Chapter 27
Interaural Time Difference Thresholds as a Function of Frequency

William M. Hartmann, Larisa Dunai, and Tianshu Qu

Abstract Different models of the binaural system make different predictions for the just-detectable interaural time difference (ITD) for sine tones. To test these models, ITD thresholds were measured for human listeners focusing on high- and low-frequency regions. The measured thresholds exhibited a minimum between 700 and 1,000 Hz. As the frequency increased above 1,000 Hz, thresholds rose faster than exponentially. Although finite thresholds could be measured at 1,400 Hz, experiments did not converge at 1,450 Hz and higher. A centroid computation along the interaural delay axis, within the context of the Jeffress model, can successfully simulate the minimum and the high-frequency dependence. In the limit of medium-low frequencies (f), where $f \cdot \text{ITD} \ll 1$, mathematical approximations predict low-frequency slopes for the centroid model and for a rate-code model. It was found that measured thresholds were approximately inversely proportional to the frequency (slope $= -1$) in agreement with a rate-code model. However, the centroid model is capable of a wide range of predictions (slopes from 0 to -2).

W.M. Hartmann (✉)
Department of Physics and Astronomy,
Michigan State University,
East Lansing, MI 48824, USA
e-mail: hartmann@pa.msu.edu

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

T. Qu
Key Laboratory on Machine Perception-Ministry of Education,
Peking University, Beijing 100871, China

1 Introduction

The interaural time difference (ITD) is an important acoustical property that is used by humans and other animals to localize the sources of sound. This chapter studies the ability to discriminate between positive and negative ITDs for sine tones as a function of tone frequency. The chapter focuses on the functional dependence in high- and low-frequency limits with the goal of testing different models of binaural hearing.

2 High Frequencies

It is well known that human listeners are not able to detect interaural time differences (ITDs) in sine tones with frequencies greater than about 1,500 Hz (Zwislocki and Feldman 1956; Klumpp and Eady 1956). Those articles showed that although the smallest ITD thresholds occurred near 1,000 Hz, thresholds became unmeasurably high when the frequency increased above 1,300 Hz. The results point to a dramatic failure in ITD processing over a very small frequency range. Our high-frequency experiments explored this dramatic dependence in detail and determined the functional dependence of the high-frequency failure.

2.1 Methods

The listener heard two tones and was required to say whether the second tone appeared to be to the left or the right of the first. The tones were 460 ms in overall duration including a 140-ms rise duration and a 140-ms fall. The long rise time was intended to prevent the onset (identical for the two ears) from affecting ITD judgements. Tones were presented to the listener by headphones at a level of 60 dB SPL – the same in both ears. The level of the electrical signals was increased at lower frequencies to compensate for the headphone frequency response.

The application of ITDs was symmetrical about zero. For example, in a *left-right* trial with a nominal ΔITD of 20 μs, the tone led in the left ear by 10 μs during the first interval, and it led in the right ear by 10 μs during the second interval.

The experiment used a three-down, one-up adaptive staircase procedure with variable increments. The trials of an experimental run continued until the staircase had made 14 turnarounds, and the final 10 were averaged to obtain a ΔITD threshold for the run. Six runs were averaged to find the final threshold at any given frequency.

2.2 Results

There were four listeners in the experiment. Their thresholds appear in Fig. 27.1. The new information in Fig. 27.1 is that in the high-frequency limit, the threshold ΔITD grows faster than exponentially with increasing frequency.

Fig. 27.1 ΔITD thresholds for four listeners are shown by symbols. Data for L2, L3, and L4 are offset vertically by 20, 50, and 60 μs, respectively. 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 $1/(f_c - f)^\eta$
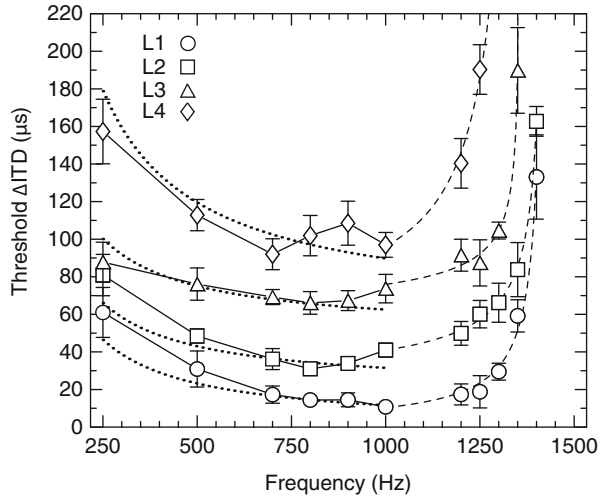
It is possible to fit the dramatic lateralization failure at 1,450 Hz with a signal processing theory that extends the Jeffress (1948) model of the binaural system. The theory has two main parts. One part is an array of coincidence cells in the midbrain operating as cross-correlators, as observed physiologically in the medial superior olive (MSO) (e.g., Goldberg and Brown 1969; Yin and Chan 1990; Coffee et al. 2006). 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 with only a weak frequency dependence.

