Audio Engineering Society

Convention Paper 10412

Presented at the 149th Convention
Online, 2020 October 27-30

A Formula for Low-Frequency Interaural Level Difference

Rahulram Sridhar and Edgar Y. Choueiri

3D Audio and Applied Acoustics Lab, Princeton University, Princeton, NJ, 08544, USA

Correspondence should be addressed to Rahulram Sridhar (rahulram@princeton.edu)

ABSTRACT

A formula for low-frequency interaural level difference (LF-ILD) as a function of source distance and direction is derived from rigid-sphere head-related transfer function (RS-HRTF) theory. Since ILD at low frequencies (typically $f < 700\,\text{Hz}$) is a dominant cue for near-field distance perception in the free-field, a simple formula akin to the Woodworth formula for high-frequency interaural time difference, may be beneficial to easily and efficiently implement distance- and direction-dependent LF-ILDs for real-time spatial audio applications on devices with limited computational resources. By evaluating the limit as $f \to 0$ of the infinite series representation of the RS-HRTF for finite source distances, an exact, closed-form expression for the "DC gain" as a function of source distance and direction is derived. For a given source location, using this expression to compute the DC gain at the left and right "ears," and then taking the ratio of the two quantities gives the desired LF-ILD. As an example, it is shown that the derived formula may be used to extrapolate far-field ILD spectra of a rigid sphere to the near-field exactly for low frequencies and with sufficiently high accuracy for higher frequencies. Furthermore, the derived formula, like the Woodworth formula, is well-suited to individualization.

1 Introduction

Many virtual reality (VR) applications use head-related transfer functions (HRTFs) to synthesize binaural signals for spatial audio reproduction over headphones or crosstalk-cancelled loudspeakers. While the HRTF is a listener-specific function of frequency, source direction and distance, the distance-dependence is significant only in the near-field (generally within 1 m from the listener). This enables an HRTF to be classified as either far-field or near-field, with distance-dependence being ignored in the former.

At low frequencies (typically less than 700 Hz), interaural level differences (ILDs) computed from a far-field HRTF are small (close to, if not exactly, 0 dB) irrespective of source direction and distance, but those computed from a near-field HRTF increase significantly with decreasing distance, especially for lateral source directions [1–4]. In contrast, interaural time differences (ITDs) and high-frequency spectral features are, even from a perceptual standpoint, largely independent of source distance, especially when compensating for the so-called *acoustic parallax* [2, 5–7]. Therefore, although low-frequency ILD (LF-ILD) is only one of many cues that contribute to near-field distance perception in reverberant environments, it is the dominant one in the free-field [3–6]. For applications where moving near-field sources are to be rendered smoothly in real-time, an accurate and efficient method for computing and reproducing these ILDs may be beneficial.

1.1 Background and Previous Work

One way to ensure that correct LF-ILDs are reproduced is to use an accurate near-field HRTF. There are many

methods for obtaining such an HRTF (for summaries, see the works by Enzner et al. [8] and Xie [9], for example), of which a few are well-suited for real-time applications.

Spagnol et al. [10] propose a parametric filter composed of a first-order, high-frequency shelving filter and a second-order rational function that together approximate the so-called *distance variation function* (DVF) proposed by Kan et al. [11].[1] The DVF is essentially the ratio between a near-field and far-field rigid-sphere HRTF (RS-HRTF) and is used to extrapolate a far-field HRTF to the near-field. The rational function is used to compute the "DC gain" (i.e., the gain at frequency, $f = 0$) of the near-field RS-HRTF as a function of source distance and direction (since the gain of the far-field RS-HRTF is 0 dB irrespective of source direction), while the shelving filter is used to transform the far-field frequency response to the near-field one, normalized by the DC gain. Since the shelving filter parameters and the rational function coefficients are computed from data corresponding only to a discrete set of source directions, linear interpolation is used to compute the parameter values of the shelving filter and the DC gains for other directions. We note that implementing the parametric filter in practice essentially amounts to applying an ILD correction that is a function of frequency, source distance and direction.

