Modeling the cochlear nucleus: A site for monaural echo suppression?

Moritz Bürck[a] and J. Leo van Hemmen
Physik Department and BCCN—Munich, Technische Universität München, 85747 Garching bei München, Germany

(Received 11 September 2006; revised 19 June 2007; accepted 19 July 2007)

Echo suppression plays an important role in identifying and localizing auditory objects. One can distinguish between binaural and monaural echo suppression, although the former is the one commonly referred to. Based on biological findings we introduce and analyze a mathematical model for a neural implementation of *monaural* echo suppression in the cochlear nucleus. The model's behavior has been verified by analytical calculations as well as by numerical simulations for several types of input signal. It shows that in the perception of a pair of clicks the leading click suppresses the lagging one and that suppression is maximal for an interclick interval of 2–3 ms. Similarly, ongoing stimuli will be affected by the suppression mechanism primarily a couple of milliseconds after onset, resulting in a reduced perception of a sound shortly after its start. Both effects match experimental data. © *2007 Acoustical Society of America.* [DOI: 10.1121/1.2770545]

PACS number(s): 43.66.Ba, 43.66.Qp, 43.66.Lj [RYL] Pages: 2226–2235

I. INTRODUCTION

In any natural environment sound is reflected. The consequence is that every sound signal is followed by countless reflections, that is, echoes from many directions. This means that any biological system processing acoustic signals has to cope with reflections interfering with the original signal. Hence, a neuronal mechanism suppressing the confusing information of echoes is advantageous for sound perception.

Humans do not consciously perceive echoes arriving less than about 20 ms after the original signal.[1] A simple experiment shows that echo suppression is in part binaural and in part monaural. In a large room, for example a lecture room, one can perceive the otherwise suppressed echoes by covering one ear. In a small room, where the echoes are faster, this does not work.[2] Thus, neuronal echo suppression consists of a slower, binaural mechanism using both ears and a faster, monaural mechanism using only the cues one ear can provide. The focus of the research presented here is on the monaural mechanism for echo suppression.

II. MONAURAL ECHO SUPPRESSION: PSYCHOPHYSICAL AND BIOLOGICAL EVIDENCE

In 1963 Harris, Flanagan, and Watson investigated the binaural interaction of a click with a click pair.[3] In their experiment, a click pair was presented to one ear of the subjects while a single click was presented to their other ear. The subjects were asked to adjust the single click in time so that they perceive an auditory event straight ahead. The results imply that the second click of a click pair is not perceived, even unconsciously, if the interval between leading and lagging click is 2 ms. The second click *is* perceived if the interclick interval is *more* than 2 ms (in the experiment: 4, 6, and 8 ms) and, more interestingly, also if the interclick interval is *less* than 2 ms (in the experiment: 0.5 and 1 ms). The authors conclude from these results that the suppression of the second click is therefore neither the result of a neuronal refractory period nor a function of mechanical properties of the basilar membrane. Instead they propose a *neural gate* that closes at about 1 ms after the start of the neuronal activity and opens after another 2 ms. Harris *et al.* suggested the cochlear nucleus as the site of this neural mechanism.

In 1980, Zurek referred to the idea of a neural gate as a possible explanation for the results he had obtained.[4] In his experiments, Zurek measured the time course of just-noticable differences in interaural time and intensity using pairs of short (less than 5 ms) as well as long (50 ms) noise bursts. His principle finding was that the interaural sensitivity to changes in both time and intensity follows a nonmonotonic course after the onset of a sound. For about 0.5 to 10 ms after the beginning of the signal, the perception is degraded. *This degradation is maximal at* 1–5 ms.

A couple of years later, Hafter, in cooperation with several colleagues, published a number of experiments concerning the perception of clicks in a click series, especially as a function of click rates.[5–7] They found that clicks start influencing each other at interclick intervals shorter than 10 ms. This interference can be explained neither by neural refractoriness nor by narrow-band filtering within the auditory system nor by dependence of successive samples of internal noise. The authors suggested a "neural saturation process." This neural saturation process was the topic of further research where Hafter and colleagues have found the saturation to be a *monaural, frequency-specific* phenomenon.[7] As a location for the neural mechanism, they proposed the cochlear nucleus.

That is, the cochlear nucleus plays a key role in what we now call monaural echo suppression. Two major functional subdivisions of the cochlear nucleus complex are the anteroventral and the dorsal cochlear nucleus (AVCN and DCN). Both receive input from the auditory nerve (AN) and

[a]Electronic mail: mbuerck@ph.tum.de

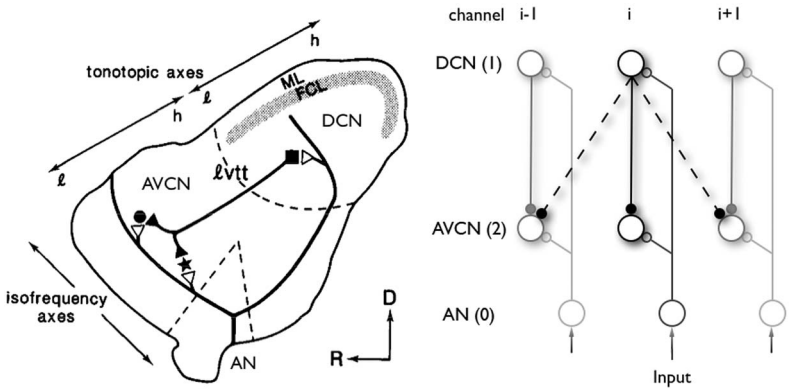

FIG. 1. Neuroanatomical structure in the cochlear nucleus as considered here. Left: sketch of Wickesberg and Oertel.[2] Both subdivisions of the cochlear nucleus are organized tonotopically. The auditory nerve (AN) is connected to AVCN (black circle and star) and DCN (black square) cells via excitatory synapses (white triangles). Cells of the DCN are connected to cells of the AVCN receiving input from the same fiber of the auditory nerve by inhibitory synapses (black triangles). Right: model of a neural frequency channel. Inhibitory synapses are represented by small black circles, excitatory synapses by small white circles. The inhibitory connections shown in dashed lines are considered in the numerical calculations only. They represent the spreading of inhibition into neighboring frequency channels, which is characterized by the quantity IS (*I*nhibitory *S*pread). For IS=1, there is no spreading, whereas for IS=n the neighboring $n-1$ nerve fibers receive inhibitory input. In the present figure (right), IS=3.

there is also a projection from the DCN to the AVCN. From the AVCN, the signals are transmitted to higher centers in the auditory brainsteam (see Fig. 1, left).

From 1988 to 1990, Wickesberg and Oertel carried out several elucidating physiological experiments in the cochlear nucleus.[2,8] They investigated the properties of the projection from DCN to AVCN, which they found to be *frequency specific*. That means that cells in the DCN and in the AVCN that are connected to each other receive input from the same fiber of the auditory nerve. Furthermore, the projection was *inhibitory* for all cells in their experiments. The results are illustrated in Fig. 1, left. As for the timing of the inhibition, they found that inhibitory postsynaptic potentials reach the AVCN 2 ms after the stimulation of the auditory nerve. Excitatory postsynaptic inputs, however, reach the AVCN 0.7 ms after stimulation. The authors concluded that action potentials in the auditory nerve trigger an inhibition that can suppress later signals and that *the suppression is maximal if the delay of the second signal is* 2 ms. They proposed a role in echo suppression.