2.3 Centroid Theory

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) to fit the lateralization data of Schiano et al. (1986). In this model display, a sine tone with angular frequency ω and an ITD of Δt excites midbrain cross-correlators represented by a cross-correlation function $c(\omega\tau)$, where τ is the lag and values of τ have a density distribution $p(\tau|\omega)$, centered on $\tau=0$.

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

$$\bar{\tau} = \frac{\int d\tau\, p(\tau|\omega)\, \tau\, c(\omega\tau - \omega\Delta t)}{\int d\tau\, p(\tau|\omega)\, c(\omega\tau - \omega\Delta t)}. \tag{27.1}$$

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

Fig. 27.2 Interaural delay
centroid as computed in the
centroid display model for six
tone frequencies. An
illustrative value of centroid
threshold τ_T is shown at 8 ms.
It predicts, for example, that
for a 250-Hz tone, the ΔITD
threshold is 76 μs, and that
for 1,450 Hz, there is no
threshold at all
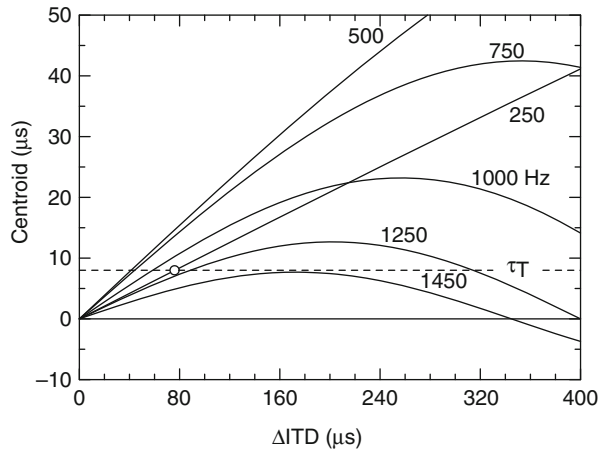
Values of $\bar{\tau}$ can be computed given reasonable choices for $c(\omega\tau)$ and $p(\tau|\omega)$.
A reasonable choice for $c(\omega\tau)$ is a cosine function, at least for frequencies of 750 Hz
and higher,

$$c(\omega\tau) = 1 + m\cos(\omega\tau). \tag{27.2}$$

Parameter m ($m \le 1$) is the "rate-ITD modulation." The density of lag values can
be modelled as a constant for very small τ followed by an exponentially decaying
function of τ, independent of frequency as per Colburn (1977).

As pointed out by Stern and Shear, if the width of $p(\tau)$ is chosen correctly, then
as the tone frequency increases, more and more cycles of $c(\omega\tau - \omega\Delta t)$ fit within the
range of lags given by $p(\tau)$. This has the effect of preventing the centroid from
increasing much as Δt increases because of partial cancellation of the positive lobes
of $c(\omega\tau - \omega\Delta t)$ by the negative lobes. 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 lateralize at 1,450 Hz and above.

Values of centroid $\bar{\tau}$ computed from the model are shown in Fig. 27.2 for $m = 0.4$
and for six different tone frequencies. Predictions for a threshold ΔITD can be made
if it is assumed that there is a threshold value of centroid τ_T. For example, if it is
assumed that the centroid threshold is $\tau_\mathrm{T} = 8\ \mu$s, as shown by the dashed horizontal
line in Fig. 27.2, then the model predicts that the threshold disappears altogether as
the frequency approaches 1,450 Hz. However, in order to model the faster than
exponential increase in threshold, the rate modulation m must decrease rapidly with
increasing frequency.

