

# Human interaural time difference thresholds for sine tones: The high-frequency limit

Andrew Brughera

Department of Biomedical Engineering, Center for Hearing Research, Boston University, Boston,  
Massachusetts 02115

Larisa Dunai

Departamento de Ingeniería Gráfica Universitat Politècnica de València, Camino de Vera,  
46022 València, Spain

William M. Hartmann<sup>a)</sup>

Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

(Received 4 December 2011; revised 8 February 2013; accepted 20 February 2013)

The smallest detectable interaural time difference (ITD) for sine tones was measured for four human listeners to determine the dependence on tone frequency. At low frequencies, 250–700 Hz, threshold ITDs were approximately inversely proportional to tone frequency. At mid-frequencies, 700–1000 Hz, threshold ITDs were smallest. At high frequencies, above 1000 Hz, thresholds increased faster than exponentially with increasing frequency becoming unmeasurably high just above 1400 Hz. A model for ITD detection began with a biophysically based computational model for a medial superior olive (MSO) neuron that produced robust ITD responses up to 1000 Hz, and demonstrated a dramatic reduction in ITD-dependence from 1000 to 1500 Hz. Rate-ITD functions from the MSO model became inputs to binaural display models—both place based and rate-difference based. A place-based, centroid model with a rigid internal threshold reproduced almost all features of the human data. A signal-detection version of this model reproduced the high-frequency divergence but badly underestimated low-frequency thresholds. A rate-difference model incorporating fast contralateral inhibition reproduced the major features of the human threshold data except for the divergence. A combined, hybrid model could reproduce all the threshold data.

© 2013 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4795778>]

PACS number(s): 43.66.Ba, 43.66.Pn, 43.66.Qp [RYL]

Pages: 2839–2855

## I. INTRODUCTION

It is a well-known fact of binaural hearing that human listeners are not able to detect interaural time differences (ITDs) in sine tones with frequencies greater than about 1500 Hz. The standard reference to this fact is an article by [Zwislocki and Feldman \(1956\)](#). That article reported ITD threshold measurements for three listeners at octave sine-tone frequencies, 250, 500, and 1000 Hz, and also at the highest frequency at which data could be obtained, approximately 1300 Hz. The results agreed well with data reported by [Klumpp and Eady \(1956\)](#), which showed an ITD threshold of 24  $\mu$ s at 1300 Hz, but unmeasurably high thresholds at 1500 Hz. Both of these articles found that the lowest threshold ITD, about 10  $\mu$ s, occurred at 1000 Hz, though the frequency resolution of the experiments was coarse.

These classic measurements from the mid 1950s offer an intriguing view of the human binaural system. The best performance occurred at 1000 Hz, but when the frequency was increased beyond 1300 Hz, the task became impossible. Thus, performance went from best to impossible in the span of less than half an octave. We are aware of no aspect of

human hearing that shows a more dramatic dependence on frequency.

The purpose of the experiments reported in this article was to investigate the high-frequency dependence of ITD detection in detail, using sine tones with a fine mesh of frequencies to trace out the high-frequency dependence. We then compare the experimental ITD thresholds with the predictions of two types of binaural processing models: a lateralization centroid model and a rate-difference model. The centroid model is a particular example of a place model ([Jeffress, 1948](#)); the rate-difference model minimizes the role of place encoding and relies on the difference in firing rates in left and right sides of the binaural system ([McAlpine et al., 2001](#)). Both models assume that the critical binaural interaction occurs in coincidence detector cells, embodied in mammals as the principal neurons of the medial superior olive (MSO) ([Goldberg and Brown, 1969](#); [Yin and Chan, 1990](#)). The ability of MSO neurons to preserve ITD information as a function of frequency is computationally simulated in a biophysically based model, and used as a basis to calculate ITD thresholds in the centroid and rate-difference models. Relevant methods and results are described separately in each experiment/model section, and the unified discussion and conclusions follow at the end.

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: [hartmann@pa.msu.edu](mailto:hartmann@pa.msu.edu)

## II. ITD DISCRIMINATION EXPERIMENTS

The experiments measured sensitivity to ITD in the fine-structure for sine tones having identical amplitudes in the two ears. The measurements depended on the fact that a change in ITD ( $\Delta$ ITD) causes the perceived image of the sound to move from right to left or conversely.

### A. Methods

As in the experiments by Zwislocki and Feldman (ZF), the method was two-interval forced choice. The listener was required to say whether the tone on the second interval appeared to be to the left or the right of the tone on the first. The stimuli in left and right ears had synchronous onsets and offsets so that the tones differed only in fine structure delay.

#### 1. Stimuli

The tones were 500 ms in duration including a 100-ms rise duration and a 100-ms fall. A gap of 400 ms separated the two tones of a trial. The long rise time was intended to prevent the onset (identical for the two ears) from affecting lateralization judgments. According to Rakerd and Hartmann (1986), 100 ms is adequate. Tones were presented to the listener via Sennheiser (Wedemark, Hanover, Germany) HD-410 headphones at a level of 70 dB SPL (sound pressure level)—the same in both ears, a level that ZF found to be about optimum for high frequencies. Although interaural differences in time and level both vary in realistic environments, the lowest ITD thresholds occur when there is no interaural level difference (ILD) (Domnitz and Colburn, 1977), and our goal was to obtain the lowest possible thresholds. Levels were measured with an A-weighted sound level meter and a flat-plate coupler. The listener was seated in a double-walled sound attenuating room and made responses by pressing buttons on a response box.

Our stimuli differed from those of ZF in the application of ITDs. In the ZF experiments, the first tone of the pair always had an ITD of zero, and the second tone had an ITD that was either positive or negative. In our experiments, like those of Haftner *et al.* (1979) and Henning (1983), the application of ITDs was symmetrical about zero. For example, in a typical *right-left* trial the tone led in the right ear by 10  $\mu$ s (ITD = 10) during the first interval, and led in the left ear by 10  $\mu$ s (ITD = -10) during the second interval. The magnitude of the difference in ITD values was then  $\Delta$ ITD = 20  $\mu$ s. Because listeners make their decisions based on the difference between the two intervals, all the data in this article will be presented in terms of  $\Delta$ ITD.

The advantage of a symmetrical experimental method is that larger  $\Delta$ ITD changes can be presented. Logically, the largest ITD that can be present in any tone,  $ITD_{max}$ , is somewhat less than half a period of the tone,  $T/2$ . Practically,  $ITD_{max}$  is equal to the so called “reversal point” identified by Sayers (1964). At the reversal point, occurring at an ITD of  $T/3$  or less, the perceived lateral position averaged over many trials is a maximum. As the ITD increases beyond the reversal point, the lateral image moves back toward the center. In an asymmetrical experiment, trials are limited to a

$\Delta$ ITD of  $ITD_{max}$ . In a symmetrical experiment, such as ours, the left can lead by  $ITD_{max}$  on one interval and the right can lead by  $ITD_{max}$  on the other interval, permitting a trial to access  $\Delta$ ITDs as large as  $2ITD_{max}$ .

Our tones originated in a Tucker-Davis (Alachua, FL) DD1 digital to analog converter, running at a sample rate of 100 kps (kilosamples/s) in each channel. A delay of a single sample would correspond to an ITD increment of 10  $\mu$ s—too large for a careful experiment. Therefore, the stimuli were recomputed prior to every trial by a Tucker-Davis AP2 array processor, controlled by an experiment program written in C. That procedure permitted our hardware to present arbitrarily small ITD values. Tones were lowpass filtered at 20 kHz by the two channels of a lowpass filter, -115 dB/octave. The interaural phase shift attributable to small differences between the two filters was reduced to a negligible value by making the sample rate so high that the filter cutoff frequency could be well above the tone frequencies.

The experiment used a three-down, one-up adaptive staircase procedure (Levitt, 1971), estimating the 79.4% correct point on a psychometric function. After three correct responses, the experiment  $\Delta$ ITD was decreased by the increment. After one wrong response, the  $\Delta$ ITD was increased by the increment. The increment itself was 17  $\mu$ s for the first four turnarounds. Thereafter, the increment was 5  $\mu$ s. However, if the experiment  $\Delta$ ITD became less than 11  $\mu$ s, the increment was reduced to 2  $\mu$ s. The minimum allowed  $\Delta$ ITD was 1  $\mu$ s. The starting value of the  $\Delta$ ITD for a run was set to various values from 100 to 500  $\mu$ s, depending on the listener and the frequency of the tone. The trials of an experimental run continued until the staircase had made fourteen turnarounds. The first four turnaround values of  $\Delta$ ITD were discarded, and the remaining ten were averaged to obtain a threshold for the run. Runs continued over the months of experimenting until it appeared that stable performance had been reached. The final result for the threshold at any given frequency was the mean of the thresholds for the final five runs and the standard deviation ( $N - 1 = 4$  weight). Although feedback was given by pilot lamps on the response box in early runs, no feedback was given on the final five runs.

Final results were accepted as reliable if each of the five runs converged. Runs were classified as convergent if the run threshold  $\Delta$ ITD was less than the starting value. Runs were classified as divergent if the staircase turnaround values tended to increase monotonically as the run progressed. Spot checks were run to ensure that final thresholds did not depend on the starting value of  $\Delta$ ITD.

The frequency dependence of the threshold  $\Delta$ ITD was explored with high resolution. Low test frequencies were 250, 500, and 700 Hz. Mid test frequencies were 700, 800, 900, and 1000 Hz—closely spaced to try to find the minimum threshold and its frequency. High test frequencies were 1200, 1250, 1300—separated by only 50 Hz to obtain a precise estimate of the highest frequency at which threshold  $\Delta$ ITDs could be measured, and to trace out the functional dependence of the approach to that limit. The different frequencies were tested in haphazard order except that more runs were done for frequencies of greater intralister variability.

## 2. Listeners

There were five listeners in the experiment. Listener L1 was female, the second author, age 31. Listeners L2–L5 were male undergraduates between the ages of 18 and 22. Listeners had normal audiograms from 250 to 8000 Hz, as measured with the Békésy tracking technique. Their pure-tone detection thresholds were within 15 dB of nominal in both ears.

## B. Results

### 1. Most sensitive listeners

Listeners were not all equally sensitive to changes in ITD. Listeners L1 and L2 were the most sensitive. They had thresholds with a well-defined minimum as a function of frequency, not much larger than  $10 \mu\text{s}$ , and measurable thresholds at 1400 Hz. Their results are shown in Figs. 1(a) and 1(b), and thresholds from previous experiments using filled symbols for comparison are also shown.

**Listener L1:** Listener L1 had more practice than any of the others, having completed 71 runs prior to data collection. The minimum threshold,  $\Delta\text{ITD} = 10.8 \mu\text{s}$ , occurred at 1000 Hz. The highest frequency for which staircases converged was 1400 Hz, where the threshold was  $\Delta\text{ITD} = 133 \mu\text{s}$ .

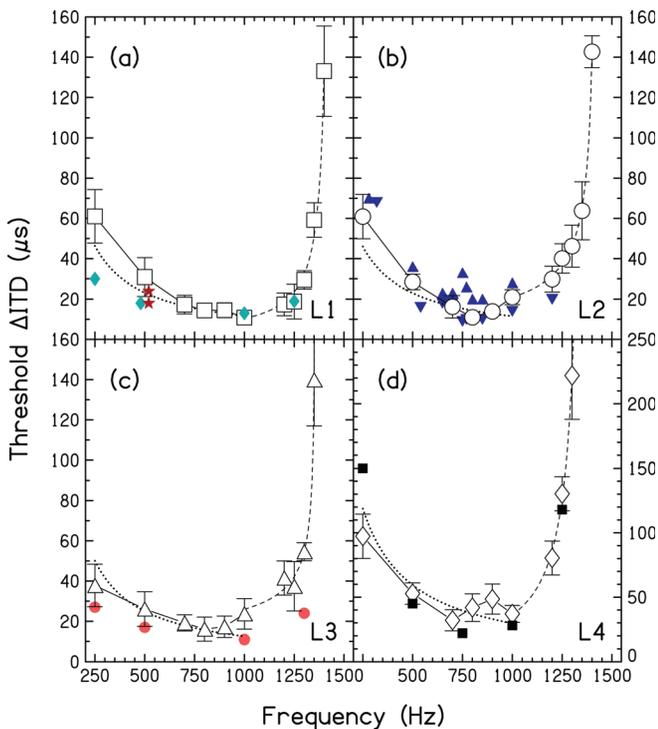


FIG. 1.  $\Delta\text{ITD}$  thresholds for four listeners, L1–L4, are shown by open symbols. Between 1200 and 1400 Hz, data points are separated by 50 Hz. Error bars are two standard deviations in overall length. The dotted line is the maximum-likelihood fit to a  $1/f$  law. The dashed line is the maximum-likelihood fit to the form  $d/(f_c - f)^k$ . The vertical axis scale is enlarged for L4 compared to the other three panels. Data from previous articles are shown by filled symbols: cyan diamonds for Zwislocki-Feldman (1956); red stars for Hershkowitz and Durlach (1969) and Domnitz (1973); orange circles for Klumpp and Eady (1956); blue triangles for two listeners from Dye (1990); black squares for Henning (1983).

Converging and diverging staircases are described in Appendix A.

**Listener L2:** Listener L2 completed 26 runs prior to data acquisition. The minimum threshold  $\Delta\text{ITD}$  of  $11.0 \mu\text{s}$  occurred at 800 Hz. The highest frequency for which listener L2 could obtain a threshold was 1400 Hz, a  $\Delta\text{ITD}$  value of  $141 \mu\text{s}$ .