In an audio plugin developed by Poirier-Quinot and Katz [13], extrapolation of a far-field HRTF to the near-field is accomplished in real-time through the use of acoustic parallax correction and a pair of cascaded biquad filters for implementing distance- and frequency-dependent ILD corrections. Although details are lacking, Poirier-Quinot and Katz indicate that the filters reproduce ILD variations "via a spherical head model," which suggest that they are at least conceptually similar to the parametric filter proposed by Spagnol et al. [10]. The biquad filter coefficients appear to be computed for a finite set of source distances and/or directions, since Poirier-Quinot and Katz state that linear interpolation between neighboring biquad values is performed when necessary.

Both methods are based on the RS-HRTF, making them conceptually-simple, computationally-efficient, and well-suited to individualization. However, recalling that *low-frequency* ILD is the perceptually-dominant distance cue [3–6], we suggest that computing a frequency-dependent ILD correction may not be necessary in practice to extrapolate a far-field HRTF to the near-field. Furthermore, as we shall later demonstrate, far-field and near-field ILD spectra corresponding to the RS-HRTF are largely similar except for an overall gain that is equal to the LF-ILD. Consequently, correcting for this ILD alone may be sufficient to compute perceptually accurate near-field ILD spectra from corresponding far-field ones. In addition to being easier to implement, further computational savings may be had by ignoring the frequency-dependence and instead using a simple formula, akin to the extended Woodworth formula for high-frequency ITD [14], for example, to compute a distance-dependent LF-ILD. Even should a frequency-dependent ILD correction be required, the parametric filter proposed by Spagnol et al. [10], for example, may benefit from a simple formula instead of the rational function to compute the DC gain exactly for any source distance and direction, without the need for any interpolation.

1.2 Objectives and Approach

Our objective is to derive a formula for LF-ILD as a function of source distance and direction. Drawing inspiration from the well-known Woodworth formula for high-frequency ITD [15, 16] and its extended version [14], both of which are based on modeling the head as a rigid sphere, we derive our formula by solving the RS-HRTF in the low-frequency limit. Specifically, we show that the limit, as $f \to 0$, of the infinite series representation of the RS-HRTF for finite source distances may be represented exactly as a closed-form expression. By using this expression to compute the DC gains for left and right "ears" and taking the ratio of these two quantities (or, equivalently, their difference in dB), the LF-ILD is readily obtained. Unlike the rational function by Spagnol et al. [10] and the biquad filters by Poirier-Quinot and Katz [13], our formula is exact and does not require additional interpolation.

1.3 Outline

We begin by reviewing RS-HRTF theory in Section 2, which we then use in Section 3 to derive our formula for LF-ILD. Finally, in Section 4, we give an example application where we use our formula to compute

[1] The DVF is identical to the *difference filter* proposed earlier by Romblom and Cook [12].

LF-ILDs for a selection of source positions in the near-field, and use these ILDs to extrapolate ILD spectra of a far-field RS-HRTF to these positions. We then evaluate the accuracy of the extrapolated spectra relative to benchmarks.

2 Review of RS-HRTF Theory

Let μ denote a non-dimensional frequency such that

$$\mu = 2\pi f \frac{r_{\mathrm{RS}}}{c}, \tag{1}$$

where f is frequency in Hz, c is the speed of sound, and r_{RS} is the radius of the sphere. The RS-HRTF, H, for a sphere centered at the origin and for a source at a finite distance, r_{S}, from the origin, is given by [1]

$$H(\mu,\rho,\Theta) = -\frac{\rho}{\mu} e^{i\mu\rho} \sum_{m=0}^{\infty} B_m^*(\mu,\rho) P_m(\cos\Theta), \tag{2}$$

where P_m is the m^{th}-degree Legendre polynomial, Θ is the so-called *angle of incidence*, $\rho = r_{\mathrm{S}}/r_{\mathrm{RS}} > 1$ is the non-dimensional distance to the source, and where

$$B_m(\mu,\rho) = (2m+1)\frac{h_m(\mu\rho)}{h_m'(\mu)}, \tag{3}$$

with B_m^* denoting its complex conjugate, h_m the m^{th}-order spherical-Hankel function (of the first kind), and h_m' its first derivate with respect to the argument. Note that $\Theta \in [0, 180^\circ]$ represents the angle between two rays from the center of the sphere, one to the source and another to one of the "ears." In the limit $\rho \to \infty$, we have [1, 17]

