

Analysis of spatial cue distortions in HRTFs using a simplified head model

Shaima'a Doma, Janina Fels

Institute for Hearing Technology and Acoustics, RWTH Aachen University, 52074 Aachen, Germany

Email: {shaimaa.doma, janina.fels}@akustik.rwth-aachen.de

Introduction

The question of accuracy of Head-Related Transfer Functions (HRTFs) arises both in the context of individual measurements as well as approximations thereof. Not only is it interesting to quantify inter-individual differences, but also to note the influence of different HRTF acquisition methods. Numerous distance metrics have been introduced for this purpose, evaluating the dissimilarity of data sets with respect to certain aspects, be it global or local spectral and/or temporal behavior. Nonetheless, a more detailed understanding on the nature of the error is hardly possible, which is rendered even harder by the geometric complexity of artificial heads, commonly used for evaluations of this kind. The present work suggests a simplified model of the head and conchal cavity to bypass this difficulty. Resonance behavior is examined for different cavity sizes and microphone displacements, both in acoustic measurements and Boundary Element Method (BEM) simulations. Discrepancies are highlighted and implications for HRTF acquisition are discussed.

The Resonator Sphere

A simplified head was constructed as a sphere with hollow cylinder extensions, all consisting of polyvinyl chloride (PVC), for which acoustic rigidity can be assumed in a first approximation. The sphere radius was chosen based on an effective head radius defined in [1] as

$$r_{\text{eff}} = 0.51w + 0.019h + 0.18d + 32 \text{ mm}, \quad (1)$$

with w , d and h corresponding to head width, depth and height, respectively. Inserting mean anthropometric values of the ITA HRTF database [2] lead to an approximate radius $r_{\text{sphere}} = 8 \text{ cm}$.

To enable excitation of different modes, cylinders of two sizes were designed (see Figure 1). They can be attached at $\pm 70^\circ$ azimuth relative to the frontal direction, with adapters for a Sennheiser KE-4 microphone in

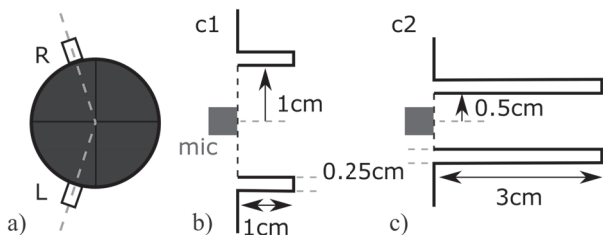


Figure 1: Top view of the *resonator sphere* (a) and dimensions of the two available cylinder types c1 (b) and c2 (c).

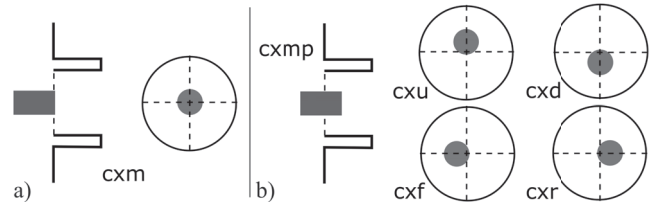


Figure 2: Microphone positions within the cylinder. Centered (a) and displaced (b) cases are defined for cylinder type $x = [1, 2]$ by abbreviations (m: mid, u: up, d: down, f: front, r: rear) and an additional p for axial protrusion.

each. These adapters allow for axial displacement (along the cylinder axis), as well as radial misalignment of the microphone (away from the cylinder axis), see Figure 2.

Analytical Approach

Three kinds of modes can be excited in a hollow cylinder: longitudinal (axial), transverse (radial and circumferential), as well as combinations thereof. The resonance frequencies depend on the cylinder radius r_{cyl} , its length L , and the boundary conditions surrounding the hollow space. A purely longitudinal mode, assuming an ideally rigid cylinder base and soft opening, would ensue at odd multiples of a $\lambda/4$ -resonance, at frequencies

$$f_{\text{long}} = \frac{(2n-1) \cdot c_0}{4L}, n \in N. \quad (2)$$

To account for effects at the orifice, a so-called "end correction" [3] is needed, substituting L for an effective length

