

Finding the direction of a sound

N. Giordano and J. Wang

Department of Physics, Purdue University,

West Lafayette, Indiana 47907-1396

Abstract

It is believed that many animals determine the direction of a low frequency sound by comparing the times at which the sound arrives at its ears. Some years ago Jeffress proposed that this is accomplished by a network of coincidence detection neurons together with a system of delay line-like axons connecting these neurons to the two ears. There is substantial experimental evidence in favor of many aspects of this model, although some recent experiments, particularly in mammals, seem to be at odds with certain aspects of the Jeffress picture. We point out that from an engineering standpoint the original Jeffress approach to coincidence detection is not a very efficient way to accomplish sound localization. Recent experiments which seem difficult to reconcile with the Jeffress picture, might be understood in terms of a different coincidence detection scheme.

I. INTRODUCTION AND BACKGROUND

The problem of how an animal determines the direction of a sound has been discussed and studied for many years. Lord Rayleigh (1907) pointed out that for low frequency sounds, directional information is contained in the difference in the phases of the sound waves at the two ears, the so-called interaural phase difference (IPD). Some years ago Jeffress (1948) proposed that the IPD is ‘computed’ by an array of coincidence detector neurons. According to his picture (as translated into current terminology), each of these neurons receives input from both ears through axons which act as delay lines, with the difference in the delays associated with the left and right connecting axons, Δ_{LR} , varying systematically across the array of neurons. Sound from a particular direction will reach the closer ear first, so there will be an IPD which is a function of the direction of the sound. Jeffress proposed that for most of the coincidence detector neurons, the signals from the two ears will arrive at different times, and give rise to a firing rate which is zero or very small. However, there will be some neurons for which the external phase delay associated with the direction of the sound, Δ_{IPD} , is just compensated by the internal delay Δ_{LR} . For these neurons the two signals will arrive at the same time, producing a high firing rate. The location of these neurons within the array thus identifies the direction of the sound.

There have been a number of experimental and computational studies of the sound localization problem, and to the best of our knowledge, all are built upon the basic Jeffress picture. There is now good experimental evidence for a system of delay lines and the associated topographic layout of neurons in certain birds (Carr and Konishi, 1988; Carr, 1993). Computational studies have explored various issues, including the details of the coincidence detection process (Softky, 1994; Agmon-Snir et al. 1998; Zimon et al., 1999), and how a neural network might ‘learn’ the proper connections to perform localization with very high accuracy (Gerstner et al., 1996). However, a recent study in mammals seems to be at odds with the Jeffress picture (Brand et al., 2002), and there are other aspects of the model (see below) about which little is known either computationally or physiologically.

In this paper we make a few general observations about the processing of IPD, and point out that there are ways to accomplish coincidence detection which are inherently more efficient and accurate than the Jeffress scheme sketched above. A similar point has been made by Skottun (1998), although he did not emphasize the generality of the argument.

We then use a biophysical model to make a quantitative estimate of the gain in accuracy obtained by employing a different approach to coincidence detection. Recent experiments involving mammals (McAlpine et al., 2001; Brand et al., 2002) can be understood in terms an approach to IPD processing and coincidence detection like the one we propose.

II. THE JEFFRESS MODEL

Figure 1 is a schematic of the Jeffress model as it would apply to a mammal such as a gerbil or a cat (a similar picture applies to other animals, such as birds, but the relevant regions of the brain are different; Carr, 1993). Sound is received at the two cochlea, which project onto the medial and lateral cochlear nuclei (CN), and these in turn project (eventually) onto the medial superior olive (MSO), Fig. 1(a). According to the model, the MSO contains an array of neurons, labeled here as N1, N2, ... in Fig. 1(b), each of which receives inputs from both the medial and lateral CN. These auditory signals arrive as spikes (either action potentials or excitatory postsynaptic potentials) which are phased locked to the sound waves which excite the cochlea. The spike trains which leave the two CN are phase shifted relative to each other by an amount Δ_{IPD} which depends on the direction of the sound source. If this source is directly in front of the head (or directly behind) $\Delta_{IPD} = 0$, while it will increase in magnitude as the source moves farther and farther to one side of the head. The network of neurons determines Δ_{IPD} by comparing it with a second phase shift Δ_{LR} which is generated by the axons which come to the neurons from the CN. A neuron such as N1 will have a relatively short connection to the left CN and a longer connection to the right CN. This will cause the signal from the left CN to reach N1 sooner than the signal from the right CN, resulting in a phase shift Δ_{LR} caused by these two unequal delay lines. The neuron N2 as pictured here is slightly farther from the left CN (and closer to the right CN) and will have a correspondingly smaller value of Δ_{LR} . Hence, the neurons in this array will have values of Δ_{LR} which vary systematically with position in the MSO.

