Two-Tone Suppression and Combination Tone Generation as Computations Performed by the Hopf Cochlea

R. Stoop and A. Kern

Institute of Neuroinformatics, University/ETH of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

(Received 12 December 2003; published 20 December 2004)

Recent evidence suggests that the compressive nonlinearity responsible for the extreme dynamic range of the mammalian cochlea is implemented in the form of Hopf amplifiers. Whereas Helmholtz's original concept of the cochlea was that of a frequency analyzer, Hopf amplifiers can be stimulated not only by one, but also by neighboring frequencies. To reduce the resulting computational overhead, the mammalian cochlea is aided by two-tone suppression. We show that the laws governing two-tone suppression and the generation of combination tones naturally emerge from the Hopf-cochlea concept. Thus the Hopf concept of the cochlea reproduces not only local properties like the correct frequency response, but additionally accounts for more complex hearing phenomena that may be related to auditory signal computation.

DOI: 10.1103/PhysRevLett.93.268103 PACS numbers: 87.19.Dd, 05.45.Xt, 43.64.+r

Introduction.—The determination of the working principles of the mammalian cochlea presents a scientific and technological challenge. Helmholtz was the first to come up with an anatomically motivated cochlea model, hypothesizing a place-frequency mapping along the cochlea duct [1]: a specific place in the cochlea would respond optimally to only one particular frequency, much like the strings on a piano (the so-called tonotopic principle). The essentials of this model were confirmed by von Békésy's physiological measurements [2] half a century later. He observed traveling waves along the cochlear basilar membrane (BM) that assumed a maximal amplitude at a place determined by the frequency of the stimulation tone. These findings were later explained from first principles by hydrodynamic models of the cochlea [3].

In the following decades, more precise physiological measurements revealed a number of hearing phenomena that cannot be explained using the aforementioned linear theory. Eguíluz *et al.* [4] were the first to point out that the incorporation of Hopf oscillators in cochlea modeling should enable the reproduction of the basic nonlinear active hearing phenomena. Indeed, the correct frequency response to single tones can be obtained using a biomorphic implementation of the Hopf amplifier concept [5]. They also proposed that more of the nonlinear phenomena of hearing might be explained if stimulation was with more than one frequency. In particular, they noted the missing fundamental frequency problem and the fact that the intensity of combination tones is not suppressed in a linear fashion. Whereas the problem of the missing fundamental can easily be explained by locking among limit-cycle oscillations, the proper emergence of combination tones has not as yet been shown.

In this Letter, we show theoretically and numerically that the most pertinent nonlinear hearing phenomena emerge naturally from the Hopf-cochlea model. In particular, we derive in a unified approach the laws that govern the physiologically measured auditory combination tones and two-tone suppression. Two-tone suppression [6–8] describes the phenomenon that, if more than one tone are presented to the cochlea, the effect is an attenuated response to the tones. By completely suppressing small contributions in close proximity, the representation of the large components is significantly enhanced, and the spectrum in a complex sound is substantially reduced. Combination tones are autonomously generated in the cochlea, when more than one pure tone arrive at the cochlea. Therefore, the combination tones and two-tone suppression phenomena are crucial for the sparse representation of auditory signals. Modern space-minimizing formats of auditory information storage systematically exploit these properties (e.g., *MP3* and *Vorbis* audio formats).

From this point of view, the notion that auditory computations are already performed at the level of the cochlea, before the signal is fed into higher (neuronal) processing stages, is justified.

Biomorphic Hopf-Cochlea model.—In our model of the cochlea, active contributions elicited by the auditory stimulations are locally injected into the passive hydrodynamic waves. It was shown [5] that this leads to the following initial value problem for the one-dimensional energy density $e(x, \omega)$:

$$\frac{\partial e(x,\omega)}{\partial x} = -\frac{1}{v_G(x,\omega)}\left[\frac{\partial v_G(x,\omega)}{\partial x} + d(x,\omega)\right]e(x,\omega) + \frac{a(x, e(x,\omega), \omega)}{v_G(x,\omega)}, \quad (1)$$

in which the dissipation d is counteracted by the power a delivered by the active process. v_G is the group velocity. Equation (1) originates in the steady-state situation, where an energy-balance argument between dissipation and active amplification can be made (for a detailed biophysical derivation of the model see [5], where it is also shown how second order couplings [9] lead to a fine-tuning of the response). The active power $a(\cdot)$ is determined from the

fact that Hopf oscillators (i.e., hair cells active at location x), deliver a force whose amplitude is proportional to the Hopf response R; cf. Eq. (5),

$$a(e, x, \omega) \sim (R(\sqrt{e(x, \omega)}))^2. \quad (2)$$

The locally generated forces lead [5] to a BM displacement

$$A(x, \omega) = (2e(x, \omega)/E(x))^{1/2}, \quad (3)$$

where $E(x)$ is the exponentially decaying BM stiffness. Equations (1)–(3) establish the connection between the cochlea differential equation and physiological measurements. For single tones, using realistic biophysical parameters, even the simplest variant of our model leads to responses that compare very well with physiological measurements [Fig. 1(a); see also [5]].

