Origin of frequency dependence of interaural time difference

Makoto Otani¹,*; Tatsuya Hirahara² and Daisuke Morikawa²

¹Graduate School of Engineering, Kyoto University, Kyoto daigaku-Katsura, Nishikyo-ku, Kyoto, 615–8540 Japan
²Faculty of Engineering, Toyama Prefectural University, 5180 Kurokawa, Imizu, Toyama, 939–0398 Japan

(Received 18 February 2021, Accepted for publication 5 April 2021)

Abstract: The interaural time difference (ITD) plays an important role in spatial hearing, particularly in azimuthal localization of sound images. Although the ITD is essentially determined by the geodesic distance between two ears, researchers have reported that the ITD is greater for lower frequencies. However, the origin of this frequency-dependence has not been revealed. This study investigates how the ITD is physically characterized to have a frequency-dependent nature by conducting measurements and numerical simulations. Dummy head measurements show that the ITD varies with frequency because the apparent propagation time to the ipsilateral ear decreases for low frequency. Dummy head simulations confirmed this phenomenon and revealed that the apparent propagation time decreases because of a sound pressure phase shift due to reflections from the head. Circular plate simulations revealed that the circular profile including its lateral surface and edge produces reflections that are relevant to the phase shift, yielding the frequency-dependence of the apparent propagation time. Furthermore, rigid sphere simulations showed that such reflections are produced even by smooth convex surfaces without clear-cut edges. These results strongly suggest that a major factor in the production of the frequency-dependence of ITDs is backscatter diffractions from convex surfaces of the head and the pinna.

Keywords: Interaural time difference, Frequency dependence, Apparent propagation time, Backscatter diffraction

1. INTRODUCTION

Binaural cues such as interaural time differences (ITDs) and interaural level differences (ILDs) play an important role in the azimuthal localization of sound images [1]. The ITD, which is the difference in the propagation time between two ears, provides essential information, particularly at frequencies below 1,400 Hz, whereas ILD provides essential information at frequencies above 1,400 Hz [2].

Woodworth proposed an ITD prediction model that determines an ITD by an interaural difference of path length along geodesic lines on the surface of the head and the ambient speed of sound [3]. This results in frequency-independent ITDs. However, ITDs have been reported to be frequency-dependent, i.e., ITDs are greater at lower frequencies [4–6]. To resolve the differences between Woodworth’s model and the frequency dependence of ITDs suggested by Abbagnaro [4], Kuhn conducted an empirical and analytical investigation [7]. Kuhn showed that Woodworth’s model reasonably parallels high-frequency ITDs measured using a manikin. An analytical solution of a scattering sound field around a rigid sphere, e.g. Rzhevkin’s solution [8], is able to reasonably reproduce low-frequency ITDs. Kuhn also showed that low- and high-frequency ITDs can be expressed as $3(a/c_0)\sin \theta_{inc}$ and $2(a/c_0)\sin \theta_{inc}$, respectively, where $a$ is the radius of the sphere or equivalent head radius, $c_0$ is the ambient speed of sound, and $\theta_{inc}$ is the angle of incidence between the median plane of the head and incident plane wave direction.

Although the frequency dependence of ITDs has been investigated empirically and theoretically, very little is known about why ITDs differ between low and high frequencies. Kuhn suggested that the larger ITDs at low frequencies may be explained by a larger equivalent radius of the head at low frequencies; that is, at low frequencies, sound waves must travel over protruding objects such as pinnae and the nose while being attenuated sufficiently to be neglected relative to the direct wave at high frequencies [7]. In addition, some researchers have reported the effect
of the pinnae in animal models [9–11] indicating that the ITD is larger when the model has pinnae [11]. However, as reported in [12], ITD measurement for $\theta_{\text{pin}} = 90^\circ$ indicates that the apparent propagation time from the sound source shows a frequency dependence not only at the contralateral ear but also at the ipsilateral ear. In fact, this result shows that the apparent propagation time demonstrates an even stronger frequency dependence at the ipsilateral ear than at the contralateral ear. This result implies that the larger equivalent radius of the head at low frequencies, as indicated by Kuhn, cannot completely explain the reason for the frequency dependence of ITDs, because, for the ipsilateral ear, sound waves do not need to travel along the head.

In this study, measurements and numerical simulations using a dummy head are conducted to clarify what produces the frequency dependence of ITDs. Numerical simulations utilizing the boundary element method (BEM) are also performed with the same head shape to validate measurement and simulation results and investigate the frequency dependence of the apparent propagation time and ITD. Furthermore, a circular plate model and a rigid sphere model are employed to facilitate further discussion of the physical characteristics of the phenomenon.

2. MEASUREMENT

2.1. Methods

Measurements were performed two times in two different anechoic chambers. One is at Industrial Research Institute of Ishikawa, Japan, and another is at Kyoto University, Japan. Hereafter, these measurement sites are labeled as IRII and KU, respectively. It should be noted that the anechoic chamber at KU is less reflective than that at IRII.

