

PERIPHEAR: A NONLINEAR ACTIVE MODEL OF THE AUDITORY PERIPHERY

Arnaud Robert¹ and Jan Eriksson²

¹CIRC Group, Swiss Federal Institute of Technology, Lausanne, Switzerland

²Physiology Institute, University of Lausanne, Switzerland *

ABSTRACT

This paper describes a phenomenological model of the auditory periphery which consists of a bank of nonlinear time-varying parallel filters. Realistic filter shapes are obtained with the all-pole gammatone filter (APGF) which provides both a good approximation of the far more complex wave-propagation or cochlear mechanics models and a very simple implementation. The model also includes an active, distributed feedback that controls the damping parameter of the APGF. As a result, the model reproduces several observed phenomena including compression, two-tone suppression and suppression of tones by noise.

1. INTRODUCTION

Understanding how spectral and temporal resolution can be achieved, in our hearing system, for a wide range of intensities and in different contexts, as encountered in real situations, is a major challenge.

In the peripheral auditory system the basilar membrane (BM), part of the cochlea, is the first main processing unit. Its dynamic behaviour is influenced by the outer hair cells (OHCs) which are believed to be involved in a complex active feedback mechanism which enhances the passive BM filtering properties and leads to BM nonlinear responses such as compression of dynamic range and suppression (related to psychoacoustic masking phenomena) by lateral interactions on the BM [4, 9]. Extensive quantitative and qualitative characterisation of mechanical frequency analysis achieved by the cochlea provides information with which to develop a cochlear model. As a whole, the BM can be regarded as a bank of overlapping asymmetrical nonlinear time-varying filters.

Many cochlear models have been proposed in the past (among others: [6, 1, 5]). The main improvement in recent years has been the introduction of nonlinear feedback loop meant to adjust some filtering parameter at the BM level. Their main limitations are a hard-to-set feedback parameters and restricted suppression area.

In order to cope with some of these limitations, we have developed a phenomenological model¹ which is simple, has an easy-tunable feedback process – thanks to the use of

simple filters – and incorporates influence of neighbouring BM filters.

2. THE COCHLEAR MODEL

2.1. Overview

In the model's design, the main concern was to allow a variation of filtering properties as a function of time and stimulus waveform and to obtain responses that were close to recorded BM responses. The external and middle ear complex is reduced to a single linear transfer function implemented as a bandpass (Butterworth) filter with low and high cutoff frequencies of 1.4 and 20 kHz, respectively.

The cochlea is modelled as a set of parallel sections: each local modelled BM section is composed of two band-pass filters: (a) a passive one with fixed parameters and (b) an active filter whose properties evolve with time in a stimulus dependent manner. The parameters tuning is done via a nonlinear feedback. As a result of the peripheral section design, complex waveform compression and suppression phenomena, observed in the cochlea, are intended to be represented by this model.

2.2. The basilar membrane

Each BM section is composed of a filtering process tuned to a specific characteristic frequency CF (representing a location on the BM) and includes an active feedback. The diagram of one such BM section, where transformation of an incoming sound into a BM displacement is achieved, is shown in figure 1.

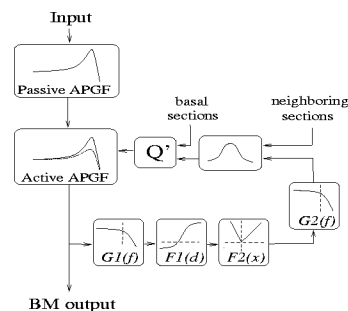


Figure 1: Model of one BM section.

For the filtering process we decided to use the all-pole gamma tone filter (APGF), thoroughly analysed in [7]. Their equations are derived by discarding zeros from the gamma-tone filter, known for providing an excellent fit

*This work is supported by the Swiss National Scientific Foundation, grant number 2150-045689.95.

¹the adjective phenomenological, opposed to biophysically inspired, means that we seek to reproduce cochlear-like responses without explicit guidance from detailed anatomical and physiological data

to the impulse response of auditory nerve fibres [7]. APGFs present many advantages such as (a) a small number of free parameters (the filter can be characterised by its damping parameter Q and the natural frequency w_n), (b) a better controlled behaviour of the frequency response tail, (c) an easier way to model level-dependent gain, bandwidth, asymmetry and centre-frequency shift.