The remarkable match of biological and psychophysical findings has led us to model the circuitry Wickesberg and Oertel discovered, leading to a possible neural process of monaural echo suppression. The scope of the present work is thus to validate the intuitive interpretation of the biological evidence on a quantitative basis, providing an answer to the question of what the *intrinsic* properties of the neuronal projection from DCN to AVCN are.

III. MODELING THE COCHLEAR NUCLEUS

The layout of the model is simple (see Fig. 1, right). We have three different populations of neurons: the neurons of the auditory nerve (population 0), of the dorsal (pop. 1), and of the anteroventral cochlear nucleus (pop. 2). In order to keep the model as simple as possible, we will use identical parameters for the neurons in the AVCN and DCN for analytical as well as numerical calculations, in this way reducing the number of model parameters significantly. Our generic neurons can be considered to be bushy cells. This is not far-fetched since Wickesberg and Oertel,[2] whom we owe our cell-parameter values, studied both bushy and stellate cells in the AVCN and found similar inhibitory response types among both cell types. The connections between the different populations have been modeled to match the findings of Wickesberg and Oertel[2,8] as well. Each neuron of the auditory nerve is connected excitatorily to one neuron of the DCN and to one neuron of the AVCN while the neuron in the AVCN is receiving inhibitory input from the cell in the DCN. Every three neurons that are connected as described above form one separate neuronal channel i. Such a channel is receiving input of only one frequency f. In other words, signal processing is frequency specific. Since a strict separation of frequencies may not be realized in actual biological systems, we allow the inhibition to spread into the neighboring channels (dashed lines in Fig. 1, right). Such a spreading of inhibition has indeed been reported in the DCN.[9,10] In order to find an analytical solution, however, we will assume strict frequency separation in the next subsection.

A. Analytical calculations

1. Methods

For our analytical calculations, we assume the firing of a neuron to be a statistical process. That is, an *inhomogeneous* Poisson process, which is defined by three properties. First, the probability of finding a spike between t and $t+\Delta t$ is $\lambda(t)\Delta t$, so $\lambda(t)$ is the time-dependent firing probability density or rate function. Second, the probability of finding two or more spikes there is $o(\Delta t)$, which means that we ignore their occurrence for small Δt. Third, a Poisson process has independent increments, i.e., events in disjoint intervals are independent.

A neuron i obeying such statistics with rate function

$$\lambda_i(t) = \nu^0 + v(t) = \nu^0 + \sum_{j,f} J_{ij}\epsilon(t - t_j^f - \Delta t_{ij}; \tau), \tag{1}$$

is then called a *Poisson neuron*; (see Refs. 11 and 12 for details). Here ν^0 is a spontaneous firing rate (which, for the sake of simplicity, we set to zero) and $v(t)$ is the cell poten-

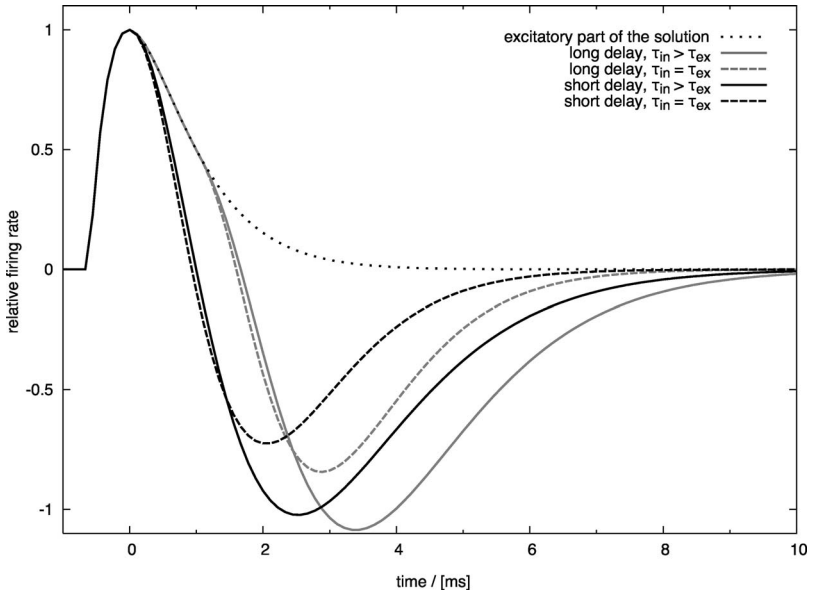

FIG. 2. Solution of Eq. (3) for a click (i.e., the impulse response). We have plotted the relative firing rate in arbitrary units versus time in milliseconds. The peaks of the solutions are aligned at 0 ms. Dotted line: excitatory part of the solution. Solid and dashed lines: solution for different time constants, black for a short and gray for a long delay of the inhibition. After rising, the firing rate drops below the spontaneous activity for all solutions, thus suppressing subsequent stimuli. The behavior is consistent with the psychophysical experiments[2] described before. $J_{10}=J_{20}=1$, $J_{21}=-0.9$, $\Delta t_{20}=\Delta t_{10}=\Delta t_{21}=0.6$ ms (black lines) and 1.6 ms (gray lines) and $\tau_{ex}=0.6$ ms. $\tau_{in}=0.6$ ms (dashed lines) and 1 ms (solid lines).

tial generating the spikes. It is determined by a sum over all input neurons j and their firing times t_j^f (abbreviated by f in the sum). The axonal delay between neuron j and neuron i is given by Δt_{ij}. Furthermore, J_{ij} is the synaptic strength of the coupling from neuron j to neuron i; it is positive for an excitatory and negative for an inhibitory coupling. The response of the cell potential to an incoming spike, the postsynaptic potential (PSP), is described by $\epsilon(t;\tau)$. We model the PSP for excitatory as well as inhibitory inputs by an alpha function,

$$\epsilon(t;\tau)=\begin{cases}(t/\tau)\exp(1-t/\tau), & t\geqslant 0,\\ 0, & t<0,\end{cases} \tag{2}$$

with τ as the characteristic time constant, usually of the order of milliseconds. Note that τ may be different for inhibitory PSPs and excitatory PSPs (in our case $\tau_{ex}=0.6$ ms or $\tau_{in}=1$ ms, respectively).

The firing rate as defined above refers to a stochastic quantity. We are interested in the expectation value of the rate rather than a specific realization of the stochastic firing process since the process considered here is self-averaging.[12] It can be shown that instead of summing over the spikes in Eq. (1), we can then integrate over the expectation value of the input spikes.[12] As a consequence, the expectation value of the rate function of the neurons in the AVCN is

$$\begin{aligned}\lambda_2(t)=&J_{20}\int_0^\infty ds\,\epsilon(s;\tau_{ex})F(t-\Delta t_{20}-s)\\&+J_{10}J_{21}\int_0^\infty ds\,\epsilon(s;\tau_{ex})\int_0^\infty ds'(s';\tau_{in})\\&\times F(t-\Delta t_{10}-\Delta t_{21}-s-s'),\end{aligned} \tag{3}$$

where we assume that the rate function of the neurons in the auditory nerve is proportional to the input function $F(t)$. We define an input function to be the firing rate in the auditory nerve induced by an external stimulus.