The high-frequency thresholds for listeners L1 and L2 were remarkably similar. For 1400 Hz, thresholds were 133 and  $141 \mu\text{s}$ , respectively (approximately 70 deg of phase shift), and all runs converged. For 1450 Hz, thresholds were 473 and  $508 \mu\text{s}$ , and all runs diverged. The difference in performance between 1400 and 1450 Hz was remarkable. At 1400 Hz, runs not only converged, but every staircase turnaround was less than the starting value for these listeners. By contrast, at 1450 Hz and 1500 Hz, every staircase was highly divergent by the measure of Appendix A. At 1550 Hz and above, the threshold was unmeasurably high for both listeners, as the value of  $\Delta\text{ITD}$  approached a complete period of the tone before 14 reversals had occurred.

### 2. Less-sensitive listeners

Listeners L3, L4, and L5 did not match the performance of listeners L1 and L2. Runs continued to completion for listeners L3 and L4.

**Listener L3:** Listener L3 completed 85 runs, 50 of which were used to obtain final data at the ten standard frequencies between 250 and 1350 Hz inclusive. Eight runs at 1400 Hz did not converge, nor did runs at 1450 and 1500 Hz. Several attempts to train L3 at 1400 Hz using feedback were unsuccessful. As shown in Fig. 1(c), the lowest threshold occurred at 800 Hz,  $\Delta\text{ITD} = 16.1 \mu\text{s}$ .

**Listener L4:** Listener L4 completed 64 runs, 45 of which were used for final data at the nine standard frequencies between 250 and 1300 Hz, inclusive. His lowest thresholds were  $32 \mu\text{s}$  and  $36 \mu\text{s}$  at 700 and 1000 Hz, respectively, about a factor of 3 larger than for listener L1. Thresholds for L4, shown in Fig. 1(d), have the same general shape as those for L1 and L2. Staircases for L4 did not obtain thresholds at 1350 and 1400 Hz consistently across successive runs.

**Listener L5:** Listener L5 was extensively tested. In 61 runs at the important frequencies of 700, 800, 900, and 1000 Hz, his thresholds were generally higher than those for the other listeners. The means of the final five runs at each of those frequencies were 36, 73, 83, and  $45 \mu\text{s}$ , respectively, roughly a factor of 4 or 5 larger than for listener L1. This listener may fall into the class of listeners who are relatively insensitive to ITD compared to ILD as described in studies summarized by McFadden *et al.* (1973).

## C. Analysis

Figure 1 shows thresholds from previous experiments using filled symbols. They include measurements by Zwislocki and Feldman (1956), Klumpp and Eady (1956),

Dye (1990), and Henning (1983). Hershkowitz and Durlach (1969) and Domnitz (1973) reported their data in a way that makes their thresholds a factor of 2 smaller than thresholds using our definition. Therefore, their thresholds were multiplied by a factor of 2 before plotting in Fig. 1(a).

Where comparison is possible, previously measured thresholds are usually similar to ours. Some differences may be attributable to different methods. For example, Hershkowitz and Durlach (1969) and Domnitz (1973) used a method of constant stimuli and found the 75%-correct points on the psychometric functions. Our staircase technique estimated the 79.4%-correct on a psychometric function, a somewhat more demanding criterion. Detailed discussion follows according to frequency range, low, medium, and high.

### 1. Low-frequency theory

Previous experiments have found that as the frequency is decreased below 500 Hz, the measured functional dependence of the threshold ITD approximately follows a  $1/f$  law (e.g., Yost, 1974). This law corresponds to a constant threshold phase shift,  $\Delta\phi = f \cdot \Delta\text{ITD}$ , but that does not necessarily indicate a special role for interaural phase in neurophysiological computation. Constant  $\Delta\phi$  is an expected behavior of a computation based on interaural time delay because the characteristic time scale for an excitation pattern on a lag axis is the stimulus period  $T$ . Relative changes in the excitation pattern caused by introducing an ITD ( $\Delta t$ ) scale as  $\Delta t/T$ , i.e., as  $f\Delta t$ . For instance, the zero-lag cross-correlation, for two sines with a relative delay of  $\Delta t$  is  $\cos(2\pi\Delta t/T)$ . At low frequency where the excitation pattern is broad on an internal ITD axis, a recognizable change needs to be proportionately broad. Other pattern comparison processes, such as threshold-crossing detectors, depend on the local slope of the excitation pattern. For these too, constant performance can be expected for constant values of  $\Delta t/T$ , i.e., for constant  $\Delta\phi$ . Skottun *et al.* (2001) made similar arguments for point processes in the range where variability is frequency independent.

Formally, the low-frequency region can be usefully defined as the region where neural synchrony at binaural comparison centers, modeled as cross-correlators, is high and approximately independent of frequency. Then the theoretical predictions for the frequency dependence of  $\Delta\text{ITD}$  depend on the distribution of the binaural centers as a function of their interaural time lag (best delay), and on the form of the binaural display. Therefore, the low-frequency thresholds are of more than passing interest.

### 2. Low-frequency experiments

Our experiment did not explore the low-frequency range as thoroughly as the high-frequency range. In order to maximize the amount of data bearing on low-frequency questions, our analysis took the low-frequency range to be the range where  $\Delta\text{ITD}$  threshold decreased with increasing frequency.

The frequency range for each listener is shown in Table I. A maximum-likelihood fit to the equation  $\Delta\phi/f$  is

shown by the dotted curves in Fig. 1. In this procedure, parameter  $\Delta\phi$  was varied to minimize the error,  $E$ ,

$$E = \sum_{i=1}^N (\Delta\text{ITD}_i - \Delta\phi/f_i)^2 / \sigma_i^2, \quad (1)$$

where  $N$  is the number of values of measured  $\Delta\text{ITD}$  having variances,  $\sigma_i^2$ . Constant parameter  $\Delta\phi$  is the corresponding  $\Delta\text{IPD}$  in units of cycles. In the last column of Table I, that  $\Delta\text{IPD}$  is converted to degrees by multiplying by 360.

A comparison between the actual data points and the dotted lines shows that thresholds for listeners L1 and L2 decrease more rapidly with increasing frequency than the  $1/f$  law, but thresholds for listeners L3 and L4 decrease less rapidly than  $1/f$ .<sup>1</sup> Because of an unusually low threshold at 250 Hz, the low-frequency threshold function for L3 was particularly flat compared to that of other listeners.

Alternatives to the  $1/f$  law can be identified by the slope of a log-log plot, for which the slope for a  $1/f$  law is  $-1$ . Table I shows that the slopes are steeper than  $-1$  for L1 and L2, and less steep for the other listeners.

Table I also shows slopes obtained by previous experimenters. Most of the data in the table indicate slopes shallower than  $-1$ , in contrast to our most sensitive listeners L1 and L2—especially L2 with a slope of  $-1.74$ . It is possible that our experiment focused so intently on the high-frequency region that listeners were in some way unprepared for low-frequency runs. To check this suspicion, we reran listener L2 using the same protocol except that the frequencies were limited to 250, 500, 700, and 800 Hz. In that experiment the slope became  $-1.04$ , much less steep than before. It is evident that a better experiment than ours, using more and lower frequencies, would be needed to make definitive statements about deviations from the  $1/f$  law.

### 3. Mid-frequency minima

Because much experimental and theoretical research in binaural hearing has been done with 500-Hz tones, it is worth noting that the minimum,  $\Delta\text{ITD}$  thresholds occur at higher frequencies – 1000 Hz for L1 and 800 Hz for L2. For both listeners, these minimum thresholds were 11  $\mu\text{s}$ . According to a one-tailed  $t$ -test, these minima are

TABLE I. Maximum-likelihood low-frequency slopes and  $\Delta\text{IPD}$  difference limens,  $\Delta\phi$ , for four listeners in this experiment compared with those from previous human studies (Zwislocki and Feldman, 1956; Ricard and Hafter, 1973; Nordmark, 1976). Data from Shackleton *et al.* (2003) describe guinea pig IC recordings.

| Subject | Frequency range (Hz) | Slope | $\Delta\text{IPD}$ (deg) |
|---------|----------------------|-------|--------------------------|
| L1      | 250–1000             | –1.33 | 4.2                      |
| L2      | 250–800              | –1.74 | 4.2                      |
| L3      | 250–800              | –0.78 | 4.5                      |
| L4      | 250–1000             | –0.56 | 10.7                     |
| ZF      | 250–1000             | –0.61 | 3.1                      |
| RH      | 250–1000             | –0.79 | 4.3                      |
| NORD    | 100–400              | –0.90 | 1.2                      |
| SHAK    | 50–850               | –0.98 | 15.0                     |

significantly less than the corresponding thresholds at 500 Hz ( $p < 0.005$ ). A minimum occurred at 800 Hz for Listener L3, lower than thresholds at 500 Hz ( $p < 0.05$ ). Minima occurred at both 700 and 1000 Hz for L4, lower than thresholds at 500 Hz ( $p < 0.025$ ).

Both Zwislocki and Feldman (1956), and Klumpp and Eady (1956) reported thresholds to be lower at 1000 Hz than at 500 Hz, as did Henning (1983) and Dye (1990). Data from Nordmark (1976) appear to show a minimum threshold near 700 Hz, though the value of the minimum itself, about 6  $\mu$ s, is atypically low and its definition seems to be unclear.

#### 4. High-frequency experiments

The high-frequency regime is of special interest in this article. Mills (1958) identified 1400 Hz as the upper limit for effective ITDs. Experiments with two listeners at 65 phons by Nordmark (1976) found that the threshold for  $\Delta$ ITD increased rapidly between the frequencies of 1200 and 1400 Hz. According to Nordmark, “Neither subject could make any discrimination based on phase for frequencies above 1430 Hz.” The data for our most sensitive listeners, L1 and L2—converging staircases at 1400 Hz and diverging staircases at 1450 Hz—agree remarkably well with the conclusions of Mills and Nordmark. Given the agreement among the different experiments and the analysis of divergence from Appendix A, it may not overstate the precision to say that the highest frequency for human ITD discrimination is near 1400 Hz. It is not 1300 Hz, and it is not 1500 Hz.

Attempts were made to fit the dependence of our measured thresholds as a function of frequency. It was found that thresholds grew faster than exponentially for listeners L1, L2, and L4. Growth was also faster than exponential for listener L3 when the anomalous points at 1200 and 1250 Hz were averaged and plotted at 1225 Hz. Because of the rapid growth, and because no finite ITD threshold could be found at higher frequency, we fitted our data with a non-analytic function typical of critical phenomena

$$\Delta\text{ITD}(f) = d/(f_c - f)^\kappa, \quad (2)$$

where  $f_c$  is the critical frequency,  $\kappa$  is the critical exponent, and  $d$  is a third fitting parameter.

The fitting procedure minimized a weighted least squares discrepancy between the formula for  $\Delta$ ITD and all the measurable thresholds at 1000 Hz and above. The fitted functions are shown by the dashed curves in Fig. 1.

### III. NEURAL MODEL

In the remainder of this article, we will compare the experimental ITD thresholds with the predictions of models of the binaural system. We consider two kinds of models, a lateralization centroid model and a rate-difference model. This section presents a computational simulation of the ability of MSO neurons to preserve ITD information as a function of frequency, which is used in Secs. IV and V as a basis to calculate ITD thresholds in the centroid and rate-difference models.

## A. Neural modeling methods

The model MSO neuron in the computational simulation is defined by its cell and membrane parameters, its input and synaptic parameters, and the acoustic stimuli assumed to drive its inputs. Basic data analysis methods are also defined, and parameter values of the model MSO neuron are provided in Tables II and III.

### 1. Cell model

The model MSO neuron is based on existing multi-compartment Hodgkin-Huxley models for a principal MSO neuron and its ion-channel dynamics (Zhou *et al.*, 2005; Scott *et al.*, 2010; Mathews *et al.*, 2010; Fischl *et al.*, 2012) with modifications from the Zhou *et al.* (2005) model described below. The four cylindrical compartments of the model comprise a contralateral dendrite, ipsilateral dendrite, soma, and axon. The two dendrites connect to opposite ends of the soma, and in this study the axon connects to the ipsilateral half of the soma (at 75% of the distance along the soma from its contralateral to ipsilateral end). While the model is used to estimate human ITD thresholds, it is based on the physiology of ITD-sensitive MSO neurons in the Mongolian gerbil. To be conservative in predicting a loss in ITD sensitivity at higher frequencies, model parameters were chosen to match or be slightly faster than those measured in the gerbil.

