Idea for Sign-Change Retrieval in Magnitude Directivity Patterns

Franck Zagala[1], Franz Zotter[2]
[1] *University of Music and Performing Arts, Graz, Austria, Email: franckzagala@gmail.com*
[2] *Institute of Electronic Music and Acoustics, Univ. Music and Performing Arts, Graz, Austria, Email: zotter@iem.at*

Introduction

Our contribution discusses ideas to decompose magnitude radiation patterns after manipulation with optimal signs in order to simplify their spherical harmonic decomposition. Directivity patterns, e.g. of musical instruments, are often measured on a grid of spherically surrounding directions [1, 2]. Across the relevant frequency range, the magnitude can be assumed to most reliably represent the essential information about directivity. By contrast, the phase is error-prone and complicated, especially if at high frequencies, where there is poor acoustic centering, and even its simplification [3, 4, 5, 6, 7] can be ill-conditioned, so that disposal of the measured phase values can be considered to avoid interference and spatial aliasing effects. And yet, despite all-positive, zero-phase radiation patterns are surely robust, they aren't delivering optimal spherical harmonic interpolation in any case. For instance, accuracy and simplicity suffers for a first-order dipole pattern. While it might be acceptable that an all-positive dipole pattern is difficult to reconstruct close to its zero with a high-order spherical harmonics expansion, a low-order expansion is inaccurate more or less everywhere, cf. Fig. 1.

This contribution gives a simple overview of techniques based on the idea that introducing a sign change, as a first approach to solve this problem. Several other strategies are outlined in the thesis [8] of the first author.

Sign change idea

Assume we are given samples of a magnitude directivity pattern $p_\lambda = |p_\lambda|$, measured by microphones $\lambda = 1, \dots, \Lambda$ of a spherical array surrounding the source. Without changing the magnitude and for optimal interpolation, we may modify each sample by a sign $z_\lambda = \pm 1$

$$\mathbf{x} = [p_\lambda \, z_\lambda] = \mathrm{diag}\{\mathbf{p}\}\mathbf{z}. \tag{1}$$

To prepare interpolation, we stack the $(\mathrm{N}+1)^2$ real-valued spherical harmonics $Y_n^m(\boldsymbol{\theta}_\lambda)$ up to a given order N, sampled at the $\Lambda = (\mathrm{N}+1)^2$ microphone locations $\boldsymbol{\theta}_\lambda$ into a matrix

$$\mathbf{Y} = \begin{pmatrix} Y_0^0(\boldsymbol{\theta}_1) & Y_1^{-1}(\boldsymbol{\theta}_1) & Y_1^0(\boldsymbol{\theta}_1) & \cdots & Y_N^N(\boldsymbol{\theta}_1) \\ Y_0^0(\boldsymbol{\theta}_2) & Y_1^{-1}(\boldsymbol{\theta}_2) & Y_1^0(\boldsymbol{\theta}_2) & \cdots & Y_N^N(\boldsymbol{\theta}_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Y_0^0(\boldsymbol{\theta}_\Lambda) & Y_1^{-1}(\boldsymbol{\theta}_\Lambda) & Y_1^0(\boldsymbol{\theta}_\Lambda) & \cdots & Y_N^N(\boldsymbol{\theta}_\Lambda) \end{pmatrix},$$

to formulate the microphone signals as $\mathbf{x} = \mathbf{Y}\,\boldsymbol{\chi}$ and get the coefficients by inversion $\boldsymbol{\chi} = \mathbf{Y}^{-1}\mathbf{x}$, so the directivity can be evaluated at any continuous direction $\boldsymbol{\theta}$

$$x(\boldsymbol{\theta}) = \mathbf{y}(\boldsymbol{\theta})^{\mathrm{T}}\mathbf{Y}^{-1}\,\mathrm{diag}\{\mathbf{p}\}\mathbf{z}, \tag{2}$$

with the continuous-direction spherical harmonics vector

$$\mathbf{y}(\boldsymbol{\theta}) = \begin{pmatrix} Y_0^0(\boldsymbol{\theta}) & Y_1^{-1}(\boldsymbol{\theta}) & Y_1^0(\boldsymbol{\theta}) & \cdots & Y_N^N(\boldsymbol{\theta}) \end{pmatrix}^{\mathsf{T}},$$

Our discussion is based on the 64 maximum determinant[1] points `md007.00064`, see also [1], that accurately interpolates with the order N = 7.