Let us just recall two main equations of the APGF, its gain response H (Eq. 1, for a N th order APGF) and the relation between the quality factor Q and the maximal gain value of H (Eq. 2). Figure 2 (left) shows $H(w)$ for different values of Q and for a given characteristic frequency of $CF = 5\text{kHz}$.

$$|H(w)|^2 = \frac{1}{\left(1 + \left(\frac{1}{Q^2} - 2\right) \cdot \left(\frac{w}{w_n}\right)^2 + \left(\frac{w}{w_n}\right)^4\right)^N} \quad (1)$$

$$Q = Hmax^{\frac{1}{N}} \cdot \sqrt{0.5 \cdot \left(1 + \sqrt{1 - Hmax^{\frac{-2}{N}}}\right)} \quad (2)$$

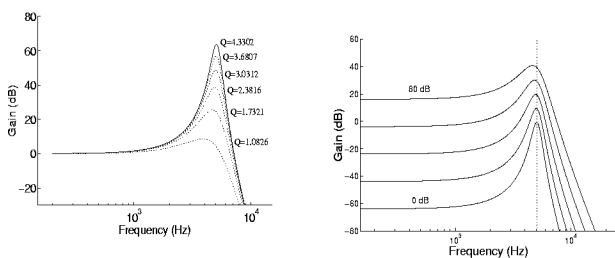


Figure 2: Responses of a APGF for different values of Q (left) and combined passive and active APGFs (right). $CF=5\text{kHz}$.

The use of two cascade APGFs, one passive and one active (whose damping factor is changed via the feedback loop) is motivated by the fact that in the BM, the filters center frequencies change with stimulus level. The responses of the cascade of the passive and active filters are given for different stimulus intensities, from 0 to 80 dB, in figure 2(right).

2.3. The Feedback

The goal of the feedback loop is to better model the BM nonlinear behaviour by continuously modifying the active filter's parameters in a way dependent on the stimulus temporal and spectral properties. We based the feedback, in part, on physiological data of OHC functioning [2, 10, 4].

The feedback loop, shown in figure 1, is composed of (1) an OHC model consisting of (1.a) a low-pass filter $G1(f)$, (1.b) an asymmetric saturation, (1.c) a rectification function, (1.d) a smoothing filter $G2$; (2) a Gaussian summation of outputs from neighbouring sections and (3) a computation of the effective quality factor. The principle is similar to the one used in [1].

The presence of the first low-pass filter ($G1$) with cut-off frequency below the section's characteristic frequency (CF), gives more weight to frequencies lower than CF . We used an order 4 APGF with low quality factor ($Q = 0.8$) and natural frequency of $w'_n = w_n/1.4$, where w_n is the natural frequency of the BM section considered.

To correctly model the compressive BM nonlinearity, the feedback must have no effect at low sound pressure levels (SPLs), increasing effect for medium to high SPLs (from 30-40 to 90 dB) and saturate at higher level. This is taken care of by the asymmetric saturation which is related to the OHC membrane potential. Its modelling equation is :

$$F_1(d) = \frac{c_1}{(1 + e^{(d+c_2)/c_3}) \cdot (1 + e^{(d+c_4)/c_5})} - c_6 \quad (3)$$

where c_n are fixed parameters and d is the BM displacement.

The rectification function expresses the fact that deflections in both directions are affected by the compressive nonlinearity. Rectification is given by:

$$F_2(x) = \begin{cases} \frac{k}{\min(F_1)} \cdot y & \text{if } y < 0, \\ \frac{k}{\max(F_1)} \cdot y & \text{if } y > 0. \end{cases} \quad (4)$$

where $y = F_1(d)$. The parameters k is a scaling factor that is adjusted to produce the desired compression and is slightly greater for those sections with high gains. The output of the rectifier is smoothed using a low-pass filter ($G2$) (first order Butterworth with cut-off frequency of $f = 1500\text{ Hz}$). The output F_2 is considered as representing a reduction in the gain of the active filter (its units are in dB), in other words:

$$\log(Hmax_{i,t}) = \log(Hmax_{i,0}) + F_2 \quad (\text{dB})$$

where $Hmax_{i,0}$ is the initial value of $Hmax_i$, index i referring to the considered BM section. A local quality factor $q(i, t)$ can be derived from $Hmax_i$ using Eq. 2. The decision to convert this output into local quality factor at this stage was rather arbitrary. It may have been done before computing the *effective* quality factor (see below). However, the results are likely to have been only marginally different since the parameters for neighbouring sections vary smoothly and can be considered equal locally.