According to Jeffress, these neurons have the property that their firing rate is highest when spikes from the left and right CN arrive simultaneously (or nearly so), and very small otherwise. A neuron for which the two phase delays Δ_{IPD} and Δ_{LR} cancel will thus exhibit the highest firing rate. In this way the neurons N1, N2, ... act as coincidence detectors. Figure 1(c) shows in schematic form the variation of the firing rate of several neurons as

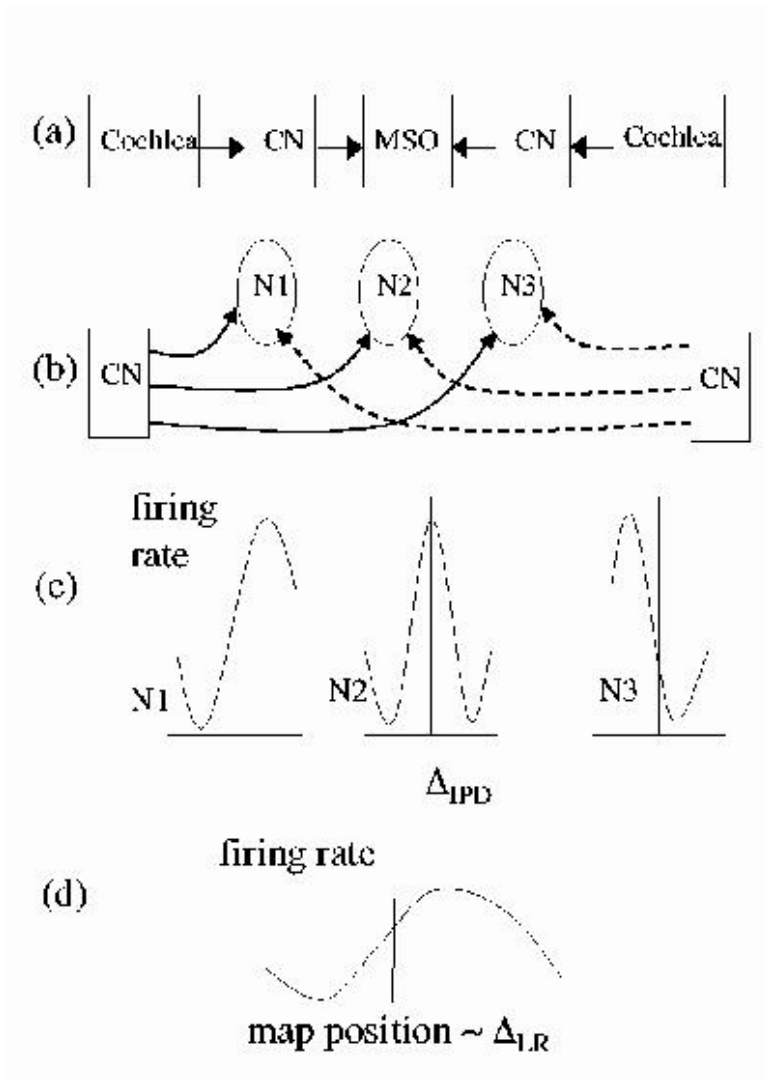


FIG. 1: The Jeffress model in a mammal. (a) Projection of auditory signals from the ears (cochlea) to coincidence detector neurons in the MSO. (b) Axonal delay line connections from the CN to the MSO neurons. (c) Firing rate functions for several MSO neurons as a function of the IPD. (d) Firing rate as a function of position in the MSO neuronal map.

a function of Δ_{IPD} . These three neurons have different values of the internal phase delay, Δ_{LR} , so their maximal firing rates occur at different values of Δ_{IPD} . For a given sound, the external phase difference Δ_{IPD} will be fixed, and the output of the coincidence detector array will appear as shown schematically, in Fig. 1(d). According to Jeffress, the location of the neuron with the highest firing rate within this neuronal place map identifies the direction of the incoming sound.