$$L' = L + 0.8 \cdot r_{\text{cyl}}. \quad (3)$$

Purely transverse resonance frequencies can be derived from the solution of the Helmholtz equation, considering the boundary condition at the cylinder walls and the periodicity of 2π of the circumferential mode component [6]. The l -th solution, at which the m -th order Bessel function of the first kind fulfills the relation

$$\frac{J_m(\xi r_{\text{cyl}})}{J_{m+1}(\xi r_{\text{cyl}})} = \frac{\xi r_{\text{cyl}}}{m+1}, \quad (4)$$

defines the resonance frequency

$$f_{\text{trans}} = \frac{\xi_{l,m} \cdot c_0}{2\pi}, \quad (5)$$

with indices l and m describing the radial and circumferential behavior, respectively. The exemplary mode

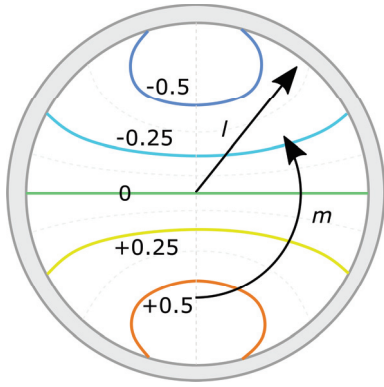


Figure 3: Ideal excitation of transverse mode $(l, m, n) = (1, 1, 0)$ in a hollow cylinder. Indices l and m indicate the magnitude variation in radial and circumferential direction, respectively. Adapted after [6].

($l = 1, m = 1$), depicted in Figure 3, shows two maxima separated by a spatial notch (running horizontally). It should be noted that the 2π periodicity allows for a rotation around the cylinder axis; this notch or other similar structures may therefore occur at an arbitrary angle, depending on the excitation.

Modes combining a transverse and a longitudinal component ($n > 0$) occur at

$$f_{\text{res}} = \sqrt{f_{\text{long}}^2 + f_{\text{trans}}^2}. \quad (6)$$

Measurements and Simulations

Acoustic measurements were performed in the hemi-anechoic chamber of the Institute for Hearing Technology and Acoustics, RWTH Aachen University. A 64-loudspeaker array (HRTF arc) rotated continuously around the object, while emitting time-interleaved exponential sweeps, achieving a resolution of 2.5° in azimuth and elevation, respectively [4]. To obtain free-field transfer functions, the KE-4 microphones were additionally mounted on a stand at the arc center for a reference measurement, the result of which was used in a subsequent division step.

The commercial software COMSOL Multiphysics 5.6 [5] was applied for BEM simulations. The *resonator sphere* was modeled as an object with an ideally rigid boundary surface, surrounded by air ($T = 18^\circ\text{C}$). A frequency dependent mesh allowed a maximum edge size of $\lambda/5$ on the cylinders and a more tolerant upper limit of $\lambda/4$ on the sphere surface, while complying with additional geometric constraints. Following the reciprocity principle, the virtual microphone – a circular surface at the base of the left cylinder – served as a sound source, with a radius of $r_{\text{mic}} = 2\text{ mm}$ approximating the radius of the effective KE-4 microphone surface.

Results and Discussion

Expected resonance behavior and the results observed in measurements and simulations are discussed in the following.

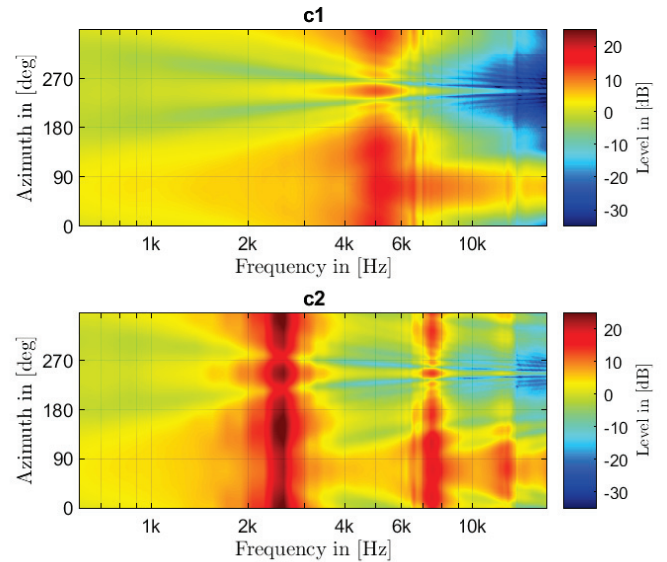


Figure 4: Horizontal plane transfer functions of the *resonator sphere* with both cylinder sizes (c1 and c2). Longitudinal modes show almost omni-directional behavior, superimposed by the sphere directivity.

