ACTA ACUSTICA UNITED WITH ACUSTICA
Vol. 104 (2018) 887 – 890

DOI 10.3813/AAA.919247

Auditory Filter Derivation at Low Levels where Masked Threshold Interacts with Absolute Threshold

Toshio Irino[1)], Kenji Yokota[1)], Toshie Matsui [1,2)], Roy D. Patterson[3)]
[1)] Faculty of Systems Engineering, Wakayama University, Japan. irino@sys.wakayama-u.ac.jp
[2)] Department of Computer Science and Engineering, Toyohashi University of Technology, Japan. tmatsui@cs.tut.ac.jp
[3)] Department of Physiology, Development and Neuroscience, University of Cambridge, UK. rdp1@cam.ac.uk

Summary
This paper reports a notched noise (NN) experiment with a large proportion of low-level, wide-notch conditions where masked threshold asymptotes to a low level that is near, but distinctly above, absolute threshold. The effects of this low level limit on the derivation of auditory filter shape are traditionally mitigated by the inclusion of an arbitrary constant, P_0, in the model of NN masking. We show that the threshold limit can be explained by assuming that there is a noise floor in the cochlea. This more physiological model of masking improves the reliability of filter shape fitting and, if the noise floor is stimulus-level dependent, the fit of the gammachirp (GC) auditory filter model is remarkably good.

PACS no. 43.66.-x

1. Introduction

In an early NN filter-shape experiment that included very wide notches, threshold asymptoted to a low level somewhat above absolute threshold [1]. In subsequent papers, to prevent these asymptotic thresholds from distorting the derived filter shape, some researchers simply avoided wide notch conditions when the noise spectrum level was below 40 dB SPL [2], and some included a threshold floor parameter, P_0, in the power spectrum model (PSM) of masking[3].

This paper reports a NN experiment with a large proportion of low-level, wide-notch conditions to test the assumption that the threshold limit is produced by a cochlear noise floor in the region of the signal. This modification makes it possible to include absolute threshold explicitly in the PSM of NN masking.

2. The experiment

A standard two-alternative, forced-choice, NN experiment was performed with noise spectrum levels ranging from a moderate 38 dB down to a very low −12 dB in 10-dB steps. The signal frequency f_s was 2.0 kHz. The width of the noise bands was $0.4 f_s$.

2.1. Notched-noise conditions

The normalized frequency distances from the signal to the nearer edges of the lower and upper noise bands $\{\Delta f_l/f_s, \Delta f_u/f_s\}$ were $\{0, 0; 0.1, 0.1;\ 0.2, 0.2;\ 0.4, 0.4;\ 0.3, 0.5;\ 0.5, 0.3\}$. The same notches were used at each spectrum level. Initially, threshold was measured for five noise spectrum levels between 38 and −2 dB in random order (for all notch widths and levels). Subsequently, threshold was also measured for −12 dB to ensure that the data set included a substantial number of thresholds in the region of absolute threshold.

2.2. Equipment and listeners

Eight normal-hearing (NH) listeners participated in the experiment after giving informed consent. They all had hearing levels (HLs) less than 20 dB between 125 and 8000 Hz. The experiment was approved by the local ethics committee of Wakayama University.

The stimuli were presented over headphones (PM-1, OPPO) via a USB interface (HA-1, OPPO) at a 48-kHz sampling rate and 24-bit resolution. The listeners were seated in a sound attenuated room.

Received 28 February 2018,
accepted 3 August 2018.