The Jeffress model has several important ingredients: (1) signals from the CN are trans-

mitted as spikes which are phase locked to the incoming sound waves; (2) a system of axons (or perhaps dendrites) act as delay lines as they carry these phase locked signals from the two CN to a region where the signals come together and are processed; (3) a collection of neurons act as coincidence detectors in processing the signals from the delay lines; (4) the coincidence detector neurons form a place map in which the direction of the sound is represented by the spatial location of the neuron with the highest firing rate. There seems to be good experimental evidence for (1-4) in several birds (Carr and Konishi, 1988; Carr, 1993). In mammals (1) and (2) have strong support, but there is less evidence for (3), and much less for (4) (Grothe, 2000). Indeed, as we will see below, recent work in gerbils (Brand et al., 2002) and other mammals (McAlpine et al., 2001) seems to be at odds with (4). In this paper we make a few observations about coincidence detection in general, and consider views of (3) and (4) which differ somewhat from the Jeffress picture. Our alternative approach to coincidence detection may explain the puzzling results for mammals, and might be relevant to birds as well.

III. COINCIDENCE DETECTION: AN ENGINEERING PERSPECTIVE

Figure 1(d) shows the variation of the firing rate with position within the map, or equivalently, as a function of the relative phase delay produced by the delay lines Δ_{LR} . This is the presumed output of the coincidence detector neurons. It is then assumed implicitly that the coincidence detector outputs project onto another set of neurons which ‘calculate’ the location of the maximum firing rate, and hence finish the task of computing the direction of the sound. We now wish to consider in a little more detail how the firing rate information in Fig. 1(d) is processed, and limits on the *accuracy* with which the maximum firing rate can be determined. The question of accuracy is a very important one. For example, owls are able to locate the direction of a sound with high precision, and the design of neural networks which can achieve the necessary precision is a non-trivial task (Gerstner et al, 1996).

The schematic firing rate curves in Fig. 1 ignore the influence of ‘noise.’ The firing rates exhibit fluctuations, i.e., noise, due to a number of sources, as have been discussed in numerous theoretical studies (Joris et al., 1994; Softky, 1994; Gerstner et al., 1996; Agmon-Snir et al., 1998). There are several places where noise can enter into the coincidence detector system. One important source is connected with the initial coding of the sound into a time

series of phase locked spikes. While this process does not seem to be precisely understood, experiments have shown that the phase locking of the spike train arriving at the coincidence detector neurons has a ‘jitter’ relative to the incident sound wave of typically $100 \mu\text{s}$ in mammals (Joris et al., 1994) and perhaps half of this (or even less) in some birds. The effect of this jitter can be made small by averaging the firing rate outputs over a sufficiently long time. However, it appears that animals are able to determine the direction of a sound within about 0.1 s, so the averaging time that one has to work with must be of this order (Gerstner et al., 1996). As a result of this jitter and other sources of noise, there will be an uncertainty $\pm\Delta F$ associated with the firing rate curve in Fig. 1(d). This uncertainty will result in a corresponding uncertainty in locating the maximum of the firing rate curve, and will thus limit the precision with which the sound can be localized.

The firing rate function in Fig. 1(d) is only a schematic; we will calculate its precise functional form, along with ΔF , for a particular biophysical model in the next section. However, the point we wish to make here depends only on the fact that this curve will generally have a maxima as in Fig. 1(d).

The output of the coincidence detector array will consist of a firing rate F which fluctuates within a range $F \pm \Delta F$, for every neuron in the array. According to the Jeffress picture, there must be a subsequent processing center in the brain that locates the maximum of this noisy function, and in this way computes the location of the sound. From an engineering perspective, this is a very inefficient way to use data of this form. It is hard to accurately determine which neuron has the maximum firing rate since the firing rate function is flat (i.e., has a vanishing slope) in this neighborhood. A better approach is to base detection on a point on the firing rate curve where the slope of this function is large. There is no fundamental reason to declare that the direction of a sound is the ‘place’ in the neuronal map where the firing rate is largest. Any other point can in principle be used, and when noise is present, the point where the slope is largest can be located most accurately.

The precise advantage in resolution depends on the shape of the firing rate function and the magnitude of the noise. However, we propose that nature will always be better served if the processing center that analyzes the firing rate data ‘locks on’ to a point on the curve where the slope is large, and uses this as a measure of the direction of the sound.