Active amplification is modeled by the generic Hopf system [4]

$$\dot{z} = (\mu + i\omega_0)z - |z|^2 z + Fe^{i\omega t}, \qquad z(t) \in \mathbb{C}. \quad (4)$$

In this equation, $\omega \equiv 2\pi f$ is the frequency of the external stimulation and $\omega_0(x)$ is the frequency for which at location x along the cochlear duct the amplification is maximal (this is how the tonotopic principle is implemented). $z(t)$ can be considered the amplification of the input signal $F(t)$. In the absence of external forcing ($F = 0$), the equation displays a Hopf bifurcation [10]: For parameter $\mu < 0$, the solution $z(t) = 0$ is a stable fixed point, which for $\mu > 0$ becomes unstable and is replaced by a stable limit cycle $z(t) = \sqrt{\mu}e^{i\omega_0 t}$. The steady-state solution for periodic forcing is obtained with the ansatz $z(t) = Re^{i\omega t + i\phi}$, where a 1:1 locking between signal and system is assumed. The response amplitude R is then determined from a cubic equation in R^2,

$$F^2 = R^6 - 2\mu R^4 + [\mu^2 + (\omega - \omega_0)^2]R^2. \quad (5)$$

Below, we show how the detailed understanding of the input-output relation resulting from this equation provides the basis for understanding the most salient nonlinear hearing phenomena, which are two-tone suppression and combination tone generation. For $\mu = 0$ and close to resonance $\omega = \omega_0$, the response $R = F^{1/3}$ emerges, which forces the gain $G = R/F = F^{-2/3}$ to increase towards infinity as F approaches zero. This implies a compressive nonlinearity for stimuli of any intensity. For $\mu < 0$ and still near resonance, we obtain for weak stimuli F the response $R = F/|\mu|$. As F increases, the term R^6 in (5) can no longer be neglected, and the compressive nonlinear regime is entered as $R^6 \approx \mu^2 R^2$. The transition occurs at $F_{cnl} \approx |\mu|^{3/2}$. Therefore, for weak stimuli F, the response R is almost linear, whereas for moderate stimuli the differential gain dR/dF decreases with increasing stimulus intensity. Away from the resonance, the last term in Eq. (5) dominates, leading to $R \approx F/|\omega - \omega_0|$, which implies a linear response to any input.

Presence of several or two tones.—In the presence of two pure tones (F_1, ω_1) and (F_2, ω_2), the driving term of Eq. (1) becomes $F(t) = \sum_{l=1,2} F_l e^{i\omega_l t + \psi_l} + \sum_{CT} F_{CT} e^{i\omega_{CT} t + \psi_{CT}}$, where CT indexes combination tones of frequencies $\omega_{CT}(n, m) = n\omega_1 + m\omega_2$, and their phases ψ. Letting $\omega_2 > \omega_1$, due to structural properties of the cochlea, only the frequencies $\omega_{CT_0} = 2\omega_1 - \omega_2$ and $\omega_{CT_1} = \omega_2 - \omega_1$ are able to propagate, where the first contribution is greater than the second [11]. Consequently, we use the Fourier ansatz

$$z(t) = R_1 e^{(i\omega_1 t + i\phi_1)} + R_2 e^{(i\omega_2 t + i\phi_2)} + R_{CT_0} e^{(i\omega_{CT_0} t + i\phi_{CT_0})} + \sum_j R_j e^{(i\omega_j t + i\phi_j)}, \quad (6)$$

where $\{R_1, R_2\}$ are the response components of frequencies $\{\omega_1, \omega_2\}$ and where the terms $\{R_j, \omega_j = n\omega_1 + m\omega_2\}$, $\{n, m\} \in \mathbb{Z}^2 \setminus \{2, -1\}$ account for the remaining combination tones [4].

