Audio Engineering Society

Convention Express Paper 97

Presented at the 154th Convention
2023 May 13–15, Espoo, Helsinki, Finland

This Express Paper was selected on the basis of a submitted synopsis that has been peer reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This express paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Dataset and analysis of acoustic intensity vector measurements around an upscaled ear model

Aaron Geldert[1], Marton Marschall[1], and Ville Pulkki[1]

[1]*Acoustics Lab, Department of Information and Communications Engineering, Aalto University*

Correspondence should be addressed to Aaron Geldert (aarongeldert@gmail.com)

ABSTRACT

This work presents a dataset of acoustic intensity vector measurements taken around a 3D-printed, upscaled model of an ear. Impulse responses of the acoustic vector were collected in an anechoic chamber in a dense grid of measurement positions around the pinna. Automated interleaved swept-sine measurements were carried out using an arc of loudspeakers, a robotic mover, a probe consisting of a pressure sensor and three anemometer particle velocity sensors, with the ear mounted on a turntable. In total, 145 unique directions of arrival and 225 measurement points were considered. Results show that the measured data is suitable for high resolution time, frequency, and spatial analyses. The full dataset containing impulse responses, ear model geometry, calibration data, and example processing scripts is shared in an open-access repository. Potential applications include the analysis and visualization of pinna-related transfer functions, as well as the validation of numerical acoustical modeling techniques.

1 Introduction

For decades, the measurement, analysis, and application of the head-related transfer function (HRTF) has been of interest to researchers in spatial audio technologies. The HRTF characterizes the response of the ears, head, and torso on sound arriving from different directions. Often, the HRTF is obtained by measuring human subjects (e.g. [1]), artificial models (e.g. [2]), numerical simulation of the acoustics in a model geometry (e.g. [3]), or combinations of these methods (e.g. [4]).

Much research has focused on studying the peaks and notches in the HRTF that strongly depend on the sound direction of arrival (DOA). as well as each individual's ear anthropometry [5, 6], but open questions remain about the true sound field around the ear. The exact mechanism that produces an elevation-dependent notch in the HRTF has not been well explained. Destructive interference from sound reflected off of the pinna is a suggested theory [7], supported by the weak correlation between the frequency of the lowest notch (N1) in the HRTF and the depth of the pinna over different source DOA [8]. However, distinct reflections in the pinna are challenging to discriminate with human-scaled measurements.

In this work, an upscaled ear model was measured to perform higher resolution measurements that otherwise would have been disturbed by the measurement equipment at normal scale. Scale acoustic measurements are often performed for architectural acoustics, where downscaled models are used to predict the responses of much larger (and costly) structures [9]. In either case, a direct rescaling of frequencies and wavelengths allows measurements of different dimensions to be compared analogously.

The ear model was 3D printed from nylon at seven times the original scale of a scanned human ear. The model contains the full pinna and ear canal with a rigid termination. A previous work [10] used the same ear model and probe examined a smaller set of source DOAs and measurement positions, demonstrating the suitability of the model for more detailed study.

Measurement of the acoustic vector of pressure and three-dimensional particle velocity allows for a detailed understanding of the local sound field over time. The instantaneous acoustic intensity vector is found via the product of the pressure and velocity to examine how acoustic energy propagates. This work aims to facilitate future analysis and visualization of the sound field in and around the anatomical structures of the pinna by collecting a detailed set of acoustic intensity measurements near the pinna.

This paper is structured as follows: Section 2 details the methodology of the measurement and processing of the responses. Section 3 shows features of the collected data and results from time-frequency analyses. Section 4 summarizes the relevance of the dataset for future research applications. Section 5 explains how to use this open-access dataset and its accompanying scripts.

2 Methods

2.1 Setup and coordinate system

Measurements were conducted in anechoic chamber "Lampio" at the Aalto Acoustics Labs in Espoo, Finland. The internal dimensions are 8.2 x 8.2 x 7.2 meters and has a background noise level below 0 dB SPL. The setup consisted of the 3D printed ear model mounted to a turntable in the center of the room, with a suspended arc of seven loudspeakers, a robotic mover, and an probe consisting of a pressure sensors and three particle velocity sensors.

