Zwicker Tone Illusion and Noise Reduction in the Auditory System

Jan-Moritz P. Franosch[1], Richard Kempter[1], Hugo Fastl[2], and J. Leo van Hemmen[1]
[1]*Physik Department, TU München, 85747 Garching bei München, Germany,*
[2]*Lehrstuhl für Mensch-Maschine-Kommunikation, TU München, 80333 München, Germany*
(April 22, 2003)

The Zwicker tone is an auditory aftereffect. For instance, after switching off a broad-band noise with a spectral gap, one perceives it as a lingering pure tone with the pitch in the gap. It is a unique illusion in that it cannot be explained by known properties of the auditory periphery alone. Here we introduce a neuronal model explaining the Zwicker tone. It is shown that a neuronal noise-reduction mechanism in conjunction with dominantly unilateral inhibition explains the effect. A pure tone's 'hole burning' in noisy surroundings is given as an illustration.

43.66.Ki,43.64.Bt,43.66.Ba,43.66.Ed

Zwicker [1] discovered an intriguing auditory aftereffect in 1964, now called the Zwicker tone. The typical sound generating it is a broad-band noise containing a spectral gap, which is presented for several seconds. After the noise has been switched off, a faint, almost pure, tone is audible for one up to six seconds. It is decaying and has a sharp pitch in the spectral gap [1,2], where no stimulus was available.

Both the localization of the Zwicker tone in the brain and its origin are long standing open problems. Here we present a neuronal model explaining both. Through computer simulations based on our model of first processing stages of the auditory pathway following the cochlea, we show that both noise reduction and dominantly unilateral, in short, *asymmetric* inhibition along the tonotopic (frequency) axis are inherent properties of parts of the auditory system so as to allow the Zwicker tone to arise. We also explain why quite a few noise configurations (Fig. 1) do or do not generate a Zwicker tone [2–5]. Furthermore, we have successfully tested the model in psychoacoustic experiments described in [6]. All cases of Fig. 1 are explained by the model, as has been verified through extensive simulations.

The Zwicker tone, as an almost pure tone, has a totally different quality from its generating noise. In Zwicker's experimental setup [1] of Fig. 1(b), the tone was *disjoint* from the spectral range of the noise generating it. That is why Zwicker called it a "negative" afterimage. We now show that this aftereffect is double-negative in that it is due to both noise reduction and asymmetric inhibition – and *not* to habituation, as in the visual system. Surprisingly, the Zwicker tone is inherently asymmetric. In the presence of broad-band noise with a spectral gap, it exists above the lower band edge but *not* below the upper band edge of the gap; cf. Fig. 1(b). Neither does it exist at the high-pass edge in Fig. 1(f).

The Zwicker tone has been found neither in the cochlea nor in the auditory nerve and cannot be explained by known properties of the auditory periphery either. Hence it is thought to have its origin in the central auditory system but it is not clear yet where precisely. It has been demonstrated to some extent in primary auditory cortex both physiologically [7] and through magneto-encephalographic methods [8].

To explain the Zwicker tone by means of a neuronal model of the early auditory system, we start with a few facts, which are widely assumed to be basic. First, neurons are tonotopically ordered along a frequency axis corresponding to the cochlea [9]. Second, neurons can habituate to a steadily stimulating sound so that their spontaneous rate after switching off the sound is lower than in their resting state. As we will see, habituation as an explanation for the Zwicker tone is questionable. Third, there is lateral inhibition. Fourth, spontaneous activity in the auditory system is high, and can even exceed 100 Hz. It is just below threshold and, hence, just not noticeable. A sound is 'perceived' if neurons fire at a rate above their spontaneous one; cf. Fig. 2.

Habituation with symmetric lateral inhibition can only explain cases (a), (g) and (h) of Fig. 1. To account for cases (b) and (f), we are bound to assume that inhibition between neurons is effective from low to high frequencies but not, or hardly, conversely. In short, it must be *asymmetric*. The range (half width at half maximum) of the asymmetric inhibition is taken to be 0.8 octaves.

Let us now consider a pure tone at a low-pass noise edge; see Fig. 1(e). The pure tone habituates the neurons in its tonotopic surroundings even more than noise alone does. With asymmetric inhibition, a Zwicker tone is therefore bound to be above the pure tone. In experiment, however, we have found one *below* the pure tone (provided its amplitude is high enough). Why is that?

