

Paul Friedel · Moritz Bürck · J. Leo van Hemmen

Neuronal identification of acoustic signal periodicity

Published online: 24 August 2007

Abstract Acoustic signals transmit information by temporal characteristics and envelope periodicity as well as by their frequency content. Many animals can extract the frequency content of a signal by means of specialized organs such as the cochlea but for the detection and identification of higher-order periodicity, e.g., amplitude modulations, this type of organ is useless. In addition, many animals do not have a cochlea but still depend on a reliable identification of different frequencies in the vast variety of acoustic signals they perceive in their natural environment. Hence, neural mechanisms to decode periodicity information must exist. We present a detailed mathematical analysis of a recurrent and a feedforward model of neuronal periodicity extraction and discuss basic constraints for neuronal circuitry performing such a task in a biological system. Both the recurrent and the feedforward model perform well using neuronal parameters typical for the auditory system. Performance is limited mainly by the temporal precision of the connections between the neurons.

Keywords periodicity detection · Auditory signal processing · Neuronal modeling

1 Introduction

Sound and vibration play a very important role in communication throughout the natural world. Vertebrates possess a highly specialized vibration detector, the cochlea, which is used to detect and to decompose signals into their constituent frequencies.

Many naturally occurring vibrations carry information about their source in the form of frequency content *and* temporal structure [45]. To be able to effectively use this information, it is important that a perceiving animal has the ability to decode both spectral and temporal cues present in the signal. Obviously animals without a mechanosensory

frequency decomposer, viz, a cochlea, at their disposal need to extract this information by neuronal means. But even animals *with* a cochlea can only extract *spectral* information using this organ — *temporal* information, slow variations of amplitude over time, still needs to be decoded neuronally. Several examples of neural temporal information extraction are known. In the human auditory system, these effects include speech recognition [62, 65], the identification of acoustic events (the “cocktail party effect” [21]), subjective pitch perception [37, 6], and the “missing fundamental” effect [6, 63]. A spectacular example of temporal information extraction is the nearly perfect recognition of speech under conditions of greatly reduced spectral information. With only three bands of noise modulated with the temporal envelopes of speech, the recognition rate of sentences is still above 80% [62]. It has often been suggested that detection of signal periodicity is essential in the recognition of auditory objects [70, 13].

Within the animal kingdom, echo-locating bats provide an interesting example of temporal information extraction. Several species of bat discriminate different insect species by their characteristic wing beat frequency which leads to a species-specific time-varying Doppler shift in the echo [61]. It has been shown that the bullfrog can extract the periodicity of complex stimuli using temporal cues [53].

Surface feeding fish can detect the frequency of vibration of the water surface with an accuracy of about 10% [11] and they use the dispersion of surface waves to determine prey distance [39]. However, the vibration-sensitive *cupulae* located on the skin of the animal [23] have no frequency specificity apart from low-pass filtering the signal [38]. The clawed frog, too, is able to detect small frequency differences in water surface wave trains [29]. All these examples clearly show that a neural mechanism to detect temporal information is a common feature shared by many animals.

The above statement is supported by physiological findings. In spiders, frequency-specific neurons have been found in the central nervous system [66], although the vibration-detecting *slit sensilla*, located in the joints of their feet [4], do not exhibit frequency tuning. The sensilla rather function as acceleration detectors with a frequency-independent

threshold sensitivity [5]. In the mammalian auditory system, neurons sensitive to specific modulation frequencies of auditory signals have been identified. There is a very strong sensitivity to temporal modulation of input in the Inferior Colliculus (IC) [37]. In the IC of the cat a topographical arrangement of amplitude modulation-sensitive neurons has been demonstrated [60].

Below, we will mathematically describe two fundamental neuronal architectures for detecting signal periodicity. We have restricted ourselves to a minimalistic implementation of the models and do not take into account specific physiological details. There are two reasons for doing so. First, discussing simple models allows a detailed mathematical treatment leading to comprehension of the abilities and limitations of the circuitry. Second, since periodicity detection is a capability present in many animals, it is important to understand general mechanisms rather than specific realizations in a certain group of animals.

The organization of this paper is as follows. Section 2 gives a short introduction to the characteristics of vibratory signals. We introduce two different models of neuronal periodicity detection in section 3, and study their mathematical properties in section 4. We consider numerical simulations in section 5, and in section 6 we discuss our results.

2 Characterization of vibratory signals

To arrive at a clear understanding of the models discussed in this paper, it is necessary to briefly describe some properties of vibratory signals. We have to bear in mind that vibratory signals are not limited to air-borne sound but may propagate in a variety of substrates such as sand [14, 1], the water surface [10, 8], spider webs [50, 43], or leaves [48]. All vibratory signals consist of a time-dependent change in pressure or medium deflection. Fast periodical variations in signal strength are normally designated as *spectral* content, or frequencies, and slow variations are denoted as *temporal* content. In general the distinction between temporal and spectral content is a matter of convention. In our setting, we consider all periodic signal fluctuations which can be resolved *neuronally* as temporal content. That is, signal variations with frequencies lower than approximately 500 Hz are temporal.

Natural signals are composed of a mixture of spectral and temporal components, leading to complex wave forms (Fig. 1). The slowly varying amplitude of the signal is called the signal *envelope*. Even if no explicit modulation is imposed on a signal, slow modulations of the signal strength generally occur. This is because of interference effects between the frequency components present in the signal, which lead to a “beating” effect. In vertebrates, the cochlea is responsible for extracting spectral information; temporal clues need to be decoded neuronally. In animals lacking a cochlea or similar structure, both spectral and temporal information must be extracted neuronally.

The simplest example of a nontrivial vibration is sinusoidal amplitude modulation (SAM). An SAM signal $s(t)$ is

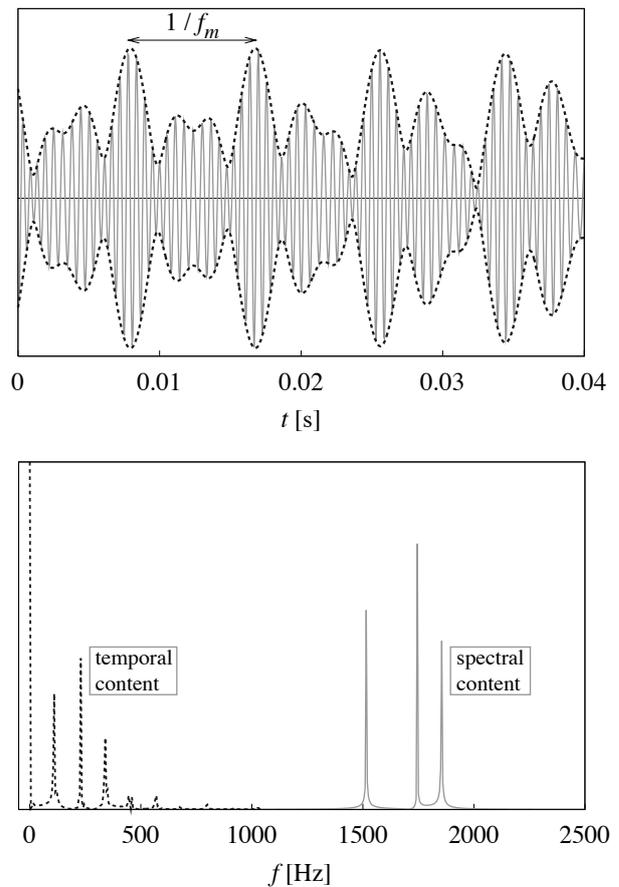


Fig. 1 The top panel shows a complex wave signal, composed of three frequency components (solid gray) and the signal envelope, or instantaneous amplitude (dotted black). The interference of the frequency components causes amplitude modulation on a slow scale. One of the modulation frequencies ($f_m \sim 100$ Hz) has been indicated in the plot. The bottom panel shows the Fourier transform of the signal (solid gray) and the envelope (dotted black). Although the signal consists of three frequencies in the 1500 – 2000 Hz range, the envelope shows only slow variations, mainly below 500 Hz. Because the envelope has a nonzero mean value the Fourier spectrum shows an additional peak at 0 Hz.

described by

$$s(t) = A \cos(2\pi f_c t) [1 + m \sin(2\pi f_m t)], \quad (1)$$

with $f_m \ll f_c$. The amplitude of the signal is A , f_c is the carrier frequency, f_m is the modulation frequency and m is the modulation depth of the signal. Of course, an SAM signal is not a realistic input. Natural signals typically consist of some superposition of frequencies that are comodulated [55] by a *characteristic profile* instead of a simple sine wave. The fundamental period of this profile is then called the modulation frequency.