A "local average" of the outputs $q(i, t)$ of neighbouring sections is then computed according to:

$$\bar{q}(i, t) = \sum_{k=i_{base}}^{i_{apex}} e^{-\left(\frac{x(CF_k) - x(CF_i)}{0.33}\right)^2} \cdot q(k, t)$$

where the gaussian weighing function is expressed in terms of the distance x from the base of the cochlea. We use $x = 1/0.06 \cdot \log(f/456 + 0.8)$ where x is in mm and f in Hz. It is centered at $f = CF_i$ and has a "distance" standard deviation of 0.33mm . This "smoothing" of the feedback factor may be said to reflect either longitudinal coupling or the effect of efferent feedback from the brainstem involving neurons with less sharp tuning than afferent fibers.

If we consider the response to a stimulus as a traveling wave, it is appropriate to also consider the fact that the output of one BM section is the input to the next. If the gain of a section basal to section i is reduced, it will influence the output of all stimulus frequencies including $f = CF_i$. This influence is measured by δ_k , a gain modification factor. Using Eq. (1), we can derive δ_k as the logarithmic difference between H_k evaluated at frequency $f = CF_i$ for a quality factor $Q = \bar{q}_k$, and H_k evaluated at the same frequency but for its initial quality factor $Q_{k,0}$. We have:

$$\delta_k = \log\left(\frac{H_k(CF_i; \bar{q}_k, t)}{H_k(CF_i; Q_{k,0})}\right) \quad (\text{dB})$$

Adding the gain modifications of basal sections we obtain:

$$\log(H' \max_{i,t}) = \log(H \max_{i,0}) + \sum_{k \in S} \delta_k \quad (5)$$

The set S includes all BM sections that are *basal* to the section i . Hence, the *effective* gain is the sum of small gains in sections basal to i . The *effective* quality factor of the BM section, Q'_i , is deduced from Eq. (2). Its value will determine the active filter's properties - its bandwidth and gain.

2.4. IHC-AN synapse

The transformation of the mechanical basilar membrane filtering into a firing pattern in the auditory nerve (AN) is accomplished by the inner hair cells (IHCs), as each BM section's response excites an ensemble of IHCs. We use the three reservoirs model of inner hair cell/auditory nerve synapse proposed by Meddis [8], that reproduces several characteristics of auditory nerve firing patterns. We did however modify the permeability factor, $k(t)$, an instantaneous nonlinear function of the filtered BM displacement, adding a quadratic term to model low spontaneous rate IHC/AN synapses (see [3] for a discussion). Thus, we have:

$$k(t) = k_0 + k_1 \cdot \max(0, V_{IHC}(t) - V_0) + k_2 \cdot (\max(0, V_{IHC}(t) - V_0))^2 \quad (6)$$

Having free parameters, we enable modelling of auditory nerve fibers with different characteristics. As an example, fibers with high (ANH) and low (ANL) spontaneous rates have been obtained, with values of respectively 60 and 2 spikes/sec.

Spike generation. The output of the IHC model provides a time-varying function for probabilistic generation of excitatory post-synaptic potentials (EPSPs) in auditory nerve fibers, achieved through a uniform random number generator. Threshold was set so that an EPSP would generate a spike with 100% probability except when the cell was in a refractory state. Absolute and relative (decreasing exponential) refractory periods were included, respectively fixed to $\tau_{abs} = 0.7ms$ and $\tau_{rel} = 10ms$.

3. RESPONSES OF THE MODEL

Stimuli. Stimuli consisted of pure tones of different frequencies and intensities. The full stimulus duration is 250 ms, decomposed as follows: 100 ms of silence, followed by a signal during 100 ms, and finally 50 ms of silence again. The signal had an on/off ramp of 2.5 ms.

3.1. BM Tuning curves

The locus of frequency-intensity combinations of tones that cause a just measurable increase in BM displacement is known as the *threshold tuning curves*. In our simulations the tuning curves were obtained by stimulating each BM section by pure tones of varying frequencies and levels. The result is shown at figure 3 and is quite in accordance with physiological data.