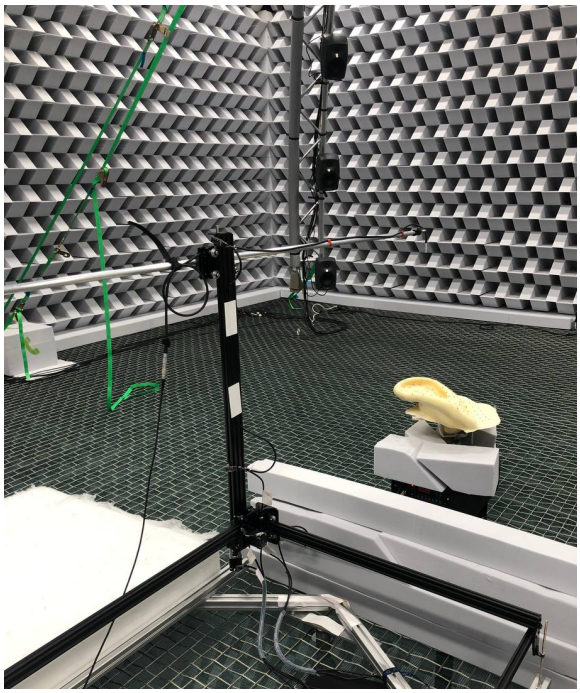


Fig. 1: Setup of the mover and probe (raised to the calibration position), loudspeakers, and ear model in the anechoic chamber.

Measurements were collected over a set of directions of arrival of the sound source, with the probe positioned at a set of 225 measurement points near the ear for each DOA. A right-handed cartesian coordinate system was defined relative to the ear (X "forward", Y "up", Z "right"), with the origin defined at a reference point near the tragus, on the interaural axis. A polar-lateral angular coordinate system was used to define the DOAs relative the ear, in which the polar angle ϕ represents the rotation around the interaural axis, and the lateral angle θ represents the angle away from the median plane, as shown in Figure 2. Note that the coordinates are centered around the entrance to the ear canal, rather than the center of the head.

The ear was mounted on the turntable to be rotated about the interaural axis. Rotations were performed in 15° increments from $\phi = [0, 330]^\circ$ for a total of 24 polar DOA angles.

The sound sources were seven loudspeakers (8030A, Genelec Oy, Iisalmi, Finland) mounted along an arc suspended from the ceiling. All sources were at a radius of 3.2 meters from the center of the ear model, at 15° intervals from $\theta = [0, 90]^\circ$. The vertically aligned ($\theta =$

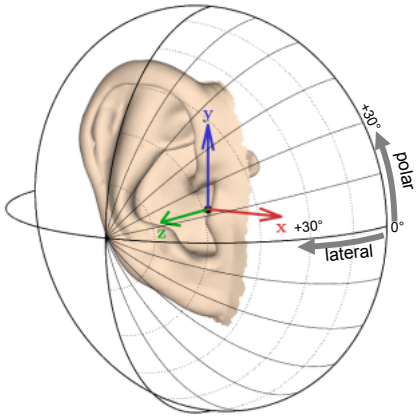


Fig. 2: The coordinate systems being used: DOAs are defined using polar-lateral angles, while the origin and measurement grid are defined with right-handed cartesian axes.

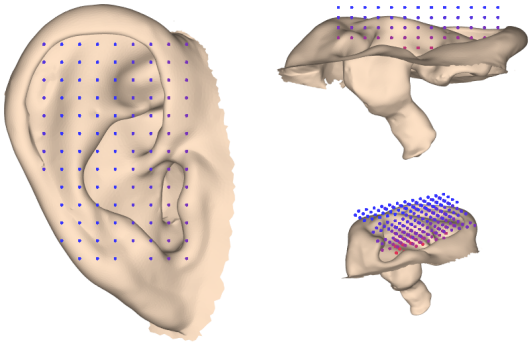


Fig. 3: The 225 measurement points, colored by depth in the Z axis. The grid has 25 mm intervals in the X and Y axes, and 19 mm intervals along the Z axis, extended to the deepest plane accessible by the probe.

90°) loudspeaker collected redundant measurements with each turntable rotation.