3 Low Frequencies

All models for the neurophysiological processing of ITD depend on the
cross-correlation between the neural inputs from the left and right ears.
The cross-correlation function provides a measure of the difference the phase

or timing of signals arriving at the two ears and thereby encodes the azimuth of a source. But although there is a general agreement about the importance of cross-correlation, there are differences of opinion about how it is applied functionally. The 1948 Jeffress model imagines a doubly tuned array of cross-correlators, tuned in best interaural delay and tuned in frequency. The tuning in best delay is normally thought to be influenced by the largest possible delay in free field given the head size, but is otherwise rather broad, enabling a *place* model for localization. Tones with different ITD cause different neurons in the central auditory system to light up.

An alternative model abandons the concept of place process. Physiological studies of single units in the inferior colliculus of guinea pigs show a strong correlation between the best interaural delay Δt and the best frequency f. The relationship is such that the phase angle $f\Delta t$ is in the neighborhood of an interaural phase of 45° (McAlpine et al. 2001). The two different models correspond to different mathematical forms for the density $p(\tau)$.

The goal of the low-frequency experiments reported here was to test models of ITD encoding against experimental measurements of just-detectable ITDs. The advantage of low-frequency experiments is that the synchrony of inputs to the cross-correlators becomes very high and stable (Joris et al. 1994) so that the remaining frequency dependence can be attributed to $p(\tau|\omega)$ and to differences between binaural display models.

3.1 Low-Frequency Theory

Binaural theories require three elements: a model cross-correlator c, a distribution function for cross-correlation units $p(\tau|\omega)$, and a model display. Function c is always a function of phase, which means that temporal parameters such as the ITD, Δt, always enter in the form $\omega\Delta t$.

In the low-frequency limit, the functional dependence on frequency can be extracted from c by expanding it in a Taylor series about $\omega\tau$. Then details of this cross-correlation function become unimportant, and attention can be focused on the distribution p and the display model. For our thresholds, the expansion in IPD, $\omega\Delta t$, is valid because the largest value obtained was always less than 0.07 cycles.

3.2 Centroid Model

The centroid model of Stern and Colburn (1978) and Stern and Shear (1996) is a model of the Jeffress type. A priori, it incorporates cells with a wide range of best delay at any frequency, though the effective extent of the array is limited by $p(\tau|\omega)$.

If the best delays, τ, scale with frequency such that $\omega\tau$ is constant, then density p can be written in terms of phase lag only, that is,

$$p(\tau\,|\,\omega)d\tau = p(\omega\tau)\mathrm{d}(\omega\tau) \qquad (27.3)$$

Then because c and its derivative c' are also functions of phase, the low-frequency limit of ΔITD threshold is independent of the tone frequency.

A second simple case occurs when the distribution of best delays does not depend at all on frequency, as for the high-frequency calculations in the previous section. Then the predicted threshold becomes

$$\Delta t = \Delta \overline{\tau}_T \, / [2m < \tau^2 > \omega^2], \tag{27.4}$$

where $< \tau^2 >$ is the second moment of $p(\tau|\omega)$. Therefore, the low-frequency limit of the ΔITD threshold varies inversely as the square of the tone frequency. When $p(\tau|f)$ includes parts that scale inversely with ω and parts that are independent of ω, as in the form investigated by Stern and Shear (1996), low-frequency dependences that are intermediate between flat and $1/\omega^2$ are possible.

3.3   Rate-Code Models

If function $p(\tau|\omega)$ is narrowly distributed on one side, [e.g., $p(\tau|\omega) \gg p(-\tau|\omega)$], the only available encoding for ITD is the difference in firing rates from left and right midbrain centers, $E_R - E_L$, presumably computed at a higher center.

A development parallel to the centroid model above, applied to the same symmetrical discrimination experiment, uses the change in $E_R - E_L$ between the two intervals, here defined as $\overline{\Delta}$, as the discrimination statistic. In the low-frequency limit,

$$\overline{\Delta} = 4\omega \Delta t \int_0^\infty d\tau \left[p(\tau|\omega) - p(-\tau|\omega) \right] | c'(\omega\tau) | \tag{27.5}$$

