INVESTIGATION OF THE EFFECT OF IMPEDANCE ON AZIMUTH CUES DERIVED FROM SPHERICAL HEAD MODELS

Bradley E Treeby, Roshun Paurobally, Jie Pan

Centre for Acoustics, Dynamics and Vibration
School of Mechanical Engineering
University of Western Australia
35 Stirling Highway, Crawley, WA 6009, Australia
treebs@mech.uwa.edu.au

ABSTRACT

Recent implementations of virtual sound using head related transfer functions often include an analytical low frequency diffraction model to reproduce stable azimuth cues which are independent of the quality and accuracy of high frequency pinnae data. This paper investigates the previously neglected effects of surface acoustic impedance on the salient azimuth cues produced by a single sphere diffraction model. Results indicate an increase in interaural level difference with both a decreasing impedance magnitude and an increasingly negative impedance phase component. The interaural time difference appears to be less sensitive to impedance changes although an increase is evident with a decreasing impedance magnitude, particularly for sources near the interaural axis. These changes produce offset interaural cues which correspond to a shift in source location in the virtual environment. The effect of incorporating a frequency dependent impedance based on the approximate acoustic properties of hair is also discussed.

1. INTRODUCTION

Extensive work over the last decade into implementations of virtual sound environments has demonstrated the importance of using accurately matched individualised head related transfer functions (HRTFs) to maintain the auditory cues required for a realistic and spatially accurate virtual environment [1, 2]. Generally the availability of laboratory environments in which to measure personalised HRTF data has remained unchanged over that period; however the ability of artificial or non-individualised functions to create realistic spatial impressions has greatly increased. In particular recent implementations of artificial HRTFs combining a low frequency head and torso (HAT) model with high frequency pinnae data have had considerable success [3, 4]. The addition of an anthropometric HAT model provides stable azimuth cues consistent with psychophysical expectations whilst reducing the reliance on coherent low frequency measurements or accurately matched HRTF sets.

These HAT models generally combine a pinnae-less spherical head with a separated spherical torso to compute the required low frequency HRTF information from analytical diffraction models [5]. Note that previously more common implementations have simply used a spherical head model – the term HAT is used here to refer to both. Such analytical HAT models have the ability to be recomputed rapidly to suit the perceptual requirements of any individual user, however the ability of these models to accurately account for acoustic impedance properties, in particular the effect of hair, has thus far been limited. The incorporation of a non-rigid head surface into previous studies of HRTF data using boundary element models has shown the importance of accurately describing the boundary conditions, with up to a 6 dB difference in the computed HRTF due to the inclusion of impedance [6].

This paper provides a comprehensive investigation into the effect of impedance on the salient azimuth cues (interaural time and level differences) provided by the displaced position of the pinnae on the human head. These cues are calculated using a spherical diffraction model based on general impedance boundary conditions. The effect of including a frequency dependent impedance term which is based on the approximate acoustic characteristics of human hair is also presented.

2. GENERAL IMPEDANCE PRESSURE MODEL

The derivation of a sphere diffraction model is not an inimitable task with many authors previously providing extensive derivations [7, 8]. Most derivations and discussions however are based on Neumann (rigid) boundary conditions and are unable to be trivially expanded to include a specific acoustic impedance term.

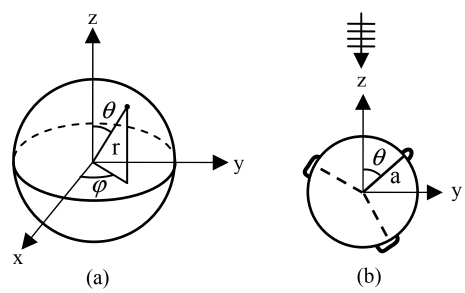

Figure 1. (a) *General spherical coordinate system* (b) *Symmetrical coordinate system with incident wave in the negative Z-direction.*

The complete relationship for the scattered sound pressure on and around spherical objects is found by assuming that an incoming wave will produce a secondary outgoing spherical wave at each scatterer, the summation of all being the total response of the system. The constants in the scattered wave expressions are solved by examining the boundary conditions imposed on each sphere. The general spherical coordinate system is shown in figure 1.

