Audio Engineering Society

Convention Paper

Presented at the 131st Convention
2011 October 20–23 New York, USA

Two Physical Models for Spatially Extended Virtual Sound Sources

Jens Ahrens and Sascha Spors

Deutsche Telekom Laboratories, Technische Universität Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany

Correspondence should be addressed to Jens Ahrens (jens.ahrens@telekom.de)

ABSTRACT

We present physical models for the sound field radiated by plates of finite size and spheres vibrating in higher modes. The intention is obtaining a model that allows for controlling the perceived spatial extent of a virtual sound source in model-based sound field synthesis. Analytical expressions for the radiated sound fields are derived and simulations of the latter are presented. An analysis of interaural coherence in a virtual listener, which has been shown to serve as an indicator for perceived spatial extent, provides an initial proof of concept.

1. INTRODUCTION

The control of the perceived spatial extent of a phantom source in Stereophony is a widely exploited aesthetic feature in various styles of music [1]. A number of different audio effects such as reverberation and chorus – or a combination thereof – can be applied in order to achieve the desired result. This rather intuitive approach is not well suited for more advanced spatial audio presentation methods like sound field synthesis where a parametric description is desired such as provided by the MPEG-4 AudioBIFS standard [2].

Finding an appropriate terminology for the present context is not straightforward. The term *apparent source width*, the linguistic meaning of which seems relevant at first sight, is not applicable here. It rather describes the influence of early reverberation on the broadening of an auditory event, mostly in the context of concert hall acoustics, e.g. [3]. In a similar, yet not identical, meaning *spaciousness* is used [4, Sec. 4.5.1, p. 348].

Tab. 2-1 in [5] suggests *diffuseness* and *blur* for naming the perception that is desired to be evoked by the methods presented in this paper. Though, these terms may be assumed not to be optimal. A related

and extensive discussion can be found in [6], where the term *individual source width* is proposed. We are not aware of indications that spatial extent of an auditory event is restricted to width. We therefore propose the term *perceived spatial extent* of an auditory event similarly to [7]. Spatial extent may refer to width, height, and potentially also to depth.

A number of recent approaches that allow for the parametric control of the perceived spatial extent of a virtual or phantom source split a given source into several sources, which are driven with decorrelated versions of the input signals. Refer e.g. to [8, 7] and references therein. Results can indeed be stunning [9].

In this contribution, we provide an alternative approach that employs physical models of extended sound sources vibrating in complex spatial modes. Such an approach might provide more explicit control over properties like the orientation of a given extended source. Although we focus on sound field synthesis methods like Wave Field Synthesis [10], Near-field Compensated Higher Order Ambisonics [11], or the Spectral Division Method [12], the presented results may also be beneficial in other audio presentation methods like Stereophony or traditional Ambisonics [13].

An essential point to mention here is the fact that we are aiming at modeling a sound source that *sounds* spatially extended. There are indications that some physical models of extended sources may not achieve this result. An example is a pulsating sphere, the sound field of which closely related to that of a monopole source [14, Eq. (6.119), p. 213]. It may be assumed that such a pulsating sphere also sounds similar to a monopole source.

Besides others, low *interaural coherence* has been shown to be an indicator for large perceived spatial extent [4]. In order to provide a first proof of concept, we analyze the interaural coherence in a virtual listener exposed to the sound fields treated in the present study. Additionally, audio examples are provided online.

2. PLATES VIBRATING IN HIGHER MODES

The first type of extended sound source we treat is a plate of finite size vibrating in higher modes. In order to model the higher-mode vibration, we divide the plate into sections of equal size, which vibrate with equal amplitude but with alternating algebraic sign. Of course, a complex amplitude can also be assigned to each vibrating section.

For convenience, we assume no z-dependency of the vibration. Though, the presented results can be straightforwardly extended to include also higher modes in z-direction.

We assume the plate to be located in the x-z-plane and to extend from $-L/2$ to $L/2$ in x-direction. Other source positions an orientations can be straightforwardly achieved by an appropriate translation and rotation of the coordinate system. Refer to Fig. 1 for an illustration of the latter.

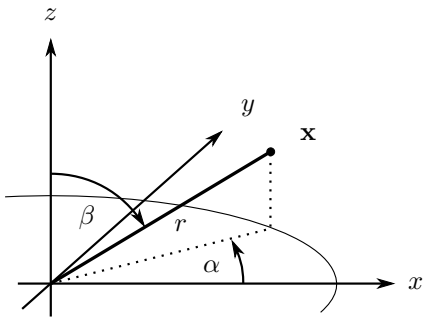


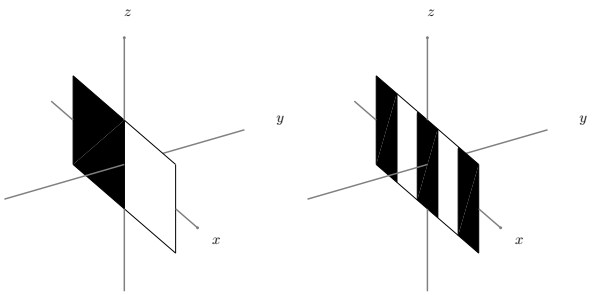
Fig. 1: The coordinate system used in this paper.

