NONLINEAR DYNAMICS OF COCHLEAR INFORMATION PROCESSING

A. Kern and R. Stoop

Institute of Neuroinformatics,
Swiss Federal Institute of Technology Zürich,
Zürich, Switzerland
e-mail: albert@ini.phys.ethz.ch
WWW: http://www.stoop.net/group

Abstract— *The nonlinear amplification process in the mammalian cochlea gives rise to a variety of phenomena, which manifest as two-tone suppression and combination tone generation. These nonlinear effects show that, besides mere mechanical-to-neural transduction, the cochlea performs significant information processing on a biophysical, pre-neural level. As nonlinear cochlear processing is a precondition for successful feature extraction at higher neural stages, its profound understanding is of interest for the design of intelligent acoustic sensors. In this contribution, we provide a thorough explanation of suppression and combination tone generation, where we rely on Hopf-type cochlear amplifiers. The underlying cochlear model can be implemented as an electronic circuit.*

I. INTRODUCTION

The first theory of the mammalian hearing organ, the fluid-filled cochlea, was put forward by H.L.F. Helmholtz in 1863 [1]. Based on anatomical investigations, Helmholtz proposed that each segment of the basilar membrane (BM), which separates the cochlear fluid, acts as a tuned oscillator. A sound of given frequency thus ellicits maximum oscillations at a specific location in the cochlea (characteristic place), so that the cochlea acts as a spatial Fourier analyzer (tonotopic principle). Mechano-sensitive cells on the BM then transduce the mechanical vibrations into neural signals. In 1928, the tonotopic principle has been verified experimentally by von Békésy [2]. In contrast to Helmholtz' original theory, however, the tonotopic principle is correctly deduced from the exponentially decaying transversal BM stiffness $E(x) = E_0 \exp(-\alpha x)$, by applying linear hydrodynamical theory [3], [4]. Cochlear hydrodynamics has also been described in terms of electronic circuit analogs [5].

In the early 1970s, increasing evidence was furnished that the cochlear response is strongly nonlinear [6], which was in stark contrast to the prevalent linear theory. With the detection of otoacoustic emissions [7] it became clear that a nonlinear force-generating mechanism must be present in the cochlea. In 1985, the outer hair cells (OHC), which reside on the BM, have been identified as the source of this mechanical amplification, and as the origin of cochlear nonlinearity [8]. This discovery has triggered intensive research in the following two decades [9]. In particular, it has been shown that a degeneration of OHC causes cochlear hearing loss. In this case, even the use of sophisticated hearing aids often results only in partial improvement of auditory performance; especially the capability for auditory scene analysis frequently remains severely hampered. It thus follows that, in addition to mechano-to-neural transduction, the cochlea performs significant information processing by means of the nonlinear amplification mechanism. Cochlear information processing applies mainly to the frequency domain, while processing of time information is performed on the neural level.

For two nonlinear phenomena – two-tone suppression and combination tone (CT) generation – there exist ample physiological measurements. Both phenomena arise if two tones are applied simultaneously to the ear. 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. 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). By this approach, the experimental observations can be explained by a variation of the effective Hopf bifurcation parameter in the presence of a second tone.

II. The Hopf Cochlea Model

Recently, it has been shown [10] that the basic characteristics of hearing can be explained from the mathematical properties of the driven Hopf oscillator,

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

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,

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

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 – linear BM response for weak stimuli ($\lessapprox$ 30 dB SPL), and a compressive nonlinearity for moderately intense responses – motivated the development of a Hopf-type cochlea model (for details see [11]). From energy-balance arguments [12], the cochlea differential equation,

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

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 [11])

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

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).

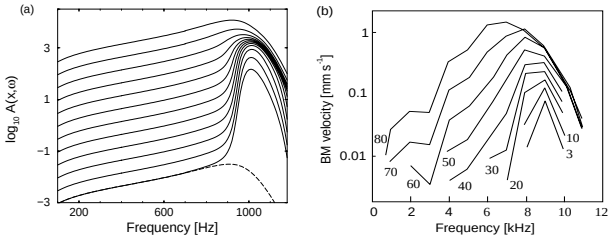

Fig. 1. Frequency response at fixed location on BM. (a) Hopf cochlea model, Eqs. (2-4). (c) Experimental measurements [13].

The frequency response of the cochlea model (measured at fixed location on BM) displays remarkable coincidence with experimental measurements (Fig. 1). Optimal responses are obtained if feedforward couplings between the Hopf amplifiers are taken into account [11]. In the following analysis, however, we use the simpler version of the model (Fig. 1a).

III. Nonlinear Cochlear Signal Processing

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

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

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

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

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

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

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,

$$F_{CT}^2 + R_1^4R_2^2 - 2R_1^2R_2F_{CT}\cos(2\phi_1 - \phi_2 + \psi_{CT}) = R_{CT}^6 - 2\mu_{eff,CT}R_{CT}^4 + [\mu_{eff,CT}^2 + (\omega_{CT} - \omega_0)^2]R_{CT}^2. \quad (8)$$

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 plays a crucial role in CT generation. Secondly, the term $R_1^4R_2^2$ expresses CT generation in the absence of external driving, F_{CT}. From the discussion in Sec. II, it is seen that the CT response is given by $R_{CT} \approx R_1^2R_2/\mu$, if $R_1^2R_2 < |\mu|^{3/2}$ (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,

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

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 $R_1^2R_2$ 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.