TABLE II. Parameter values of the model MSO neuron in each compartment. (n/a is non-applicable.)

| Parameter (unit)                         | Dendrites (2) | Soma         | Axon    |
|--|---------------|--------------|---------|
| Temperature ( $^{\circ}$ C)              | 37            | 37           | 37      |
| Number of segments, nseg                 | 20            | 2            | 51      |
| Diameter ( $\mu$ m)                      | 3.5           | 20           | 2       |
| Length ( $\mu$ m)                        | 150           | 40           | 400     |
| Resistivity, $R_a$ (ohm $\times$ cm)     | 150           | 150          | 150     |
| $C_M$ ( $\mu$ F/cm $^2$ )                | 1             | 1            | 1       |
| $E_K$ (mV)                               | -106          | -106         | -106    |
| $E_{Na}$ (mV)                            | n/a           | 62.1         | 62.1    |
| $E_h$ (mV)                               | -43           | -43          | -43     |
| $E_{PAS}$ (mV)                           | -60           | n/a          | n/a     |
| $E_{LeakNa}$ (mV)                        | n/a           | -60          | -65     |
| $G_{maxK_{LT}}$ (S/cm $^2$ )             | 0.0022        | 0.054        | 0.0595  |
| $G_{maxNa}$ (S/cm $^2$ )                 | n/a           | 0.072        | 0.25    |
| $G_{maxh}$ (S/cm $^2$ )                  | 0.0011        | 0.0216       | 0.0025  |
| $G_{PAS}$ (S/cm $^2$ )                   | 0.00005       | n/a          | n/a     |
| $G_{LeakNa}$ (S/cm $^2$ )                | n/a           | 0.0004       | 0.00005 |
| $E_E$ (mV)                               | 0             | n/a          | n/a     |
| $\zeta_E$ (nS)                           | 18–220        | n/a          | n/a     |
| $\tau_{E_{rise}}$ (ms)                   | 0.39996       | n/a          | n/a     |
| $\tau_{E_{decay}}$ (ms)                  | 0.4           | n/a          | n/a     |
| $E_I$ (mV)                               | n/a           | -90          | n/a     |
| $\zeta_I$ (nS)                           | n/a           | 30–72, 3–8   | n/a     |
| $\tau_{I_{rise}}$ (ms)                   | n/a           | 0.39996, 0.4 | n/a     |
| $\tau_{I_{decay}}$ (ms)                  | n/a           | 0.4, 2.0     | n/a     |
| $V_{AP-THRESH}$ (mV, set)                | n/a           | n/a          | -20     |
| $V_{REST}$ (mV, measured)                | -60.3         | -60.3        | -64.3   |
| $\tau_M$ (ms, calculated at $V_{REST}$ ) | 0.36          | 0.29         | 0.79    |

TABLE III. Synaptic strengths vs frequency.

| Frequency (Hz) | $\zeta_E$ (nS) | $\zeta_I$ slowly decaying (nS) | $\zeta_I$ rapidly decaying (nS) |
|----------------|----------------|--------------------------------|---------------------------------|
| 250            | 27             | 8                              | 36                              |
| 500            | 18             | 6                              | 30                              |
| 750            | 80             | 8                              | 72                              |
| 1000           | 108            | 5                              | 36                              |
| 1250           | 180            | 3                              | 40                              |
| 1500           | 220            | 4                              | 40                              |

The ion-channel dynamics of the model are characterized using existing equations derived from MSO neurons and ventral cochlear nucleus (VCN) neurons: for sodium (Na), the equations for gerbil MSO neurons in [Scott et al. \(2010\)](#); for the hyperpolarization-activated cation (*h*), the equations for VCN neurons in guinea pig ([Zhou et al., 2005](#); [Rothman and Manis, 2003](#)); and for low-threshold potassium ( $K_{LT}$ ), the equations for gerbil MSO neurons in [Mathews et al. \(2010\)](#) with one modification for a faster inactivation (contained in the dynamics equations received courtesy of [Fischl and colleagues, 2012](#)). The time constant of  $K_{LT}$  inactivation ( $\tau_z$ ) as a function of membrane potential ( $v_M$ ) is given in ms by

$$\tau_z = 10.7 + 170 / \{5 \exp[(v_M + 60)/10] + \exp[(70 - v_M)/8]\}. \quad (3)$$

The time constant values for all gating variables were divided by the  $Q_{10}$  temperature factor of  $3.0^{(T_b - 22)/10}$ , where  $T_b$  equals the human body temperature in Celsius,  $37^\circ\text{C}$ . The net effect on  $\tau_z$  is to be nearly constant at 2.1 ms for  $v_M$  between  $-100$  and  $40$  mV, compared with  $\tau_z$  in [Mathews et al. \(2010\)](#), which increases from 2.1 ms at  $-10$  mV to a maximum of 13.8 ms at  $-72$  mV. While the difference in  $\tau_z$  is large, the limited range of steady-state  $K_{LT}$  inactivation ( $z_\infty$ , ranging from 0.45 at  $-60$  mV to 0.27 at  $40$  mV) decreases the effect of the faster  $\tau_z$ .

The model MSO neuron in this study had minimal leakage currents at rest while maintaining resting membrane potential ( $V_{REST}$ ) and resting membrane time-constant ( $\tau_M$ ) values consistent with physiology ([Scott et al., 2005](#); [Mathews et al., 2010](#)). Reversal potentials of the ion channels were set as in [Mathews et al. \(2010\)](#):  $E_K = -106$  mV,  $E_h = -43$  mV; and in [Scott et al. \(2010\)](#):  $E_{Na} = 62.1$  mV. To produce near-zero leakage currents at rest, the reversal potentials for leakage currents ( $E_{PAS}$  in dendrites,  $E_{LeakNa}$  in soma and axon) were set approximately to  $V_{REST}$  in each compartment. In this condition, at  $V_{REST}$  the outward current  $i_{K_{LT}}$  is offset by inward currents  $i_{Na}$  and  $i_h$ . By choosing the ratio of  $i_{Na}$  to  $i_h$  (0, 0.25, 1 in dendrites, soma, and axon, respectively) and computing the gate variable activations at  $V_{REST}$ , the ratios of the maximum conductance for  $K_{LT}$ , Na, and  $h$  ( $G_{maxK_{LT}}$ ,  $G_{maxNa}$ , and  $G_{maxh}$ ) were found in each compartment. The  $G_{max}$  values were then scaled proportionally for the desired  $\tau_M$  given the membrane capacitance ( $C_M = 1 \mu\text{F}/\text{cm}^2$ ), and in some cases altered slightly to enhance ITD performance. The model  $\tau_M$  values reported

below were calculated based on the actual  $V_{REST}$  values ( $v_m$  after 500 ms in a simulation without inputs). In the dendrites,  $E_{PAS} = -60$  mV,  $V_{REST} = -60.3$  mV, and  $\tau_M = 0.36$  ms. In the soma,  $E_{LeakNa} = -60$  mV,  $V_{REST} = -60.3$  mV, and  $\tau_M = 0.29$  ms. These  $\tau_M$  values are in the faster range of measured somatic  $\tau_M$  values in MSO neurons ([Scott et al., 2005](#); [Scott et al., 2007](#)). In the axon, for both stability and ITD sensitivity,  $E_{LeakNa} = -65$  mV,  $V_{REST} = -64.3$  mV, and  $\tau_M = 0.79$  ms. Action potentials were counted at the midpoint of the axon, and the voltage threshold for counting action potentials ( $V_{AP-THRESH}$ ) was held constant at  $-20$  mV. Simulations were performed in freely available NEURON software ([www.neuron.yale.edu](http://www.neuron.yale.edu); last viewed March 26, 2013), which supports linear space-gradients in  $v_m$ .

## 2. Input model and synapses

The input model consisting of periodic rate functions of Poisson-like processes was the same as in [Zhou et al., \(2005\)](#), but with reduced numbers of inputs to more closely reflect anatomy ([Couchman et al., 2010](#)). The eight excitatory inputs represent bilateral inputs to the MSO neuron, four each from the ipsilateral and contralateral anteroventral cochlear nuclei (AVCN). The four contralaterally-driven inhibitory inputs represent glycinergic inputs to the MSO neuron from the medial nucleus of the trapezoid body ([Smith et al., 2000](#)). Each excitatory input is connected to the dendrite on the same side by an excitatory synapse, at distances to the soma between 40% and 60% of the dendritic length (one synapse at each of 42.5, 47.5, 52.5, and 57.5%). The inhibitory inputs connect to the contralateral half of the soma (at 25% of the distance along the soma from its contralateral to ipsilateral end).

Excitatory parameters are denoted with subscript, *E*, and inhibitory parameters with subscript, *I*. While parameter values differ between excitatory and inhibitory synapses, each synapse is modeled as a variable conductance  $\sigma(t)$  in series with a fixed reversal potential ( $E_E = 0$  mV,  $E_I = -90$  mV). Each synaptic conductance is augmented by a time-varying increment in response to each action potential from its input. This increment,  $\Delta\sigma(t)$ , is the difference of two exponentials, having a faster rise time constant,  $\tau_{rise}$ , and a slower decay time constant,  $\tau_{decay}$  ( $\tau_{decay} > \tau_{rise}$ ), and a peak conductance,  $\zeta$ , at time,  $t_p$ , after the input action potential at time  $t = 0$ ;  $\Delta\sigma(t) = [\zeta/s(t_p)] [\exp(-t/\tau_{decay}) - \exp(-t/\tau_{rise})]$  for  $t \geq 0$ , else  $\Delta\sigma(t) = 0$ . The normalization factor,  $s(t_p)$ , is equal to  $\exp(-t_p/\tau_{decay}) - \exp(-t_p/\tau_{rise})$ , and the time of the peak is given by  $t_p = [\tau_{rise}\tau_{decay}/(\tau_{decay} - \tau_{rise})] \ln(\tau_{decay}/\tau_{rise})$ . Synaptic time constants were set according to measured values from MSO neurons ([Fischl et al., 2012](#); [Magnusson et al., 2005](#)). For excitatory synapses,  $\tau_{E_{rise}} = 0.39996$  ms and  $\tau_{E_{decay}} = 0.4$ , and for slowly decaying inhibition,  $\tau_{I_{rise}} = 0.4$  ms and  $\tau_{I_{decay}} = 2$  ms. In addition, to produce contralateral-leading best-ITDs similar to those measured in [Brand et al. \(2002\)](#), rapidly decaying inhibition was applied in separate simulations using  $\tau_{I_{rise}} = 0.39996$  ms and  $\tau_{I_{decay}} = 0.4$  ms. The amplitudes of conductance increments,  $\zeta_E$  and  $\zeta_I$ , were varied between simulations and held constant within each simulation (dynamic synaptic depression was not included in this model).

For each synapse there is at most one input event per stimulus period, and events are generated by a two-stage process for each period. First, the occurrence (or not) of an input spike is determined with a fixed probability and second, the temporal location,  $t_k$ , within the period is determined. Input patterns are characterized with three parameters: the stimulus period,  $T$ , the average input spike rate to each synapse,  $R_{ave}$ , and the input synchrony index, SI. Parameters  $R_{ave}$  and SI, respectively, control the rate and the temporal aspects of input spike trains. The probability of an input event within a period is equal to the lesser of  $R_{ave}T$  and 1. The temporal location,  $t_k$ , within the period is drawn from a Gaussian distribution with mean,  $T/2$ , and standard deviation,  $T/(2F)$ , where  $F = \pi/\sqrt{2 \ln(1/SI)}$ , i.e., the inverse of the coefficient of variation of the jitter distribution.

### 3. Stimuli and input parameters

Responses of the model MSO neuron to 500-ms tones with ITD were simulated for input frequency ( $f = 1/T$ ) equal to 250, 500, 750, 1000, 1250, and 1500 Hz. There were five repetitions at each ITD, with the reported discharge rates being the mean rate at each ITD; the standard deviation of discharge rate over the five repetitions was also recorded. In all simulations, the sample rate was 120 kHz, with an integer number of samples per ITD step at all applied stimulus frequencies. The ITD-resolution was  $T/20$ , except for the stimulus frequency of 1250 Hz where ITD-resolution was  $T/24$ . The input SI was set to physiologically measured values (Joris *et al.*, 1994) from AVCN projections in the trapezoid body stimulated at their characteristic frequencies (CFs), i.e., SI = 0.93 at 250 Hz, 0.9 at 500 Hz, 0.85 at 750 Hz, 0.8 at 1000 Hz, 0.75 at 1250 Hz, and 0.7 at 1500 Hz. Input rate,  $R_{ave}$ , to each synapse reflects measured spike rates in AVCN projections at each frequency (Joris *et al.*, 1994) such that the input to each model synapse entrains to the stimulus frequency up to 600 Hz, and then saturates at 600 spikes/s for higher frequencies. The input rates and synchrony not only reflect the AVCN responses to acoustic tones at 60 dB SPL and higher, they maintain a reasonably high input rate to the model MSO neuron per stimulus period, allowing the possibility of continued ITD sensitivity as the stimulus frequency increases to 1500 Hz.

### B. Neural modeling results

Figure 2 displays rate-ITD functions of model MSO neurons, showing discharge rates in action-potentials per second (spikes/s) as ITD varied from  $-1$  to  $+1$  ms for stimulus tones at CF from 250 to 1500 Hz.<sup>2</sup> [The frequency legend in Fig. 2(B) applies to all panels.] Each panel of Fig. 2 shows responses of the model MSO in one of three conditions simulating the absence and presence of contralaterally driven glycinergic inhibition: Fig. 2(A), bilateral excitation (EE) only; Fig. 2(B), excitation and slowly decaying inhibition ( $\tau_{I_{rise}} = 0.4$  ms;  $\tau_{I_{decay}} = 2.0$  ms); Fig. 2(C), excitation and rapidly decaying inhibition ( $\tau_{I_{rise}} = 0.39996$  ms;  $\tau_{I_{decay}} = 0.4$  ms).