On the surface of the water [3, 7, 9, 44], in spider webs [49], and on plant leaves [3, 33], discrimination of prey is achieved by using only the frequency content of the incoming signals. In both cases prey-generated signals contain high frequencies ($\gtrsim 50$ Hz). In contrast, background noise such

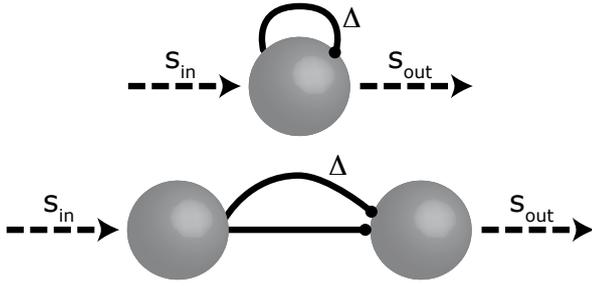


Fig. 2 There are basically two possible ways to extract frequency or timing information from a signal using spiking neurons. The first method (upper panel) uses a recurrent loop with time delay Δ . This we call the *recurrent model*. The neuron is driven by a continuous input function s_{in} . If the neuron emits a spike at time $t = t_0$, the firing probability is enhanced at time $t = t_0 + \Delta$. Signal periodicity with characteristic time Δ then leads to a higher number of spikes in the output signal s_{out} .

The second method (lower panel) is based on the same idea, but uses a feedforward network, and is called the *feedforward model*. The first neuron, again driven by s_{in} , sends two spikes to the output neuron with a delay differing by an amount Δ , e.g. using interneurons. Again, correlations in the input signal with period Δ lead to an augmented firing probability for the output neuron.

as abiotic signals and vibrations caused by the movement of the animal itself is limited to low frequencies ($\lesssim 15$ Hz). To detect prey, the animal needs to know whether the signal has predominantly high-frequency content or low-frequency content. Amplitude modulations are believed to be of less importance in this case.

3 The models

The goal of our models will be to identify slow fluctuations present in a specific input signal. Mathematically, periodic features of a signal $s(t)$ can be detected by calculating its autocorrelation χ (see e.g. [51]), defined by

$$\chi(\Delta) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T d\tau s(\tau)s(\Delta + \tau). \quad (2)$$

The autocorrelation has maxima for correlation times Δ corresponding to the frequencies present in the signal, but also for the periods of the envelope fluctuations (Fig. 1). The above calculation immediately suggests two neuronal mechanisms for detecting periodicity (Fig. 2).

The first model consists of a neuron that receives an input signal $s_{in}(t)$. As the neuron spikes, the output spike is fed into a pathway that ultimately projects onto the neuron itself with a particular delay Δ , corresponding to the correlation time above. This pathway need not be a direct connection from the axon onto the neuron's own dendritic tree. One or more processing steps may occur before the output from the neuron returns but a well-defined delay can be associated with the pathway. This matter will be further discussed in section 6.1. Because of the delay loop, the neuron detects correlations on a time scale Δ . An array of such neurons, all

with different Δ , can then function as a periodicity analyzer. We call this model the *recurrent model*.

The second model consists of a two-neuron network. If the input neuron fires, its spikes are fed into two pathways to the output neuron. The temporal durations of these pathways differ by an amount Δ . The output neuron will have a high firing probability if spikes arrive from the two different pathways at the same time. Again, the network reveals correlations on time scale Δ . We call this model the *feedforward model*.

Both types of networks have been discussed before in the literature. To our knowledge, the first author to propose a network of delay lines to detect signal periodicity has been Licklider [47]. More recent work on feedforward-like models has been done by Borst et al. [12] and Meddis and O'Mard [52]. Both articles proposed a very detailed model, based on specific properties of neuronal circuitry found in the mammalian auditory system. Although Borst et al. and Meddis and O'mard have used neuronal oscillators instead of delay lines, our analysis can be applied to their models as well. Cariani [19,20] discussed a recurrent-like model. He has used an overly simple model in which formal neurons manipulating strings of 0s and 1s are used. None of these authors have included a detailed mathematical analysis of their models. In this paper, we provide this missing analysis and use fairly realistic neuron models for our simulations without settling on a specific neuronal architecture. We will now analyse the characteristics of our models in more detail.

3.1 Detailed description of the recurrent model

The recurrent model consists of N_{out} output neurons that all receive the same external continuous input $s_{in}(t)$. All input neurons have a recurrent connection that feeds output spikes back into the neuron itself. The recurrent spikes are characterized by a delay Δ that is different for each neuron and has strength J . The feedback current is described by a general function g (see also section 4.1) for which we will take an α -function in our simulations [31]

$$g(t) = \theta(t - t_0 - \Delta) \frac{t - t_0 - \Delta}{\tau^2} e^{-(t - t_0 - \Delta)/\tau}. \quad (3)$$

The width of the α -function is given by τ , t_0 is the spiking time of the neuron and θ denotes the Heaviside step function, i.e., $\theta(t) = 0$ for $t < 0$ and $\theta(t) = 1$ for $t \geq 0$.

The neurons are simulated as leaky integrate-and-fire (LIF) neurons; see appendix A.2. To get the model at work the output neurons must fire a first spike to start with, since the feedback loop needs input, which can only come from the neurons themselves. It is not possible to use supra-threshold input since this would imply that *all* output neurons would fire in response to the input, regardless the length of their delay loop. The solution is to use subthreshold input with added internal neuronal noise. Every now and then the neuron will fire. But only if the delay loop length has the right value the neuron will be able to resonate in response to the

input. The mechanism described here is called *stochastic resonance* for which there exists a detailed review [30].

3.2 Detailed description of the feedforward model

The feedforward model consists of N_{in} input neurons, which we simulate as Poisson neurons; see appendix A.1. This is convenient but by no means necessary. The input neurons are driven by an external input $s_{\text{in}}(t)$. If one of the input neurons fires its spike is fed into an axon branching off to N_{out} different output neurons. One spike reaches the output neurons directly and another spike resulting from the same event reaches the output neuron with a delay Δ . A specific delay Δ is associated with every output neuron. In this way, every output neuron will turn out to encode a particular frequency $f = 1/\Delta$.

The output neurons are simulated as leaky integrate-and-fire (LIF) neurons without noise; see Appendix A.2. If a spike is emitted at time $t = t_0$ by any of the input neurons it leads to two postsynaptic current injections arriving at the output neurons, again in the form of α -functions,

$$\epsilon_{\text{direct}} = J\theta(t - t_0) \frac{t - t_0}{\tau^2} e^{-(t - t_0)/\tau} \quad (4)$$

and

$$\epsilon_{\text{delayed}} = J\theta(t - t_0 - \Delta) \frac{t - t_0 - \Delta}{\tau^2} e^{-(t - t_0 - \Delta)/\tau}. \quad (5)$$

The former spike travels to the output neuron without delay and the latter arrives with a delay Δ . The synaptic coupling strength is again given by the parameter J .

4 Mathematical discussion of the models

In this section we will mathematically discuss the behavior of the two types of periodicity detector. Explicit analysis of LIF neurons is generally quite difficult (for an extensive review, see [16,17]). We will see, however, that no explicit analysis of LIF neuron dynamics is needed to gain valuable insight into the dynamics of our models. In fact, the key properties of the models are independent of the specific type of neurons that are used.

4.1 Recurrent model

The problem of analytic calculations using integrate-and-fire neurons lies in the nonlinearity of the spike-generation. In the case of the recurrent network we discuss here, the problem is even more difficult than usual since the feedback introduces an extra complication into the system. We therefore simplify our discussion by considering Poisson neurons; see appendix A.1. We will later compare the findings obtained here with the simulations (section 5) to see whether the calculations using Poisson neurons can serve to understand the dynamics of the LIF neurons used here.

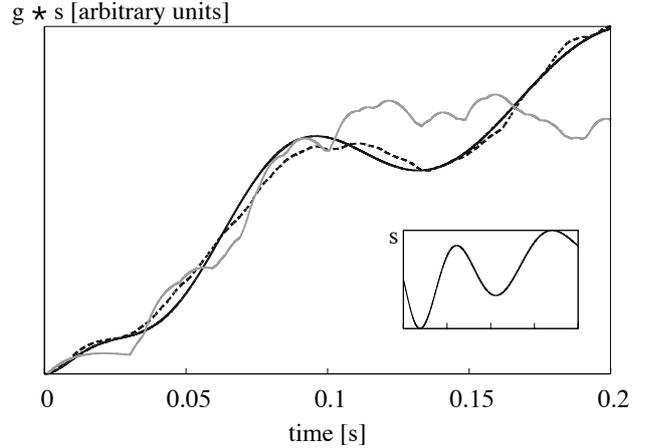


Fig. 3 Stochastic fluctuations in the response of a Poisson neuron are smaller if the firing rate is higher. The convolution of a signal s (inset) with the response kernel g in black is compared with two explicit realizations of the firing process (normalized for comparison). The gray curve was obtained using about 40 spikes, the black dotted curve results from about 300 spikes. Clearly, a high firing rate (or, mathematically equivalent, a large population of Poisson neurons) is needed for Eq. (6) to apply.

