Auditory nerve model for predicting performance limits of normal and impaired listeners

Michael G. Heinz

Speech and Hearing Sciences Program, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and Hearing Research Center, Boston University, Boston, Massachusetts 02215

mgheinz@alum.mit.edu

Xuedong Zhang

Hearing Research Center and Biomedical Engineering, Boston University, Boston, Massachusetts 02215

zxd@bu.edu

Ian C. Bruce

Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21205

ibruce@bme.jhu.edu

Laurel H. Carney

Hearing Research Center and Biomedical Engineering, Boston University, Boston, Massachusetts 02215

carney@bu.edu

Abstract: A computational auditory nerve (AN) model was developed for use in modeling psychophysical experiments with normal and impaired human listeners. In this phenomenological model, many physiologically vulnerable response properties associated with the cochlear amplifier are represented by a single nonlinear control mechanism, including the effects of level-dependent tuning, compression, level-dependent phase, suppression, and fast nonlinear dynamics on the responses of high, medium, and low spontaneous-rate (SR) AN fibers. Several model versions are described that can be used to evaluate the relative effects of these nonlinear properties.

PACS numbers: 43.64Bt

Date Received: February 9, 2001 **Date Accepted:** May 19, 2001

1. Introduction

Zhang *et al.*[1] have developed a physiologically based auditory nerve (AN) model (Fig. 1) that uses a nonlinear control path to account for broadened tuning with increases in level, including the associated compressive-magnitude and nonlinear-phase responses, suppression, and the fast dynamics of these properties. The use of a single control path to account for all of these physiologically vulnerable nonlinear response properties represents the finding that these properties are concomitantly absent in many forms of sensorineural hearing loss that are thought to be related to outer hair cell (OHC) dysfunction.[2] This phenomenological implementation was used here to develop an AN model for evaluating normal and impaired human psychophysical performance. The present model includes five versions, differing in the control-path implementation, which allow evaluation of the relative importance of sensitivity, spectral resolution, magnitude compression and level-dependent phase responses, and suppression.

The present model has a modified description of frequency resolution as a function of characteristic frequency (CF: most sensitive frequency of an AN fiber) that is derived from estimates of human rather than cat tuning. The present model extends the Zhang *et al.*[1] model by including a simple implementation of all three spontaneous-rate (SR) groups of AN fibers.[3] Low SR (LSR), high threshold fibers are important for psychophysical predictions because of their potential role in encoding high sound levels,[4,5] for which the majority of AN fibers are saturated.

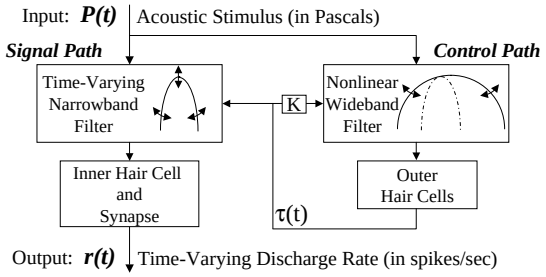

Fig. 1. Schematic block diagram of the phenomenological nonlinear auditory-nerve model for a single CF.

Finally, the present model is specifically designed to describe the time-varying discharge rate of the AN fiber, $r(t)$, rather than producing discharge times with a modified Poisson renewal process,[1] which is more computationally intensive. Psychophysical performance limits based on the Poisson variability of AN discharges can be calculated directly from $r(t)$ for either average rate or temporal information.[6,7]