The excitatory synaptic time constants ( $\tau_{E_{rise}} = 0.39996$  ms;  $\tau_{E_{decay}} = 0.4$  ms) and the excitatory synaptic strength ( $\zeta_E$ ) as a function of frequency were maintained across all three

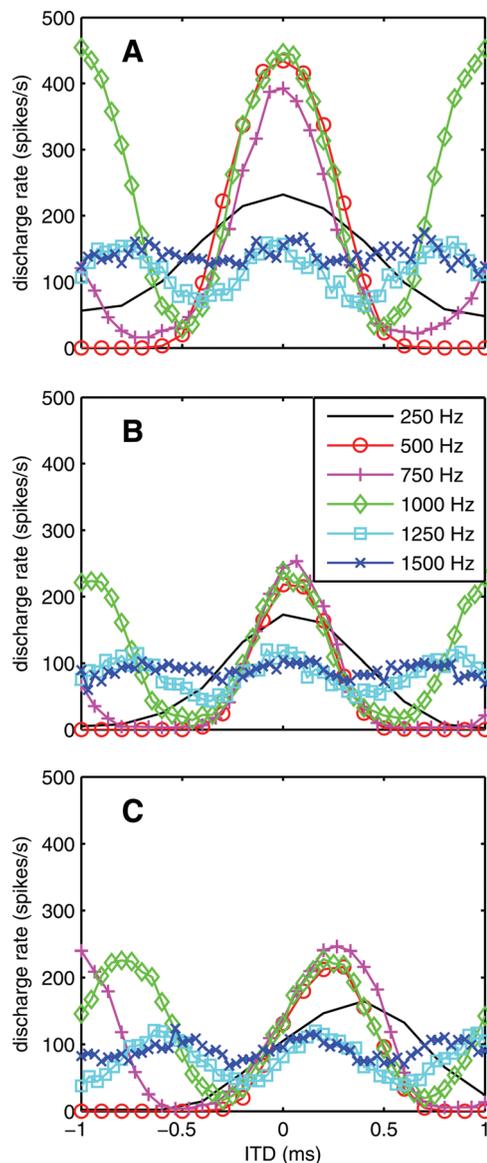


FIG. 2. Discharge rate as a function of ITD in acoustic tones from 250 to 1500 Hz for the model MSO neuron with (A) purely excitatory inputs, (B) excitatory and slowly decaying inhibitory inputs, and (C) excitatory and rapidly decaying inhibitory inputs.

conditions. At each stimulus frequency, a single value of  $\zeta_E$  was selected that produced both an unsaturated rate-ITD function in the purely excitatory condition, and a relatively steep rate-ITD function at zero ITD with a contralateral-leading best-ITD in the rapidly decaying inhibition condition. Inhibitory synaptic strength ( $\zeta_I$ ) at each frequency was adjusted independently for slowly decaying or rapidly decaying inhibition. Synaptic strength values across input frequencies and inhibition-conditions are provided in Table III. The synchrony index and rate of synaptic input events are functions of frequency, given above in Sec. III A 3.

#### 1. Rate-ITD functions

The rate-ITD functions in Fig. 2 served as the inputs to the binaural display models. Key features of these functions include the approximate periodicity corresponding to the stimulus frequency, the maximum and minimum firing rates,

TABLE IV. Modulation indexes ( $m$ ) of fitted rate-ITD functions under selected conditions of excitation and inhibition.

| Frequency (Hz) | EE only | I slowly decaying + EE | I rapidly decaying + EE |
|----------------|---------|------------------------|-------------------------|
| 250            | 0.79    | 1.00                   | 1.00                    |
| 500            | 1.00    | 1.00                   | 1.00                    |
| 750            | 0.92    | 0.98                   | 0.97                    |
| 1000           | 0.88    | 0.89                   | 0.87                    |
| 1250           | 0.43    | 0.47                   | 0.51                    |
| 1500           | 0.22    | 0.27                   | 0.26                    |

and the modulation index,  $m$ , which indicates the sensitivity to ITD

$$m = (\max(\text{rate}) - \min(\text{rate})) / (\max(\text{rate}) + \min(\text{rate})). \quad (4)$$

A fifth key feature is the half-width of the peak, defined as the shortest time difference between crossings of the mean excursion value  $[\max(\text{rate}) + \min(\text{rate})]/2$ . Modulation index and half-width are given in Tables IV and V.

In all conditions, the rate-ITD functions were highly modulated at and below 1000 Hz, and became progressively less modulated with increasing frequency above 1000 Hz. Maximum spike rates decreased markedly between 1000 and 1250 Hz, and minimum spike rates increased steadily between 1000 and 1500 Hz. The rate-ITD functions remained reasonably well-modulated at 1250 Hz, and became relatively flat at 1500 Hz. Compared with the purely excitatory condition, both types of contralateral inhibition reduced overall discharge rates. Slowly decaying inhibition increased the modulation percentage, except at 500 Hz, where  $m$  held steady at its maximum possible value of 1. At 1000 Hz and below, slowly decaying inhibition decreased half-widths by more than 12%, indicating that the sharpness of ITD-tuning was increased by the inhibition; at 1250 Hz, the half-width decreased slightly, and at 1500 Hz, the half-width increased slightly.

Rapidly decaying inhibition, across all input frequencies, shifted the best-ITD from zero (in the purely excitatory condition) to contralateral-leading ITD, such that the steepest slope in each rate-ITD function occurred near the midline (zero ITD). Rapidly decaying inhibition increased the modulation index, except at 500 Hz, where  $m$  held steady at its maximum value of 1, and at 1000 Hz where  $m$  decreased by 1%. Rapidly decaying inhibition also decreased half-widths by more than 12% at 500 Hz and below, but had less effect

TABLE V. Half-widths (in ms) of fitted rate-ITD functions under selected conditions of excitation and inhibition.

| Frequency (Hz) | EE only | I slowly decaying + EE | I rapidly decaying + EE |
|----------------|---------|------------------------|-------------------------|
| 250            | 1.06    | 0.93                   | 0.93                    |
| 500            | 0.58    | 0.47                   | 0.50                    |
| 750            | 0.50    | 0.40                   | 0.54                    |
| 1000           | 0.50    | 0.40                   | 0.49                    |
| 1250           | 0.40    | 0.38                   | 0.39                    |
| 1500           | 0.30    | 0.33                   | 0.33                    |

on ITD tuning at 750 Hz and above. The shift in best-ITD and the sharpening of ITD-tuning were two other key features of the rate-ITD functions incorporated in the binaural display models.

Best-ITDs also tended to be slightly contralateral-leading for slowly decaying inhibition [Fig. 2(B)], which is most easily observable at 750 Hz and below. Although the relatively low electrical impedance of the soma produced nearly identical membrane potentials in its ipsilateral and contralateral halves (Fig. 3), the locations of the axon and the inhibitory synaptic current inputs affected responses to ITD in the model neuron. Where there was no significant inhibitory shift in best-ITD in a symmetrical model neuron with contralaterally driven inhibition applied to the center of the soma (not shown), the observed inhibitory shifts to contralateral-leading best-ITDs were facilitated by the asymmetrical location of the inhibitory synapses (at the contralateral side of the soma) and the axon extending from the ipsilateral side of the soma.

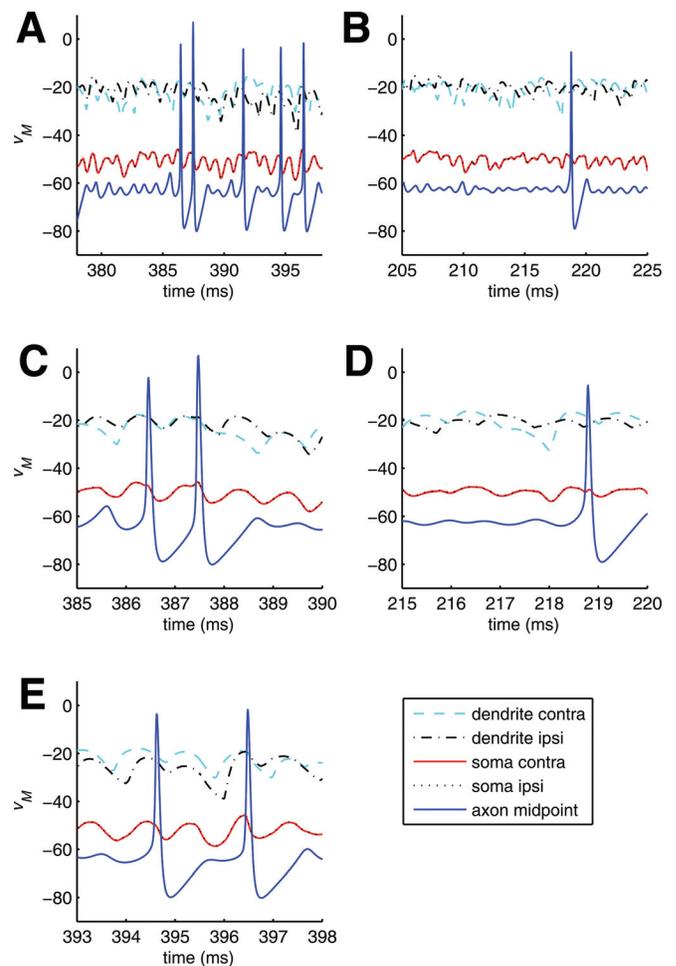


FIG. 3. Membrane potentials,  $v_M$ , as a function of time in the model MSO neuron with rapidly decaying inhibition for a stimulus tone at 1000 Hz:  $v_M$  in the contralateral dendrite (light blue dashed line), ipsilateral dendrite (black dash-dotted line), contralateral (red line) and ipsilateral (black dotted line) halves of the soma (curves overlap), and the axon (dark blue line). (A), (C), (E) ITD = 200  $\mu$ s (best-ITD, bilateral inputs in-phase). (B), (D) ITD = -300  $\mu$ s (bilateral inputs out-of-phase). Axonal action potentials occur frequently for (C), (E) coincident binaural EPSPs, and rarely for (D) a large monaural EPSP.

## 2. Membrane potentials

Membrane potential,  $v_M$ , as a function of time in the model MSO neuron with rapidly decaying inhibition is shown in Fig. 3 for a stimulus tone at 1000 Hz, and in Fig. 4 for separately presented tone stimuli at 1250 and 1500 Hz. Potential  $v_M$  is plotted for the midpoint of the axon (dark blue line); the contralateral half of the soma (red line) and ipsilateral half of the soma (black dotted line), where  $v_M$  is nearly equal and the two curves practically overlap; and the contralateral and ipsilateral dendrites (light blue dashed line, and black dash-dotted line, respectively, each recorded near its excitatory synapses, at 37.5% of the dendritic length from the soma).

In Fig. 3 at 1000 Hz, in Figs. 3(A), 3(C) and 3(E), ITD = 200  $\mu$ s (best-ITD, bilateral inputs in-phase), and in Figs. 3(B) and 3(D), ITD = -300  $\mu$ s (bilateral inputs out-of-phase). Figures 3(A) and 3(B) show 20-ms samples illustrating that the discharge rate in the axon was high for the in-phase condition, and low for the out-of-phase condition. The ratio of average spike rates between in-phase and out-of-phase conditions, equal to 14, was even higher than suggested by the figure—due to the actual out-of-phase spike rate being three times lower than suggested by the single spike in Fig. 3(B). Figures 3(C) and 3(E) show close-ups of excitatory post-synaptic potentials (EPSPs) and resulting axonal action-potentials in the in-phase condition. In Fig. 3(C), the EPSPs in the contralateral and ipsilateral dendrites were of moderate amplitude and highly synchronous. In Fig. 3(E), the bilateral dendritic EPSPs were reasonably synchronous and of high amplitude. As expected, highly synchronous bilateral EPSPs of high amplitude (not shown) were also sufficient to trigger an action potential. Figure 3(D) shows a close-up of EPSPs and an axonal action potential in the out-of-phase condition, where a contralateral dendritic EPSP of high-amplitude was sufficient to trigger the action potential. Large action potentials occurred only in the axon, not in the soma, and the lack of dendritic sodium currents prevented any form of action potential in the dendrites. During some axonal action potentials, such as that of Fig. 3(D) and the second of Fig. 3(C), there was a slight momentary increase in somatic  $v_M$  that occurred late in the somatic EPSP, during the plateau or downward slope of the EPSP, suggesting back-propagation of the action potential from the axon to the soma.

The bilaterally in-phase conditions that produced action potentials at 1000 Hz continued to do so at 1250 and 1500 Hz, but less frequently, due to the increased shunting effects of more consistently activated  $K_{LT}$  currents at higher frequencies (Colburn *et al.*, 2008). Figures 4(A) and 4(B) show  $v_M$  at 1250 Hz, and Figs. 4(C)–4(E) show  $v_M$  at 1500 Hz. At 1250 Hz, Fig. 4(A) shows the best-ITD in-phase condition at ITD = 160  $\mu$ s, and Fig. 4(B) shows the out-of-phase condition at ITD = -240  $\mu$ s. The ratio of spike rates between the in-phase and out-of-phase conditions fell sharply above 1000 Hz, but at 1250 Hz this ratio remained reasonably high at 3:1. At 1500 Hz, Fig. 4(C) shows the best-ITD in-phase condition at ITD = 133  $\mu$ s, and Fig. 4(D) shows the out-of-phase condition at ITD = -200  $\mu$ s, where

