SOFiA Sound Field Analysis Toolbox

Benjamin Bernschütz[1,3], Christoph Pörschmann[1], Sascha Spors[2], Stefan Weinzierl[3]

[1]Cologne University of Applied Sciences, Germany - benjamin.bernschuetz@fh-koeln.de, christoph.poerschmann@fh-koeln.de

[2]Deutsche Telekom Laboratories / Technical University of Berlin, Germany - sascha.spors@telekom.de

[3]Technical University of Berlin, Germany - Audio Communication Group - stefan.weinzierl@tu-berlin.de

Abstract

Spherical microphone arrays and sound field decomposition using spherical harmonics have recently come into the focus of interest for the analysis of sound fields, room acoustics or spatial audio recording applications. For this purpose, the Sound Field Analysis Toolbox (SOFiA) has been developed and implemented. It comprises various functions for the analysis of sound fields using data from spherical microphone arrays. SOFiA is designed for MATLAB and consists of native m-functions and externals written in C/C++ yielding an efficient and consistent processing chain. An overview of the SOFiA processing chain is given as well as a closer discussion of selected details. As sound field analysis involves a considerable amount of signal processing, a detailed verification is crucial. A general approach for a gradual verification of the processing chain for sound field analysis is presented and applied to SOFiA. This verification process includes several steps from an analytic computation of spatial Fourier coefficients to the complete evaluation of predefined real measurement scenarios. The SOFiA toolbox software is freely available under the GNU GPL v3 license. For the evaluation of SOFiA, exemplary datasets from a VariSphear scanning array system are available under the Creative Commons license.

1 Introduction

We present a sound field analysis toolbox for MATLAB called SOFiA. The corresponding theory of sound field analysis is presented e.g. in [1], [2], [3]. The toolbox offers capabilities to process real data from spherical microphone arrays and to simulate the impact of a plane wave on a spherical array. The latter is useful for any kind of verification or simulation of spherical array problems. The first release presented here focuses on the fundamental processing toolchain and enables the user to render directional room impulse responses and to visualize sound field responses. Further extension of the toolbox is planned. The toolbox is released as free software under the GNU GPL v3 license and is available at [4] including an online documentation. Apart from the basic signal processing routines, various application examples are included and illustrated. Furthermore real microphone array datasets captured with a VariSphear scanning array system [5] are available under the Creative Commons license for experiments and evaluation. SOFiA is intended to serve academic or research purposes and is focused on sound field analysis of measured data, simulation of spherical arrays and teaching or studying the methods of sound field analysis. The first part of this paper introduces the software structure and gives a short description of the modules. In the second part a proposal for the gradual verification of a sound field processing chain is presented and applied to SOFiA.

2 Software architecture

SOFiA consists of different processing modules that can be individually arranged depending on the intended purpose. Some of the modules are implemented in M-code (the native MATLAB language). The computationally intensive modules are written in C/C++ and can be accessed through MEX files within MATLAB. These modules are called cores. The latter internally access the BOOST C++ library [6] for the computation of higher mathematical functions such as Bessel and Hankel functions, Legendre polynomials and spherical harmonics. Fig. 1 gives an overview of the structure and the features of the first release.

2.1 Data Import and Preprocessing

Different options are available to import microphone array data into SOFiA. The *readVSAdata* module is used to import VariSphear [5] datasets. For other types of microphone arrays the *mergeArrayData* module aligns the data to be processable by further structures. Both modules process time domain data and enable temporal downsampling and normalization. The downsampling option can be used to avoid spatial aliasing although at the expense of decreasing the audio bandwidth. The *F/D/T* (Frequency Domain Transform) module transforms the signals into the frequency domain using the fast Fourier transform, whereas only the half-sided spectrum is handed over as discussed in section 3.4. Temporal segments of the time signals to be included in the transform operation can be defined. This way the data can be analyzed in blocks or by implementing a running window, thereby facilitating the extraction of the temporal sound field structure as discussed in section 3.5.

2.2 Plane wave generators

Plane waves can be simulated using the plane wave generators *I/W/G* (Ideal Wave Generator) and *S/W/G* (Sampled Wave Generator). Both generate a full spectrum unity gain plane wave. The *I/W/G* module delivers ideal analytical Fourier coefficients as discussed in section 3.2 whereas the *S/W/G* module involves discrete spatial sampling as demonstrated in section 3.3. The wave generators are useful to test the signal processing chain or to set up simulation environments for spherical array problems.
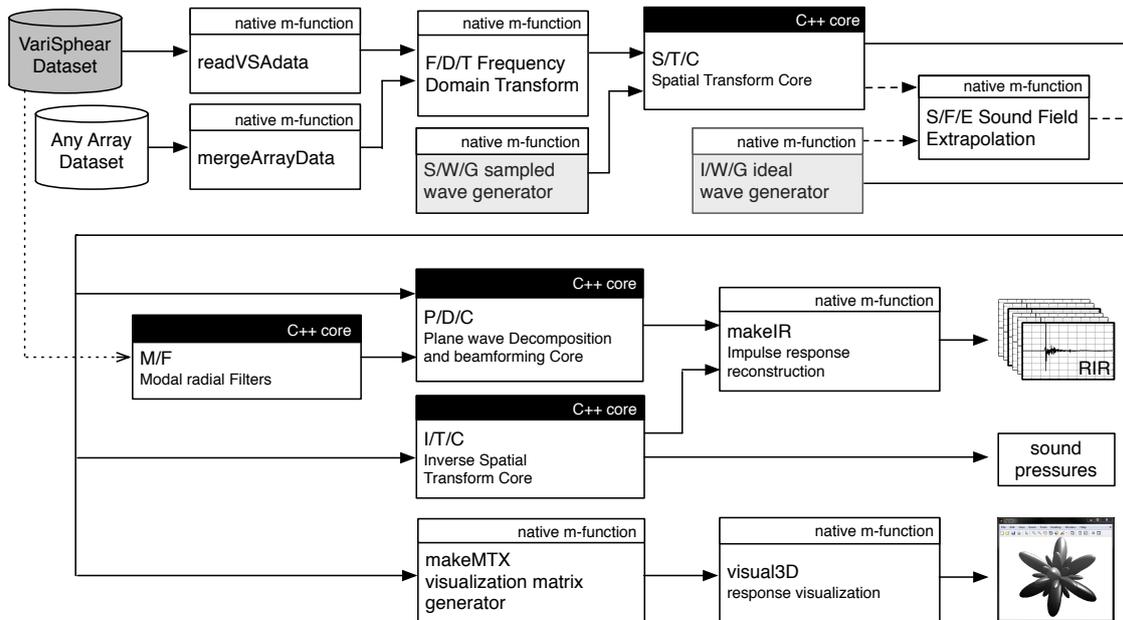
Figure 1: SOFiA module and signalflow overview of the first release.

