A developmental learning rule for coincidence tuning in the *barn owl* auditory system

Wulfram Gerstner*, Richard Kempter J.Leo van Hemmen
Institut für Theoretische Physik,
Physik-Department der TU München
D-85748 Garching bei München, Germany
and
Hermann Wagner
Lehrstuhl für Zoologie/Tierphysiologie
Institut für Biologie II, RWTH Aachen
Kopernikusstr. 16, D-52074 Aachen, Germany

Binaural coincidence detection is essential for the localization of external sounds and requires auditory signal processing with high temporal precision. We present an integrate-and-fire model of spike processing in the third order nucleus laminaris of the auditory pathway of the barn owl. Each input spike generates an excitatory postsynaptic potential with a width of 250 μs. Output spikes occur with a tenfold enhanced temporal precision. This is possible since neuronal connections are fine tuned during a critical period of development as has been suggested by recent experiments. This rule does not only explain the temporal presicion in the output, but causes also a tuning to interaural time difference. The learning rule is of the Hebbian type: A synaptic weight is increased if a presynaptic spike arrives at about the same time as or slightly before postsynaptic firing. A presynaptic spike arriving after postsynaptic firing leads to a decrease of the synaptic efficacy.

1 Introduction

Owls are able to locate acoustic signals based on the extraction of interaural time difference (ITD) by coincidence detection [1, 2, 3]. The spatial resolution of sound localization found in behavioral experiments corresponds to a temporal resolution of auditory signal processing in the range of a few μs. It follows that

*present address: Swiss Federal Insititute of Technology, Center for Neuromimetic Systems, EPFL-DI, CH-1015 Lausanne, email: wgerst@di.epfl.ch

both the firing of spikes and their transmission along the so-called time pathway of the auditory system must occur with high temporal precision.

Each neuron in the nucleus laminaris, the third processing stage in the ascending auditory pathway, responds to signals in a narrow frequency range. Its spikes are phase locked to a stimulation tone for frequencies up to 8 kHz [4, 5]. The nuclues laminaris is of particular interest, because it is the first processing stage where signals from the right and left ear converge. Owls use the interaural phase difference (that is, the interaural time difference modulo the period T of the stimulus) for azimuthal sound localization. For a successful localization, the temporal precision of spike encoding and transmission must be at least in the range of some 10 μs.

This poses at least two severe problems. First, the neural architecture has to be adapted to operating with high temporal precision. Considering the fact that the total delay from the ear to the nucleus magnocellularis is approximately 2-3 ms [4], a temporal precision of some 10 μs requires some fine tuning, possibly based on learning. Recent experiments have shown that the tuning to ITDs evolves during a critical period of the development [6]. Here we suggest that Hebbian learning is an appropriate mechanism. Second, neurons must operate with the necessary temporal precision. A firing precision of some 10 μs seems truly remarkable considering the fact that neuronal time constants are probably in the range of 100-300 μs [7, 8, 9]. It is shown below that neuronal spikes can be transmitted with a temporal precision of 25 μs despite the width of a single excitatory postsynaptic potential (EPSP) of 250 μs. This is possible because spikes are always triggered during the rise time of an EPSP.

2 Neuron model

We concentrate on a single frequency channel of the auditory pathway and model a neuron of the nucleus laminaris. Since synapses are directly located on the *soma* or very short dendrites, the spatial structure of the neuron can be reduced to a single compartment. In order to simplify the dynamics, we take an integrate-and-fire unit. Its membrane potential changes according to

$$\frac{d}{dt}u = -\frac{u}{\tau_0} + I(t) \tag{1}$$

where $I(t)$ is some input and τ_0 is the membrane time constant. The neuron fires, if $u(t)$ crosses a threshold ϑ. This defines a firing time $t_{\rm out}$. After firing u is reset to an initial value $u_0 = 0$. Since auditory neurons are known to be fast, we assume a membrane time constant of 100 μs.

The laminaris neuron receives input from several presynaptic neurons $1 \leq k \leq K$. Each input spike at time t_k^f generates a current pulse which decays exponentially with a time constant $\tau_r = 100$ μs. The magnitude of the current pulse depends on the coupling strength J_k. The total input is

$$I(t) = \sum_{k=1}^{K} \sum_{f=1}^{F_k} J_k \exp(\frac{t - t_k^f}{\tau_r})\, \theta(t - t_k^f) \tag{2}$$

where $\theta(x)$ is the unit step function and the sum runs over all synapses $1 \leq k \leq K$ and all input spikes $1 \leq f \leq F_k$. Combining Eqs. (1) and (2) we find that each input evokes an EPSP with a width of 250 μs at half maximum. This decay is twice as fast as in chicken [7, 8, 9], but seems reasonable for an expert auditory-cue based hunter like the barn owl.