Figure 1 illustrates the measurement system at KU. For the measurement at KU, both ears of the dummy head was located at a height of 1,850 mm from the floor. A loudspeaker (TD508MK3, Eclipse) was located on the interaural axis of the dummy head at a distance of 1,070 mm at KU from the left ear. A small reference microphone $M_{\text{SP}}$ (ECM77B, SONY) was located at a distance of 70 mm from the front of the loudspeaker. Two electret condenser microphones (ECM-77B, SONY) $M_L$ and $M_R$, embedded in ear plugs made from silicone impression material, were installed in both ear canals of the dummy head, as shown in Fig. 2(b). An amplifier (Evolution 50A, Creek) was used to amplify a computer generated input signal to the loudspeaker. Microphone amplifiers (1021, Earthworks) were used to amplify output signals from the microphones. Signals were D/A and A/D converted using AD/DA converter (Fireface UFX, RME) at a sampling rate of 192-kHz and 24-bit quantization.

The measurement at IRII employed almost the same experimental setup as one at KU, except using a different loudspeaker (MG105D09-08, Vifa), amplifier (TA-F501, SONY) and AD/DA converter (UA-101, Roland), instead. In addition, the reference microphone $M_{\text{SP}}$ was located at a distance of 60 mm from the front of the loudspeaker.

A preliminary measurement confirmed that signals received by the three microphones were synchronized when the three microphones were placed at the same position. Temperature, relative humidity, and atmospheric pressure were recorded at the time of measurement. The test signal was a pure tone burst with duration of five or ten periods. For the measurement at IRII, the tone burst signals were not tapered in order to observe waveform deformations of measured signals at their rises, whereas, for the measurement at KU, they were tapered at their onsets and offsets for accurate estimation of ITDs. In this study, the origin of the coordinate system is located at the midpoint between the centers of the entrances to the two ear canals, at which the two microphones $M_L$ and $M_R$ are located. In this coordinate system, the loudspeaker, $M_L$ and $M_{\text{SP}}$ are located at the azimuth of 90°, and $M_R$ is located at the azimuth of $-90^\circ$.
2.1.1. Dummy head

The dummy head used in this study was developed using magnetic resonance imaging (MRI) and three-dimensional (3D) printing (stereolithography). Cross-sectional images of a subject’s head were captured using an MRI scanner (Marconi Magnex Eclipse 1.5 T, Shimadzu) at ATR-BAIC (Brain Activity Imaging Center, Advanced Telecommunications Research Institute International), Japan. The surface geometry of the head was extracted from the images to yield a 3D polygon model of the subject’s head. The head was replicated from the polygon model as a hard-resin dummy via 3D printing. Figures 2(a) and 2(b) show the 3D polygon model and dummy head, respectively.

2.1.2. Analysis

Let the signals received at the three microphones $M_{S}$, $M_{L}$, and $M_{R}$ be $S_{ref}(t)$, $S_{L}(t)$, and $S_{R}(t)$, respectively. The propagation times from the sound source to both ears, $\tau_{L}$ and $\tau_{R}$, were calculated by determining the delay time that yields the maximum value of a cross correlation function between $S_{ref}(t)$ and $S_{L}(t)$ and between $S_{ref}(t)$ and $S_{R}(t)$, respectively. Subsequently, ITDs were obtained as differences between $\tau_{L}$ and $\tau_{R}$, namely $\tau_{R} - \tau_{L}$.

Figures 3(a)–3(c) illustrate the received signals at the three microphones for non-tapered tone bursts at frequencies of 3 kHz, 1 kHz, and 500 Hz, respectively, from the measurement at IRII. Amplitudes of the $S_{L}(t)$ and $S_{R}(t)$ are normalized to see the waveforms in detail, i.e. actual amplitudes of $S_{L}(t)$ and $S_{R}(t)$ are of course smaller than $S_{ref}(t)$. As shown in Fig. 3(a) for 3 kHz, sounds reflected from the room, presumably from the floor grid, were observed starting at approximately 8 ms and later, although the measurements were performed in the anechoic chamber. Such reflected sounds were superimposed on the direct sound for 1 kHz and 500 Hz as shown in Figs. 3(b) and 3(c), respectively, which deformed the received signals $S_{L}(t)$ and $S_{R}(t)$. However, such reflected sounds do not have a prominent effect on $S_{ref}(t)$, because $M_{SP}$ is located close to the loudspeaker. Thus, at low frequencies, cross correlation could not yield the correct delay time.

Figures 3(b) and 3(c) also show that, for 500 Hz and 1 kHz, the first peaks of $S_{L}(t)$ and $S_{R}(t)$ have smaller amplitudes than the following peaks. Considering that such distortion is also observed in $S_{ref}(t)$, we can say that this is due to a transient response of the loudspeaker. However, in addition, the first period of the $S_{L}(t)$ waveform for 500 Hz and 1 kHz is further deformed. It has a steep onset and a jagged first peak. These deformations cannot be due to the transient response of the loudspeaker.