We can describe the rate function λ of a single Poisson neuron projecting back to itself with a particular delay time Δ by the integral equation

$$\begin{aligned} \lambda(t) &= s_{\text{in}}(t) + J \int_{-\infty}^{\infty} ds g(s, \Delta) \lambda(t - s) \\ &= s_{\text{in}}(t) + J(g \star \lambda)(t). \end{aligned} \quad (6)$$

The rate function consists of the sum of the external input s_{in} and the delayed input from the recurrent loop, “smeared out” by the kernel g . The feedback strength is given by J and we choose g to ensure causality ($g(t) = 0$ if $t < 0$) and to have unit weight

$$\int_{-\infty}^{\infty} dt g(t) = 1. \quad (7)$$

Since g has a finite width the delay loop smears out the feedback.

In (6) the convolution integral of λ with the kernel g assumes that we may use the expectation value of the firing rate λ to describe the neuron output instead of a specific realization of the output. We thereby ignore the “spiky” character of the neuron output. This approach is only correct for very high firing rates or, mathematically equivalent, a large number of Poisson neurons with a low firing rate. The total amount of output spikes must be high enough so that the output signal is reliably sampled by the output spikes. An example of such a smooth convolution of λ and g is shown in figure 3.

To solve Eq. (6) for the output firing rate λ we take the Fourier transform of the equation. The Fourier transform of a function h is defined by

$$H(\omega) = \mathcal{F}[h(t)](\omega) := \int_{-\infty}^{\infty} dt e^{-i\omega t} h(t) \quad (8)$$

and has the useful property that, when transformed, a convolution becomes an ordinary product. Denoting the Fourier transform of each input term by a capital letter we obtain

$$\Lambda(\omega) = S_{\text{in}}(\omega) + JG(\omega)\Lambda(\omega). \quad (9)$$

The solution is then given by

$$\Lambda = \frac{S_{\text{in}}}{1 - JG}. \quad (10)$$

The solution as a function of time can then be found by taking the inverse Fourier transform

$$\lambda(t) = \mathcal{F}^{-1}[\Lambda(\omega)](t) := \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega t} \Lambda(\omega). \quad (11)$$

Given any input function s_{in} and response function $g(t)$ we can now explicitly calculate the firing probability of the neuron. In our simulations we will use an α -function for g to model the response function [see Eq. (3)]. The Fourier transform of this response function is given by

$$G(\omega) = \frac{e^{-i\omega\Delta}}{(1 + i\omega\tau)^2}. \quad (12)$$

Since we are interested in identifying periodicity, we must know which frequency f corresponds to a certain delay time Δ . A first guess would be to set $f = 1/\Delta$, but since the response function transforms the recurrent signal this relation cannot be expected to hold exactly. We therefore consider the response of the system to an incoming pure sine wave of frequency f and find the corresponding Δ that maximizes the amplitude of the response. We then have an explicit connection between the delay Δ and the signal frequency that is decoded optimally through this delay.

For harmonic input given by

$$s_{\text{in}}(t) = A \cos(\omega t) = A \cos(2\pi f t) \quad (13)$$

we calculate the response to be

$$\lambda(t) = L \cos(\omega t + \phi), \quad (14)$$

where ϕ is a phase that is not relevant for our further calculations and L is an amplitude given by

$$L = \frac{2(1 + \xi^2)^2}{\sqrt{J^2 + (1 + \xi^2)^2 - J[2(1 - \xi^2)\cos(\omega\Delta) - 4\xi\sin(\omega\Delta)]}}, \quad (15)$$

with the definition $\xi := \omega\tau$. The amplitude L of the response is maximal if the relation

$$2(1 - \xi^2)\cos(\omega\Delta) - 4\xi\sin(\omega\Delta) = 0 \quad (16)$$

holds. The delay must therefore satisfy

$$\Delta = \omega^{-1} \left[\arctan\left(\frac{2\xi}{\xi^2 - 1}\right) + n\pi \right], \quad (17)$$

with $n = 1$ if $\xi > 1$ and $n = 2$ for $\xi < 1$. If the width of the kernel g approaches zero ($\xi \rightarrow 0$) this relation indeed reduces to

$$\Delta = \frac{2\pi}{\omega} = \frac{1}{f}. \quad (18)$$

As a more complicated and realistic example let us consider an input of the form

$$s_{\text{in}}(t) = \int_0^{\infty} d\sigma B(\sigma) \cos[\sigma t + \phi(\sigma)]. \quad (19)$$

Instead of a single harmonic component we now describe the input by a distribution of input frequencies with arbitrary amplitude and phase. The Fourier transform of the input is given by

$$S_{\text{in}}(\omega) = \int_0^{\infty} d\sigma B(\sigma) \pi e^{i\phi(\sigma)\omega/\sigma} \times [\delta(\sigma - \omega) + \delta(\sigma + \omega)], \quad (20)$$

with $\delta(\cdot)$ the Dirac delta function. Plugging this result into (10) and (11) gives the solution for the firing rate of the output neuron

$$\lambda(t) = \int_0^{\infty} d\sigma B(\sigma) \mathcal{R} \left[\frac{e^{i(\phi(\sigma) + \sigma t)}}{1 - J e^{-i\sigma\Delta} / (1 + i\sigma\tau)^2} \right], \quad (21)$$

where $\mathcal{R}[x]$ denotes the real part of x . We note that the signal function (19) need not be positive, although a negative firing rate certainly does not make sense for a Poisson neuron. In the simulations we always use half-wave rectified signals. Unfortunately exact calculations are not feasible in this case. In spite of this drawback Eq. (21) captures the essence of the network response. If the solution (21) is plotted for various input spectra $B(\sigma)$ the amplitude of λ is largest if the length of the delay loop Δ corresponds to a frequency that is present in the input signal (plot not shown).

4.2 Feedforward model

For the analytic description of the feedforward model we will use an input population of Poisson neurons which are, just as before, driven by an input $s_{\text{in}}(t)$ identical for each neuron. We consider LIF neurons as output neurons; every output neuron receives input from the Poisson neurons via two distinct pathways: a direct connection and a connection with a delay Δ which is different for each output neuron.

If we calculate the expectation value of the current that arrives at the output neurons a sinusoidal function results. The response of LIF neurons to harmonic input is difficult to calculate but several exact results have been presented by Burkitt [18] as will be discussed below.

We start our calculations by considering input given by

$$s_{\text{in}}(t) = \frac{A}{2} [1 + \cos(\omega t)]. \quad (22)$$

The input current that one output neuron with a particular delay time Δ receives from the set of N_{in} input neurons is given by [referring to (4) and (5)]

$$\varepsilon_{\text{total}} = \varepsilon_{\text{direct}} + \varepsilon_{\text{delayed}} . \quad (23)$$

The expectation value of the current to the output neurons is given by (see appendix A.1)

$$\langle I \rangle = \int_{-\infty}^{\infty} ds s_{\text{in}}(s) \varepsilon_{\text{total}}(t-s) \quad (24)$$

and the variance of the current is given by

$$\text{var}_I = \int_{-\infty}^{\infty} ds s_{\text{in}}(s) \varepsilon_{\text{total}}^2(t-s) . \quad (25)$$

Equations (24) and (25) can be evaluated exactly for the given input function (22). The current is

$$\langle I \rangle = N_{\text{in}}AJ \left\{ 1 + \frac{\cos(\omega\Delta/2)}{(1+\xi^2)^2} \left[(1-\xi^2) \cos(\omega(t-\Delta/2)) + 2\xi \sin(\omega(t-\Delta/2)) \right] \right\} \quad (26)$$

with $\xi = \omega\tau$. The amplitude (current arriving at the output neuron) is thus maximal for integer

$$\Delta \cdot \frac{\omega}{2\pi} \in \mathbb{N} . \quad (27)$$

That is, a maximal response of the output neurons is to be expected if the input frequency matches the delay of the system. If the input signal contains a periodicity with frequency \bar{f} the neuron with a delay time $\bar{\Delta}$ corresponding to this frequency will respond optimally. All neurons sensitive to a *subharmonic* frequency (\bar{f}/n , with $n \in \mathbb{N}$) will also respond, as can be seen from (27). This is because an input signal with a periodicity \bar{f} is automatically also periodic with frequency \bar{f}/n .