If $p(\tau|\omega)$ is a function of the product $\omega\tau$, as it appears to be in the guinea pig, the integral is independent of ω, and the frequency dependence of the ITD threshold is given by

$$\Delta t \propto \frac{\overline{\Delta}_T}{[4m\omega]}, \tag{27.6}$$

where $\overline{\Delta}_T$ is a constant, the threshold value of the discriminator. Therefore, the ΔITD is inversely proportional to the first power of the frequency.
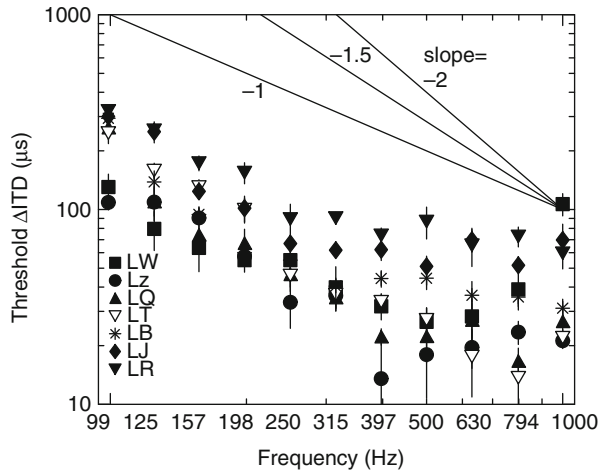
If $p(\tau|\omega)$ is independent of ω, then

$$\Delta t = \frac{\overline{\Delta}_T}{[4m < \tau > \omega^2]}, \tag{27.7}$$

where $< \tau >$ is the first moment of $p(\tau|\omega) - p(-\tau|\omega)$. The conclusions of the calculations in the low-frequency limit are given in Table 27.1.

Table 27.1 Predicted low-frequency slopes of ΔITD as a function of frequency

Model	$p(\tau|\omega)=p(\tau)$	$p(\tau|\omega)=p(\omega\tau)$
Centroid	−2	0
Rate	−2	−1

Fig. 27.3 Low-frequency thresholds for seven listeners. Three slopes are shown

3.4 Low-Frequency Experiment

Experimental stimuli were sine tones with 11 different frequencies spaced by one-third octave from 1,000 Hz down to 99 Hz. Methods were the same as for the high-frequency experiment. There were seven listeners in the experiment, with thresholds shown in Fig. 27.3. Maximum-likelihood slopes for the seven listeners were as follows: −0.97, −1.15, −1.43, −1.43, −1.08, −1.25, and −0.83 μs. The average is −1.16 and the standard deviation is 0.23. This result is most evidently consistent with the rate-code prediction with a distribution of the form $p(\tau|\omega)=p(\omega\tau)$, namely, a slope of −1. However, a final conclusion awaits the results of further experiments at still lower frequencies where the slopes appear to become steeper.

4 Conclusion

Mathematical models based on cross-correlation can successfully fit features of the mid- and high-frequency dependence of ITD thresholds, including the divergence at 1,450 Hz. They also predict the low-frequency limit of thresholds. The models gain simplicity by postulating thresholds internal to the binaural system. Incorporating variance, as in signal detection theory, can be expected to make the predictions fuzzier and more complicated.

References

Coffee CS, Ebert CS, Marshall AF, Skaggs JD, Falk SE, Crocker WD, Pearson JM, Fitzpatrick DC (2006) Detection of interaural correlation by neurons in the superior olivary complex, inferior colliculus, and auditory cortex of the unanesthetized rabbit. Hear Res 221:1–16

Colburn HS (1977) Theory of binaural interaction based on auditory-nerve data II. Detection of tones in noise. J Acoust Soc Am 61:525–533

Goldberg JM, Brown PB (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

Jeffress LA (1948) A place theory of sound localization. J Comp Physiol Psychol 41:35–39

Joris PX, Carney LH, Smith PH, Yin TCT (1994) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. J Neurophysiol 71:1022–1036

Klumpp RB, Eady HR (1956) Some measurements of interaural time difference thresholds. J Acoust Soc Am 28:859–860

McAlpine D, Jiang D, Palmer AR (2001) A neural code for low-frequency sound localization in mammals. Nat Neurosci 4:396–401

Schiano JL, Trahiotis C, Bernstein LR (1986) Lateralization of low-frequency tones and narrow bands of noise. J Acoust Soc Am 79:1563–1570

Stern RM, Colburn HS (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 RM, Shear GD (1996) Lateralization and detection of low-frequency binaural stimuli: effects of distribution of internal delay. J Acoust Soc Am 100:2278–2288

Yin TCT, Chan JCK (1990) Interaural time sensitivity in medial superior olive of cat. J Neurophysiol 64:465–488

Zwislocki J, Feldman RS (1956) Just noticeable differences in dichotic phase. J Acoust Soc Am 28:860–864