Figures 4(a)–4(c) exhibit the received signals from the measurement at KU, for which tapered ten-cycle tone bursts were employed as test signals. Similar to the measurement at IRII, reflected sounds are observed after the direct sounds, which are prominently shown especially for $S_{R}(t)$ at 3 kHz as shown in Fig. 4(a). Although such reflected sounds are superimposed on the direct sound for cases of the lower frequencies, the amplitudes of reflected sounds are smaller than those at IRII. Therefore, $\tau_{L}$, $\tau_{R}$, and ITD are derived from the measurement results at KU.

2.2. Results

Figure 5 illustrates the frequency characteristics of $\tau_{L}$, $\tau_{R}$, and ITD for a frequency range from 200 Hz to 8 kHz, which are derived from the measurement at KU employing tapered ten-cycle tone bursts. The filled circles denote measured values obtained by cross correlation. Generally, $\tau_{L}$ and $\tau_{R}$ decrease with lower frequencies at frequencies higher than 250 Hz. However, at frequencies below 200 Hz, they are prominently large. This might occur because reflected sound was superimposed on the direct sound. Furthermore, $\tau_{R}$ is much greater at 5 and 6 kHz than other neighboring frequencies, thereby resulting in greater values.
of ITD. These deviations of $\tau_R$ at 5 and 6 kHz are attributable to an erroneous estimation due to low signal to noise ratio at these frequencies where head-related transfer functions demonstrate prominent spectral notches at a contralateral ear, the right ear in this case. Figure 6 depicts the frequency characteristics observed at the left and right ear of the dummy head derived from the numerical simulation, described later in the Sect. 3. The figure demonstrates that, at the right ear (contralateral), there is a spectral notch around 5 and 6 kHz where the sound pressure level is smaller compared to lower frequencies. Such spectral notch leads to the erroneous estimation of $\tau_R$ and, subsequently, the ITD at these frequencies. In Fig. 5, the solid lines represent regression lines for the middle frequency range (250 Hz–1.25 kHz) and high frequency range (1.25 kHz–8 kHz) calculated from measured $S_L(t)$ and $S_R(t)$. A regression line was not calculated for the low frequency range (150 Hz–250 Hz), because the delay time was not computed correctly due to the superimposed reflected sound at these frequencies as mentioned in the previous subsection. Slopes of these regression lines were calculated as $\Delta t / \Delta \log_{10} f$. The ratio of the slope of $S_L(t)$ to that of $S_R(t)$ was 1.68 and 1.17 for the middle- and high-frequency ranges, respectively.

The ITD was approximately 0.75 ms at frequencies above 2 kHz, whereas it was greater than 0.8 ms below 600 Hz. This result agrees well with those of Kuhn [7] and Suzuki et al. [5] confirming that the ITD is larger at lower frequencies. In addition, the results show that $S_L(t)$ and $S_R(t)$ are smaller at lower frequencies except for frequencies below 300 Hz. In particular, at frequencies over 250 Hz, $S_L(t)$ decreases markedly with decreasing frequency, which ultimately results in a larger ITD at lower frequencies.

Figures 3(b) and 3(c) show that, compared to $S_{\text{ref}}(t)$ and $S_R(t)$, the sound pressure of $S_L(t)$ rapidly rises at its beginning and the entire waveform of $S_L(t)$ is shifted to the left, thereby resulting in a smaller apparent propagation time, namely, a greater apparent speed of sound at lower frequencies. As described in the following section, such deformation of the waveform of $S_L(t)$ is caused by sounds...
reflected from the head that are superimposed on the direct sound.

3. NUMERICAL SIMULATION

The measurement results show that larger ITDs were observed at lower frequencies; this phenomenon may be explained by $\tau_L$ at the left ear, which is on the source side, varying with frequency. Further investigation was performed by utilizing the BEM simulations. The BEM is capable of reproducing sound wave behavior in ideal conditions and observing direct and reflected sound waves separately.

3.1. Methods

The BEM was applied to the wave equation whose boundary conditions are given by the surface geometry of the head. According to the Helmholtz integral equation which is derived from the wave equation, a sound pressure field around an object is determined by sound pressures on the object’s surface. In the BEM simulation, firstly, sound pressures on the object’s surface are calculated by solving simultaneous equations whose coefficient matrices are determined from discretized surface geometry and its boundary conditions, namely, the acoustic impedance or admittance for a certain frequency. Subsequently, sound pressures at arbitrary positions in the sound field are obtainable from the calculated surface sound pressures. The direct and indirect components of sound pressures are calculated respectively as the free-field Green’s function and as a summation of effects from all the discretized object’s surface elements, thereby enabling a separate observation of direct and reflected sound pressures. Details of the BEM can be found in Otani et al. [13].