The mode number $\eta \in \mathbb{N}_0$ reflects the number of vibration nodes apparent in x-direction apart from the boundaries of the plate. The case of $\eta = 0$ thus represents a plane that vibrates "in-phase". Refer to Fig. 2 for an illustration.

From a physical point of view, we want to dictate the sound velocity in y-direction at the surface of the plate and calculate the evolving sound pressure field. We thus impose a *Neumann boundary condition* [14].

The relationship between the surface velocity $V(\mathbf{x}_0, \omega)$ of an infinite plate and the evolving sound field $S(\mathbf{x}, \omega)$ is given by *Rayleigh's first integral formula* as [14, (2.75), p. 36]

$$S(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V(\mathbf{x}_0, \omega)\, G_0(\mathbf{x} - \mathbf{x}_0, \omega)\; dx_0\, dz_0\,. \tag{1}$$



(a) $\eta = 1$ (b) $\eta = 4$

Fig. 2: Schematic illustration of higher-mode vibration of a plate of finite size for different mode numbers η. Black areas denote positive sign of vibration; white areas denote negative sign of vibration.

$\mathbf{x}_0$ denotes a position on the plate $\mathbf{x}$ a position in space and $G_0(\mathbf{x}-\mathbf{x}_0,\omega)$ the *free-field Green's function*, which is given by

$$G_0(\mathbf{x}-\mathbf{x}_0,\omega) = \frac{1}{4\pi}\frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|}}{|\mathbf{x}-\mathbf{x}_0|} \;, \tag{2}$$

Note that, strictly spoken, the quantity $V(\mathbf{x}_0,\omega)$ in (1) does not represent the surface velocity in y-direction but is rather directly proportional to it. For convenience, we will nevertheless speak of a velocity.

Since most sound field synthesis systems are restricted to horizontal-only synthesis, we apply the well-know stationary phase approximation on (1) [14, 10] in order to obtain an approximation of (1) in the horizontal plane. The latter is given by

$$S\left(\mathbf{x}\big|_{z=0},\omega\right) \approx \sqrt{\frac{2\pi d}{i\frac{\omega}{c}}} \int_{-\infty}^{\infty} V(x_0,\omega)\, G_0\left(\mathbf{x}\big|_{z=0} - \mathbf{x}_0\big|_{z_0=0},\omega\right)\, dx_0 \;, \tag{3}$$

whereby d denotes a reference distance. Note that (3) is termed a *2.5-dimensional approximation* of the Rayleigh integral and may be interpreted as the sound field of a line source of infinite length that is located along the x-axis and vibrates section-wise with alternating algebraic sign. Explicitly modeling spatial extent in z-direction influences the distance attenuation of the emitted sound field, which is not of special interest in the present study. Note that we omit the square root in (3) in the remainder for notational simplicity.

Recall that we want to model a source of finite size vibrating at mode number η. We denote the radiated sound field $S_\eta(\mathbf{x},\omega)$. We define a number of $\eta+1$ sections with index l, which extend along $x_l \le x \le x_{l+1}$ with

$$x_l = -\frac{L}{2} + l\frac{L}{\eta+1} \;\forall 0 \le l \le \eta \;. \tag{4}$$

The sound field $S_\eta^l(\mathbf{x},\omega)$ emitted by the section with index l is modeled as a spatial windowing of a line source of infinite extent. $S_\eta^l(\mathbf{x},\omega)$ is given by [15, 14]

$$S_\eta^l(\mathbf{x},\omega) = \int_{-\infty}^{\infty} w_l(x_0)\, \underbrace{V(x_0,\omega)}_{=(-1)^l}\, G_0(\mathbf{x}-\mathbf{x}_0,\omega)\, dx_0 \;, \tag{5}$$

whereby $w(x_0)$ denotes a window function that is given by

$$w_l(x) = \begin{cases} 1 & \text{for } x_l \le x \le x_{l+1} \\ 0 & \text{elsewhere .} \end{cases} \tag{6}$$

Eq. (5) transformed to wavenumber domain reads [16]

$$\tilde{S}_\eta^l(k_x,y,z,\omega) = (-1)^l\, \tilde{w}_l(k_x)\, \tilde{G}_0(k_x,y,z,\omega) \;. \tag{7}$$