In the case of identical time constants for excitatory and inhibitory PSPs ($\tau:=\tau_{ex}=\tau_{in}$), the resulting "impulse response" (the solution to a delta function as approximation of a click) is given by

$$\lambda_2(t)=J_{20}\epsilon_\tau(t-\Delta t_{20})+J_{10}J_{21}(\epsilon_\tau * \epsilon_\tau)(t-\Delta t_{10}-\Delta t_{21}), \tag{4}$$

with $\epsilon_\tau(t):=\epsilon(t;\tau)$ and "$*$" denoting the convolution; calculated explicitly,

$$\begin{aligned}\lambda_2(t)=&J_{20}\epsilon_\tau(t-\Delta t_{20})+J_{10}J_{21}\frac{e^1(t-\Delta t_{10}-\Delta t_{21})^2}{6\tau}\\&\times\epsilon_\tau(t-\Delta t_{10}-\Delta t_{21}).\end{aligned} \tag{5}$$

The first term on the right in (3)–(5) represents the excitatory influence of auditory-nerve activity and the second term stands for the inhibitory influence of the activity in the DCN, which is driven itself by an excitatory input from the auditory nerve. In (4) and (5) we can see that the inhibition, in comparison to excitation, is delayed by the time shifting and smeared out by the convolution with the second PSP. It nevertheless follows a similar time course as the excitation, viz., that of an EPSP. The dashed lines in Fig. 2 show realizations of (4) and (5) (see the next section for a more detailed description).

We have solved (3) for two input functions, one representing a click and one representing the onset of a pure tone. We have chosen the delta function to simulate the click, viz., $F(t)=\delta(t)$. As for the pure tone, we had to choose a slightly more complex function. The onset is described by a Heaviside stepfunction; $\Theta(t)=1$ for $t\geqslant 0$ and 0 for $t<0$. For an oscillation we have to take into account that only the compressional part of the signal, that means an increase in the air pressure, evokes neural activity in the auditory system which, in addition, adapts rapidly to an ongoing signal. So, the half-wave rectified part of a sine is used to model the activity in the auditory nerve in response to a pure tone in combination with a decaying exponential term. Since this would lead to quite a bulky expression in the analytical so-

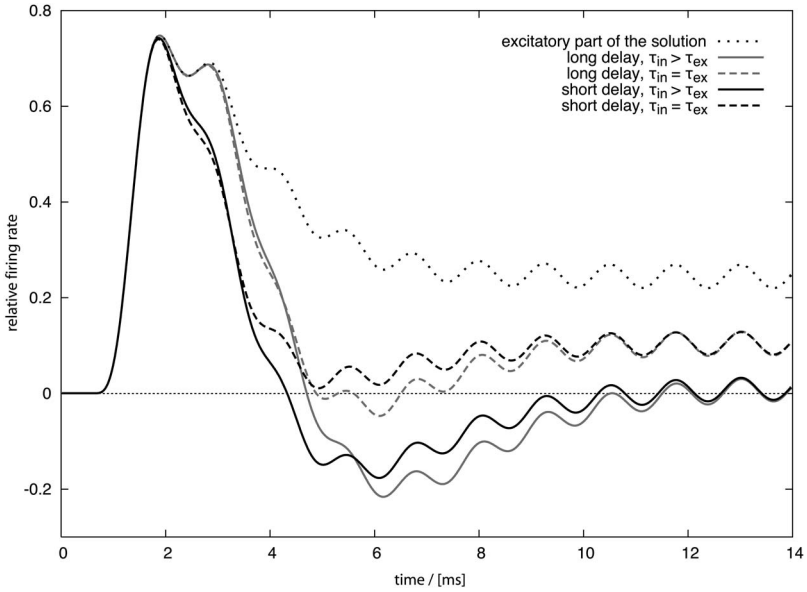

FIG. 3. Solution of Eq. (3) for the onset of a pure tone. We plot the relative firing rate in arbitrary units versus time in milliseconds. The curves are aligned similarly to Fig. 2. Dotted line: excitatory part of the solution. Solid and dashed lines: solution for different time constants, black for a short and gray for a long delay of the inhibition. The inhibition results in a drop of the firing rate shortly after the onset of the signal, which is equivalent to the decreased sensitivity after the beginning of a signal reported in psychophysical experiments.[4] Note that the amplitude of the oscillation is identical for all three curves, which means that the phase information of the signal is not influenced by the inhibition. $J_{10}=J_{20}=1$, $J_{21}=-0.35$, $\Delta t_{20}=\Delta t_{10}=\Delta t_{21}=0.6$ ms (black lines) and 1.6 ms (gray lines), $\tau_{ex}=0.6$ ms and $f=800$ Hz. $\tau_{in}=0.6$ ms (dashed lines) and 1 ms (solid lines).

lution, yet another function was used, $F(t)=[1-\cos(\omega t)][0.15+0.85\exp(-t/\tau_{adapt})]$ with $\tau_{adapt}=1.1$ ms.[30] By using this function we assume the neurons in the auditory nerve to be slower and less precise than they actually are.

2. Results

Four solutions for a delta-function click as input are shown in Fig. 2 in comparison to their purely excitatory part (solid and dashed lines in comparison to the dotted line). The solid lines represent the solutions for identical time constants of inhibition and excitation whereas for the dashed lines the time constant of the inhibition was larger. Based on Wickesberg and Oertel's *in vitro* experiments,[2] the axonal delay from DCN to AVCN was set to 1.6 ms (gray lines). The shorter axonal delay of 0.6 ms (black lines) is in accordance with *in vivo* experiments performed by Wickesberg in 1996.[13]

After the click has been fed into the neural network, the firing rate begins to increase slowly, then grows faster and reaches a clear maximum before decreasing again. For the relevant case, viz., the case of a short axonal delay (0.6 ms), the inhibition results in a significant sharpening of the peak as discovered before experimentally[13] and in a drop of the firing rate below the spontaneous firing rate shortly after the signal. Most important, there is a *pronounced minimum* of the firing rate corresponding to a maximum of the inhibition. For identical time constants of excitation and inhibition, this minimum is found at about 2 ms after the firing rate reaches its maximum (dashed black line in Fig. 2). As the time constant of the inhibition increases, the minimum moves away from the maximum (solid black line in Fig. 2) and gets stronger. We note that a longer axonal delay from auditory nerve to AVCN via DCN ($\Delta t_{10}+\Delta t_{21}$) does not influence the strength or the form of the excitation or inhibition, but only shifts the start of the IPSP in time (gray lines in Fig. 2).

It is important to realize that except for the weights, which have not been determined physiologically, the parameters used to obtain the results are not the result of any arbitrary fitting. They were, instead, taken from the publications by Wickesberg and Oertel[2] and by Wickesberg[13] and are in the typical range for neurons in the cochlear nucleus.[14–16]

Since the system is effectively linear, the solution for two subsequent clicks is the sum of the solutions for the individual clicks. That means *a second click following the first one will be suppressed* and the suppression will be maximal at a delay of 2 to 2.5 ms, depending on the time constant of the inhibition. The behavior of the model thus matches the experimental results of Harris, Flanagan, and Watson very well.