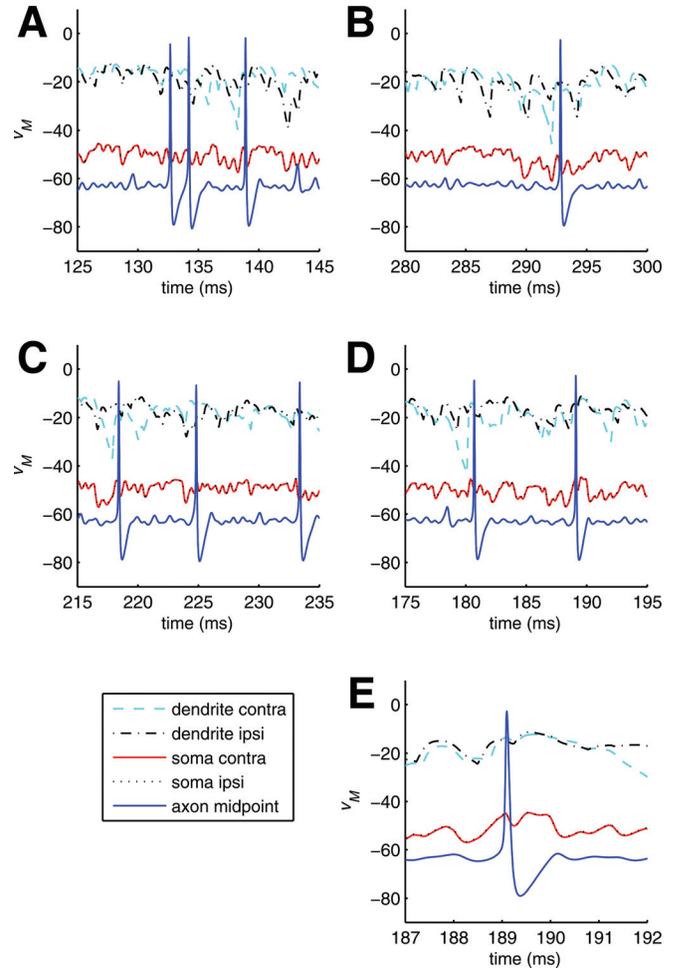


FIG. 4. Membrane potentials,  $v_M$ , as a function of time in the model MSO neuron with rapidly decaying inhibition for a stimulus tone at 1250 Hz and 1500 Hz:  $v_M$  in the contralateral dendrite (light blue dashed line), ipsilateral dendrite (black dashed-dotted line), contralateral (red line), and ipsilateral (black dotted line) halves of the soma (curves overlap), and the axon (dark blue line). As at lower frequencies, action potentials were triggered by in-phase binaural EPSPs at best-ITD, but less frequently. (A) 1250 Hz, bilaterally in-phase at best-ITD = 160  $\mu$ s. (B) 1250 Hz, bilaterally out-of-phase at ITD = -240  $\mu$ s. (C) 1500 Hz, bilaterally in-phase at best ITD = 133  $\mu$ s. (D), (E), 1500 Hz, bilateral out-of-phase at ITD = -200  $\mu$ s, where an axonal action potential was triggered by errantly coincident bilateral EPSPs.

the spike rate approximately doubled (compared with the out-of-phase condition at 1250 Hz), such that ratio of spike rates between in-phase and out-of-phase conditions decreased to approximately 3:2 at 1500 Hz. In the out-of-phase condition at 1500 Hz, action potentials triggered by errantly coincident bilateral EPSPs, such as the spike shown in Fig. 4(E), became more frequent compared with 1250 Hz. A contributing mechanism to this increase in errant binaural coincidences at higher frequencies is the combination of a shorter stimulus period and decreased input synchrony.

The high dendritic voltages (around -20 mV) recorded near the model excitatory synapses were due to the strong synapses at 750 Hz and above. At 500 Hz, with much weaker synapses,  $v_M$  was in the range of -45 mV (not shown). At all frequencies, EPSPs decreased significantly as they traveled through the dendrite toward the soma, such that  $v_M$  in the proximal segment of the dendrite (i.e.,  $v_M$  at 3.75  $\mu$ m or

2.5% of dendritic length from the soma, not shown) was virtually indistinguishable from  $v_M$  in the soma.

#### IV. CENTROID MODEL

It is possible to account for features of the experimental thresholds, including the dramatic lateralization failure at 1450 Hz, with a signal processing theory that extends the Jeffress (1948) model of the binaural system. The theory has two main parts. One part is an array of coincidence cells (MSO cells) in the brainstem operating as cross-correlators, as modeled in Sec. III. The second part is a hypothetical binaural display that is a nexus between the coincidence cells and a spatial representation that is adequate to determine laterality for a listener. The display is imagined to have a wide distribution of best delays, and the distribution depends only weakly on the cell best frequency.

##### A. Centroid lateralization display

The centroid lateralization display was introduced by Stern and Colburn (1978) and applied to the lateralization of 500-Hz tones with interaural time and level differences. It was modified and extended to other frequencies by Stern and Shear (1996). In this display, a sine tone with an ITD of  $\Delta t$  excites brainstem cross-correlator cells represented by a cross-correlation function,  $c(\tau)$ , where  $\tau$  is the lag (or best interaural delay of the cell), and values of  $\tau$  have a wide range, limited only by the density distribution  $p(\tau)$ , centered on  $\tau = 0$ .

The operative measure of laterality is the centroid of the density-weighted cross-correlation,

$$\bar{\tau}(\Delta t) = \frac{\int d\tau p(\tau) \tau c(\tau - \Delta t)}{\int d\tau p(\tau) c(\tau - \Delta t)}, \quad (5)$$

and the integrals are over the range of minus to plus infinity.

Values of  $\bar{\tau}$  were computed from model functions  $c(\tau)$  and  $p(\tau)$ . Function  $c(\tau)$  was an analytic fit to the rate-ITD functions in Fig. 2(A), chosen mainly because it is symmetric about ITD = 0. The fit captured the period, minimum, and maximum, as well as the narrowing of the peaks at low frequencies. Fits to all of the rate-ITD functions, including those in Fig. 2(A), are described in Appendix B. Density function  $p(\tau)$  was a simplified version of the form introduced by Colburn (1977),

$$p(\tau) = C \quad (|\tau| \leq T_o), \quad (6)$$

$$p(\tau) = C \exp[-(|\tau| - T_o)/\tau_o] \quad (|\tau| > T_o), \quad (7)$$

where  $C$  normalizes the integrated density to 1.0. As noted by Stern and Shear (1996), Colburn's  $p(\tau)$  decays too slowly to successfully model the lateralization of tones at high frequency. Therefore, we chose a more rapidly decaying function with  $T_o = 0.2$  ms, and  $\tau_o = 0.22$  ms.

With these choices for  $c(\tau)$  and  $p(\tau)$ , the nature of the calculation in Eq. (5) is that as the tone frequency increases, more and more cycles of  $c(\tau - \Delta t)$  fit within the range of

lags given by  $p(\tau)$ . This has the effect of preventing the centroid from increasing much as  $\Delta t$  increases because of partial cancellation of the positive side-lobes of  $c(\tau - \Delta t)$  by the negative side-lobes. Similar behavior was noted by Stern and Shear (1996) in their calculation of lateralization as a function of frequency in fitting the data of Schiano *et al.* (1986). Because the centroid is the cue to laterality available to the listener, limiting the centroid in this way limits the perceived laterality. That limit could be a key to the failure to discriminate ITDs at 1450 Hz and above.

##### B. Centroid threshold calculation

Values of centroid  $\bar{\tau}$  computed from Eq. (5) using the excitation-only rate-ITD functions from Fig. 2(A) for  $c(\tau)$  are shown in Fig. 5 for seven different tone frequencies. For those frequencies where no MSO model calculations were done, functions were interpolated. Figure 5 leads to predictions for a threshold  $\Delta ITD$  if it is assumed that there is a threshold value of centroid  $\bar{\tau}_T$ . For example, if it is assumed that the centroid threshold is  $\bar{\tau}_T = 9 \mu\text{s}$ , as shown by the dashed horizontal line in Fig. 5, then the model predicts a threshold of  $\Delta ITD = 56.5 \mu\text{s}$  for a frequency of 1250 Hz, as shown by the open circle in Fig. 5. Because of the choice of model parameters, there is no intersection for the  $\bar{\tau}$  function for 1450 Hz, and the  $\Delta ITD$  threshold is found to diverge, consistent with experiment.

Computed thresholds from the centroid model, to be compared with experimental thresholds, were calculated by starting with the excitation-only rate-ITD functions from Fig. 2(A) and varying the modulation,  $m$ , over a range of  $\pm 10\%$ . The range of predicted thresholds is shown by the shaded region in Fig. 6(a), which can be compared with the experimental values of  $\Delta ITD$  for the four listeners. The calculation agrees with the experimental values in four important ways: (i) The threshold increases as the frequency is

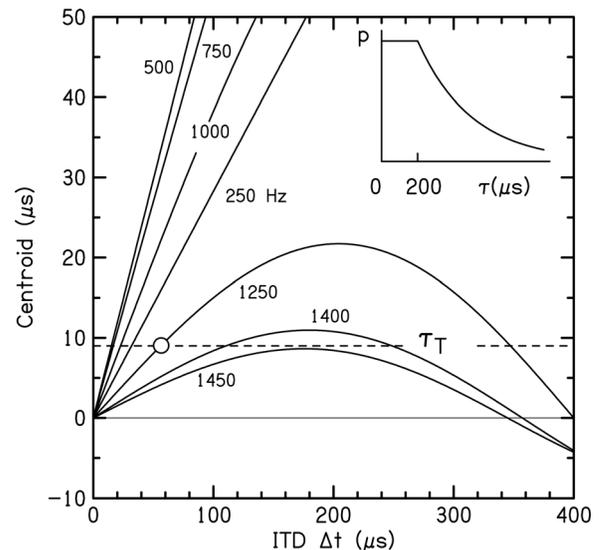


FIG. 5. Interaural delay centroid as a function of the interaural time difference,  $\Delta t$ , as computed in the centroid display model for seven tone frequencies. An illustrative value of centroid threshold  $\tau_T$  is shown at  $9 \mu\text{s}$ . The open circle shows the predicted threshold for 1250 Hz. The density  $p(\tau)$  for  $\tau > 0$  is shown by the inset in the upper right corner.

reduced to 250 Hz. At this low frequency there is unusual sensitivity to the model synaptic strength. The two low-frequency branches of the shaded region in Fig. 6(a) arise from a difference of only 10% in synaptic strength. (ii) The threshold shows a broad minimum ( $\Delta\text{ITD}$  near  $20\ \mu\text{s}$ ) between 500 and 1000 Hz. (iii) The threshold rises faster than exponentially between 1000 and 1400 Hz, successfully mimicking the rapid rise seen experimentally. (iv) Most importantly, the threshold disappears altogether as the frequency reaches 1450 Hz. There are two reasons for the vanishing threshold—the reduced modulation of the rate-ITD function at high frequency, and the canceling of side-lobes in the region of  $p(\tau)$ .

Figures 6(b) and 6(c) reveal features of the centroid calculation by selectively removing the frequency dependence of modulation or the side-lobe effect. In Fig. 6(b), the side-lobe cancellation in Fig. 6(a) was retained while modulation of the rate-ITD function ( $m$ ) was fixed, either 80% to 100% (blue) or 20% to 35% (red). Because thresholds remain finite at high frequency when the modulation is large, one knows that the decreased modulation at high frequency was essential for the divergence seen in Fig. 6(a). Side-lobe cancellation by itself is not adequate given these values of  $T_o$  and  $\tau_o$ . High-frequency thresholds diverge when the modulation is small (red region), but then the model usually overestimates the thresholds at mid frequencies.

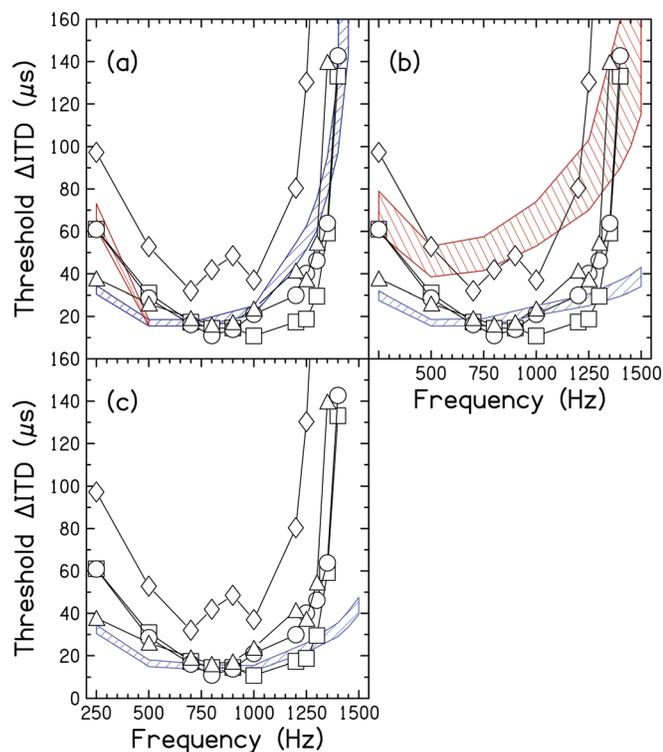


FIG. 6. Computed thresholds  $\Delta\text{ITD}$  for the centroid model with internal centroid threshold of  $9\ \mu\text{s}$  are shown by hatched regions. (a) All inclusive. (b) All inclusive except that the rate-ITD function modulation is constant,  $m = 0.8\text{--}1.0$  (blue) or  $m = 0.25\text{--}0.35$  (red). (c) All inclusive except that side-lobe cancellation is excluded by limiting the internal delay line to interaural phases in the range  $-180$  to  $+180$  deg. Symbols show experimental thresholds copied from Fig. 1.

In Fig. 6(c), the modulation parameters of Fig. 6(a) were retained while the effect of side-lobes was reduced because there were no cross-correlator cells with best interaural phase differences greater than  $180$  deg—a model feature known as the *pi-limit* (Thompson *et al.*, 2006). Therefore, interaural delay lines were shorter for higher frequencies. As expected, eliminating cells with large phase delay had no effect on thresholds at low frequency—Fig. 6(c) looks like Fig. 6(a) at low frequency. But thresholds in Fig. 6(c) remain finite at high frequency, demonstrating that side-lobe cancellation played a critical role in the high-frequency divergence in Fig. 6(a).

Figures 6(a)–6(c) show that model threshold increases as the frequency decreases from 500 to 250 Hz. Part of this increase is caused by reduced modulation at 250 Hz, but most of it is caused by the broader rate-ITD function at 250 Hz—the  $1/f$  effect.

Our experiments found that thresholds decreased as the frequency increased from 500 to 800 or 1000 Hz. Only the *pi-limit* calculations in Fig. 6(c) reproduced that feature. The inability of complete centroid calculations in Fig. 6(a) to reproduce that feature is partly due to the side-lobe cancellation and partly due to reduced modulation of model rate-ITD functions at 1000 Hz.

