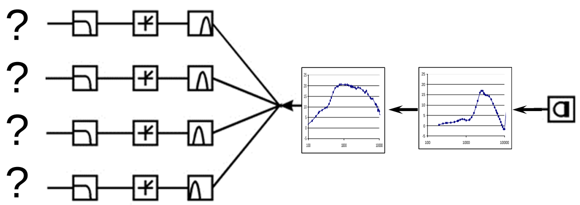





# A cochlear simulation

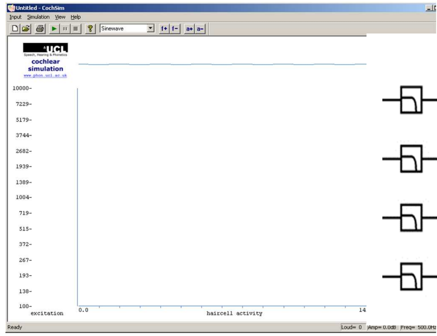


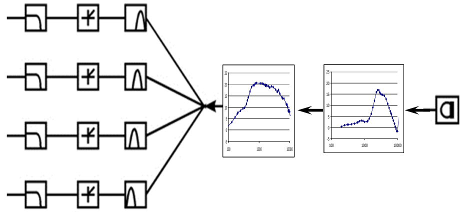
# Flip it around






# A cochlear simulation



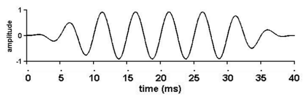



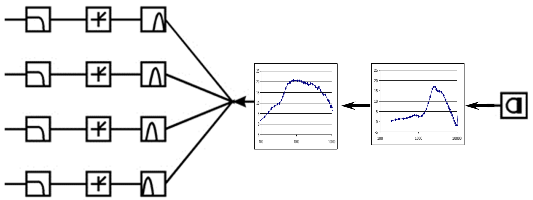
How should we look at the output of the model?