Without noise no Zwicker tone [2]. Until now we did not exploit the fact that noise plays a key role. We define noise to be a sound of roughly constant amplitude over a broad frequency range (exceeding about 0.3 octaves) and with a duration greatly exceeding 100 ms. The deficiencies of the above habituation argument are overcome by a model that incorporates noise detection but drops habituation. Asymmetric inhibition is a key element dangling in the background. In addition, we assume a tonotopic array of *noise-detection neurons*: only with noisy input around their best frequency do they become active and inhibit 'output neurons' projecting to higher-order centers; cf. Fig. 2(d). They are slow in responding so as to

catch the noise characteristics. That is to say, their inhibition lasts longer than any other integration time in the auditory brain stem and, thus, is of the order of seconds after the noise has been switched off. We analyze threefold evidence in favor of a noise-reduction mechanism: physiological, psychophysical, and computational.

The first possible origin of the Zwicker tone is the dorsal cochlear nucleus (DCN). Here one finds strong lateral inhibition [10,11] and several types of neuron [12,13]. We list them according to the shape of their receptive field, i.e., their firing rate plotted as a function of a pure tone's frequency f and intensity I: Type I with a V-shaped response domain in the (f, I) plane, the 'V' pointing at the best frequency, type II with the same receptive field but now flanked by inhibitory side-bands [14] so that their activity is suppressed by broadband noise, and type IV responding to broadband noise. Moreover, type II neurons inhibit their tonotopic type IV companions [13] so that the latter do not function in the presence of a pure tone that is strong enough. Though the DCN structure is not completely known yet, we tentatively associate type II with 'feature detectors' and type IV with 'noise detectors'; see Fig. 3 for the complete setup. Type IV neurons are known to have time constants of the order of seconds [11], so they are hypothesized to be responsible for the duration of the Zwicker tone in our model. Noise-detection neurons may, and we suppose will, inhibit the output neurons. In fact, broadband noise causes additional inhibition of DCN neurons that could not be detected without noise [15].

We now return to psychophysical experiments that cannot be explained by habituation, e.g., a pure tone at a noise edge; see Fig. 1(e). The noise-reduction model, on the other hand, provides a solution. As shown in Fig. 2(f), during sound presentation noise-detection neurons react against the noise, except where they are inhibited by the pure tone. In the tone's tonotopic neighborhood *hole burning* arises because the pure tone excites feature-detector neurons, which inhibit the noise detectors. Hence, the latter are not active in the neighborhood of the pure tone, neither before nor after the sound has been switched off. Conversely, noise detectors further to the left on the frequency axis are active during the stimulus and also for a second or so after its termination. Shortly after the termination, the corresponding output neurons still feel inhibition from noise detectors, fire less and, thus, generate less (asymmetric) inhibition to their right. Hence the output neurons fire there at a higher rate than the spontaneous one and generate a Zwicker tone below the pure tone.

Further computational evidence for the case of a pure tone at a low-pass noise edge is provided by Fig. 4. It explains the characteristics of the Zwicker tone, in particular, position, low amplitude, and being a pure tone. It is a critical test since it rules out all all previous explanations such as symmetric inhibition and habituation, plausible as they look at a first sight.

For a pure tone embedded in white noise (Fig. 1(d)), hole burning is even more true and explains why a Zwicker tone is perceived. If only a pure tone is present, as in Fig. 1(j), noise-detection neurons are not excited and no aftereffect occurs. Figures 1(f) & (k) are mirror images of (a) & (e) but no Zwicker tone arises because inhibition is asymmetric so that neurons 'to the left' do not notice what happens above the frequency edge.

We now describe the biophysical model that we used to compute all the responses in Figs. 2 & 4. Single neurons have been modeled by the Spike Response Model [17], details can be found in [18,19]. The membrane potential $v_i(t)$ of neuron i at time t is taken to be

$$v_i(t) = \sum_j J_{ij} \sum_{\{t_j\}} \varepsilon(t - t_j) + \eta(t - t_i(t)) \qquad (1)$$

where J_{ij} is the synaptic efficacy between neuron j and i, $\{t_j\}$ is the set of firing times of neuron j while $t_i(t)$ is neuron i's most recent firing time preceding t. Neuron i spikes at time t_i if and only if $v_i(t_i) = \theta$ with $v_i'(t_i) > 0$. The postsynaptic potential $\varepsilon(t)$ vanishes for $t < 0$ and is given by $\varepsilon(t) := (t/\tau_{\text{syn}}) \exp(1 - t/\tau_{\text{syn}})$ elsewhere.