$$\lim_{\rho\to\infty} H(\mu,\rho,\Theta) = \frac{1}{\mu^2} \sum_{m=0}^{\infty} A_m^*(\mu) P_m(\cos\Theta), \tag{4}$$

where

$$A_m(\mu) = \frac{(-i)^{m-1}(2m+1)}{h_m'(\mu)}. \tag{5}$$

We note that Eqs. (2) and (4) are only valid for $\mu > 0$.

3 Derivation of LF-ILD Formula

We begin by defining LF-ILD, L, as

$$L(\rho,\Theta_{\mathrm{L}},\Theta_{\mathrm{R}}) = \lim_{\mu\to 0} \frac{H(\mu,\rho,\Theta_{\mathrm{L}})}{H(\mu,\rho,\Theta_{\mathrm{R}})}, \tag{6}$$

where Θ_{L} and Θ_{R} denote the angles of incidence corresponding to the left and right "ears," respectively. Although customary, we don't need to take the absolute value of the ratio shown in Eq. (6) since the limit is always real-valued.

For $\mu \ll 1$, the limiting form of $h_m(\mu)$ is given by [18]

$$h_m(\mu) \sim \frac{(2m-1)!!}{i\mu^{m+1}}, \tag{7}$$

where !! denotes a double factorial, such that $m!! = m(m-2)(m-4)\ldots$. Using the recurrence relation [18]

$$h_m'(\mu) = -h_{m+1}(\mu) + \frac{m}{\mu} h_m(\mu), \tag{8}$$

we have, for $\mu \ll 1$,

$$\begin{aligned} h_m'(\mu) &= \frac{i[(2m+1)!! - m(2m-1)!!]}{\mu^{m+2}} \\ &= \frac{i(m+1)(2m-1)!!}{\mu^{m+2}}. \end{aligned} \tag{9}$$

Similar to Eq. (7), the limiting form of $h_m(\mu\rho)$ for finite-valued ρ is given by [18]

$$h_m(\mu\rho) \sim \frac{(2m-1)!!}{i(\mu\rho)^{m+1}}. \tag{10}$$

Substituting for $h_m'(\mu)$ and $h_m(\mu\rho)$ from Eqs. (9) and (10), respectively, into $B_m(\mu,\rho)$ in Eq. (3) gives, for $\mu \ll 1$ and $\rho > 1$,

$$B_m(\mu,\rho) = \frac{-\mu(2m+1)}{\rho^{m+1}(m+1)}, \tag{11}$$

which, when substituted into Eq. (2), gives

$$H(\mu,\rho,\Theta) = e^{i\mu\rho} \sum_{m=0}^{\infty} \frac{(2m+1)P_m(\cos\Theta)}{\rho^m(m+1)}, \tag{12}$$

from which we get

$$\lim_{\mu\to 0} H(\mu,\rho,\Theta) = \sum_{m=0}^{\infty} \frac{(2m+1)P_m(\cos\Theta)}{\rho^m(m+1)}. \tag{13}$$

For reasons that will soon become clear, we may rewrite Eq. (13) as

$$\begin{aligned} \lim_{\mu\to 0} H(\mu,\rho,\Theta) = 2\sum_{m=0}^{\infty} \rho^{-m} P_m(\cos\Theta) - \\ \sum_{m=0}^{\infty} \frac{\rho^{-m} P_m(\cos\Theta)}{m+1}. \end{aligned} \tag{14}$$