Assuming a locally reactive surface and a monochromatic incident wave (harmonic time dependence), the general boundary condition can be obtained by relating the pressure and radial velocity at the sphere surface:

$$\frac{\partial p}{\partial r}+\left(i\rho_0\omega\frac{1}{Z(\omega)}\right)p\Bigg|_{sphere\,surface}=0. \qquad (1)$$

Here ρ_0 is the mean medium density, ω is the angular frequency, and $Z(\omega)$ is the specific acoustic impedance of the sphere surface.

For a plane wave incident on a single sphere in a uniform non-viscous medium, the propagation of both the incident and reflected waves are governed by the scalar wave equation (using the spherical Laplacian). If the plane wave is travelling in the negative z-direction, the solution is axially symmetric and will only depend on one spherical angle (θ). The boundary conditions are imposed on the sphere surface ($r = a$) and used to calculate the constants in the scattered wave equation. The following general equation for sound pressure on and around a single sphere including impedance is thus derived by summing the incident and scattered waves:

$$(p_i+p_s)=p_0e^{-i\omega t}\sum_{l=0}^{\infty}(-i)^l(2l+1)P_l[\cos\theta]B_l, \qquad (2)$$

where

$$B_l=j_l[kr]-\frac{h_l^{(1)}[kr]\left(k\,j_l'[ka]+i\rho_0\omega\frac{1}{Z}j_l[ka]\right)}{k\,h_l'^{(1)}[ka]+i\rho_0\omega\frac{1}{Z(\omega)}h_l^{(1)}[ka]}.$$

Here p_0 is the mean pressure value (corresponding to the strength of the incident plane wave), P_l is a Legendre polynomial of order l, θ and r define the location of the pressure point to be calculated, j_l and h_l are spherical Bessel and Hankel functions of order l, k is the wave number, and a is the radius of the sphere.

To investigate the effect of impedance on the salient azimuth cues (which are derived from pressure values on the sphere surface), (2) can be significantly simplified using the recursion and cross product relationships for spherical Bessel functions. This yields the surface pressure function (3) which is similar in form to the previously published solution for the Robin boundary condition [8]. For a rigid sphere the impedance value becomes infinite and (3) reduces to yield the solution for the Neumann boundary condition as presented and discussed by previous authors [7].

$$\frac{(p_i+p_s)_{r=a}}{p_0e^{-i\omega t}}=\frac{1}{(ka)^2}\sum_{l=0}^{\infty}\frac{(-i)^{l-1}(2l+1)P_l[\cos\theta]}{h_l'^{(1)}[ka]+i\rho_0c\frac{1}{Z(\omega)}h_l^{(1)}[ka]}. \qquad (3)$$

The analytical interaural level difference (ILD) is obtained using (3) by examining the difference in sound pressure level between the two ear locations, each offset from the frontal median axis by 100° [9]:

$$ILD=20\log_{10}\left(\left|\frac{(p_i+p_s)_{right\,ear}}{(p_i+p_s)_{left\,ear}}\right|\right). \qquad (4)$$

Similarly the interaural time difference (ITD) for each angle of incidence is calculated from the difference in arrival time between the two ear locations. The actual value of ITD differs slightly depending on the definition of arrival time. Interaural phase delay and group delay are frequently used and exhibit similar characteristics but both are frequency dependent and display an increase in ITD at low frequencies [10]. Ray tracing algorithms [9], amplitude threshold based time constants, or comparisons to minimum phase reconstructions [11] can alternatively be used to derive a single value for interaural time difference for each source angle.

Due to the ease of computation, trends in ITD due to the inclusion of impedance are examined here by using a frequency averaged interaural phase delay. The phase delay for a particular location on the sphere is computed by the negative of the phase response $\Theta(\omega)$ divided by angular frequency as shown in (5).

$$D(\omega)\triangleq-\frac{\Theta(\omega)}{\omega}. \qquad (5)$$

For a HRTF measured on a subject with pinnae, the monaural resonant effects of the pinnae cavity only begin to appear above 3 kHz. Implementations of composite HRTFs generally use the analytical HAT model exclusively for frequencies below 500 Hz, then progressively introduce the measured HRTF for frequencies up to 3000 Hz after which the HAT model is no longer used [4, 12]. ITD values are thus calculated here from interaural phase delays frequency averaged over 100 to 3000 Hz. This has the effect of proportionally increasing the ITD values when compared to those obtained from averages over a larger frequency range, or those found using some of the alternative measures mentioned previously. This increase is due to the augmented effect of the larger phase delay below 1000 Hz; however observed trends are considered consistent regardless of this selection.

