On the detection of dispersion in the head-related transfer function

Zachary A. Constan[a]
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

William M. Hartmann[b]
Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824 [c]

(Received 22 November 2002; revised 21 May 2003; accepted 21 May 2003)

Because of dispersion in head-related transfer functions (HRTFs), the interaural time difference (ITD) varies with frequency. This physical effect ought to have consequences for the size or shape of the auditory image of broadband noise because different frequency regions of the noise have different ITDs. However, virtual reality experiments suggest that human listeners are insensitive to head-related dispersion. The experiments of this article test that suggestion by experiments that isolate dispersion from amplitude effects in the HRTF and attempt to optimize the opportunity for detecting it. Nevertheless, the experiments find that the only effect of dispersion is to shift the lateralization of the auditory image. This negative result is explained in terms of the cross-correlation function for head-dispersed noise. Although the broad-band cross-correlation function differs considerably from 1.0, the cross-correlation functions within bands characteristic of auditory filters do not. A detailed study of the lateralization shifts show that the experimental shifts can be successfully calculated as an average of stimulus ITDs as weighted by Raatgever's frequency-weighting function (Thesis, Delft, The Netherlands, 1980). [DOI: 10.1121/1.1592159]

PACS numbers: 43.66.Pn, 43.66.Ba, 43.66.Qp [LRB]

I. INTRODUCTION

As a sound wave approaches a listener's ear, it is diffracted by the listener's head. Diffraction causes the sound wave to be filtered, as characterized by the head-related transfer function (HRTF). The HRTF is dispersive, i.e., the phase shift is not a linear function of the frequency. Instead, the phase shift increases more slowly than linearly with increasing frequency. Because the ratio of the phase shift to frequency is the phase delay, the phase delay is not constant but decreases with increasing frequency. It is as though high-frequency sound waves traveled faster around the head than low-frequency waves.

The effects of diffraction by the head, including dispersion, can be well approximated by a diffraction formula for the sound pressure on the surface of a sphere. The formula for an incident plane wave from a distant source (Kuhn, 1977, 1979, 1982, 1987; Kuhn and Guernsey, 1983) appears as Eq. (A1) in the Appendix. This series formula gives the complex sound pressure as a function of azimuth, θ_i, frequency, f, and head radius, a. The frequency and radius enter only as the dimensionless variable $c/(2\pi af)$, where c is the speed of sound. For an incident wave from a nearby source, a corresponding series formula was given by Brungart (1999), Brungart and Rabinowitz (1999), and Brungart *et al.* (1999). In that case the distance from the source to the head is another variable.

In connection with binaural hearing, where two ears are involved, it is evident that diffraction will have an effect on *interaural* properties. In particular, dispersion affects the interaural time difference (ITD), defined here as the interaural phase delay. The ITD is an important cue used by listeners to determine the azimuth of a sound source (Strutt, 1907), and dispersion causes the ITD to depend, not only on the azimuth, but also on the frequency.

Figure 1 shows the ITD as a function of frequency for seven different azimuths of incidence (measured from the forward direction) as calculated from the spherical-head formula. The high-frequency limit of the ITD, shown by filled squares on the right of Fig. 1, is two-thirds of the low-frequency limit, shown by open circles on the left. A comparison between the curves and the high- and low-frequency limits shows that most of the change in ITD occurs in a rather narrow frequency region. For frequencies around 1 kHz, where much of the dispersion occurs, the wavelengths of the sound waves are large compared to head anatomical details and very large compared to details of the pinna. Therefore, as observed by Kuhn (1982) or Brungart and Rabinowitz (1999), the ITDs calculated from the spherical-head formula turn out to be in reasonable agreement with ITDs measured on real heads or on artificial heads such as KEMAR.

The question posed in this article is whether human listeners are sensitive to the dispersion created by the head. One line of reasoning suggests that dispersion should, in-

[a]Present address: Department of Physics, Albion College, Albion, MI 49224.
[b]Author to whom correspondence should be addressed. Electronic mail: hartmann@pa.msu.edu
[c]Present address.

 0001-4966/2003/114(2)/998/11/$19.00

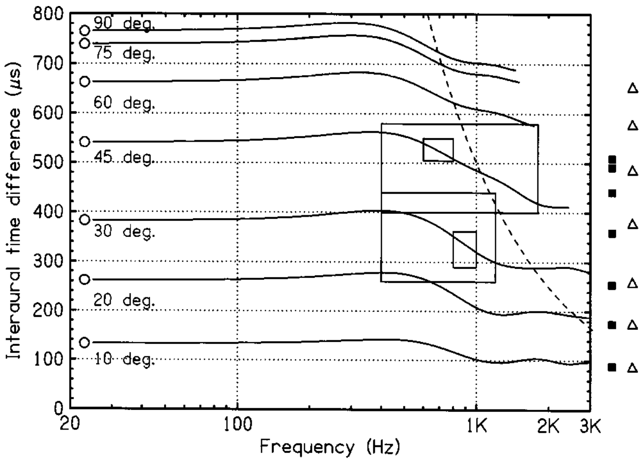

FIG. 1. Head-related dispersion according to Kuhn's spherical-head model with the ears at antipodes. This graph plots the ITD as a function of log frequency for sound incident on a spherical head from various azimuths, measured from the forward direction (nose). The dashed curve indicates ITDs that equal one-half period for a particular frequency. Points to the left of this line are lateralized as expected according to the sign of the ITD. Points on this line have ambiguous lateralization. The open circles on the left show the low-frequency limit, $\mathrm{ITD}=(3a/c)\sin(\theta_i)$. The filled squares on the right show the high-frequency limit, $\mathrm{ITD}=(2a/c)\sin(\theta_i)$. The open triangles on the right show the predictions of the Woodworth formula, $\mathrm{ITD}=(a/c)[\sin(\theta_i)+\theta_i]$, included for reference only. Rectangles show the frequency ranges used in experiment 1: Broadband: 20–3000 Hz. midband 30°: 400–1200 Hz. Midband 45°: 400–1800 Hz. Narrowband 30°: 800–1000 Hz. Narrowband 45°: 600–800 Hz.

deed, be perceptible in human binaural hearing. Figure 1 shows that the ITD depends strongly on frequency for an incident azimuth of 30°. The ITD changes by about 80 μs (from 399 to 320 μs) as the frequency increases from 500 to 1000 Hz. By comparison, it is known that listeners can detect a change in the ITD as small as 10 μs when the baseline ITD is 400 μs (Domnitz, 1973; Domnitz and Colburn, 1977). The ITD shown in Fig. 1 is the *phase delay*, which describes the delay in the fine structure of a narrowband signal. Listeners are sensitive to such ITDs in waveform fine structure at frequencies near 800 Hz, where the dispersion occurs for 30°, and even more sensitive at lower frequencies where dispersion occurs for larger azimuths. Therefore, the variation in ITD caused by head dispersion is clearly in the range of phase delays that are perceptible.

By numerical differentiation of Eq. (A1) in the Appendix we learned that head dispersion causes a frequency variation in *group delay* about 50% greater than the variation in phase delay. The group delay describes the delay of peaks and valleys in the envelope of the signal. Therefore, the interaural comparison of envelope features would be at least as inconsistent across frequency as the comparison of waveform features.

The frequency dependence of the ITD leads to binaural incoherence, characterized by a normalized interaural cross-correlation function with a peak that is less than unity. In turn, binaural incoherence leads to a broadened auditory image, an effect known as apparent source width (ASW) (Blauert and Lindemann, 1986; Ando, 1998). Listeners are extremely sensitive to the ASW introduced by small amounts of binaural incoherence (Gabriel and Colburn, 1981; Constan, 2002). Therefore, one might expect that head-related dispersion would become evident to listeners through a change in the ASW.