2.3 (Inverse) Spatial Fourier Transform

An important operation in sound field analysis is the spatial Fourier transform [1]. SOFiA includes a spatial Fourier transform core ($S/T/C$) and an inverse spatial Fourier Transform core ($I/T/C$). Both are optimized for sound field analysis and offer fast calculation speed on the one hand and high numerical accuracy on the other hand.
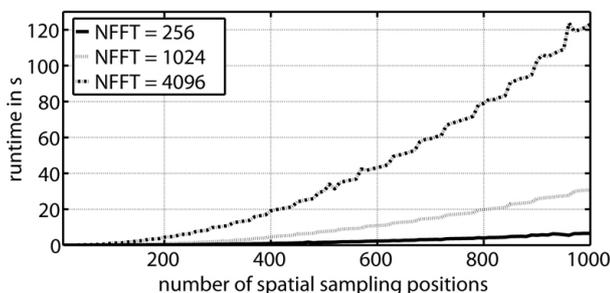

Figure 2: Runtime example of the SOFiA S/T/C module on $M=10-1000$ spatial sampling positions (microphones) for different FFT-Blocksizes (N_{FFT}). The maximum order of decomposition is $N = \sqrt{M/2} - 1$. System: Intel Core2Quad Q9550, 2.8 GHz, 8GB RAM. Microsoft Windows 7 64 Bit, Matlab 2010a (64Bit). All results are averaged by 10 runs.

2.4 Plane wave decomposition/beamforming

The included plane wave decomposition and beamforming core ($P/D/C$) can process a large amount of decomposition directions simultaneously, which leads to fast processing e.g. for visualization or auralization purposes. The core accepts radial filter coefficients either coming from the SOFiA M/F core (see section 2.5) or generated externally. Furthermore it offers an input for beamforming coefficients. These can either have a constant directivity or be frequency dependent.

2.5 Modal Radial Filters

The SOFiA M/F (Modal Filter) core generates modal radial filters for spherical arrays of different configurations:

- Open sphere with omnidirectional microphones,
- Open sphere with cardioid microphones, and
- Rigid sphere with omnidirectional microphones.

Furthermore the core includes amplification soft-limiting and free field powerloss compensation as proposed in [7]. Amplification limiting is vital for working with real measurement data from microphone arrays with a finite signal to noise ratio. The core has an output channel to observe the on-axis ($\frac{\omega}{c} r$)-response when using limiter functions.
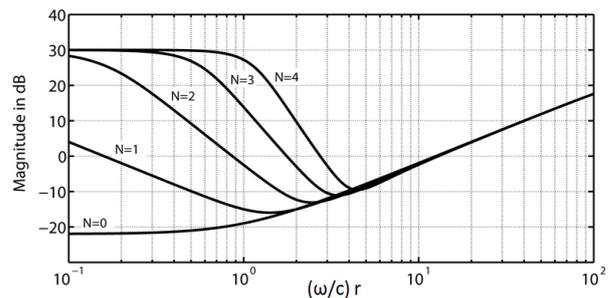

Figure 3: Exemplary radial filter magnitudes for orders $N = 0...4$ using amplification soft-limiting to $+30$ dB. The filters are designed for a rigid sphere configuration.

2.6 Sound field extrapolation

A very simple and basic module for the extrapolation of sound fields named $E/X/P$ covering both exterior and interior domain problems is included in the basic release. The module is implemented straightforwardly as defined in [1] but care has to be taken in practice because the extrapolation process involves possible poles, a problem
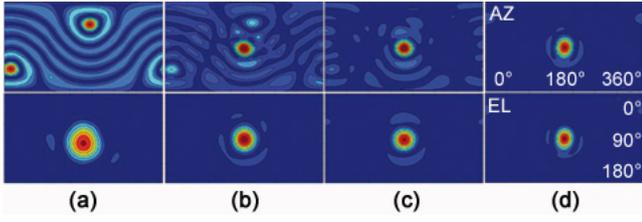
Figure 4: Response to a measured wavefront coming from $\phi_w = 180°$ and $\theta_w = 90°$ in the anechoic chamber at a decomposition order of $N = 7$. The upper row shows the response for different $(\frac{\omega}{c} r)$ using radial-filters without amplitude limiting. At low $(\frac{\omega}{c} r)$ the array response becomes unstable due to measurement noise and a high amplification ratio of higher modes. The bottom row shows the same results with SOFiA amplitude soft limiting to a maximum of $+5\,\mathrm{dB}$. The spatial response expands with falling $(\frac{\omega}{c} r)$ but the beam maintains stable. Values: (a) $(\frac{\omega}{c} r) = 1.5$, (b) $(\frac{\omega}{c} r) = 3$, (c) $(\frac{\omega}{c} r) = 3.5$ and (d) $(\frac{\omega}{c} r) = 5$.

similar to the radial filter inversion problem when dealing with open sphere configurations using pressure microphones. Nevertheless, the presented module can serve as a basis for extrapolation experiments or optimized implementation.