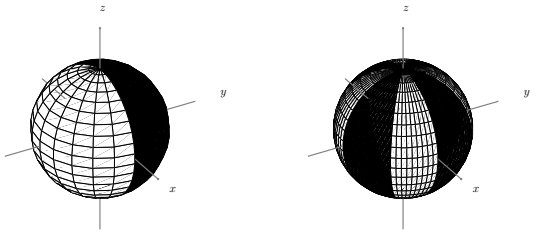
The Fourier transformation $\tilde{w}_l(k_x)$ of $w_l(x)$ is given by [14]

$$\tilde{w}_l(k_x) = \begin{cases} x_{l+1} - x_l & \text{for } k_x = 0 \\ \frac{1}{ik_x}\left(e^{ik_x x_{l+1}} - e^{ik_x x_l}\right) & \text{elsewhere} \end{cases} . \tag{8}$$

The sound field $\tilde{S}(k_x,y,z,\omega)$ in wavenumber domain evolving when all vibrating sections are combined is finally given by

$$\tilde{S}_\eta(k_x,y,z,\omega) = \tilde{G}_0(k_x,y,z,\omega) \sum_{l=0}^{\eta} \tilde{S}_\eta^l(k_x,y,z,\omega) \;. \tag{9}$$

$S_\eta(\mathbf{x},\omega)$ and $s_\eta(\mathbf{x},t)$ can then be obtained via numerical Fourier transforms. $\tilde{G}_0(k_x,y,z,\omega)$ is given by (17) in App. A.



(a) $\eta = 1$ (b) $\eta = 4$

Fig. 3: Schematic illustration of higher-mode vibration of a sphere for different mode numbers η. Black areas denote positive sign of vibration; white areas denote negative sign of vibration.

3. SPHERES VIBRATING IN HIGHER MODES

Another simple geometry of an extended sound source is the sphere, which may yield different properties than the vibrating plate e.g. with respect to the perceived orientation and directivity. Again, we want to dictate the velocity at the surface of the vibrating sphere and thus impose a Neumann boundary condition. In this case however, it is not obvious how a 2.5D simplification can be performed. The 3D case is therefore considered whereby we assume that the surface velocity at the surface of the spherical source under consideration is independent of the colatitude β. A denotes the radius of the sphere, which we assume to be centered around the coordinate origin. Other source positions can be straightforwardly achieved by an appropriate translation of the coordinate system.

In the following, $\eta \in \mathbb{N}_0$ refers to the vibration mode number of the spherical source under consideration. For $\eta = 0$, the classical case of a pulsating ("breathing") sphere evolves, which has been solved e.g. in [14, 17]. In higher modes, we split the sphere into 2η sections of equal size that vibrate with equal amplitude but with alternating algebraic sign. Fig. 3 schematically depicts two such spheres that vibrate with $\eta = 1$ and $\eta = 4$.

For notational clarity, we assume that all vibration modes exhibit equal amplitude in the following derivation. In practice, any complex amplitude can be assigned to each vibrating section of each vibration mode. Additionally, the sections of a given mode can be arbitrarily rotated along the azimuth.

The relationship between the velocity $V(\beta, \alpha, \omega)$ at the surface of a sphere of radius A and the radiated sound field $S_\eta(\mathbf{x}, \omega)$ is given in [14, Eq. (6.106), p. 210] as

$$S_\eta(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{i \breve{V}_n^m(\omega)}{\frac{\omega}{c} h_n^{(2)'}\left(\frac{\omega}{c} A\right)} h_n^{(2)}\left(\frac{\omega}{c} r\right) Y_n^m(\beta, \alpha) \; . \quad (10)$$

$h_n^{(2)'}(\cdot)$ denotes the derivative of the Hankel function with respect to the argument; $\breve{V}_n^m(\omega)$ denotes the spherical harmonics expansion coefficients of $V(\beta, \alpha, \omega)$ defined as [14]

$$V(\beta, \alpha, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \breve{V}_n^m(\omega) \, Y_n^m(\beta, \alpha) \; . \quad (11)$$

Note that in the present case $V(\beta, \alpha, \omega) = V(\alpha, \omega)$ and the coefficients $\breve{V}_n^m(\omega)$ can be determined to be [14]

$$\begin{aligned} \breve{V}_n^m(\omega) &= \int_0^{2\pi} \int_0^{\pi} V(\alpha, \omega) Y_n^{-m}(\beta, \alpha) \sin\phi \; d\phi \; d\theta \\ &= (-1)^m \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} \; \Psi_n^m \; \chi^m(\eta) \end{aligned} \quad (12)$$

with

$$\Psi_n^m = \int_0^{\pi} P_n^{|m|}(\cos\beta) \sin\beta \; d\beta \quad (13)$$

and

$$\chi^m(\eta) = \sum_{l=0}^{2\eta-1} (-1)^l \int_{\alpha_l}^{\alpha_{l+1}} e^{-im\alpha} \; d\alpha \quad (14)$$

$\forall \eta > 0$ and with $\alpha_l = l\frac{\pi}{\eta}$. Ψ_n^m and $\chi^m(\eta)$ are given by (21) and (22) derived in App. B.