A logical gap in the above reasoning is that it concerns two different kinds of incoherence. The ASW experiments showing keen sensitivity to binaural incoherence have studied incoherence that occurs within a single frequency region, perhaps producing moment-to-moment variations in the lateral position of an image. By contrast, the incoherence introduced by dispersion is across different frequencies, perhaps producing different lateral positions for different frequency regions. One can view the present article as an attempt to discover whether what is learned about perceived incoherence from ASW experiments can be applied across bands to dispersion.

Although the auditory sensitivities described above suggest that listeners might be able to detect head-related dispersion, the evidence from several virtual reality experiments indicates that listeners cannot do so. Experiments by Kistler and Wightman (1992), Hartmann and Wittenberg (1996), and Kulkarni *et al.* (1999) asked listeners to discriminate between accurate HRTFs and HRTFs in which the ITDs were made frequency independent. The experiments found that it does not seem to matter perceptually if the details of ITD frequency dependence are suppressed in favor of frequency-independent ITDs. Equivalently, these experiments show that it does not matter if the true interaural phase shifts, determined experimentally from HRTFs, are replaced by phase shifts that increase linearly with frequency with a suitable slope.

The virtual reality experiments suggest that listeners cannot detect the frequency dependence of the ITD typical of human heads. However, additional work is needed to test this generalization because the virtual reality experiments have not been conclusive. First, head diffraction and corresponding virtual reality experiments do not present dispersion in isolation. Instead, head dispersion is always accompanied by an interaural level difference (ILD), and diffraction causes the ILD itself to depend on frequency. Second, the virtual reality experiments have not been optimized for revealing the effects of dispersion.

The experiments of the present article test the ability to detect dispersion when it is the only frequency-dependent variable and under conditions expected to be optimum. The experiments use headphones to produce isolated dispersion, i.e., frequency-dependent ITDs according to the spherical-head calculation without the corresponding ILDs. The experiments use noise bands with frequency ranges and simulated azimuths of incidence intended to give the listener the best chance of detecting dispersion.

II. EXPERIMENT 1: DISCRIMINATION AT TWO INCIDENT AZIMUTHS

In experiment 1, listeners compared noise bands with spherical-head-related dispersion, as shown in Fig. 1, against noise bands with ITDs that were constant (zero dispersion). Listeners had the task of recognizing the dispersive (head-delayed) noise. It was expected that our listeners would use differences in apparent source width to recognize dispersion.

A. Method

Experiment 1 featured two incident azimuths for the head-delayed noise, 30° and 45°. Figure 1 shows that both have an interesting ITD characteristic. Over the frequency range of interest (20–3000 Hz), the 30° plot exhibits the steepest change in ITD as a function of frequency while the 45° plot has the most extreme ITD variation (from 563 to 411 μs). The large variations for these two azimuths would be expected to give listeners a good chance of detection.

The experiment tested listeners at three different bandwidths, identified as broadband, midband, and narrowband. The purpose of the three was to provide the listener with different perspectives on the head-delayed interval. The broadband (BB) (20–3000 Hz) stimulus included the dispersive frequency region along with higher and lower frequencies, where the ITD was nearly constant. The midband (MB) stimulus was defined by the region where the ITD varied from its maximum value to a much lower high-frequency shelf value, presenting listeners with only the band of greatest change. Finally, the narrowband (NB) stimulus was limited to the steepest part of the ITD change, the region of greatest slope. Because the experimental signals were produced digitally, it was easy to produce precise noise bands with frequency ranges shown by rectangles in Fig. 1 with numerical values given in the caption.

Because our goal was to determine whether listeners can detect dispersion present in head-delayed noise and not present in constant-ITD noise, it was important to control for the fact that listeners could distinguish between the two noises by the lateralization of their images if the constant ITD was not chosen correctly. The correct value of constant ITD would be that value which leads to the same lateral position as the head-delayed noise. However, we did not know what the correct value was. Therefore, we used ten different equally spaced constant ITDs in random order during the course of an experiment run. The constant ITD ranges were as follows: 30° BB/MB: 260–440 μs; 30° NB: 290–362 μs; 45° BB/MB: 400–580 μs; and 45° NB: 505–550 μs. These regions are shown by rectangles in Fig. 1. For each range the largest and smallest of the ten values were outside the range of ITDs spanned by the head-delayed noise bands.

Both constant-ITD noises and head-delayed noises were generated digitally by a Tucker-Davis Technologies Array Processor, AP2. The processor constructed noises in the frequency domain, setting the upper and lower band limits and filling between them with equal-amplitude random-phase components. For a head-delayed interval, the processor then introduced the precalculated interaural phases. For a constant-ITD interval, it imposed one of the constant ITD values. Both head-delayed and constant-ITD noises were recomputed for each experimental trial by using a different basis noise waveform and adding frequency-dependent or frequency-independent delay as required. The frequency dependence of the ITD was verified with a digital delay line, signal subtractor, and a spectrum analyzer, independent of the signal generating equipment.

B. Listeners and procedure

There were four listeners, each identified by a single letter: T was a male age 24 with normal hearing and no previous experience in listening experiments. W was a male age 60 with typical middle-age high-frequency hearing loss but normal hearing in the range of the present experiments and with extensive previous listening experience. X was a male age 26 with normal hearing and previous experience. Z was a male age 27 with normal hearing and previous experience. Listeners Z and W were the authors.

During runs, listeners sat in a double-walled sound room listening with Sennheiser HD-480 II headphones. The experimental runs consisted of 100 two-interval forced-choice (2IFC) trials (ten for each constant ITD value). On each trial, the program presented the listener with two 500-ms intervals in random order. One was a constant-ITD stimulus, and the other was head delayed. Both intervals had rise/fall times of approximately 30 ms, and had simultaneous onsets/offsets in both ears.

After presenting both intervals, the program prompted for a response, and the listener pressed one of two buttons to indicate which interval was head delayed. It was expected that listeners would learn to recognize the head-delayed noise by feedback given by pilot lamps on the listener's response box after every trial. Following several training runs, listeners participated in four runs at each combination of incident azimuth (30° and 45°) and bandwidth (BB, MB, NB).

C. Results—45°

Experiment 1 found the percentage of trials in which the listener correctly identified the head-delayed interval as a function of the ITD of the alternative, namely the constant-ITD interval. The experimental results for 45° are shown in Fig. 2, one panel for each listener. Each panel includes three plots, one for each bandwidth. The data are averaged over the four runs for each bandwidth and show the percentage of trials answered correctly. The error bars on each point have an overall length of two standard errors of the mean, based on the four runs. The horizontal bars labeled ''BB/MB range'' and ''NB range'' delineate the set of ITDs included in head-delayed stimuli of those bandwidths. The dashed vertical line labeled ''LFL'' corresponds to the low-frequency ITD limit of the head-delayed noise, as represented by the open circles in Fig. 1. The dashed horizontal line indicates the 50% correct level (guessing), which is the performance expected if listeners cannot distinguish between constant-ITD and head-delayed intervals.

The percentages of correct responses in Fig. 2 show that listeners were often confused. The data for listeners W and Z tended to be a U-shaped function of the constant ITD, with best performance at the extremes of the range where the constant ITD most differed from ITDs characteristic of the head-delayed noise. These listeners achieved performance rates well in excess of 75% for extreme ITDs but only chance performance (50%) near the center of the range. The data for listeners T and X show that performance hovered around chance, exhibiting only a partial U shape.