The surface mesh of the 3D polygon model, shown in Fig. 2(a), was optimized for BEM simulation. It has more than six elements per wavelength for frequencies up to 15 kHz with approximately 3-mm element length in average. The entrances of the ear canals were blocked. The receiving points were located at a distance of approximately 1 mm from the surface of the blocked ear canal entrances. A point source was located on the interaural axis at a distance of 1 m from the receiver at the left ear. In the simulation, the speed of sound was 340 m/s, which is almost the same as that in the measurement. The surface of the head was assumed to be acoustically rigid. Frequency responses at the receiving points were calculated at intervals of 86 Hz between 86 Hz and 20 kHz, which yielded 512-pt impulse responses with a sampling rate of 44.1 kHz.

Tapered tone-burst signals consisting of a sine wave with a duration of five periods were convolved with the calculated impulse responses to yield the tone-burst responses at both ears. Thus, $\tau_L$ and $\tau_R$, and subsequently the ITD ($= \tau_R - \tau_L$), were estimated by determining the maximum of the cross correlation of the source signal and tone-burst responses. To increase temporal resolution, 8-time upsampling was applied to the source signal and tone-burst responses in advance.

3.2. Results

Figure 7 shows $\tau_L$, $\tau_R$, and ITD from the simulation as a function of frequency. The filled circles indicate measured ITD values obtained by cross correlation, from the measurement at KU. The measured and simulated ITDs agree well at frequencies from 400 Hz to 7 kHz except 5 and 6 kHz.

The results show that the simulated $\tau_R$ decreases as the frequency decreases from 3 kHz to 250 Hz; the simulated $\tau_L$ decreases rapidly as the frequency decreases from 3 kHz to 400 Hz, while exhibiting approximately the same slope as that of the simulated $\tau_R$ at other frequencies. As a result, the ITD increases as the frequency decreases from 3 kHz to 400 Hz, whereas it is almost constant at other frequencies.

Both the measurement and simulation results show the same ITD frequency characteristics, i.e., the ITD increases as the frequency decreases. Reflections from the head are likely to have an effect on this phenomenon. Therefore, direct and indirect components of impulse responses were...
simulated separately to clarify how the frequency dependence of ITD occurs.

Figure 8(a) demonstrates waveforms of the simulated impulse response (direct and indirect) and its direct and indirect components, i.e., reflected sounds from the head. Figure 8(b) depicts magnified waveforms between 2.8 and 3.0 ms in order to focus on their rises. These figures exhibit that the amplitude of the first peak of the indirect component is greater than that of the direct sound, thereby leading to a steeper rise of the indirect component than the direct component. Figure 9 shows waveforms of a tone-burst response and its direct and indirect components at 500 Hz, respectively obtained by convolving non-tapered tone bursts with the impulse response and its direct and indirect components. This figure shows that the direct and indirect waveforms rise almost simultaneously; however, the waveform of the indirect component is prominently deformed at its rise and the peaks of its subsequent waveform appear earlier than those of the direct waveform, which reflects the steeper rise of the indirect component of the impulse response shown in Fig. 8(b). As a result, the peaks of the tone-burst response, which is the sum of the direct and indirect waveforms, also appear earlier than those of the direct waveform. Therefore, when estimating propagation times by the maximum of cross-correlation function, the propagation time is estimated to be smaller than that in the case without a head, which leads to a frequency dependence of the ITD. It should be noted that, as mentioned in the previous section and shown in Fig. 3(c), the same type of distortion was also observed in the measured $S_L(t)$ for 500 Hz.

Figure 10 shows the phase differences and amplitude ratios of the indirect and direct components. The results demonstrate that, at frequencies below 3 kHz, the indirect component is out of phase with the direct component between $-50^\circ$ and $-90^\circ$, which shifts the peak position of the tone-burst response to an earlier time. The results also demonstrate that, at frequencies below 3 kHz, the amplitude ratio decreases monotonically with decreasing frequency. As illustrated in Fig. 7, the variations of simulated $\tau_L, \tau_R$ and ITDs are quite small below 400 Hz. This is because, in this frequency range, the indirect components, i.e., the reflections from the head, have relatively small amplitude and, therefore, do not have a prominent influence on the phase of the tone-burst response.

4. CIRCULAR PLATE MODEL SIMULATION

The measurements and numerical simulations using the dummy head, as described above, revealed that the apparent propagation time from the sound source to the ipsilateral ear varies with frequency, which leads to a frequency dependence of the ITD. This phenomenon
occurs because sounds reflected from the head shift the phase of the tone-burst response. However, it is unknown which part of the head produces the reflection relevant to this phenomenon. Therefore, a simplified head shape, i.e., a circular plate, was employed for further investigation.