We focus on a laminaris neuron which is tuned to a 5 kHz signal (period $T = 200\mu$s) and stimulate both ears with a tone of the optimal frequency and vanishing interaural time difference (ITD=0). Input spikes from presynaptic terminals are phase locked to the stimulus tone with a jitter of 40 μs. The mean phase of spikes arriving at a synapse k is different from the mean phase at another synapse k' because the transmission lines from the ear to the nucleus laminaris have different delays. It has been estimated from experimental data that the signal transmission delay Δ_k from the ear to the nucleus laminaris varies between 2 and 3 ms [4]. Note that a delay as small as 0.2 ms shifts the signal by a full period. Thus, if the delays Δ_k for the synapses k with $1 \leq k \leq K$ are chosen at random between 2 and 3 ms, then the periodicity of the signal is lost after integrating over different presynpatic inputs. We therefore need a tuning process which selects transmission lines with matching delays.

3 Temporal tuning through learning

We assume a developmental period of unsupervised learning during which a fine tuning of the temporal characteristics of signal transmission takes place. Before learning the laminaris neuron receives many inputs ($K = 600$, 300 from each the left and the right ear) with weak coupling ($J_k = 1$). Due to the broad distribution of delays the total input (2) has, apart from fluctuations, no temporal structure. The output is not sensitive to the interaural time difference (ITD); cf. Fig. 1, top. During learning some connections increase their strength to a maximum of $J_{\max} = 3$, others are removed. A slight sensitivity to the ITD develops as seen in experiments [6]; cf. Fig. 1, middle. After learning, the laminaris neuron receives 75 inputs from each ear [5]. The connections to those neurons have become very effective; cf. Fig. 1, bottom. The surviving inputs all have nearly the same time delay modulo the period T of the stimulus. Thus, after learning the number of synapses is reduced and only inputs with matching delays survive. On the other hand, with 75 inputs remaining from each ear there is a high degree of convergence or 'pooling' which leads to a significant increase in temporal precision. To see this, we compare the phase jitter of the output spikes with that of the presynaptic input lines. The phase jitter of the presynaptic spike trains corresponds to a temporal precision of 40 μs. In contrast, after learning the output spikes have a temporal precision of 25 μs. Thus our learning rule leads to the fine tuning of the neuronal connections necessary for precise temporal coding.

We use a variant of Hebbian learning. In standard Hebbian learning, synaptic weights are changed if pre- and postsynaptic activity occurs simultaneously. In the context of temporal coding by spikes, the concept of 'simultaneous activity' has to be refined. We assume that a synapse k is changed, if a presynaptic spike t_k^f and a postsynaptic spike t_{out} occur within a *learning window* $W(t_k^f - t_{\text{out}})$. More precisely, each pair of presynaptic and postsynaptic spikes changes a synapse J_k by an amount

$$\Delta J_k = W(t_k^f - t_{\text{out}}). \tag{3}$$

Depending on the sign of $W(x)$, a contact to a presynaptic neuron is either increased or decreased, in agreement with recent experimental results [10]. A decrease below $J_k = 0$ and an increase above $J_{\max}$ is not allowed. In our model, we assume a function $W(x)$ with two phases. For $x \approx 0$, the function $W(x)$ is positive. This leads to a strengthening (potentiation) of the contact with a presynaptic neuron k which is active shortly before or after a postsynaptic spike. Synaptic contacts which become active more than 3 ms later than the postsynaptic spike are decreased. Finally, an active synapse is always increased by a small amount γ, even if no output spike occurs. The effect of learning on a synapse k after F_k input spikes and N_{out} output spikes is

$$\Delta J_k = F_k\,\gamma + \sum_{n=1}^{N_{\text{out}}} \sum_{f=1}^{F_k} W(t_k^f - t_{\text{out}}^n)\,. \tag{4}$$

We require $\int W(s)ds = -W_0 < 0$. The combination of decrease and increase then balances the average effects of potentiation and depression and leads to a normalization of the number and weight of synapses.

Learning occurs in three phases. At the beginning, input and output spikes are not correlated. In particular, output spikes are not phase locked to the stimulation tone. During this phase, all weights J_k approach a transitory equilibrium (defined by $\langle \Delta J_k \rangle = 0$) which leads to a mean output firing rate $r = \gamma / W_0$. We recall that each synapse k corresponds to a delay line with a delay Δ_k. In Fig. 1 we have plotted the value J for each synapse as a function of the delay Δ. The initial distribution is flat.

Shot noise due to the spike nature of the signals leads to spontaneous symmetry breaking and starts the second phase of learning (Fig. 1, middle). As a result, $J(\Delta)$ develops a T-periodic modulation. In the third and final stage of learning, all values $J(\Delta)$ approach a limiting value, either 0 or $J_{\max} = 3$. The output spikes now show a high degree of phase locking measured by the vector strength. Moreover, the output firing rate r depends on the interaural time difference (ITD), as found in experiments with adult owls.