In order to calculate the response generated at the loci corresponding to $\{\omega_1, \omega_2\}$ for two-tone suppression (and to ω_{CT_0} for combination tones), this Fourier ansatz is inserted into the generic Hopf equation. The combination tones are of order $m + n > 2$ in the responses R_1, R_2. For the responses at frequencies $\{\omega_1, \omega_2\}$, we therefore collect contributions to ω_1 and ω_2 up to third order. This yields

$$i\omega_k R_k = (\mu + i\omega_0)R_k - R_k(|R_k|^2 + 2|R_l|^2) + F_k \exp(-i\phi_k), \quad (7)$$

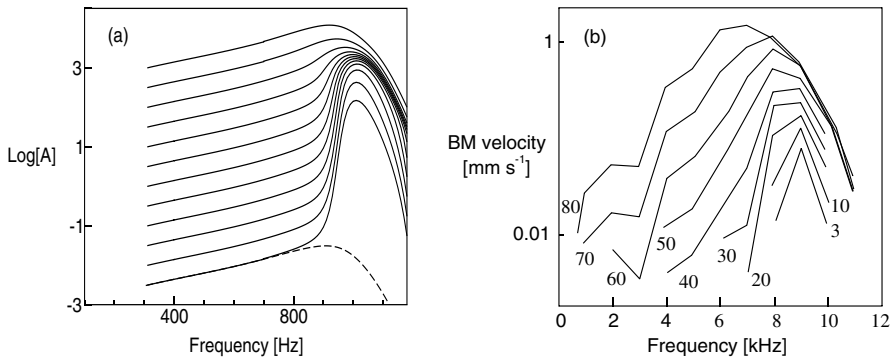

FIG. 1. Local BM frequency responses: (a) Hopf-cochlea model Eqs. (1)–(3), where longitudinal BM coupling has been included only. Dashed line: passive response (stimulus frequency: $\omega/2\pi = 1000$ Hz). For adjacent lines, the stimulus intensity differs by 10 dB. (b) Experimental measurements [13].

where $k \neq l \in \{1, 2\}$. After isolating the phases and multiplying by their complex conjugates, the responses

$$F_k^2 = R_k^6 - 2(\mu - 2R_l^2)R_k^4 + [(\mu - 2R_l^2)^2 + (\omega_k - \omega_0)^2]R_k^2, \tag{8}$$

with $k \neq l \in \{1, 2\}$, are obtained.

The response corresponding to ω_{CT_0} leads, analogously, to

$$F_{CT_0}^2 + R_1^4 R_2^2 - 2R_1^2 R_2 F_{CT_0} \cos(2\phi_1 - \phi_2 + \psi_{CT_0}) = R_{CT_0}^6 - 2[\mu - 2(R_1^2 + R_2^2)]R_{CT_0}^4 + \{[\mu - 2(R_1^2 + R_2^2)]^2 + (\omega_{CT_0} - \omega_0)^2\}R_{CT_0}^2. \tag{9}$$

This equation shows, using $F_{CT_0} = 0$, that combination tones are generated even in the absence of an external driving. This is the case pertinent to the emergence of combination tones.

It is essential to note that both phenomena, two-tone suppression and combination tones, are captured by a modified Hopf equation

$$F^2 = R^6 - 2\tilde{\mu}R^4 + [\tilde{\mu}^2 + (\omega - \omega_0)^2]R^2, \tag{10}$$

where $\tilde{\mu}$ is an effective bifurcation parameter. For two-tone suppression, we have

$$\tilde{\mu}_{S,k} = \mu - 2R_k^2; \tag{11}$$

for the combination tone, we have

$$\tilde{\mu}_{CT_0} = \mu - 2(R_1^2 + R_2^2). \tag{12}$$

Both modifications of μ drive the bifurcation parameter farther away to the left of the bifurcation point, which has a natural interpretation as a suppression effect, as $R \approx F/|\tilde{\mu}|$. This observation provides an explanation of the two phenomena from the same fundamental principle, Hopf's equation [Eq. (5) and Eq. (10), respectively]. For the comparison between the model and experimental measurements, the BM response A of Eq. (3) is identified with the Hopf response R, and the stimulation intensity I_k with the square of the forcing F_k^2, $k \in \{1, 2\}$.