As for a pure tone, several solutions for a beginning tone of 800 Hz are shown in Fig. 3. All the solutions have in common the drop of the firing rate shortly after the onset of the signal due to the inhibition. For the solutions with a large inhibitory time constant, the firing rate drops below the spontaneous firing rate, consistent with psychophysical data indicating decreased sensitivity immediately after the onset of an ongoing stimulus.[4] Then the firing rate increases and reaches a level that depends on the time constant of the inhibition. By a change of the strength of the inhibition, the time constant therefore shifts the overall level of activity, which is an effect of the normalization of the PSPs. On closer examination, however, it becomes evident that the amplitude of the oscillation is identical for all three graphs plotted in the figure. This means that the phase information of the signal is preserved. Thus the inhibition does not degrade the information on the frequency of the pure tone but, in our model, even improves the signal-to-noise ratio.

Altogether, the behavior of the model matches the psychophysical results cited above. A single click will suppress a subsequent click and the suppression is maximal a couple of milliseconds after the first click.[3] The perception of an ongoing stimulus is distorted shortly after its onset, whereas its perception later on is not disturbed.[4]

B. Numerical calculations

The consistency of our solution and the psychophysical evidence is remarkable, but it is still hard to estimate how close our analytical results are to the behavior of the real

biological system. Though the parameters used are biologically plausible, our neuron model has ignored an essential characteristic of any neuron, the threshold. The threshold, however, will play an important role at the beginning of any signal processed so that our results are encouraging but do not necessarily provide a biologically relevant explanation of the effect described in Sec. I. In the next section, we will use another method for modeling the neurons in our system. The new model will feature an explicit threshold and therefore overcome the limitations of the Poisson neuron. The tradeoff is in the loss of the analytical solution.

1. Methods

The model we use in this section is an extension of the leaky integrate-and-fire neuron, the *spike response model* (SRM).[12] In the SRM description, spikes are represented by delta functions and absolute as well as relative refractory periods are taken explicitly into account. The cell potential $v_i(t)$ of a cell i receiving input from the cell j is given by

$$v_i(t) = \sum_f \eta(t - t_i^f) + \sum_{j,f} J_{ij}\epsilon(t - t_j^f - \Delta_{ij}^{\text{axon}}). \tag{6}$$

As before, the sum over f is a sum over the firing times t^f of the neurons. The new terms $\eta(t-t_i^f)$ represent the refractory behavior of the neuron i we are focussing on and are described by

$$\eta(t) = \begin{cases} -\infty, & t < \tau_{\text{abs}}, \\ -N\exp(-t/\tau_{\text{rel}}), & t \geqslant \tau_{\text{abs}}. \end{cases} \tag{7}$$

In our simulations τ_{abs} is taken to be 0.25 ms and N equals two times the maximum value a single EPSP reaches. As before, J_{ij} is the strength of the coupling from neuron j to neuron i and ϵ is the postsynaptic response to a spike. Again τ will be different for excitatory and inhibitory PSPs (0.6 and 1 ms). The axonal delay $\Delta_{ij}^{\text{axon}}$ is composed of the synaptic delay and the finite speed of propagation of a spike along the axon. A spike will be fired by neuron i if the potential $v_i(t)$ exceeds a threshold ϑ which is set to 0.9 of the maximum value an EPSP reaches.

Consequently, for both parameter sets the DCN is primarily acting as a relay of the AN spikes, delaying them temporally and making their action inhibitory. The free parameters are the strength of the inhibition, J_{in}, and the spreading of the inhibition in neighboring frequencies, IS (*I*nhibitory *S*pread). For IS=1, there is no spreading whereas for IS=n, the neighboring $n-1$ nerve fibers will receive inhibitory input (see Fig. 1). The strength of the inhibition decreases exponentially with the distance between the nerve fibers. If not specified otherwise, we have used a minimal spreading (IS=5) and a medium strength of inhibition ($J_{\text{in}}=-0.8$).

The cell parameters were extracted from Refs. 2 and 17. With regard to the different values for the relative delay of inhibition with respect to excitation in the AVCN found *in vitro*[2] (1.6 ms) and *in vivo*[13] (0.6 ms), we have checked the behavior of our model for both short and long delay (data not shown). In addition we have tested the robustness of the model through two different values for the time constant of the relative refractory period, $\tau_{\text{rel}}=0.3$ ms as it has been measured by Wu and Oertel[17] and $\tau_{\text{rel}}=1$ ms (data not shown). The parameters are summarized in Table I.

TABLE I. The parameters as used for the numerical simulations. The first parameter set is consistent with Wickesberg and Oertel[2], Wu and Oertel[17], and Wickesberg[13]. The parameters in brackets lead to very similar (axonal delay) or virtually identical behavior (time constant of the relative refractory period) as the primary parameter set (data not shown). The strength of the inhibition was varied to optimize the behavior of the model.

τ_{rel}	0.3(1.0) ms
τ_{ex}	0.6 ms
τ_{in}	1 ms
τ_{abs}	0.25 ms
ϑ	0.9
J_{ex}	1
J_{in}	−0.8
$\Delta_{ij}^{\text{axon}}$	0.6(1.6) ms

As for modeling the periphery of the auditory system, we have taken advantage of the C++ package LUTEar,[18] which has been developed at the university of Essex. LUTEar reproduces the spike times in the auditory nerve fibers corresponding to any wav-file. Version 2.0.9 was used to calculate the spike times in 500 frequency channels, which cover a frequency range from 200 Hz to 16 kHz. The channels are distributed uniformly on a logarithmic scale, resulting in a doubling of frequency every 79 channels. Each frequency channel consists of one nerve fiber. The output of one nerve fiber is the input for one cell in the AVCN and DCN of our model (Fig. 1). We use the spike times of clicks, pairs of clicks, and white noise as well as a pure tone (data not shown).

2. Results

As in the analytical calculations, the behavior of the model is consistent with the results of Harris, Flanagan, and Watson.[3] Figures 4 and 5 illustrate the behavior of the model for a click series. A single click is followed by eight pairs of clicks with interclick intervals (ICIs) of 0.5, 1, 2, 3, 4, 6, 8, and 10 ms, each pair being separated from its neighboring pairs by 40 ms.

As the interclick interval increases in the auditory nerve (Fig. 4), the bands of activity become broader until, at a separation of 2 ms, we can distinguish the individual clicks. Every click is followed by an oscillating activity in the lower frequencies, which is the result of natural oscillations of the basilar membrane. At the stage of the cochlear nucleus, these oscillations almost disappear (see Fig. 5).