Result for $\eta = 0$, i.e. the sound field of a pulsating sphere, is given by [14, Eq. (6.119), p. 213]

$$S_0(\mathbf{x}, \omega) = \frac{h_0^{(2)}\left(\frac{\omega}{c} r\right)}{\frac{\omega}{c} h_0^{(2)'}\left(\frac{\omega}{c} A\right)} V(\omega) \quad (15)$$

Note that rotation of vibrating sections along the azimuth is achieved by replacing $\breve{V}_n^m(\omega)$ with

$\breve{V}_n^m(\omega)e^{-im\alpha_{\text{rot}}}$ in (12), whereby α_{rot} denotes the rotation angle [18, Eq. (3.3.31), p. 127].

4. RESULTS

In this section, we present a set of simulations of spatially extended sources as treated above in order to illustrate their basic properties.

4.1. Emitted Sound Fields

Fig. 4 and 6 show sample monochromatic sound fields emitted by a truncated line source and a spherical source respectively for different vibration modes.

It can be observed that a sound field with strong spatial variation can indeed be achieved. This is especially true if more than one mode are excited for a given frequency. Refer to Fig. 4(c), which shows a weighted superposition of $S_1(\mathbf{x},\omega)$ and $S_4(\mathbf{x},\omega)$.

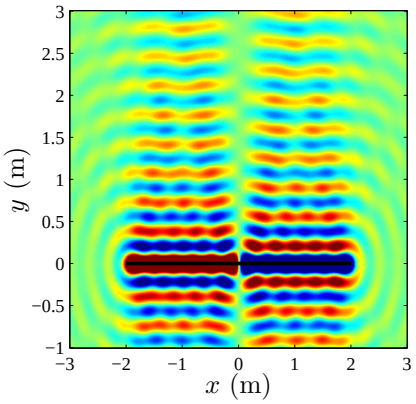
We emphasize that the ratio between the dimensions of the considered vibrating section and the wavelength λ at the considered frequency f is crucial.

At short wavelengths λ (high frequencies) and low vibration mode numbers η, the individual sections tend to emit highly directional wave fronts that do hardly overlap. Therefore, no considerable spatial variation arises. Refer to Fig. 4(a) for an example involving the line source. The properties of the spherical source are similar.

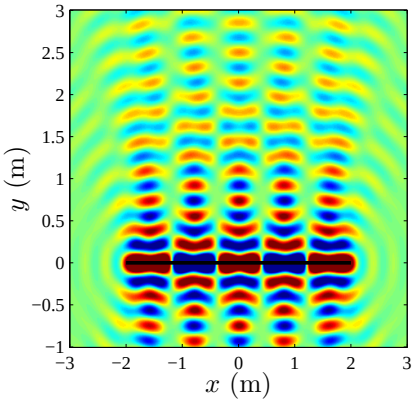
At long wavelengths λ (low frequencies), a high vibration mode numbers η will lead to vibrating sections that radiate similarly to point sources. Since adjacent sections exhibit opposite algebraic sign, they tend to cancel out each other's sound field and the complex source turn into an end-fire array [19]. Refer to Fig. 5 for an example involving the line source.

The ratio between wavelength and the length of the vibrating sections in Fig. 4(a) is approx. 0.17 and in Fig. 4(b) it is approx. 0.43. The ratio between wavelength and the radian measure along the azimuth of the vibrating sections in Fig. 6(a) is approx. 0.11 and in Fig. 6(b) it is approx. 0.44.

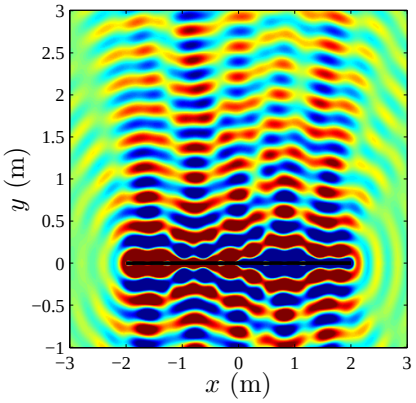
Informal experiments show that a ratio between 0.3 and 0.5 is reasonable when a single mode shall evoke a sound field with a complex structure. When several modes are added for a given frequency, then also other ratios can be useful.



(a) $S_1(\mathbf{x},\omega)$



(b) $S_4(\mathbf{x},\omega)$



(c) $S_1(\mathbf{x},\omega) + (0.5 + 1.5i)S_4(\mathbf{x},\omega)$

Fig. 4: Sound fields in the horizontal plane evoked by higher-mode vibration of a truncated line source of length $L = 4$ m located along the x-axis for $f = 1000$ Hz and different mode numbers η.