It is of interest to compare the ITD of minimum perfor-

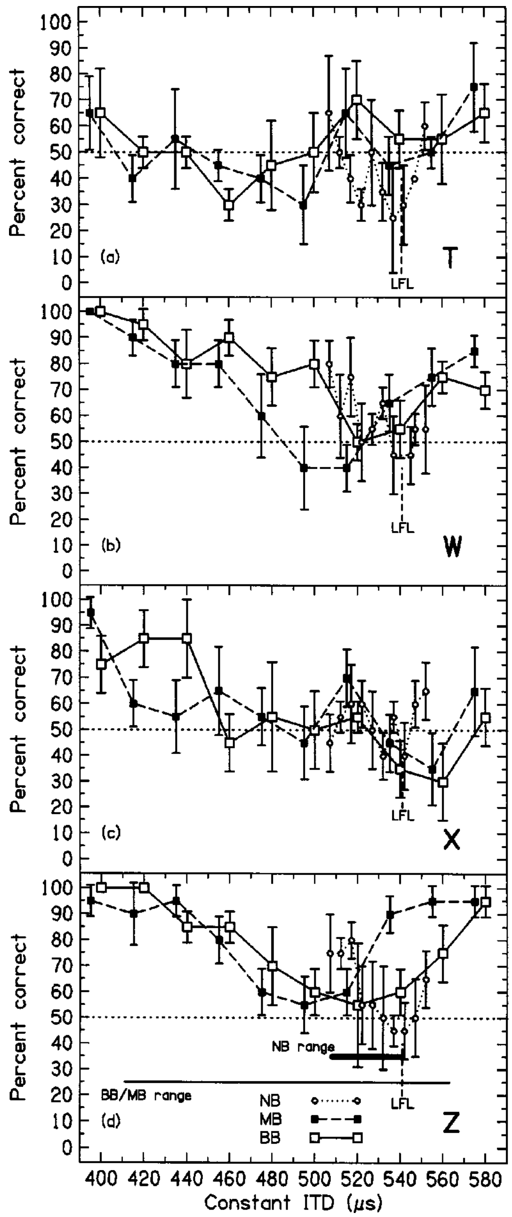

FIG. 2. Results of experiment 1 for 45° head dispersion for listeners T, W, X, and Z. This graph plots the percentage of trials in which the listener correctly identified the head-delayed stimulus versus the constant-ITD interval. The three functions represent results from broadband (BB), midband (MB), and narrowband (NB) conditions. The dashed horizontal line at 50% correct indicates chance performance (guessing). The thin horizontal line marks the range of ITDs included in the BB and MB head-delayed stimuli, while the thick horizontal line indicates those ITDs included in the NB head-delayed stimulus. The ten experimental values of constant ITD span these ranges for each bandwidth. The dashed vertical line labeled "LFL" marks the low-frequency limit ITD for a sound incident at 45°. Error bars on data points have an overall length of two standard errors.

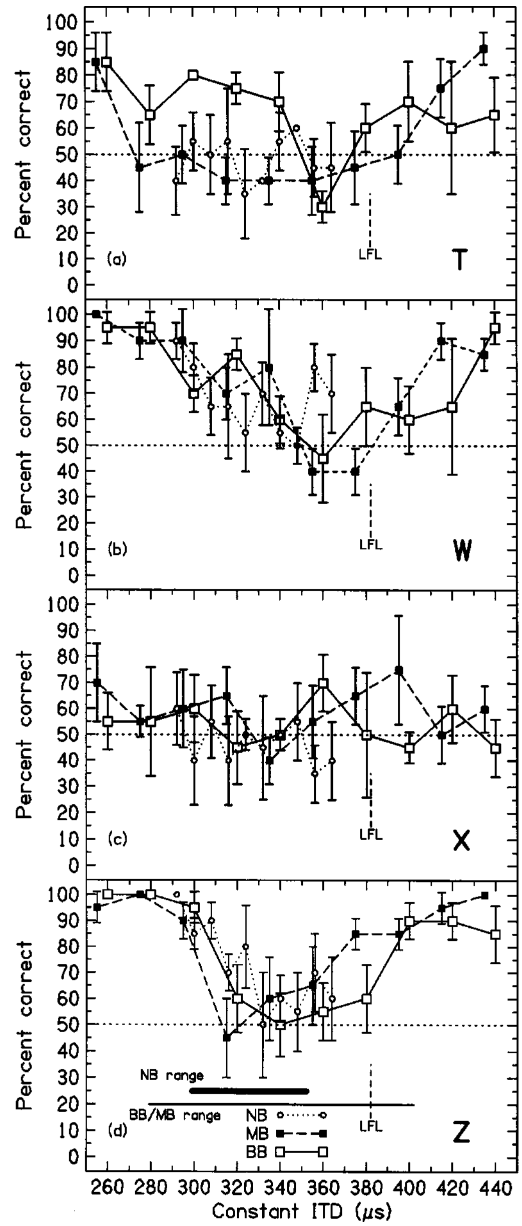

FIG. 3. Same as Fig. 2 but for a 30° head dispersion.

mance for listeners W and Z with the low-frequency limit, indicated by the vertical dashed line (LFL), especially for the broadband (BB) case because the BB noise included the low-frequency region. The data show that the minimum occurs to the left of the vertical line, i.e., the most confusable constant ITD was less than the low-frequency limiting value of head-delay. A similar conclusion can be drawn for listener T, though with less confidence. The data for listener X do not show any clear frequency dependence.

D. Results—30°

The experimental results for an azimuth of 30° are shown in Fig. 3, entirely parallel to Fig. 2 for 45°. The data for 30° resembled the data for 45° in that performance was at chance near the middle of the range of constant ITD. For listeners W and Z, the U-shape was narrower and better defined for 30° than for 45°. The data for listener T were also better described as U-shaped for 30° than for 45°.

Similar to the 45° data, the minima in the U-shaped functions did not coincide with the low-frequency limit. Instead, the BB data for listeners T, W, and Z had a minimum that was again to the left of the low-frequency limit, i.e., the most confusable constant ITD was less than the low-frequency limit of the head-delayed ITD. Because the U-shaped functions were somewhat narrower for 30° compared to 45°, the shift in the minimum away from the low-frequency limit was more convincing for 30° than for 45°.

E. Discussion

The results shown in Figs. 2 and 3 suggest a simple interpretation of experiment 1. The minima of the U-shapes near chance performance for listeners W and Z, and the generally chance performance for listeners T and X, suggest that listeners cannot distinguish between head-delayed noise and noise with a constant ITD. The high performance for listeners W and Z for ITDs at the extremes of the range suggest that these listeners made decisions based on the different lateralizations of the two kinds of noise. Specifically, noises with constant ITDs that differed greatly from those ITDs characteristic of the head-delayed noise were systematically heard to the right or the left of the head-delayed noise.

Informally, all the listeners reported that they heard lateralization differences in the noises and that their judgments were affected by that cue. Apparently listeners W and Z, who were more familiar with the experiment, made more consistent use of the lateralization cue.

III. EXPERIMENT 2: ROVED AZIMUTH

Experiment 1 suggested that listeners are unable to distinguish noise with head-related dispersion from noise with constant ITD, consistent with the results of virtual reality experiments. However, experiment 1 is not entirely satisfying because it included a useful lateralization cue, and it was evident that some listeners were influenced by that cue. The worry about experiment 1 is that in paying attention to the lateralization of the noise images, listeners may have missed more subtle differences associated with apparent source width. Therefore, we are not prepared to say that listeners cannot detect dispersion solely on the basis of experiment 1.

A better experiment would ask listeners to discriminate between head-delayed noise and noise with constant ITD in a context in which lateralization cues are of no use. Therefore, we designed experiment 2 to eliminate the usefulness of lateral position, leaving listeners with only interaural correlation as a means for discrimination. Our technique comprised several features. First, we roved the azimuth of the head-delayed noise perceptibly but slightly to prevent listeners from using an arbitrarily-small shift in perceived location as a cue. Second, we chose values of the constant ITD to be maximally confusable with the set of azimuthal locations. Third, we provided feedback after the response. The combination of the small rove in azimuth and the feedback was intended to actively discourage the listener from attempting to use laterality as a cue in the task.