2. Model overview

The implementation of the present model is identical to the Zhang *et al.*[1] model (Fig. 1), except as described in Section 3. Level-dependent tuning is represented in the signal path by a nonlinear, third-order, time-varying narrowband filter followed by a linear, first-order, broadband filter (not shown in Fig. 1). The gain and bandwidth of the nonlinear filter are controlled via the time constant $\tau(t)$, which is varied according to the control-path output. The control path has a nonlinear wideband filter followed by an asymmetric saturating nonlinearity. The saturating nonlinearity, which represents OHC transduction properties, is followed by a third-order lowpass filter with an 800-Hz cutoff frequency. The lowpass filter determines the dynamics of the cochlear amplifier (time constant of roughly 200 μs).[1] The lowpass-filter output is mapped into $\tau(t)$ according to a nonlinearity that equals τ_{narrow} for zero input and τ_{wide} for the maximum dc output of the OHC nonlinearity, where $\tau_{narrow} > \tau_{wide}$. The low-level time constant τ_{narrow} is specified from data describing low-level tuning, and the high-level time constant τ_{wide} is specified as a fraction of τ_{narrow} according to the cochlear-amplifier gain for the given CF. The linear broadband filter following the nonlinear signal-path filter has a time constant of τ_{wide}. Thus, the control path broadens the nonlinear signal filter [i.e., decreases $\tau(t)$] as the level of the stimulus that passes through the control-path filter increases. This implementation produces the desired decrease in gain and level-dependent phase shifts near CF associated with broadened tuning. Two-tone suppression[8,9] is accounted for by making the control-path filter wider than the excitatory filter (Fig. 1), thus allowing nonexcitatory stimuli to reduce the gain of the cochlear amplifier.

The control path can be modified to allow the relative effects of different nonlinear properties to be evaluated separately. Five useful versions of the AN model are described. (1) The standard AN model (*nonlinear with compression and suppression*) represents a normal-hearing subject. (2) The effects of suppression can be removed by implementing the control-path filter with the same bandwidth and CF as the nonlinear signal-path filter. This AN model version (*nonlinear with compression and without suppression*) is unrealistic physiologically, but is useful for isolating suppression effects. (3) The *linear sharp* AN model corresponds to $\tau(t) = \tau_{narrow}$ for the narrowband signal-path filter and retains full cochlear-amplifier gain. This model version has the same low thresholds and low-level tuning as the nonlinear AN models, but has sharp tuning at all levels, similar to many classic auditory models. (4) The effects of reduced frequency selectivity can be isolated from sensitivity effects with the *linear broad, low threshold* version, which corresponds to $\tau(t) = \tau_{wide}$ and full cochlear-amplifier gain, independent of stimulus level. (5) The *linear broad, impaired* AN model corresponds to $\tau(t) = \tau_{wide}$ with no cochlear-amplifier gain and represents a cochlea that has completely

dysfunctional OHCs but fully functioning inner hair cells (IHCs).

The time-varying AN discharge rate, $r(t)$, is calculated by passing the output of the signal-path filter through an asymmetric saturating nonlinearity, a lowpass filter, and a synapse model. The saturating nonlinearity and lowpass filter produce response properties associated with IHC transduction, whereas the synapse model includes adaptation effects such as the extended dynamic range at onset relative to the steady-state response.[1]

3. Implementation details

This section describes the implementation details of the present model that differ from the Zhang *et al.*[1] model (see http://earlab.bu.edu/ for the source code for both models). The cochlear map [i.e., the CF distribution along the basilar membrane (BM), used for AN population studies] in the present model was based on a human cochlear map[10] (Table 1 in ref 6), rather than cat. Psychophysical estimates of human frequency selectivity at low levels[11] were used to specify the low-level tuning of the model filters (Eq. 4 in ref 1) as a function of CF. The low-level time constant of the nonlinear signal-path filter was specified by

$$\tau_{narrow}(CF) = 1.2\{2\pi[1.019][24.7(4.37CF/1000+1)]\}^{-1}, \tag{1}$$

where the expression $24.7(4.37CF/1000+1)$ describes mid-level psychophysical estimates of equivalent rectangular bandwidth (ERB), the factor of 1.019 converts between gammatone-filter bandwidth and ERB,[12] and the factor of 1.2 increase in τ is a simplified conversion from mid- to low-level tuning that is consistent with the ratio of low- to mid-level human ERBs.[11]