4.1. Methods

Boundary element models of a circular plate with various thicknesses \( d \) and radii \( r \) were developed. Figure 11 shows the configuration of the source and receiver. The source and receiver were located at distances of 1,000 mm and 1 mm from the surface of the circular plate, respectively. Other numerical conditions were the same as those in the simulation described in the previous section. Five circular plates were used: \((r, d) = (75, 2), (75, 10), (75, 20), (100, 10), \) and \((150, 10)\), where 75-mm radius roughly corresponds to the radius of the head; the variation in the thickness is to investigate an effect of circular profile of the plate. BEM models for these circular plates have approximately 3 mm of element length in average. For the model of \((r, d) = (75, 2)\) having 2-mm plate thickness, the element lengths for the lateral surface are smaller than 3 mm, for appropriate shaping of the model. The procedure outlined in the previous section was used to estimate the propagation time \( \tau \) from the source to the receiver.

4.2. Results

4.2.1. Comparison between head and circular plate

Figure 12(a) shows the estimated propagation time \( \tau \) for circular plate models with a fixed thickness \( (d = 10 \text{ mm}) \) and \( r = 75, 100, \) and 150 mm, and \( \tau_0 \) for the head model. The results for both the circular plate and head models show that in a high-frequency range (above several kHz), the apparent propagation time does not vary prominently with frequency and is almost identical to the actual propagation time, approximately 2.9 ms for the circular plate model. In the middle-frequency range (several hundreds Hz to several kHz), the apparent propagation time decreases with decreasing frequency. In the low-frequency range (below several hundreds Hz), the apparent propagation time does not vary prominently with frequency.
frequency, but is notably less than the actual propagation time. These results confirm that, similar to the head model, the apparent propagation time also decreases at lower frequencies for the circular plate model. Furthermore, frequency bounds for the above-mentioned characteristics of apparent propagation time vary with the radius of the circular plate. Among the examined radii, i.e., 75, 100, and 150 mm, the circular plate with $r = 150$ mm showed the closest frequency bounds to those of the head model (2 kHz and 400 Hz); hence, a smaller radius leads to higher frequency bounds. Such variation in the frequency bounds corresponds to a variation in the ITD transition frequency observed in the various head sizes.

Figure 12(b) shows the modeled impulse responses at the receiver for the circular plate models with radii of 75, 100, and 150 mm. The results show that, regardless of the radius, the impulse responses have a negative peak after a positive peak (i.e., a direct sound), indicating that these negative peaks make the apparent propagation time less than the actual propagation time. These negative peaks appear at 0.23 ms ($r = 75$), 0.30 ms ($r = 100$), and 0.48 ms ($r = 150$) after the positive peak appears. These are proportional to the distance difference between the direct path and the indirect path via the edge of the circular plate. Accordingly, these negative peaks are suggested to be reflections from the edge or circular profile of the circular plate.

Figures 13(a)–13(c) illustrate the waveforms of the direct and indirect components of the tone-burst responses for 3 kHz, 1 kHz, and 500 Hz, respectively. Similar to those for the head model shown in Fig. 9, Fig. 13(c) for 500 Hz demonstrates that the indirect waveform rises approximately 0.2 ms after the direct waveform rises; it has the shape of a phase-reversed version of the direct component with a 0.2 ms delay. Such an indirect component shifts the tone burst response to the left of the direct component of the tone burst response. Thus, the apparent propagation time is estimated to be shorter than the actual propagation time. Consequently, the result reveals that such shortened apparent propagation time is attributable to the negative peak appearing in the impulse response for the circular plate model, because the negative peak, shown in Fig. 12(b), is the only indirect component of the impulse response for the circular plate model. Also at 1 kHz, shown in Fig. 13(b), the indirect component of the tone burst response makes the entire waveform of the tone-burst response shift left. In contrast, at 3 kHz, as shown in Fig. 13(a), the indirect component of the tone-burst response makes the apparent propagation time greater than that of the direct component. This is because the 0.2 ms delay of the indirect component is greater than half the period at 3 kHz. In such a case, the tone burst response appears later than its direct component, thereby resulting in a greater apparent propagation time than the actual propagation time. Actually, such frequency characteristics were shown in Fig. 12(a), demonstrating that, at higher frequencies, the apparent propagation time is greater than the actual propagation time, approximately 2.9 ms.

The negative peak, or the phase-reversed indirect component, is a backscatter diffraction from edges of the circular plate. Kawai and Toyoda [14] showed theoretically that reflection/diffraction field due to a wall is predicted by particle velocities (dipole source) on a virtual surface which is an extension of the wall; consequently, the backscatter diffraction wave is phase-reversed in the source-side region, as indicated in Eq. (2) of Ref. [14].