### C. Centroid staircase calculation

The calculation in Sec. IV B assumed a fixed centroid threshold internal to the binaural system. A calculation that is more consistent with signal detection theory would abandon such an internal threshold and compare the computed centroid with the variance intrinsic to the model binaural system. To do this alternative calculation, we ran simulated adaptive staircases with response decisions based on the rate-ITD functions of the model MSO neuron, with mean and variability, as described in Sec. III A 3. A similar psychophysically motivated test of the ITD information capability of a single-neuron (inferior colliculus) was made by Shackleton *et al.* (2003).

Centroids were computed for simulated forced-choice trials consisting of left-leading and right-leading tones. The calculations used excitation-only rate-ITD functions from Sec. III, normally distributed about their mean values shown in Fig. 2(A), with the standard deviation determined over the five computations. If the centroid for the right-leading tone was further to the right than the centroid for the left-leading tone, the response to the simulated trial was taken to be correct; otherwise, it was wrong. Sequences of simulated trials like this became simulated runs, obeying all the rules of our real staircase runs, as described in Sec. II. Dozens of simulated runs for each frequency led to model thresholds, depending only on the model MSO calculations and the  $p(\tau)$  function.

The results of the staircase simulations using the centroid display model and the model MSO cell, with its mean rate-ITD function and variability, are shown in Fig. 7 for three different decays of the  $p(\tau)$  function. All the calculations failed to agree with the experimental  $\Delta\text{ITD}$  thresholds at low frequency. Simulated staircase average thresholds

were only a few microseconds, with the staircases often down at the floor value of  $1\ \mu\text{s}$ . Thus, the model greatly underestimated human ITD thresholds.

An explanation for the failure of the centroid model at low frequency is not hard to find. The centroid model is an integral over the internal ITD axis, as weighted by the  $p(\tau)$  function. The random variations introduced into the calculation by including the variance of the rate-ITD function are positive and negative with equal probability and tend to cancel in the integration. As a result, the variance has far too little effect on the calculated lateralization to agree with the experiment. However, provided that the decay of  $p(\tau)$  is not too rapid, there is an opportunity for side lobes to cancel centroid strength, and that can lead to high and diverging thresholds at high frequency, as shown by the blue and green regions in Fig. 7.

The fraction of the cells under the  $p(\tau)$  function with a best delay in the range  $-T_o - \tau_o$  to  $+T_o + \tau_o$  is given by

$$\begin{aligned} & [1 + (1 - e^{-1})\tau_o/T_o] / [1 + \tau_o/T_o] \\ & = [1 + 0.632\tau_o/T_o] / [1 + \tau_o/T_o]. \end{aligned} \quad (8)$$

For the green region,  $T_o = 0.2\ \text{ms}$  and  $\tau_o = 0.4\ \text{ms}$ . Therefore, for this region, 75% of the cells have best delays between  $-600$  and  $600\ \mu\text{s}$ .

## V. RATE-DIFFERENCE MODEL

An alternative to a place model of ITD encoding is a model in which binaural cross-correlator cells have a narrow distribution of best interaural delay. The peak firing rate for

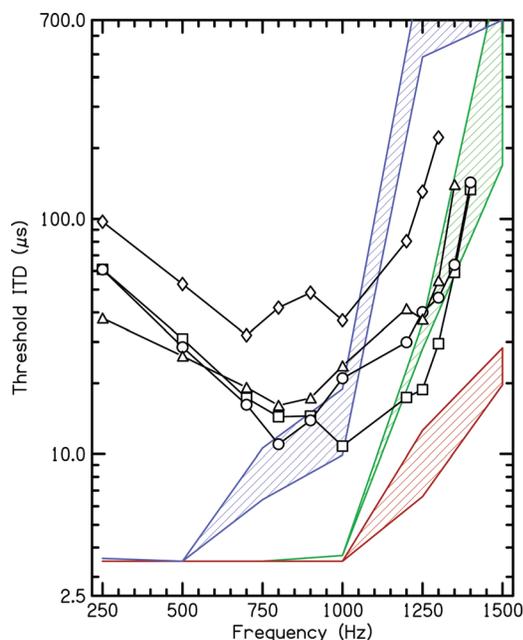


FIG. 7. Computed thresholds  $\Delta\text{ITD}$  for the centroid model from staircase simulations are shown by hatched regions—centered on the mean over 20 simulated runs and two standard deviations in width. In function  $p(\tau)$  parameter  $T_o$  was always  $0.2\ \text{ms}$ . For the blue region,  $\tau_o = 0.75\ \text{ms}$ . For the green region,  $\tau_o = 0.4\ \text{ms}$ . For the red region,  $\tau_o = 0.22\ \text{ms}$ . Experimental data are shown by open symbols. The vertical scale is logarithmic to accommodate the wide range of the calculations. The positive curvature of the data plots on this scale show the faster than exponential frequency dependence.

cells in the left brain stem is expected when waveform features occur in the right ear prior to the left. Thus, the rate-ITD function resembles Fig. 2(C), where the peak response is shifted to the right by rapidly decaying inhibition. A shift of the same sign but smaller magnitude appears in Fig. 2(B) because of slow inhibition. Figures 2(A)–2(C) show that the best delay (peak of the function) depends on the frequency of the tone. The functions for the cells in the right brain stem are assumed to be similar, except for a reversal of the ITD axis. Consequently, when a tone leads in the left ear, the excitation of the cells in the right brain stem is greater than for the cells on the left. A more central process that registers the difference in excitation can then determine the ITD and the laterality of the tone. Such a model is a “rate-difference” model for lateralization.

## A. Methods

Staircase simulations, as for the centroid model described in Sec. IV C, were done for a rate-difference model beginning with the rate-ITD functions in Figs. 2(A)–2(C). The rate-ITD functions and their standard deviations were fitted as described in Appendix B. For each simulated tone interval (right-leading and left-leading) excitation was computed for right and left model cells, including random variation consistent with the model cell standard deviation. If the difference between right and left cells was greater on the right-leading interval than on the left-leading interval, the response was said to be correct; otherwise, it was wrong. Simulated staircases were run using different values of the starting  $\Delta\text{ITD}$ , ranging from  $100\ \mu\text{s}$  to  $600\ \mu\text{s}$ . As for the real experiments, it was important that the final threshold did not depend on the starting  $\Delta\text{ITD}$ .

## B. Results

Simulated staircases for the excitation-only rate-ITD function shown in Fig. 2(A) did not converge. Because these functions are approximately even functions of the ITD, there was no reason for the excitation to be greater on one side compared to the other, whatever the  $\Delta\text{ITD}$ , and convergence would not be expected.

Thresholds computed from the slow-inhibition rate-ITD functions from Fig. 2(B) are shown by the hatched region in Fig. 8(a). The hatched region is centered on the mean threshold, computed over 100 staircase runs, and is two standard deviations in overall width. Because of the small displacement of these rate-ITD functions away from zero, staircases converged to thresholds in the range of human experiments only for 500, 750, and 1000 Hz. For other frequencies, individual staircases converged, but for 1250 Hz and 1500 Hz, the threshold increased with increasing starting ITD value. Consequently, sequences of staircases diverged.

Thresholds computed from the rapid-inhibition rate-ITD functions from Fig. 2(C) are shown in Fig. 8(b). Again, the hatched region is centered on the mean threshold, computed over 100 staircase runs, and is two standard deviations in overall width. All staircases converged for all frequencies and all starting values of the ITD. Importantly, staircases failed to diverge at 1500 Hz, contrary to experiment. The

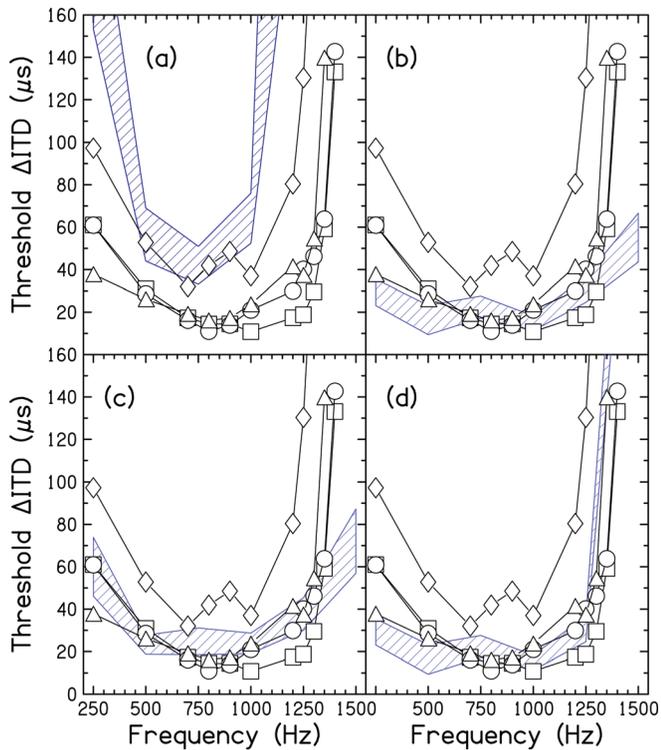


FIG. 8. Computed thresholds  $\Delta$ ITD for the rate-difference model are shown by hatched regions—centered on the mean over 100 simulated runs, and two standard deviations in width. (a) With rate-ITD functions from Fig. 2(B), slow inhibition. (b) With rate-ITD functions from Fig. 2(C), fast inhibition. (c) Same as (b) except that the standard deviations for the rate-ITD functions were computed from the mean and the duration. See the text. (d) Same as (b) except that the synaptic strength was reduced for the 1500-Hz calculation. See the text. Symbols show experimental thresholds copied from Fig. 1.

thresholds in Fig. 8(c) come from identical simulated staircases except that the standard deviation from the Hodgkin-Huxley cell calculations was replaced by the square root of the mean firing rate divided by the tone duration (0.5 s), as expected for a Poisson process. Again, all staircases converged, and thresholds were similar to those in Fig. 8(b), but usually somewhat larger.

The thresholds in Fig. 8(d) were obtained from staircases that were identical to those for Fig. 8(b) except that the effect of the rapid inhibition was reduced for 1500 Hz. The rapid inhibition for the calculation of Fig. 8(b) led to a best phase of 64 deg at 1500 Hz. In Fig. 8(d) that was reduced to 10 deg, roughly similar to the effect of slowly decaying inhibition. With that replacement, staircases converged to large threshold values, and these always depended on the starting value of  $\Delta$ ITD. Therefore, sequences of staircases did not converge as sometimes seen for human listeners at high frequencies such as 1450 Hz. However, for human listeners, individual staircases often failed to converge also.

## VI. GENERAL DISCUSSION

Interaural time difference thresholds ( $\Delta$ ITD) for sine tones were measured for four listeners to determine the detailed dependence of thresholds on tone frequency. The experimental thresholds were compared to model thresholds calculated from a Hodgkin-Huxley model for an MSO cell and several different binaural display models.

## A. Experimental summary

In the low-frequency region, 700 Hz and below, measured thresholds were approximately inversely proportional to frequency, corresponding to a constant interaural phase difference threshold. This kind of scaling can be expected based on the increasing width of rate-ITD functions with decreasing frequency. Departures from  $1/f$  scaling might be attributed to frequency-dependent responses of MSO cells or to the best-delay distribution of these cells. However, low-frequency experiments to establish these properties would have to be better than ours because our experiments encountered across- and within-individual differences too large to come to a conclusion. It was noted that the low-frequency dependence observed historically tends to be shallower than the inverse first power law predicted by scaling.

Minimum  $\Delta$ ITD thresholds occurred in the mid-frequency region between 700 and 1000 Hz. Our interpretation of this minimum is that a frequency dependence similar to the low-frequency  $1/f$  law continues to apply in this region and tends to cause thresholds to decrease with increasing frequency. However, loss of synchrony in binaural cross-correlator (MSO) cells causes thresholds to increase with increasing frequency. The tradeoff between these two effects leads to the minimum.

In the high-frequency region above 1000 Hz, thresholds grew faster than exponentially with increasing frequency until they became unmeasurable. Measurable thresholds were found at 1400 Hz for two of our four listeners, but none were found at 1450 Hz. The implication of the high-frequency data for neural models of ITD processing is that the ability of the binaural system to encode ITD does not just fade away as frequency increases. Instead, it disappears abruptly. The data suggest a neural process that suddenly stops at a critical frequency. This kind of behavior is difficult to simulate in a theoretical model.

## B. Binaural model summary

A physiologically based ITD-sensitive MSO neuron model was developed in which large action potentials are limited to the axon with only minor back-propagation to the soma, similar to real MSO neurons (Scott *et al.*, 2007). While in previous multi-compartment models (Zhou *et al.*, 2005; Mathews *et al.*, 2010; Fischl *et al.*, 2012) ITD-sensitivity decreases for inputs above 500 Hz, the frequency-range of ITD-sensitivity in the present multi-compartment model was extended upward to 1000 Hz, using specific ratios of ion-channel currents in combination with realistically fast membrane time constants at realistic resting potentials (details provided in Sec. III A). In the present study, the physiologically based excitatory synaptic time constant of 0.4 ms (Fischl *et al.*, 2012) is slower than modeled previously (Zhou *et al.*, 2005; Brand *et al.*, 2002), which helped facilitate the relatively low ITD-sensitivity at 1500 Hz that was only partially duplicated using faster synapses. Responses of the present model decreased additionally above 1000 Hz because with increasing stimulus frequency, although the input spike rate was held constant, there was an inherently reduced number of input spikes per stimulus period. The resting axonal membrane time

constant of 0.8 ms may have also contributed to the frequency selectivity of the model neuron, though it also became difficult to obtain responses across frequency in slower axons, and difficult to avoid spontaneous activity in faster axons.

