2004 International Symposium on Nonlinear
Theory and its Applications (NOLTA2004)
Fukuoka, Japan, Nov. 29 - Dec. 3, 2004

Acoustic Nonlinear Sensory Processing

A. Kern and R. Stoop

Institute for Neuroinformatics
University / ETH Zürich
Winterthurerstr. 190, 8055 Zürich
Switzerland
Email: albert@ini.phys.ethz.ch, ruedi@ini.phys.ethz.ch

Abstract—In most realistic situations, the intensity of input signals to a sensor encompasses many orders of magnitude (in the case of acoustic signals: 10^{12}). In order to accommodate such a large dynamic range, the sensor is required to operate nonlinearly. The sensor's nonlinearity is usually expected to imply severe signal distortion. In biological sensor's, however, nonlinearities are used to perform processing of sensory information. In this contribution, we demonstrate how the mammalian auditory system performs computations by means of an arrangement of coupled Hopf-type nonlinearities, which represent cochlear outer hair cells. The emerging phenomenon of two- (or multi-) tone suppression effectuates an increase in the signal-to-noise ratio of the cochlear sensor and is a prerequisite for the auditory system's solution of the scene segmentation problem.

1. Introduction

Future intelligent acoustic devices, as well as biological auditory systems, are confronted with the problem of auditory scene segmentation [1]: A mixture of incoming acoustic stimuli must be decomposed into different sound sources, and the information which is relevant for the task to be performed must be extracted. Up to now, a general solution to the auditory (as well as the visual) scene segmentation problem has not been devised. However, pathways towards an at least partial solution can be found by investigating the principles of auditory processing in biology.

In the mammalian hearing organ, the fluid-filled cochlea, sound-induced mechanical vibrations are transduced into neural excitation patterns, which involves significant information processing at the biophysical (pre-neuronal) level. The spiral-shaped cochlea is divided into two parts by the basilar membrane (BM). An incoming sound stimulus gives rise to a traveling wave along the BM, whose exponentially decreasing transversal stiffness leads to a wave maximum at a frequeny-specific resonance point (characteristic place). This place-frequency map (*tonotopic map*), which has first been proposed by H.L.F. von Helmholtz in 1863 [2] and was experimentally verified by von Békésy in 1928 [3], is the basis for the frequency selectivity of the auditory system.

Later, measurements of primary auditory nerve excitation [4] revealed a very sharp frequency selectivity that could not be explained by the structural properties of the cochlea alone, and a nonlinear growth of the responses was observed. With the improvement of measurement techniques, nonlinear BM vibrations could be measured [5]. In 1978, the discovery of otoacoustic emissions [6] (the phenomenon that the hearing system is actually able to *produce* sounds) reactivated a hypothesis originally brought forward by Gold in 1948 [7], that the cochlea acts as a regenerative *active amplifier* in order to compensate for viscous friction. This hypothesis was substantiated by the observation that the outer hair cells (OHC), which reside on the BM, amplify the BM movement by their ability to contract and elongate (motility) [8]. For low sound intensities (less than 30 dB SPL), the OHC's effectuate an enhancement of BM vibration of up to 50 dB; with increasing sound level, this amplification is reduced, which leads to a pronounced *compressive nonlinearity*; for large sound levels (beyond 90 dB SPL), the BM response is essentially passive and linear. If the physiological condition of the experimental animal deteriorates, or if cochlear damage has been induced by acoustic trauma (*hearing loss*), amplification and compressive nonlinearity are severely reduced.

The detection of OHC motility triggered efforts towards a deeper understanding of the cochlear amplification mechanism. At present, its working principles are seen as follows: it is believed that a displacement of the BM leads to a deflection of the OHC hair bundle (which protrudes into the cochlear fluid). This causes the opening or closure of ion channels in the bundle, which changes the cell's transmembrane potential. Voltage sensitive proteins in the lateral wall thus change their conformation and thereby change the area of the cell's lateral surface. As the cell volume must stay constant, this results in an elongation or contraction of the cell soma.

That the cochlea performs significant information processing by means of the nonlinear amplification mechanism becomes evident if the nonlinear interactions between *two* tones are investigated. If two tones are applied simultaneously to the ear, two phenomena arise: *suppression* and the generation of *combination tones*. In the case of suppression, the BM response to a single tone of frequency f_1 is reduced (*suppressed*) in the presence of a second tone of

frequency f_2. Evidently, the suppressive effect of the f_1- and f_2-tones is mutual. Suppression leads to a selective enhancement of signal features, which is a precondition for successful auditory scene analysis on later stages of auditory processing.