3. GENERAL EFFECT OF SPHERE IMPEDANCE ON THE PRIMARY AZIMUTH CUES

Two primary aspects of impedance are investigated here, firstly the impedance magnitude (which is approximately inversely related to the amount of boundary absorption) and secondly the phase (which encompasses the effect of different ratios of resistive to reactive impedance components). For convenience impedance values are initially assumed independent of frequency to examine general trends in interaural time and level difference. Discussion relating to frequency dependent impedance is given in section 4.

3.1. Effect of Impedance on Interaural Level Difference

Figure 2 shows the effect of decreasing the magnitude of a purely resistive impedance. A clear increase in the ILD can be seen, particularly for incident angles near the interaural axis. This is primarily due to the reduction in pressure level at the contralateral ear which results from the increased absorption of the sphere surface reducing the intensity of the sound which is diffracted to the rear of the sphere. As with previously published ILD plots [11] the decrease in level difference at 80° is a result of symmetrically diffracted waves arriving in phase at the rear of the sphere creating a bright spot at the contralateral ear.

When the surface impedance also contains a reactive (imaginary) component, the interaction with the boundary causes a phase difference in the reflected wave. This alters the resultant pressure field on and around the sphere and thus also potentially affecting the interaural cues. Figure 3 (corresponding to a plane wave source coincident with the interaural axis) illustrates the change in ILD with the impedance phase angle whilst keeping the overall impedance magnitude constant at a value of 1000 Pa s m^{-1}. For a negative impedance phase angle (positive resistive component, negative reactive component – this corresponds to most common acoustic materials), there is a further increase in ILD.

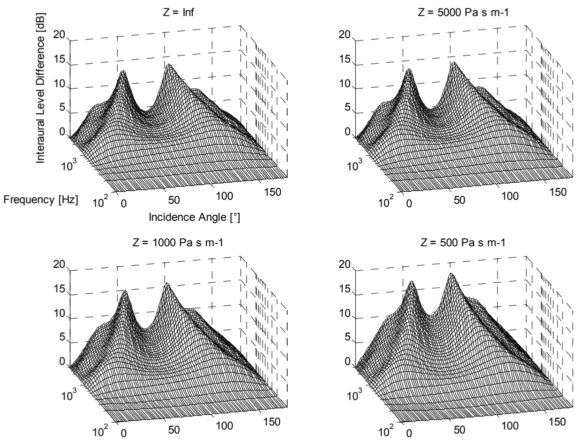

Figure 2. *Increase in the interaural level difference with a decrease in resistive impedance.*

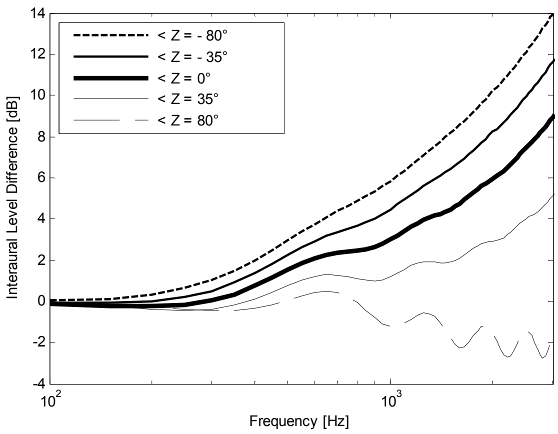

Figure 3. *Change in interaural level difference with impedance phase angle for a constant impedance magnitude of 1000 Pa s m^{-1} and a source angle of 90°.*