A. Method

Experiment 2 used six different stimuli—three head-delayed bands associated with various azimuths and three corresponding constant-ITD bands. According to the design, each trial selected the two intervals randomly, one from each group so there would be a comparison between head-delayed and constant-ITD stimuli. In this way, the lateral positions of the two were randomized, and lateralization provided no consistent cues for the listener. It was expected that listeners would quickly learn from the feedback that small lateralization cues were useless in performing the task. Aside from the roving image positions and the adjusted constant ITDs, experiment 2 was identical to experiment 1, with the same task—identify the head-delayed interval.

TABLE I. Ideal performance in experiment 2A, equivalence of lateral position. In this two-interval task, complete confusion between head-delays at azimuths of 42°, 45°, and 49° and three constant ITDs is shown by 50% in the diagonal cells. This indicates that the noise with constant ITD appears to the left of the head-delayed noise on exactly half the trials. In addition the adjustments attempt to get the off-diagonal corners to be complementary and fairly close to zero and 100%.

	ITD1	ITD2	ITD3
42°	50%		ϵ%
45°		50%	
49°	$(100-\epsilon)$%		50%

The experiment included three azimuths, near 45°, for head-delayed stimuli based on previous experience in experiment 1. From Fig. 1 it is clear that 45° data exhibit a large overall variation in ITD, producing considerable decorrelation over a relatively small frequency range and a good chance for our listeners to distinguish head-delayed noises from constant-ITD noises.

To set up experiment 2 we first needed to choose incidence azimuths above and below 45°. We chose 49° and 42° because for both of these azimuths the ITDs differed from the ITD for 45° by about 30 μs, a small but noticeable difference. Next, we needed to choose a set of corresponding constant ITDs. According to results from experiment 1, when listeners attempted to identify the 45° head-delayed stimulus, the ITD of greatest confusion was 520 μs. That value established a starting point for a preliminary experiment, experiment 2A.

B. Experiment 2A—Equivalence of lateral position

The purpose of experiment 2A was to choose three constant ITDs that would lead to lateral positions equivalent to head-delayed lateral positions for azimuths of 42°, 45°, and 49°. In experiment 2A the listeners were asked to choose which of two noises was heard further to the left. One of the noises was head-delayed, the other had constant ITD.

The goal of finding three constant ITDs (ITD1, ITD2, and ITD3) that were lateralization-equivalent to head delays could be expressed in the form of an ideal response matrix with azimuths in the rows and constant ITDs in the columns as shown in Table I. Each cell of the matrix gives the percentage of responses indicating that the constant-ITD noise is heard to the left of the head-delayed noise. Most important, the diagonal elements of the matrix are 50% indicating complete confusion about whether the constant-ITD stimulus is to the left or right of the head-delayed stimulus. Within that 50% constraint for the diagonal elements, further centering leverage can be obtained by having the left-most constant-ITD reliably appear to the left of the right-most head-delay, and to have the right-most constant ITD appear to the right of the left-most head delay by an equal amount.

With the ideal of Table I in mind, we performed runs of 99 trials, 11 repetitions of all possible combinations of constant ITD and head delay in random order. The listeners were

TABLE II. Results of experiments 2A and 2B for three bandwidths and four listeners (T, W, X, and Z). The results of experiment 2A are shown as the best matching constant ITDs (units of μs) for three azimuths (42°, 45°, and 49°). The results of experiment 2B are the diagonal elements of the confusion matrices, percentages showing the listener's ability to distinguish between a head-delayed stimulus (azimuth) and the constant ITD. A score of 50% is expected for random guessing. Average ITDs over the four listeners and the most confusing ITD from experiment 1 (45°) are given for comparison.

Listener	42° vs ITD1	45° vs ITD2	49° vs ITD3
		Broadband (20–3000 Hz)	
	2A 2B	2A 2B	2A 2B
T	475 μs⇒49%	510 μs⇒52%	545 μs⇒55%
W	490 μs⇒59%	520 μs⇒47%	550 μs⇒53%
X	490 μs⇒53%	520 μs⇒52%	550 μs⇒52%
Z	480 μs⇒51%	510 μs⇒54%	540 μs⇒56%
AV	484 μs⇒53%	515 μs⇒51%	546 μs⇒54%
Experiment 1		520 μs	
		Midband (400–1800 Hz)	
	2A 2B	2A 2B	2A 2B
T	460 μs⇒52%	505 μs⇒48%	550 μs⇒56%
W	490 μs⇒43%	525 μs⇒54%	545 μs⇒67%
X	485 μs⇒60%	515 μs⇒51%	545 μs⇒45%
Z	480 μs⇒50%	510 μs⇒49%	540 μs⇒50%
AV	479 μs⇒51%	514 μs⇒50%	545 μs⇒54%
Experiment 1		510 μs	
		Narrowband (600–800 Hz)	
	2A 2B	2A 2B	2A 2B
T	490 μs⇒57%	535 μs⇒48%	580 μs⇒49%
W	500 μs⇒50%	530 μs⇒58%	560 μs⇒43%
X	490 μs⇒47%	520 μs⇒51%	550 μs⇒52%
Z	500 μs⇒45%	530 μs⇒48%	560 μs⇒55%
AV	495 μs⇒50%	529 μs⇒51%	562 μs⇒50%
Experiment 1		540 μs	

the same as in experiment 1. We made adjustments to the various ITD values until the listener produced two runs with results corresponding approximately to the ideal. Different listeners required slightly different constant ITDs to achieve lateralization equivalence. The final values of ITD for each bandwidth and listener from experiment 2A are given in Table II (dimensions of μs). The last line of each section of the table shows the most confusing constant ITDs from experiment 1 for comparison with the 45° results of experiment 2A. The results are very similar.

The percentages of constant-ITD-on-the-left responses for experiment 2A were collected into three-by-three confusion matrices, in all, a set of 12 arrays—four listeners times three bandwidth conditions, broadband (BB, 20–3000 Hz), midband (MB, 400–1800 Hz), and narrowband (NB, 600–800 Hz). For listeners W and Z, the arrays resembled Table I rather well. The mean of the 18 diagonal elements for these two listeners was 47% ($\pm$16), and the corner elements were equal or close to 100% and 0%. For listeners T and X, the arrays approximated the form of Table I, but the table entries all tended more toward a mean value of 50%, as might be expected from the results of experiment 1. Therefore, it is probable that the sets of constant ITDs in Table II effectively eliminated lateralization as a useable cue in discrimination experiments for both pairs of listeners. Although changes in noise image location might have been perceptible for listeners W and Z, those changes were unlikely to be of any value in performing the task because of the symmetry of the design.

C. Experiment 2B—Discrimination

After the appropriate noise signals were established in experiment 2A, the four listeners were presented with those noises, in pairs, in experiment 2B. The task in experiment 2B was the same as in experiment 1—in a randomly ordered trial, discriminate the head-delayed interval from the constant-ITD interval by any means possible. With the utility of a lateralization cue likely eliminated, it was expected that listeners could only depend on the stimulus dispersion.

Experiment 2B comprised four runs for the three bandwidth conditions (BB,MB,NB)—a total of 12 runs for each listener. Each run consisted of 99 trials, 11 occurrences of the 9 possible pairs of incident azimuth and constant ITD, presented in random order.

The results for experiment 2B were displayed in 12 arrays, as for experiment 2A, a total of 108 values. Of these, only one combination of head-delay and constant ITD led to more than 75% correct (77%) and only one combination led to more than 75% wrong (76%). The 12 arrays are summarized in the 12 lines of Table II by giving only the diagonal elements of the arrays. If listeners cannot distinguish between constant ITDs and head-delays, these elements should be 50%. The table entries are, in fact, close to 50%. None of the entries approach threshold values of 25% or 75%. The conclusion of experiment 2B was that listeners cannot distinguish between head-delayed noise and constant-ITD noise when these noises are similarly lateralized.