Combination tones (distortion products) with frequencies $f_{CT} = nf_1 + mf_2$ ($n, m \in \mathbb{Z}$) are generated by the nonlinear interaction between the two frequency components. Due to the structural properties of the cochlea, only the frequencies $2f_1 - f_2$ and (to a lesser extent) $f_2 - f_1$ ($f_2 > f_1$) are able to propagate to their respective characteristic places.

In this contribution, we give a detailed explanation for the observed nonlinear phenomena, based on nonlinear dynamical systems theory. Specifically, the cochlear amplification mechanism is described in terms of oscillators undergoing a Hopf bifurcation (Hopf oscillators). The experimental observations can then be explained by a variation of the effective Hopf bifurcation parameter in the presence of a second tone.

2. The Hopf Cochlea Model

Recently, it has been shown [9] that the basic characteristics of hearing can be explained from the mathematical properties of the driven Hopf oscillator,

$$\dot{z} = (\mu + i\omega_0)z - |z|^2 z + F(t), \quad z(t) \in \mathbb{C}, \tag{1}$$

where ω_0 is the natural frequency of the oscillation, $\mu \in \mathbb{R}$ denotes the bifurcation parameter, and $F(t) = Fe^{i\omega t}$ is an external periodic forcing with frequency ω. In the absence of external forcing, (1) describes the generic differential equation displaying a Hopf bifurcation. For an input $F(t)$, $z(t)$ can be considered as the amplified signal. The steady-state solution for periodic forcings is obtained by the ansatz $z(t) = Re^{i\omega t + i\phi}$, which leads to a cubic equation in R^2,

$$F^2 = R^6 - 2\mu R^4 + [\mu^2 + (\omega - \omega_0)^2]R^2. \tag{2}$$

Assuming $\omega = \omega_0$ and $\mu < 0$, for $F \lessapprox |\mu|^{3/2}$, the response is linear, $R \approx -F/\mu$. If $F \gtrapprox |\mu|^{3/2}$, the R^6-term becomes dominant, and the compressive nonlinear regime is entered, $R \approx F^{1/3}$, with the amplification gain decreasing like $F^{-2/3}$. For $\omega \neq \omega_c$, $R \approx F/\sqrt{\mu^2 + (\omega - \omega_c)^2}$, and the response is always linear. If $\mu > 0$, stable limit cycles emerge, which explains the generation of otoacoustic emissions.

The fact that the properties of (2) explain the observed characteristics of hearing motivated the development of a Hopf-type cochlea model (for details see [10]). From energy-balance arguments, the cochlea differential equation,

$$\frac{\partial e(x,\omega)}{\partial x} = -\frac{e(x,\omega)}{v(x,\omega)}\left[\frac{\partial v(x,\omega)}{\partial x} + d(x,\omega)\right] + \frac{a(x,e,\omega)}{v(x,\omega)}, \tag{3}$$

was derived. $e(x,\omega)$ denotes the one-dimensional energy density of the cochlear fluid, $v(x,\omega)$ is the group velocity of the BM traveling wave, $d(x,\omega)$ encompasses viscous losses, and $a(\cdot)$ denotes the nonlinear active amplification by OHC. Based on cochlear biophysics (see [10])

$$a(e,x,\omega) = L(R(\sqrt{\sigma e(x,\omega)}))^2, \tag{4}$$

where L and σ are constants, and $R(\cdot)$ is determined by (2). The connection between the cochlea model and experimentally measured BM response A is given by the relation $A(x,\omega) = (2e(x,\omega)/E(x))^{1/2}$ ($E(x) = E_0 \exp(-\alpha x)$ denotes the BM stiffness). The response of the cochlea model displays remarkable coincidence with experimental measurements [10].

3. Nonlinear Cochlear Signal Processing

In the presence of a tone consisting of two frequencies, the driving term of (1) reads

$$F(t) = F_1 e^{i\omega_1 t + i\psi_2} + F_2 e^{i\omega_2 t + i\psi_1} + F_{CT} e^{i\omega_{CT} t + i\psi_{CT}}, \tag{5}$$

where we allow for phases ψ_k of the two frequency components, $F_k > 0$, and $\omega_k = 2\pi f_k$, $k = \{1, 2\}$. When CT responses at frequency $\omega_{CT} = 2\omega_1 - \omega_2$ ($\omega_2 > \omega_1$) are generated at a certain site on the BM, these constitute a component of the input to Hopf oscillators at neighboring BM locations. For the Hopf cochlea model, the last term in (5) must therefore be considered.