Furthermore, we notice that the overall activity is reduced almost by a factor of 2 in comparison to the activity in the auditory nerve [15 794 spikes in the AN (Fig. 4) and 8893 in the AVCN (Fig. 5)]. At a second glance, it becomes apparent that the first four pairs of clicks are indistinguishable. It is only at an interclick interval of 4 ms that the second click appears, which gets stronger as the separation between the clicks increases. In Figs. 4 and 5, the time constant of the relative refractory period is 0.3 ms and the axonal delay is 0.6 ms, but for $\tau_{\text{rel}}=1$ ms the graphical representa-

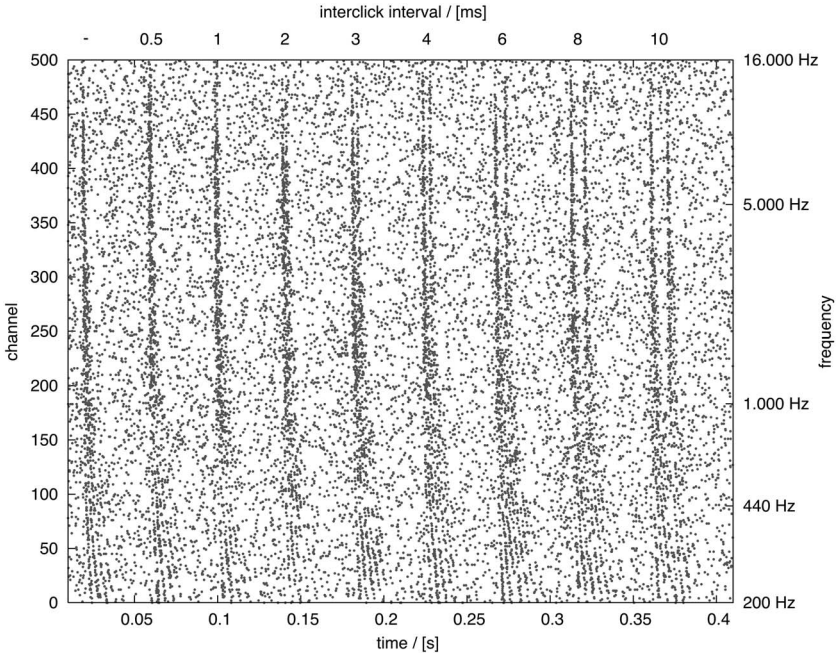

FIG. 4. Spike times for a series of clicks in the auditory nerve. The horizontal axis denotes time, whereas the vertical axis exhibits channel number (left) and corresponding frequency (right). One single click is followed by eight pairs of clicks with an interclick interval of 0.5, 1, 2, 3, 4, 6, 8, and 10 ms. Each pair of clicks is separated from the neighboring pairs by 40 ms of silence. The activity corresponding to the individual clicks as well as the natural oscillation of the basilar membrane can be discerned clearly. Calculated with LUTEar.[18]

tion of the neural activity in the cochlear nucleus is practically indistinguishable (data not shown). Increasing the axonal delay to 1.6 ms leads to a very similar result where the second click is suppressed for ICIs of 2 and 3 ms but at the same time the response to the single click as well as to the click pairs for an ICI of 0,5 and 1 ms is slightly broader in time (data not shown). An increase in the spreading of the inhibition (IS$>$5) results in an almost identical behavior if the strength of the inhibition is reduced simultaneously (data not shown).

The spike histogram in Fig. 6 provides a more detailed view of the click suppression illustrated in Figs. 4 and 5. We see the spike activity evoked by a single click (light transparent gray) compared to that evoked by a click pair (dark gray) for the same set of interclick intervals as used by Harris *et al.*[3] in the auditory nerve (left) and the cochlear nucleus (right). The data have been obtained by adding ten runs and calculating the running average over 0.0228 ms in the upper 250 frequency channels, covering the frequency range from 1790 Hz to 16 kHz. In the auditory nerve the two clicks of a click pair can be discerned clearly for all interclick intervals but do not form distinct events for ICIs of 1 and 0.5 ms. In the cochlear nucleus the second click is suppressed for interclick intervals of 2 and to some extent also of 1 ms but not of 0.5 ms (width at half height is 1.7 ms for an ICI of 2 ms, 1.8 ms for an ICI of 1 ms, and 1.7 ms for a single click). The suppression in our model is therefore not due to the refractory period of the neurons and our results agree with the results of Harris *et al.*[3] in that the second click of a click pair can be perceived for *very* short and medial but not for short interclick intervals. We note that Harris *et al.*[3] reported that the second click of a click pair is perceived consciously only

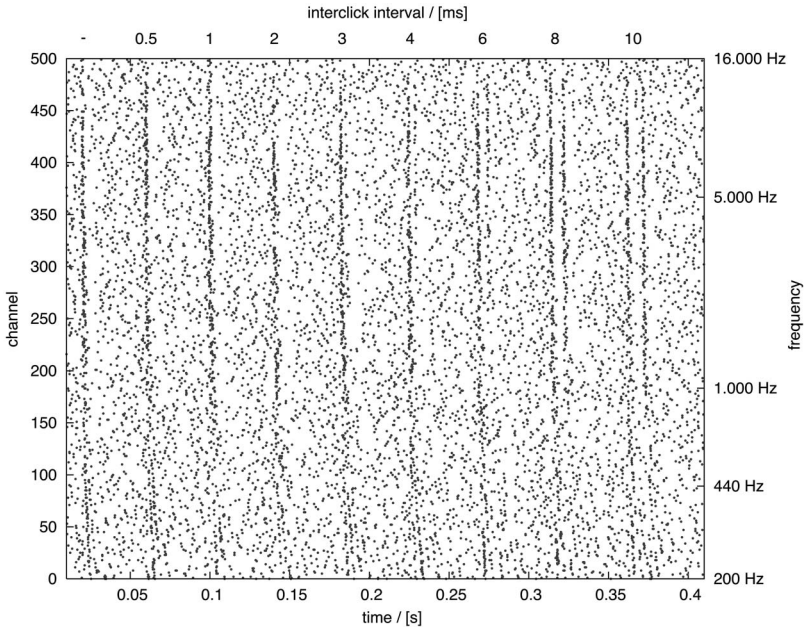

FIG. 5. Spike times for a series of clicks in the cochlear nucleus, plotted in a way identical to the one of Fig. 4. The natural oscillations of the basilar membrane almost vanish and the first three pairs of clicks are indistinguishable from each other. Especially the click pair with an interclick interval of 2 ms differs from that of Fig. 4 in that the second click has disappeared clearly. The behavior of the model matches the results of Harris *et al.*,[3] where the second click is not perceived at an interclick interval of 2 ms. Δt_{12}=0.6 ms, τ_{rel}=0.3 ms, and IS=5. The inhibitory spread (IS) has been defined in Fig. 1.