Complexity/accuracy measure: How complicated the interpolated result evolves between the sampling points is now affected by the signs in $\mathbf{z}$. As cost function in terms of the unknown sign modifiers $\mathbf{z}$ we consider the reconstruction accuracy

$$J = \left\| \,\mathrm{diag}\{\mathbf{p}\}\mathbf{z} - \mathbf{Y}\,\mathrm{diag}\{\mathring{\mathbf{w}}\}\mathbf{Y}^{-1}\,\mathrm{diag}\{\mathbf{p}\}\mathbf{z} \right\|^2, \tag{3}$$

with further limited or tapered order using the weights $\mathring{\mathbf{w}}$ that penalize higher-order content (overshoots, ripples) by introducing reconstruction errors.

Zero-phase reconstruction

The simplest approach uses zero phase, $z_\lambda = +1$. However, this might require higher orders than necessary to accurately reconstruct the magnitude pattern at each sampling point, see Fig. 1 for a dipole and a window $\mathring{\mathbf{w}}$ keeping components up 2nd order, or 7th order.

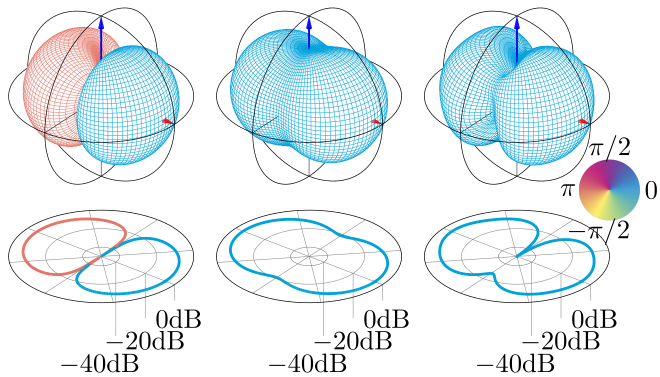

Figure 1: Original dipole directivity and its zero phase reconstruction with 2nd order and 7th order from the magnitude of 64 points

Naive minimization/full search

To test the cost function wrt. all sign combinations for the example array of $\Lambda = 64$ mics, one would need to evaluate the measure J for 2^Λ times. Optimistically assuming one test to take 1 ns, a full search would take $2^{64} \cdot 10^{-9}\,\mathrm{s} \approx 5.9$ centuries, is therefore unproductive. For 32 and 20 mics, full search would only require $2^{32} \cdot 10^{-9}\,\mathrm{s} \approx 4.2\,\mathrm{s}$, or $2^{20} \cdot 10^{-9}\,\mathrm{s} \approx 1\,\mathrm{ms}$, respectively. (A non-binary phase $e^{\mathrm{i}\frac{2\pi}{P}k}$ would even require to search P^Λ combinations.)

[1] `https://web.maths.unsw.edu.au/~rsw/Sphere/Extremal/`

Reduction to regions of common sign
Assuming that neighbouring microphones share the same polarity if they belong to the same lobe, one could try grouping the microphones into regions and test the polarity combinations only of these regions. Hence, the dimensionality of the problem is drastically reduced.

This contribution proposes an algorithm to decompose the microphones into regions according to their magnitude, which is described in Algo. 1.

The algorithm accepts a $F \times 3$ convex-hull triangulation matrix $\mathbf{T}$ as geometric input containing 3 microphone indices composing each of hull's F facets, and the vector $\mathbf{p}$ containing the magnitudes of each microphone as measured input.

Two sets $\mathcal{S}$ and $\mathcal{T}$ are initialized with the index of all microphones. $\mathcal{T}$ simply contains all microphones indices, while $\mathcal{S}$ contains the indices of the unassigned microphones. Iteratively, a new region $\mathcal{R}_{regIdx}$ is created by starting with an unassigned microphone of $\mathcal{S}$ with the highest magnitude. To this region, all new microphones of the region are iteratively searched whether they have new neighbours in $\mathcal{T}$ of decreasing magnitude that are not yet from the region, which are then added to the region. $\mathcal{I}$ is a helper set keeping track of microphones added last to the region, thus enabling the expansion of the region without repeated search through the neighbourhood of one microphone twice. After the peak region is completed (i.e. there are no more neighbouring microphone of smaller magnitude), all its microphones are removed from the unassigned set $\mathcal{S}$. As $\mathcal{S}$ empties, all microphones are assigned to at least one region, and the algorithm terminates. Some microphones may be part of different regions simultaneously (doublons), which could yield the definition of their polarity conflicting. The conflict can be circumvented by declaring those doublons to "single-microphone regions" of their own.