The inclusion of different SR groups and the SR dependence of AN thresholds required modifying the transformation from IHC-lowpass-filter output to immediate permeability in the synapse model. The slope parameter, V_{sat},[1] was adjusted so that the rate-level curves had shapes and relative thresholds that correspond well with those reported from physiological studies,[5,13] and varies with SR as well as with CF. Thus, Eq. A17 in ref 1 was modified to

$$V_{sat} = 60\left(\frac{1+SR}{6+SR}\right)P_{Imax}K_{CF}; \qquad K_{CF} = \max[1.5, 2+1.3\log_{10}\left(\frac{CF}{1000}\right)], \tag{2}$$

where P_{Imax} was unchanged. This modification changed the value of the parameters p_1 and p_2 given in Table I and Eq. 18 of ref 1, which were simplified based on the previous SR-independent V_{sat}. The appendix in ref 1 describes the dependence of p_1 and p_2 on SR. The parameter PTS (peak-to-steady-state discharge ratio) in Table I of ref 1 depends on SR as $PTS = 1+9SR/(9+SR)$, consistent with AN data.[14] In addition, the parameter A_{SS} was changed from 350 to 130 spikes/sec to maintain reasonable saturated rates in $r(t)$ without refractoriness. The cutoff frequency of the IHC lowpass filter, cut_{ihc}, was changed from 3.8 to 4.5 kHz to maintain a proper rolloff in phase locking with CF for $r(t)$.[15] A 500-kHz sampling rate was often necessary due to strong nonlinearities in the model;[1] lower sampling rates may suffice in some conditions.

4. Model properties

Broadened tuning with increases in stimulus level is illustrated in Fig. 2. Normalized magnitude responses of the signal-path filter (Fig. 2A) are qualitatively consistent with BM magnitude responses to tones.[16] Compressive magnitude responses occur near CF (Figs. 2A,B), with maximum compression occurring at CF and more linear responses well away from CF. The compressive response begins to occur at 20 dB SPL and is fully compressive by 40 dB SPL (Fig. 2B), consistent with physiological data.[16] Model compression begins to decrease above 80 dB SPL; however, filter responses are still slightly compressive up to 120 dB SPL. The cochlear amplifier boosts the response to low level stimuli (by 30 dB in Fig. 2B), and compression represents the reduction in amplification as stimulus level increases. The amount of compression (or cochlear-amplifier gain) in the model is largest for high CFs (Figs. 2C,D); there is 55 dB

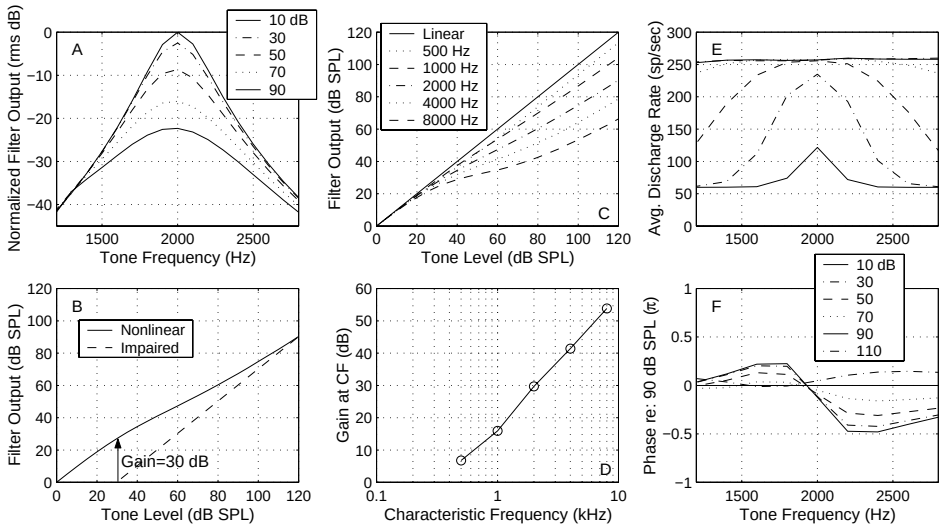

Fig. 2. Model response properties associated with nonlinear tuning. (A) Nonlinear signal-path-filter responses for a 2-kHz CF for various levels. RMS output re: 10-dB-SPL response. (B) Cochlear amplification at 2-kHz CF. (C) Magnitude compression for various CFs. (D) Cochlear-amplifier gain vs. CF, defined as the reduction in filter response at 120 dB SPL relative to a linear response. (E,F) AN response areas for 2-kHz CF. Phase plotted re: 90-dB-SPL response for same levels as (E). Any difference from zero represents a phase response that changes with level.