The variance of the current is given by

$$\text{var}_I = \frac{N_{\text{in}}AJ^2}{\tau} \left\{ \frac{1}{4} + \frac{2\cos(\omega\Delta/2)}{(4+\xi^2)^3} \times \left[(8-6\xi^2) \cos(\omega(t-\Delta/2)) + \xi(12-\xi^2) \sin(\omega(t-\Delta/2)) \right] + M \right\} \quad (28)$$

where M is given by

$$M = e^{-\Delta/\tau} \left\{ \frac{1-\Delta/\tau}{4} + \frac{1}{(4+\xi^2)^3} \times \left[\left(16-8\xi^2 + \frac{\Delta}{\tau(16-\xi^4)} \right) \cos(\omega(t-\Delta)) + 2\xi \left(12-\xi^2 + \frac{2\Delta}{\tau(4+\xi^2)} \right) \sin(\omega(t-\Delta)) \right] \right\} . \quad (29)$$

In order to allow correct periodicity detection, the time scale of the periodicity must clearly exceed the time scale τ of the individual current response functions ε . We thus expect the system to work best if the relation $\Delta \gg \tau$ holds, meaning that the time scale of the periodicity is much larger than that of the post-synaptic response. In the auditory system we can expect this condition to hold. M can then be neglected because of the exponential prefactor $e^{-\Delta/\tau}$ in (29). If low-frequency input is presented, we have $\omega \ll 1/\tau$ and thus $\xi = \omega\tau \rightarrow 0$. The current and its variance are then given by

$$\langle I \rangle = N_{\text{in}}AJ [1 + \cos(\omega\Delta/2) \cos(\omega(t-\Delta/2))] \quad (30)$$

and

$$\text{var}_I = \frac{4N_{\text{in}}AJ^2}{\tau} [1/16 + \cos(\omega\Delta/2) \cos(\omega(t-\Delta/2))] . \quad (31)$$

The relative variation of the current is proportional to

$$\frac{\delta I}{I} = \frac{\sqrt{\text{var}_I}}{I} \propto \frac{1}{\sqrt{N_{\text{in}}A\tau}} , \quad (32)$$

which also holds if we do *not* assume $\Delta \gg \tau$ and $\xi \rightarrow 0$. As expected, the current is less sensitive to random fluctuations if the number of input neurons or the input amplitude increases. The fact that the current fluctuates more if τ gets smaller is because a very short synaptic time scale tends to enhance the ‘‘spiky’’ character of the current. The system, however, does not become less reliable since a short post-synaptic current enables better coincidence detection by the output neurons [40,41].

The expression (26) for the mean current, which is a good approximation if there are enough input neurons, shows that all output neurons receive a harmonic (sinusoidal) current. The amplitude of the current is largest if the delay matches the periodicity of the input signal. The response of integrate-and-fire neurons to harmonic input is difficult to calculate but it has been done for a slightly different system in [18]. The results show that the periodicity of the input current is retained in the firing of the output neuron. This means that the output signal is phase locked to the current. The vector strength V , which is defined as the absolute value of the first Fourier coefficient of the signal divided by the zeroth Fourier coefficient, measures the amount of synchronization or phase locking. For perfect phase locking $V = 1$. For a random distribution of phases (complete absence of phase locking) we find $V = 0$. V tends to be larger in the output neuron than in the current itself. Due to the phase locking the modulation present in the input signal s_{in} tends to be enhanced by the system in accordance with physiological findings in the mammalian auditory pathway [37].

First summary. Stepping back for an overview we would like to quickly summarize what we have obtained so far. We have derived analytical expressions for the network response to simple periodic input signals [equations (21) and (26)]. As expected the response of the output neurons is maximal if the periodicity of the input signal matches the time delay Δ . The relationship between delay time Δ and the optimal decoding frequency f is given by a simple inverse function in the feedforward model (27), and by a more complicated expression in the recurrent model (17). We now turn to numerical simulations to characterize the response of the models to more realistic input.

5 Simulation results

In this section we discuss results obtained by numerical simulations. The neural networks as described in section 3 have been implemented through the C++ programming language. To test the performance of the models we have provided the networks with three different kinds of input: amplitude-modulated (AM) input, a Gaussian distribution of frequency components, and input mimicking the “missing fundamental” effect, as explained below. The response of the system was characterized by counting the number of output spikes that occurred during one second of input presentation as a function of the coding frequency of the output neuron. The coding frequency of the output neurons was calculated using (17) for the recurrent network and (27) for the feedforward network. We will see that in both networks the neurons encoding the periodicity present in the input signal respond maximally. The networks are thus able to convert a *periodicity code* into a *rate code*.

Half-wave rectification of the signals was always performed before presenting them to the network. Hair cells, the basic receptor units of the ear and the lateral line system, depolarize following one direction of displacement and hyperpolarize if displacement is in the other direction [34]. Half-wave rectification is therefore automatically performed upon detection in many sensory systems.

The input signal to the network was normalized to deliver the same time-integrated input power in each case. Obviously, it is not realistic to expect external input to a vibration detection system to be normalized but several mechanisms of neuronal *gain adaptation* have been shown to exist; e.g., in the auditory pathway [64, 69, 36, 25]. Such mechanisms are thought to keep neuronal firing rates within an optimal range. In our case power normalization is needed to keep the output firing under control. If the input power is too low, the output neurons cannot fire at all. If, on the other hand, the input power is too high all neurons will fire at a high rate and the discriminative capacity of the system is lost.

The numerical values of the parameters used in the computations are given in appendix B.

5.1 Amplitude-modulated input

We consider two types of AM input signals. First, we present a modulated pure tone

$$s_{\text{in}}(t) = \frac{A}{2} [1 + \cos(2\pi f_m t + \phi)] \cos 2\pi f_c t \quad (33)$$

with modulation frequency $f_m = 50$ Hz or $f_m = 200$ Hz and random modulation phase ϕ . The carrier frequency is $f_c = 2000$ Hz.

In the second case we consider noise by composing a signal from 50 sinusoidal components with frequencies f_{rand}^n chosen from a uniform distribution on $[0, 1000$ Hz] and random phases ϕ_{rand}^n with uniform distribution on $[0, 2\pi]$ so as to obtain

$$s_{\text{noise}}(t) = \sum_{n=1}^{50} \cos(2\pi f_{\text{rand}}^n t + \phi_{\text{rand}}^n) . \quad (34)$$

We then modulate this signal with modulation frequency $f_m = 50$ Hz

$$s_{\text{in}}(t) = \frac{A}{2} [1 + \cos(2\pi f_m t + \phi)] \times s_{\text{noise}} . \quad (35)$$

In both cases the amplitude has been chosen in such a way that the rectified input signal is normalized appropriately.

The results of these simulations are displayed in Fig. 4. Obviously both network types succeed very well in detecting the periodicity of the input signal. Clear peaks in the response occur for the correct frequencies. The response peaks for the subharmonic frequencies are also distinctly recognizable. Although the response of the recurrent model is quite noisy, this drawback may be overcome easily by combining input from several close-by channels.

5.2 Gaussian frequency distribution

The second test for the feedforward and recurrent models consists of taking a *distribution* of frequencies as input. To mimic the real biological situation we have built an input signal from 30 different frequency components chosen randomly from a Gaussian probability distribution with a center frequency μ and a width σ . The components were added together with random phases and the resulting signal was half-wave rectified and presented to the network. This signal can be considered as a rough model for a strampling insect on the water surface or in a spider web. The results of these simulations are displayed in Fig. 5. It is pretty evident that both frequency profiles with $(\mu, \sigma) = (30$ Hz, 5 Hz) and $(\mu, \sigma) = (100$ Hz, 20 Hz) are correctly identified by the two networks.