2.7 Data Visualization

SOFiA delivers tools for a response visualization. In a first step the *makeMTX* module generates a matrix of $[360 \times 180]$ nodes leading to a resolution of $1°$ for the full globe. The module internally uses the $P/D/C$ core, section 2.4. Once the matrix is generated, the data is shifted to the *visual3D* module that directly generates a plot in MATLAB. Different visualization styles are available, see Fig. 5.
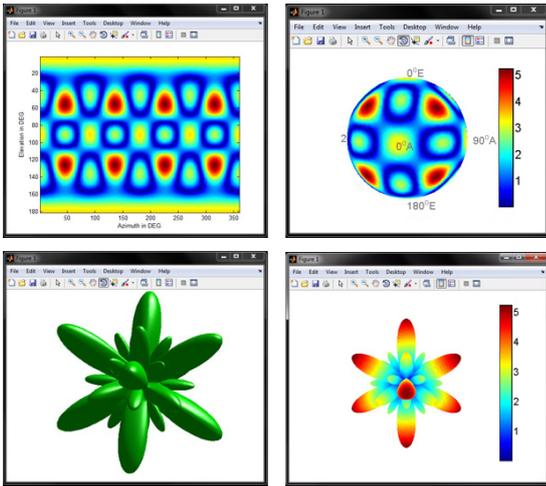

Figure 5: Different visualization styles using the *visual3D* module. All plots show the same dataset.

2.8 Impulse response rendering

The *makeIR* module generates impulse responses or time domain signals from data, either coming from the $P/D/C$ core or the $I/T/C$ core. It recombines the previously omitted negative part of the spectrum, applies a fast inverse Fourier transform and enables resampling

to the original sampling rate in case the signal was downsampled in *readVSAdata* or *mergeArrayData*. A windowing option is included.

3 Verification

Sound field analysis comprises of a considerable amount of processing and therefore a detailed verification is crucial. The verification of a sound field analysis chain can be performed in various stages of increasing complexity concerning the processing steps and the required experimental environment. A general verification approach is proposed, starting with an emulation of spatial Fourier coefficients using a generator module and ending up evaluating measurements taken in rooms. Apart from delivering a verification of the toolchain the following steps can also be useful to have a comprehensive and practical introduction to the basic principles of sound field analysis using spatial decomposition techniques as presented in [1], [2], [3]. Therefore all steps will be discussed in detail.

3.1 Stage I: Spatial Fourier Coefficients

At the first stage spatial Fourier coefficients are generated that only consider the angular solutions of the wave equation given in terms of Legendre polynomials and exponential functions. The latter can be merged into a single expression represented by the surface spherical harmonics [1]:

$$Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} \, P_n^m(\cos\theta) \, e^{im\phi}. \quad (1)$$

P_n^m denote Legendre functions of order n and mode m, θ and ϕ describe the elevation and azimuth angles. As the radial part has not been considered so far, the solution does not depend on the frequency, and the resulting coefficients do not directly correspond to those expected from real spatial waves. But in a first step it is useful to evaluate the angular functions separately. The Fourier coefficients $\mathring{G}_{mn}$ for a direction (θ_w, ϕ_w) directly correspond to the complex conjugate spherical harmonics:

$$\mathring{G}_{mn} = Y_n^{m*}(\theta_w, \phi_w). \quad (2)$$

The output signal S for a specific look direction (θ_l, ϕ_l) is generated using:

$$S(\theta_l, \phi_l) = \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^m(\theta_l, \phi_l) \, \mathring{G}_{mn}. \quad (3)$$

This output signal represents an ideal radial compensated plane wave array response of order N.

The spatial sharpness improves with an increased order N. In theory for $N \to \infty$ a spatial dirac impulse will arise. In measurement applications the maximum order is limited due to different factors such as a finite number
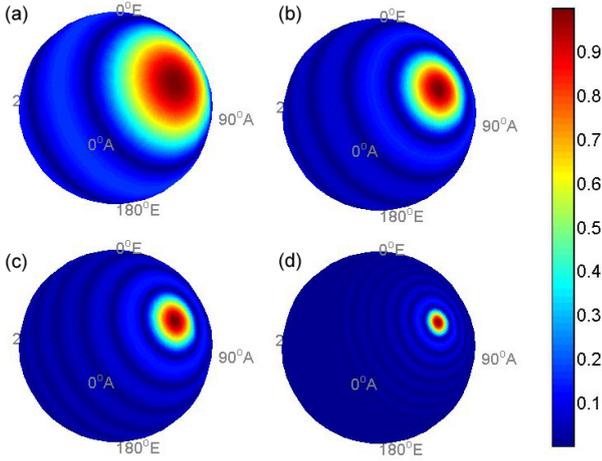
Figure 6: Ideal plane wave response to a wave from $\phi_w = 60°$ and $\theta_w = 60°$ at a single frequency for different orders: (a) $N = 3$, (b) $N = 6$, (c) $N = 9$ and (d) $N = 20$.

of discrete spatial sampling nodes, a limited signal to noise ratio or microphone positioning errors [3].

Involved SOFiA modules: *I/W/G*, *P/D/C*, *makeMTX* and *visual3D*. Optional: *makeIR*.

3.2 Stage II: Including the radial solution

On a second stage the radial part of the wave equation is included, involving frequency dependence. Now the radius, the sphere configuration and the microphone characteristics of the array have to be considered, and the resulting coefficients correspond to those expected for a real plane wave. A radial function $b_n(\frac{\omega}{c} r)$ is introduced, where ω denotes the angular frequency, c the speed of sound and r the array measurement radius. The design of the radial function depends on the sphere configuration and microphone types. An open measurement sphere with pressure transducers leads to [3]:

$$b_n\left(\tfrac{\omega}{c} r\right) = 4\pi \, i^n \, j_n\left(\tfrac{\omega}{c} r\right). \tag{4}$$

Using an open sphere with cardioid transducers yields [8]:

$$b_n\left(\tfrac{\omega}{c} r\right) = 4\pi \, i^n \, \tfrac{1}{2}\left(j_n\left(\tfrac{\omega}{c} r\right) - i \, j_n'\left(\tfrac{\omega}{c} r\right)\right). \tag{5}$$