The steady-state solution of (1) is obtained from the Fourier series ansatz

$$\begin{aligned} z(t) \quad = \quad & R_1 e^{i\omega_1 t + i\phi_1} + R_2 e^{i\omega_2 t + i\phi_2} \\ & + R_{CT} e^{i\omega_{CT} t + i\phi_{CT}} + \sum_j R_j e^{i\omega_j t + i\phi_j}. \end{aligned} \tag{6}$$

The third term denotes the propagating combination tone with frequency $\omega_{CT} = 2\omega_1 - \omega_2$ ($\omega_2 > \omega_1$), and the sum includes all higer-order contributions $\omega_j = n\omega_1 + m\omega_2$, $\{n, m\} \in \mathbb{Z}^2 \setminus \{2, -1\}$.

After some calculations, the response to frequencies ω_1, ω_2 is obtained as

$$F_k^2 = R_k^6 - 2\mu_{eff,k} R_k^4 + [\mu_{eff,k}^2 + (\omega_k - \omega_0)^2]R_k^2, \tag{7}$$

where $k = \{1, 2\}$ and $j \neq k$. These equations can be interpreted as single Hopf equations with effective bifurcation parameters $\mu_{eff,k} = \mu - 2R_j^2$ (cf. Eq. (2) and note that $\mu < 0$). Since the small-signal gain is given by $1/|\mu_{eff}|$, it becomes evident that the *suppressive effect* in the presence of a second tone is captured by a shift of the effective bifurcation parameter away from the bifurcation point.

The response at ω_{CT} is obtained in the same way,

$$\begin{aligned} & F_{CT}^2 + R_1^4 R_2^2 - 2R_1^2 R_2 F_{CT} \cos(2\phi_1 - \phi_2 + \psi_{CT}) \\ & = R_{CT}^6 - 2\mu_{eff,CT} R_{CT}^4 + [\mu_{eff,CT}^2 + (\omega_{CT} - \omega_0)^2]R_{CT}^2. \end{aligned} \tag{8}$$

If comparing (8) with (7), three points attract our attention. First, we note from the emergence of an effective bifurcation parameter $\mu_{eff,CT} = \mu - 2(R_1^2 + R_2^2)$, that suppression

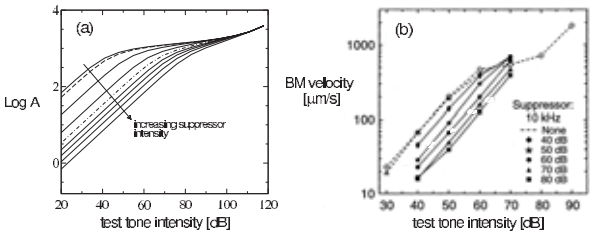

Figure 1: Two-tone suppression: a) Model response: suppressor intensity increases from 10 dB to 110 dB in steps of 10 dB. The 10, 20, and 30 dB lines coincide. b) Experimental measurements [11].

plays a crucial role in CT generation. Secondly, the term $R_1^4 R_2^2$ expresses CT generation in the absence of external driving, F_{CT}. From the discussion in Sec. 2, it is seen that the CT response is given by $R_{CT} \approx R_1^2 R_2/\mu$, if $R_1^2 R_2 < |\mu|^{3/2}$ (assuming $\omega_{CT} = \omega_0$). If R_2 is kept fixed and R_1 is increased, we thus assume a 2 dB/dB increase of R_{CT}.

As a third point, we observe that the presence of an external driving F_{CT} at frequency ω_{CT} not only gives rise to the term F_{CT}^2. In addition, a phase-dependent term is induced, where ϕ_k ($k = 1, 2$) denote the phase differences between R_k and the driving force,

$$\phi_k = \arctan \frac{\omega_k - \omega_0}{\mu - R_k(R_k^2 - 2R_j^2)} - \psi_{CT}, \quad j \neq k. \qquad (9)$$

For a single Hopf oscillator, the CT response is easily computed from (8) and (9). In the cochlea model, however, the phase ψ_{CT} is determined by the cochlear hydrodynamic wave. The computation of the cos-term in (8) thus becomes difficult, but fortunately, its contribution to CT generation can be neglected for the following arguments. Firstly, if f_1 and f_2 are not too close, either F_{CT} or $R_1^2 R_2$ dominate on the left hand side of (8), so that the cos-term always remains small. This has been verified by numerical simulations for the frequencies used. Secondly, the interaction with the hydrodynamic wave causes rapid changes of ψ_{CT} along the BM, so that the contributions by the phases are effectively averaged out.

