

COVER PAGE

Document downloaded by @DAEL

Sat Nov 28 06:23:39 2026
For personal use

When automatic English translation is provided, only the original document is authentic.
The EAA cannot be held responsible of any translation error

Bibliographical reference

A Nonlinear Transmission Line Model of the Cochlea With Temporal Integration Accounts for Duration Effects in Threshold Fine Structure, Jesko L. Verhey, Manfred Mauermann and Bastian Epp, Acta Acustica **vol. 103** (Number 5), 2017, pp. 721-724

DOI
https://doi.org/10.3813/AAA.919100

Acta Acustica united with Acustica
Vol. 103 (2017) 721 – 724

DOI 10.3813/AAA.919100

A Nonlinear Transmission Line Model of the Cochlea With Temporal Integration Accounts for Duration Effects in Threshold Fine Structure

Jesko L. Verhey[1)], Manfred Mauermann[2)], Bastian Epp[3)]

[1)] Department of Experimental Audiology, Otto von Guericke University Magdeburg, Leipziger Str. 44, 39120 Magdeburg, Germany. email: jesko.verhey@med.ovgu.de

[2)] Medical Physics and Cluster of Excellence Hearing4All, University of Oldenburg, 26111 Oldenburg, Germany.

[3)] Hearing Systems group, Department of Electrical Engineering, Technical University of Denmark, Ørsteds Plads Building 352, Kongens Lyngby, Denmark

Summary

For normal-hearing listeners, auditory pure-tone thresholds in quiet often show quasi periodic fluctuations when measured with a high frequency resolution, referred to as threshold fine structure. Threshold fine structure is dependent on the stimulus duration, with smaller fluctuations for short than for long signals. The present study demonstrates how this effect can be captured by a nonlinear and active model of the cochlear in combination with a temporal integration stage. Since this cochlear model also accounts for fine structure and connected level-dependent effects, it is superior to filter-based approaches and hence allows the investigation of the contributions of cochlear- and retro-cochlear processing on behavioural data, including stimulus-duration dependent effects of threshold fine structure.

© 2017 The Author(s). Published by S. Hirzel Verlag · EAA. This is an open access article under the terms of the Creative Commons Attribution (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Normal-hearing listeners often show a quasi periodic fluctuation in threshold level when thresholds are measured with a high frequency resolution. This is commonly referred to as microstructure [1] or fine structure [2] of the threshold in quiet. Other psychoacoustical effects such as binaural diplacusis have been argued to be linked to this threshold fine structure [3]. Since a minimum in threshold fine structure often coincides with the frequency of spontaneous otoacoustic emissions [4] it is assumed that the fine structure and the spontaneous otoacoustic emissions originate from the same cochlear mechanism. A nonlinear cochlear transmission-line model can predict these effects [5].

Threshold fine structure depends on duration. Cohen [6] showed that the threshold versus duration curve was steeper for a signal with a frequency coinciding with a minimum of the threshold fine structure (trough frequency) than for a signal with a frequency coinciding with a maximum of the threshold fine structure (peak frequency).

Received 13 January 2017,
accepted 14 July 2017,
published online 2 August 2017.

Cohen [6] argued that the different slopes for the threshold curves were due to the energy spread associated with decreasing signal duration. According to this explanation, the energy spread due to the shortening of the signal to surrounding frequencies could not be used when the frequency was equal to a trough of the fine structure since thresholds were higher for the surrounding frequencies. In contrast, for a peak frequency, as the bandwidth widens with decreasing signal duration, the energy spread to frequencies for which the detection threshold was lower than that of the signal frequency, i.e., they contributed to detection threshold.

This explanation assumes that each frequency (associated with a specific cochlear place) is linked to a certain sensitivity for this frequency. This can be realized by assuming a frequency-dependent gain of the cochlear amplifier with a higher gain at fine-structure minima than at maxima, i.e., a fine-structure filter. However, fine structure effects on modulation detection thresholds are difficult to reconcile with this linear filter approach. Modulation detection thresholds were higher for a carrier frequency equal to the peak of the threshold fine structure than for a carrier frequency equal to the trough of the threshold fine structure [7]. The size of the effect depended on the carrier level being highest when the carrier level was close to the threshold in quiet. This level-dependence cannot be predicted on the basis of a linear filter approach.