The model MSO neuron produced peak-type responses to ITD (as described in [Batra et al., 1997](#)). The model neuron, with excitatory inputs only or with both excitatory and inhibitory inputs, was incorporated into several different binaural display models. Tests of the display models, comparing their predictions with human threshold data, make the assumption that the model neuron properties are realistic.

### 1. Centroid display with hard centroid threshold

When the centroid display ([Stern and Colburn, 1978](#)) was given a hard internal threshold and combined with an MSO model having excitatory inputs only, it proved possible to fit the experimental data including the low-frequency rise and the high-frequency divergence. However, the hard threshold is inconsistent with signal detection theory, and ignoring inhibition is also unrealistic.

### 2. Centroid model: Staircase simulation

Staircase simulations consistent with signal detection theory found that the centroid display model failed at low frequency, predicting ITD thresholds far lower than observed experimentally. However, the model successfully reproduced the divergence seen at high frequency. The divergence was caused by side-lobe cancellation. Additional calculations, not shown, discovered that, apart from the divergence, the faster than exponential threshold dependence required both side-lobe cancellation and frequency-dependence of the rate-ITD function, just as found for the hard centroid threshold assumption. Success with the centroid model depended on the selection of parameters, particularly in the neuron density function  $p(\tau)$ .

### 3. Rate-difference model: Excitation only

Given the physiological basis of the MSO neural model, all our rate-difference model calculations might be said to have no adjustable parameters at all. In a rate-difference model, image lateralization depends on a comparison of spike rates from MSO cells in right and left brainstems. Our rate-ITD functions for excitation only are symmetrical about zero internal delay with no displacement to either side. Consequently, the model has no broken symmetry that would lead to a rate difference and the model fails to converge to thresholds.

### 4. Rate-difference model: Slowly decaying inhibition

Incorporating inhibition reduced the MSO output rate, sharpened the temporal response, and displaced the peak of the rate-ITD function along the delay axis. However, for slowly decaying inhibition the displacement was too small to generate adequate binaural rate differences except for a few frequencies where predicted thresholds were in the range of human data. Slow inhibition was an attractive model feature because only slowly decaying inhibitory post-synaptic potentials have been recorded in MSO neurons ([Smith, 1995](#); [Magnusson et al., 2005](#)). There are, however, other sources of internal delay

apart from inhibition, notably axonal delay, as originally suggested by [Jeffress \(1948\)](#), possibly modified by different axon morphology ([Seidl et al., 2010](#)), and cochlear delays ([Shamma et al., 1989](#); [Bonham and Lewis, 1999](#)). Calculations (not shown) found that for every frequency it was possible to find some *ad hoc* additional delay (between 25 and 140  $\mu$ s) which could bring the calculated thresholds down to the measured values and below. The relationship between best-fitting additional delay and frequency was not systematic.

### 5. Rate-difference model: Rapidly decaying inhibition

An MSO model with rapidly decaying contralateral inhibition displaced the best-ITD to significant contralateral-leading ITD values, as previously demonstrated in point-neuron models ([Brand et al., 2002](#); [Zhou et al., 2005](#)). Combined with the rate-difference display, the cell with fast inhibition led to predicted thresholds in reasonable agreement with human  $\Delta$ ITD thresholds except that the model failed to exhibit the divergence at high frequency seen experimentally. The divergence could be recovered by an *ad hoc* reduction in synaptic strengths, which reduced the displacement.

## C. Perspective

### 1. Ecological advantage

The experiments of this article showed that human ITD thresholds increase faster than exponentially with increasing frequency, finally diverging just above 1400 Hz. Possibly there is some ecological advantage to this dramatic change in sensitivity. In contrast with the gerbil, which appears to employ a range of neural mechanisms to extend the upper frequency range of ITD sensitivity ([Day and Semple, 2011](#)), human listeners may benefit from reduced sensitivity to ITD fine-structure at frequencies above 1400 Hz. This reduction would mitigate the increasing ambiguity for humans in encoding fine-structure ITD in narrow-band stimuli as sound frequencies increase—where, due to the relatively large human head-size, the ITDs become greater than half a period of the stimulus, and ITD images appear on the wrong side of the midline ([Sayers, 1964](#)).

### 2. Alternative neurophysiological levels

The focus of the models presented in this article has been on low-level mechanisms in the human MSO and its inputs, and in the initial binaural display. Limiting factors include maintaining a bias toward peak-type MSO neurons across CF, sharply reduced synchrony in the high-frequency inputs from the AVCN to MSO, and a practical upper limit to excitatory synaptic strength in MSO neurons. All these factors contribute to reduced modulation in rate-ITD functions and smaller displacement of the rate-ITD function from  $ITD = 0$  leading to decreased ITD sensitivity.

Alternatively, there may be high-level suppression of fine-structure ITD responses in favor of the more reliable and behaviorally relevant cues for sound localization at high frequency: interaural-level-difference and envelope-ITD ([Strutt, 1907](#); [Henning, 1974](#); [Macpherson and Middlebrooks, 2002](#); [Joris, 2003](#)). High-level effects discounting ITDs at high

frequency may reflect evolutionary pressure based on many generations of experience with large heads.

There are also intermediate-level possibilities beyond the MSO. A potential underlying mechanism for abrupt changes in sensitivity may involve a variation in neural response type within the inferior colliculus (IC) (Sivaramakrishnan and Oliver, 2001) and a simple division of ITD-sensitive projections from the MSO within the tonotopic organization of the IC. At lower CFs, ITD-sensitive inputs from the MSO may innervate IC neurons that respond in a strong sustained manner to ongoing inputs. At higher CFs, ITD-sensitive inputs from the MSO may innervate adapting neurons that respond much more strongly to the onsets of their inputs than to their ongoing components. Such an arrangement would still enable the high-CF IC neurons to respond to wideband stimuli of transient or time-varying nature. However, although the proposed transition across CFs may be sharp, the fairly broad frequency-tuning of auditory filters would make the transition of IC responses across input frequencies more gradual and would not exhibit the sharpness seen in human thresholds.

### 3. Hybrid model

Although limitations on ITD encoding might arise from different levels of the auditory system, the modeling presented in this article makes it highly plausible that important limitations arise already at the lowest brainstem level. The explanation for the observed sharp ITD cutoff at high frequency may ultimately be traced to biological limits in low-level binaural processing of the MSO and previous stages of the binaural system. If it is granted that it is reasonable to search for a low-level explanation, it then becomes a problem that neither the centroid model nor the rate-difference model can explain all the data. The centroid model can explain the high-frequency divergence, but it fails dramatically at low frequency. The rate-difference model is successful at low and intermediate frequencies but reproduces the high-frequency divergence only with *ad hoc* assumptions that considerably reduce the displacement of the peak response along the delay axis. A hybrid model, rate code at low frequency and centroid at high frequency, could account for the observed human threshold data. Such a hybrid model is economical. At low frequencies, it avoids the need for the long internal interaural delays required by the centroid model or other variants of the Jeffress model. At high frequencies it avoids the need for a very tight distribution of internal delays required to maintain a small range of best interaural phase responses. The hybrid model has enough flexibility that it is possible to imagine that the entire frequency dependence of human ITD thresholds, including the high-frequency divergence, arises from the properties of MSO cells.

### ACKNOWLEDGMENTS

We are grateful to Dr. Les Bernstein for a useful discussion about the centroid display and to Dr. Steve Colburn for discussions about modeling. Zane Crawford provided valuable statistical help. This research was supported by The Vicerectorado de Profesorado y Ordenación Académica of the Universitat Politècnica de València (Spain), which

brought L.D. to Michigan State, by the NIDCD Grant No. DC-00181 and the AFOSR Grant No. 11NL002. A.B. was supported by NIDCD Grant Nos. DC-00100 (H. S. Colburn) and P30-DC04663 (Core Center).

## APPENDIX A: DIVERGING STAIRCASES

### 1. Forms of divergence

In the course of our experimental work we encountered experiment runs with three kinds of staircase divergence. In order of decreasing severity, these are

*a. No-threshold.* A staircase produced no threshold when, as the run progressed, the value of  $\Delta$ ITD approached a period of the tone before 14 turnarounds had occurred. Some of these runs were stopped by the experimenter.

*b. Staircase divergence.* According to our definition, a converging staircase produces a final threshold estimate for  $\Delta$ ITD that is less than the starting value. A diverging staircase produces a threshold that is larger than the starting value.

*c. Diverging staircase sequences.* A sequence of staircases diverges when thresholds increase systematically as the starting value of  $\Delta$ ITD increases. The individual staircases may or may not converge. We often found thresholds only a few dozen microseconds above the starting point, but when the starting point was increased by 50 or 100  $\mu$ s, the staircase again diverged. It became evident that there were wide regions of  $\Delta$ ITD less than the period where sequences of staircases did not converge.

### 2. Diverging staircases

Runs with a staircase divergence normally showed an increasing trend in staircase bottoms and in staircase tops. In our staircase runs with 14 turnarounds, there were 7 opportunities for staircase bottoms to decrease, and 6 opportunities for staircase tops to decrease, a total of 13 opportunities per run. Mathematically, a convergent staircase may have as few as one decreasing value, but in practice, an examination of 50 randomly chosen, converging staircases showed that the average was 5.7 (standard deviation = 1.5) decreases. By contrast, for those staircases that we identified as divergent in Sec. II B, the majority had *no decreases*. The average of 23 such divergent staircases was 0.7 decreases, far fewer than counted for converging staircases. But although staircases at high frequencies persistently diverged, that does not mean that listeners gained no information from the ITDs. Percentages of correct responses often exceeded the random guessing value of 50%.

To explore the divergence of staircases for different hypothetical observers, we ran millions of staircases, following all our rules, using responses from a random number generator. We defined the *divergence* as the difference between a staircase threshold and an arbitrary starting value. The cumulative distribution in Fig. 9 shows the percentage of staircases

with divergence less than the value given on the horizontal axis. For instance, the solid line shows that for a random-guessing observer ( $P_c = 50$ ), 50% of the staircases will diverge by  $300 \mu\text{s}$  or less. The dashed line shows that for an observer who chooses correctly on two-thirds of the trials ( $P_c = 67$ ), 50% of staircases will diverge by  $90 \mu\text{s}$  or less.<sup>3</sup>

Divergences were studied for individual listeners at the smallest frequencies for which no convergence occurred:

a. *L1 at 1450 Hz.* For nine staircases, the median divergence was  $144 \mu\text{s}$ ; the mean was  $181 \mu\text{s}$ .

b. *L2 at 1450 Hz.* For six staircases, the median divergence was  $130 \mu\text{s}$ ; the mean was  $140 \mu\text{s}$ .

c. *L3 at 1400 Hz.* For nine staircases, the median divergence was  $68 \mu\text{s}$ ; the mean was  $62 \mu\text{s}$ .

d. *L4 at 1350 Hz.* For four staircases, the median divergence was  $120 \mu\text{s}$ ; the mean was  $136 \mu\text{s}$ .

Most of these divergences correspond to a percentage of correct responses somewhere between 50% and 67%. Therefore, although no convergence occurred according to our protocol, there would exist some staircase, targeting a low, but finite, positive value of  $d'$ , which would converge. At the same time, except for L3 at 1400 Hz, the listeners would probably not produce valid thresholds at these high-frequency limits in protocols that require performance greater than 70% correct. This analysis of divergence leads to additional confidence that ITD discrimination for sine tone stops between 1400 and 1450 Hz.

## APPENDIX B: RATE-ITD FITS

The simulations of this article required analytic forms for the rate-ITD functions shown in Fig. 2. These functions resemble offset cosine functions, as expected for the cross-

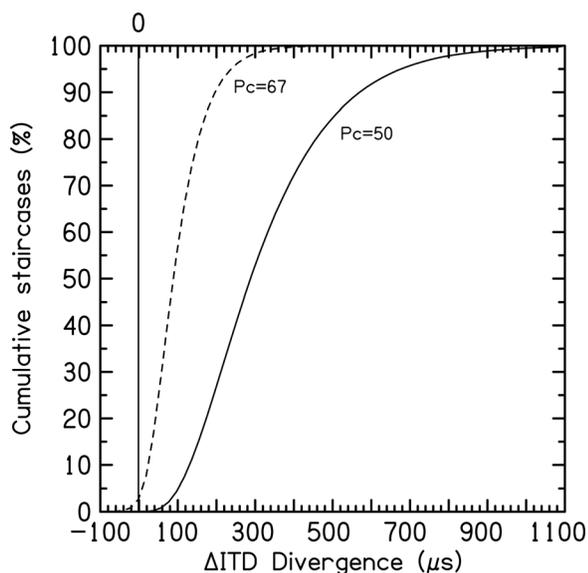


FIG. 9. Cumulative histogram for divergence for different hypothetical listeners: Solid line 50% correct. Dashed line 67% correct. In the latter case, 2% of runs converge, i.e., the divergence is negative.