A 3D sound intensity probe (Microflown USP Regular, Microflown Technologies, Arnhem, Netherlands), was mounted on a robotic mover capable of positioning it in three dimensions around the ear model. The probe consists of an omnidirectional pressure microphone (2.56 mm electret condenser) and three Microflown MEMS particle velocity sensors based on a heated-wire anemometer design. The four sensors are mounted near-coincident within the cylindrical body of the probe (12.7 mm diameter), allowing for uniquely high resolution, localized measurement.

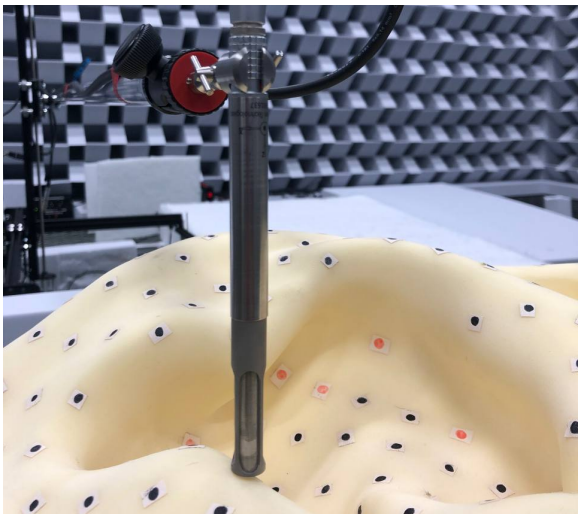


Fig. 4: The mover with probe being calibrated with the ear model. The sensors are 10 mm above the bottom of the probe housing.

The mover was assembled from an Open Builds ACRO 1010 frame, with X, Y, and Z axis movement ranges of 800 mm. The accuracy of positioning the probe using the mover, and its alignment with the ear coordinate system at different turntable rotations was tested extensively. The probe position was never found to exceed 1 mm of error during measurement movement sequences.

Signals from the probe were amplified with a 4-channel preamplifer (MFPA-4, Microflown Technologies) and then routed to an audio interface (Fireface UFX+, RME Audio) for analog-to-digital conversion, and recorded at a 48kHz sample rate in MATLAB.

To reduce the impact of the measurement setup on the free-field conditions of the anechoic chamber, damping material was placed next to the structures of the loudspeaker arc and mover, as well as inside of the arc's metal trusses and probe's boom arm.

2.2 Measurement procedure

Measurements were conducted in 24 batches, corresponding to each turntable rotation (polar DOA angle). In each batch, impulse response measurements were collected at each of the 225 measurement positions around the ear for each of the seven source loudspeakers at a given turntable orientation. An interleaved 15 second exponential sine sweep from 10 Hz to 24 kHz was used, with a 1 second interleaving delay between

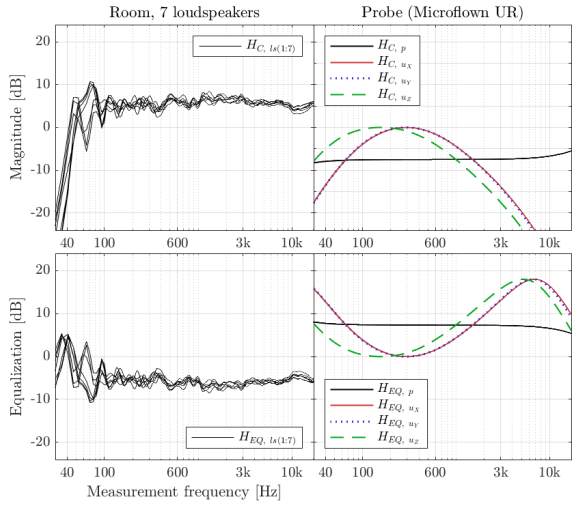


Fig. 5: The responses (top) of the room and 7 source loudspeakers (left), pressure and velocity sensors (right) to be calibrated. The corresponding regularized equalization filters (bottom) were applied to measurements during preprocessing.

loudspeakers. After capturing each measurement, the robotic mover automatically moved the probe to the next position, pausing to allow mechanical settling before collecting the next measurement. Between batches, the turntable rotation was completed, the absorption material was manually repositioned, and the mover was recalibrated.