Results.—Physiological measurements of combination tones use a variety of paradigms [12]. We restrict ourselves to the case where the intensity $I_2(\omega_2)$ is held fixed and intensity $I_1(\omega_1)$ is varied. Figure 2 shows that with the active components implemented as discussed, our model matches well with the experimentally measured physiological membrane response. The main features of this response can be explained from the most fundamental ingredient of our cochlea model, the Hopf amplifier, in a simple way. In the physiological experiments and our cochlea model simulations, a prominent 2 dB per dB increase of the combination tone response is observed. At about $\omega_{CT_0} = \omega_0$, we have $R_{CT_0} \approx R_1^2 R_2/\tilde{\mu}_{CT_0}$ for $R_1^2 R_2 < |\tilde{\mu}_{CT_0}|^{3/2}$ [see Eq. (11)]. Since, for $R_1^2 \ll R_2^2$, $\tilde{\mu}_{CT_0}$ remains essentially constant when R_1 is increased, this explains why $R_{CT_0}(I_1)$ increases by 2 dB per dB. As soon as $R_1 \gtrsim R_2$, however, the increase of R_1 drives $\tilde{\mu}_{CT_0}$ farther to the left, leading to a decreased response. This explains the part of the curves left of the maxima. The decrease to the right of the maxima is the effect of the Hopf amplifiers serial cascade; the details of this more difficult explanation is omitted. The explanation of the observed right shift of the curves for increased I_2 is simple again: The maximum of the combination tone response is attained for $R_1^2 \approx R_2^2$.

Also for two-tone suppression, an excellent agreement between physiological measurements [6] and the model response is obtained. This is shown in Fig. 3, where the maximal BM response (cf. Fig. 1) is plotted vs the test-tone intensity I_1. For obtaining this correspondence, only the dB-scale origin and the proportionality constant in Eq. (2) determining the gain and the width of the compressive nonlinearity had to be chosen appropriately in the modeling. Low-dB BM input-output curves slowly detach from the zero suppressor curve, staying in its neighborhood up to an intensity of 30 dB. This is the consequence of the strong compressive nonlinearity inherited from the single frequency response, for both modeling and experimental results. A regime of strong suppressor efficacy—indicated by large intercurve distances—that terminates in strongly reduced suppressor efficacy is found adjacent to it. Again, the Hopf response Eq. (5) provides a simple qualitative and quantitative explanation: at sufficiently weak intensity I_2 of the suppressor, we are in the linear regime, where

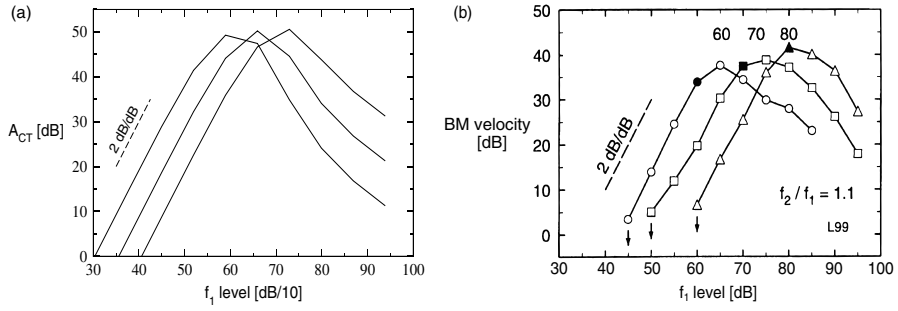

FIG. 2. Combination tone generation: BM response at a characteristic place for a tone with frequency $f_{CT_0} = 2f_1 - f_2$, as a functon of f_1 intensity. (a) Model response (curves for $f_2 = 60$, 70, and 80 dB; $f_1 = 930$ Hz, $f_2 = 1000$ Hz, $f_2/f_1 = 1.05$). (b) Experimental measurements [12].

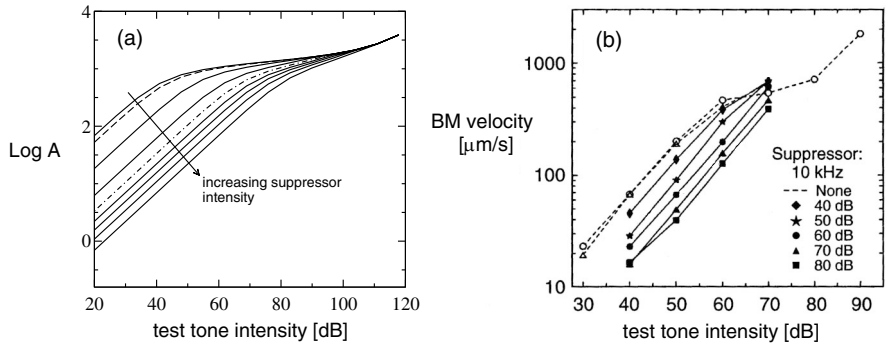

FIG. 3. High-side suppression: (a) Model response [Eq. (3)] at resonance. Suppressor intensities from 10 to 110 dB, in steps of 10 dB. The 10, 20, and 30 dB lines coincide ($\omega_t/2\pi =$ 0.9 kHz, $\omega_s/2\pi = 1.0$ kHz). (b) Experimental measurements [14] ($\omega_t/2\pi =$ 8 kHz).