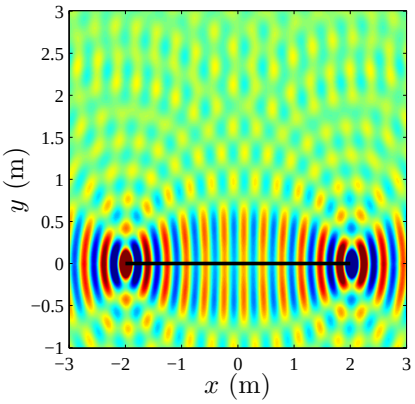


Fig. 5: Sound field in the horizontal plane evoked by higher-mode vibration of a truncated line source of length $L = 4$ m located along the x-axis for $f = 1000$ Hz and $\eta = 25$.

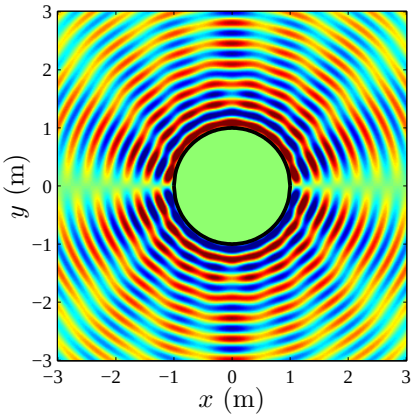
4.2. Interaural Coherence

As mentioned in the Introduction of this paper, it is essential that the source under consideration *sounds* spatially extended. In this section, we analyze the scenarios depicted in Fig. 4 and 6 with respect to the coherence of the ear signals of a virtual listener. The interaural coherence has been shown to be an indicator for a large perceived spatial extent [20]. The *magnitude squared interaural coherence estimate* $C_{lr}(\omega)$ may be defined as [21]

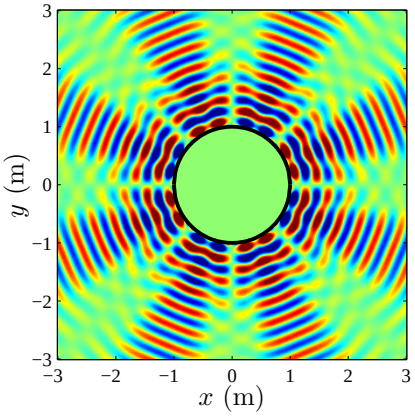
$$C_{lr}(\omega) = \frac{|P_{lr}(\omega)|^2}{P_{ll}(\omega) P_{rr}(\omega)} \,, \tag{16}$$

with $P_{lr}(\omega)$ denoting the cross power spectral density between the left-ear and right-ear signals and $P_{ll}(\omega)$ and $P_{rr}(\omega)$ denoting the power spectral density of the left and right ear signal respectively. $C_{lr}(\omega)$ is bounded between 0 and 1 whereby 0 indicates no coherence at all and 1 indicates perfect coherence.

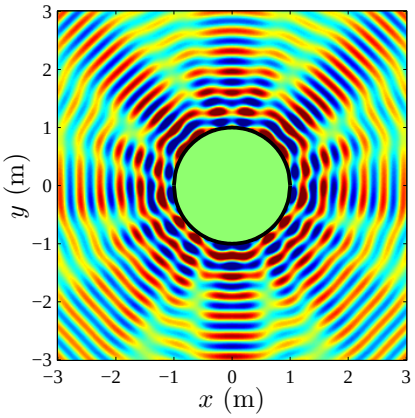
In order to obtain an estimate for the ear signals of a listener we simulate the impulse response of the path between the complex source under consideration and two locations that correspond to the locations of the listeners ears. More explicitly, we evaluated the positions $\mathbf{x} = [\pm 0.08\ 3]^T$ m, which are located at 3 m distance from the source and 16 cm apart from each other.



(a) $\eta = 1$



(b) $\eta = 4$



(c) $S_1(\mathbf{x}, \omega) + (0.5 + 1.5i) S_4(\mathbf{x}, \omega)$; $S_4(\mathbf{x}, \omega)$ has been rotated by $\alpha_{\text{rot}} = \pi/8$

Fig. 6: Sound field in the horizontal plane evoked by higher-mode vibration of a spherical source of radius $A = 1$ m centered around the coordinate origin for $f = 1000$ Hz and different mode numbers η.

Note that this procedure completely neglects scattering and diffraction at the listener's body. The deviation of these simulated data from the true ear signals depends on the orientation of the listener and can not be quantified. It has been shown in [22] that the coherence of the sound field at two locations in space, more precisely a measure derived from that coherence, does indeed exhibit a strong relation to the perceived spatial extent. We therefore assume that the results presented below hold also qualitatively for the actual ear signals.