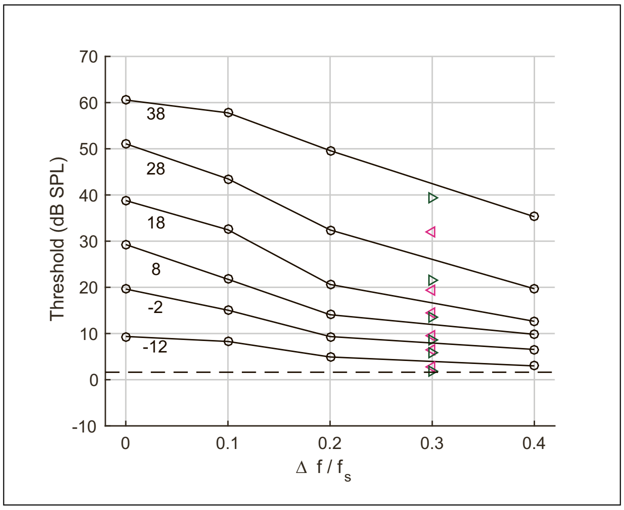

Figure 1. Average NN threshold (solid lines) for eight listeners, and their average absolute threshold (dashed line). The signal frequency was 2.0 kHz. The abscissa is normalized notch width ($\Delta f / f_s$). The circles (○) show symmetric notch conditions; the right-pointing triangles (▷) show conditions with additional shifting of the upper noise band by 0.2; the left-pointing triangles (◁) show conditions with additional shifting of the lower noise band by 0.2. The parameter beneath each threshold curve is noise spectrum level which was the same for the lower and upper bands. The noise levels for the triangles are the same as the levels of the threshold curves just above.

2.3. Results

Figure 1 shows average NN threshold for the eight listeners at the six masker levels (solid lines), along with the average absolute threshold (dashed line). The thresholds associated with the two highest noise levels, 28 and 38 dB, remain well above absolute threshold out to the widest notches. At lower noise levels (18, 8 and −2), however, threshold is limited by the proximity of absolute threshold, and threshold for the −12 dB noise level converges onto absolote threshold at the wider notch widths. The set of curves suggests that NN threshold should converge onto absolute threshold in the masking model. In earlier studies, absolute threshold was not directly represented in the model used to derive the shape and gain of the auditory filter.

3. Auditory filter derivation

The power spectrum model of masking produces an estimate of signal threshold, $\hat{P}_s$ (on a dB scale), using the following pair of equations,

$$\hat{P}_s = K + \hat{P}_{ext}, \tag{1}$$

$$\hat{P}_{ext} = N_0 + 10\log_{10}\left[\int_{f_{l_{min}}}^{f_{l_{max}}} W(f)\,\mathrm{d}f + \int_{f_{u_{min}}}^{f_{u_{max}}} W(f)\,\mathrm{d}f\right], \tag{2}$$

where K is the signal-to-noise ratio at the output of the auditory filter and $\hat{P}_{ext}$ is an estimate of the external noise that passes through the auditory filter. N_0 is the spectrum level of the noise in the bands that produce the notch, and $W(f)$ is the power weighting function of the auditory filter. When the auditory filter is modeled with the gammachirp, $G_c(f)$ [4], the filter weighting function, $W(f)$ becomes $|T(f)G_c(f)|^2$ where $T(f)$ is the transfer of sound from the free field or headphone to the input of the cochlea.

3.1. The P_0 threshold limit

Glasberg and Moore [3] introduced a term, P_0, to represent the lower limit of NN threshold and prevent it from distorting the representation of the tails of the auditory filter. In this case,

$$\hat{P}_s^{(P_0)} = 10\log_{10}\left\{10^{\hat{P}_s/10} + 10^{P_0/10}\right\}. \tag{3}$$

The coefficients of the auditory filter were estimated using the least-squared method to minimize the error between the measured thresholds, P_s, and the thresholds predicted by the model, $\hat{P}_s$; that is

$$\mathbf{c}_{gc}^{(P_0)} = \underset{\mathbf{c}_{gc}}{\operatorname{argmin}}\left\{\frac{1}{N}\sum_{i=1}^{N}(P_{s_i} - \hat{P}_{s_i}^{(P_0)})^2\right\}, \tag{4}$$

where $\mathbf{c}_{gc}^{(P_0)}$ is a vector of the GC coefficients, $\{b_1, c_1, f_{rat}^{(0)}, f_{rat}^{(1)}, b_2, c_2\}$, plus the constants $\{K\&P_0\}$. Glasberg and Moore showed that P_0 is effective in reducing rms error. They suggested that P_0 is related to absolute threshold but they did not explain the relationship in detail. This standard model of auditory filter derivation will be referred to as the "P_0 model" in what follows.