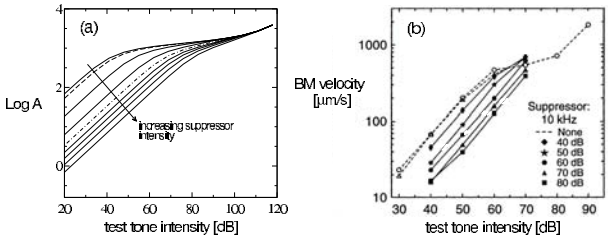

Fig. 2. Two-tone suppression: a) Model response: suppressor intensity increases from 10 dB to 110 dB in steps of 10 dB. The 10, 20, and 30 dB lines coincide. b) Experimental measurements [15].

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) [14].

A. 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. 2). In this representation, the BM response at characteristic place (the location of maximum BM response, cf. Fig. 1) is plotted as a function of sound intensity. For suppressor levels I_s below 40 dB (top curve in Fig. 2) 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. 2a), 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. 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. 2. If the suppressor enters the compressive non-

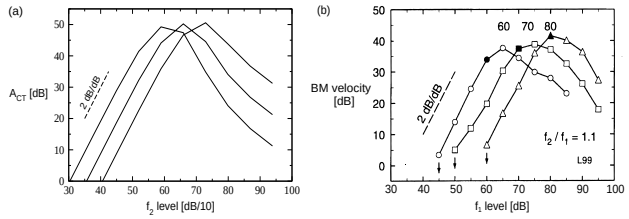

Fig. 3. Combination tone generation: BM response at characteristic place for a tone with frequency $f = 2f_1 - f_2$, as a functon of f_1-intensity. a) Model response (curves for $f_2 = 60, 70, 80$ dB; $f_1 = 930$ Hz, $f_2 = 1000$ Hz, $f_2/f_1 = 1.05$). b) Experimental measurements [16] ($f_2/f_1 = 1.1$).

linear 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.

B. Combination Tones

CT measurements are performed in a variety of experimental settings [16]. 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. (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 $F_{CT} \approx R_2 R_1^2/\mu_{eff,CT} \sim R_1^2$, 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.

IV. 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 [17]. We therefore expect that the Hopf approach to cochlear modeling will be of great benefit for developping sound-processing devices. Hopf oscillators can be implemented as electronic circuits [18].

References

[1] H.L.F. Helmholtz, *Die Lehre von den Tonempfi ndungen als physiologische Grundlage für die Theorie der Musik* (Vieweg, Braunschweig, 1863).
[2] G. von B´ek´esy, *Phys. Z.* **29**, 793 (1928).
[3] O.F. Ranke, *Die Gleichrichter-Resonanztheorie* (Lehman, M¨unchen, 1931).
[4] B.P. Peterson and L.C. Bogert, *J. Acoust. Soc. Am.* **22**, 369 (1950).
[5] R.L. Wegel and C.E. Lane, *Physical Review* **23**, 266 (1924).
[6] W.S. Rhode, *J. Acoust. Soc. Am.* **49**, 1218 (1971).
[7] D.T. Kemp, *J. Acoust. Soc. Am.* **64**, 1386 (1978).
[8] W.E. Brownell, C.R. Bader, D. Bertrand, and Y. de Ribaupierre, *Science* **227**, 194 (1985).
[9] W.E. Brownell, A.A. Spector, R.M. Raphael, and A.S. Popel, *Annu. Rev. Biomed. Eng.* **3**, 169 (2001).
[10] V.M. Egu´ıluz, M. Ospeck, Y. Choe, A.J. Hudspeth, M.O. Magnasco, *Phys. Rev. Lett.* **84**, 5232 (2000).
[11] A. Kern and R. Stoop, *Phys. Rev. Lett.* **91**, 128101 (2003).
[12] G.B. Whitham, *Linear and Nonlinear Waves* (Interscience Publishers, New York, 1999).
[13] M.A. Ruggero, *Curr. Opin. Neurobiol.* **2**, 449 (1992).
[14] For the model, it is assumed that the wave propagation along the cochlea can be described by linear hydrodynamic theory. The energy densities associated with the different frequency components (e_1, e_2 and e_{CT}) are therefore exclusively coupled through the active process (term $a(\cdot)$ in (3)).
[15] M.A. Ruggero, L. Robles, N.C. Rich, *J. Neurophysiol.* **68**, 1087 (1992).
[16] L. Robles, M.A. Ruggero, and N.C. Rich, *J. Neurophysiol.* **77**, 2385 (1997).
[17] J. Tchorz and B. Kollmeier, *J. Acoust. Soc. Am.* **106**, 2040 (1999).
[18] J.-J. van der Vyver, A. Kern, and R. Stoop, In: *Proc. of the IEEE European Conference on Circuit Theory and Design (ECCTD)* vol. III, 285 (2003).