REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE Dec 30, 2002	3. REPORT TYPE AND DATES COVERED Final 01/Oct/00 - 30/Sep/02

4. TITLE AND SUBTITLE
Nonlinear Nonlocal Cochlear Models, Multitones, Noises and Masking Thresholds

5. FUNDING NUMBERS
DAAD19-00-1-0524

6. AUTHOR(S)
Jack Xin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Dept. of Mathematics & TICAM
University of Texas at Austin, Austin, Tx 78712

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

41670.2-MA

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

An important part of voice signal processing is to perform a nonlinear operation along frequency on the short time spectrogram, while the nonlinear adaptation along time is better understood. We developed, computed and analyzed a class of nonlinear nonlocal cochlear models to approximate this nonlinear aspect. The model is mechanical in nature, and outputs the acoustic responses on the basilar membrane. In case of two or three tones, our results are in qualitative agreement with existing data. We prove that the model is well-posed in Sobolev spaces for all time, and admits exact multi-frequency solutions (quasi-periodic in time) if the nonlinearity is cubic and weak enough. We upscale the model output towards modeling psychoacoustic responses, to help direct applications in signal processing based on first principles. For input of tone plus a banded noise, we calibrate the model with absolute masking thresholds (on noise only), then rely on model nonlinearity to capture tonal masking of noise and modified thresholds resulting from their interactions.

14. SUBJECT TERMS
Dispersive nonlinear waves, interactions of multitones and noises, analysis and computation, masking thresholds

15. NUMBER OF PAGES

8

16. PRICE CODE

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRL-RO-BI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

- [ ] Reprint (Orig + 2 copies)
- [ ] Manuscript (1 copy)
- [ ] Technical Report (Orig + 2 copies)
- [x] Final Progress Report (Orig + 2 copies)
- [ ] Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER:

REPORT TITLE: Nonlinear Nonlocal Cochlear Models, Multitones, Noises and Masking Thresholds.

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

Jack Xin

Jack Xin

Enclosure 3

ARO Final Progress Report on Grant DAAD 19-00-1-0524.

Nonlinear Nonlocal Cochlear Models, Multitones, Noises and Masking Thresholds

PI: Jack Xin
Department of Mathematics and TICAM,
University of Texas at Austin,
Austin, TX 78712.
Email: jxin@math.utexas.edu.

Abstract

An important part of voice signal processing is to perform a nonlinear operation along frequency on the short time spectrogram, while the nonlinear adaptation along time is better understood. We developed, computed and analyzed a class of nonlinear nonlocal cochlear models to approximate this nonlinear aspect. The model is mechanical in nature, and outputs the acoustic responses on the basilar membrane. In case of two or three tones, our results are in qualitative agreement with existing data. We prove that the model is well-posed in Sobolev spaces for all time, and admits exact multi-frequency solutions (quasi-periodic in time) if the nonlinearity is cubic and weak enough. We upscale the model output towards modeling psychoacoustic responses, to help direct applications in signal processing based on first principles. For input of tone plus a banded noise, we calibrate the model with absolute masking thresholds (on noise only), then rely on model nonlinearity to capture tonal masking of noise and modified thresholds resulting from their interactions.

1 Statement of the problem studied

Given an input sound signal, we studied the mechanical responses at basilar membrane (BM) with a nonlinear nonlocal model of transmission line type:

$$p_{xx} - Nu_{tt} = \epsilon_s(x)u_t, \quad x \in (0, L), \tag{1.1}$$

$$p = mu_{tt} + r(x, |u|)u_t + s(x)u, \tag{1.2}$$

where p is the fluid pressure difference across the BM, u the BM displacement, L the longitudinal length of BM; N a constant depending on fluid density and cochlear channel size; $\epsilon_s(x)$ is a selective damping term operative near $x \sim L$; m, r, s are the mass, damping, and stiffness of BM per unit area, with m a constant, s a continuously differentiable nonnegative function of x. The function(al) r of x, u is of the form:

$$r(x, |u|) = r_a + \gamma \int_0^L P(|u(x', t)|)K(x - x')\, dx'. \tag{1.3}$$

Here r_a is the local part of BM damping, taken as positive constant for simplicity. In the nonlocal BM damping: $K = K(x)$ is a localized Lipschitz continuous kernel function with total integral over $x \in R^1$ equal to 1; $P(\cdot)$ is a nonnegative continuously differentiable function such that

$$P(0) = 0, P(q) \leq C(1 + q^2), \quad C > 0, \quad \forall q \geq 0. \tag{1.4}$$

The common choice is $P(q) = q$ (giving overall quadratic nonlinearity) or $P(q) = q^2$ (overall cubic nonlinearity).