For a rigid sphere with pressure transducers the wave that is scattered by the array surface is taken into consideration [2]:

$$b_n\left(\tfrac{\omega}{c} r\right) = 4\pi \, i^n \left(j_n\left(\tfrac{\omega}{c} r\right) - \frac{j_n'\left(\tfrac{\omega}{c} r\right)}{h_n^{(2)'}\left(\tfrac{\omega}{c} r\right)} h_n^{(2)}\left(\tfrac{\omega}{c} r\right)\right). \tag{6}$$

4π is introduced for normalization and $i = \sqrt{-1}$ denotes the imaginary unit. The radial functions contain different compositions of spherical Bessel functions $j_n(\frac{\omega}{c} r)$ and spherical Hankel functions $h_n(\frac{\omega}{c} r)$, providing solutions to the radial part of the wave equation in spherical coordinates [1]. Taking into account the radial functions, the analytic spatial Fourier coefficients for a plane wave

arriving from the direction (θ_w, ϕ_w) can be synthesized according to:

$$\mathring{P}_{mn} = Y_n^{m*}(\theta_w, \phi_w) \, b_n\left(\tfrac{\omega}{c} r\right). \tag{7}$$

When running the plane wave decomposition we have to compensate for the radial structure of the sound field to achieve the same response observed in the previous stage. This is where the radial filter $d_n(\frac{\omega}{c} r)$ is introduced. For simulation purposes we can simply use the radial filters defined by:

$$d_n\left(\tfrac{\omega}{c} r\right) = \frac{1}{b_n\left(\tfrac{\omega}{c} r\right)}. \tag{8}$$

But in measurements these filters do not work on the entire frequency range. First of all in measurements only a limited signal to noise ratio is available and the modal amplification demanded by these filters becomes very high at low $(\frac{\omega}{c} r)$. Thus the filters will amplify noise rather than signal at lower frequencies. This may lead to an unstable array response. Hence in practice the amplification of higher modes must be limited to a reasonable value in order to obtain a stable array response, e.g. as proposed in [7]. This limiting operation leads to loss of spatial resolution at lower frequencies. Furthermore, there is another problem when inverting the radial function given in Eq. (4), as the spherical Bessel function has roots leading to poles in the response. However, this can be solved e.g. by a dual-radius approach [3] or by choosing a different configuration. For the following plane wave decomposition we use the radial filters shown in Eq. (8):

$$S\left(\theta_l, \phi_l, \tfrac{\omega}{c} r\right) = \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^m(\theta_l, \phi_l) \\ \mathring{P}_{mn} \, c_n\left(\tfrac{\omega}{c} r\right) d_n\left(\tfrac{\omega}{c} r\right). \tag{9}$$

The generated output signal represents a perfect plane wave response of order N captured by an ideal microphone array with a number of microphones tending to infinity and an infinite signal to noise ratio. The coefficients $c_n(\frac{\omega}{c} r)$ are introduced in order to enable beamforming and can be configured to hold a constant directivity or to be frequency-dependent. In this paper $c_n = 1$ is assumed leading to a plane wave decomposition using a regular beam pattern. Therefore in the following the $c_n(\frac{\omega}{c} r)$ coefficients will be neglected in order to simplify the equations. The result is identical to stage I, see Fig. 6, but the scenario is much closer to a sound field sampled by an array. The sphere configuration, radius and observed frequency range are not of very great importance in this case, as the virtual radial filters can compensate for any required amplification in most of the cases as no signals containing noise are used here. The current stage is suitable to observe the influence of amplitude limited radial filters on the array response.

Involved SOFiA modules: *I/W/G*, *P/D/C*, *makeMTX* and *visual3D*. Optional: *makeIR*.

3.3 Stage III: Discrete sampling positions

The third stage introduces discrete spatial sampling positions as the number of microphones in a practical array is limited. For that purpose a sound field is generated using the analytical definition of a unity gain plane wave. The complex pressure P for a microphone at position $(\theta_j,\ \phi_j,\ r)$ in an open sphere array with pressure transducers and a specified angular frequency ω is determined by:

$$P(\theta_j,\phi_j,r_j,\tfrac{\omega}{c}) = e^{i\,\frac{\omega}{c}\,r_m(sin\theta_j cos\phi_j\,sin\theta_w cos\phi_w\,+}$$
$$\phantom{P(\theta_j,\phi_j,r_j,\tfrac{\omega}{c}) = e^{i}}\ _{sin\theta_j sin\phi_j\,sin\theta_w sin\phi_w\,+\,cos\theta_j\,cos\theta_w)}. \tag{10}$$

At this point the spatial Fourier transform [1] must be introduced in order to transform the pressure in the spatial domain to spatial Fourier coefficients $\mathring{P}_{mn}$ in the spherical wave domain:

$$\mathring{P}_{mn}\!\left(\tfrac{\omega}{c}\,r\right) = \int\!\int P\!\left(\theta,\phi,r,\tfrac{\omega}{c}\right) Y_n^{m^*}(\theta,\phi)\,sin\theta\,d\theta\,d\phi. \tag{11}$$

As discrete microphone positions are considered in this section the integration of the complete sphere is replaced by a summation of all available microphone positions (e.g. defined by a quadrature grid) on the sphere, compare [3]:

$$\mathring{P}_{mn}\!\left(\tfrac{\omega}{c}\,r\right) = \sum_{j=1}^{M} \beta_j\, P\!\left(\theta_j,\phi_j,r,\tfrac{\omega}{c}\right) Y_n^{m^*}(\theta_j,\phi_j). \tag{12}$$