The Hopf model response for a two-frequeny tone is obtained by resolving a system of three differential equations of the form (3). This provides the energy densities e_k and e_{CT}. As R_k and R_{CT} must be substituted in (4), these equations are coupled by Eqs. (7) and (8).

3.1. Two-Tone Suppression

In two-tone suppression experiments, the response to one tone (the *test tone*) is measured in the presence of a *suppressor* tone (indexing by t and s). The test-tone input-output function obtained by the Hopf cochlea model, determined for increasing suppressor intensity, shows nearly perfect agreement with experimental measurements (Fig. 1). In this representation, the BM response at characteristic place (the location of maximum BM response) is plotted as a function of sound intensity. For suppressor levels I_s below 40 dB (top curve in Fig. 1) we recognize the strong compressive nonlinearity which is characteristic for the single-frequency cochlear response. For $I_s \gtrapprox 40$ dB (dashed line in Fig. 1a), the small-signal gain of the test tone becomes significantly reduced, with constant separations between the curves. If $I_s > 70$ dB (dashed-dotted line), these are reduced by a factor of about $1/3$.

The Hopf cochlea model provides an explanation for these observations (see also [12]). Since the small-signal response of the test tone is given by $R_t = F_t/|\mu_{eff,t}|$, and $\mu_{eff,t} = \mu - 2R_s^2$, we conclude that suppressive effects become appreciable if $\mu_{eff,t}$ deviates significantly from μ, which is the case if $F_s \sim R_s \gtrapprox \sqrt{|\mu|}$. The spacing between the curves reflects the compressive nonlinearity of the suppressor response, R_s. For $I_s < 70$ dB, $R_t \sim F_t/R_s^2 \sim F_t/I_t$, which explains the constant spacings between the curves in Fig. 1. If the suppressor enters the compressive nonlinear regime, $R_s^2 \sim F_s^{2/3} \sim I_s^{1/3}$ holds, which leads to a reduction of the spacing by $1/3$. It is remarkable that the same effect is observed in the experiment.

In Fig. 2, the BM response of the test tone is displayed as a function of suppressor intensity, where model and experiment again show remarkable coincidence. For suppressor frequencies larger than the test tone frequency, the nonlinear compressive regime leads to a more moderate dependence of the test tone response on the suppressor intensity.

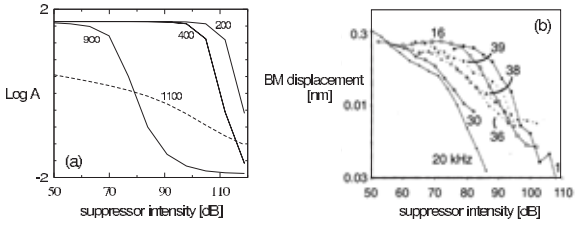

Figure 2: Two-tone suppression as a function of suppressor intensity: a) Model ($\omega_t/2\pi = 1$ kHz). b) experimental measurements [13] ($\omega_t/2\pi = 34$ kHz), where the labels indicate the suppression frequencies.

3.2. Combination Tones

CT measurements are performed in a variety of experimental settings [14]. We restrict our analysis to the situation where the CT response is measured as a function of the intensity of the f_1-component, while the level of the f_2-component is kept fixed (Fig. 3). We observe a close agreement of the model results with the experimental measurements.

An explanation of Fig. 3 is again provided by the Hopf cochlea model. At the increasing branches of the curves, the slope is exactly 2 dB/dB, as was predicted from Eq.

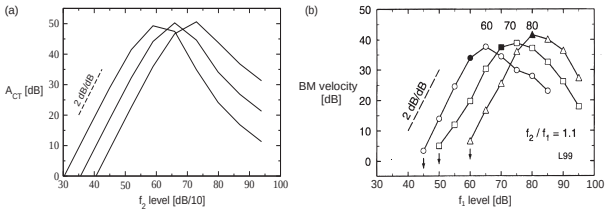

Figure 3: Combination tone generation: BM response at characteristic place for a tone with frequency $f = 2f_1 - f_2$, as a functon of f_1-intensity. a) Model response (curves for $f_2 = 60, 70, 80$ dB; $f_1 = 930$ Hz, $f_2 = 1000$ Hz, $f_2/f_1 = 1.05$). b) Experimental measurements [14] ($f_2/f_1 = 1.1$).