3.2. Compression

Compression is the phenomena by which the intensity dynamic range of the input signal is reduced to a smaller response dynamic range at the cochlear level. Compression

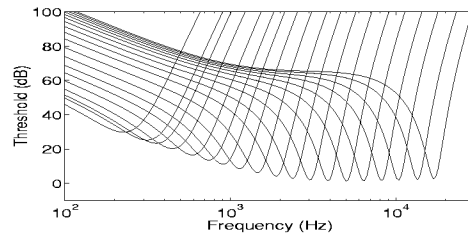


Figure 3: Threshold tuning curves.

is known to result from the time-variations of the BM filters gains as a function of sound level.

As could be expected, the model, by including a nonlinearity in its feedback loop that influences the filters' tuning properties, offers a good representation of the compressive behavior. This is shown in figure 4, for tones of different frequencies, recorded at the CF location, and for stimulus levels ranging from 0 dB to 100 dB. Compression factors of up to 40 dB were simulated, just as in the real cochlea.

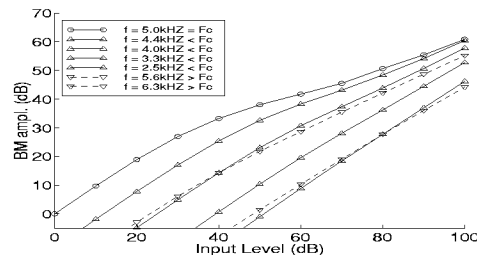


Figure 4: Compressive responses of the BM. CF=5kHz.

3.3. Suppression

Two-tone suppression is the reduction of the BM displacement in the region most sensitive to a probe tone by the addition of a second (suppressor) tone at a different frequency. It is attributed to the saturation of the active mechanism in the cochlea. Masker tones of varying frequency and intensity delivered simultaneously with the probe can result in both excitatory and suppressive masking of the probe. In the model, we expect to occur because of the saturating nonlinear feedback and the addition of influence from neighbouring sections.

In order to show the model's behaviour in suppression conditions, a probe tone of 8 kHz was delivered at 40 dB simultaneously with masker tones at varying frequencies (200-20kHz) and intensity (from 30 to 90 dB in 5 dB steps). The results are presented at figure 5. The change in amplitude is recorded; positive and negative amplitudes represent excitatory and suppressive masking, respectively. Below CF, first suppression then eventually excitation can be caused by frequencies much lower than CF if of sufficient intensity, whereas above CF suppression is limited to frequencies less than about 1.5 octaves above CF.

Results observed in physiological recordings are well represented in our model: (1) for high masker levels, around CF, the BM response increases since it now primarily responds to the masker (no suppression occurs), (2) at equal levels, suppression is stronger for below-CF masker than for above-CF ones, (3) for low to medium masker intensity, suppression is strong near CF (4) high-side (masker's

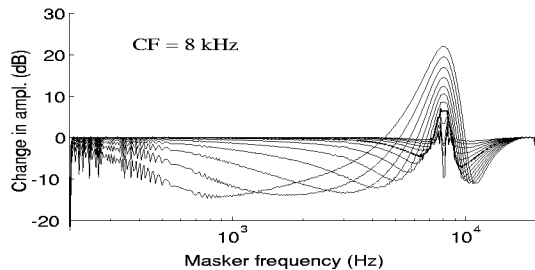


Figure 5: Suppressive and excitatory masking. CF=8kHz.

frequency is above CF) suppression drops off more rapidly than does low-side (masker's frequency is below CF) suppression. Also, the addition of influence from neighbouring sections is expected to enable better modelling of suppression by wide-band signals such as noise.

A frequency-intensity response map for pure tones with background noise (masked response area) is shown at figure 6 for two ANFs with high (ANH) and low (ANL) spontaneous firing rates. The level of the background noise was adjusted to evoke a rate of activity intermediate between baseline and saturation, i.e. 50 and 60 dB for ANH and ANL fibers. Each point on the surface of these plots represents the percentage change in mean firing rate when the pure tone stimulus is added to the noise background. Regions of suppressed activity are indicated by lighter areas, that is pure tones at these frequencies will cause a decrease in the mean firing rate in response to the noise. Both high- and low-frequency suppression regions were obtained, most evident in the ANL fibers on the right. For high-spontaneous rate fibers, suppression by high-frequency tones was more pronounced than suppression by low-frequency tones.