IV. A BIOPHYSICAL CALCULATION

To make our arguments more quantitative, we have calculated the firing rate function using a biophysical model of the coincidence detector neurons in the MSO. The model we use is the same as that developed by Rothman et al. (1993) to describe the bushy cells in a mammalian MSO (see also Brand et al., 2002). These cells possess fast and slow K channels, along with Na channels, with kinetic (and other) parameters as given by Rothman et al. (1993). In our model, each coincidence detector bushy cell has two short dendrites, and each dendrite contains a single synapse at which input from a CN is received as an excitatory postsynaptic potential (EPSP). These EPSPs are phase locked to the incoming sound signal, with one of these signals being delayed an extra amount Δt_{LR} , which is proportional to the phase delay Δ_{LR} . In addition, a random jitter of 0.1 ms is added to the arrival time of each EPSP. The calculation assumes that the sound is directly in front of the animal's head, so $\Delta_{IPD} = 0$.

Experiments show that the brain can determine the direction of a sound within a time of order 0.1 s, so we measure the firing rate F by simply counting the number of action potentials generated at the bushy cell soma during a time interval of this size. This measuring interval, combined with the jitter, determines the firing rate noise ΔF . Figure 2 shows results for the firing rate across the coincidence detector array, i.e., as a function of the phase difference Δ_{LR} . At each value of Δ_{LR} we have ‘measured’ F many times, with many simulations of length 0.1 s. These values of F were distributed in an approximately Gaussian manner with a mean \bar{F} , and a standard deviation σ which provides us with a measure of the noise ΔF . The solid curve in Fig. 2 shows \bar{F} while the dotted curves show $\bar{F} \pm \Delta F$.

From Fig. 2 we can see that if F is in the neighborhood of the maximum, then the uncertainty in determining the value of the phase delay at the one standard deviation level will be $\Delta_{LR} \approx \pm 0.18$ radian. On the other hand, if a point on the firing rate curve near half of the maximum rate is used as the reference point, then the uncertainty is ± 0.06 radian. In this way, nature can thus gain a factor of 3 in accuracy with the same basic neural circuitry.

This accuracy gain is, of course, dependent on many factors, including the properties of the coincidence detector neurons, the precise manner in which the firing rate information from the coincidence detector array is processed, and the amount of noise in the system. For example, if we reduce the phase locking jitter in our model calculation to 0.05 ms (a

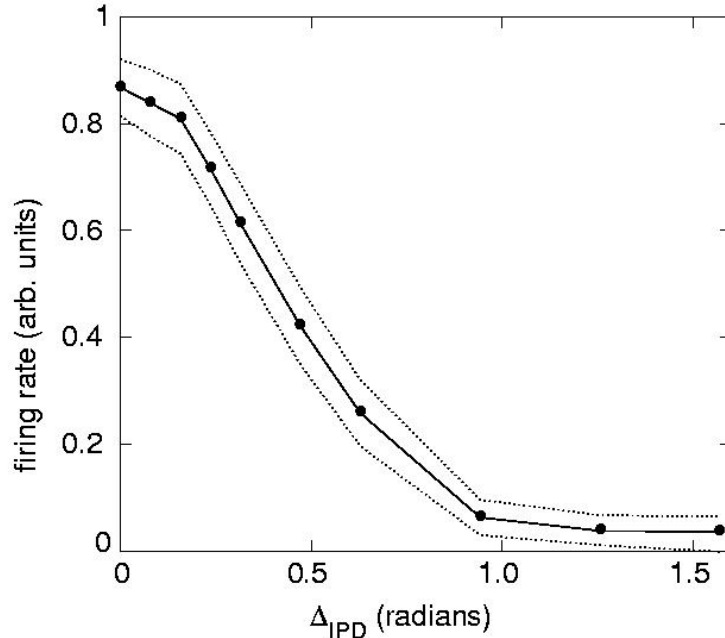


FIG. 2: Biophysical calculation of the firing rate function across the neuronal map. The solid symbols and connecting line is the average firing rate \bar{F} , while the dotted lines show $\bar{F} \pm \Delta F$.

factor of 2 smaller than assumed above), the gain in accuracy is increased to a factor of 5. However, our main point is independent of such details. It will always be more accurate to locate a point on the firing rate curve if the slope in that region is large, as opposed to using a region where the curve is flat.