### 4.2.2. Effect of the circular plate thickness

Figure 12(c) shows the apparent propagation time for the circular plate models with a fixed radius ($r = 75$ mm) and various thicknesses ($d = 2, 10, \text{and } 20$ mm). The results show that, for $d = 10$ and 20 mm, the apparent
propagation time varies with frequency; there is no prominent difference between \( d = 10 \) and \( 20 \) mm. However, for \( d = 2 \) mm, the apparent propagation time is almost identical to the propagation time that is observed for a distance between the source and receiver positions without the circular plate (approximately 2.9 ms), indicating no variation with frequency. Figure 12(d) shows the impulse responses for the circular plate model with various thicknesses. As shown in Fig. 12(d), negative peaks appear for all thicknesses. However, the amplitude of the negative peak for \( d = 2 \) mm is smaller than that for \( d = 10 \) and 20 mm. Therefore, in the case of the thin circular plate, the reflected sound does not have a prominent effect on the apparent propagation time, thereby resulting in a frequency-independent apparent propagation time, which is almost identical to the actual propagation time. These results suggest that the negative peak is produced by a reflection from the circular profile of the circular plate including its lateral surface and edge. Because the frequency characteristics of the apparent propagation time for the circular plate model are similar to those for the head model, the same type of reflection is likely to produce a frequency dependence of apparent propagation time and subsequently ITDs, even though the head is a more complicated shape than the circular plate.

5. RIGID SPHERE SIMULATION

In the previous section, the results of circular plate models suggest that reflections from the circular profile of the circular plate have an impact on a frequency dependence of the apparent propagation time and ITDs. The circular profile of the circular plate consists of a lateral surface and a clear-cut edge, while a human head does not have such clear-cut edges. Therefore in this section, a further BEM simulation is performed by introducing a rigid sphere model in order to explore if smooth surface without clear-cut edges produces a frequency dependence of apparent propagation time.

5.1. Methods

A rigid sphere model with a radius of 65 mm, which corresponds to the interaural distance of the head model, was constructed as a simplified head. The source and receiver positions were defined in the same manner as for the head model. Two receiving points, which correspond to left and right ears, were located at a distance of approximately 1 mm from the surface of the sphere. A point source was located on the left side at a distance of 1 m from the left receiving point. The surface of the sphere was assumed to be acoustically rigid. After calculating impulse responses at the two receiver points, \( t_L \), \( t_R \), and the ITD (\( = t_R - t_L \)) were estimated by applying the same procedure as used for the head model.

5.2. Results

Figures 14(a)–14(c) illustrate the waveforms of the direct and indirect components of the tone-burst responses for 3 kHz, 1 kHz, and 500 kHz, respectively, for the rigid sphere model. The results show that, as in the cases of the dummy head model and the circular plate model (\( r = 75, d = 10 \)), the waveform of the indirect component appears earlier than that of the direct component. As a result, the indirect component shifts the tone burst response to the left of the direct component. Figure 15 shows simulated \( t_L \), \( t_R \), and ITD values with the rigid sphere along with the apparent propagation time for 1 m propagation in the case without the rigid sphere, namely, in a free field. The filled circles represent the corresponding values from the measurement of dummy head at KU. The results show that, in the rigid sphere simulation, the apparent propagation time to the left ear (ipsilateral ear) \( t_L \) decreases with
decreasing frequency at frequencies higher than 400 Hz whereas it exhibits a slope that is less steep than that using the head model as shown in Fig. 7. On the other hand, $\tau_{C28}$ for the contralateral ear decreases slightly with decreasing frequency above 1 kHz. However, its slope is less steep than that of $\tau_{L}$, thereby resulting in the same frequency dependence of ITDs as observed in the case of measurements and simulations of a human head.

Because the sphere does not have a clear-cut edge, the rigid sphere model demonstrates that the frequency dependence of apparent propagation time to the ipsilateral ear is produced without such a clear-cut edge. Figure 16 depicts the simulated impulse responses at the left and right ear positions, respectively corresponding to ipsilateral and contralateral to the sound source, for the rigid sphere model. The impulse response at the left ear, indicated by the solid line, demonstrates no prominent negative peak around 3.2 ms as observed for the circular plate model, which originates from a clear-cut edge of the circular plate. Instead, it has a weak negative peak appearing right after the direct component around 3 ms, with a following gentle slope that slowly converges to zero amplitude. This negative peak and slope are the backscatter diffractions from the surface of ipsilateral hemisphere. Namely, these negative pressures are backscatter diffractions from the convex surface of the ipsilateral hemisphere, which are directionally accumulated, temporally continuous, and out of phase with the direct component. Consequently, the rigid sphere model indicates that the anti-phase backscatter diffraction is produced not only by a clear-cut edge but also by a smooth convex surface such as a spherical surface.

Here, revisiting Fig. 8 that depicts the indirect component of the simulated impulse response for the dummy head, it is likely that the negative peak appearing after 3 ms is an anti-phase backscatter diffraction from the pinna outline, because it reaches to the ear canal entrance approximately 0.1 ms after the direct component, corresponding to a path length difference between the direct sound and the reflection from the pinna outline, i.e. several centimeters. Pinna also does not have a clear-cut edge but has convex surfaces of smaller curvature radius than the sphere model, rather close to a circular plate in shape, thereby leading to a greater amplitude of the negative peak than the spherical surface. According to these simulation results, it is indicated that the head itself, as can be approximated by a sphere, produces relatively weak anti-phase backscatter diffractions; in addition, the pinna produces prominent anti-phase backscatter diffractions in the same way as the circular plates do. In total, the convex surfaces included in the head and pinna yield anti-phase backscatter diffractions, leading to the greater apparent propagation time to the ipsilateral ear and, consequently, the frequency-dependence of ITDs.