IV. EXPERIMENT 3—BROADBAND, WRONG HEAD

Experiments 1 and 2 showed that listeners are insensitive to dispersion resulting from head diffraction. We suspect that this insensitivity arises because the head dispersion does not introduce a detectable form of incoherence. There is another possibility, however. It is possible that listeners are insensitive to head-related dispersion because they are so familiar with it. In real life, this dispersion, and the resulting binaural incoherence, is an unavoidable consequence of having a head. Experiment 3—broadband, wrong head was performed to test the second possibility.

A. Method

To test the possibility that listeners did not detect the head-related dispersion because it was so natural, we presented listeners with dispersion that had the same magnitude and yet was unnatural. The wrong-head experiment used an inverted dispersion curve. The stimuli were created by starting with the values for 45° incidence, as shown in Fig. 1, and simply reversing the sign of the difference from the low-frequency limit. The procedure is illustrated in Fig. 4, which shows the 45° part of Fig. 1 together with the wrong-head function. The same transformation was applied to the 42° and 49° stimuli as well to produce three perceptually-separate wrong-head images. Only the broadband frequency range (20–3000 Hz) was used.

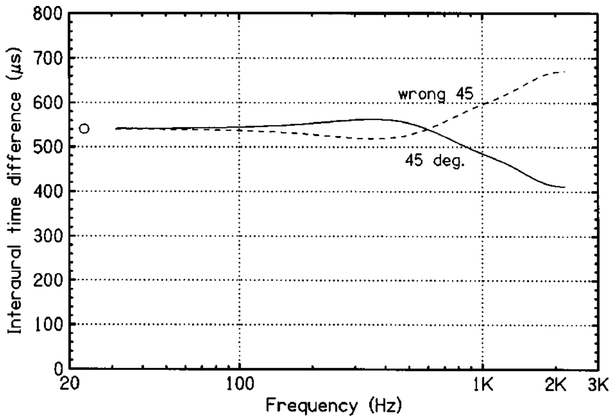

FIG. 4. The wrong-head stimulus for 45° is a reflection of the stimulus for 45° about the low-frequency limit for the ITD. Thus, the magnitude of the dispersion is the same but the signs are opposite.

Aside from changes in the stimuli, experiment 3 ran exactly the same as experiment 2, with parts A and B. Again, runs consisted of 11 presentations of each of the 9 pairings. The four listeners each did four runs in 3B.

B. Results of Experiment 3A—Equivalence of lateral position

As for experiment 2A, constant ITDs were adjusted to approach the ideal in Table I. After a series of runs established the three values of constant ITDs that were most confused with the three wrong-head delays, the percentage of judgments with the constant ITD on the left were put into three-by-three confusion matrices. Ideally the diagonal elements of these matrices should be 50%. The averages of diagonal elements actually obtained for the four listeners were T 55 (±17), W 50 (±9), X 55 (±4), and Z 48 (±5). The final constant-ITD values appear in Table III.

It is of interest to compare the optimized constant ITDs in Table III (wrong head) with the optimized constant ITDs in the broadband part of Table II (head-related) because the low-frequency limits were the same for these two cases. The constant ITDs that optimally matched the wrong-head delays were always greater, by about 8% on the average. This result indicates that the perceived equivalent ITD is not established by the low-frequency limit. Instead, the equivalent ITD seems to reflect some form of average of ITDs present in the band.

TABLE III. Results of experiments 3A and 3B (wrong head) for four listeners (T, W, X, and Z). The results of experiment 3A are shown as the best matching constant ITDs (units of μs) for three azimuths (42°, 45°, and 49°). The results of experiment 3B are the diagonal elements of confusion matrices, percentages showing the listener's ability to distinguish between a wrong-head-delayed stimulus (azimuth) and the constant ITD. A score of 50% is expected for random guessing. Average ITDs and percentages of correct identifications over the four listeners are given to summarize.

	Wrong head—Broadband (20–3000 Hz)					
	42° vs ITD1		45° vs ITD2		49° vs ITD3	
Listener	3A	3B	3A	3B	3A	3B
T	500 μs	⇒53%	545 μs	⇒46%	590 μs	⇒58%
W	530 μs	⇒21%	560 μs	⇒54%	590 μs	⇒53%
X	530 μs	⇒55%	565 μs	⇒48%	600 μs	⇒56%
Z	530 μs	⇒53%	560 μs	⇒57%	590 μs	⇒63%
AV	522 μs	⇒46%	558 μs	⇒51%	592 μs	⇒58%

C. Results of Experiment 3B—Discrimination

After choosing appropriate constant ITDs in experiment 3A, we asked listeners to discriminate between constant ITDs and head-delays in experiment 3B, parallel to experiment 2B. The results for experiment 3B were displayed in three-by-three confusion matrices as for experiment 3A. It was expected that the diagonal elements of these matrices would be about 50% because lateralizations should be most equivalent for those elements. The values of diagonal elements actually obtained appear in Table III. Except for one anomalously low value (21% for listener W) the values were close to 50% as expected. In fact, the off-diagonal elements were also close to 50%. Of 24 such elements for the four listeners, the largest was 67% and the smallest was 37%. The mean (±sd) was 52% (±9%). There was only one combination of head-delay and constant ITD that led to more than 75% correct or incorrect. The conclusion of experiment 3B was the same as for experiment 2B, namely that listeners cannot distinguish between head-delayed noise and constant-ITD noise when these noises are similarly lateralized. Therefore, the conjecture that head-related dispersion is not detected only because it is familiar is disproved.

V. LATERAL POSITION

The lateral position experiments, 2A and 3A, were necessary preliminary experiments for the discrimination experiments that followed. However, the lateral position experiments proved to have value in their own right. We used these data to find the value of the constant-ITD that best matched the simulated 42°, 45°, and 49° head-delayed functions.

Because the percentage of constant-ITD-left responses in the lateralization confusion matrices (not shown here) did not agree exactly with the ideal in Table I, we used the data in the matrices to make straight line fits. We performed best fits independently for the four experimental conditions: broadband, midband, narrowband, and wrong-head. The best-matching ITD was obtained by drawing a straight line through the two percentages closest to 50% and interpolating or extrapolating to the 50% point. The three azimuths, four experimental conditions, and four listeners led to a total of 48 such calculations.[1] The 48 best-matching ITDs are shown by open symbols in Fig. 5.

The best-matching ITDs from Fig. 5 can be compared with expectations about lateralization based on the stimulus. The dashed lines in the figure give the values of the ITDs at the highest frequency and the lowest frequency in the stimulus as well as the maximum value of ITD in the stimulus. (It is the minimum value for the wrong head.) Figure 5 shows that the best-matching ITDs do not agree with any of these limiting values except, perhaps, for the narrowbands, where everything is close together. It is particularly interesting that the best-matching ITDs do not agree with the low-frequency values in the stimulus. A similar conclusion was reached in experiments 1 and 3.