For convenience, we present an analysis of the linear source exclusively. It can be shown that the results hold qualitatively for the spherical source as well. We examine broadband signals obtained from sampling (3) in time-frequency domain and performing a numerical inverse Fourier transform. Note that for all scenarios discussed in this section, audio examples can be downloaded from [23].

We superpose two vibration modes for a given frequency. The mode number η and the weights of the individual modes were chosen randomly but bounded to a useful range as specified in Sec. 4.1. Note that we experienced large variation of the results repeatedly calculated using constant parameters.

We divided the audible frequency range into static bands, the widths of which corresponds to the width of the critical bands at the given frequency range [4]. For all these bands, η and the weights were kept constant.

Fig. 7 and 8 show analyses of the sound field evoked by a line source of length $L = 4$ m and $L = 0.5$ m, respectively, vibrating in higher modes as described above. It can be seen that, for the chosen parameters, the $L = 4$ m source can indeed evoke low $C_{lr}(\omega)$ between $f = 600$ Hz and $f = 4000$ Hz, which may be considered as an important frequency range in terms of the perceived spatial extent [22, 20]. Although, $C_{lr}(\omega)$ is rather high for very low and for high frequencies in Fig. 7(b), preliminary experiments show that a careful choice of parameters can significantly reduce $C_{lr}(\omega)$ also in those regions. The impulse responses shown in Fig. 7(a) and 8(a) show that the vast part of the energy arrives within a few milliseconds.

Reducing the length L of the line source to 0.5 m does indeed increase the interaural coherence and might therefore lead to a smaller perceived spatial extent. Though, we emphasize that the details of the relationship between coherence and perceived extent are not known, e.g. [20].

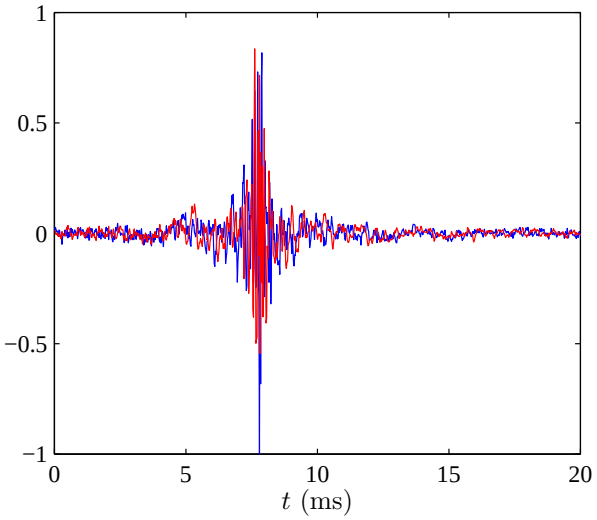
The origin of the notches in the interaural coherences shown in Fig. 7(b) and 8(b) is not clear at this stage. However, there are indication that a suitable choice of parameters avoids them.

5. CONCLUSIONS AND OUTLOOK

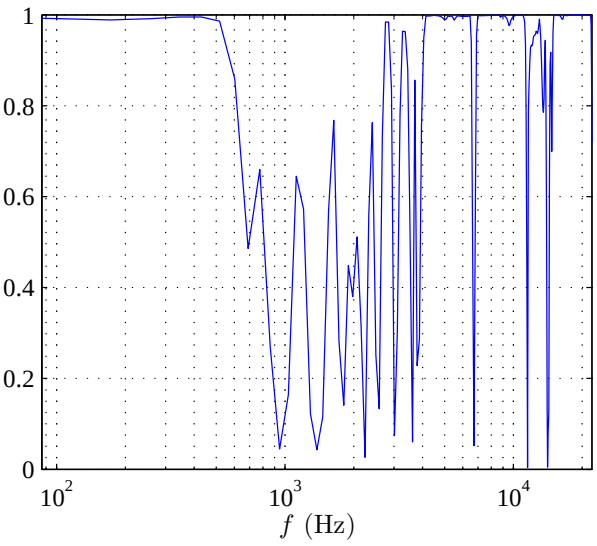
We have presented two physically motivated models for spatially extended virtual sound sources that are intended to be employed in model-based sound field synthesis. Explicitly, the two proposed models are an approximation of a plate as well as a sphere, both vibrating in higher spatial modes. A first analysis of the coherence of the resulting sound pressure between different locations in space suggests that the perception of spatial extent can indeed be evoked by the proposed sound source models. This statement has also been confirmed by informal listening. A crucial point to be investigated in future work is the appropriate selection of the involved parameters, such as the size of the model sound source and mode numbers and weights, that evoke a specific perception.

It is emphasized here that the proposed models are intended to be seen as an intermediate step on the way to the solution to the problem. This is mainly due to the fact that the implementational and computational complexity of the current models is considerable. Once the perception of spatial extent has been confirmed formally and the appropriate choice of the involved parameters is clear, it might be preferable to analyze the properties of the according sound fields in order to potentially deduce those properties that are related to the perception of spatial extent. It might then be possible to find simpler means of creating similar sound fields avoiding the presented complex models but preserving their statistical properties.