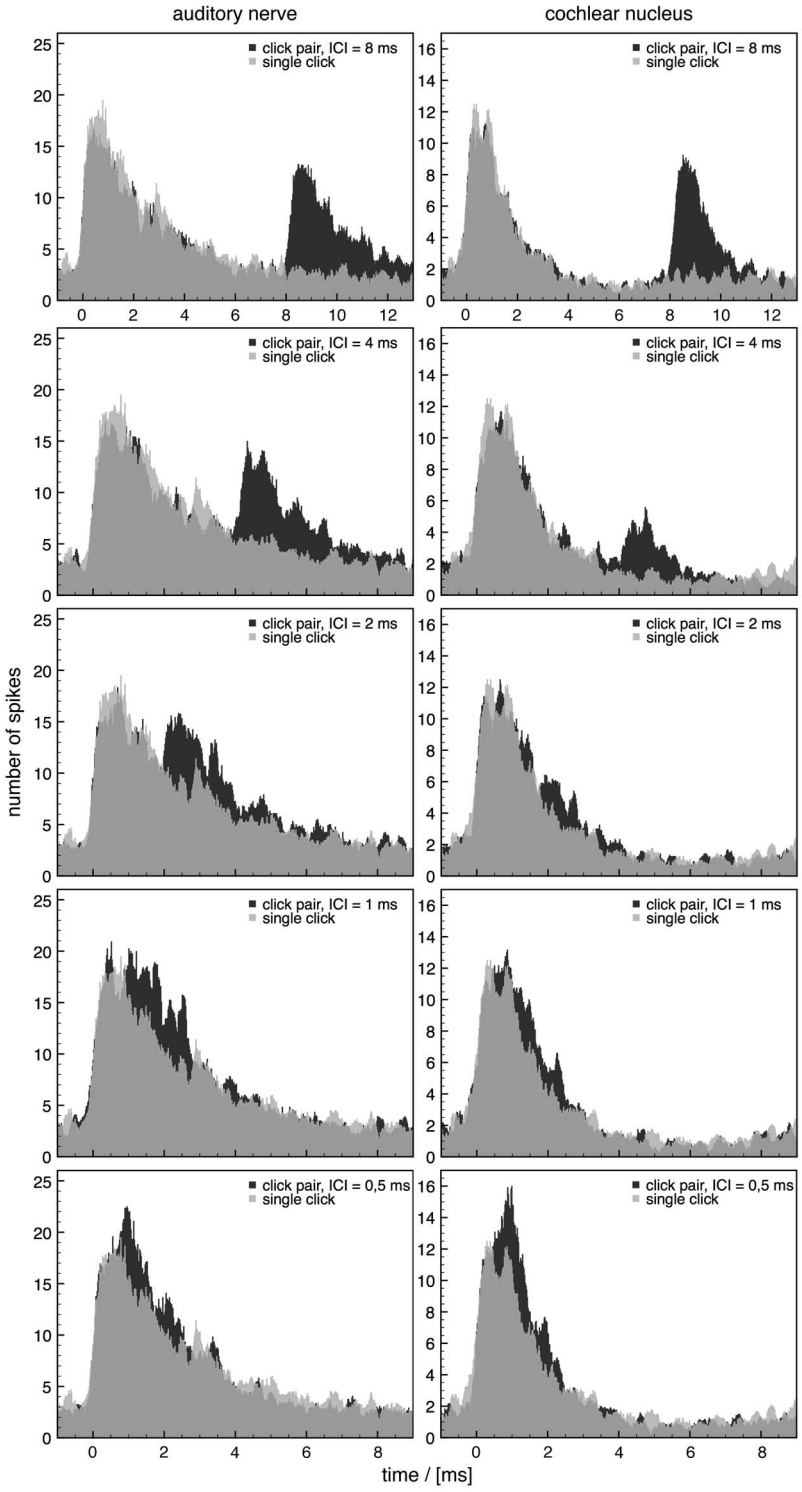

FIG. 6. Comparison of the spike activity evoked by a single click (light transparent gray) to that evoked by click pairs (dark gray) in the auditory nerve (left) and in the cochlear nucleus (right) for interclick intervals (ICI) of 0.5, 1, 2, 4, and 8 ms (bottom to top). The data have been obtained by adding several runs (ten repetitions, bin size=0.0019 ms) and calculating a running average (12 bins=0.0228 ms); only the activity in channels 250 to 500 (1790 Hz to 16 kHz) has been taken into account. All pairs of clicks can be discerned clearly from a single click in the auditory nerve, whereas in the cochlear nucleus the second click of a click pair is suppressed at an ICI of 2 ms. The visible presence of the second click at an ICI of 1 and 0.5 ms in the cochlear nucleus proves that the suppression at an ICI of 2 ms is not due to a neuronal refractory period or the like. Parameter values are $\Delta t_{12}=0.6$ ms, $\tau_{rel}=0.3$ ms, and IS=5.

at interspike intervals of 4 and 8 ms and that in our simulations the two clicks form a clearly distinct event only at the very same click delays.

A closer examination of the activity corresponding to a single click reveals two peaks of activity in the cochlear nucleus but *not* in the auditory nerve (e.g., top row of Fig. 6). Interestingly, 1 ms after a *single click*, that is even when there was *no second click*, a click was perceived in a significant portion of the trials in Harris *et al.*'s experiment.[3] The suppression of activity in the trial of each click results in a sharpening of the peaks, which was described by Wickesberg before.[13]

In addition to the response of the model to a click series, we have evaluated the response to a pure tone (data not shown) and to white noise. In order to understand the behavior of the model we have analyzed at the interspike interval distribution in the cochlear nucleus across all frequency channels during half a second amidst a lasting pure tone of 440 Hz (data not shown) and ongoing white noise. The sound pressure level was 100 dB(A), which is relatively loud

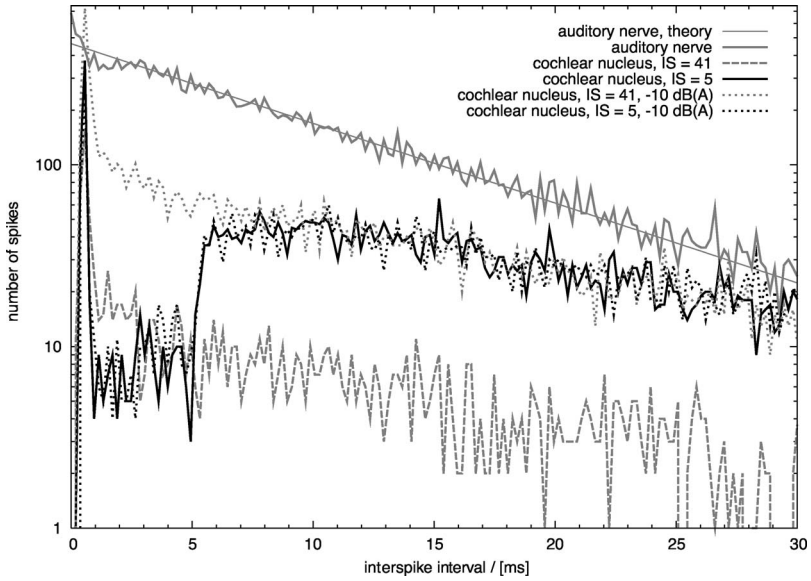

FIG. 7. Response to white noise. Distribution of spikes summed over all frequency channels in the auditory nerve and the cochlear nucleus, plotted as a response (vertical axis) to ongoing white noise in dependence upon interspike intervals (horizontal axis). Solid gray lines: activity in the auditory nerve (thick line) and the theoretical interspike interval distribution for random spikes (straight thin line). Dashed gray and solid black line: activity in the cochlear nucleus. Dotted gray and black lines: activity in the cochlear nucleus for a sound pressure level reduced by 10 dB (A). In the case of small spreading of the inhibition (IS) into neighboring frequency channels, spikes with interspike intervals less than 5 ms are subject to significant inhibition whereas spikes with larger interspike intervals are only slightly inhibited (black line). This is consistent with psychophysical data indicating that the perception of an ongoing stimulus is degraded a few milliseconds after its onset.[4] Greater spreading inhibition results in equal spike suppression for both small and large interspike intervals; the characteristic inhibition at small interspike intervals is lost (dashed and dotted gray line). A reduction in the sound pressure level results in a significant increase of activity for larger inhibitory spreading (dashed and dotted gray line) but does not change the activity for small inhibitory spreading (solid and dotted black line). Parameter values are $\tau_{rel}=0.3$ ms, $J_{in}=-0.8$, and IS=5 (black lines) and $J_{in}=-0.05$ and IS=41 (gray dashed and dotted line). IS=41 means that the inhibition spreads into the neighboring 40 frequency channels where 79 channels correspond to one octave (see also Fig. 1).

but improves the statistics by increasing the number of spikes. The interspike interval distribution in Fig. 7 represents ongoing white noise; the distribution for a lasting pure tone bears the same characteristics. The gray solid lines represent the interspike interval distribution in the auditory nerve, the thick line the actual data and the thin line the distribution expected for a stochastic point process. The black solid line corresponds to the distribution in the cochlear nucleus with IS=5 and the dashed gray line to that with IS=41. The dotted lines represent the spike activity in response to white noise at a sound pressure level of 90 dB(A), the gray dotted line for a large (IS=41) and the black dotted line (almost identical to the solid black line) for a small (IS=5) inhibitory spreading.