© 2017 The Author(s). Published by S. Hirzel Verlag · EAA.
This is an open access article under the terms of the CC BY 4.0 license.

Thus, a more realistic model of cochlear mechanics is required. Epp *et al.* [5] showed that a realistic cochlear model predicts fine structure effects, including those found in modulation detection experiments.

The present study used the cochlear model of [5] as a basis for the prediction of the effect of duration on temporal fine structure. The aim of the present study was to investigate the ability of the model to account for duration effects in threshold fine structure, and to what extent the effect was cochlear or retro-cochlear in origin. To this end, experimental data on threshold fine structure measured in dB SPL were compared to model predictions using the cochlear model in it's original implementation, and in combination with a leaky integrator as a retro-cochlear process.

2. Methods

2.1. Psycoacoustical experiment

Six normal-hearing subjects with audiometric thresholds < 15 dB HL at octave frequencies between 0.125 kHz to 8 kHz participated in the experiments. Subjects were seated in a double-walled sound attenuating booth. Signals were presented via Sennheiser HDA200 headphones. Before measuring thresholds as a function of stimulus duration, the individual fine structure was measured in various octave bands until a region showing sufficient fine structure was determined for each subject. As a screening procedure, the modified von Békésy tracking method of [2] was used. Adjacent peak and trough frequencies between 1 and 3 kHz were selected for each subject and used in the following temporal-integration experiments.

In the temporal-integration experiment, thresholds in quiet were measured using an adaptive three-interval three-alternative forced-choice paradigm with an one-up two-down rule [8]. Consecutive intervals of a trial were separated by 400 ms of silence. One randomly chosen interval contained the signal. The task of the subject was to indicate the interval containing the signal. Depending on the response of the subject, the signal level for the next trial was increased or decreased. The initial step size was 6 dB. After the first upper reversal it was reduced to 3 dB and after the second upper reversal to 1 dB. The adaptive run continued for another eight reversals with this minimum step size. The mean of the levels at these eight reversal was taken as threshold estimate. The final estimate of the individual threshold was the mean of the threshold estimates of three runs.

Signals were gated with linear 4-ms ramps at onset and offset. Signal duration was 8, 16, 32, 64, 128, 256, or 512 ms. The signal duration was defined using the 50% down points of the ramps.

2.2. Simulations

A nonlinear and active transmission line model of the cochlea with 1000 segments representing discrete cochlear partitions (CP) was used [5]. The active process in the cochlea was implemented as a combination of velocity-dependent damping and feedback stiffness in the equation of motion for each CP. Note that the implementation of the model in its current form only phenomenologically describes the macromechanics of the BM and does neither model, nor allow to infer physiologically realistic details about cochlear micromechanics. To account for fine structure effects, a random variation of the place-frequency map (roughness) was introduced. This approach leads to self-sustained oscillations of the CPs and has successfully been used to predict fine structure of the threshold in quiet [5]. To this end, a velocity threshold was assumed. When the threshold was reached at any CP, the excitation was considered to be sufficiently strong to detect the signal. This "velocity-threshold model" was used here to investigate to what extent cochlear processing contributes to the effect of stimulus duration. In addition to the velocity-threshold model, four modified versions of the model (referred to as "integrator models") were used. In the integrator models, a temporal integration stage was added to process the nonlinearly transformed cochleogram, i.e., the temporal output of each CP of the simulated cochlea over time: the absolute value of the cochleogram was raised to an exponent α, summed over all CPs c_j and low-pass filtered using an integration window with a time constant τ. The four versions of the integrator model differed with respect to the time constant of the integration window (τ , 100 or 200 ms) and the exponent used to process the absolute value of the segment velocity (α, 2 or 3). The two time constant have both been used for modeling of psychoacoustical data [9]. Temporal integration models commonly assume an integration of the intensity (i.e., $\alpha = 2$). Some authors, however, argue that an exponent of 3 provides a better prediction of temporal integration data [10]. To show the influence of this parameter on the predictions, both values were included in the simulations. The summation over the output of all CPs c_j represents an across-frequency integration. The temporal window was implemented as a low-pass filter with an exponentially decaying impulse response,