The anechoic chamber was remotely monitored via a webcam. Occasionally, a batch of measurements was restarted due to a software failure. Each batch took about two hours to complete, and all measurements were completed over six consecutive days. Temperature and humidity conditions were not monitored but are believed to have not varied substantially.

2.3 Signal processing

The measured responses were convolved with an inverse swept-sine signal to extract the impulse response in a standard method [11]. The impulse responses were time windowed to a uniform length of 8192 samples (171 ms). The start of the window was aligned with a loopback measurement to preserve the acoustic propagation time. A raised cosine fade-in and fade-out was applied to the first 256 samples and last 1023 samples of each response.

Preprocessing was performed to reduce the influence of the loudspeaker, room, and probe responses. Calibration measurements of the loudspeaker and room responses $H_{C,\,ls}$ were taken for each of the seven loudspeakers using a GRAS 46AF 1/2" free-field measurement microphone (GRAS S&V, Holte, Denmark) positioned at the origin of the ear coordinate system, oriented in the direction of each loudspeaker, without the turntable or ear model present. The responses of each of the probe's four sensors, $H_{C,\,p}$ and $H_{C,\,u_{XYZ}}$, were estimated based on a parametric model provided by the manufacturer during a prior device calibration, and particularly show that the velocity sensors have a pronounced band-pass response.

The frequency band of interest is from 100 Hz to 3000 Hz. Equalization allows for a nearly flat (white) magnitude response in this bandwidth, at the cost of noise amplification.

An equalization filter, H_{EQ}, was calculated based on each response to be equalized, H_C, using a standard Tikhonov regularization [12]:

$$H_{EQ} = \left| \frac{H_C^*}{|H_C|^2 + \beta^2} \right|, \tag{1}$$

where β is the regularization parameter and $*$ denotes the complex conjugate. The regularization parameter β is calculated as in [13]:

$$\beta = \sqrt{\frac{1-\sqrt{1-\alpha^{-2}}}{1+\sqrt{1-\alpha^{-2}}}}, \quad \alpha = 10^{G/20}, \tag{2}$$

in which G maximum allowable white noise gain, in decibels. For loudspeaker and room responses, $G_{ls} = 6$ dB, and for the probe, $G_{probe} = 18$ dB. The corresponding equalization filters were applied to all measurements during preprocessing.

The 3D particle velocity channels were then rotated to align the sensor coordinate system with the ear coordinate system for each measurement DOA. The relative orientation of the probe and loudspeaker arc was kept fixed during all measurements, with the turntable and ear rotating relative to the probe as the polar angle of the DOA varied. Therefore, two correcting signal rotations were needed: a fixed correcting rotation to align the probe response with the source's DOA (which was found by aligning the time-average intensity vector of the response), and a counterrotation of the signal according to the turntable orientation. The X and Y

probe responses were rotated according to the rotation matrix of the necessary combined rotation angle for each polar response. As the rotation was applied via a linear combination of the X and Y sensor responses, it was performed only after the probe response had been calibrated.

The polarity of the velocity signals was reversed to align the sensors' responses with the direction of the traveling wave. This step ensures that the time-average intensity vector points away from the sound source.

The analogous measurements at the dimensions of human anthropometry can be obtained through a direct rescaling of frequencies and wavelengths by a factor of seven. The relevant frequency band in the given dataset is suggested to be between 100 Hz and 3 kHz, corresponding to an anthropometrically scaled 700 Hz to 21 kHz, which contains the relevant features attributed to the pinna-related transfer function.

3 Results and analysis

3.1 Estimation of the PRTF

The measured and preprocessed pressure responses were examined at measurement point 76, the point deepest in the cavum concha and considered to be the closest to the ear canal entrance. This point was chosen to be comparable to the microphone position during blocked meatus HRTF measurements. For source DOAs in the median plane ($\theta = 0°$), the tragus obstructs the line-of-sight path and the response is comprised of diffracted and reflected waves.

The mean response near the entrance to the ear canal was calculated as a weighted average of the responses over all source directions from the ipsilateral hemisphere. The weighting corrected for the non-uniform spatial sampling of the dataset, where points are clustered near the interaural axis.

When rescaled to HRTF frequencies, the response remains near unity below 1kHz, and exhibits the characteristic ear canal peak at 3kHz, notch at 9kHz, and concha peak near 6kHz described by Shaw [14].