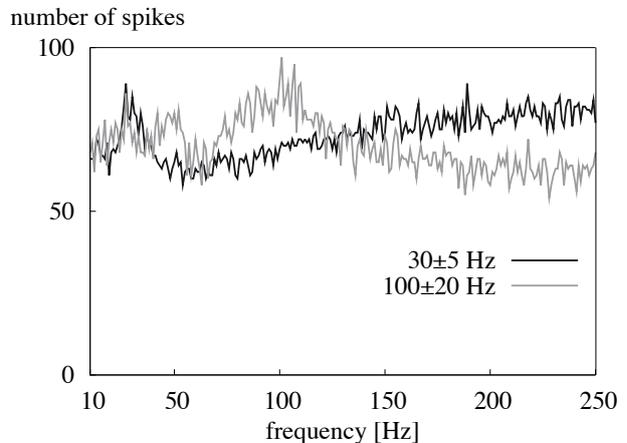
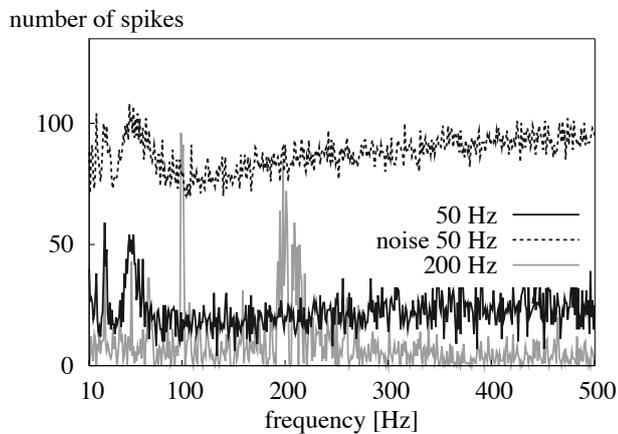
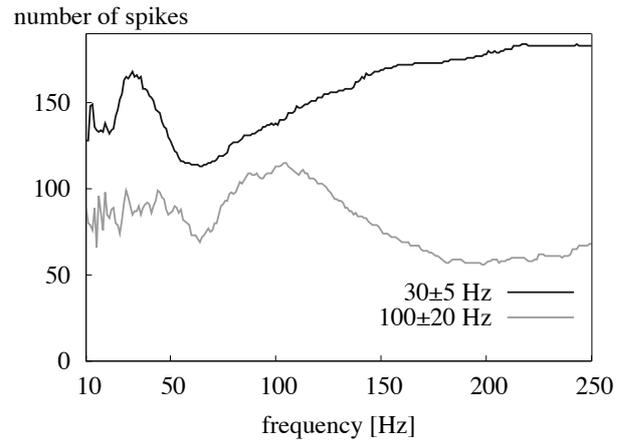
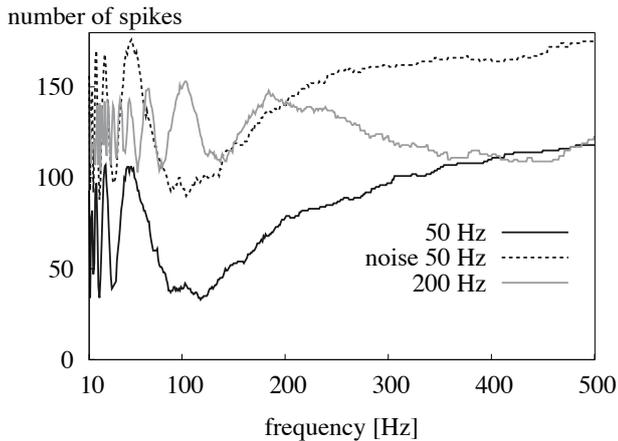


Fig. 4 Response to AM input of an array of neurons, each with a different delay and corresponding frequency (horizontal axis). The total number of spikes in one second is shown vertically. Top panel: feedforward model; bottom panel: recurrent model. The peaks corresponding to the input periodicity clearly appear in the graphs. Evidently, both networks correctly identify the signals.

Fig. 5 Response of an array of neurons to a distribution of frequencies. Input was presented to an array with neurons, each with a different delay and specific frequency (horizontal axis). The total number of spikes in one second is shown vertically. Top panel: feedforward model; bottom panel: recurrent model. Similar to figure 4 the signals are reliably identified.

5.3 Missing fundamental

If several pure tones with a *common* fundamental frequency are presented to a listener, the subject often perceives a tone with a pitch corresponding to this fundamental frequency, even though the fundamental frequency itself is not present in the input signal. Nonetheless a clear neuronal representation of this frequency is formed by the subject. To mimic such an experiment we give both models input consisting of three harmonics

$$s_{\text{in}}(t) = \sum_{n=1}^3 \cos(2\pi f_n t + \phi_n), \quad (36)$$

with $f_1 = 200$ Hz, $f_2 = 300$ Hz, $f_3 = 400$ Hz and the phases random. The response of the feedforward model is shown in Fig. 6, together with the response to a pure tone of 100 Hz. Although the peak is not as clear as with pure tone stimulation, a pitch of 100 Hz is still easily recognizable.

Because of the noisy response, the missing fundamental effect is not reproduced very well by the recurrent model. A

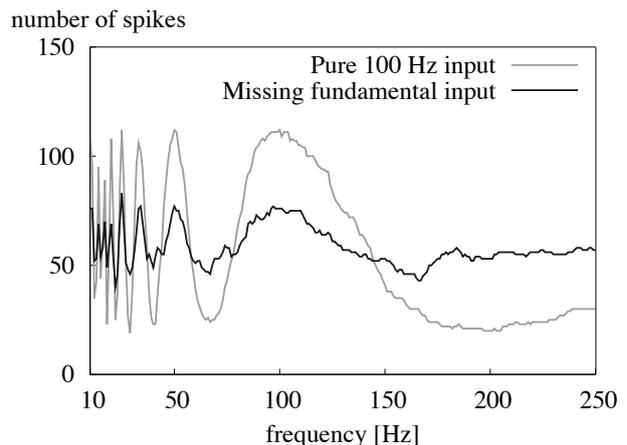


Fig. 6 Response of the feedforward network to “missing fundamental” input as in Eq. (36) with three frequencies 200, 300 and 400 Hz compared to the response to a pure 100 Hz tone. The peak at 100 Hz is clearly recognizable.

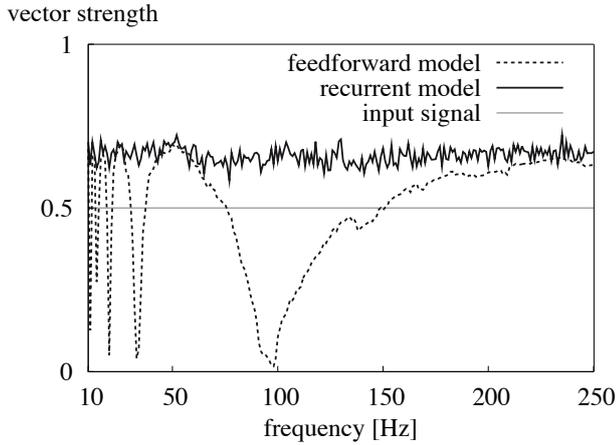


Fig. 7 Phase locking strength as a function of best frequency for the feedforward and the recurrent model. 50 Hz modulated input as in Eq. (35). Vector strength of the input signal is 0.5, as indicated by the horizontal line. Surprisingly, output phase locking is stronger than input phase locking in the relevant frequency range.

very good response can sometimes be obtained but this crucially depends on the precise values of the phases ϕ_n , which is not realistic biologically. Results for the recurrent network are therefore not shown.

5.4 Phase locking

A very important concept in auditory or vibratory processing is phase locking. Phase locking describes the capability of neurons to spike preferentially at a specific phase of the input signal. Phase locking is especially important to extract precise temporal clues from a signal; for instance, in sound localization [56, 32]. The amount of phase locking is characterized by the vector strength V as discussed above.

For AM noise input, as in (35), the vector strength has been displayed in Fig. 7. Interestingly, phase locking is quite good in the recurrent model although the output firing rate fluctuates a lot. This behavior results from the subthreshold input dynamics of the recurrent model. Only the presence of noise in the input assures that every now and then a spike occurs. The occurrence of a spike is of course much more likely if the input amplitude is large and consequently the output firing tends to be phase-locked to the input periodicity. For the feedforward model phase locking is good if the decoding frequency of the output neurons matches the periodicity of the input. Again, spike generation is most likely when the input amplitude is large *and* the delay time matches the frequency of the input signal. Phase locking results. Remarkably, phase locking of the output is significantly stronger than in the input signal for both models.

5.5 Temporal jitter of delays

The ability to identify signal periodicity crucially depends on the timing of the delays Δ . We therefore investigate the effect of temporal jitter in the delays on identification performance. We present four different pure tones with frequency f_{in} to both networks and add stochastic jitter to the delay time for every emitted spike. The jitter is Gaussian distributed with mean 0 and a standard deviation from 0.2 ms to 20 ms. For each trial (a specific combination of input frequency and jitter strength) we calculate the *selectivity* Q defined by

$$Q = \frac{|\sum_j r_j e^{2\pi i \Delta_j f_{in}}|}{\sum_j r_j}. \quad (37)$$

Here Δ_j is the temporal delay corresponding to output neuron j and r_j is its firing rate. This definition has again the form of a vector strength. If the output firing rate peaks for neurons with the correct delay ($\Delta_j f_{in}$ integer) the value of the numerator in Eq. (37) will be large. If much temporal jitter is present all output neurons will respond, even if their delay does *not* match the input signal frequency. In this case the phases in the numerator of Eq. (37) will cancel out and Q will have a low value.