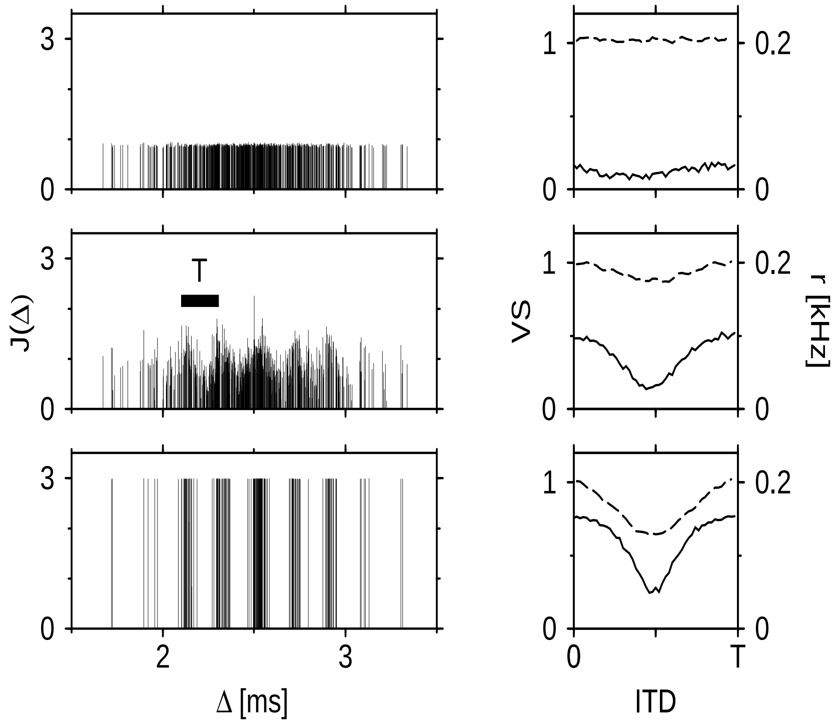

Figure 1: *Development of tuning to a 5kHz tone.* In the left column, we show the strength of synaptic contacts as a function of the delay Δ. On the right-hand side, we show the vector strength (vs, solid line) and the output firing rate (r, dashed) as a function of the interaural time delay (ITD). **Top**. Before learning, there are 600 synapses (300 from each ear) with different delays, chosen randomly from a Gaussian distribution with mean 2.5 ms and variance 0.3 ms. The output is not phase-locked ($vs \approx 0.1$) and shows no dependence upon the ITD. **Middle**. During learning, some synapses are strengthened others decreased. Those synapses which increase have delays which are similar to each other or differ by multiples of the period $T = 0.2$ ms of the stimulating tone. The vector strength of the output increases and starts to depend on the ITD [6]. **Bottom**. After learning, only 150 synapses (75 from each ear) survive. Both the firing rate r and the vector strength vs show the characteristic dependence upon the ITD as seen in experiments with adult owls [5]. The neuron has the maximal response ($r = 200$ Hz) for ITD=0 which is the stimulus used during the learning session of the model neuron. The vector strength at ITD=0 is $vs \approx 0.8$ which corresponds to a temporal precision of 25 μs. By definition, the vector strength is proportional to the amplitude of the first Fourier component of the distribution of output phases, normalized so that $vs = 1$ indicates perfect phase locking (infinite temporal precision).

Acknowledgments: R.K. holds scholarship of the state of Bavaria. W.G. has been supported by the Deutsche Forschungsgemeinschaft (DFG) under grant number He 1729/2-2, H.W. by a Heisenberg fellowship of the DFG.

References

[1] L. A. Jeffress, J. Comp. Physiol. Psychol. **41**, 35 (1948).

[2] M. Konishi, Trends in Neurosciences **9**, 163 (1986).

[3] C. E. Carr, Annual Rev. Neurosci. **16**, 223 (1993).

[4] W. E. Sullivan and M. Konishi, J. Neurosci. **4**, 1787 (1984).

[5] C. E. Carr and M. Konishi, J. Neurosci. **10**, 3227 (1990).

[6] C. E. Carr, in *Proceedings of the 10th International Symposium on Hearing* (Irsee, Bavaria, 1994), pp. 21–26.

[7] E. M. Overholt, E. W. Rubel, and R. L. Hyson, J. Neurosci. **12**, 1698 (1992).

[8] A. D. Reyes, E. W. Rubel, and W. J. Spain, J. Neurosci. **14**, 5352 (1994).

[9] A. D. Reyes, E. W. Rubel, and W. J. Spain, J. Neurosci. **16**, 993 (1996).

[10] H. Markram, J. Lübke, M. Frotscher, and B. Sakmann, preprint (1996)