The refractory function η describes absolute refractory behavior for $0 < t \leq \tau_{\text{ref}}$ through $\eta(t) := -\infty$ and relative refractory follow-up for $t > \tau_{\text{ref}}$, here by means of $\eta(t) = \gamma_{\text{ref}}/(t - \tau_{\text{ref}})$. Habituation is taken into account by an *additional* refractory function $\eta_{\text{hab}}(t) = -v_{\text{hab}} \exp(-t/\tau_{\text{hab}})$, where τ_{hab} is the neuron's habituation time constant and v_{hab} is the habituation strength.

To get a realistic input, the basilar membrane has been modeled as a set of fourth-order linear gamma-tone filters so as to account for the slopes of the spectra at the edges as shown by the dashed lines in Fig. 1. The amplitude of the basilar membrane calculated by these filters is coupled to the stereocilia of a Meddis inner-hair-cell model. This peripheral part was simulated by using the software package 'LUTEar' [16]. Its output consists of probabilities for spikes of the auditory nerve fibers, an inhomogeneous Poisson process.

In conclusion, the Zwicker tone is a psychoacoustic aftereffect providing essential information on mechanisms in the auditory system that cope with noise. We have presented, and analyzed, a model explaining psychoacoustic experiments that eliminate the usual habituation argument. In so doing, we have provided strong evidence for our hypothesis that a neuronal noise-reduction mechanism in conjunction with asymmetric inhibition generates the aftereffect. Furthermore, the model fully explains all Zwicker-tone experiments known at present and summarized in Fig. 1. We suggest that type IV neurons in the cochlear nucleus (DCN) play the noise-reduction role.

The Zwicker tone as a transient auditory sensation is often thought of as a short-term tinnitus. Tinnitus, on the other hand, is a long-term auditory phantom percept. Central tinnitus and Zwicker tone are related in that both can be perceived as a pure tone that is generated in the central auditory system. Both can be induced by a 'spectral gap' in the auditory-nerve activity. A common type

of central tinnitus develops over days following peripheral hearing loss, and the perceived pitch of the illusionary pure tone often matches frequencies of the hearing loss – similar to the relation between a notched-noise stimulus and the Zwicker tone. In contrast to the latter, central tinnitus might be the result of a persistent activation of a noise-reduction mechanism. By incorporating synaptic learning rules into our current model of the Zwicker tone, we hope to gain further insights into the mechanism underlying the generation of tinnitus, which can lead to new strategies for central tinnitus therapies.

This work was DFG-supported through GRK 267/1-96, FOR 306, and Ke 788/1-1.

---

[1] E. Zwicker, J. Acoust. Soc. Am. **36**, 2413–2415 (1964).

[2] R. C. Lummis and N. Guttman, J. Acoust. Soc. Am. **51**, 1930–1944 (1972).

[3] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, 2nd ed. (Springer, Berlin, 1999).

[4] H. Fastl, Acustica **67**, 177–186 (1989).

[5] H. Fastl and G. Krump, in *Advances in Hearing*, edited by G. A. Manley (World Scientific, Singapore, 1995), pp. 457–466.

[6] H. Fastl, D. Patsouras, J.-M. P. Franosch, and J. L. van Hemmen, in *Proc. of the 12th International Symposium on Hearing* (Shaker, Maastricht, 2001), pp. 67–74.

[7] R. W. W. Tomlinson, U. Biebel, and G. Langner, in *Göttingen Neurobiology Report 1998*, edited by N. Elsner and U. Eysel (Thieme, Stuttgart, 1998), Vol. 2, p. 345.

[8] E. S. Hoke, M. Hoke, and B. Ross, Audiol. Neurootol. **1**, 161–174 (1996).

[9] Auditory excitation through the cochlea is taken to be the logarithm of the amplitude of the basilar membrane at a certain frequency. Excitation is then proportional to the firing rate of the corresponding auditory nerve fibers [3]. A pure tone's input width is 0.5 octaves.

[10] *The Mammalian Auditory Pathway: Neurophysiology*, edited by A. N. Popper and R. R. Fay (Springer, New York, 1992).

[11] E. D. Young and W. Brownell, J. Neurophysiol. **39**, 282–300 (1976).

[12] H. F. Voigt and E. D. Young, Hear. Res. **6**, 153–169 (1982); J. Neurophysiol. **64**, 1590–1610 (1990).

[13] I. Nelken and E. D. Young, J. Basic Clin. Physiol. Pharmacol. **7**, 199 (1996).

[14] I. M. Winter and A. R. Palmer, J. Neurophysiol. **73**, 141–159 (1995).

[15] W. S. Rhode and S. Greenberg, J. Neurophysiol. **71**, 493–514 (1994).