The generating function, G, for Legendre polynomials is given by [18]

$$G(x,z) = \sum_{m=0}^{\infty} P_m(x) z^m = \frac{1}{\sqrt{1-2xz+z^2}} \tag{15}$$

provided $|z| < 1$. From Eqs. (14) and (15), and recalling that $\rho > 1$, we get

$$\lim_{\mu\to 0} H(\rho,\mu,\Theta) = 2G(\cos\Theta, 1/\rho) - \rho \sum_{m=0}^{\infty} \frac{\rho^{-(m+1)} P_m(\cos\Theta)}{m+1}, \tag{16}$$

where we extract a factor of ρ out of the infinite sum for a reason that will soon become apparent. By integrating the generating function, G, in Eq. (15), we get

$$\int_0^z G(x,\zeta)\,\mathrm{d}\zeta = \sum_{m=0}^{\infty} \frac{P_m(x) z^{m+1}}{m+1}, \tag{17}$$

where ζ is the variable of integration. Then, from Eqs. (16) and (17), we get

$$\lim_{\mu\to 0} H(\mu,\rho,\Theta) = 2G(\cos\Theta, 1/\rho) - \rho \int_0^{1/\rho} G(\cos\Theta,\zeta)\,\mathrm{d}\zeta. \tag{18}$$

Substituting the closed-form expression shown in Eq. (15) for G and evaluating the integral then gives

$$\lim_{\mu\to 0} H(\mu,\rho,\Theta) = \frac{2\rho}{d(\rho,\Theta)} - \rho \times \log\left(\frac{d(\rho,\Theta)+1-\rho\cos\Theta}{\rho(1-\cos\Theta)}\right), \tag{19}$$

where $\log(\cdot)$ denotes taking the natural logarithm of the argument and d given by

$$d(\rho,\Theta) = \sqrt{\rho^2 - 2\rho\cos\Theta + 1} \tag{20}$$

is the non-dimensional Euclidean distance between the source and the chosen ear. It may be easily verified that Eq. (19) is not valid for $\Theta = 0$. Consequently, we take the limit of the expression on the right-hand-side of Eq. (19) as $\Theta \to 0$ to obtain the overall result

$$\lim_{\mu\to 0} H(\mu,\rho,\Theta) = \frac{2\rho}{d(\rho,\Theta)} - \rho \times \begin{cases} \log\left(\dfrac{d(\rho,\Theta)+1-\rho\cos\Theta}{\rho(1-\cos\Theta)}\right), & \Theta \neq 0, \\ \log\left(\dfrac{\rho}{\rho-1}\right), & \Theta = 0. \end{cases} \tag{21}$$

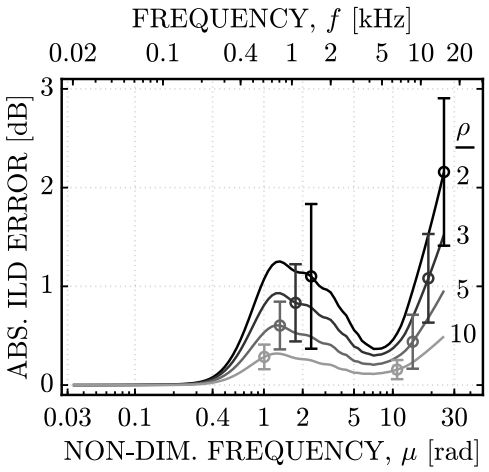

Fig. 1: Mean absolute ILD error versus μ. The errors are averaged over all Θ and the error bars show 1 standard deviation. The top axis shows frequencies for a rigid sphere of radius 0.09 m.

Equation (21) is valid for all $\rho > 1$ and it may be easily verified that

$$\lim_{\substack{\mu\to 0 \\ \rho\to\infty}} H(\mu,\rho,\Theta) = 1, \tag{22}$$

which is a well-known result. The desired formula for L is then given by Eqs. (6), (20), and (21) taken together, noting that the limit of a ratio is equivalent to the ratio of the limits.

4 Example Application

We suggest, in Section 1.1, that using LF-ILDs to extrapolate far-field ILD spectra to the near-field may be sufficient to enable distance perception because 1) LF-ILD is the dominant cue for this purpose [3–6], and 2) the frequency-dependence of ILD does not change significantly with source distance except for an overall gain given by the LF-ILD. To validate the latter claim, we quantify the errors in near-field ILD spectra computed from far-field spectra using only LF-ILD computed using our formula as a correction factor. Specifically, we use this approach to compute near-field spectra for a few values of ρ, and 181 equally-spaced values of Θ between 0 and 180°, and compute the absolute error between these spectra and benchmarks computed directly from the near-field RS-HRTF given by Eq. (2).