Centered Microphone

Longitudinal resonances in a hollow cylinder ideally only exhibit axial changes in magnitude. Within each cross-section, respectively, a constant sound pressure level (SPL) can be assumed. Accordingly, the directivities at f_{long} merely show a pressure amplification with little spatial variation, besides the typical effects of the sound propagation around a sphere. Interference patterns occur on the averted side, including a localized increase in magnitude, the so-called *bright spot*. Shadowing mainly affects higher longitudinal modes (harmonics).

In the measured data, longitudinal resonance behavior is in alignment with these idealized assumptions, as Figure 4 shows for states c1m and c2m, measured with centered microphones. This also holds true for the numerical simulations, which, however, differ from the measured results in two aspects: they possess more pointed spectral peaks and a shift in resonance frequencies. The first deviation can be explained by the non-ideal boundary conditions of the real object, as opposed to the perfectly rigid model surface. The measurement thereby captures a superposition of multiple resonances lying closely together, leading to lower resonance quality. To investigate the second discrepancy, the resonance frequencies, extracted from the measured and simulated spectra, were contrasted with the analytical solution, see Table 1. The analytical solutions with the original cylinder lengths ($L_1 = 1\text{ cm}$ and $L_2 = 3\text{ cm}$) deviate from the measured results by 68% and 13% for c1 and c2, respectively. Applying orifice correction, the analytical solution with effective lengths $L'_1 = 1.8\text{ cm}$ and $L'_2 = 3.4\text{ cm}$ is well-aligned with the measurements. The residual error for the shorter cylinder indicates that the corrective factor, adjusting the length based on the radius, might not be optimal when the radius and length are in a similar range.

Table 1: Comparison of fundamental frequencies (first longitudinal resonances). Analytical solutions are based on eq. (2) with and without orifice correction.

	c1m	c2m
Measured	5.10 kHz	2.53 kHz
Numerical	6.37 kHz	4.54 kHz
Analytical (L)	8.56 kHz	2.86 kHz
Analytical (L')	4.76 kHz	2.52 kHz
Measured (axial shift)	5.34 kHz	2.70 kHz
Numerical (axial shift)	6.55 kHz	5.00 kHz

A deviation between the numerical solution and its measured and analytical counterparts remains an open question. A mismatch in the speed of sound can be ruled out as a reason, since a room temperature of 18 °C was applied in both calculations and an approximately equal temperature can be assumed for the hemi-anechoic room at the time of measurement. While longitudinal modes play an incidental role in HRTFs, they may appear at high frequencies. The shown discrepancy might thus be present between measured and simulated HRTFs, yet largely undetected due to spatial and spectral complexity.

Radial Displacement

As mentioned in the introduction, the cylinders in this model are meant to resemble a simplified conchal cavity. When HRTF measurements are performed, a microphone is placed at the ear canal entrance, which is typically situated more closely to the lower end of the conchal cavity. The centered microphone is therefore a less accurate representation, as opposed to a microphone shifted along the sphere surface. In terms of longitudinal modes, the transfer functions are expected to show no change of behavior due to this radial shift. Hence, the focus is, in the following, laid on transverse modes, where changes in magnitude within the cylinder cross-section should be observed upon excitation.

As opposed to the longitudinal modes discussed in the previous section, the first transverse resonance for cylinder c1 can be identified at an identical frequency for both the measurement and simulation. Figure 6 depicts the spatial behavior of the spectral magnitude at a frequency $f_{\text{trans}} = 11.6$ kHz. What stands out first is that the centered microphone fails to detect the mode. From a reciprocal viewpoint, as applied in BEM, this is understandable: the sound source imposes a maximum at the cylinder axis, where the mode requires a spectral minimum to form. Once the source position is moved from the center (here by 2 mm), excitation is possible; this is visible in the middle and right sub-figures for a downward (c1d) and frontal (c1f) displacement, respectively. The directivity in the c1d case clearly resembles the mode $(l, m, n) = (1, 1, 0)$, previously shown in Figure 3. In the c1f case, the pattern seems to be rotated around the cylinder axis. This is again comprehensible for the reciprocal case; the notch line forms vertically, "dodging" the