3.2. Effect of Impedance on Interaural Time Difference

The interaural time delay takes precedence as the primary azimuth cue at low frequencies (where there is a very minimal interaural level difference) or when the ITD and ILD cues conflict. Previous studies have shown that the time difference is more impervious to system changes than the level difference (for example for near-field sources [11]) and this trend continues here. As Figure 4 illustrates, the frequency averaged phase delay ITD shows a general increase as the impedance magnitude is reduced. For sources near the interaural axis, the difference in ITD between a completely rigid and a highly absorptive sphere is around $70\mu s$. This difference is enough to move the spatial cue by as much as 20° – a value significantly greater than previously published values of 10° for spatial resolution around the interaural axis [9]. For source directions closer to the median axis, the change in ITD and thus spatial direction is much less than this, however the minimum audible angle threshold in this region is also reduced. Altering the impedance phase angle in either direction slightly reduces the value of ITD.

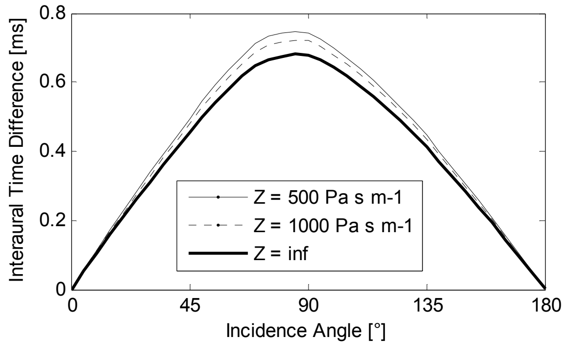

Figure 4. *Change in interaural time difference with resistive impedance magnitude.*

4. APPROXIMATE CHANGES TO AZIMUTH CUES DUE TO HAIR IMPEDANCE

Intrinsically coupled with the absence of impedance considerations in most HAT models is the lack of any comprehensive study or experimental data encompassing the required specific acoustic impedance measurements of the various surfaces which constitute the human head and torso. This is largely because of the lack of techniques available which are able to accurately record impedance when impedance tube, reverberation room or flat samples can not be readily obtained. This makes it difficult to quantitatively define the effect the inclusion of impedance has on the interaural level and time differences at low frequencies.

Previous studies have provided basic absorption coefficient data measured using an impedance tube for both human skin (found to be approximately rigid) and hair which had impedance characteristics likened to a typical foam or mineral wool [13]. Using these results the specific acoustic impedance characteristics of hair can be approximated as an exponentially decaying impedance magnitude with a linearly increasing phase component as shown in (6).

$$Z(\omega) = \left(4000\,e^{-0.0015\omega/\pi} + 700\right)e^{i\pi/180(\omega/150\pi - 60)}. \qquad (6)$$

Incorporating this equation into (3) and (4) yields the possible change in ILD due to the inclusion of hair-like impedance. Figure 5 illustrates the resultant ILD curves at 3000 Hz, with up to an 6 dB increase for source angles near the interaural axis. Although hair generally does not cover the entire head surface (thus the expected ILD change may be somewhat smaller), the difference due to the inclusion of impedance is nonetheless significant. A change in ITD similar to that shown in Figure 4 is evident using an impedance function as described by (6) with a maximum difference of $43\mu s$. The reduced maximum ITD difference using the frequency dependent impedance is due to the contribution of the phase delay at low frequencies, where the impedance value approaches that of a rigid sphere. As the ITD cue normally takes precedence over the ILD when the cues conflict, it is unknown as yet to what extent the shifted level and time differences will affect the perceived source direction. The changes however are significantly above previously published thresholds for interaural differences.

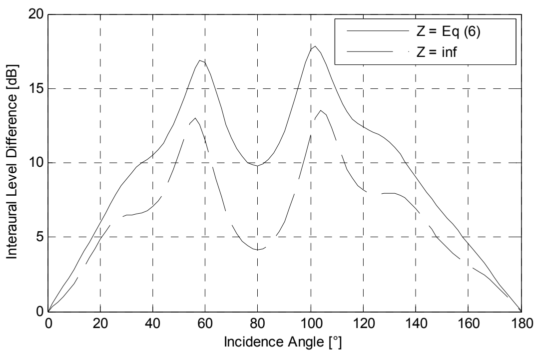

Figure 5. *Change in ILD at 3000 Hz due to the approximate acoustic characteristics of hair.*