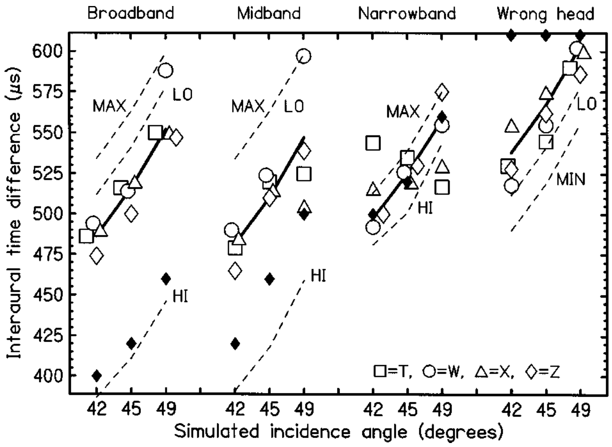

FIG. 5. Open symbols show the best-matching ITDs, i.e., values of constant ITD matching the lateralization of noise synthesized with incident azimuths of 42°, 45°, and 49° for four conditions: broadband, midband, narrowband, and broadband wrong head. The four open symbols are for the four different listeners. (Some points have been slightly displaced horizontally for clarity.) Dashed lines are labeled ''HI,'' ''LO,'' and ''MAX.'' These show the ITD at the highest frequency in the stimulus (HI), the ITD at the lowest frequency in the stimulus (LO), and the maximum value of the ITD in the stimulus (MAX). (For midband noise the MAX and LO values are the same.) (The maximum is turned into a minimum by the wrong-head manipulation.) The highest-frequency values for the wrong-head stimuli are above 600 μs. The solid line shows the weighted average of stimulus ITDs, as weighted by Raatgever's asymmetrical Gaussian function. The solid diamonds show the peaks of cross-correlation functions measured on the stimuli. Solid diamonds on the top axis are off the chart at ITD values (from left to right) 620, 660, and 700 μs.

An alternative estimate of the expected best-matching ITD based on the stimulus is the value of the lag at the peak of the cross-correlation function. The filled diamonds in Fig. 5 show this lag of the peak measured on the experimental stimuli. It is evident that the peaks of the cross-correlation functions agree with the best-matching ITDs only for the narrowbands. For the broadband and midband cases, the peaks of the cross-correlation occur at ITD values that are much lower than the best-matching constant ITD values. As might be expected, for the wrong-head case the cross-correlation peaks occur at ITD values that are much higher.

The problem with the peak in the cross-correlation function is that the cross-correlation calculation gives too much weight to the high frequencies, as noted by Kulkarni *et al.* (1999). An improved estimate would weight lower frequencies preferentially prior to cross-correlation (e.g., Bernstein and Trahiotis, 1996). Alternatively, a prediction for the best-matching ITD can be obtained from a frequency-weighted average of the stimulus ITDs. The solid line in Fig. 5 was calculated from a weighted average of the values of ITD present in the bands. The weighting function emphasized the frequencies around 600 Hz and rolled off with different slopes above and below 600. The function was an asymmetrical Gaussian:

$$u(f)=\exp[-(f/600-1)^2], \quad f\geqslant 600,$$

$$u(f)=\exp[-(f/300-2)^2], \quad f<600.$$

This function was developed by Raatgever (1980) to describe broadband noise experiments in which one-third octave bands of noise were given delays different from the rest of the noise spectrum and then increased in intensity until they dominated the perception of lateralization. The frequency bands that required the least increase were given the greatest weight in function $u(f)$. It is evident that Raatgever's weighting procedure provides an excellent fit to the lateralization results in Fig. 5. This weighting function has also been incorporated into models that have successfully accounted for the lateralization of bands of noise narrower than our broadband noise. (Stern *et al.*, 1988; Shackleton *et al.*, 1992).

VI. COHERENCE

Data from the spherical-head experiments showed that listeners could not distinguish between noise bands of constant-ITD and head-delayed noises with equivalent lateral positions. That result is somewhat surprising because the head-delayed stimuli led to a large change in ITD across frequency.

To further pursue the binaural physical properties of the stimulus we measured the coherence (defined as the height of the peak of the normalized cross-correlation function) of 1 s samples of our stimuli. The coherence values for head-delayed noises at 45° were as follows: for broadband noise, 0.963; for midband noise, 0.972; and for narrowband noise, 0.999. Values for 42° and 49° differed by less than 0.01 from the values at 45°. By contrast, the measured coherence values of our constant-ITD stimuli were all 1.000, as expected.

A. Comparative coherence

The measured values of reduced coherence can be compared with the just-detectable incoherence found in experiments designed to study the detection of incoherence *per se*. (Note: Quantitatively, one can think of incoherence as 1.0 minus the coherence.) These experiments (called ''added-noise'' experiments) begin with a common noise in both ears (reference coherence of 1.0) and then reduce the coherence by adding an independent noise to one or both ears. The just-detectable level of the independent noise immediately leads to the threshold for incoherence detection.

In such experiments, Pollack and Trittipoe (1959a,b) found that listeners could detect a decorrelation in broadband noise that reduces the coherence to 0.96. Gabriel and Colburn (1981) found detectable decorrelation when coherence was reduced to 0.975. Constan (2002) found detectable decorrelation when coherence was reduced to 0.96 for nine listeners out of nine. For two of the listeners who participated in the present study decorrelation could be detected at a coherence value greater than 0.98. These threshold coherence values are comparable to the coherence of our head-delayed noises—0.963 for broadband and 0.972 for midband. Therefore, one might reasonably expect the head-delayed incoherence to have been detected in experiments 1–3.

A possible explanation for our negative detection result lies in the fact that experiments 1–3 required coherence discrimination when the baseline ITDs were near 500 μs. Pollack and Trittipoe (1959a,b) reported some loss in resolution for images lateralized by an ITD, and the review by Durlach and Colburn (1978) noted that coherent noise begins to lose its compact character for delays greater than 1000 μs. Such an effect is expected on the basis of models of the binaural delay line in which the density of EE cells becomes progressively sparser as their characteristic delay increases. The consequent emphasis on cells with small delays has been called ''centrality'' (Stern and Trahiotis, 1995).

An experimental study of the ability to detect decorrelation as a function of ITD was made by Constan (2002). The results showed that the value of coherence required for detection did decrease slightly as the ITD increased from 0 to 500 μs, as expected, but for two listeners out of three, the reduced coherence remained above 0.97. Therefore, an appeal to the baseline ITD as an explanation for the results of experiment 1–3 was not persuasive.

B. Incoherence within bands and across bands

There is an important difference between incoherence caused by dispersion and incoherence caused by adding independent noise. Dispersion creates incoherence because different frequency regions lead to different ITDs, perhaps perceived as different directions by the listener. However, within any narrowband, the change in coherence tends to be very small. There are mathematical reasons why this is so. Over a narrowband, it can be assumed that the ITD varies linearly with frequency. If the band is centered at frequency f_0 and the extent of the variation in ITD over the band is $\Delta\tau$, then it can be shown that the peak of the normalized cross-correlation function differs from unity by an amount proportional to $(f_0\Delta\tau)^2$. Thus, incoherence varies only as the square of the dispersion, tending to make it small. Dramatic evidence of this effect occurred for the narrowbands used in our experiments. These bands were chosen to have the greatest possible concentrated variation in ITD, and yet when their cross-correlation functions were measured the coherence turned out to be 0.999 or 0.998.

The small incoherence found in the narrowbands was surprising, but it might have been predicted. In the limit of an infinitesimal bandwidth the signal becomes a pure tone, and the coherence for a pure tone is always unity. The moral of this story is that it is difficult to get much incoherence from a smooth ITD variation over a small bandwidth.

In fact, the narrowbands used in our experiments were not particularly narrow by the standards of auditory filtering. The band limits were 600 and 800 Hz, a bandwidth of 200 Hz centered on 700, for a relative bandwidth of 29%, or 0.42 oct, or 2.0 Cams.[2] Therefore, even with a large amount of dispersion there is very little incoherence within each critical band—even less than we measured in our narrowband stimuli. By contrast, when noise is made binaurally incoherent by adding a little independent noise, the incoherence is statistically identical in each critical band and equal to the overall incoherence.