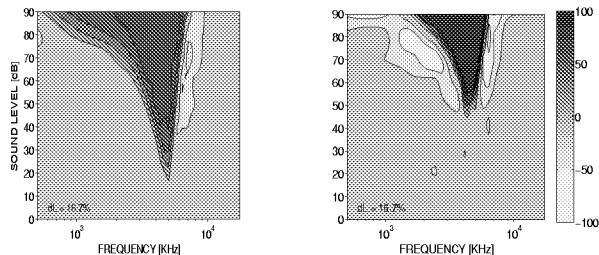


Figure 6: Masked response areas for two ANFs (CF=5kHz). Left: ANH, right:ANL. dL is the percentage change between two contour curves.

4. DISCUSSION

General summary. We have presented in this paper a nonlinear active model of the cochlea whose design was influenced by physiological observations on BM filtering and OHC functioning. The basilar membrane is modelled as an ensemble of parallel filtering sections, each composed of two narrow-band filters, a passive and an active one. The latter's filtering properties are set (via its Q) by a feedback loop in which neighbouring BM sections also contribute. The model is presently used to study responses to complex stimuli in models of the auditory nerve and cochlear nucleus neurons and to provide physiologically plausible front-end for speech analysis.

Model responses. Model responses presented in this

paper – tuning and compression curves, suppression phenomena – are quite in accordance with physiological data. The use of two APGFs in each BM sections, the nonlinearity in the model's feedback process and the distributed form of feedback have all contributed to render possible the simulation of various known BM responses.

In our feedback implementation, added to the commonly used saturating function, two important elements enable better modelling of the suppression phenomena. First, the low-pass filter $G1$ provides a higher threshold for suppression than for excitation at CF. This would correspond to the intuitive reasoning that the area of the BM where positive feedback is strongest lies just basally from the point of maximal response and, hence, that this is where the saturation threshold is lowest. Second, the distributed form of feedback has improved modelling of the suppression phenomenon, particularly to broadband signals and to lower than CF frequencies.

Comparison to other models. The main differences between the presented model and those of, among others, [6, 1, 5] are the use of APGFs and the distributed feedback. We hope that these new features will enable a better modelling of the AN responses to more complex stimuli such as speech.

Further work. Among other things, we believe the suppression modelling could be further improved by enabling fast suppression effects that seem to exist[9]. This is hindered in our model by the $G1$ low-pass filter.

5. REFERENCES

1. L. H. Carney. A model for the responses of low-frequency auditory-nerve fibers in cat. *Journ. Acoust. Soc. Am.*, 93:401–417, 1993.
2. P. Dallos. Response characteristics of mammalian cochlear hair cells. *Journal of Neuroscience*, 5:1591–1608, 1985.
3. D. C. Geisler and S. Greenberg. A two-stage nonlinear cochlear model possesses automatic gain control. *Journ. Acoust. Soc. Am.*, 80:1359–1363, 1986.
4. D. C. Geisler, G. K. Yates, R. B. Patuzzi, and B. M. Johnstone. Saturation of outer hair cell receptor currents causes two-tone suppression. *Hear. Res.*, 44:241–256, 1990.
5. C. Giguère and P. C. Wooland. A computational model of the auditory periphery for speech and hearing research. I. Ascending path. *Journ. Acoust. Soc. Am.*, 95:331–342, 1994.
6. J. L. Goldstein. Modeling rapid waveform compression on the basilar membrane as multiple bandpass nonlinearity filtering. *Hear. Res.*, 49:39–60, 1990.
7. R. F. Lyon. The all-pole gammatone filter and auditory model. In *Forum Acusticum '96, Antwerp, Belgium*, April 1996.
8. R. Meddis. Simulation of mechanical to auditory-neural transduction: Further studies. *Journ. Acoust. Soc. Am.*, 83:1056–1063, 1988.
9. M. A. Ruggero, L. Robles, and N. C. Rich. Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. *Journal of Neurophysiology*, 68:1087–1099, 1992.
10. I. Russel, A. Cody, and G. Richardson. The responses of inner and outer hair cells in the basal turn of the guinea-pig cochlea and in the mouse cochlea grown in vitro. *Hear. Res.*, 22:199–216, 1986.