In Fig. 8 the selectivity for different input frequencies and jitter magnitudes is plotted, normalized to the selectivity without jitter. As could be expected, the selectivity deteriorates if jitter is present in the delays. For high input frequencies the sensitivity to temporal jitter is largest. For low frequencies, say $\lesssim 25$ Hz, both models are quite robust and can cope with temporal jitter up to ~ 10 ms. A jitter of about 20% of the input periodicity leads to a 50% decrease in selectivity. The amount of jitter thus determines the fastest input periodicity that can still be identified. For a temporal jitter of 1 ms this upper limit is approximately 200 Hz.

6 Discussion

6.1 Performance limits

The capability to distinguish different frequencies hinges on the fact that the delay times Δ are well known and constant. Only then is it possible to reliably assign a particular frequency to the output neurons. In reality the time it takes for the signal to complete the delay line may vary somewhat.

As shown in section 5.5 the system works best for relatively low frequencies up to about 200 Hz. Even if the delay were much more accurate than has been suggested here, it would not be possible to detect very high frequencies reliably. The width of the post-synaptic current response presents a fundamental limit to the delay time that can be detected. In the experimental literature it has been found that AM sensitivity reaches a frequency as high as 1000 Hz but the vast majority of neurons is sensitive to modulation frequencies in the range of 10 – 300 Hz, most of them lying in the even more restricted range of 30 – 100 Hz. This finding is valid

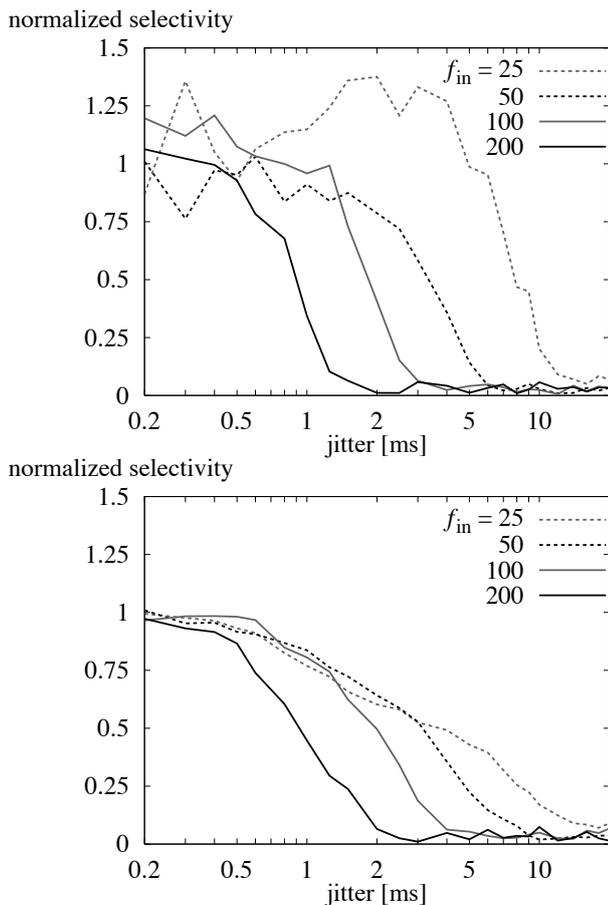


Fig. 8 The selectivity as defined in (37) for the feedforward (top) and recurrent (bottom) model for several input signal frequencies f_{in} as a function of jitter. The selectivity Q is normalized with respect to the value in the absence of jitter. As expected, increasing jitter leads to a decrease in selectivity. For both the feedforward and the recurrent model a jitter of 20% of the input period leads to a 50% decrease in selectivity.

for various animals [46, 42, 57–59]. Relevant biological stimuli on the water surface and in spider webs also tend to contain most of their information in the low frequency range $\lesssim 250$ Hz [10, 43]. These findings agree very well with our calculations.

In section 3 we have seen that the delay times Δ need not necessarily arise from a direct connection between two neurons but that they could be the result of a number of interneurons. These interneurons then have to be driven by a very reliable synapse: every input spike should trigger an output spike, and the delay between input and output spike should be fixed. A very prominent example of such a reliable “one-to-one” synapse in the auditory pathway is the so-called *Calyx of Held* at the end of the auditory nerve.

Although this specific type of reliable synapse is only found in the lower auditory pathway its existence demonstrates that fast and reliable synapses are present in the auditory system, a neuronal system of exceptional acuity. For example, in the mammalian auditory brain-stem nuclei neu-

rons can preserve the relative timing of action potentials passed through sequential synaptic levels [67]. In the avian auditory system, too, single presynaptic stimuli can produce short (and thus precise) suprathreshold spikes with a time constant of about 0.5 ms resulting in reliable information transmission [71]. Another possibility to reliably transfer precisely-timed signals is the use of synfire chains [28]. Depending on the input strength, synfire chains can relay information with a temporal precision around 1 ms, accurate enough for use in long-delay feedback and feedforward loops.

We conclude that relatively long and well-defined delay times Δ can be realized in biological systems by means of interneurons, thus posing no fundamental problem to our model.

6.2 Conclusions

In this work we have quantitatively analyzed two different models of periodicity detection. We have shown that both a feedforward architecture and a recurrent loop architecture can be used to extract periodic modulation from input signals. Furthermore, we have provided an extensive mathematical characterization. It has been shown that for both approaches the basic constraints are the same.

As expected, neuronal time constants are a limiting factor for recognizing the periodicity of the input modulation. The non-zero width of the response kernel theoretically limits modulation recognition to about $\lesssim 1000$ Hz. In real biological systems, however, the limiting factor will probably be the accuracy of the delay Δ . Using an estimate $\delta\Delta \approx 0.5$ ms cuts down the accessible detection range to about 200 Hz. This value is reasonably close to the reported limitations of about 300 Hz in biological systems.

Another limitation common to both feedforward and recurrent circuitries is that they detect only the highest modulation frequency components in any signal. Since activity in high-frequency channels also excites low-frequency channels it is not possible to distinguish subharmonics of a high-frequency signal from a direct low-frequency input. The known phenomenon of the missing fundamental fits well into the behavior of such a simple network for periodicity extraction. Equivalent to the above is the fact that every neuron responds not only to its own specific frequency but also to all of its harmonics. Consequently the perceived similarity between tones one octave apart from each other [35, 27, 26] and the interference of harmonic target-distractor combinations at low frequencies [15] are a natural side-effect of the proposed architecture.

The two models differ in their behavior as far as their robustness is concerned. By design, the recurrent network is much more susceptible to noise and, as a consequence, can be disturbed by noise more easily than the feedforward model. This is a common problem of excitatory recurrent networks in general since in such networks perturbations tend to amplify themselves.

We have shown that without any specialized architectural features—a generic neuron model and simple delay

lines— the modulation frequency components of a signal can be resolved neurally. Considering the simplicity of the setup, the models already show remarkable properties that are in good agreement with experimental data. We stress that no parameter tuning was needed to obtain our results.

A small number of input neurons suffices to sample the input signal with a high enough accuracy, and neuronal parameters lie in the range of milliseconds, comparable to typical auditory time scales. If a cochlea is present, signal processing is further facilitated by the prior frequency decomposition but the presence of a cochlea is certainly no prerequisite for periodicity analysis. The output of the networks could be enhanced by postprocessing mechanisms. For example, lateral inhibition can be used to detect and sharpen the peaks in the output (consider Fig. 4, 5 and 6 in this regard).

At this point we want to emphasize that *extracting* the slowly-varying envelope from an input signal is easily accomplished biologically. Half-wave rectification of the signal and low-pass filtering suffice. This can be accomplished by a slow synapse filtering out all high frequency components, a mechanism that has been demonstrated explicitly to work in the electric fish sensory pathway [54]. Even if the slow envelope of the signal has been extracted, however, the frequency of the envelope oscillations is still unknown. Our models, on the contrary, are able to *identify* the frequency of the envelope oscillations. This is of great biological importance since recognition of sounds often depends on the ability to quantitatively determine the periodicity of the input signal.

