Audio Engineering Society

Convention e-Brief

Presented at the 132nd Convention
2012 April 26–29 Budapest, Hungary

# Sound Field Synthesis Toolbox

Hagen Wierstorf[1], Sascha Spors[2]

[1]*Assessment of IP-based Applications, T-Labs, TU Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany*

[2]*Quality and Usability Lab, T-Labs, TU Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany*

Correspondence should be addressed to Hagen Wierstorf (`hagen.wierstorf@tu-berlin.de`)

## ABSTRACT

An open source toolbox for Sound Field Synthesis (SFS) is introduced. The toolbox is able to numerically simulate sound fields synthesized by SFS methods like Wave Field Synthesis or higher order Ambisonics. Various loudspeaker driving signals for the mentioned methods are provided for 2-, 2.5- and 3-dimensional synthesis. The toolbox allows mono-frequent as well as broadband excitation signals. The latter allows to generate snapshots of the spatio-temporal impulse response of a chosen reproduction technique. The toolbox furthermore includes the computation of binaural room impulse responses (BRIR) for a given SFS setup. These can be used to simulate different sound field synthesis methods via binaural resynthesis. The toolbox is provided for Matlab/Octave and comes with an online documentation.

## 1. ACCESS

The latest version of the *Sound Field Synthesis Toolbox* can be downloaded at `https://dev.qu.tu-berlin.de/projects/sfs-toolbox/files`.

It comes with detailed built-in help, available via the `help` function within Matlab or Octave. An additional wiki page provides an online help, a tutorial for the first steps and different use cases at `https://dev.qu.tu-berlin.de/projects/sfs-toolbox/wiki`.

## 2. INTRODUCTION

Sound field synthesis (SFS) offers the possibility to create a determined sound field within an extended listening area. Common methods to reach this in reality are Wave Field Synthesis (WFS) or higher order Ambisonics (HOA) which apply different kinds of driving signals to a loudspeaker array in order to control the sound field within the listening area. For real loudspeaker setups the SoundScape Renderer (SSR) [1] can provide such driving signals. The SSR is an open source software, which is also developed by our groups.

To investigate the properties of different sound field synthesis methods and implementations, it is preferable to simulate their behavior beforehand instead of measuring the whole sound field afterwards. To close this gap, the *Sound Field Synthesis Toolbox* provides functions to numerically simulate WFS and HOA for 2-, 2.5-, and 3-dimensional synthesis (see e.g. [2]. These functions provide simulations for the sound fields of mono-frequent as well as broadband virtual sources. The latter allows to generate snapshots of the spatio-temporal impulse response for a given SFS method.

### REPRODUCIBLE RESEARCH

One motivation to release the *Sound Field Synthesis Toolbox* as an open source project was the principle of reproducible research [3, 4]. Like other fields that involve signal processing, the study of sound fields implies implementing a multitude of algorithms and running numerical simulations on a computer. As a consequence, the results obtained by this method are sometimes highly vulnerable to implementation errors. To ensure that other researches can testify the correctness of our results or easily reproduce them, the most straightforward approach is to publish the used code together with the results. Furthermore, the toolbox is a good starting point for anyone who wants to enter this research field.

As an alternative to applying expensive, real loudspeaker setups, dynamic binaural synthesis can simulate different loudspeaker setups to evaluate sound field synthesis methods (e.g. [5],[6]). Therefore, head-related impulse responses (HRIRs) are convolved with driving signals to simulate a synthesized sound field via headphones. To arrive at a more realistic scenario, a dynamic binaural synthesis system tracks the head orientation of the listener and chooses the appropriate HRIR pair for convolution with the source signals. In its *binaural room scanning* mode, the SSR is capable to operate as such a dynamic binaural synthesis system. As an input the SSR needs a binaural room impulse response (BRIR) data set, containing the simulated loudspeaker array, covering all possible orientations of the listener. The *Sound Field Synthesis Toolbox* can compute these BRIR data sets by applying the driving functions that are used for the simulation of the sound fields. As a consequence, the toolbox can ensure that the simulation of the sound field and the real stimuli presented to the listener are the same.

In order to be user-friendly the toolbox is written for Matlab and Octave, and a large documentation is presented within the code of the toolbox as well as on a web site, see Section 1.

## 3. SOUND FIELD SYNTHESIS

The theory of sound field synthesis assumes the existence of a listening area surrounded by elementary sound sources, referred to as secondary sources in the remainder of this paper. The question of how to drive these secondary sources to get a determined sound field within the listening area can answered by solving the following equation as [2]

$$P(\mathbf{x},\omega) = \int_{\partial V} D(\mathbf{x}_0,\omega)\; G(\mathbf{x}-\mathbf{x}_0,\omega)\; dS \;, \qquad (1)$$

where $\mathbf{x}$ describes a position within $V$, $\mathbf{x}_0$ the position of the secondary sources, $\partial V$ the surface of $V$, where the secondary sources are located, $dS$ an infinitesimal surface element, and $\omega = 2\pi f$ with temporal frequency $f$. The functions $D$ and $G$ denote the secondary source driving signal and the sound field emitted by a secondary source, respectively. $G$ is given by the three dimensional Greens function for a monopole source. $P$ is the desired sound field. There are different solutions to this integral. An implicit solution is WFS. In this case, the local spatial derivation of the wanted sound field determines the driving signal. An explicit solution is HOA, which solves the integral equation for explicit geometries like spherical or circular loudspeaker arrays. The desired sound field $P$ can be provided in terms of a measured one, or of a physical model of the sound field. Common physical models are those of a point source, a plane wave, or a focused source.