The physical boundary and initial conditions are:

$$p_x(0, t) = T_M p_T(t) \equiv f(t), \quad p(L, t) = 0, \tag{1.5}$$

$$u(x, 0) = u_0(x), u_t(x, 0) = u_1(x), \tag{1.6}$$

where $p_T(t)$ is the input sound pressure at the eardrum; and T_M is a bounded linear operator on the space of bounded continuous functions, with output depending on the frequency content of $p_T(t)$. If $p_T = \sum_{j=1}^{J_M} A_j \exp\{i\omega_j t\} + c.c.$, a multitone input, c.c denoting complex conjugate, J_M a positive integer, then $T_M p_T(t) = \sum_{j=1}^{J_M} B_j \exp\{i\omega_j t\} + c.c$, where $B_j = a_M(\omega_j)A_j$, c.c for complex conjugate, $a_M(\cdot)$ a scaling function built from the filtering characteristics of the middle ear [8]. Established data are available in [7].

The linear part of the model is well-known [11] except for the selective damping term $\epsilon_s(x)u_t$, yet it is the nonlinear nonlocal part that plays the significant role in generating nontrivial output for multi-tone input, where there is no linear superposition. The model generalized earlier ones in [9], [6], [5], [1].

The major problem is to compute and analyze "long time" behavior of solutions, and the energy distribution over frequencies. For example, if two tones at frequencies f_1, f_2 and intensities I_1 and I_2 are taken as boundary input at $x = 0$, what are the intensities like at f_1 and f_2 after transients die out (typically after 20 ms (milleseconds)) ? Mathematically, one asks about the analytical temporal/spatial structures of solutions at long times. Similarly, we study the effect of noise on tones, and investigate how noise can be shadowed by tones at a later time (20 ms) through nonlinear interactions.

2 Summary of the most important results

(I) In [14], we used a semi-implicit finite difference method to compute quasi-steady-state responses under one or two or three tone input at $x = 0$. For one tone input, we recover isodisplacement curves similar to those in [10], showing sensitivities of BM responses near the input tones. For two and three input tones, we vary the intensity and frequency of one tone, and computed the responses of other tones. In particular, increasing intensity of one tone reduced the intensities of other tones, in agreement with experimental findings in [2], [4], [3]. Moreover, we observed propagation of dispersive long waves passing through quasi-steady-states and accumulating near $x = L$, which we devised the term $\epsilon_s(x)\, u_t$ to damped out near the right boundary point.

The numerical computation of three tone responses is hard to find in the literature, and our contribution [14] appeared to be the first.
(II) In [15], we proved that:

Theorem 2.1 (Global Well-posedness) *The model cochlear system (1.1)-(1.5) has unique solutions:*

$$(u,p) \in C([0,\infty); H^1([0,L]) \times H^3([0,L])).$$

If additionally:

$$s(x) \geq s_0 > 0; \;\; \|\epsilon_s\|_2 < \frac{2r_a}{3L^{\frac{3}{2}}}, \tag{2.7}$$

for some positive constant $s_0 > 0$, then:

$$\|(u,u_t)\|_2 \leq O(t^{1/2}), \; t^{-1}\int_0^t \|w_2\|_2^2(t')\, dt' \leq O(1), \tag{2.8}$$

for all $t > 0$.

The proof is based on writing the pressure p in terms of displacement u, and perform energy estimates on (u, u_t), $\| \cdot \|_2$ is the standard L^2 norm. We also constructed exact solutions with multi-frequencies in time:

Theorem 2.2 (Multitone Solutions) *Let $P(u) = u^2$, the overall nonlinearity is cubic; and let the input boundary condition be:*

$$f(t) = \sum_{j=1,\cdots,m} a_j \exp\{i\omega_j t\} + c.c,$$

for $m \in Z$ (any positive integer). Fix $\rho \geq 1$. Then for γ small enough, the cochlear system (1.1)-(1.5) has a solution of the form:

$$u(x,t) = \sum_{k \in Z^m} U_k(x) \exp\{ik \cdot \omega t\}, \tag{2.9}$$

where:

$$\omega = (\omega_1, \omega_2, \cdots, \omega_m),$$

$U_k(x) \in H^2([0, L])$ such that:

$$\sum_{k \in Z^m} \rho^{|k|} U_k(x) < \infty,$$

uniformly in $x \in [0, L]$. The pressure p is similar.