Finally, it has been proven in the literature that reverberation can exhibit the ability to increase the *apparent source width*, e.g. [3]. The combination of the proposed models with reverberation that supports large apparent source width seems promising and will be investigated by the authors.

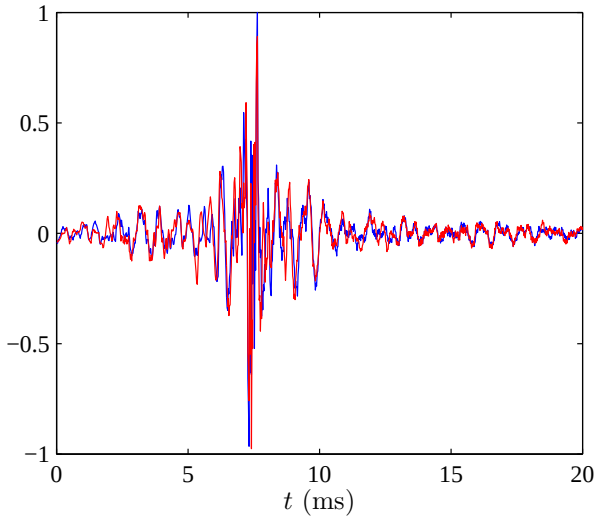


(a) Impulse responses

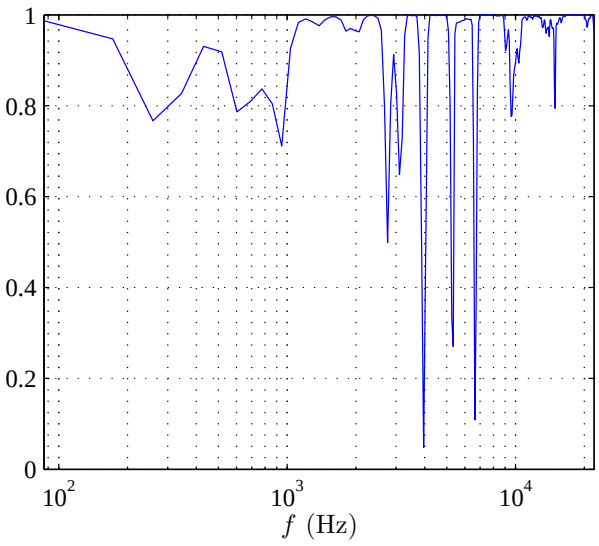


(b) $C_{lr}(\omega)$

Fig. 7: Analysis of the sound field evoked by higher-mode vibration of a line source of length $L = 4$ m. The observation points are $\mathbf{x} = [\pm 0.08 \;\; 3]^T$ m.



(a) Impulse responses



(b) $C_{lr}(\omega)$

Fig. 8: Analysis of the sound field evoked by higher-mode vibration of a line source of length $L = 0.5$ m. The observation points are $\mathbf{x} = [\pm 0.08 \;\; 3]^T$ m.

6. REFERENCES

[1] R. Izhaki. *Mixing Audio - Concepts, Practices and Tools*. Focal Press, Oxford, 2007.

[2] E.D. Scheirer, R. Väänänen, and V. Houpaniemi. AudioBIFS: Describing audio scenes with the MPEG-4 multimedia standard. *IEEE Trans. on Multimedia*, 1(3):237–250, September 1999.

[3] D. Griesinger. The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces. *Acustica*, 83(4):721–731, 1997.

[4] J. Blauert. *Spatial Hearing*. Springer, New York, 1997.

[5] H. Wittek. Perceptual differences between wavefield synthesis and stereophony. PhD thesis, University of Surrey, 2007.

[6] F. Rumsey. Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. *JAES*, 50(9):651666, 2002.

[7] M.-V. Laitinen, T. Pihlajamäki, C. Erkut, and V. Pulkki. Parametric time-frequency representation of spatial sound in virtual worlds. *submitted to ACM Trans. on Appl. Perception*.

[8] C. Verron, M. Aramaki, R. Kronland-Martinet, and G. Pallone. A 3-D immersive synthesizer for environmental sounds. *IEEE Trans. on Audio Speech and Lang. Proc.*, 18(6):1550–1561, August 2010.

[9] V. Pulkki. New spatial audio coding methods based on time-frequency processing. In *Workshop presented at the 40th Conference of the AES*, Tokyo, Japan, October 2010.

[10] A. J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *JASA*, 93(5):2764–2778, May 1993.

[11] J. Daniel. Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia [Representations of sound fields, application to the transmission and reproduction of complex sound scenes in a multimedia context]. PhD thesis, Université Paris 6, 2001. text in French.