Depending on the geometry of the secondary source setup, 2-, 2.5-, or 3-dimensional synthesis is possible. For a 3-dimensional setup, a 2-dimensional surface of monopole sources synthesizes a sound field within a 3-dimensional listening area. Common loudspeakers approximate monopole sources, but normally only a 1-dimensional surface, like a linear loudspeaker array, is present. In such a case the 2.5-dimensional synthesis is applied. It is derived from the case of a 2-dimensional synthesis, but exchanging the line sources needed for the 2-dimensional case by monopole sources. As a consequence the 2.5-dimensional case exposes a wrong amplitude decay in comparison to the desired sound field.

As an example, the case of a plane wave as a source model for a 2.5-dimensional WFS setup is considered in the following. Its driving function is given by

$$D(\mathbf{x}_0,\omega) = \hat{S}(\omega) a(\mathbf{x}_0) g_0 \sqrt{\frac{\omega}{ic}}\, \mathbf{n}_{\mathbf{x}_s} \mathbf{n}\; e^{i\frac{\omega}{c}\mathbf{n}\mathbf{x}_0} \;, \qquad (2)$$

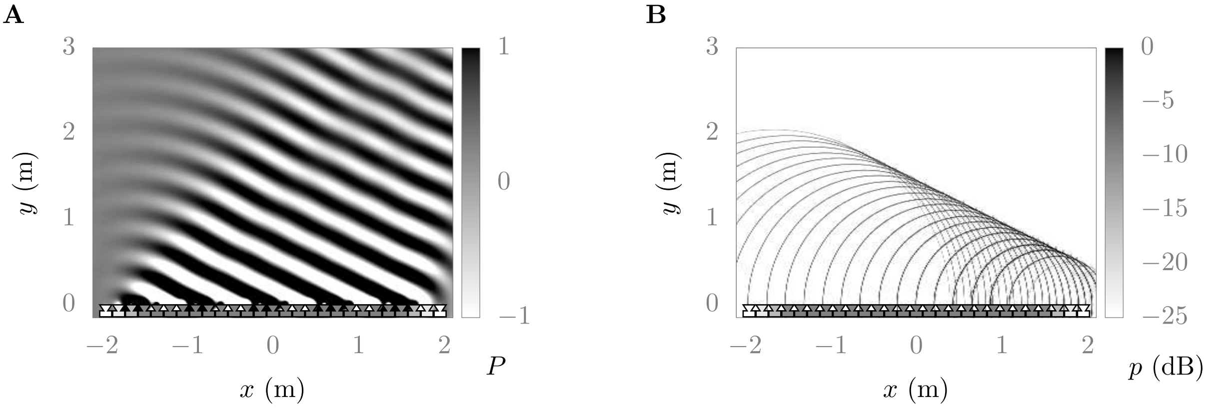

Fig. 1: The mono-frequent (A) and a snapshot in time of the spatio-temporal impulse (B) response of the loudspeaker array for the WFS driving function given in (2) is shown. For the mono-frequent part the frequency was $f = 1000\,\mathrm{Hz}$. The direction of the plane wave is $\frac{(1,0.5)\,\mathrm{m}}{\|(1,0.5)\|}$.

where $\hat{S}$ denotes the frequency spectrum of the plane wave, $a$ is a window function for weighting the secondary sources, $g_0$ is a constant chosen in such a way that the amplitude of the synthesized sound field is correct at a reference point, $\mathbf{n}_{\mathbf{x}_\mathrm{s}}$ the direction of the plane wave, and $\mathbf{n}$ the inward pointing normal vector of $\partial V$. For a detailed discussion of the window function and WFS driving functions for other source models and dimensions, see [7]. Fig. 1 A shows the result for a mono-frequent sound field with a frequency of $f = 1000\,\mathrm{Hz}$. Fig. 1 B presents a snapshot in time of the spatio-temporal impulse response, excited with a broadband click. In both cases the direction of the plane wave was $\mathbf{n}_{\mathbf{x}_0} = \frac{(1,0.5)\,\mathrm{m}}{\|(1,0.5)\|}$.

In order to simulate such sound fields, a numerical solution of (1) has to be derived. The implementation of this solution is discussed in the next section.

## 4. IMPLEMENTATION

| | WFS | | | HOA | | |
|---|---|---|---|---|---|---|
| | 2D | 2.5D | 3D | 2D | 2.5D | 3D |
| plane wave | x | x | | | x | |
| point source | x | x | | | x | |
| focused source | x | x | | | | |

Table 1: Driving functions for different source models implemented so far within the toolbox.