The proof depends on estimates of linear operators indexed by $k \in Z^m$, which are elliptic with complex coefficients. The exact solution is found by contraction mapping theorem in the Banach space whose elements have the expansions as in (2.9). We see that much more frequencies are generated by the m input ones $\omega_1, \cdots, \omega_m$. For example, the commonly observed $2\omega_1 - \omega_2$ etc.

(3) In [16], we relate the mechanical output on BM to masking thresholds in psychoacoustics (perception). The absolute hearing thresholds for banded noise is known, see Fig 1. We are interested in producing new thresholds when tones are added. For example, if we insert a tone at 2 kHz and 80 dB, then ask what the maximal level of noise is in order not to be heard (in the presence of the tone). The new threshold curve (masking curve) is shown in Fig 2, plotting masked noise intensity vs. noise frequency. The idea is: (step 1) use Fig. 1 to determine the corresponding BM displacement thresholds for noise input. Next (step 2) when both tones and noises are sent in, we extract the noisy components of BM displacements, and calculate the masking threshold levels with the BM displacement thresholds of step 1. The overall model exploits information at both the microscopic (BM) level and the macroscopic (psychoacoustic) level. We are performing extensive validation tests on this first principle based multi-scale PDE approach to psychoacoustics.

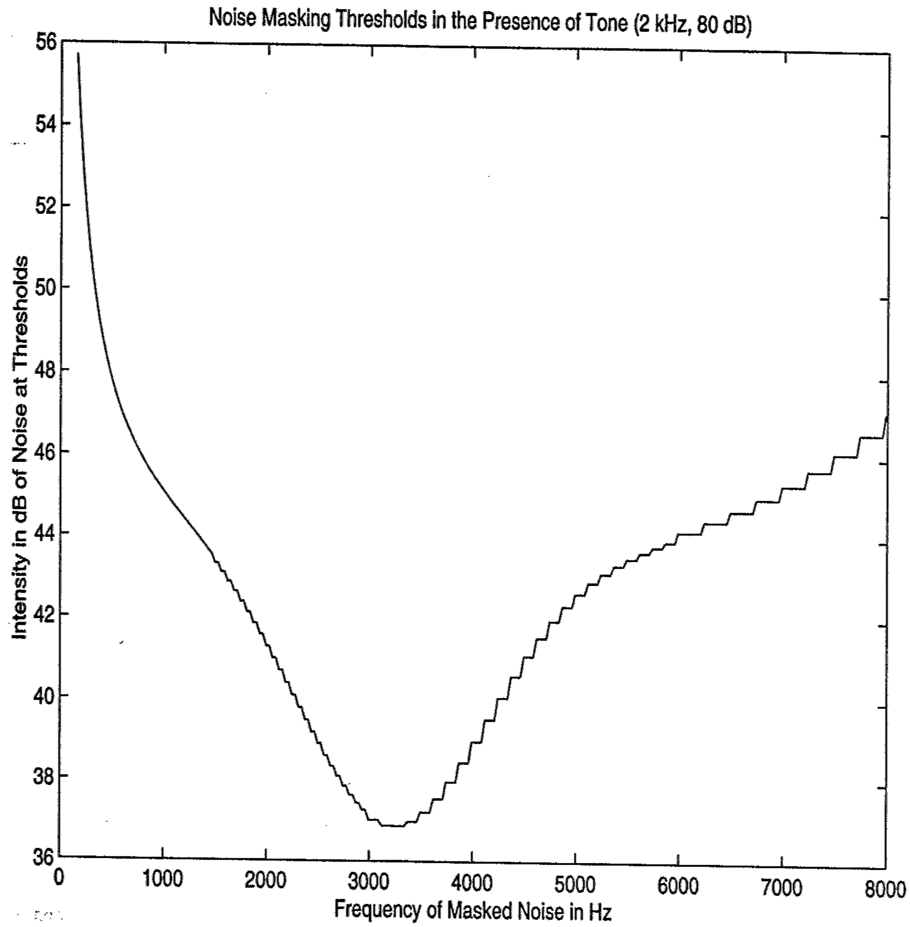

Figure 1: Absolute masking thresholds for banded noise (bandwidth 300 Hz).