3.2. A cochlear-noise threshold limit

One method of incorporating absolute threshold into the power spectrum model of masking is to assume that there is some form of noise floor in the cochlea, $N_c(f)$. In that case, the power spectrum model has two components (Equation 5), as in the P_0 model (Equation 3), and the first component, $\hat{P}_{ext}$, is produced by the external noise (Equation 2). The second component, $\hat{P}_{int}$, is produced by an internal, cochlear noise, as in Equation (6).

$$\hat{P}_s = K + 10\log_{10}\left\{10^{\hat{P}_{ext}/10} + 10^{\hat{P}_{int}/10}\right\}, \tag{5}$$

$$\hat{P}_{int} = 10\log_{10}\left[\int_{f_{a_{min}}}^{f_{a_{max}}} \left|N_c(f)G_c(f)\right|^2 df\right]. \tag{6}$$

$G_c(f)$ is the filter function used in the iterative fitting process. $N_c(f)$ is the spectral distribution of the cochlear noise and the integral is performed over the human audible range. We assumed $N_c(f)$ is similar to the 0-dB HL function [5] which is close to that of the internal "self-generated noise" recently reported by Buss *et al.* [6]. But the selection of function would not affect the resulting fit because, in this study, the data were all gathered at one signal frequency (2 kHz). Any variation associated with different parameter values will be absorbed by a constant $N_c(f_{ref})$ associated with the reference frequency, f_{ref}.