In order to solve (1) for a given set of the driving function $D$ and Greens function $G$, the integral is transformed to a summation by a spatial sampling of the surface $\partial V$. The solution is identical with the spatial continuous solution if the spatial sampling frequency is sufficiently high to handle the frequency content of the source.

Another sampling occurs of course in the time domain. For some applications of sound field synthesis, the typical sampling rate of 44100 Hz is not adequate. For such cases the toolbox implements a fractional delay filter as an alternative to fixed integer delays (e.g. WFS with a large number of loudspeakers).

Presently the toolbox contains driving functions for the source models *plane wave*, *point source*, and *focused source*. Tab. 1 presents the status of the implementation of these driving functions for WFS and HOA, depending on the dimensionality of the setup. The toolbox will incorporate the missing driving functions in a future version.

As mentioned in the introduction, the toolbox enables, besides the simulations shown in Fig. 1, binaural simulations by applying impulse response measurements. The result can then be presented via headphones to a listener. In this way, the properties of a sound field can be explored in the same way as when putting a dummy head at the same position within a real setup. In addition, listeners can evaluate different sound field synthesis methods and loudspeaker setups, of course with the restrictions binaural synthesis currently imposes [6, 8]. In order to investigate the influence of different rooms on the synthesis of the loudspeaker array, it is desirable to easily exchange the used HRIRs/BRIRs easily. The toolbox comes with its own Matlab/Octave based impulse response format and a set of functions to convert freely available HRIRs sets [9, 10, 11, 12] into this format.

Besides its main functionality, the *Sound Field Synthesis Toolbox* provides auxiliary functions, which can for example examine the interaural time differences (ITDs) and interaural level differences (ILDs) of a given HRIR set. For further perceptual analysis, the *Auditory Modeling Toolbox* [13] is a complementing instrument. Various plotting routines for Matlab/Octave and Gnuplot allow a presentation of the results in papers, as the example in Fig. 1 shows.

## 5. SUMMARY AND OUTLOOK

Numerical simulations of sound field synthesis methods on a computer enables to investigate these methods in a manner that is not possible in a purely analytical way and instead of setting up these systems. As a consequence, binaural synthesis based simulations allows such systems to be perceptually evaluated. All these simulations of course need the implementation of the algorithms to simulate the sound field synthesis methods and are therefore vulnerable to implementation errors. To allow the reproduction of the results gained by the simulations, it is desirable to publish the used code together with the results. The published code is also a good starting point for every one joining the research field of sound field synthesis. With the release of the *Sound Field Synthesis Toolbox* we hope to contribute to these requirements.

As Tab. 1 indicates, the implementation work has not been completed yet and future versions of the toolbox will contain the missing driving functions. The incorporating of HRIRs to the toolbox have indicated the problem of a missing common and open format for impulse responses. There is ongoing work and discussion on such a format [14, 15]. As soon as the community has agreed on one unique format, it will be included in the toolbox.

## 6. ACKNOWLEDGMENTS

This work was supported by DFG RA 2044/1-1.

## 7. REFERENCES

[1] Geier M, et al. The SoundScape Renderer: A Unified Spatial Audio Reproduction Framework for Arbitrary Rendering Methods. In *124$^{th}$ AES Conv.* 2008.

[2] Ahrens J. *Analytic Methods of Sound Field Synthesis*. Springer, Berlin, 2012.

[3] Donoho DL, et al. Reproducible Research in Computational Harmonic Analysis. *Comput Sci Eng*, 11(1):8–18, 2009.

[4] Ince DC, et al. The case for open computer programs. *Nature*, 482(7386):485–488, 2012.

[5] Wierstorf H, et al. Perception of Focused Sources in Wave Field Synthesis. *J Audio Eng Soc*, accepted.

[6] Völk F, et al. Simulation of wave field synthesis. *J Acoust Soc Am*, 123(5):3159, 2008.

[7] Spors S, et al. The Theory of Wave Field Synthesis Revisited. In *124$^{th}$ AES Conv.* 2008.

[8] Wierstorf H, et al. Psychoakustik der Wellenfeldsynthese: Vor- und Nachteile binauraler Simulation. In *DAGA*. 2012.

[9] Gardner WG and Martin KD. HRTF measurements of a KEMAR. *J Acoust Soc Am*, 97(6):3907–3909, 1995.

[10] Algazi VR, et al. The CIPIC HRTF database. In *WASPAA*, pp. 99–102. 2001.

[11] Kayser H, et al. Database of Multichannel In-Ear and Behind-the-Ear Head-Related and Binaural Room Impulse Responses. *EURASIP J Adv Signal Process*, 2009:1–11, 2009.

[12] Wierstorf H, et al. A Free Database of Head-Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances. In *130$^{th}$ AES Conv.* 2011.

[13] Søndergaard PL, et al. Towards a binaural modelling toolbox. In *FORUM ACUSTICUM*, pp. 2081–2086. 2011.

[14] Wefers F. A free , open-source software package for directional audio data. In *DAGA*, pp. 1059–1060. 2010.

[15] Andreopoulou A and Roginska A. Towards the Creation of a Standardized HRTF Repository. In *131$^{th}$ AES Conv.* 2011.