$$w_\tau(t) = \frac{1}{\tau} \mathrm{e}^{-t/\tau}. \tag{1}$$

The impulse response $\mathrm{w}_\tau(t)$ was sampled and truncated to a length of five times the time constant τ to obtain the discrete impulse response $w_\tau[n]$. Low-pass filtering was realized as

$$c_{\tau,\alpha}[n] = \left(\sum_{j=1}^{1000} \left| c_j[n] \right|^\alpha \right) * w_\tau[n], \tag{2}$$

where $c_{\tau,\alpha}[n]$ respresents the output of the integrator, $c_j[n]$ the velocity of the CP with index j, $w_\tau[n]$ the integration window with time constant τ. The operator $*$ denotes filtering including appropriate initial conditions to avoid onset effects in the filtering process. The initial conditions were obtained as the final state of the filter after processing a one-second period of self-sustained activity of the simulated CPs (without external stimulation). The thresholds

in quiet were simulated in each of the four investigated integrator models by assuming that a stimulus signal was detected if:

$$\left| 1 - \frac{c_{\tau,\alpha}[n]}{\kappa} \right| \leq \delta \quad \text{for} \quad n \in \left[0, f_s \cdot T - 1\right], \qquad (3)$$

i.e., if the value of $c_{\tau,\alpha}[n]$ approached a critical value κ within a tolerance of $\delta = 3\%$ within the duration T of the stimulus. The value of κ was derived empirically as a scaled version of the output of the model integration stage after processing a one-second period of self-sustained activity (without external stimulation). The thresholds of the models were determined by varying the stimulus level in steps of 0.25 dB until the criterion value κ was reached. The signals were the same as in the psychoacoustical experiment.

3. Results

Figure 1 shows the level at threshold as a function of signal duration. Each panel shows individual data for the trough frequency (filled circles) and adjacent peak frequency (open circles) of the individual threshold fine structure. Thresholds are given in decibels sound pressure level (dB SPL) of the signal prior to signal gating. For all subjects, thresholds decreased as the duration increased. The effect was larger for the trough frequency than for the peak frequency, in qualitative agreement with the data of [6]. Individual differences were observed with respect to the size of the effect and the difference between the two threshold curves. The maximum threshold difference between the two frequencies ranged from 6 dB (S6) to 11 dB (S1).

Panel A of Figure 2 shows the average threshold-duration curves for the six subjects. The mean threshold for the trough and the peak frequencies were about the same (14 dB) at a signal duration of 8 ms. For the peak frequency, thresholds decreased by about −2 dB per doubling of the duration down to 2 dB. In the range from 8 to 32 ms, the slope of the threshold curve for the trough frequency was considerably steeper (about -4 dB per doubling of duration) than for the peak frequency. A similar slope for the threshold curves of the two frequencies was observed for durations larger than 64 ms. The minimum threshold was −4.5 dB for the trough frequency and a duration of 512 ms. A two-way ANOVA (SPSS) was used to examine the effect of stimulus duration ("duration", seven levels) and position with respect to the threshold fine structure ("position", two levels: peak or trough). The effect of "duration" on the thresholds was significant ($F(6, 30) = 488.3$, $p < 0.001$). The factor "position" also had a significant effect on the thresholds ($F(1,5) = 30.4$, $p < 0.01$). There was a statistically significant interaction between the effects of interaction of "position" and "duration" on the thresholds ($F(6,30) = 21.8$, $p < 0.001$), i.e., the position with respect to the threshold fine structure had a significant effect on the shape of the threshold curve.