As an example, human speech consists of several frequency bands that are comodulated by a guttural fundamental frequency called the *voicing frequency* or *fundamental frequency* [2]. A speaker can be distinguished by his own voicing frequency and this immediately suggests a mechanism for the identification and separation of people speaking simultaneously. It is possible to separate contributions from different speakers by focussing attention on specific periodicity frequencies. The now isolated sound source can then be localized using standard mechanisms based on interaural time differences (ITD) recognition, since phase locking to the input remains intact throughout the processing in our models. In such a setting, auditory object recognition would therefore occur before the localization of the object. This agrees with previous experimental work showing that spatial separation of sounds is indeed linked to comodulation of the signal amplitude across several frequency channels [24, 22, 45].

Acknowledgements M.B. and J.L.v.H. were partially supported by the BCCN–Munich and P.F. was funded by the Deutsche Forschungsgemeinschaft.

A Model neurons

A.1 Poisson neuron

A Poisson neuron is a simple model neuron with stochastic firing dynamics. The neuron lacks a threshold and may therefore be seen as biologically unrealistic. On the other hand, precisely because of the lacking threshold, the neuron can be described very easily mathematically [68]. The firing dynamics of a Poisson neuron is determined by an instantaneous rate function, or spike probability density λ , which in general depends on time. The probability that the neuron fires in the small time interval $[t, t + \delta t)$ is given by

$$P_{\text{spike in } [t, t + \delta t)} = \lambda(t) \delta t . \quad (38)$$

The stochastic firing dynamics is *completely* determined by the rate function λ , which means that firing events do not influence the future firing probability. If the size of the time interval δt is taken small enough the probability of getting more than one spike within that interval (even for very high firing rates) is taken to be negligible *viz.*, $o(t)$. The firing of the neuron is thus a *renewal process* and refractoriness is taken into account through the $o(t)$ -condition.

The expectation value of the output from a Poisson neuron can be calculated analytically. If a spike occurring at time $t = t_0$ of the Poisson neuron gives rise to a response $g(t - t_0)$, a post-synaptic current, the total response is given by [68]

$$\int_{-\infty}^t ds g(t - s) \lambda(s) . \quad (39)$$

If we choose $g(t)$ so that causality is ensured, i.e. $g(t) = 0$ for $t < 0$, we can write the expectation of the output

$$\int_{-\infty}^{\infty} ds g(t - s) \lambda(s) . \quad (40)$$

Since a Poisson neuron fires stochastically two realizations of the firing dynamics governed by some particular $\lambda(t)$ will always differ. The variance of the output response can also be calculated explicitly and is given by [68]

$$\int_{-\infty}^{\infty} ds g^2(t - s) \lambda(s) . \quad (41)$$

A.2 Leaky integrate-and-fire neuron

The firing dynamics of a leaky integrate-and-fire (LIF) neuron is governed by a differential equation for the membrane potential V [31],

$$\frac{dV}{dt} = -(V - V_0) / \tau_{\text{mem}} + (I_{\text{ext}} + I_{\text{noise}}) / C_{\text{mem}} . \quad (42)$$

The potential changes under influence of an external input current I_{ext} that drives the neuron. If no input current is present the potential relaxes to a resting value V_0 with characteristic membrane time constant τ_{mem} . The last term, I_{noise} , takes into account internal noise of the neuron that will be needed in the case of the recurrent model. The constant C_{mem} is the membrane conductance of the neuron determining how effectively the current can change the membrane potential.

If the potential in (42) reaches a certain threshold value V_θ a spike occurs and the potential is reset to a value V_R . Refractoriness of the neuron can be taken into account by disallowing the neuron to fire for a certain period after spiking, by changing the threshold voltage temporarily to a higher value, or by temporarily ignoring the input current (see also [31]).

B Numerical implementation of the models

Both neuronal models have been implemented using the C++ programming language. Below, all model parameters are listed, as well as the internal noise mechanism adopted for the neurons in the recurrent model.

B.1 Recurrent model

parameter	value
number of output neurons	$N_{\text{out}} = 491$
output frequency range	10 – 500 Hz
synaptic time constant	$\tau = 1$ ms
synaptic strength	$J = 2.5 \times 10^{-5}$
input normalization	$1/T \int_0^T s_{\text{in}} dt = 300$
output neuron	
membrane time	$\tau_{\text{mem}} = 1.25$ ms
absolute refraction time	$\tau_{\text{ref}} = 1.0$ ms
rest potential	$V_0 = 0$
reset potential	$V_R = V_0 = 0$
threshold	$V_\theta = 1$
capacitance	$C_{\text{mem}} = 1$

Internal noise of the neurons has been implemented by adding a noise term I_{noise} to the input of each neuron given by

$$I_{\text{noise}} = \sum_{n=1}^{50} A_{\text{noise}} \cos(2\pi f_{\text{noise}}^n t + \phi_{\text{noise}}^n) \quad (43)$$

where the frequencies are chosen from a uniform distribution $f_{\text{noise}}^n \in [0 - 1000$ Hz]. Phases are uniformly distributed in $\phi_{\text{noise}}^n \in [0 - 2\pi]$ and the amplitude of every component is given by $A_{\text{noise}} = 0.01/50$. For each neuron, independent noise is assumed and the noise is then added linearly to the input for each neuron.

B.2 Feedforward model

parameter	value
number of input neurons	$N_{\text{in}} = 25$
number of output neurons	$N_{\text{out}} = 491$
output frequency range	10 – 500 Hz
input neuron mean rate	20 Hz
synaptic time constant	$\tau = 1$ ms
synaptic strength	$J = 3.5 \times 10^{-4}$
output neuron	
membrane time	$\tau_{\text{mem}} = 1$ ms
absolute refraction time	$\tau_{\text{ref}} = 0.25$ ms
rest potential	$V_0 = 0$
reset potential	$V_R = V_0 = 0$
threshold	$V_\theta = 1$
capacitance	$C_{\text{mem}} = 1$

References

- Aicher, B., Tautz, J.: Signal transmission through the substrate. *J Comp Physiol A* **166**, 345 (1990)
- Avendaño, C., Deng, L., Hermansky, H., Gold, B.: The analysis and representation of speech. In: S. Greenberg, W. Ainsworth, A. Popper, R. Fay (eds.) *Speech Processing in the Auditory System*, chap. 2, p. 63. Springer, New York (2004)
- Barth, F.: Neuroethology of the spider vibration sense. In: F. Barth (ed.) *Neurobiology of Arachnids*, chap. 11, p. 203. Springer, New York (1985)
- Barth, F.: The vibrational sense of spiders. In: R. Hoy, A. Popper, R. Fay (eds.) *Comparative Hearing: Insects*, chap. 7, p. 228. Springer, New York (1998)
- Barth, F., Geethabali: Spider vibration receptors: Threshold curves of individual slits in the metatarsal lyriform organ. *J Comp Physiol A* **148**, 175 (1982)
- Bendor, D., Wang, X.: The neuronal representation of pitch in primate auditory cortex. *Nature* **436**, 1161 (2005)
- Bleckmann, H.: Prey identification and prey localization in surface-feeding fish and fishing spiders. In: J. Atema, R. Fay, A. Popper, W. Tavolga (eds.) *Sensory Biology of Aquatic Animals*, chap. 24, p. 619. Springer, New York (1988)
- Bleckmann, H.: Reception of Hydrodynamic Stimuli in Aquatic and Semiaquatic Animals. Gustav Fischer, Stuttgart (1994)
- Bleckmann, H., Borchard, M., Horn, P., Görner, P.: Stimulus discrimination and wave source localization in fishing spiders (*Dolomedes triton* and *D. okefinokensis*). *J Comp Physiol A* **174**, 305 (1994)
- Bleckmann, H., Breithaubt, T., Blickhan, R., J., T.: The time course and frequency content of hydrodynamic events caused by moving fish, frogs and crustaceans. *J Comp Physiol A* **168**, 749 (1991)
- Bleckmann, H., Waldner, I., Schwartz, E.: Frequency discrimination in the surface-feeding fish *Aplocheilichthys lineatus* – A prerequisite for prey localization? *J Comp Physiol A* **143**, 485 (1981)
- Borst, M., Langner, G., Palm, G.: A biologically motivated network for phase extraction from complex sounds. *Biol Cybern* **90**, 98 (2004)
- Bregman, A.: *Auditory Scene Analysis*. MIT Press, Cambridge MA (1990)
- Brownell, P.: Compressional and surface waves in sand: Used by desert scorpions to locate prey. *Science* **197**, 479 (1977)
- Buell, T., Hafter, E.: Combination of binaural information across frequency bands. *J Acoust Soc Am* **90**, 1894 (1991)
- Burkitt, A.: A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biol Cybern* **95**, 1 (2006)
- Burkitt, A.: A review of the integrate-and-fire neuron model: II. Inhomogeneous synaptic input and network properties. *Biol Cybern* **95**, 97 (2006)
- Burkitt, A., Clark, G.: Synchronization of the neural response to noise periodic synaptic input. *Neural Comp* **13**, 2639 (2001)
- Cariani, P.: Neural timing nets. *Neur Netw* **14**, 737 (2001)
- Cariani, P.: Recurrent timing nets for auditory scene analysis. *Proc Int Joint Conf Neural Netw* p. 1575 (2003)
- Cherry, E.: Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* **25**, 975 (1953)
- Cooke, M., Ellis, D.: The auditory organization of speech and other sources in listeners and computational models. *Speech Comm* **35**, 141 (2001)
- Coombs, S., Görner, P., Münz, H.e.: *The Mechanosensory Lateral Line: Neurobiology and Evolution*. Springer, New York (1989)
- Culling, J., Summerfield, Q.: Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *J Acoust Soc Am* **98**, 785 (1995)
- Dean, I., Harper, N., McAlpine, D.: Neural population coding of sound level adapts to stimulus statistics. *Nature Neurosci* **8**, 1684 (2005)
- Demany, L., Semal, C.: Dichotic fusion of 2 tones one octave apart: Evidence for internal octave templates. *J Acoust Soc Am* **83**, 687 (1988)
- Deutsch, D.: Octave generalization of specific interference effects in memory for tonal pitch. *Percept Psychophys* **13**, 271 (1973)
- Diesmann, M., Gewaltig, M.O., Aertsen, A.: Stable propagation of synchronous spiking in cortical neural networks. *Nature* **402**, 529 (1999)
- Elepfandt, A.: Wave frequency recognition and absolute pitch for water waves in the clawed frog *Xenopus laevis*. *J Comp Physiol A* **158**, 235 (1986)
- Gammaitoni, L., Hänggi, P., Jung, P., Marchesoni, F.: Stochastic resonance. *Rev Mod Phys* **70**, 223 (1998)