of gain at 8 kHz and less than 20 dB of gain below 1 kHz, consistent with data from the basal[16] and apical[17] turns of the chinchilla cochlea. Model compression (in dB/dB from 40 to 80 dB SPL, Fig. 2C) increases with frequency by a factor of 1.9 over 3 octaves, similar to psychophysical estimates of human compression (2.1 factor over 3-octaves).[18] Nonlinear cochlear tuning influences AN-fiber response areas in both average rate and phase (Figs. 2E,F). The level-dependent phase changes below and above CF (Fig. 2F) are consistent with nonlinear BM[16] and AN[19] phase responses, but are not included in most AN models.

Two-tone suppression tuning curves (Fig. 3) illustrate frequency and level combinations of a suppressor tone that reduce the response to a CF tone.[8,9] Suppression occurs below and above CF, consistent with physiological data.[9] Growth of suppression with level in the present model matches Zhang *et al.* (Fig. 13 in ref 1). The asymmetric growth is qualitatively consistent with physiological AN data for which suppression grows more rapidly below than above CF; however, the degree of model asymmetry is less than in the physiological data.[1,9]

Threshold tuning curves and rate-level curves for five versions of the AN model are compared in Fig. 4. The effects of the external and middle ear are not considered in this model; thus all AN-model versions except the impaired version have CF thresholds near 0 dB SPL. Loss of the cochlear-amplifier gain (Fig. 2D) in the impaired version results in a sloping high-frequency hearing loss, with less than 20-dB loss at low frequencies and roughly 60-dB loss at

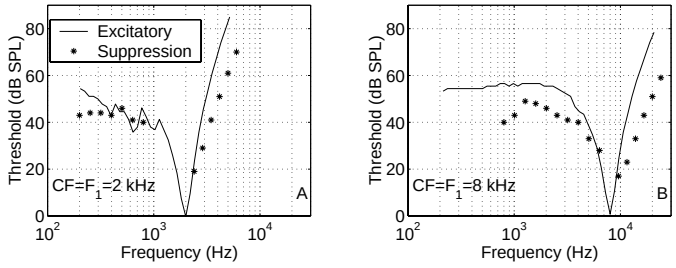

Fig. 3. Two-tone-suppression tuning curves for the nonlinear AN model (A: CF=2 kHz; B: CF=8 kHz). Solid line: excitatory tuning (threshold: 10 sp/sec > SR). Stars: suppression thresholds (-10 sp/sec criterion) for a second tone presented in addition to 15-dB-SPL CF-tone (as in ref 9).