The weighting factors β_j account for the sampling grid. The response signal for $(\theta_l,\ \phi_l)$ and the angular frequency ω can be achieved using:

$$S\!\left(\theta_l,\phi_l,\tfrac{\omega}{c}\,r\right) = \sum_{n=0}^{N}\sum_{m=-n}^{n} Y_n^m(\theta_l,\phi_l)\,\frac{\mathring{P}_{mn}}{4\pi\,i^n\,j_n(\tfrac{\omega}{c}\,r)}, \tag{13}$$

with the radial filter $4\pi\,i^n\,j_n(\tfrac{\omega}{c}\,r)$ corresponding to the open sphere design with pressure transducers. At this point spatial aliasing may be observed, caused by the discrete sampling within the space domain. For any realistic processing application the maximum order N should be limited to $N < \tfrac{\omega}{c}\,r$ to keep the aliasing contributions related to sampling low [9]. This implies effectively an upper frequency limit up to which meaningful results can be achieved. But for simulation purposes or studies on aliasing artifacts it is quite interesting to exceed this limit as illustrated in Fig. 7 and Fig. 8. Furthermore this analytically defined synthetic sound field and the discrete sampling positions placed within, facilitate different error simulations where e.g. the impact of a microphone position deviation or a variance of the air temperature on the array response can be analyzed.

Involved SOFiA modules: *S/W/G, S/T/C, P/D/C, makeMTX, visual3D* and *makeIR*.

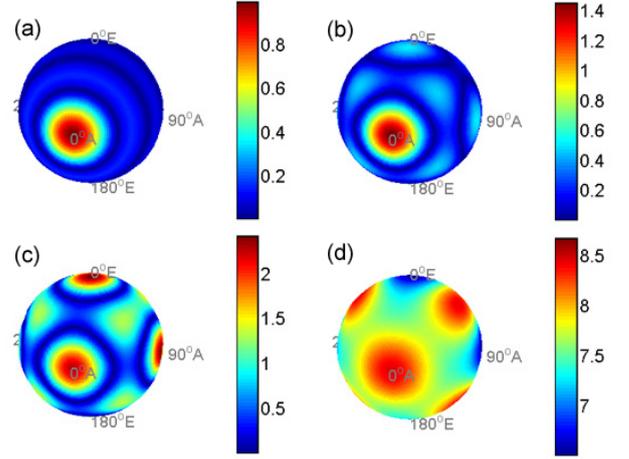**Figure 7:** Aliasing artifacts for (a) $(\tfrac{\omega}{c}\,r) < N$ with very low alias contribution, (b) $(\tfrac{\omega}{c}\,r) > N$ with some alias contribution, (c) and (d) examples of $(\tfrac{\omega}{c}\,r) \gg N$ with high alias contribution. The directional information is disarranged when exceeding the $(\tfrac{\omega}{c}\,r) \approx N$ limit and very high output levels arrise.
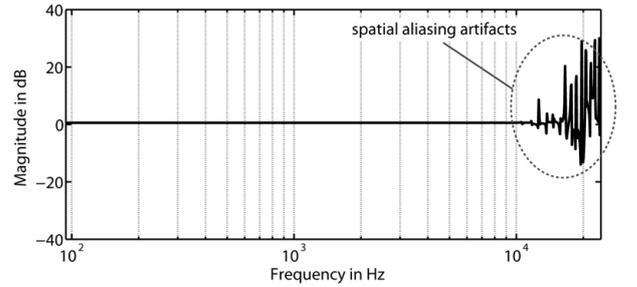

Figure 8: Exemplary simulation to illustrate the impact of spatial aliasing on the frequency response for an array looking into the direction of a full audio spectrum unity plane wave. For frequencies where $(\tfrac{\omega}{c}\,r) < N$ the response is flat as expected. At a region where $(\tfrac{\omega}{c}\,r) \gg N$ severe peaks and notches arise due to spatial aliasing.

3.4 Stage IV: The measured sound field

At the next stage we leave the simulation environment behind and step into real measurements. The aim of this stage is to establish an experimental setup that is widely controllable and to deliver results that are easily comparable to those observed in the previous simulative stages. As now real physical parameters and signals paired with discrete digital signal processing are coming up, a detailed view on the necessary steps is presented. For the experiment a spherical microphone array including preamplifiers and audio interfaces is required. The measurements take place in the anechoic chamber, in order to enable capturing a single wave, expecting only a comparatively low contribution of reflections. The microphone array is placed in the far field of a speaker system (source). Impulse responses $h_j(\theta_j,\phi_j,r,t)$ are captured for all array microphones $j = 1...M$ e.g. by using techniques introduced in [10] as depicted in Fig. 9. A conventional Fourier transform from time domain to frequency domain is applied to the impulse responses
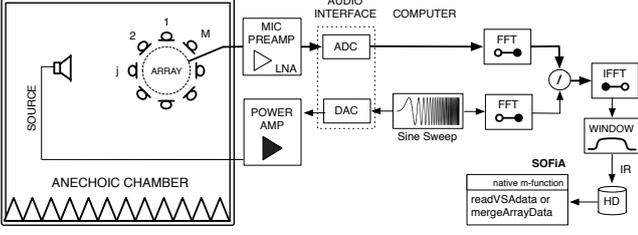
Figure 9: Capturing the impulse responses using a microphone array in the anechoic chamber in stage IV.

delivering complex-valued pressures:

$$P(\theta_j, \phi_j, r, f) = \int_{-\infty}^{\infty} h_j(\theta_j, \phi_j, r, t)\, e^{-i\, 2\pi f\, t}\, dt, \quad (14)$$

where f denotes the frequency and t the time. For digital processing a DFT (Discrete Fourier Transform) is used instead, combining discrete sampling in both the temporal and spectral domain and assuming a causal signal of finite length:

$$P(\theta_j, \phi_j, r, q) = \sum_{u=0}^{K-1} h_j(\theta_j, \phi_j, r, u)\, e^{-i\, 2\pi u\, q/K}, \quad (15)$$

where K is the total number of samples in the time domain, u describes the temporal sampling index and q a frequency bin. The range of q is $[0; K-1]$. For the subsequent processing steps q can be limited to the half-sided spectrum in the range $[0; \lceil K/2 \rceil]$ to save computational workload without suffering from any loss of information. The value of the angular frequency ω is related to the discrete frequency q as follows:

$$\omega = \frac{2\pi\, q\, f_s}{(K-1)}, \quad \omega \in [0\, ; 2\pi f_s], \quad (16)$$

depending on the sampling frequency f_s. We stay in the temporal discrete q notation leading to the following expression for the estimation of the spatial Fourier coefficients:

$$\mathring{P}_{mn}(r, q) \approx \sum_{j=1}^{M} \beta_j\, P(\theta_j, \phi_j, r, q)\, Y_n^{m^*}(\theta_j, \phi_j), \quad (17)$$

being similar to Eq. (12). A discrete formulation of Eq. (9) is:

$$S(\theta_l, \phi_l, r, q) = \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^m(\theta_l, \phi_l)\, \mathring{P}_{mn}(r, q)\, d_n'(qr), \quad (18)$$

and returns a Nth-order spatial decomposition of the incoming sound field refering to the direction (θ_l, ϕ_l). It is necessary to use radial filters $d_n'(qr)$ with modal amplification limiting here, as discussed in section 2.5. The resulting output signal $S(\theta_l, \phi_l, r, q)$ can be used to extract a superdirective plane wave decomposition of the measured sound field or to visualize the impact of plane

waves. The directional impulse response or time signal is calculated using the inverse Fourier transform to return from the frequency domain to the time domain:

$$h(\theta_l, \phi_l, r, t) = \int_{-\infty}^{\infty} S(\theta_l, \phi_l, r, f)\, e^{i\, 2\pi f\, t}\, df, \quad (19)$$

which is beeing replaced again by a discrete formulation using the inverse discrete Fourier transform (IDFT):

$$h(\theta_l, \phi_l, r, u) = \Re\left\{ \frac{1}{K} \sum_{q=0}^{K-1} S(\theta_l, \phi_l, r, q)\, e^{i\, 2\pi u\, q/K} \right\}. \quad (20)$$

When using the half-sided spectrum for $q \in [0; K']$, $K' = \lceil K/2 \rceil$, as discussed above, the dismissed part of the spectrum has to be reconstructed. For an even number of samples K the impulse response or time signal can be computed using:

$$h(\theta_l, \phi_l, r, u) = \Re\Bigg\{ \frac{1}{K} \Bigg(\sum_{q=0}^{K'} S(\theta_l, \phi_l, r, q)\, e^{i\, 2\pi u\, q/K} + \\ \sum_{q'=1}^{K'-1} S^*(\theta_l, \phi_l, r, K' - q')\, e^{i\, 2\pi u\, (K'+q')/K} \Bigg) \Bigg\}, \quad (21)$$

and for an odd number of samples K using:

$$h(\theta_l, \phi_l, r, u) = \Re\Bigg\{ \frac{1}{K} \Bigg(\sum_{q=0}^{K'} S(\theta_l, \phi_l, r, q)\, e^{i\, 2\pi u\, q/K} + \\ \sum_{q'=0}^{K'-1} S^*(\theta_l, \phi_l, r, K' - q')\, e^{i\, 2\pi u\, (K'+q'+1)/K} \Bigg) \Bigg\}. \quad (22)$$

Where only the real part $\Re\{\cdot\}$ of the expressions is taken into account. The theoretical result should not have any imaginary components but in practice a very tiny imaginary part can appear due to numerical issues. To visualize the response at a specific frequency Eq. (18) must be applied to a desired matrix of visualization angles. For instance, $[360 \times 180]$ nodes have to be calculated using Eq. (18) to obtain a step resolution of $1°$ on the full globe, as described in section 2.7. Furthermore the wavefronts generated by the speaker are assumed to be plane when arriving at the array. This simplification leads to an error that decreases when increasing the source distance as the wave tends to become more plane for larger distances. Generally neglecting errors and limitations that arise in a realistic scenario, we now should be able to observe an array response that is very close to the simulated plane wave responses of the previous stages. The response is illustrated in Fig. 10. Apart from errors, we encounter discrete spatial sampling positions on the one hand and a limited signal to noise
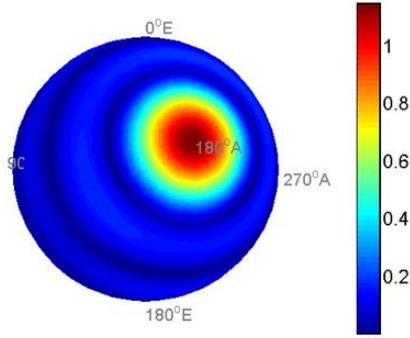
Figure 10: Array response to a plane wave coming from $\phi_w = 180°$ and $\theta_w = 90°$ at a frequency of around 1.6 kHz. The decomposition order is $N = 4$. The array runs on a 86 points Levedev grid in open sphere configuration at a diameter of $d = 50$cm. A large diaphragm cardioid microphone (Microtech Gefell M900) and an AD-Systems Flex15 PA speaker system at a distance of 4 m are used. The result looks very similar to the simulations.

ratio on the other hand. These two inevitable factors limit the possible frequency operation range of the array as has already been discussed in the previous sections and [3]. The optimum operation range for spherical array is around $(\frac{\omega}{c} r) \approx N$ [3].

Involved SOFiA modules: *readVSAdata, F/D/T, S/F/T, P/D/C, makeMTX* and *visual3D*. Optional: *makeIR*.

3.5 STAGE V - Spatiotemporal decomposition

For the next experiment we also take measurements from the anechoic chamber. Various sources are placed around the array as depicted in Fig. 11 and Fig. 12. The sources are all placed on the same radius referring to the center of the array and are calibrated to a common sound pressure level. A loadspeaker management processor is used to control levels and delay times for all sources independently. This way it is possible to generate multiple waves differing in arrival directions, sound pressure levels and arrival times in an accurately defined and widely controllable scenario.
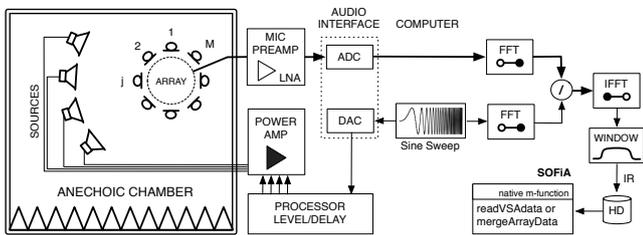

Figure 11: Setup for capturing multiple sources demonstrating the spatiotemporal resolution in stage V.