VII. SUMMARY AND CONCLUSION

After presenting listeners with a range of theoretical head-shift stimuli (covering several different azimuths, three bandwidths, and one inversion), this study concludes that listeners cannot distinguish a signal with head-related dispersion from one that is perfectly coherent. Thus, it seems that listeners are insensitive to the decorrelation introduced by the presence of a spherical head. It is possible for listeners to detect dispersion based on lateral position, as suggested by Kulkarni *et al.* (1999), but when lateral position is eliminated as a useable cue listeners cannot discriminate between noises with constant ITD and noises with dispersion. Unfortunately, it is statistically difficult to demonstrate a negative result. Confidence in the conclusion reported here derives partly from the consistency of results from three rather similar experiments.

Initially, this conclusion was surprising because the experimental head-delayed stimuli had been chosen to try to maximize the role of dispersion in several ways, and the values of incoherence measured on the stimuli were comparable to known thresholds for incoherence detection. These thresholds were based on added-noise experiments in which incoherence was created by starting with diotic noise as a reference (zero incoherence) and adding independent noise. However, dispersion creates incoherence that is quantitatively evident only when computed across a frequency range that is wide compared to auditory filter widths. By contrast, the added-noise experiments create incoherence in each critical band. We conclude that listeners are sensitive to incoherence within critical bands, but are sufficiently insensitive to incoherence across bands that they cannot detect head-related dispersion.

Additional experiments studied the lateralization of dispersive noises. These experiments bear directly on the shift in the location of an auditory image caused by head dispersion in free-field listening. The shifts were found to be large enough to recommend that future work on localization or lateralization of broadband sounds cannot simply associate an ITD with direction without considering the range of the power spectrum. The low-frequency limit is not accurate enough. Further, the effective ITD cannot be estimated successfully from the peak of the cross-correlation function. Instead, the effective ITD for image location turns out to be a weighted average of ITDs actually present in the stimulus. An asymmetrical Gaussian weighting function, proposed by Raatgever (1980) for a different purpose, gives a good account of our lateralization data.

ACKNOWLEDGMENTS

Dr. Brad Rakerd wrote the software to measure cross-correlation functions, and Tim McCaskey helped write the programs to compute spherical-head-related transfer functions. The authors are also grateful to Dr. H. S. Colburn for helpful discussions of the results. This work was mainly supported by the NIDCD Grant No. DC00181. The article was written while the second author was visiting the Hearing Research Center at Boston University and partially supported by NIDCD Grant No. DC00100.

APPENDIX: ROUND-HEAD MATHEMATICS

The dispersion calculations of this article follow Kuhn (1977) in assuming a plane wave incident on a spherical head with ears separated by 180°.

Kuhn's formulation of the pressure at the surface of a sphere is as follows:

$$\left(\frac{p_i+p_s}{p_o}\right)_{r=a}=\left(\frac{1}{ka}\right)^2\sum_{n=0}^{n_{\max}}\frac{i^{n+1}(2n+1)P_n(\cos\theta)}{j_n'(ka)-iy_n'(ka)}. \tag{A1}$$

The symbols p_i, p_s, and p_o refer to incident, scattered, and free-field pressures respectively. The argument ka in this equation consists of the wave number k, defined as 2π divided by the wavelength, and the head radius a, which is nominally 0.0875 m. Function P is a Legendre polynomial. Functions j' and y' are derivatives of spherical Bessel functions and spherical Neuman functions respectively (Abramowitz and Stegun, 1964).

Angle θ is the azimuth of the source, measured with respect to the interaural axis. For antipodal ears, one can set θ to be 90° or less for the near ear, and set it to the complement $(180-\theta)$ for the far ear. Because Eq. (A1) only involves $\cos(\theta)$, evaluations for the near ear and far ear only require the calculation of Legendre polynomials for $c=\cos(\theta)$ and $-c=\cos(180-\theta)$, respectively. The usual definition of incident angle in binaural research is with respect to the forward direction, called θ_i. Angle θ in Eq. (A1) is computed from the equation: $\theta=90-\theta_i$.

Equation (A1) is a transfer function H for a given incident angle θ_i and angular frequency ω, where ω is k times the speed of sound. With subscript e indicating either left L or right R ears, the transfer function can be written in terms of magnitude and phase,

$$H_e=|H_e(\omega)|\exp[i\phi_e(\omega)]. \tag{A2}$$

The log magnitude and phase can be extracted by taking the natural log,

$$\ln H_e(\omega)=n_e(\omega)+i\phi_e(\omega), \tag{A3}$$

where n_e is the transfer function gain in nepers. Interaural properties can be completely described by the interaural transfer function, which is the ratio H_R/H_L, or

$$\ln[H_R(\omega)-H_L(\omega)]=\Delta n(\omega)+i\Delta\phi(\omega), \tag{A4}$$

where

$$\Delta n(\omega)=n_R(\omega)-n_L(\omega), \tag{A5}$$

and

$$\Delta\phi(\omega)=\phi_R(\omega)-\phi_L(\omega). \tag{A6}$$

The interaural time difference ($\text{ITD}=\Delta t$) then is the phase delay for a sine of frequency ω, $\Delta t=\Delta\phi/\omega$.

1. Our stimuli

Stimuli used in this paper always took $\Delta n=0$ because the focus was on the ITD and its frequency dependence. Our experiments contrasted two forms of ITD, head-delayed—as computed from the phase shifts in Eq. (6)—and constant, $\Delta t=\Delta t_0$. The constant ITD condition was called "linear phase" by Kulkarni *et al.* (1999) because it leads to an interaural phase that is a linear function of frequency, $\Delta\phi=\Delta t_0\omega$. This condition was also used by McKinley and Ericson (1997).

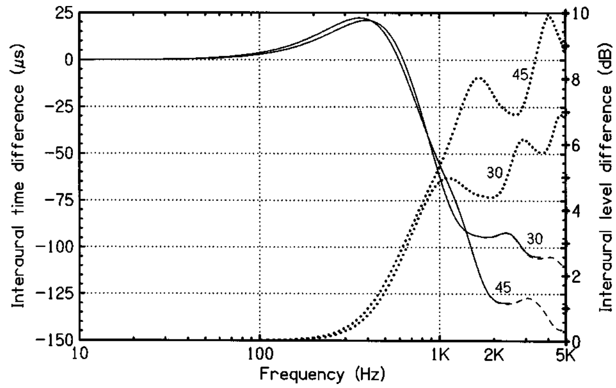

FIG. 6. Interaural time difference (ITD or Δt) for 30° and 45°. The ITD shown by the solid line is from the imaginary part of the log of the interaural transfer function [Eq. (A6)]. The ITD shown by the dashed line is calculated from the Hilbert transform of the real part of the log of the interaural transfer function [Eq. (A7)]. The low-frequency limit is subtracted off in both calculations. The agreement between the two calculations is perfect, evidence that the interaural transfer function is minimum phase. The dotted lines show the interaural level difference (ILD or Δn), the real part of the log of the interaural transfer function. These ILDs were used in the Hilbert transform calculation.

2. Minimum phase

The concept of minimum phase is not particularly helpful in the context of our work. It is discussed here in order to make contact with previous work (Kulkarni *et al.*, 1999; Kistler and Wightman, 1992). Minimum phase values of phase shifts for left and right ears, ψ_L and ψ_R, are Hilbert transforms of n_L and n_R (e.g., Hartmann, 1997, pp. 580 ff). Because the Hilbert transform is a linear operation, the minimum phase function for the interaural phase shift is

$$\Delta\psi=\psi_R-\psi_L=\mathcal{H}(\Delta n), \tag{A7}$$

where $\mathcal{H}$ is the Hilbert transform operator.