[16] L. P. O'Mard, M. J. Hewitt, and R. Meddis, *LUTEar*

[17] W. Gerstner and J. L. van Hemmen, in *Models of Neural Networks II*, edited by E. Domany, J. L. van Hemmen, and K. Schulten (Springer, New York, 1994), pp. 39–47.

[18] In the habituation model (Figs. 2d, 3 & 4), a synaptic weight J_{ij} [see also Eq.(1)] depends on the overall efficacy J_0, the width d, and the lateral offset x_0 of the arborization function: $J_{ij} = J_0\,\{1+\cos[2\pi(j-i-x_0)/d]\}/2000$ for $-d \le 2(j-i-x_0) \le d$ and $J_{ij} = 0$ elsewhere. For pairs of neurons in four layers [input fibers (in), feature detectors (feature), noise detectors (noise), output neurons (out)], the triplets (J_0, d, x_0) are: $(0.34, 80, 0)$ for in→noise, $(1.1, 80, 0)$ for in→out, $(-0.6, 40, 0)$ for noise→out, $(-0.05, 160, 80)$ for out→out ($J_{ii} = 0$), $(1.15, 80, -80)$ and $(-0.82, 80, 80)$ for in→feature. For feature→noise we have $J_{ij} = -0.005$ for $127 < j-i < 213$ and 0 elsewhere. In Fig. 2a, parameters are $(1.8, 200, 0)$ for in→out and $(-0.06, 250, 125)$ for out→out.

[19] Here, all Spike-Response neurons [17] have a uniform threshold $\theta = 1$, and, furthermore, are described through the triplet $(\tau_{\rm syn}, \tau_{\rm ref}, \gamma_{\rm ref})$. In the noise-reduction model (Figs. 2d, 3, 4) we have used (400 ms, 10 ms, 0.3 ms) for noise detectors, (200 ms, 1 ms, 3 ms) for output neurons, and (250 ms, 1 ms, 3 ms) for feature detectors. In the habituation model (Fig. 2a), we have assumed (150 ms, 1 ms, 3 ms), $\tau_{\rm hab} = 7$ s and $v_{\rm hab} = 0.006$.

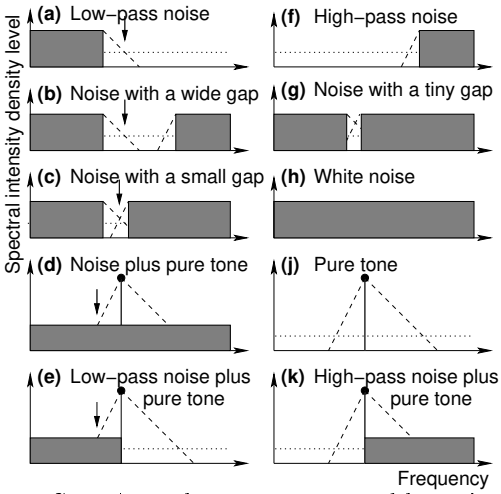

FIG. 1. A Zwicker tone is generated by various noise configurations. The horizontal axis denotes increasing frequency, the vertical one indicates the corresponding sound amplitude (arbitrary units). Shaded areas denote noise and vertical lines ending in a filled circle indicate pure tones. Zwicker tone generating sounds (left column) are low-pass noise (a), noise with a gap (b) & (c), or noise plus a pure tone (d) & (e). The pitch of the Zwicker tone is indicated by a downward arrow. Dashed lines denote auditory excitation [3,9] and dotted lines indicate the threshold in silence. No Zwicker tone has been observed in (f)–(k). Cases (e) and (k) represent new experimental results [6], motivated by the *hole burning* concept of the noise-reduction model presented here. As the gap decreases, the Zwicker tone is first perceived less clearly (c) until it is no longer audible (g).