(8). The role of suppression is twofold: For low f_1-levels, suppression of the CT stems exclusively from the f_2-component. This explains the *decrease* of the CT response upon increase of the f_2-level (while f_1-intensity remains fixed), which is observed when CT responses at different curves are read off for fixed f_1-intensity. For the same reason, the 2 dB/dB-slope remains unaffected: From Eq. (8) follows $R_{CT} \approx R_2 R_1^2/\mu_{eff,CT} \sim R_1^2$, as $\mu_{eff,CT}$ is only a function of R_2 for small f_1-intensities. Since $\mu_{eff,CT} = \mu - 2(R_1^2 + R_2^2)$, the contribution of the f_1-component to suppression becomes significant if $R_1 \gtrapprox R_2$, which is the case when the intensity of the f_1-component exceeds the f_2-level. This explains the decrease of the CT response for large f_1-intensities.

4. Conclusion

In the preceeding section we have demonstrated that the Hopf cochlea model provides an successful description of cochlear nonlinear phenomena. The role of suppression in cochlear information processing consists in the reduction of the response to small-amplitude signals (which can be considered as noise). This leads to a pattern-sharpening effect, analog to the increase in resolution of neural receptive fields, which is achieved by lateral or surround inhibition. The role of combination tones is less clear. Possibly, they may help in signal identification (scene analysis) if several signals of comparable magnitude are present; if the signal intensities differ, the combination tone is readily suppressed. CT generation sometimes plays a role in music – the phenomenon has been described for the first time by the violinist Tartini in 1714,

For the design of intelligent acoustic devices, which perform signal identification and scene analysis tasks, a profound understanding of the nonlinear phenomena in mammalian hearing may provide helpful. For example, if a speech recognition system is endowed with a simple cochlea model as a front end, its performance increases significantly [15]. We therefore expect that the Hopf approach to cochlear modeling will be of great benefit for developping sound-processing devices.

References

[1] A.S. Bregman, *Auditory Scene Analysis* MIT Press, Cambridge, 1999.

[2] H.L.F. Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik.* Vieweg, Braunschweig, 1863.

[3] G. von Békésy, "Zur Theorie des Hörens. Die Schwingungsform der Basilarmembran", *Phys. Z.* 29, pp. 793–810, 1928.

[4] Y. Katsuki, T. Sumi, H. Uchiyama, and T. Watanabe, "Electric responses of auditory neurons in cat to sound stimulation", *J. Neurophysiol.* 21, pp. 569–588, 1958.

[5] B.M. Johnstone and A.J. Boyle, "Basilar membrane vibration examined with the Mössbauer technique", *Science* 158, pp. 389–390, 1967.

[6] D.T. Kemp, "Stimulated acoustic emission from within the human auditory system", *J. Acoust. Soc. Am.* 64, pp. 1386–1391, 1978.

[7] T. Gold, "Hearing. II. The physical basis of the action of the cochlea", *Proc. R. Soc. London B Biol. Sci.* 135, pp. 492–498, 1948.

[8] W.E. Brownell, C.R. Bader, D. Bertrand, and Y. de Ribaupierre, "Evoked mechanical responses of isolated cochlear outer hair cells", *Science* 227, pp. 194–196, 1985.

[9] V.M. Eguíluz, M. Ospeck, Y. Choe, A.J. Hudspeth, and M.O. Magnasco, "Essential nonlinearities in hearing", *Phys. Rev. Lett.* 84, pp. 5232–5235, 2000.

[10] A. Kern and R. Stoop, "Essential role of couplings between hearing nonlinearities", *Phys. Rev. Lett.* 91, pp. 128101 1–4, 2003.

[11] M.A. Ruggero, L. Robles, N.C. Rich, "Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory nerve-rate suppression", *J. Neurophysiol.* 68, pp. 1087–1099, 1992.

[12] R. Stoop and A. Kern, "Essential auditory contrast-sharpening is preneuronal", *Proc. Natl. Acad. Sci. U.S.A.* 101, pp. 9179–9181, 2004.

[13] W.S. Rhode and N.P. Cooper, "Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo", *Auditory Neurosci.* 3, pp. 101–121, 1996.

[14] L. Robles, M.A. Ruggero, and N.C. Rich, "Two-tone distortion on the basilar membrane of the chinchilla cochlea", *J. Neurophysiol.* 77, pp. 2385–2399, 1997.

[15] J. Tchorz and B. Kollmeier, "A model of auditory perception as front end for automatic speech recognition", *J. Acoust. Soc. Am.* 106, pp. 2040–2050, 1999.