5. CONCLUSION

This paper presents a general form of the solution to the single sphere diffraction problem for a general impedance boundary condition. This is then used to examine the trends in the two primary auditory cues (used to locate azimuthal displaced sounds) with changes in complex impedance. Both the interaural level and time differences exhibit non-trivial increases with decreasing impedance magnitude. These changes produce offset interaural cues (particularly for sources near the interaural axis) that are greater than the minimum audible angles and thus correspond to a shift in source direction in the virtual spatial environment. Results for the change in interaural differences due to the inclusion of impedance based on a hair-type material also show significant changes, particularly for sources located towards the interaural axis.

Given that the effect of neglecting impedance appears significant, implementations of virtual sound using low frequency head and torso models should exhibit an increase in accuracy by the inclusion of an equivalent impedance value which results in similar low frequency azimuth cues as a sphere of multifaceted impedance matched to the human head. A comparison of cues derived from experimentation or boundary element models with those produced from analytical models could be used to derive this value. A useful annexure would be to further extend the impedance based analytical model to account for location dependent boundary conditions (multifaceted surfaces) as is the case for finite and boundary element computational models. This would allow a more accurately defined impedance surface to be used in the calculation of interaural cues.

6. ACKNOWLEDGEMENTS

The first author would like to acknowledge the assistance of the Robert and Maude Gledden, and F S Shaw Memorial Postgraduate Scholarships.

7. REFERENCES

[1] Bronkhorst, A. W., 'Localisation of real and virtual sources', *J. Acoust. Soc. Am.*, vol. 98, no. 5, pp. 2542-2553, (1995).

[2] Wenzel, E. M., Arruda, M., Kistler, D. J. and Wightman, F. L., 'Localization using nonindividualized head-related transfer functions', *J. Acoust. Soc. Am.*, vol. 94, no. 1, pp. 111-123, (1993).

[3] Algazi, V. R., Duda, R. O., Duraiswami, R., Gumerov, N. A. and Tang, Z., 'Approximating the head-related transfer function using simple geometric models of the head and torso', *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2053-2064, (2002).

[4] Algazi, V. R., Duda, R. O., Morrison, R. P. and Thompson, D. M., 'The use of head-and-torso models for improved spatial sound synthesis', in *Proc. AES 113th Convention*, AES, Los Angeles, CA, p. preprint 5712, (2002).

[5] Gumerov, N. A. and Duraiswami, R., 'Computation of scattering from N spheres using multipole reexpansion', *J. Acoust. Soc. Am.*, vol. 112, no. 6, pp. 2688 - 2701, (2002).

[6] Katz, B. F. G., 'Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparison to real measurements', *J. Acoust. Soc. Am.*, vol. 110, no. 5, pp. 2449-2455, (2001).

[7] Morse, P. M. and Ingard, K. U., *Theoretical Acoustics*, McGraw-Hill, New York, (1968).

[8] Gumerov, N. A. and Duraiswami, R., *Fast multipole methods for the Helmholtz equation in three dimensions*, Elsevier, Oxford, UK, (2004).

[9] Blauert, J., *Raumliches Horen. (English) Spatial hearing: the psychophysics of human sound localization*, Rev. edn, MIT Press, Cambridge, Mass., (1997).

[10] Kuhn, G. F., 'Model for the interaural time differences in the azimuthal plane', *J. Acoust. Soc. Am.*, vol. 62, no. 1, pp. 157-167, (1977).

[11] Duda, R. O. and Martens, W. L., 'Range dependence of the response of a spherical head model', *J. Acoust. Soc. Am.*, vol. 104, no. 5, pp. 3048-3058, (1998).

[12] Zotkin, D. N., Hwang, J., Duraiswami, R. and Davis, L. S., 'HRTF personalization using anthropometric measurements', in *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on.*, IEEE, New Paltz, NY, USA, pp. 157-160, (2003).

[13] Katz, B. F. G., 'Acoustic absorption measurement of human hair and skin within the audible frequency range', *J. Acoust. Soc. Am.*, vol. 108, no. 5, pp. 2238 - 2242, (2000).