Figure 1 shows a plot of mean absolute ILD error versus μ for a few values of ρ and where the errors are averaged over all Θ. The errors indicate the penalties incurred when using only an LF-ILD to extrapolate ILD spectra computed from the far-field RS-HRTF given by Eq. (4) to the near-field. Alternatively, they show the extent to which near-field ILD spectra differ from far-field ones, ignoring any overall gain (i.e., the LF-ILD).

We see that the mean error is less than 1 dB for $\rho \geq 2$ and for most values of μ up to 12 ($f \approx 8$ kHz for $r_{\mathrm{RS}} = 0.09$ m). The mean error doesn't exceed 3 dB across all μ up to $\mu \approx 30$ (shown up to $\mu \approx 25$ for clarity) and increases with decreasing ρ. For $\mu < 0.4$, the mean error is 0 indicating that ILDs for these values of μ do not vary with ρ except for an overall gain given by the LF-ILD computed using our formula. Since $\rho = 10$ (i.e., a source distance of approximately 1 m for typical head sizes) is often considered in practice to be the beginning of the far-field, using the ILD spectra for $\rho = 10$ as the far-field benchmark (instead of $\rho \to \infty$ used here) will result in mean absolute ILD errors that are smaller than those shown in Fig. 1.

Subjective listening tests, which are beyond the scope of the present work, are required to definitively evaluate the influence of the ILD spectral errors shown in Fig. 1 on near-field distance perception in the free-field. However, knowing that humans are not capable of resolving changes in ILD smaller than 1 dB for a wide range of baseline ILDs [19], and for frequencies up to approximately 5 kHz [20], our results indicate that, in general, the variation of ILD spectra with source distance is not perceptually significant except for an overall gain given by the LF-ILD. From these results and the previously demonstrated dominance of low-frequency ILD as a distance cue [3–6], we suggest that using our formula for LF-ILD may be sufficient to render distance cues for many practical applications.

Finally, it is worth noting that our formula, along with the Woodworth formula for high-frequency ITD and the far-field spherical head-shadowing filter proposed by Brown and Duda [21] (and recently revised by Romblom and Bahu [22]), provides a *complete* solution for rendering virtual sound sources that is based primarily on modeling the listener's head by a rigid sphere. Such a solution is easy to implement and computationally very efficient, making it ideal for real-time applications on devices where computational resources may be limited. An added benefit is that all components of this solution, including our formula, are readily individualized by determining a listener-specific sphere radius, r_{RS}, from anthropometric data as described by Algazi et al. [23] or Sridhar and Choueiri [24], for example.

Acknowledgements

This work was jointly sponsored by Focal-JMlab and Tesla, Inc. The authors thank the anonymous reviewers for their feedback on an early draft of this manuscript.

References

[1] Duda, R. O. and Martens, W. L., "Range dependence of the response of a spherical head model," *The Journal of the Acoustical Society of America*, 104(5), pp. 3048–3058, 1998, doi: 10.1121/1.423886.

[2] Brungart, D. S. and Rabinowitz, W. M., "Auditory localization of nearby sources. Head-related transfer functions," *The Journal of the Acoustical Society of America*, 106(3), pp. 1465–1479, 1999, doi:10.1121/1.427180.

[3] Zahorik, P., Brungart, D. S., and Bronkhorst, A. W., "Auditory Distance Perception in Humans: A Summary of Past and Present Research," *Acta Acustica united with Acustica*, 91(3), pp. 409–420, 2005, ISSN 1610-1928.

[4] Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., and Pardhan, S., "Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss," *Attention, Perception, & Psychophysics*, 78(2), pp. 373–395, 2016, doi:10.3758/s13414-015-1015-1.

[5] Brungart, D. S., Durlach, N. I., and Rabinowitz, W. M., "Auditory localization of nearby sources. II. Localization of a broadband source," *The Journal of the Acoustical Society of America*, 106(4), pp. 1956–1968, 1999, doi:10.1121/1.427943.