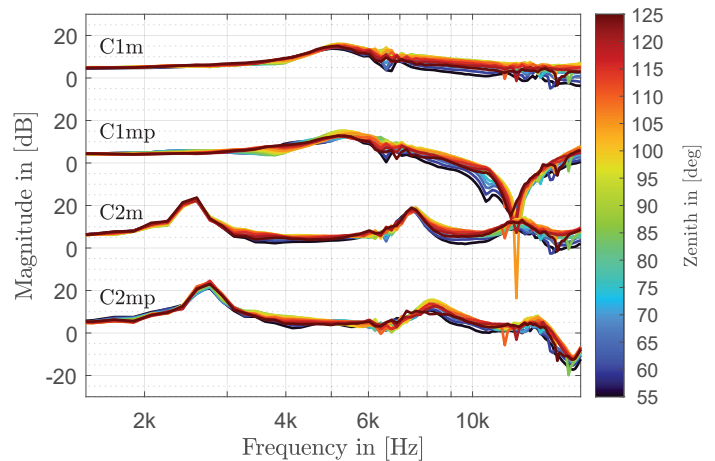


Figure 5: Transfer functions perpendicular to the orifice ($\varphi = 70^\circ$), measured with (cxmp) and without (cxm) an axial protrusion of the microphone by 3 mm. The axial shift causes a change in fundamental frequency and enables transverse mode excitation in the measured data.

pressure maximum imposed by the source.

It should be noted that, in spite of similarity to the mode, the analytical solution (eq. 5) leads to a supposed resonance frequency of $f_{\text{trans}'} = 13.1$ kHz, differing from the measured and simulated data. Non-ideal boundary conditions can be excluded as an explanation, since the simulation already assumes an ideally rigid surface, in contrast to the measurement, yet both showed excitation at the same frequency. The resonance frequency observed in the measurement and simulation corresponds to an effective cylinder radius of 0.6 cm instead of 1 cm, considering the given equation. To assess whether a generalized corrective factor would be suitable to also align higher transverse modes with the analytical solution, further investigations are needed.

Axial Displacement

Besides the position of the ear canal entrance within the conchal cavity, the depth of microphone insertion is also of relevance. A microphone protrusion into the cavity is therefore discussed in the following. Considering the compact design of the KE-4 microphone, a minor protrusion was expected to have little effect on mode development. In the case of the first longitudinal resonance, this would only mean a slightly lower magnitude, since the spatial maximum would still occur at the sphere surface. However, even an axial offset of only 3 mm alters the fundamental frequency, as Figure 5 shows. Thus, not only a change in the sampling point within the cylinder, but also a change in impedance at the closed end of the cylinder must be taken into account. This is remarkable, given that the surface of the microphone is relatively small compared to the cross-sectional area of the cylinder (4% for c1 and 16% for c2). Furthermore, the transverse mode emerges in spite of the microphone being centered within the cylinder. This was not observed for the corresponding simulated results.