TABLE VI. (a) Fitting parameters, excitation only. (b) Fitting parameters, slow inhibition. (c) Fitting parameters, fast inhibition.

| $f$ (Hz) | $A$   | $B$   | $\phi$ (deg) | $\eta$ | $A'$  | $B'$  | $\phi'$ (deg) | $\eta'$ |
|----------|-------|-------|--------------|--------|-------|-------|---------------|---------|
| (a)      |       |       |              |        |       |       |               |         |
| 250      | 129.6 | 102.4 | -1           | 0.160  | 9.37  | 5.47  | -1            | 0.025   |
| 500      | 217.2 | 217.2 | 0            | 0.135  | 6.98  | 7.48  | 0             | 0.070   |
| 750      | 204.4 | 188.4 | 0            | 0.070  | 13.05 | 10.60 | 0             | 0.030   |
| 1000     | 241.8 | 212.6 | -1           | 0.0    | 16.96 | 11.86 | -1            | 0.200   |
| 1250     | 112.0 | 48.0  | -5           | 0.0    | 20.37 | 15.0  | -5            | 0.040   |
| 1500     | 143.2 | 31.6  | 0            | 0.025  | 20.53 | 15.76 | 0             | 0.035   |
| (b)      |       |       |              |        |       |       |               |         |
| 250      | 86.4  | 86.4  | 5            | 0.260  | 6.38  | 6.38  | 5             | 0.135   |
| 500      | 109.2 | 109.2 | 8            | 0.230  | 8.09  | 8.09  | 8             | 0.155   |
| 750      | 128.4 | 125.2 | 14           | 0.125  | 10.07 | 8.98  | 14            | 0.080   |
| 1000     | 126.4 | 112.0 | 13           | 0.055  | 12.43 | 9.8   | 13            | 0.080   |
| 1250     | 81.4  | 38.6  | 5            | 0.015  | 14.2  | 7.78  | 5             | 0.075   |
| 1500     | 82.0  | 22.4  | 18           | 0.0    | 17.3  | 10.1  | 18            | 0.40    |
| (c)      |       |       |              |        |       |       |               |         |
| 250      | 82.4  | 82.4  | 30           | 0.235  | 5.26  | 5.26  | 30            | 0.110   |
| 500      | 107.8 | 107.8 | 40           | 0.180  | 7.70  | 7.70  | 40            | 0.105   |
| 750      | 125.2 | 121.2 | 66           | 0.050  | 11.5  | 9.36  | 66            | 0.005   |
| 1000     | 122.2 | 105.4 | 77           | 0.005  | 9.80  | 6.84  | 77            | 0.015   |
| 1250     | 79.6  | 40.4  | 86           | 0.010  | 10.6  | 5.95  | 86            | 0.020   |
| 1500     | 91.4  | 23.4  | 64           | 0.0    | 7.68  | 7.68  | 64            | 0.005   |

correlation of sine tones, but for low frequencies the functions are not symmetrical—the peaks are sharper than the valleys. The rate-ITD functions were fitted with a four-parameter equation

$$c(\tau) = A + B \cos[2\pi f\tau + \phi + 2\pi\eta \sin(2\pi f\tau + \phi)].$$

Here  $A$  is the average value (vertical offset), and  $B$  is the amplitude of the modulation. Parameter  $\phi$  indicates the best interaural phase, equal to the best ITD multiplied by the frequency. Parameter  $\eta$  is the asymmetry parameter which phase modulates the cosine function synchronously with the tone frequency. For  $\eta = 0$ , there is no peak sharpening. For  $\eta$  near 0.2, there is considerable sharpening, and for  $\eta$  greater than 0.2 the function acquires some oscillatory character.

Values of the parameters for the A, B, and C functions in Fig. 2 are given in Tables VI(a), VI(b), and VI(c) respectively.

Because the standard deviations of the rate-ITD functions tended to follow the functions themselves, the same form was used for the standard deviation. Logic required that the phase  $\phi$  be the same for a rate-ITD function and the corresponding standard deviation. The parameters for standard deviations are given by primed variables in Table VI.

<sup>1</sup>A function of the form  $\Delta\text{ITD} = t_o \exp(-f/f_o)$ , where  $t_o$  and  $f_o$  are adjustable parameters, leads to an excellent fit to the low-frequency data for all four listeners—from 250 to 800 Hz for L1, L2, and L3 and from 250 to 700 Hz for L4. It is a much better fit for every listener than the function with constant interaural phase difference,  $\Delta\text{ITD} = c/f$ . Of course, the exponential fit has two adjustable parameters and the constant IPD has only one. That may make the comparison unfair.

<sup>2</sup>At 250 Hz, with a stimulus period of 4 ms, spike rates continued to decrease for ITDs from  $-1$  to  $-2$  ms and from  $1$  to  $2$  ms (not shown). In the conditions with inhibition, for this expanded range of ITD, spike rates were near zero (maximum 8 spikes/s). In the purely excitatory model, spike rates decreased from 56 to 27 spikes/s as ITD decreased from  $-1$  to  $-2$  ms, and decreased from 49 to 28 spikes/s as ITD increased from  $1$  to  $2$  ms.

2 ms. Our calculated ITD discrimination thresholds include the simulation results from this expanded range of ITD at 250 Hz.

<sup>3</sup>The computer model in this appendix is unusual because the percentage of correct responses is held constant and does not grow as the ITD grows. However, this may be an appropriate assumption in the region of large interaural phase shifts where this model is applied.

- Batra, R., Kuwada, S., and Fitzpatrick, D. C. (1997). "Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. II. Coincidence detection," *J. Neurophysiol.* **78**, 1237–1247.
- Bonham, B. H., and Lewis, E. R. (1999). "Localization by interaural time difference (ITD): Effects of interaural frequency mismatch," *J. Acoust. Soc. Am.* **106**, 281–290.
- Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (2002). "Precise inhibition is essential for microsecond interaural time difference coding," *Nature* **417**, 543–547.
- Colburn, H. S. (1977). "Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise," *J. Acoust. Soc. Am.* **61**, 525–533.
- Colburn, H. S., Chung, Y., Zhou, Y., and Brughera, A. (2008). "Models of brainstem responses to bilateral electrical stimulation," *J. Assoc. Res. Otolaryngol.* **10**, 91–110.
- Couchman, K., Grothe, B., and Felmy, F. (2010). "Medial superior olive neurons receive surprisingly few excitatory and inhibitory inputs with balanced strength and short-term dynamics," *J. Neurosci.* **30**, 17111–17121.
- Day, M. L., and Semple, M. N. (2011). "Frequency-dependent interaural delays in the medial superior olive: Implications for interaural cochlear delays," *J. Neurophysiol.* **106**, 1985–1999.
- Domnitz, R. H. (1973). "The interaural time jnd as a simultaneous function of interaural time and interaural amplitude," *J. Acoust. Soc. Am.* **53**, 1549–1552.
- Domnitz, R. H., and Colburn, H. S. (1977). "Lateral position and interaural discrimination," *J. Acoust. Soc. Am.* **61**, 1586–1598.
- Dye, R. H. (1990). "The combination of interaural information across frequencies: Lateralization on the basis of interaural delay," *J. Acoust. Soc. Am.* **88**, 2159–2170.
- Fischl, M. J., Combs, T. D., Klug, A., Grothe, B., and Burger, R. M. (2012). "Modulation of synaptic input by GABAB receptors improves coincidence detection for computation of sound location," *J. Physiol. (London)* **590**, 3047–3066.
- Goldberg, J. M., and Brown, P. B. (1969). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization," *J. Neurophysiol.* **32**, 613–636.
- Haftner, E. R., Dye, R. H., and Gilkey, R. H. (1979). "Lateralization of tonal signals which have neither onsets nor offsets," *J. Acoust. Soc. Am.* **65**, 471–477.
- Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex wave-forms," *J. Acoust. Soc. Am.* **55**, 84–90.
- Henning, G. B. (1983). "Lateralization of low-frequency transients," *Hear. Res.* **9**, 153–172.
- Hershkowitz, R. M., and Durlach, N. I. (1969). "Interaural time and amplitude jnds for a 500-Hz tone," *J. Acoust. Soc. Am.* **46**, 1464–1467.
- Jeffress, L. A. (1948). "A place theory of sound localization," *J. Comp. Physiol. Psychol.* **41**, 35–39.
- Joris, P. X. (2003). "Interaural time sensitivity dominated by cochlea-induced envelope patterns," *J. Neurosci.* 6345–6350.
- Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. T. (1994). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency," *J. Neurophysiol.* **71**, 1022–1036.
- Klump, R. B., and Eady, H. R. (1956). "Some measurements of interaural time difference thresholds," *J. Acoust. Soc. Am.* **28**, 859–860.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Macpherson, E. A., and Middlebrooks, J. C. (2002). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," *J. Acoust. Soc. Am.* **111**, 2219–2236.
- Magnusson, A. K., Kapfer, C., Grothe, B., and Koch, U. (2005). "Maturation of glycinergic inhibition in the gerbil medial superior olive after hearing onset," *J. Physiol. (London)* **568**, 497–512.
- Mathews, P. J., Jercog, P. E., Rinzal, J., Scott, L. L., and Golding, N. L. (2010). "Control of submillisecond synaptic timing in binaural coincidence detectors by  $K_v1$  channels," *Nat. Neurosci.* **13**, 603–609.
- McAlpine, D., Jiang, D., and Palmer, A. R. (2001). "A neural code for low-frequency sound localization in mammals," *Nat. Neurosci.* **4**, 396–401.
- McFadden, D., Jeffress, L. A., and Russell, W. E. (1973). "Individual differences in sensitivity to interaural differences in time and level," *Percept. Mot. Skills* **37**, 755–761.
- Mills, A. W. (1958). "On the minimum audible angle," *J. Acoust. Soc. Am.* **30**, 237–246.
- Nordmark, J. O. (1976). "Binaural time discrimination," *J. Acoust. Soc. Am.* **60**, 870–880.
- Rakerd, B., and Hartmann, W. M. (1986). "Localization of sound in rooms. III: Onset and duration effects," *J. Acoust. Soc. Am.* **80**, 1695–1706.
- Richard, G. L., and Haftner, E. R. (1973). "Detection of interaural time differences in short duration, low-frequency tones," *J. Acoust. Soc. Am.* **53**, 335.
- Rothman, J. S., and Manis, P. B. (2003). "The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons," *J. Neurophysiol.* **89**, 3097–3113.
- Sayers, B. McA. (1964). "Acoustic-image lateralization judgments with binaural tones," *J. Acoust. Soc. Am.* **36**, 923–926.
- Schiano, J. L., Trahiotis, C., and Bernstein, L. R. (1986). "Lateralization of low-frequency tones and narrow bands of noise," *J. Acoust. Soc. Am.* **79**, 1563–1570.
- Scott, L. L., Hage, T. A., and Golding, N. L. (2007). "Weak action potential backpropagation is associated with high-frequency axonal firing capability in principal neurons of the gerbil medial superior olive," *J. Physiol. (London)* **583**, 647–661.
- Scott, L. L., Mathews, P. J., and Golding, N. L. (2005). "Posthearing developmental refinement of temporal processing in principal neurons of the medial superior olive," *J. Neurosci.* **25**, 7887–7895.
- Scott, L. L., Mathews, P. J., and Golding, N. L. (2010). "Perisomatic voltage-gated sodium channels actively maintain linear synaptic integration in principal neurons of the medial superior olive," *J. Neurosci.* **30**, 2039–2050.
- Seidl, A. H., Rubel, E. W., and Harris, D. M. (2010). "Mechanisms for adjusting interaural time differences to achieve binaural coincidence detection," *J. Neurosci.* **30**, 70–80.
- Shackleton, T. M., Skottun, B. C., Arnott, R. H., and Palmer, A. R. (2003). "Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of guinea pigs," *J. Neurosci.* **23**, 716–724.
- Shamma, S. A., Shen, N., and Gopalaswamy, P. (1989). "Stereoausis: Binaural processing without neural delays," *J. Acoust. Soc. Am.* **86**, 989–1006.
- Sivaramakrishnan, S., and Oliver, D. L. (2001). "Distinct K currents result in physiologically distinct cell types in the inferior colliculus of the rat," *J. Neurosci.* **21**, 2861–2877.
- Skottun, B. C., Shackleton, T. M., Arnott, R. H., and Palmer, A. R. (2001). "The ability of inferior colliculus neurons to signal differences in interaural delay," *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14050–14054.
- Smith, A. J., Owens, S., and Forsythe, I. D. (2000). "Characterisation of inhibitory and excitatory postsynaptic currents of the rat medial superior olive," *J. Physiol. (London)* **529**, 681–698.
- Smith, P. H. (1995). "Structural and functional differences distinguish principal from non-principal cells in the guinea pig MSO slice," *J. Neurophysiol.* **73**, 1653–1667.
- Stern, R. M., and Colburn, H. S. (1978). "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," *J. Acoust. Soc. Am.* **64**, 127–140.
- Stern, R. M., and Shear, G. D. (1996). "Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay," *J. Acoust. Soc. Am.* **100**, 2278–2288.
- Strutt, J. W. (1907). "On our perception of sound direction," *Philos. Mag.* **13**, 214–232.
- Thompson, S. K., von Kreigstein, K., Deane-Pratt, A., Marquardt, T., Deichmann, R., Griffiths, T. D., and McAlpine, D. (2006). "Representation of interaural time delay in the human auditory midbrain," *Nat. Neurosci.* **9**, 1096–1098.
- Yin, T. C. T., and Chan, J. C. K. (1990). "Interaural time sensitivity in medial superior olive of cat," *J. Neurophysiol.* **64**, 465–488.
- Yost, W. A. (1974). "Discriminations of interaural phase differences," *J. Acoust. Soc. Am.* **55**, 1299–1303.
- Zhou, Y., Carney, L. H., and Colburn, H. S. (2005). "A model for interaural time difference sensitivity in the medial superior olive: Interaction of excitatory and inhibitory inputs, channel dynamics, and cellular morphology," *J. Neurosci.* **25**, 3046–3058.
- Zwislocki, J., and Feldman, R. S. (1956). "Just noticeable differences in dichotic phase," *J. Acoust. Soc. Am.* **28**, 860–864.