3.2 Responses in the median plane

The pinna related transfer function is known to exhibit prominent changes in the frequency and depth of notches above 3kHz, between sources in the median (XY) plane. Figure 7 shows that such variation is also present in the given dataset.

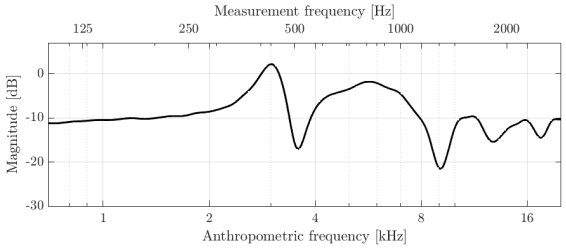


Fig. 6: Mean magnitude response, averaged over all ipsilateral source directions, near the ear canal entrance. The dual frequency axes correspond to the measured upscaled model (top) and the analogous human ear dimensions.

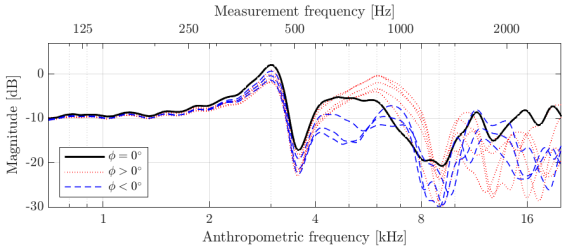


Fig. 7: Magnitude responses of the PRTF in the median plane, where the solid line is the frontal response, and DOAs above (red, dotted) and below (blue, dashed) the horizon are distinguished. The high frequency peaks and notches show variation in the median plane. The dual frequency axes correspond to the measured upscaled model (top) and the analogous human ear dimensions.

3.3 Instantaneous intensity

The instantaneous sound intensity can be directly obtained from the product of the pressure and velocity vector, as in $\vec{I}(t) = p(t)\vec{u}(t)$. Sound intensity is a vector quantity useful for examining the transfer of energy through space.

In Figure 8, the intensity vector field (of one plane of measurement points) shows the sound intensity of the arriving wavefront directed away from the source.

3.4 Near field effects

It is important to note that the sound source was positioned at 3.2 m from the center of the turntable, which corresponds to an anthropometric distance of a source 46 cm from the entrance to the ear canal. Therefore,

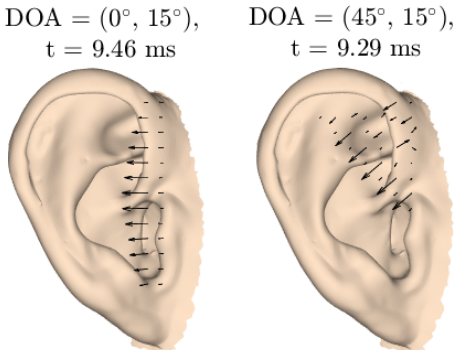


Fig. 8: The instantaneous intensity vector field of the impulse response for 2 DOAs for one plane of measurement points, at a moment showing the arrival of the wavefront.

the source cannot be assumed to be in the far-field, in contrast with many HRTF measurements.

First, the dimensions of the measurement grid mean that the true source DOA may differ between measurement points by up to 5.7°. The shape of the propagating wavefront is therefore spherical, rather than planar. This effect is partially shown on the left side of Figure 8, where the instantaneous intensity of the wavefront does not match between points aligned along the Y axis, due to differences in distance from the source. The stimulus level may also differ by up to 1 dB between points due to distance attenuation.

Further, with a spherical wave stimulus, there is an inherent phase shift between the pressure p and velocity u (in the direction of the traveling wave). If the source is assumed to be an ideal monopole, the phase shift is given by:

$$\angle\left(\frac{p}{u}\right) = \frac{ikr}{1+ikr}, \tag{3}$$

where the wavenumber $k = \frac{2\pi f}{c}$, $c = 343$ m/s, $r = 3.2$ m, and i is the imaginary unit. At 100 Hz, this phase shift is nearly 10°, and falls below 1° above 1 kHz.