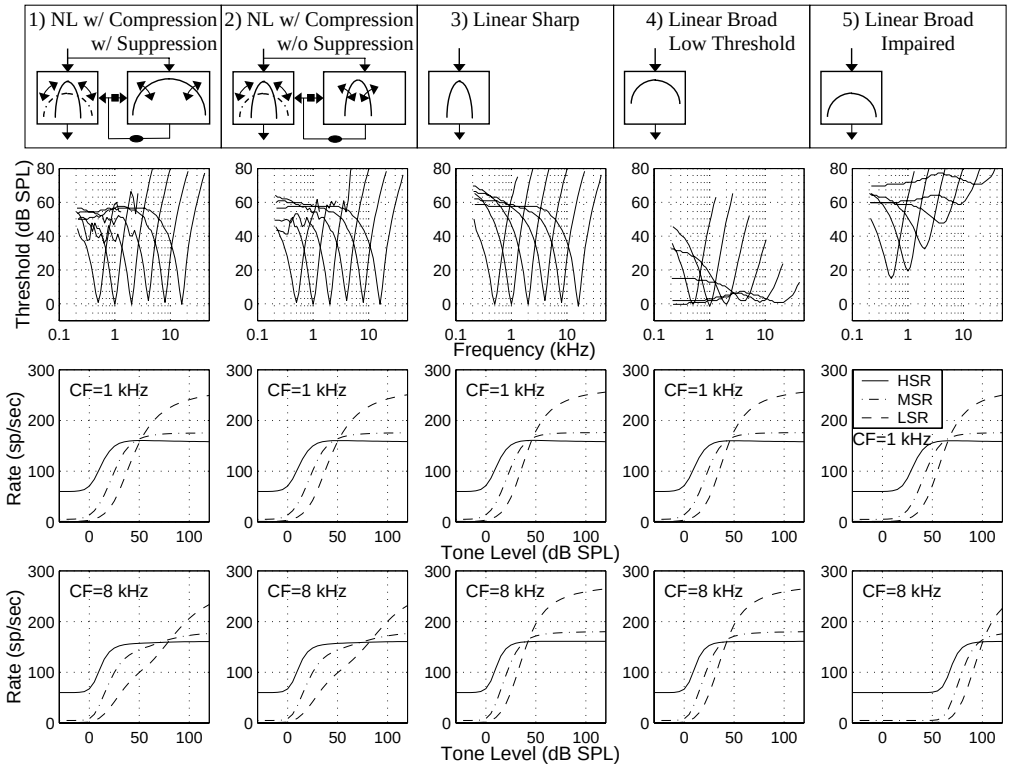

Fig. 4. Five model versions (five columns, see text) used to evaluate nonlinear properties associated with cochlear amplifier. (Row 1) Schematic diagrams. (Row 2) Excitatory tuning curves for HSR (60 sp/sec) fibers. (Rows 3,4) Rate-level curves at CF for 1- and 8-kHz CFs. Three SR groups (HSR; MSR: 5 sp/sec; LSR: 1 sp/sec). Rate over 40-240 ms for 250-ms tones.

high frequencies. Rate-level curves for high-SR (HSR) fibers have thresholds near 0 dB SPL, and a dynamic range of about 30 dB SPL. Medium SR (MSR) and LSR fibers have higher thresholds and larger dynamic ranges than the HSR fibers. The LSR rate-level curves for the nonlinear version of the AN model have "sloping-saturation" shapes at 1 kHz, and "straight" shapes at 8 kHz, consistent with guinea pig data.[5] The linear versions of the AN model have reduced dynamic ranges for MSR and LSR fibers, especially at high CFs, consistent with physiological data from impaired animals.[20] The dependence of saturated rate on SR in Fig. 4 varies with the stimulus parameters used to calculate average rate; however, the SR effects on threshold, dynamic range, and rate-level shape are consistent.

5. Discussion

The cochlear-amplifier gain as a function of CF produced "straight" rate-level curves for LSR fibers at high frequencies but not at low frequencies, consistent with guinea pig data;[5] however, straight rate-level curves have not been observed in cat.[4,13] Further studies directly relating physiological properties to psychophysical performance are necessary to verify the appropriate strength of compression as a function of CF for the human AN model. The implementation of human tuning (Eq. 1) was an intuitive choice based on psychophysical estimates of auditory filters;[11] however, these ERBs may be influenced by suppression.[7] More accurate descriptions can be easily included when they become available. The traveling wave delay correction used in the cat AN model[1] may not be appropriate for humans, and thus only the CF-dependent delay associated with the gammatone filters is present; a more accurate description could be derived from evoked-response and otoacoustic-emission studies. The simplified synapse model used to generate appropriate rate-level curves and PTS ratios for different SR groups may not capture

other effects as a function of SR, such as recovery from prior stimulation.[21] The present model does not include the effects of refractoriness, which could be incorporated as in ref 1 if desired.