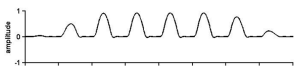
# Could look at the output waveforms

input signal



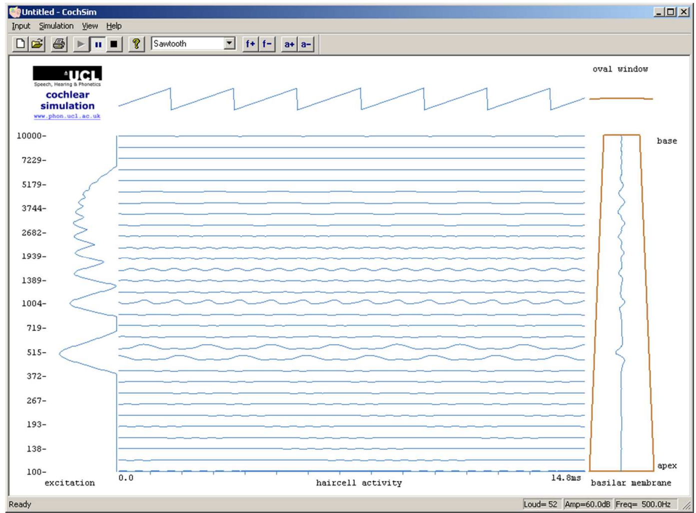



output signal

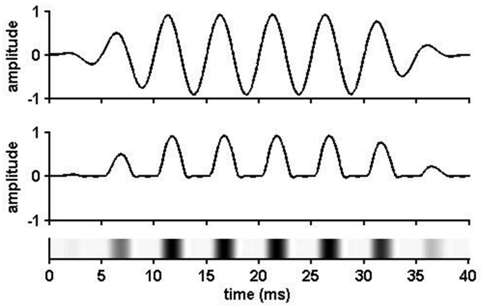





# But hard to see what is going on (especially for complex waves)



# Solution: encode wave amplitude in a different way




waveform at 200 Hz

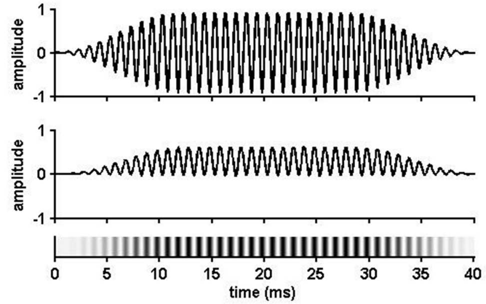
rectified & smoothed

spectrographic

waveform amplitude is recoded as
the darkness of the trace



# Encode wave amplitude as trace darkness



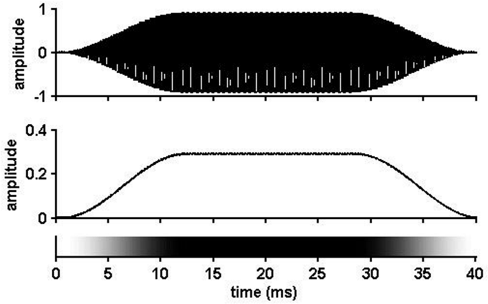

waveform at 1 kHz

rectified & smoothed

spectrographic



# Encode wave amplitude as trace darkness




waveform at 4 kHz

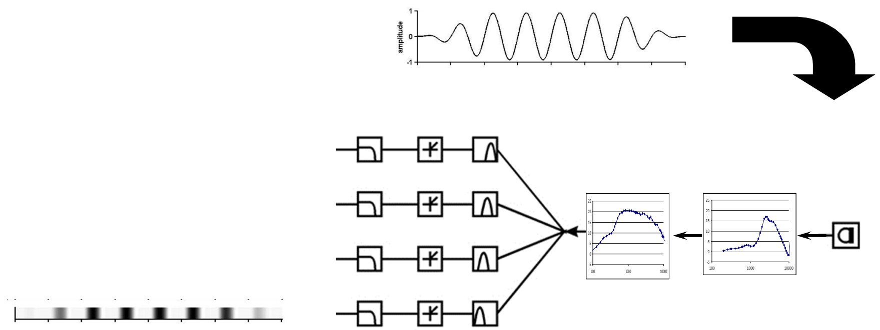
rectified & smoothed

spectrographic



# Construct the output display one strip at a time

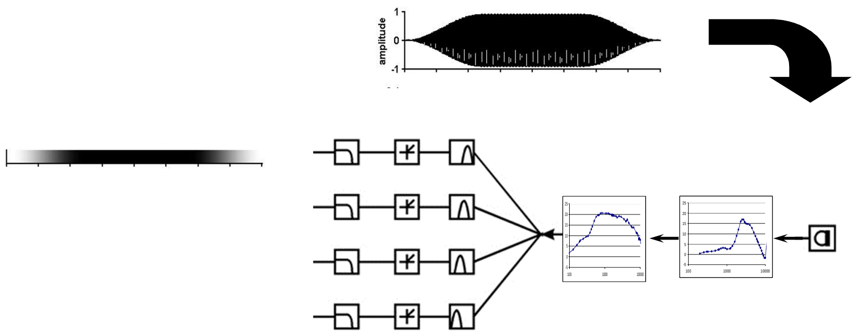
input signal at 200 Hz



output display

# Construct the output display one strip at a time

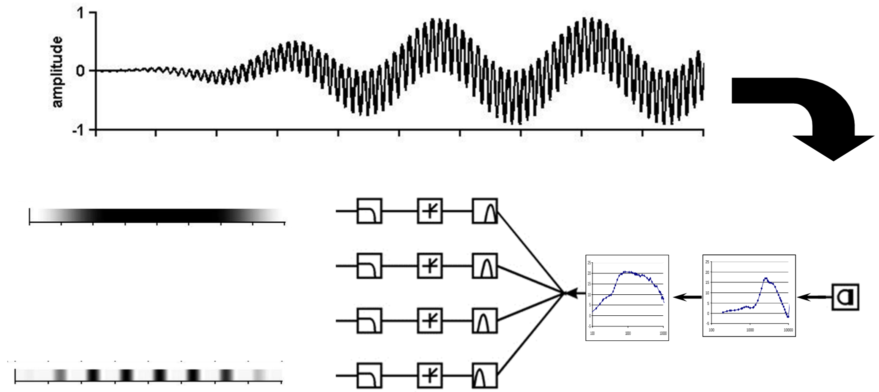
input signal at 4 kHz




output display



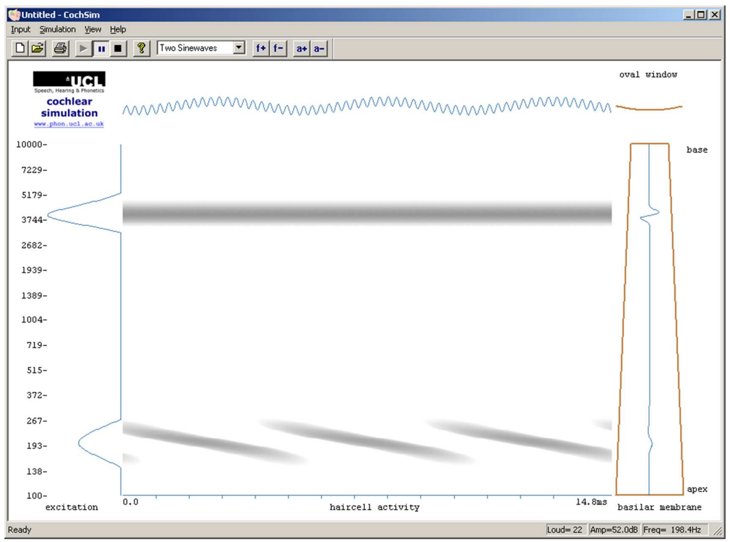
# 4 kHz + 200 Hz




output display



# 4 kHz + 200 Hz






# Auditory and ordinary spectrograms