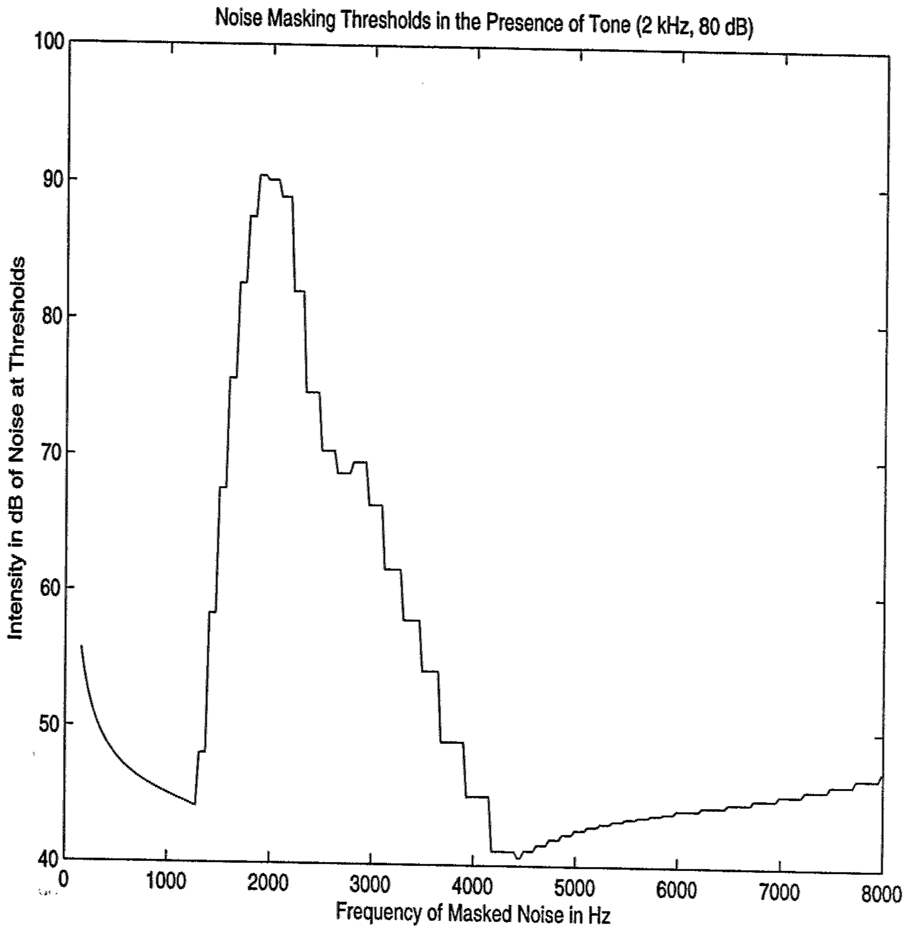

Figure 2: Masking thresholds on banded noise (bandwidth 300 Hz) in the presence of a tone (2 kHz, 80 dB).

The masking curve in Fig 2 has formed a peak centered at tone frequency 2k Hz, which helped to compress noise whose intensities fall under the curve. Such technique can be utilized for voice/speech recognition in conjunction with time adaptation. A related but less first principle based procedure (so called peak focusing) is adopted for this purpose in our earlier works [12], [13].

3 List of all publications and technical reports

(1) (with Y-Y Qi, L. Deng), "Time Domain Computation of a Nonlinear Nonlocal Cochlear Model with Applications to Multitone Interaction in Hearing", Journal of the Acoustical Soc. America, Vol. 112, No. 5, Pt. 2 of 2, p. 2300, Nov, 2002, abstract of lecture presentation at the First Pan-American/Iberian Meeting on Acoustics.

(2) (with Y-Y Qi, L. Deng), "Time Domain Computation of a Nonlinear Nonlocal Cochlear Model with Applications to Multitone Interaction in Hearing", http://arXiv.org/abs/nlin.PS/0202004, to appear in Comm. in Mathematical Sciences, Winter, 2003.

(3) (with M. D. LaMar, Y-Y Qi) "A simple hydrodynamic semi-continuous model of vocal-fold motion", Journal of the Acoustical Soc. America, Vol. 112, No. 5, Pt. 2 of 2, p. 2444, Nov. 2002, abstract of poster presentation at the First Pan-American/Iberian Meeting on Acoustics.