V. DOES NATURE DO IT THIS WAY?

In the previous section we have pointed out that it is inherently more accurate to base location on a place where the firing rate curve has a large slope. In his original paper, Jeffress seems to suggest that using the point of maximal firing rate (where the slope of the firing rate curve vanishes) is most convenient and as far as we can tell, all subsequent workers have made the same assumption. However, a recent experiment seems to be at odds with this picture. Brand and coworkers (2002) studied the firing properties of neurons in the medial superior olive (MSO) of Mongolian gerbils, an animal known to use the IPD to localize sounds. They were able to find neurons that project onto the MSO and fire AP spikes which are phase locked to the inputs from the two CN. These observations are

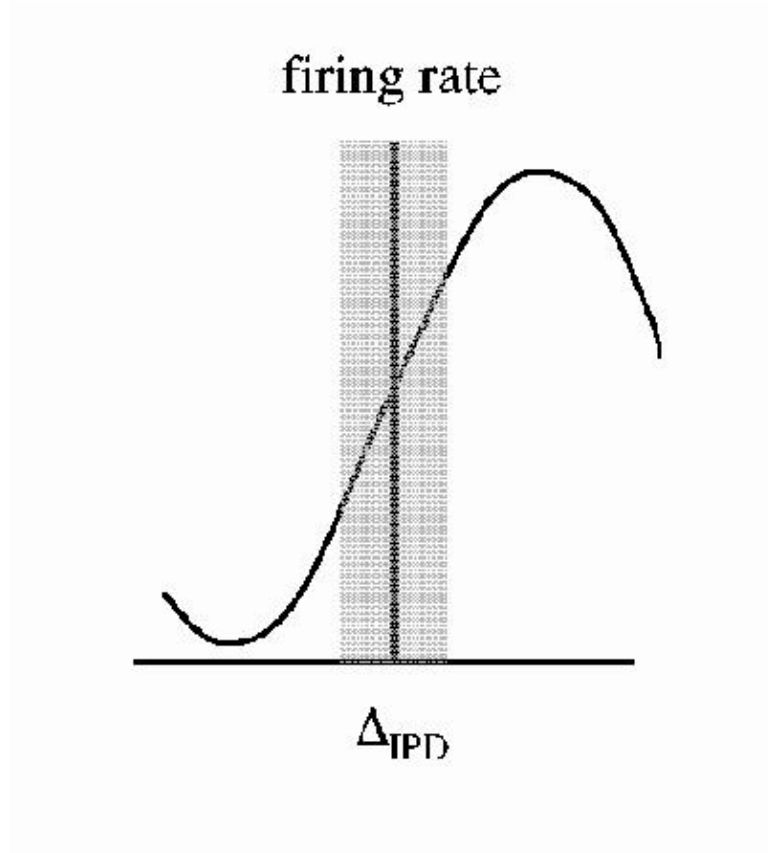


FIG. 3: Schematic of the gerbil firing rate function, after Brand et al. (2002). The shaded region shows the physiologically allowed values of Δ_{IPD} .

precisely in line with the Jeffress picture. However, Brand and coworkers found that the firing rate functions of these neurons all looked schematically like the solid curve shown in Fig. 3. While this curve has the general oscillatory form expected from the Jeffress picture, the maxima for all neurons was shifted substantially away from $\Delta_{IPD} = 0$. In fact, the maxima were shifted outside the physiologically possible range. The shaded region near the origin shows the possible values of Δ_{IPD} . This phase shift is proportional to the difference in arrival times for a sound at the two ears, and this difference is limited by the size of the head (the separation of the ears). Hence there is a maximum possible value of Δ_{IPD} , and this determines the width of the shaded region. The results of Brand and coworkers thus strongly suggest that the gerbil does not use the neuron with the maximum firing rate to determine the sound direction.

Interestingly, studies of other mammals (McAlpine et al., 2001) appear to find firing rate curves similar to those reported for gerbils, with the maximum shifted systematically away

from $\Delta_{IPD} = 0$. We believe that this is evidence that nature is not using the location of maximum firing rate to localize the sound, but is instead using a ‘reference point’ on the firing rate curve where the slope is high.

VI. RATE VERSUS TIME CODING

In a sense, the model proposed by Jeffress is incomplete, since it does not consider specifically how the outputs of the coincidence neurons N1, N2, ... are used to determine which neuron exhibits the largest firing rate. One can certainly imagine neural networks which accomplish this computation. Likewise, it is possible to conceive of neural networks that compute the location of a neuron at another spot on the firing rate curve, such as the one where the firing rate is half its maximum value. In considering the construction of such neural networks, the issue of information coding becomes important.