This section will reveal the full potential of the technique leading to a short-time spatiotemporal decomposition of a measured sound field into plane waves. The spatial decomposition has already been discussed and observed in the previous sections. To decompose the temporal structure it is necessary to introduce the short time
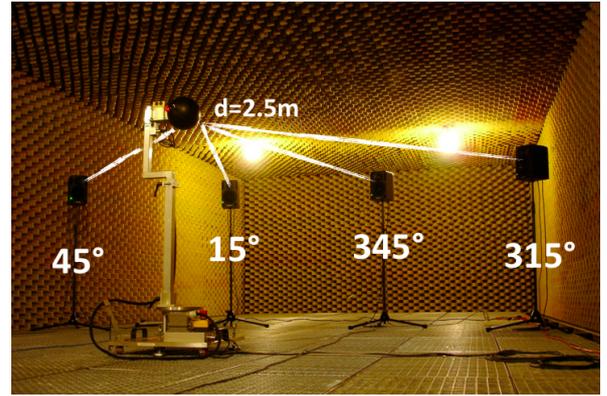

Figure 12: Experimental setup in the anechoic chamber as used for stage V: A VariSphear scanning array measurement system equipped with an Earthworks M30 microphone in a rigid sphere configuration and four Genelec 1029A speakers controlled by a XTA DP224 speaker management system are used.
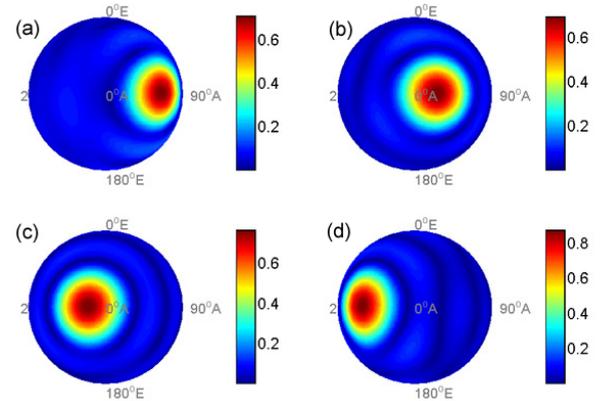

Figure 13: Result from the experiment in stage V. The sources are delayed as follows: $\phi_w = 45° : 0$ ms, $\phi_w = 15° : 16$ ms, $\phi_w = 345° : 32$ ms, $\phi_w = 315° : 48$ ms. The visual response for a decomposition of order $N = 5$ at a frequency of around 3.8 kHz is plotted for the corresponding timeshifts (a) $\tau = 0$ ms, (b) $\tau = 16$ ms, (c) $\tau = 32$ ms, (d) $\tau = 48$ ms. A rectangular window of 256 samples is applied. The array is configured to a 110 points Lebedev grid in a rigid sphere with a diameter of $d = 17.5$ cm. The sources can now be separated nicely in time and space.

Fourier transform which is quite similar to Eq. (14) but now including a window function $w(t - \tau)$:

$$P(\theta_j, \phi_j, r, f, \tau) = \int_{-\infty}^{\infty} h_j(\theta_j, \phi_j, r, t) \\ w(t - \tau) \, e^{-i\,2\pi f\,t} \, dt, \tag{23}$$

that enables to extract the spectra of different time slots shifted by τ and covered by the window function. Again considering the discreteness, a short-time DFT is used for the processing analog to Eq. (15):

$$P(\theta_j, \phi_j, r, q, T) = \sum_{u=0}^{K-1} h_j(\theta_j, \phi_j, r, u) \\ w(u - T) \, e^{-i\,2\pi u\,q/K}, \tag{24}$$

where T is the sample shift of the window function. There are several window functions available, from a simple rectangular window to more sophisticated ones like e.g. Hann, Hamming or Blackman windows, each having different impacts on the spectrum and bringing its own assets and drawbacks [11]. A convenient choice for the presented purpose is e.g. a Tukey window (tapered cosine). Let the window be symmetrical with an odd sample length L, then reaching from $T - (L-1)/2$ to $T + (L-1)/2$. Choosing a short window in time domain will give a good temporal resolution but decreases the spectral resolution and vice versa. This means that sharpness can be exchanged between the time and the frequency domain by varying L. The window can be moved in blocks or sample by sample, thereby leading to a sliding response. However for each sampleshift T we obtain a particular set of spatial Fourier coefficients $\mathring{P}_{mn}$:

$$\mathring{P}_{mn}(r, q, T) \approx \sum_{j=1}^{M} \beta_j \, P(\theta_j, \phi_j, r, q, T) \, Y_n^{m^*}(\theta_j, \phi_j),$$

$$(25)$$

where the pressure $P(\theta_j, \phi_j, r, q, T)$ is computed using Eq. (24). The number of available spectral coefficients q depends on the length of the applied window L. Further processing is done according to the last section starting with Eq. (18) et sqq. and e.g. repeating these operations for all different sets of temporal shifted $\mathring{P}_{mn}(r, q, T)$ coefficients depending on the required purpose.

Involved SOFiA modules: *readVSAdata, F/D/T, S/F/T, P/D/C, makeMTX* and *visual3D*. Optional: *makeIR*.

Before proceeding to stage VI further experiments with different configurations could be conducted inside the anechoic chamber, like e.g. placing a reflective surface. But these experiments do not reveal much more information about the verification process itself.