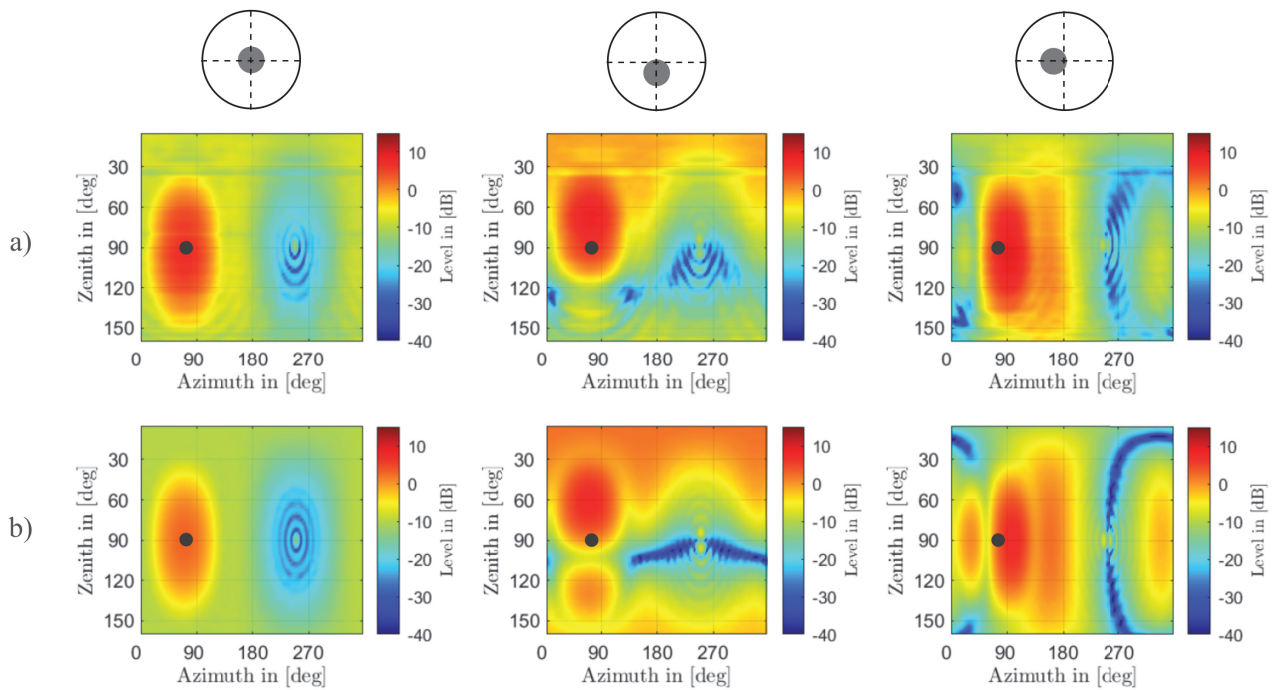


Figure 6: Directivities of transverse mode $(l, m, n) = (1, 1, 0)$ in cylinder $c1$ for a centered and radially displaced microphone. The small circles indicate the cylinder axis direction ($\varphi = 70^\circ, \vartheta = 90^\circ$). Measured (a)) and numerically simulated (b)) data show related excitation patterns for the displaced microphone cases, yet varying notch quality. This mode is not captured by a centered microphone.

Conclusion and Outlook

In this work, we used a simplified model of the head and conchal cavity to examine acoustic resonance behavior in measured and numerically simulated HRTFs. In individual HRTF measurements, the exact position of the microphone may vary between repeated runs; similarly, the properties of the virtual microphone may differ between simulations on the same object. Understandably, degrees of freedom in terms of positioning and orientation can result in discrepancies between the acquired transfer functions. The present study shed a light on these effects, which may go unnoticed in studies evaluating similarities based on regular HRTFs. In the presented data, measured longitudinal resonance frequencies corresponded to the analytical solution with orifice correction, while the numerical solutions differed unexpectedly. Transverse modes, on the contrary, showed similar behavior in the measurement and simulation, while deviating from the analytical solution of the respective mode. While regular HRTFs comprise more details than the presented object, with a superposition of different modes, similar effects can be expected, which in that case, however, would be harder to detect. Future investigations on the above-mentioned deviations, as well as attempts to align the measured, simulated and analytical results would help optimize these approaches, with the goal of minimizing the rather complex error found between HRTFs acquired in different ways.

Funding Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project no. 402811912.

References

- [1] Algazi, V.R.; Avendano, C. and Duda, R.O.: Estimation of a spherical-head model from anthropometry. *Journal of the Audio Engineering Society*, 49.6 (2001)
- [2] Bomhardt, R.; de la Fuente Klein, M. and Fels, J.: A high-resolution head-related transfer function and three-dimensional ear model database. *Proceedings of the 172nd Meeting of the Acoustical Society of America* (2016)
- [3] Kuttruff, H.: *Acoustics: an introduction*. CRC Press (2007)
- [4] Richter, J.-G. and Fels, J.: On the influence of continuous subject rotation during high-resolution head-related transfer function measurements. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2019)
- [5] COMSOL Multiphysics Reference Manual, version 5.6, COMSOL, Inc, www.comsol.com
- [6] Rona A.: The Acoustic Resonance of Rectangular and Cylindrical Cavities. *Journal of Algorithms & Computational Technology* (2007)