There have been many discussions in the literature about how information is coded in a time sequence of action potentials (Dayan and Abbott, 2001), and two main classes of codes have been discussed. Loosely speaking, in rate codes the information is contained (solely or mainly) in the neuronal firing rates, while in time codes the relative timing of action potentials carries a substantial amount of information. The neural circuit that processes the outputs of the coincidence detector neurons N1, N2, ... are generally assumed to process only the firing rates, and hence utilize rate coding. In contrast, the information processing that takes place in the neural circuit in Fig. 1 relies heavily on time coding.

An interesting possibility, which has not been considered in the literature, would be a neural circuit that processes the outputs of N1, N2, ... using a time coding approach. Such a circuit would use the timing of the action potentials generated at N1, N2, ... to locate a particular place on the firing rate curve. Such neural circuits seem quite feasible.

Methods

The model employed for MSO neurons followed the closely work of Rothman et al. (1993). The membrane currents were due to a slow K conductance, a fast K conductance, an inactivating Na conductance, and a leak conductance. The parameters describing these conductances were the same as given by Rothman, et al. (1993) at 38°C. The MSO neuron

was taken to be a single compartment with these ion channel conductances, with radius $r = 30 \mu\text{m}$ and length $\ell = 30 \mu\text{m}$, along with two passive dendrites which contained the same leak conductance as the MSO neurons, but no active conductances. The dendrites had $r = 1 \mu\text{m}$ and $\ell = 150 \mu\text{m}$, which is approximately half of the electrotonic length. The input spikes to the MSO were phase locked to a 250 Hz signal, and drove synapses described by an α function (Koch, 2002) with a maximum synaptic conductance of $0.5 \mu\text{S}$ and a time constant of 0.2 ms. These synapses were located on the dendrites a distance of $60 \mu\text{m}$ from the soma. With these parameters, a single synaptic EPSP was not able to generate an action potential at the soma, but EPSPs on both of the dendrites could do so provided that their spacing in time was not too large. The EPSPs were each shifted randomly in time with a jitter distributed in a gaussian manner with a standard deviation of 0.1 ms.

Acknowledgments

We thank C. Sahley, B. Burrell, and B. Moss for helpful suggestions. A comment by D. D. Nolte provided an important impetus for this work.

References

- Agmon-Snir, H., C. E. Carr, and J. Rinzel. 1998. The role of dendrites in auditory coincidence detection. *Nature* 393:268-272.
- Brand, A., O. Behrend, T. Marquardt, D. McAlpine, and B. Grothe, 2002. Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417:543-547.
- Carr, C. E. 1993. Processing of temporal information in the brain. *Annu. Rev. Neurosci.* 16:223-243.
- Carr, C. E., and M. Konishi. 1988. Axonal delay lines for time measurement in the owl's brainstem. *Proc. Natl. Acad. Sci.* 85:8311-8315.
- Dayan, P., and L. F. Abbott. 2001. Theoretical Neuroscience. MIT Press, Cambridge, MA.
- Gerstner, W., R. Kempter, J. L. van Hemmen, and H. Wagner. 1996. A neuronal learning rule for sum-millisecond temporal coding. *Nature* 383:76-78.
- Grothe, B. 2000. The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. *Prog. in Neurobiol.* 61:581-610.

- Jeffress, L. A. A place theory of sound localization. 1948. *J. Comp. Physiol. Psychol.* 41:35-39.
- Joris, P. X., L. H. Carney, P. H. Smith, and T. C. T. Yin. 1994. Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J. Neurophysiol.* 71:1022-1036.
- Koch, C. 1999. Biophysics of computation. Oxford University Press, New York.
- McAlpine, D., D. Jiang, and A. R. Palmer. 2001. A neural code for low-frequency sound localization in mammals. *Nature Neurosci.* 4:396-401.
- Rayleigh, Lord. 1907. On our perception of sound direction. *Phil. Mag.* 13:214-233.
- Rothman, J. S., E. D. Young, and P. B. Manis. 1993. Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: implications of a computational model. *J. Neurophysiol.* 70:2562-2583.
- Skottun, B. C. 1998. Sound localization and neurons. *Nature* 393:531-531.
- Softky, W. 1994. Sub-millisecond coincidence detection in active dendritic trees. *Neuroscience* 58:13-41.
- Zimon, J. Z., C. E. Carr, and S. A. Shamma. 1999. A dendritic model of coincidence detection in the avian brainstem. *Neurocomputing* 26-26:263-269.