3.6 STAGE VI - Analysis of a room

The last stage takes place in a room. A setup close to the one illustrated in Fig. 9 is used to capture room impulse responses. But the source characteristics should be omnidirectional in this case, e.g. using a dodecahedron. The post-processing is identical to stage V, chapter 3.5. The spatio-temporal structure of the sound field in the room can now be analyzed by using response visualization, Fig. 14 (a). Further directional impulse responses can be extracted for different directions, Fig. 14 (b). By applying simple geometries, the results can be roughly verified for their temporal integrity and general plausibility. Listening to either the raw impulse responses or impulse responses convolved with dry stimuli quickly reveals possible problems and errors in the processing chain.

Involved SOFiA modules: *readVSAdata, F/D/T, S/F/T, P/D/C, makeMTX, visual3D* and *makeIR*.
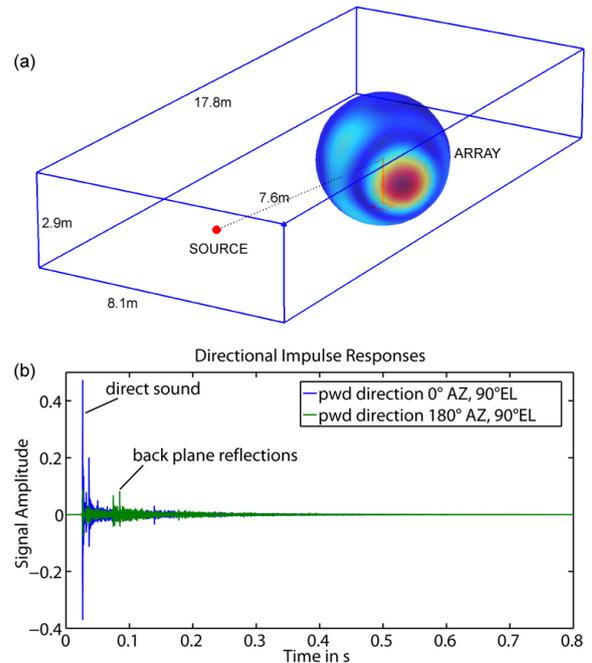

Figure 14: Exemplary results from a measurement session in a classroom for a decomposition of order of $N = 4$ sampled on a 590 points Lebedev grid in a rigid sphere configuration having a diameter of $d = 17.5\,\text{cm}$. The source is a high power wideband dodecahedron. (a) Visual response of a side wall reflection at the timeshift of $\tau \approx 33\,\text{ms}$ for a frequency of around $2\,\text{kHz}$. A rectangular window of 128 samples is applied leading to a frequency resolution of around $170\,\text{Hz}$ and a temporal resolution of approx. $3\,\text{ms}$ which corresponds to a sonic wave traveling like one meter in space. The CAD model of the room was captured using the laser unit of the VariSphear system [5]. (b) Directional impulse responses for different look directions. A high crosstalk rejection for different directions can be observed.

4 Conclusions and outlook

A sound field analysis toolbox for MATLAB has been presented. The toolbox includes the fundamental toolchain from the import of spherical microphone array data or the simulation of synthetic waves to the reconstruction of directional impulse responses or time domain signals and array response visualizations. The modular and largely abstract design allows for a flexible setup and to include external code. The SOFiA modules can be easily modified, rearranged and aligned to create a specific processing chain according to the respective problem. A general approach for a toolchain verification process has been proposed. The underlying theory has been discussed and examples were illustrated. Certainly the verification process could be expanded and arbitrary additional steps and experiments can be included. However the steps presented here are condensed to the essential problems and are characterized by a comprehensive proceeding covering a broad scale. It has been shown that the SOFiA toolbox delivers a reliable toolset for the fundamental sound field analysis processing. The presented verification process is also feasible to attain a survey on the methods of sound field analysis in a practical approach. Even if the current

version of SOFiA already covers a helpful set of tools there is unlimited potential for further development. Particularly problems concerning spatial aliasing or sound field extrapolation need further investigation from the current point of view. The development of further SOFiA modules is planned, e.g. covering visualization and auralization of sound fields. Contributions by the acoustic community are welcome.

5 Funding

References

[1] Earl G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, 1999.

[2] Jens Meyer and Gary W. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *IEEE ICASSP*, vol. 2, pp. 1781–1784, 2002.

[3] Boaz Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 135—-143, 2005.

[4] Benjamin Bernschütz, Christoph Pörschmann, Sascha Spors, and Stefan Weinzierl, "SOFiA sound field analysis toolbox," http://code.google.com/p/sofia-toolbox/, September 2011.

[5] Benjamin Bernschütz, Christoph Pörschmann, Sascha Spors, and Stefan Weinzierl, "Entwurf und Aufbau eines sphärischen Mikrofonarrays für Forschungsanwendungen in Raumakustik und Virtual Audio," *Proceedings of the 36. Deutsche Jahrestagung für Akustik (DAGA 2010)*, pp. 717–718, 2010.

[6] John Maddock, Paul Bristow, Hubert Holin, and Xiaogang Zhang, "Boost c++ libraries, math/special functions," http://www.boost.org/, September 2011.

[7] Benjamin Bernschütz, Christoph Pörschmann, Sascha Spors, and Stefan Weinzierl, "Soft-limiting der modalen Amplitudenverstärkung bei sphärischen Mikrofonarrays im Plane Wave Decomposition Verfahren," *Proceedings of the 37. Deutsche Jahrestagung für Akustik (DAGA 2011)*, pp. 661–662, 2011.

[8] Frank Melchior, Oliver Thiergart, Giovanni Del Galdo, Diemer de Vries, and Sandra Brix, "Dual radius spherical cardioid microphone arrays for binaural auralization," *Audio Engineering Society Convention 127*, 2009.

[9] Boaz Rafaely, "Spatial aliasing in spherical microphone arrays," *IEEE Transactions on Speech and Audio Processing*, vol. 55, no. 3, pp. 1003—-1010, 2007.

[10] Swen Müller and Paulo Massarani, "Transfer-function measurement with sweeps," *J. Audio Eng. Soc*, vol. 49, no. 6, pp. 443–471, 2001.

[11] F.J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, jan. 1978.