In Fig. 15, compared to the values from the measurement of the dummy head as indicated by the filled circles, the slope of $\tau_{L}$ for the rigid sphere is much less steep than that for the dummy head measurement. Furthermore, the slope of $\tau_{R}$ is rather flat regardless of frequency while that for the dummy head measurements decreases with decreasing frequency. The results also show that $\tau_{R}$ is smaller for the rigid sphere model than for the dummy head measurement, thereby resulting in much smaller ITDs of the rigid sphere model. This deviation is explainable by a difference in geodesic distance between the sphere and the dummy head, namely, a human head has a greater geodesic distance between the two ears than the sphere model. Considering that the circular plate models with various radii resulted in different slopes of $\tau$ as shown in Fig. 12(a), the difference in slopes of $\tau_{L}$ between the rigid sphere and the dummy head is most likely explained by the
size difference between the rigid sphere and the dummy head.

6. DISCUSSION

As shown in Figs. 5 and 7, the measurement and simulation results indicate that the apparent propagation time to the ipsilateral ear decreases at lower frequencies. Simultaneously, the apparent propagation time to the contralateral ear also decreases. The apparent propagation time to the ipsilateral ear as a function of frequency, however, shows a steeper slope compared with that to the contralateral ear, which subsequently results in the frequency dependence of ITDs. As shown by the circular plate model simulation, the slope of the apparent propagation time for the ipsilateral ear is explainable by reflections from surfaces such as the circular profile of the circular plate. On the other hand, the reason for the apparent propagation time to the contralateral ear decreasing with lower frequencies is still unclear. For $\theta_{\text{inc}} = 90^\circ$, which was examined in this study, no direct sound reaches the contralateral ear. The contralateral ear can observe only the indirect sounds that advance along the surface of the head. Therefore, one possible explanation is that the apparent propagation speed of sound wave that advances along the surface may depend on the frequency. As shown in Fig. 15, the results of the rigid sphere model demonstrate that $\tau_R$ for the contralateral ear increases slightly above 1 kHz. Compared to the apparent propagation time in a free field, which is constant above 300 Hz, the sphere also is likely to lead to the frequency dependence of $\tau_R$. However, when compared to that of the dummy head measurement at KU, the frequency dependence of $\tau_R$ for the rigid sphere model is prominently weak. Such discrepancies between the results for the head and rigid sphere imply that the complicated geometry of the head, including its nonsphericity and protruding objects (pinnae and nose), would be one cause to yield the frequency dependence of the apparent propagation time to the contralateral ear. It, however, remains to be revealed why the apparent propagation time to the contralateral ear has a frequency dependence. This issue must be addressed in future works.

As shown in Fig. 7, the simulated ITDs agree well with the measured ITDs at frequencies above 400 Hz, while exhibiting deviations from the measured ITDs at frequencies below 300 Hz. In addition, simulated $\tau_L$ and $\tau_R$ deviate larger from measured values at lower frequencies. As mentioned in the section “Measurement,” the measured $\tau_L$ and $\tau_R$ may be erroneous because of the reflected sounds that superimpose the direct sound at lower frequencies. This may be one cause of the differences between measured and simulated values.

Other possible explanations for the difference between simulated and measured values of $\tau_L$ and $\tau_R$ would be the difference in acoustic impedance of the surface and sound sources between the model and the dummy head. In the BEM simulation, the surface of the head model is assumed to be acoustically rigid having an infinite acoustic impedance while the dummy head was made from hard resin whose acoustic impedance is finite [15]. Such differences in acoustic impedances must have an impact on the simulation results in calculating the phase and amplitude of the reflected waveforms. As for the sound sources, an ideal point source radiating an ideal spherical wave is assumed in the simulations, whereas an actual loudspeaker cannot radiate an ideal spherical wave. Assuming that the sound source is located at a distance of 1 m from the circular plate with a 0.1-m radius, the acoustic path length of the indirect sound wave differs by approximately 15 μs between spherical and plane waves. Such differences in path lengths of indirect waves due to the sound sources would result in different apparent propagation times between the simulations and measurements. At least, above 400 Hz, it is supposed that some factors including acoustic impedance and source resulted in the differences in $\tau_L$ and $\tau_R$ between the simulations and measurements. Because the ITD corresponds to $\tau_R - \tau_L$, these differences cancel, which thereby lead to the agreement of ITDs between the simulations and measurements at this frequency range.