Panel B of Figure 2 shows simulated thresholds using the velocity-threshold model. Thresholds for the peak and

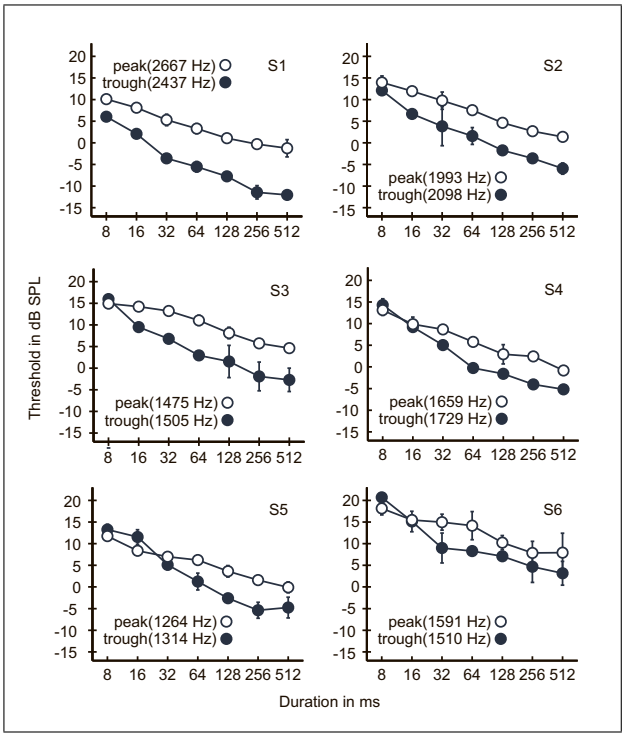


Figure 1. Detection thresholds of pure tones with signal frequencies corresponding to peaks (open circles) and troughs (filled circles) of threshold fine structure. Each panel shows the detection thresholds for one subject and the intraindividual standard deviations. Apart from the subject abbreviation, each panel also contains the peak and trough frequencies for the subject.

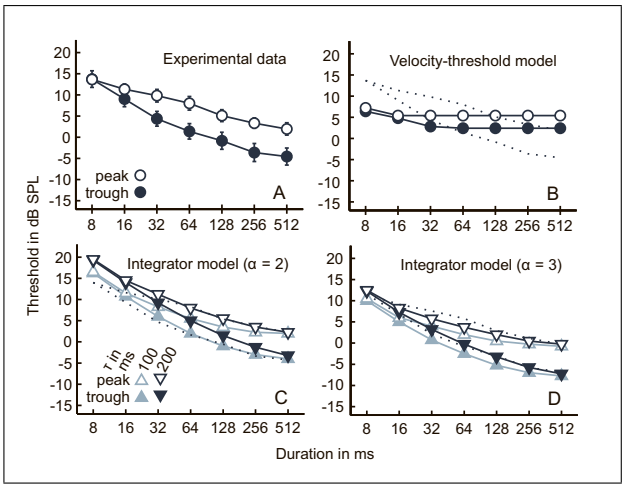


Figure 2. Same as Figure 1 but showing average detection thresholds and interindividual standard errors (panel A) and model predictions using different model versions (for details see text). The mean data of panel A is redrawn in panels B-D (dashed lines). For better comparison of the data and simulations, the mean data are normalized to the predictions for 512 ms duration and τ = 200 ms.

trough frequencies were about the same for the 8 ms signal duration. For the peak frequency, thresholds were constant for durations of 16 ms and longer. For the trough frequency, thresholds decreased up to a duration of 128 ms and did not depend on duration for longer signals. The

difference in thresholds between the peak and the trough frequencies was determined by the simulated fine structure at these frequencies. Panels C and D of Figure 2 show the predictions using the integrator models with values of $\alpha = 2$ (panel C) and $\alpha = 3$ (panel D) in combination with time constants of $\tau = 100$ (upward pointing triangles) and $\tau = 200$ *ms* (downward pointing triangles). The models predicted, consistent with the data, highest thresholds for short durations, and decreasing thresholds for increasing durations. The decrease in thresholds from 8 ms to 512 ms duration was larger for the trough frequencies than for the peak frequencies. The differences between thresholds for the trough and the peak frequencies at each duration increased with increasing stimulus duration, asymptotically approaching the value of the threshold fine structure. For the exponent $\alpha = 2$, the initial decrease between 8 and 16 ms was slightly steeper, and the decrease as a function of duration showed a larger curvature than the measured data. The initial decrease in the predictions was mainly determined by the exponent, being shallower for an exponent of $\alpha = 3$ than for an exponent of $\alpha = 2$. The curvature was determined by the time constant τ. A longer time constant led to a higher decrease in threshold with durations between 16 and 512 ms than a shorter time constant. Simulated thresholds for a time constant of $\tau = 100$ ms (light blue symbols) were slightly higher than for $\tau = 200$ ms (dark blue symbols), and the predicted decrease of thresholds was steeper for the models using $\alpha = 2$ (panel C) than for the models using $\alpha = 3$ (panel D).