31. Gerstner, W., Kistler, W.: Spiking Neuron Models. Cambridge University Press, Cambridge (2002)
32. Grothe, B., Klump, G.: Temporal processing in sensory systems. *Curr Opin Neurobiol* **10**, 467 (2000)
33. Hergenröder, R., Barth, F.: The release of attack and escape behavior by vibratory stimuli in a wandering spider (*Cupiennius salei* Keys.). *J Comp Physiol A* **152**, 347 (1983)
34. Hudspeth, A., Corey, D.: Sensitivity, polarity, and conductance change in the response of vertebrate hair cells to controlled mechanical stimuli. *Proc Natl Acad Sci USA* **74**, 2407 (1977)
35. Humphreys, L.: Generalization as a function of method of reinforcement. *J Exp Psych* **25**, 361 (1939)
36. Ingham, N., McAlpine, D.: Spike-frequency adaptation in the inferior colliculus. *J Neurophysiol* **91**, 632 (2004)
37. Joris, P., Schreiner, C., Rees, A.: Neural processing of amplitude-modulated sounds. *Physiol Rev* **84**, 541 (2004)
38. Kalmijn, A.: Hydrodynamic and acoustic field detection. In: J. Atema, R. Fay, A. Popper, W. Tavolga (eds.) *Sensory Biology of Aquatic Animals*, chap. 4, p. 83. Springer, New York (1988)
39. Käse, R., Bleckmann, H.: Prey localization by surface ray-tracing: Fish track bugs like oceanographers track storms. *Experientia* **43**, 290 (1987)
40. Kempster, R., Gerstner, W., Van Hemmen, J., , Wagner, H.: Extracting oscillations: Neural coincidence detection with noise periodic spike input. *Neural Comp* **10**, 1987 (1998)
41. Kempster, R., Gerstner, W., Van Hemmen, J.: How the threshold of a neuron determines its capacity for coincidence detection. *BioSys* **48**, 105 (1998)
42. Krishna, B., Semple, M.: Auditory temporal processing: Response to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* **84**, 255 (2000)
43. Landolfa, M., Barth, F.: Vibrations in the orb web of the spider *Nephilia clavipes*: Cues for discrimination and orientation. *J Comp Physiol A* **179**, 493 (1996)
44. Lang, H.: Surface wave discrimination between prey and non-prey by the backswimmer *Notonecta glauca* L. (Hemiptera, Heteroptera). *Beh Ecol Sociobiol* **6**, 233 (1980)
45. Langner, G.: Periodicity coding in the auditory system. *Hear Res* **60**, 115 (1992)
46. Langner, G., Schreiner, C.: Periodicity coding in the inferior colliculus of the cat. i. neuronal mechanisms. *J Neurophysiol* **60**, 1799 (1988)
47. Licklider, J.: A duplex theory of pitch perception. *Experientia* **7**, 128 (1951)
48. Magal, C., Schöller, M., Tautz, J., Casas, J.: The role of leaf structure in vibration propagation. *J Acoust Soc Am* **108**, 2412 (2000)
49. Masters, W.: Vibrations in the orbwebs of *Nuctenea sclopetaria* (Araneidae). *Beh Ecol Sociobiol* **15**, 217 (1984)
50. Masters, W., Markl, H., Moffat, A.: Transmission of vibration in a spider's web. In: W. Shear (ed.) *Spiders: Webs, Behavior, and Evolution*, chap. 3, p. 49. Stanford University Press, Stanford, CA (1986)
51. Meddis, R., O'Mard, L.: A unitary model of pitch perception. *J Acoust Soc Am* **102**, 1811 (1997)
52. Meddis, R., O'Mard, L.: Virtual pitch in a computational physiological model. *J Acoust Soc Am* **120**, 3861 (2006)
53. Megela Simmons, A., Ferragamo, M.: Periodicity extraction in the anuran auditory nerve. *J Comp Physiol A* **172**, 57 (1993)
54. Middleton, J., Longtin, A., Benda, J., Maler, L.: The cellular basis for parallel neural transmission of a high-frequency stimulus and its low-frequency envelope. *Proc Natl Acad Sci USA* **103**, 14,596 (2006)
55. Nelken, I., Rotman, Y., Bar Yosef, O.: Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* **397**, 154 (1999)
56. Oertel, D.: The role of timing in the brain stem auditory nuclei of vertebrates. *Ann Rev Physiol* **61**, 497 (1999)
57. Rees, A., Møller, A.: Responses of neurons in the inferior colliculus of the rat to am and fm tones. *Hear Res* **10**, 301 (1983)
58. Rees, A., Møller, A.: Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear Res* **27**, 129 (1987)
59. Rees, A., Palmer, A.: Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. *J Acoust Soc Am* **85**, 1978 (1989)
60. Schreiner, C., Langner, G.: Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol* **60**, 1823 (1988)
61. Schuller, G.: Natural ultrasonic echoes from wing beating insects are encoded by collicular neurons in the CF-FM bat, *Rhinolophus ferrumequinum*. *J Comp Physiol A* **155**, 121 (1984)
62. Shannon, R., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M.: Speech recognition with primarily temporal cues. *Science* **270**, 303 (1995)
63. Smith, J., Marsh, J., Greenberg, S., Brown, W.: Human auditory frequency-following responses to a missing fundamental. *Science* **201**, 639 (1978)
64. Smith, R., Zwislocki, J.: Short-term adaptation and incremental responses of single auditory-nerve fibers. *Biol Cybern* **17**, 169 (1975)
65. Smith, Z., Delgutte, B., Oxenham, A.: Chimaeric sounds reveal dichotomies in auditory perception. *Nature* **416**, 87 (2002)
66. Speck-Hergenröder, J., Barth, F.: Tuning of vibration sensitive neurons in the central nervous system of a wandering spider, *Cupiennius salei* Keys. *J Comp Physiol A* **160**, 467 (1987)
67. Trussell, L.: Synaptic mechanisms for coding timing in auditory neurons. *Ann Rev Physiol* **61**, 477 (1999)
68. Van Hemmen, J.: Theory of synaptic plasticity. In: F. Moss, S. Gielen (eds.) *Handbook of Biophysics*, Vol. 4, Neuro-Informatics and Neural Modelling, chap. 18, p. 771. Elsevier, Amsterdam (2001)
69. Westerman, L., Smith, R.: Rapid and short-term adaptation in auditory nerve responses. *Hear Res* **15**, 249 (1984)
70. Yost, W.: Auditory image perception and analysis: The basis for hearing. *Hear Res* **56**, 8 (1991)
71. Zhang, S., Trussell, L.: A characterization of excitatory postsynaptic potentials in the avian nucleus magnocellularis. *J Neurophysiol* **72**, 705 (1994)