The large inhibitory spread of about half an octave (IS =41) leads to a decreased overall activity, a reduction of the signal intensity similar to the analytical results obtained before. Compared to the activity in the auditory nerve, the characteristics of the distribution change only slightly. This contrasts with the result for a small spreading of the inhibition into neighboring frequencies (IS=5). Here the overall intensity is reduced as well, but spikes with interspike intervals less than 5 ms are suppressed disproportionately to spikes with larger interspike intervals. The drop in activity at small interspike intervals corresponds well with reduced sensitivity in the perception of an ongoing stimulus shortly after its onset.[4] For a couple of milliseconds after onset, there are simply no spikes left to evaluate. Note the peak at an interspike interval of about 0.5 ms, which agrees with Fig. 6. A rapid succession of spikes can evoke more spikes than the succession contains because the time constant of the relative refractory period is much smaller than the time constant of the EPSP. Accordingly, the peak at 0.5 ms shrinks for a longer time constant of the relative refractory period (data not shown). Interestingly, analogous simulations with a lasting pure tone show that the synchronization of the neuronal activity to the signal changes only slightly, the vector strength being 0.71 in the AN and 0.74 in the AVCN (data not shown). When the sound pressure level of the signal is reduced by 10 dB(A) the number of spikes remains constant in the case of a small inhibitory spread (compare dotted and solid black line in Fig. 7) but increases in the case of a large inhibitory spread (dotted and dashed gray line in Fig. 7). The model thus provides nonlinear gain control for frequency-specific inhibition.

We note that the equivalence of a broad weak and a narrow strong inhibition as we found it for click pairs breaks down for a narrow-band signal (as a pure tone) or a random signal (as white noise).

IV. DISCUSSION

In the previous two sections the analytical as well as the numerical approach have led to the same conclusions concerning the properties of the model presented. In both cases the second click of a click pair is suppressed and the suppression is maximal at about 2 ms after the first click (Figs. 2 and 6). As for the response to the onset of an ongoing stimulus, the spiking activity is minimal a couple of milliseconds after onset (Figs. 3 and 7). This behavior is consis-

tent with the psychophysical experiments cited in Sec. I. Moreover, it is very robust. The only free parameters are the strength of the inhibition and the firing threshold of the neurons. All other parameters arise naturally from matching both form and amplitude of the model's cell potentials to the measurements published by Wickesberg and Oertel,[2] Wickesberg,[13] and Wu and Oertel.[17] Hence our model is highly plausible from a biological point of view.

Whereas the model can reproduce the psychophysical results obtained for clicks using both a relatively strong, frequency-specific inhibition and a weaker inhibition spreading across many frequency channels, a strong, frequency-specific inhibition is needed in order to reconstruct psychophysical results for ongoing stimuli. This can be understood by the following argument: To suppress a single spike in one frequency channel, a certain amount of inhibition is needed. This inhibition can arise either from a strong frequency-specific inhibitory connection (small IS) or from an integration over many weak inhibitory inputs (large IS). In the case of a precisely timed broadband signal as a click both are possible, but as either broadband character (in case of a pure tone) or precise timing (in case of white noise) disappear, a large inhibitory spread will yield only nonspecific spike suppression.

As for the function of the frequency-specific inhibitory connection from the DCN to the AVCN, our results suggest three possibilities.

First, analytical (Fig. 3) and numerical (Figs. 4, 5, and 7) calculations show that a type of *gain control* is provided since the response is reduced significantly through inhibition. The model can even provide constant output for different input intensities in the case of an ongoing stimulus.

Second, the analytical calculations show that inhibition induces *contrast enhancement* since the dc component of the pure tone response is decreased, whereas the ac component (i.e., the amplitude of the oscillation) is unchanged (Fig. 3). This means that phase locking should be improved since the ac to dc ratio is larger with inhibition than without. An improvement in phase locking has been reported physiologically in the AVCN as compared to the AN.[19] In our numerical simulations this effect was only marginal. Presumably this is due to the fact that multiple AN fibers need to converge onto an AVCN cell in order to show enhanced synchronization, as Joris *et al.* demonstrated in their coincidence cell model.[19] Our model, however, only specifies a single AN fiber innervating a single AVCN cell.

Third, *monaural echo suppression* could be an additional function of the inhibitory circuitry we have modeled in the cochlear nucleus. A click quickly following another click will under natural circumstances most often be an echo. It is known that echoes occurring 2–20 ms after the initial signal are suppressed in the central nervous system.[1] A model for *binaural* echo suppression has been proposed previously in this journal[20] and has been extended in a dissertation available online.[21] Both versions of the model feature a persistent inhibition that suppresses echoes for up to 20 ms. A drawback of the extended version of the binaural model is a minimal suppression time of 5 ms, which means it is not capable of a suppression occurring in less than 5 ms. That is, our model with a suppression time of about 2–4 ms completes the picture.

Apart from the functions listed above, several publications hint towards yet an additional role delayed, frequency-specific inhibition might play. A neuronal circuitry very similar to the one discussed in the present paper has been found in the medial superior olive of the mustached bat.[22] There a delayed monaural inhibition is proposed as a kind of filter for amplitude-modulated (AM) sounds, important to detecting and identifying wing-beating prey. Amplitude modulations are known to play an important role in the processing of many acoustic signals. They are of special importance to transmitting information in speech[23] as well as for the perception of pitch.[24] It has been shown that identifying acoustic signal periodicity is well within the scope of neuronal hardware.[25]

An idea related to a delayed monaural frequency-specific inhibition was brought up by de Cheveigné in 1998[26] in the context of pitch perception. He showed that many pitch phenomena can be explained by a model involving an array of delay lines and inhibitory gating. Against this background, it is very interesting that in 1996 Wickesberg found a second, even faster mechanism of inhibition in the cochlear nucleus.[13] Referring to observations of Shore,[27] Wickesberg speculated that "the circuitry in the cochlear nucleus may be just the first stage in an ascending and descending[27] cascade of delayed, on-frequency inhibition." It is an important fact that the circuitry in the real cochlear nucleus does not only extend to the bushy cells that project to the superior olive complex and are involved in sound localization (where echo suppression is essential), but also to the stellate cells that project to the inferior colliculus. Stellate cells have much slower responses than bushy cells,[14] which means inhibition would follow excitation more closely. Inhibiting different types of cells could contribute to extending the range of the timing in a cascade of delayed inhibition, so we could think of a cascade of frequency-specific inhibitory loops filtering signals according to their respective AM frequency. Indeed, a phenomenological model of responses to AM tones uses the circuitry we have analyzed in the present paper as a module.[28]

Finally, it should be mentioned that the filtering of a signal obviously performed by the cochlear nucleus shows remarkable similarities with a filtering technology used in sound studios and extensively in radio as well as TV broadcasts.[29] The so-called *compressors* attenuate the signal after it reaches a certain threshold. That is, they reduce the dynamic range of the signal. If such a compressor is set, for example, to an "attack" of 2 ms and a "decay" of 10 ms, it will start to attenuate the signal 2 ms after the amplitude of the audio signal exceeds its threshold. This attenuation will gradually decay back to 0 within the next 10 ms, a behavior similar to that of our model. A compressor improves the perceived volume of an audio signal.