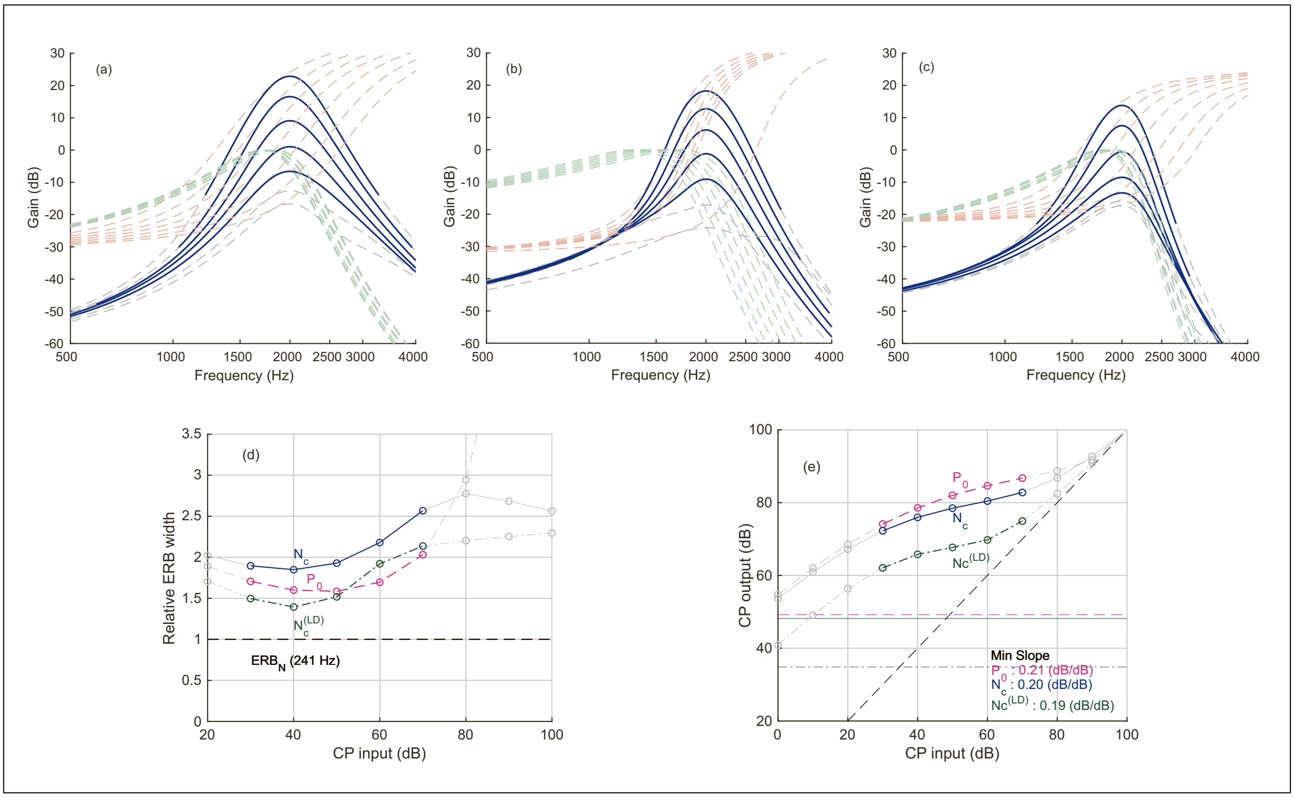

Figure 2. Upper row: GammaChirp auditory filter sets derived with (a) the N_c model, (b) the P_0 model, and (c) the $N_c^{(LD)}$ model from the 36 NN thresholds shown in Figure 1. The abscissa is warped to the $\mathrm{ERB_N}$ number scale but it is still labeled in frequency units (Hz). Filters are plotted for cochlear input levels between 30 and 70 dB in 10-dB steps. The thick solid lines show the GC filters; the thin dashed lines show the corresponding pGC and HP-AF components of the GC filter as detailed in [4]. Lower row: Relative bandwidths (d) and input-output functions (e) for the N_c , P_0, and $N_c^{(LD)}$ filter sets. CP (Cochlear Partition) input is the dB value at the input to the cochlea which is calculated from the dB SPL value at the ear drum and the middle ear transfer function. CP output is the internal excitation level when the input and output dB values are equalized at the CP input value of 100 dB.

In the absence of external noise, the masking expression (Equation 5) reduces naturally to an estimate of absolute threshold based on the PSM of masking; namely,

$$\hat{P}_{abs} = K + \hat{P}_{int}. \tag{7}$$

This effectively incorporates the observed absolute thresholds into the auditory filter fit. The fitting criterion is

$$\mathbf{c}_{gc}^{(N_c)} = \underset{\mathbf{c}_{gc}}{\operatorname{argmin}} \left\{ \frac{1}{N} \sum_{i=1}^{N} (P_{s_i} - \hat{P}_{s_i})^2 + (P_{abs} - \hat{P}_{abs})^2 \right\}, \tag{8}$$

where $\mathbf{c}_{gc}^{(N_c)}$ is a vector of gammachirp coefficients, $\{b_1, c_1, f_{rat}^{(0)}, f_{rat}^{(1)}, b_2, c_2\}$, plus the constants $\{K \& N_c(f_{ref})\}$. The process jointly minimizes the error associated with NN threshold and the error associated with absolute threshold, which allows this "N_c" model of masking to predict the continuous change from low-level NN threshold to observed absolute threshold. The N_c interpretation of the transition to absolute threshold provides a better foundation for the masking model than the ambiguous constant, P_0.

3.3. Comparison of N_c and P_0 fits

The N_c and P_0 models were initially compared in terms of the fits they provided to the full set of 36 NN thresholds shown in Figure 1. Each model was fitted to the 36 thresholds 10 times, using different initial values for the GC coefficients, chosen randomly within a range ±20% of the summary coefficient values reported in [4]. The best of the 10 filters was the one that minimized the rms error.

Figure 2a shows the best GC filter set for the N_c model. The rms error for the NN thresholds was 2.6 dB; the error for absolute threshold was 0.42 dB. Thus the N_c model of NN masking provides a good fit to both aspects of the data. Figure 2a shows that the gain and the selectivity of the GC filter increase as signal level decreases. The filter shape is more symmetric than that reported in [4].

Figure 2b shows the best GC filter set for the P_0 model. The rms error for the NN thresholds is 2.2 dB which is slightly less than that for the N_c model. The filter shape is similar to that reported in [4] where the P_0 model was used in the fitting process. The P_0 value is 3.9 dB above absolute threshold, indicating that the P_0 value cannot serve as an estimate of absolute threshold.

Figure 2d shows the bandwidths of the estimated GC filters. The bandwidth for the N_c model is about $2 \times ERB_N$ or more. The bandwidth for the P_0 model is narrower. Figure 2e shows the input-output functions. They have a broad compressive region between 30 and 70 dB where the minimum slopes are 0.20 dB/dB (N_c) and 0.21 dB/dB (P_0).

We also compared the N_c and P_0 models by testing whether they can predict the "unknown thresholds" when the fits were performed separately with the upper or lower halves of the full set of 36 NN thresholds in Figure 1 (see supplement for details). The results showed the N_c model provides more stable predictions than the P_0 model.

In summary, the N_c model can explain absolute threshold and provide stable estimation but the resultant filter shapes seem overly symmetric.