Despite these limitations, this AN model provides a useful tool for quantitatively relating physiological properties associated with the cochlear amplifier to human psychophysical performance. Performance limits based on rate or temporal information in the AN population (i.e., with physiological distributions of SR[3] and CF[10]) can be calculated directly from $r(t)$ using signal detection theory.[6,7] This approach has been used to demonstrate that compression and nonlinear phase cues extend the dynamic range for level encoding within a narrow CF range, and that suppression can produce psychophysical ERBs that overestimate peripheral tuning.[7]

Acknowledgments

This model was developed as part of a graduate dissertation in the Speech and Hearing Sciences Program of the Harvard-MIT Division of Health Sciences and Technology.[7] This work was supported by ONR-MURI, NIH-NIDCD, and by NSF-IBN.

References and links

[1]X. Zhang, M.G. Heinz, I.C. Bruce and L.H. Carney, "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression," J. Acoust. Soc. Am. **109**, 648-670 (2001).
[2]M.A. Ruggero, "Physiology and coding of sound in the auditory nerve," in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A.N. Popper and R.R. Fay (Springer-Verlag, New York, 1992) pp. 34-93.
[3]M.C. Liberman, "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**, 442-455 (1978).
[4]R.L. Winslow and M.B. Sachs, "Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle," Hear. Res. **35**, 165-190 (1988).
[5]I.M. Winter and A.R. Palmer, "Intensity coding in low-frequency auditory-nerve fibers of the guinea pig," J. Acoust. Soc. Am. **90**, 1958-1967 (1991).
[6]M.G. Heinz, H.S. Colburn and L.H. Carney, "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve," Neural Computation, in press (2001).
[7]M.G. Heinz, "Quantifying the effects of the cochlear amplifier on temporal and average-rate information in the auditory nerve," Ph.D. diss., Massachusetts Institute of Technology, Cambridge, MA (2000).
[8]M.B. Sachs and N.Y.S. Kiang, "Two-tone inhibition in auditory-nerve fibers," J. Acoust. Soc. Am. **43**, 1120-1128 (1968).
[9]B. Delgutte, "Two-tone rate suppression in auditory-nerve fibers: Dependence on suppressor frequency and level," Hear. Res. **49**, 225-246 (1990).
[10]D.D. Greenwood, "A cochlear frequency-position function for several species - 29 years later," J. Acoust. Soc. Am. **87**, 2592-2605 (1990).
[11]B.R. Glasberg and B.C.J. Moore, "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138 (1990).
[12]R.D. Patterson, I. Nimmo-Smith, J. Holdsworth and P. Rice, "An efficient auditory filterbank based on the gammatone function," paper presented at a meeting of the IOC Speech Group on Auditory Modeling at RSRE, December 14-15 (1987).
[13]M.B. Sachs and P.J. Abbas, "Rate versus level functions for auditory nerve fibers in cats: Tone burst stimuli," J. Acoust. Soc. Am. **81**, 680-691 (1974).
[14]W.S. Rhode and P.H. Smith, "Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers," Hear. Res. **18**, 159-168 (1985).
[15]D.H. Johnson, "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115-1122 (1980).
[16]M.A. Ruggero, N.C. Rich, A. Recio, S.S. Narayan and L. Robles, "Basilar-membrane responses to tones at the base of the chinchilla cochlea," J. Acoust. Soc. Am. **101**, 2151-2163 (1997).
[17]N.P. Cooper and W.S. Rhode, "Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea," J. Neurophysiol. **78**, 261-270 (1997).
[18]M.L. Hicks and S.P. Bacon, "Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing," J. Acoust. Soc. Am. **105**, 326-338 (1999).
[19]D.J. Anderson, J.E. Rose, J.E. Hind and J.F. Brugge, "Temporal position of discharges in single auditory nerve fibers within the cycle of a sinewave stimulus: Frequency and intensity effects," J. Acoust. Soc. Am. **49**, 1131-1139 (1971).
[20]R.V. Harrison, "Rate-versus-intensity function and related AP responses in normal and pathological guinea pig and human cochleas," J. Acoust. Soc. Am. **70**, 1036-1044 (1981).
[21]E.M. Relkin and J.R. Doucet, "Recovery from prior stimulation. I: Relationship to spontaneous firing rates of primary auditory neurons," Hear. Res. **55**, 215-222 (1991).