We conclude that the neural circuitry discovered by Wickesberg and Oertel[2] in the cochlear nucleus—in addition to providing gain control—can be regarded as the site of monaural echo suppression. Furthermore, it may be involved in the preprocessing of amplitude-modulated sounds.

ACKNOWLEDGMENTS

The authors thank P. Friedel, B. Grothe, H. Meffin, B. Krebs, and F. L. Occhionero for fruitful discussions. Special thanks go to V. Dasika for many helpful comments on earlier versions of the manuscript. M. B. gratefully acknowledges financial support from the Bernstein Center for Computational Neuroscience—Munich.

[1] J. Blauert, *Spatial Hearing* (MIT, Cambridge, MA, 1999).
[2] R. E. Wickesberg and D. Oertel, "Delayed, frequency-specific inhibition in the cochlear nuclei of mice: A mechanism for monaural echo suppression," J. Neurosci. **10**, 1762–1768 (1990).
[3] G. G. Harris, J. L. Flanagan, and B. J. Watson, "Binaural interaction of a click with a click pair," J. Acoust. Soc. Am. **35**, 672–678 (1963).
[4] P. M. Zurek, "The precedence effect and its possible role in the avoidance of interaural ambiguities," J. Acoust. Soc. Am. **67**, 952–964 (1980).
[5] E. R. Hafter and R. H. Dye, "Detection of interaural differences of time in trains of high frequency clicks as a function of interclick interval and number," J. Acoust. Soc. Am. **73**, 644–651 (1983).
[6] E. R. Hafter and E. M. Wenzel, "Lateralization of transients presented at high rates: site of the saturation effect," in *Hearing—Physiological Basis and Psychophysics*, edited by R. Klinke and R. Hartman (Springer, Berlin, 1983), pp. 208–220.
[7] E. R. Hafter, T. N. Buell, and V. M. Richards, "Onset coding in lateralization: Its form, site and function," in *Auditory function*, edited by G. Edelman (Wiley, New York, 1988), pp. 647–674.
[8] R. E. Wickesberg and D. Oertel, "Tonotopic projection from the dorsal to the anteroventral cochlear nucleus of mice," J. Comp. Neurol. **268**, 389–399 (1988).
[9] G. A. Spirou, K. A. Davis, I. Nelken, and E. D. Young, "Spectral Integration by Type II Interneurons in Dorsal Cochlear Nucleus," J. Neurophysiol. **82**, 648–663 (1999).
[10] G. A. Spirou and E. D. Young, "Organization of dorsal cochlear nucleus type IV unit response maps and their relationship to activation by band-limited noise," J. Neurophysiol. **66**, 1750–1768 (1991).
[11] R. Kempter, W. Gerstner, and J. L. van Hemmen, "Hebbian learning and spiking neurons," Phys. Rev. E **59**, 4498–4514 (1999).
[12] J. L. van Hemmen, "Theory of synaptic plasticity," in *Handbook of Biological Physics (Vol. 4), Neuro-informatics, Neural Modelling*, edited by F. Moss and S. Gielen (Elsevier, Amsterdam, 2001), pp. 771–823.
[13] R. E. Wickesberg, "Rapid inhibition in the cochlear nuclear complex of the chinchilla," J. Acoust. Soc. Am. **100**, 1691–1702 (1996).
[14] J. S. Isaacson and B. Walmsley, "Receptors underlying excitatory synaptic transmission in slices of the rat anteroventral cochlear nucleus," J. Neurophysiol. **73**, 964–973 (1995).
[15] A. J. Smith, S. Owens, and I. D. Forsythe, "Characterisation of inhibitory and excitatory postsynaptic currents of the rat medial superior olive," J. Physiol. (London) **529**, 681–698 (2000).
[16] M. Barnes-Davies and I. Forsythe, "Pre- and postsynaptic glutamate receptors at a giant excitatory synapse in rat auditory brainstem slices," J. Physiol. (London) **488**, 387–406 (1995).
[17] S. H. Wu and D. Oertel, "Intracellular Injection with Horseradish Peroxidase of Physiologically Characterized Stellate and Bushy Cells in Slices of Mouse Anteroventral Cochlear Nucleus," J. Neurosci. **4**, 1577–1588 (1984).
[18] L. P. O'Mard, M. J. Hewitt, and R. Meddis, "Lutear, v2.0.9," C++ core routines library, University of Essex, Hearing Research Laboratory (1997).
[19] P. X. Joris, L. H. Carney, P. H. Smith, and T. C. T. Yin, "Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. responses to tones at the characteristic frequency," J. Neurophysiol. **71**, 1022–1036 (1994).
[20] H. Cai, L. H. Carey, and S. H. Colburn, "A model for binaural response properties of inferior colliculus neurons. I. A model with interaural time difference-sensitive excitatory and inhibitory inputs," J. Acoust. Soc. Am. **103**, 475–493 (1998).
[21] T. P. Zahn, "Neural architecture for echo suppression during sound source localization based on spiking neural cell models," Ph.D. thesis, Technische Universität Ilmenau, 2003.
[22] B. Grothe, "Interaction of excitation and inhibition in processing of pure tone and amplitude-modulated stimuli in the medial superior olive of the mustached bat," J. Neurophysiol. **71**, 706–721 (1994).
[23] R. Shannon, F. G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," Science **270**, 303–304 (1995).
[24] D. Bendor and X. Wang, "The neuronal representation of pitch in primate auditory cortex," Nature (London) **436**, 1161–1165 (2005).
[25] P. Friedel, M. Bürck, and J. L. van Hemmen, "Neuronal identification of acoustic signal periodicity," Biol. Cybern. **97**(3), 247–260 (2007).
[26] A. de Cheveigné, "Cancellation model of pitch perception," J. Acoust. Soc. Am. **103**, 1261–1271 (1998).
[27] S. E. Shore, "Recovery of forward-masked responses in ventral cochlear nucleus neurons," Hear. Res. **82**, 31–43 (1995).
[28] P. C. Nelson and L. H. Carney, "A phenomenological model of peripheral and central neural responses to amplitude-modulated tones," J. Acoust. Soc. Am. **116**, 2173–2186 (2004).
[29] M. Dickreiter, *Handbuch der Tonstudiotechnik* [A Handbook of Sound Studio Technology] (Saur, Munich, 1997).
[30] M. J. Hewitt and R. Meddis, "An evaluation of eight computer models of mammalian inner hair-cell function," J. Acoust. Soc. Am. **90**, 904–907 (1991).