4. A level-dependent, cochlear-noise model

Finally, we constructed a model in which the cochlear noise level depends on the level of the external noise at frequencies distant from the signal. In this $N_c^{(LD)}$ model, the level of the internal noise spectrum, on a dB scale, varies linearly with the external NN level, N_0, as

$$N_{c(dB)}^{(LD)}(f) := N_{c(dB)}(f) + n_{LD} \cdot \left(N_0 - N_0^{(abs)}\right). \tag{9}$$

$N_{c(dB)}(f)$ is the cochlear-noise spectrum on a dB scale. n_{LD} is the slope of the level dependence. $N_0^{(abs)}$ is the N_0 level just below that which would affect absolute threshold; it is less than -12 dB and may be as low as -20 dB in the current experiment. The $N_c^{(LD)}$ model was fitted to the full set of thresholds shown in Figure 1. Figure 2c shows the auditory filters for the best $N_c^{(LD)}$ model, in which n_{LD} was employed as an optimization parameter and $n_{LD} \cdot N_0^{(abs)}$ was a constant fixed manually. The resultant GC filter has strong asymmetry with a steep high-frequency skirt. Most notably, the rms error for the fit to the 36 NN thresholds was 1.0 dB – much smaller than the rms error for either the N_c model (2.6 dB) or the P_0 model (2.2 dB). The error for absolute threshold was 1.3 dB, which is slightly greater than that of the N_c model (0.42 dB) and is less than that of the P_0 model (3.9 dB). The dashed-dotted line, labeled $N_c^{(LD)}$ in Figure 2d shows that the bandwidth values are less than those of the N_c model. The dashed-dotted line in Figure 2e shows that the input-output function has the expected broad compressive region between 30 and 60 dB; the minimum slope is 0.19 dB/dB. Interestingly, the estimated value of the slope n_{LD} in Equation (9) was 0.39 dB/dB, which is highly compressive.

These results indicate that the $N_c^{(LD)}$ model of NN masking provides an excellent fit which justifies the extra coefficient used to make the cochlear noise level dependent. The accuracy of the fit suggests that the hypothesis that the cochlear noise floor depends on the external noise level at distant frequencies warrants further investigation.

The source of the level dependence is not at all clear, however. One candidate could be distortion products emanating from the upper noise band [7]. Another candidate could be suppression [8]. It seems, however, unlikely because the internal representations of both signal and noise would be suppressed to the same degree and there would be no effect on the level dependence. Finally, it could be a nonlinearity involved with subsequent neural transduction, but that is beyond the scope of this paper. It is not clear, however, whether any of these factors could explain the form of the level dependence observed, that is, an n_{LD} value of 0.39 dB/dB.

5. Conclusions

This paper provides a detailed set of NN threshold values, including low-level noises, to show how threshold converges onto absolute threshold as notch width increases at low noise levels. To explain the data, the power spectrum model of masking was extended with the assumption that absolute threshold represents the effect of a noise floor in the cochlea which limits threshold at wide notch widths. This N_c model of masking was shown to provide better prediction of absolute threshold than the conventional P_0 model. However, the N_c-model filter shapes seem overly symmetric. When the cochlear noise floor was made level dependent, the resultant $N_c^{(LD)}$ model provided an excellent fit that appears to combine the merits of both the P_0 and N_c models, as well as a major reduction in the rms error of the filter-shape estimation. It is not clear, however, why the level of the cochlear noise should depend on the spectrum level of the external masking noise.

Acknowledgement

This work was partially supported by JSPS KAKENHI Grant Number JP16H01734.

References

[1] R. D. Patterson, I. Nimmo-Smith: Off-frequency listening and auditory-filter asymmetry. J Acoust Soc Am **67** (1980) 229–245.

[2] B. R. Glasberg, B. C. J. Moore: Derivation of auditory filter shapes from notched-noise data. Hear Res (1990) 103–138.

[3] B. R. Glasberg, B. C. J. Moore: Frequency selectivity as a function of level and frequency measured with uniformly exciting noise. J Acoust Soc Am 108 (2000) 2318–2328.

[4] R. D. Patterson, M. Unoki, T. Irino: Extending the domain of center frequencies for the compressive gammachirp auditory filter. J Acoust Soc Am (2003) 1529–1542.

[5] ANSI: ANSI S3.6-2010 Specification for audiometers, American National Standards Institute, New York, 2010.

[6] E. Buss, H. L. Porter, L. J. Leibold, J. H. Grose, J. W. Hall III: Effects of self-generated noise on estimates of detection threshold in quiet for school-age children and adults. Ear & Hearing **37** (2016) 650–659.

[7] B. C. J. Moore, B. R. Glasberg, M. van der Heijden, A. J. M. Houtsma, A. Kohlrausch: Comparison of auditory filter shapes obtained with notched-noise and noise-tone maskers. J Acoust Soc Am **97** (1995) 1175-1182.

[8] T. Houtgast: Psychophysical Evidence for Lateral Inhibition in Hearing. J Acoust Soc Am **51** (1972) 1885–1984.