[12] J. Ahrens and S. Spors. Sound field reproduction using planar and linear arrays of loudspeakers. *IEEE Trans. on Sp. and Audio Proc.*, 18(8):2038–2050, November 2010.

[13] M. A. Gerzon. With-height sound reproduction. *JAES*, 21:2–10, 1973.

[14] E. G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*. Academic Press, London, 1999.

[15] E. W. Start. Direct sound enhancement by wave field synthesis. PhD thesis, Delft University of Technology, 1997.

[16] B. Girod, R. Rabenstein, and A. Stenger. *Signals and Systems*. J.Wiley & Sons, New York, 2001.

[17] D. T. Blackstock. *Fundamentals of Physical Acoustics*. Wiley and Sons, Inc., New York, 2000.

[18] N. A. Gumerov and R. Duraiswami. *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier, Amsterdam, 2004.

[19] M.M. Boone, W.-H. Cho, and J.-G. Ih. Design of a highly directional endfire loudspeaker array. *JAES*, 57(5):309–325, May 2009.

[20] J. Blauert and W. Lindemann. Spatial mapping of intracranial auditory events for various degrees of interaural coherence. *JASA*, 79(3):806–813, March 1986.

[21] S.M. Kay. *Modern Spectral Estimation*. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[22] H. Riekehof-Boehmer and H. Wittek. Prediction of perceived width of stereo microphone setups. In *130th Convention of the AES*, London, UK, May 2011.

[23] Quality and Usability Lab. Spatial Audio Research - Audio research at Quality and Usability Lab. `http://audio.qu.tu-berlin.de/?p=664`.

[24] J. Ahrens. The single-layer potential approach applied to sound field synthesis including cases of non-enclosing distributions of secondary sources. Doctoral dissertation, Technische Universität Berlin, 2010.

[25] G. Arfken and H. Weber. *Mathematical Methods for Physicists*. Elsevier Academic Press, San Diego, 6th edition edition, 2005.

[26] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, San Diego, 2000.

7. APPENDICES

A. Statement of $\tilde{G}_0(k_x, y, z, \omega)$

From [24]:

$$\tilde{G}_0(k_x, y, z, \omega) = \begin{cases} -\frac{i}{4} H_0^{(2)}\left(\sqrt{\left(\frac{\omega}{c}\right)^2 - {k_x}^2}\sqrt{y^2+z^2}\right) \\ \qquad \text{for } 0 \leq |k_x| < \left|\frac{\omega}{c}\right| \\ \frac{1}{2\pi} K_0\left(\sqrt{{k_x}^2 - \left(\frac{\omega}{c}\right)^2}\sqrt{y^2+z^2}\right) \\ \qquad \text{for } 0 < \left|\frac{\omega}{c}\right| < |k_x| \ . \end{cases} \tag{17}$$

B. Derivation of Ψ_n^m and $\chi^m(\eta)$

Eq. (13) can be simplified via the substitution $u = \cos\beta$ as

$$\Psi_n^m = \int_{-1}^{1} P_n^{|m|}(u)\ du \ . \tag{18}$$

From the parity relation [25]

$$P_n^{|m|}(-u) = (-1)^{n+|m|} P_n^{|m|}(u) \tag{19}$$

we can deduce that the integral in (18) vanishes for $n + |m|$ being odd. Furthermore,

$$\int_{-1}^{1} P_n^{|m|}(u)\ du = 2 \int_{0}^{1} P_n^{|m|}(u)\ du \quad \forall n + |m| \text{ even} \ . \tag{20}$$

The solution to the integral on the right hand side of (20) is given in [26, 7.126-2] so that Ψ_n^m is finally given by

$$\Psi_n^m = \frac{\pi\, 2^{-2|m|}\, \Gamma(1+|m|+n)}{\Gamma\left(\frac{1}{2}+\frac{|m|}{2}\right)\Gamma\left(\frac{3}{2}+\frac{|m|}{2}\right)\Gamma(1-|m|+n)} \times {}_3F_2\left(\frac{|m|+n+1}{2}, \frac{|m|-n}{2}, \frac{|m|}{2}+1; |m|+1, \frac{|m|+3}{2}; 1\right) \quad \forall n+|m| \text{ even} \ , \tag{21}$$

and $\Psi_n^m = 0$ elsewhere. $\Gamma(\cdot)$ denotes the gamma function and ${}_3F_2(\cdot)$ the generalized hypergeometric function [25].

$\chi^m(\eta)$ can be determined to be

$$\chi^m(\eta) = \sum_{l=0}^{2\eta-1} (-1)^l \times \begin{cases} \alpha_{l+1} - \alpha_l & \forall m = 0 \\ \frac{i}{m}\left(e^{-im\alpha_{l+1}} - e^{-im\alpha_l}\right) & \forall m \neq 0 \end{cases} . \tag{22}$$