4 Applications and conclusions

The presented work demonstrates a unique approach combining acoustic scale modeling and acoustic vector measurements to study the sound field around a seven times upscaled model of an outer ear. The dataset contains impulse responses of the sound pressure and particle velocity for 145 unique source directions on one hemisphere around the ear. Measurements were performed at 225 positions in and around the cavities of the outer ear, with over 32,000 acoustic vector impulse responses captured in total.

A mesh of the ear model is included, as it may be useful for extracting anthropometric features or boundary element method (BEM) simulations to cross-validate the measurement results, in addition to visualizations.

The dataset is intended to enable many further analyses of the sound field near the surface of the outer ear, particularly for investigating the behavior of reflections and modes in the complex geometry of the pinna.

5 Dataset

The dataset is published in an open access repository at `https://doi.org/10.5281/zenodo.7564880`. It is organized as follows:

- **sofa_preprocessed:** 225 .sofa format files, corresponding to each measurement point, containing the pressure and velocity responses for all DOAs, with the preprocessing from 2.3 applied
- **mat_original:** MATLAB data files containing the original measurements, deconvolved into IRs and time windowed, without any further preprocessing
- **calibration_data:** measured loudspeaker/room responses and parametrically modeled probe responses to be used for calibration
- **model_data:** .stl and .txt files that specify the geometry of the ear model and measurement points (mm)

A repository of metadata and MATLAB scripts demonstrating loading, preprocessing, visualization, and basic analyses with the dataset are also available at `https://github.com/aaron-geldert/upscaled-ear-model-scripts`.

Acknowledgements

Thanks are extended to Aleksi Öyry for work on constructing the mover and measurement setup.

References

[1] Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C., "The cipic hrtf database," in *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575)*, pp. 99–102, IEEE, 2001.

[2] Spagnol, S., Purkhús, K. B., Unnthórsson, R., and Björnsson, S. K., "The viking HRTF dataset," in *16th Sound and music computing conference*, pp. 55–60, Sound and Music Computing Network, 2019.

[3] Takemoto, H., Mokhtari, P., Kato, H., Nishimura, R., and Iida, K., "Mechanism for generating peaks and notches of head-related transfer functions in the median plane," *The Journal of the Acoustical Society of America*, 132(6), pp. 3832–3841, 2012.

[4] Brinkmann, F., Dinakaran, M., Pelzer, R., Grosche, P., Voss, D., and Weinzierl, S., "A cross-evaluated database of measured and simulated HRTFs including 3D head meshes, anthropometric features, and headphone impulse responses," *Journal of the Audio Engineering Society*, 67(9), pp. 705–718, 2019.

[5] Møller, H., "Fundamentals of binaural technology," *Applied Acoustics*, 36(3-4), pp. 171–218, 1992, ISSN 0003682X, doi:10.1016/0003-682X(92)90046-U.

[6] Iida, K., *Head-related transfer function and acoustic virtual reality*, Springer, 2019.

[7] Spagnol, S., Geronazzo, M., and Avanzini, F., "On the relation between pinna reflection patterns and head-related transfer function features," *IEEE transactions on audio, speech, and language processing*, 21(3), pp. 508–519, 2012.

[8] Onofrei, M. G., Miccini, R., Unnthorsson, R., Serafin, S., and Spagnol, S., "3D ear shape as an estimator of HRTF notch frequency," in *17th Sound and Music Computing Conference*, pp. 131–137, Axea sas/SMC Network, 2020.

[9] Rindel, J. H., "Modelling in auditorium acoustics. From ripple tank and scale models to computer simulations," *Revista de Acústica*, 33(3-4), pp. 31–35, 2002.

[10] Rönkkö, P., *Measuring Acoustic Intensity Field in Upscaled Physical Model of Ear*, Master's thesis, Aalto University. School of Electrical Engineering, 2019.

[11] Farina, A., "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio engineering society convention 108*, Audio Engineering Society, 2000.

[12] Tikhonov, A. N. and Arsenin, V. Y., *Solution of ill-posed problems*, Winston & Sons, 1977.

[13] Moreau, S., Daniel, J., and Bertet, S., "3D sound field recording with higher order ambisonics–Objective measurements and validation of a 4th order spherical microphone," in *120th Convention of the AES*, pp. 20–23, 2006.

[14] Shaw, E. A., "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," *The Journal of the Acoustical Society of America*, 56(6), pp. 1848–1861, 1974.