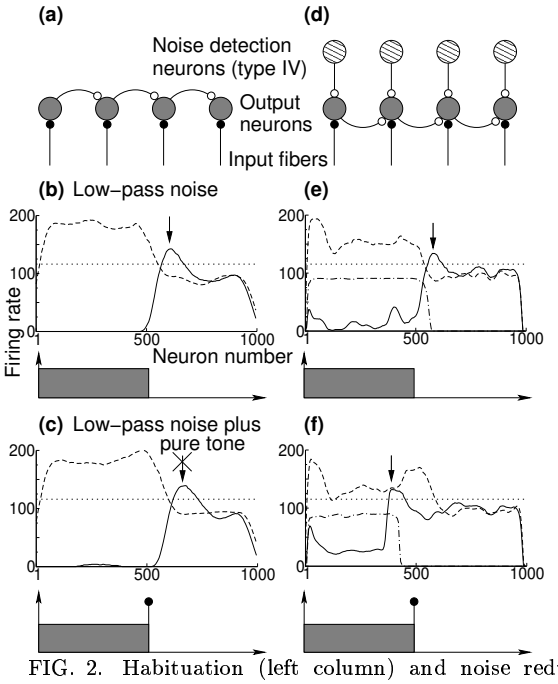

FIG. 2. Habituation (left column) and noise reduction (right column). (a) Neuronal implementation of asymmetric lateral inhibition. Grey circles denote output neurons, small filled circles indicate excitatory and small open circles inhibitory synapses. (b),(c) Response (upper panels) of the habituation model (a) to the sounds (lower panels) of Figs. 1(a),(e). Firing rates of the output neurons before (horizontal dotted line, spontaneous rate), during (dashed line) and immediately after (solid line) the sound presentation are shown. As in (e) and (f), they are the result of a numerical simulation of a single run of 1000 neurons with best frequencies between 0.2 and 16 kHz; for simulation details, see [18,19]. Downward arrows indicate Zwicker tones predicted by the model. In case (c), habituation predicts a Zwicker tone (crossed arrow) to the right of the pure tone whereas in experiment there is one *to the left*. (d) Neuronal implementation of the full model with asymmetric inhibition and noise detection. The responses in (e) and (f) correspond to (b) and (c), respectively. Dash-dotted lines stand for firing rates of noise-detection neurons. Their distribution of firing rates is shifted to the left as we go from (e) to (f), which is due to the pure tone's 'hole burning' in (f); see Fig. 3 and the text below.

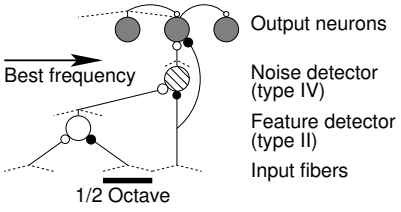

FIG. 3. Implementation of the noise-reduction model through noise-detection (type IV) neurons. Same notation as in Fig. 2(d), with best frequency (arrow) as horizontal axis. The horizontal bar indicates half an octave, which has been emulated here by 80 neurons. The feature detectors (type II) are active, if there is a steeply rising edge, i.e., a rising spectral intensity as frequency increases. Feature detectors asymmetrically inhibit noise-detection neurons, which, in turn, inhibit corresponding output neurons; see [18] for the details of the synaptic connectivity (dashed lines).

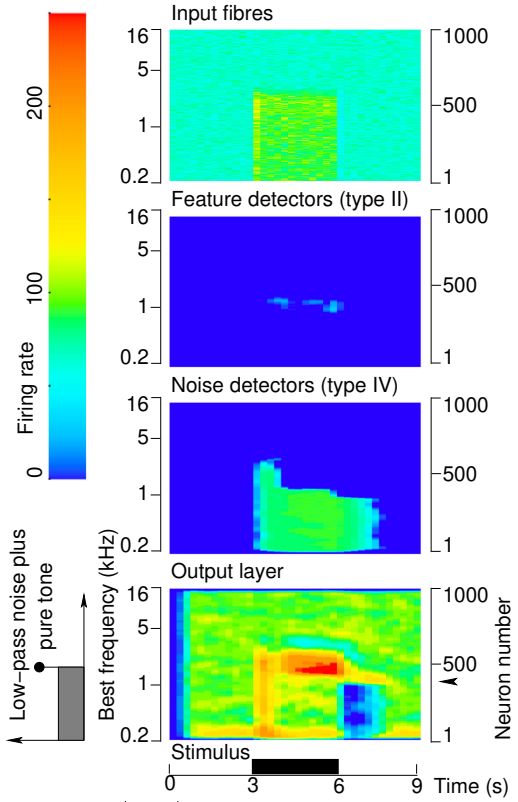

FIG. 4. (color) Numerical simulation of a Zwicker tone. Firing rates (color coded) of 1000 output neurons are a function of time and the neurons' best frequency. Frequencies are equidistant on a logarithmic scale. The stimulating sound is low-pass noise plus a pure tone at the edge, 1700 Hz; cf. Figs. 1(e) and 2(f). Sound is on between 3 and 6 s. Feature detectors respond to the pure tone, causing a 'hole burning' in the noise-detector layer since noise-detector neurons are inhibited by feature detectors; cf. Figs. 2 & 3. The Zwicker tone, brought about by the mechanism described in the main text, is perceived *below* the pure tone and is indicated by an arrow head, pointing at the yellow strip. Data are binned. A bin consists of the rates of a neuron and its two left and right neighbors. Each data point is an average over its bin and 250 ms.