[6] Brungart, D. S., "Auditory localization of nearby sources. III. Stimulus effects," *The Journal of the Acoustical Society of America*, 106(6), pp. 3589–3602, 1999, doi:10.1121/1.428212.

[7] Kim, H.-Y., Suzuki, Y., Takane, S., and Sone, T., "Control of auditory distance perception based on the auditory parallax model," *Applied Acoustics*, 62(3), pp. 245–270, 2001, ISSN 0003-682X, doi:https://doi.org/10.1016/S0003-682X(00)00023-2.

[8] Enzner, G., Antweiler, C., and Spors, S., "Trends in Acquisition of Individual Head-Related Transfer Functions," in J. Blauert, editor, *The Technology of Binaural Listening*, pp. 57–92, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, ISBN 978-3-642-37762-4, doi:10.1007/978-3-642-37762-4_3.

[9] Xie, B., *Head-related transfer function and virtual auditory display*, J. Ross Publishing, 2013, ISBN 978-1-60427-070-9.

[10] Spagnol, S., Tavazzi, E., and Avanzini, F., "Distance rendering and perception of nearby virtual sound sources with a near-field filter model," *Applied Acoustics*, 115, pp. 61 – 73, 2017, ISSN 0003-682X, doi:https://doi.org/10.1016/j.apacoust.2016.08.015.

[11] Kan, A., Jin, C., and van Schaik, A., "A psychophysical evaluation of near-field head-related transfer functions synthesized using a distance variation function," *The Journal of the Acoustical Society of America*, 125(4), pp. 2233–2242, 2009, doi:10.1121/1.3081395.

[12] Romblom, D. and Cook, B., "Near-Field Compensation for HRTF Processing," in *Audio Engineering Society Convention 125*, 2008.

[13] Poirier-Quinot, D. and Katz, B. F. G., "The Anaglyph Binaural Audio Engine," in *Audio Engineering Society Convention 144*, 2018.

[14] Aaronson, N. L. and Hartmann, W. M., "Testing, correcting, and extending the Woodworth model for interaural time difference," *The Journal of the Acoustical Society of America*, 135(2), pp. 817–823, 2014, doi:10.1121/1.4861243.

[15] Woodworth, R. S. and Schlosberg, H., *Experimental psychology*, Oxford and IBH Publishing, 1954, ISBN 9788120417083.

[16] Kuhn, G. F., "Model for the interaural time differences in the azimuthal plane," *The Journal of the Acoustical Society of America*, 62(1), pp. 157–167, 1977, doi:10.1121/1.381498.

[17] Cooper, D. H. and Bauck, J. L., "On Acoustical Specification of Natural Stereo Imaging," in *Audio Engineering Society Convention 65*, 1980.

[18] DLMF, "NIST Digital Library of Mathematical Functions," http://dlmf.nist.gov/, Release 1.0.23 of 2019-06-15, 2019, F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[19] Hershkowitz, R. M. and Durlach, N. I., "Interaural Time and Amplitude jnds for a 500-Hz Tone," *The Journal of the Acoustical Society of America*, 46(6B), pp. 1464–1467, 1969, doi:10.1121/1.1911887.

[20] Yost, W. A. and Dye, R. H., "Discrimination of interaural differences of level as a function of frequency," *The Journal of the Acoustical Society of America*, 83(5), pp. 1846–1851, 1988, doi:10.1121/1.396520.

[21] Brown, C. P. and Duda, R. O., "A structural model for binaural sound synthesis," *IEEE Transactions on Speech and Audio Processing*, 6(5), pp. 476–488, 1998, ISSN 1063-6676, doi:10.1109/89.709673.

[22] Romblom, D. and Bahu, H., "A Revision and Objective Evaluation of the 1-Pole 1-Zero Spherical Head Shadowing Filter," in *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, 2018.

[23] Algazi, V. R., Avendano, C., and Duda, R. O., "Estimation of a Spherical-Head Model from Anthropometry," *J. Audio Eng. Soc*, 49(6), pp. 472–479, 2001.

[24] Sridhar, R. and Choueiri, E., "Capturing the Elevation Dependence of Interaural Time Difference with an Extension of the Spherical-Head Model," in *Audio Engineering Society Convention 139*, 2015.