$R_2 \sim F_2/|\mu|$ is valid. This implies that $R_2^2 \ll |\mu|$, which leads to $\tilde{\mu}_{S,1} \approx \mu$, thus the generated response curve is confined to the close vicinity of the suppressor-free case. For the same reason, the response of $\tilde{\mu}_{S,1}$ to a change of F_2 will be rather mild. Only if $2R_2^2$ becomes comparable to μ [dashed curve in Fig. 1(a)], the suppressor unfolds its efficacy. When $2R_2^2$ dominates, we observe a linear logarithmic suppression. Its efficacy decreases after the suppressor has run in its compressive nonlinearity (dash-dotted curve). This is when yet another regime of linear logarithmic suppression emerges, where the spacings are reduced to about one-third. By a rigorous analysis, the factor of $1/3$ can be made precise.

Discussion.—Using a biomorphic model of the cochlea, we are able to give a detailed explanation of two-tone suppression and combination tone generation, the two most pertinent nonlinear hearing phenomena. Our results show clearly that efficient signal compression mechanisms are already performed by the biophysics of the cochlea.

Subcritical tuning of Hopf amplifiers and their embedding in a biophysically detailed cochlea model were essential for the quality of the results. At first glance, the impression may be that, in the presence of a second tone, both critically and subcritically tuned systems become subcritically tuned, rendering the latter distinction unsubstantial. This, however, is incorrect: Recall first that the suppression threshold is given by $R_2^2 \sim |\mu|/2$, for both tunings [Eq. (11)]. The transition point for critical tuning is therefore determined by $F_{1,cnl}^{\text{crit}} \approx 2^{3/2} R_2^3$, whose lower bound is zero, attained for $F_2 \to 0$. $F_{1,cnl}^{\text{crit}}$ therefore keeps changing for arbitrarily weak suppressor intensities. This implies that at about the zero suppression curve, no accumulation of curves should be obtained, which is in contradiction with the physiological data [e.g., Fig. 2(b)]. For subcritical tuning, $|\tilde{\mu}|$ is bounded below by $|\mu|$. This implies a stationary transition point $F_{1,cnl}^{\text{sub}} = |\mu|^{3/2}$ for low suppressor intensities, in agreement with the observed accumulation of low-dB BM input-output curves. Thus, subcritical tuning is, indeed, responsible for the excellent agreement between the modeling and physiology.

We thank M. Magnasco, M. A. Ruggero, T. Kohda, and W. Schwarz for beneficial exchanges on the subject, and the Swiss KTI for continued support.

[1] H. L. F. Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (Vieweg, Braunschweig, 1863).
[2] G. von Békésy, Phys. Z. **29**, 793 (1928).
[3] E. de Boer, in *The Cochlea, Springer Handbook of Auditory Research*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer, New York, 1996), p. 258.
[4] V. M. Eguíluz, M. Ospeck, Y. Choe, A. J. Hudspeth, and M. O. Magnasco, Phys. Rev. Lett. **84**, 5232 (2000).
[5] A. Kern and R. Stoop, Phys. Rev. Lett. **91**, 128101 (2003).
[6] L. Robles, M. A. Ruggero, and N. C. Rich, Nature (London) **349**, 413 (1991).
[7] R. B. Patuzzi, P. M. Sellick, and B. M. Johnstone, Hear. Res. **13**, 19 (1984).
[8] N. P. Cooper and W. S. Rhode, Hear. Res. **63**, 163 (1992).
[9] C. D. Geisler and C. Sang, Hear. Res. **86**, 132 (1995).
[10] E. Hopf, Ber. Math.-Phys., Sächs. Akad. d. Wiss. Leipzig **94**, 1 (1942).
[11] Suppose that $\omega_1 > \omega_2$, but close. Then, unlike ω_{CT_0}, ω_{CT_1} is rather small. From this, using Eq. (10), it can be inferred that ω_{CT_1} undergoes a much smaller amplification compared with ω_{CT_0}. Moreover, from physiological measurements, at small frequencies, the gain is reduced.
[12] L. Robles, M. A. Ruggero, and N. C. Rich, J. Neurophysiol. **77**, 2385 (1997).
[13] M. A. Ruggero, Curr. Opin. Neurobiol. **2**, 449 (1992).
[14] M. A. Ruggero, L. Robles, and N. C. Rich, J. Neurophysiol. **68**, 1087 (1992).