Previous authors have studied interaural phase shifts of the form

$$\Delta\phi=\Delta\psi+\omega t_0, \tag{A8}$$

i.e., minimum phase plus linear phase (frequency-independent delay), where $\Delta\psi$ is calculated from a measured log magnitude HRTF. This form has been justified because it is believed that the HRTFs are well described as minimum phase plus linear phase (e.g., Mehrgardt and Mellert, 1977). In our case, Eq. (A8) is exact because the spherical head transfer functions, as described by Eq. (A1) are minimum phase. To illustrate that idea, Fig. 6 shows the ITD for 30° and 45° calculated in two ways. First (solid line), the ITD was calculated from Eq. (A6), essentially the difference of left and right phase shifts calculated from Eq. (A1). Second (dashed line), the ITD was calculated from the Hilbert transform of the log magnitude of the interaural transfer function, i.e., from Eq. (A7). (In both cases, the low-frequency limit of the ITD was subtracted off.) The agreement between the two different kinds of calculation is so good that the dashed line cannot be seen except at the high frequencies where the solid line stops, somewhere above 2 kHz.

In connection with minimum phase calculations, a semantic difficulty has arisen when Eq. (A8) is used to describe the interaural time difference by dividing by ω,

$$\Delta t = \Delta \phi/\omega = \Delta \psi/\omega + t_0. \tag{A9}$$

This ITD, Δt, has been described in the literature as frequency independent. In fact, it is only the t_0 part that is frequency independent. The minimum phase part $\Delta\psi/\omega$ has an important frequency dependence that cannot be ignored.

[1]Of these 48 fits, 23 were interpolated, 11 were extrapolated, and 7 were exactly on 50%. There were seven comparisons that were not monotonic and defied the fitting procedure. For those seven values we took the values of ITD1, ITD2, or ITD3 from the stimuli themselves because the course of the equivalence-of-position experiments leads naturally to these associations.

[2]Auditory filter units of *Cams* come from integrating the reciprocal of the critical bandwidth, as measured on the banks of the River Cam in England. (Glasberg and Moore, 1990; Moore, 1995). The integral is not hard to do (Hartmann, 1997, p. 251). The term "Cam" is preferred by the authors as a name for units on this particular critical band scale. It is preferred over the more widely used "ERB" (stands for equivalent rectangular bandwidths) because all critical band units since the time of Barkhausen (1926) and Fletcher (1938) have been equivalent rectangular bandwidths.

Abramowitz, M., and Stegun, I. (**1964**). *Handbook of Mathematical Functions*, National Bureau of Standards Applied Math Series No. 55 (USGPO, Washington, DC).

Ando, Y. (**1998**). *Architectural Acoustics* (Springer, New York).

Barkhausen, H. (**1926**). "Ein neuer Schallmesser für die Praxis," Z. Tech. Phys. (Leipzig) **7**, 599–601.

Bernstein, L. R., and Trahiotis, C. (**1996**). "The normalized correlation: Accounting for binaural detection across center frequency," J. Acoust. Soc. Am. **100**, 3774–3784.

Blauert, J., and Lindemann, W. (**1986**). "Auditory spaciousness: Some further psychoacoustic analyses," J. Acoust. Soc. Am. **80**, 533–542.

Brungart, D. S. (**1999**). "Auditory localization of nearby sources. III. Stimulus effects," J. Acoust. Soc. Am. **106**, 3589–3602.

Brungart, D. S., and Rabinowitz, W. M. (**1999**). "Auditory localization of nearby sources—HRTFs," J. Acoust. Soc. Am. **106**, 1465–1479.

Brungart, D. S., Durlach, N. I., and Rabinowitz, W. M. (**1999**). "Auditory localization of nearby sources II: Localization of a broadband source," J. Acoust. Soc. Am. **106**, 1956–1968.

Constan, Z. A. (**2002**). "All Things Coherence," doctoral thesis, Michigan State University, unpublished.

Domnitz, R. (**1973**). "The interaural time jnd as a simultaneous function of interaural time and interaural amplitude," J. Acoust. Soc. Am. **53**, 1549–1552.

Domnitz, R., and Colburn, H. S. (**1977**). "Lateral position and interaural discrimination," J. Acoust. Soc. Am. **61**, 1586–1598.

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural Phenomena," *Handbook of Perception*, Vol. 4, edited by E. Carterette (Academic, San Francisco), pp. 365–466.

Fletcher, H. (**1938**). "Loudness, masking and their relation to the hearing process and the problem of noise measurement," J. Acoust. Soc. Am. **9**, 275–293.

Gabriel, K. J., and Colburn, H. S. (**1981**). "Interaural correlation discrimination: I. Bandwidth and level dependence," J. Acoust. Soc. Am. **69**, 1394–1401.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched noise data," Hear. Res. **47**, 103–138.

Hartmann, W. M. (**1997**). *Signals, Sound and Sensation* (AIP, Springer-Verlag, New York).

Hartmann, W. M., and Wittenberg, A. (**1996**). "On the externalization of sound images," J. Acoust. Soc. Am. **99**, 3678–3688.

Kistler, D. J., and Wightman, F. L. (**1992**). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," J. Acoust. Soc. Am. **91**, 1637–1647.

Kuhn, G. F. (**1977**). "Model for the interaural time differences in the azimuthal plane," J. Acoust. Soc. Am. **62**, 157–167.

Kuhn, G. F. (**1979**). "The pressure transformation from a diffuse sound field to the external ear and to the body and head surface," J. Acoust. Soc. Am. **65**, 991–1000.

Kuhn, G. F. (**1982**). "Towards a model for sound localization," in *Localization of Sound: Theory and Applications*, edited by R. W. Gatehouse (Amphora, Groton, CT), pp. 51–64.

Kuhn, G. F. (**1987**). "Physical acoustics and measurements pertaining to directional hearing," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer, New York), pp. 3–25.

Kuhn, G. F., and Guernsey, R. M. (**1983**). "Sound pressure distribution about the human head and torso," J. Acoust. Soc. Am. **73**, 95–105.

Kulkarni, A., Isabelle, S. K., and Colburn, H. S. (**1999**). "Sensitivity of human subjects to head-related transfer-function phase spectra," J. Acoust. Soc. Am. **105**, 2821–2840.

McKinley, R. L., and Ericson, M. A. (**1997**). "Flight demonstration of a 3-D audio display," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. A. Anderson (Erlbaum, Mahwah, NJ), pp. 683–699.

Mehrgardt, S., and Mellert, V. (**1977**). "Transformation characteristics of the external human ear," J. Acoust. Soc. Am. **61**, 1567–1576.

Moore, B. C. J. (**1995**). "Frequency analysis and masking," in *Hearing, Handbook of Perception and Cognition*, 2nd ed., edited by B. C. J. Moore (Academic, San Diego).

Pollack, I., and Trittipoe, W. J. (**1959a**). "Binaural listening and interaural noise cross correlation," J. Acoust. Soc. Am. **31**, 1250–1252.

Pollack, I., and Trittipoe, W. J. (**1959b**). "Interaural noise correlations: Examination of variables," J. Acoust. Soc. Am. **31**, 1616–1618.

Raatgever, J. (**1980**). "On the binaural processing of stimuli with different interaural phase relations," thesis, Dutch Efficiency Bureau, Pijnacker, Delft, The Netherlands.

Shackleton, T. M., Meddis, R., and Hewitt, M. J. (**1992**). "Across frequency integration in a model of lateralization," J. Acoust. Soc. Am. **91**, 2276–2279.

Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (**1988**). "Lateralization of complex binaural stimuli: A weighted image model," J. Acoust. Soc. Am. **84**, 156–165.

Stern, R. M., and Trahiotis, C. (**1995**). "Models of naural Interaction," in *Handbook of Perception and Cognition—Hearing*, edited by B. Moore (Academic, San Diego), pp. 347–386.

Strutt, J. W. (**1907**). "On our perception of sound direction," Philos. Mag. **13**, 214–232.