(4) (with M. D. LaMar, Y-Y Qi), "Modeling Vocal Fold Motion with a Hydrodynamic Semi-Continuum Model", revised version under review at Journal of Acoustic Soc. America.

(5) "Cochlear Modeling and Multitone Processing", ARO Interim Report, April 1, 2001.

(6) "A Nonlinear Nonlocal Cochlear Model and Multitone Masking in Hearing", ARO Interim Report, Jan 8, 2002.

(7) (with Y-Y Qi) "Global well-posedness and multitone solutions of a class of nonlinear nonlocal cochlear models in hearing", soon to be completed for publication.

(8) (with Y-Y Qi) "Modeling noise masking thresholds by upscaling a nonlinear nonlocal cochlear model", in preparation.

4 List of all participating personnel

Besides the PI, the reporting author, the scientific personnel are: Dr. Yingyong Qi (Qualcomm Inc, San Diego), Dr. Li Deng (Microsoft Research, Seattle), and PI's Ph.D student Mr. M. Drew LaMar.

References

[1] L. Deng, *Processing of acoustic signals in a cochlear model incorporating laterally coupled suppresive elements*, Neural Networks, 5(1), pp 19–34, 1992.

[2] L. Deng and C. D. Geisler, *Responses of auditory-nerve fibers to multiple-tone complexes*, J. Acoust. Soc. Amer., 82(6), 1989–2000, 1987.

[3] L. Deng and C. D. Geisler, *Changes in the phase of excitor-tone responses in cat auditory-nerve fibers by suppressor tones and fatique*, J. Acoust. Soc. Amer, 78(5), pp 1633–1643, 1985.

[4] C. D. Geisler, *From Sound to Synapse*, Physiology of the Mammalian Ear, Oxford Univ. Press, 1998.

[5] L. Deng, *Processing of acoustic signals in a cochlear model incorporating laterally coupled suppressive elements,* Neural Networks, Vol.5, No.1, January, 1992, pp.19–34.

[6] Y. Jau and C. D. Geisler, *Results from a cochlear model utilizing longitudinal coupling*, in Mechanics of Hearing, Martinus Nijhoff Pub., Delft Univ. Press, 1983, E. de Boer and M. Viergever, pp 169–176.

[7] J. J. Guinan and W. T. Peake, *Middle-Ear Characteristics of Anesthesized Cats*, J. Acoust. Soc. Amer, 41(5), pp 1237–1261, 1967.

[8] M. C. Liberman, *The cochlear frequency map for the cat: Labeling auditory nerve fibers of known characteristic frequency*, J. Acoust. Soc. Amer, volume 72, pp 1441–1449, 1982.

[9] R. Lyon, *A computational model of filtering, detection, and compression in the cochlear*, IEEE International Conference on Acoustics, Speech and Signal Processing, pp 1282–1285, 1982.

[10] S. Neely, *Mathematical modeling of cochlear mechanics*, J. Acoust. Soc. Amer, 78(1), pp 345–352, 1985.

[11] Sondhi, M., *The Acoustical Inverse Problem for the Cochlear*, Lecture Notes in Biomathematics, Springer-Verlag, 1980, eds M. Holmes and L. Rubenfeld, volume 43, pp 95–104.

[12] Y-Y Qi, J. Xin, *A Perception and PDE Based Nonlinear Transformation for Processing Spoken Words (Short Version)*, Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP), Vol. I, pp 445-448, 2000.

[13] Y-Y Qi, J. Xin, *A Perception and PDE Based Nonlinear Transformation for Processing Spoken Words*, Physica D 149 (2001), pp 143-160.

[14] J. Xin, Y-Y Qi, L. Deng, *Time Domain Computation of a Nonlinear Nonlocal Cochlear Model with Applications to Multitone Interaction in Hearing,* http://arXiv.org/abs/nlin.PS/0202004, to appear in Comm. in Mathematical Sciences, Winter, 2003.

[15] J. Xin, Y-Y Qi, *Global well-posedness and multitone solutions of a class of nonlinear nonlocal cochlear models in hearing,* soon to be completed for publication.

[16] J. Xin, Y-Y Qi, *Modeling noise masking thresholds by upscaling a nonlocal nonlinear cochlear model*, in preparation.