4. Discussion

The simulations (cf. Figure 2) showed that differences in the slope between trough and peak frequencies at short durations (8 to 32 ms) were accounted for by the cochlear model. However, for the steady decrease of thresholds with increasing duration, a retro-cochlear temporal-integration stage was required. This difference in slope between trough and peak frequency was more pronounced in the simulations than it was in the data. This might indicate that the used temporal integration stage was oversimplified. With the current choice of the integration window and the values of α and τ, the combination of a time constant of $\tau = 200$ ms and and $\alpha = 3$ provided the best fit to the data.

One should keep in mind that the current approach of simulating temporal intergation with one leaky intergator might not be realistic. Recently, [9] showed that temporal intergation of loudness may be better modelled by assuming two parallel leaky integrators instead of a single leaky integrator. Some models of temporal integration did even not require a leaky integrator, but used probabilistic theories of sound detectability that explained detection by a probability accumulation over time [11]. Independent of the retro-cochlear processing, fine structure effects are inherent when using a model as described in [5] and hence the combined model should be able to account for the data.

In summary, the difference in the slope of the threshold curves for the peak- and trough frequencies could be explained by two elements: A) a mechanical processing stage including inherent properties of a physiologically plausible cochlear model to account for duration- and frequency dependent spread of excitation; and B) retro-cochlear temporal integration of the transformed output of the cochlea with a time constant of about 100-200 ms to account for the monotonic decrease.

Acknowledgement

We thank Heiner Lendzian for collecting the psychoacoustical data. This work was supported by the Deutsche Forschungsgemeinschaft and the Centre for Applied Hearing research (CAHR).

References

[1] E. Elliott: A ripple effect in the audiogram. Nature **181** (1958) 1076.

[2] S. J. Heise, J. L. Verhey, M. Mauermann: Automatic screening and detection of threshold fine structure. Int. J. Audiol. **47** (2008) 520–532.

[3] G. van den Brink: The relation between binaural diplacusis for pure tones and for complex sounds under normal conditions and with induced monaural pitch shift. Acustica **32** (1975) 159–165.

[4] E. Schloth: Relation between spectral composition of spontaneous oto-acoustic emissions and fine-structure of threshold in quiet. Acustica **53** (1983) 250–256.

[5] B. Epp, J. L. Verhey, M. Mauermann: Modeling cochlear dynamics: Interrelation between cochlear mechanics and psychoacoustics. J. Acoust. Soc. Am. **128** (2010) 1870–1883.

[6] M. F. Cohen: Detection threshold microstructure and its effect on temporal integration data. J. Acoust. Soc. Am. **71** (1982) 405–409.

[7] S. J. Heise, M. Mauermann, J. L. Verhey: Investigating possible mechanisms behind the effect of threshold fine structure on amplitude modulation perception. J. Acoust. Soc. Am. **126** (2009) 2490–2500.

[8] H. Levitt: Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. **49** (1971) 466–476.

[9] J. Hots, J. Rennies, J. L. Verhey: Modeling temporal integration of loudness. Acta Acust. united Ac. **100** (2014) 184–187.

[10] P. Heil, H. Neubauer: A unifying basis of auditory thresholds based on temporal summation. PNAS **100** (2003) 6151–6156

[11] P. Heil, H. Neubauer, M. Tetschke, D. R. F. Irvine: A probabilistic model of absolute auditory thresholds and its possible physiological basis. Adv. Exp. Med. Biol. **787** (2013) 21–29.