It is commonly found that the spherical model underestimates the ITD (e.g. [6,16]). As shown in the literatures for animal models [9–11], ITDs are larger when a head has pinnae. This can be explained by a difference between the amplitudes of negative peak originating from a head and a pinna, or a sphere and a circular plate, as observed in this study. In addition, interaural path length is basically smaller for the sphere compared to actual heads. Underestimation of ITD when using the sphere model can be explainable by these geometrical and acoustical difference between a sphere and an actual head. In addition, Gourevitch and Brette [17] indicated that ITD varies when there exist reflections from nearby walls. This phenomenon may be similar to the phenomenon considered in the current work because both phenomena can be considered as effects of reflections, or backscatter diffraction, on ITDs. Although the walls assumed in [17] are located separated from the head and more distant while the current work discusses the reflections from head itself, it would be intriguing to see if there is a similar physical mechanism in future works.

7. CONCLUSIONS

This study investigated the frequency dependence of ITDs through measurements and numerical simulations using a dummy head, a circular plate, and a sphere. The measurement results for the dummy head showed that ITDs
vary with the frequency because the apparent propagation time to an ipsilateral ear varies with frequency. Numerical simulations with the dummy head showed that this phenomenon is caused by a phase shift of the received sound waveform due to reflections from the head. Numerical simulations with the circular plate revealed that the circular profile of the circular plate produces a reflection, or anti-phase backscatter diffraction, that is relevant to the phase shift, yielding the frequency dependence of the apparent propagation time. Furthermore, numerical simulations with the rigid sphere showed that such anti-phase backscatter diffractions are produced without a clear-cut edge but with a convex surface. These results strongly suggest that a major factor in the production of the frequency dependence of ITDs is the anti-phase backscatter diffraction from the convex surfaces of the head and the pinna.

ACKNOWLEDGMENTS

A part of this work was supported by Grants-in-Aid from MEXT, Japan (No. 16H02857 and No. 16H01736).

REFERENCES

[1] J. Blauert, Spatial Hearing, revised ed. (MIT Press, Cambridge, Mass., 1997), pp. 66–69.
[2] A. W. Mills, “On the minimum audible angle,” J. Acoust. Soc. Am., 30, 237–246 (1958).
[3] R. S. Woodworth, H. Schlosberg, J. W. Kling and L. A. Riggs, Woodworth & Schlosberg’s Experimental Psychology, revised ed. (Holt, Rinehart and Winston, New York, 1971), pp. 259–261.
[4] L. A. Abbagnaro, B. B. Bauer and E. L. Torick, “Measurement of diffraction and interaural delay of a progressive sound wave caused by the human head. II.” J. Acoust. Soc. Am., 58, 693–700 (1975).
[5] A. Suzuki and M. Tohyama, “Inter-aural cross-correlation coefficients of KEMAR head and torso simulator in a free field,” Tech. Rep. Inst. Electron. Inf. Commun. Eng. Jpn., 81(37), pp. 23–30 (EA81-7) (1981).
[6] V. Benichoux, M. Rébillat and R. Brette, “On the variation of interaural time differences with frequency,” J. Acoust. Soc. Am., 139, 1810–1821 (2016).
[7] G. F. Kuhn, “Model for the interaural time differences in the azimuthal plane,” J. Acoust. Soc. Am., 62, 157–167 (1977).
[8] S. N. Rzhevkin, A Course of Lectures on the Theory of Sound (Pergamon Press, Oxford, 1963), pp. 362–363.
[9] G. L. Roth, R. K. Kochhar and J. E. Hind, “Interaural time differences: Implications regarding the neurophysiology of sound localization,” J. Acoust. Soc. Am., 68, 1643–1651 (1981).
[10] N. Greene, K. L. Anbuhl, W. Williams and D. J. Tollin, “The acoustical cues to sound location in the guinea pig (Cavia porcellus),” Hear. Res., 316, 1–15 (2014).
[11] K. Koka, H. G. Jones, J. L. Thornton, J. E. Lupo and D. J. Tollin, “Sound pressure transformations by the head and pinnae of the adult Chinchilla (Chinchilla lanigera),” Hear. Res., 272, 135–147 (2011).
[12] T. Hirahara, M. Otani and D. Morikawa, “Origin of frequency dependence in ITD—Acoustical measurement,” Proc. Spring Meet. Acoust. Soc. Jpn., pp. 733–734 (2012).
[13] M. Otani, T. Hirahara and S. Ise, “Numerical study on source distance dependency of head-related transfer functions,” J. Acoust. Soc. Am., 120, 3253–3261 (2009).
[14] Y. Kawai and M. Toyoda, “Development of edge-effect suppression barriers,” Acoust. Sci. & Tech., 35, 28–34 (2014).
[15] A. Trogé, R. L. O’Leary, G. Hayward, R. A. Pethrick and A. J. Mulholland, “Properties of photocured epoxy resin materials for application in piezoelectric ultrasonic transducer matching layers,” J. Acoust. Soc. Am., 128, 2704–2714 (2010).
[16] V. R. Algazi, C. Avendano and R. O. Duda, “Estimation of a spherical-head model from anthropometry,” J. Audio Eng. Soc., 49, 472–479 (2001).
[17] B. Gourévitch and R. Brette, “The impact of early reflections on binaural cues,” J. Acoust. Soc. Am., 132, 9–27 (2012).