Input: $\mathbf{T}$, $\mathbf{p}$
Output: $\mathcal{R}$
$\mathcal{T} \leftarrow \{1, \ldots, length(\mathbf{p})\}$;
$\mathcal{S} \leftarrow \mathcal{T}$;
regIdx $= 0$;
`// create and expand regions`
while $\mathcal{S} \neq \emptyset$ **do**
  regIdx $++$;
  currPoint $\leftarrow \{s \in \mathcal{S} : \mathbf{p}_s = \max(\mathbf{p})\}$ (only one if many);
  $\mathcal{R}_{regIdx} \leftarrow \{\text{currPoint}\}$;
  $\mathcal{I} \leftarrow \{\text{currPoint}\}$;
  while $\mathcal{I} \neq \emptyset$ **do**
    $\mathcal{Q} \leftarrow \emptyset$;
    forall $i \in \mathcal{I}$ **do**
      $\mathcal{N} \leftarrow \{t \in \mathcal{T} : t$ "touches" i in $\mathbf{T}\}$;
      $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{n \in \mathcal{N} : \mathbf{p}_n \leq \mathbf{p}_i\}$;
    end
    $\mathcal{R}_{regIdx} \leftarrow \mathcal{R}_{regIdx} \cup \mathcal{Q}$;
    $\mathcal{I} \leftarrow \mathcal{Q}$;
  end
  $\mathcal{S} \leftarrow \mathcal{S} - \mathcal{R}_{regIdx}$
end
`// find doublons`
$\mathcal{D} \leftarrow \emptyset$;
forall $rIdx1 \in \{1, \ldots, |\mathcal{R}| - 1\}$ **do**
  forall $rIdx2 \in \{rIdx1 + 1, \ldots, |\mathcal{R}|\}$ **do**
    $\mathcal{D} \leftarrow \mathcal{D} \cup (\mathcal{R}_{rIdx1} \cap \mathcal{R}_{rIdx2})$;
    $\mathcal{R}_{rIdx1} \leftarrow \mathcal{R}_{rIdx1} - \mathcal{R}_{rIdx2}$
  end
end
$\mathcal{R}_{|\mathcal{R}|} \leftarrow \mathcal{R}_{|\mathcal{R}|} - \mathcal{D}$;
`// create own region for each doublon`
forall $d \in \mathcal{D}$ **do**
  regIdx $++$;
  $\mathcal{R}_{regIdx} \leftarrow \{d\}$;
end

Algorithm 1: Regions decomposition. $|\mathcal{R}|$ denotes the cardinality of $\mathcal{R}$.

Semidefinite relaxation methods
Similarly to the methods used in [9] for the design of radiation filters, one could minimize (3) with a convex optimization algorithms. The optimization problem writes

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{C}^{\Lambda}}{\text{minimize}} \quad & \mathbf{z}^{\mathsf{H}} \mathbf{C} \mathbf{z} \\ \text{subject to} \quad & \operatorname{diag}\{\mathbf{z}^*\} \mathbf{z} = \mathbf{1}, \end{aligned} \tag{4}$$

where

$$\mathbf{C} = \mathbf{B}^{\mathsf{H}} \mathbf{B}$$

$$\mathbf{B} = (\mathbf{I} - \mathbf{Y} \operatorname{diag}\{\mathring{\mathbf{w}}\} \mathbf{Y}^{\dagger}) \operatorname{diag}\{\mathbf{p}\}.$$

After rewriting $\mathbf{z}^{\mathsf{H}}\mathbf{C}\mathbf{z} = \operatorname{trace}\{\mathbf{z}^{\mathsf{H}}\mathbf{C}\mathbf{z}\} = \operatorname{trace}\{\mathbf{C}\mathbf{z}\mathbf{z}^{\mathsf{H}}\}$, we now search for an optimum matrix $\mathbf{Z}$ of rank 1. This leads to a new formulation of the optimization problem

$$\begin{aligned} \underset{\mathbf{Z} \in \mathcal{H}^{\succcurlyeq}}{\text{minimize}} \quad & \operatorname{trace}\{\mathbf{C}\mathbf{Z}\} \\ \text{subject to} \quad & \operatorname{diag}\{\mathbf{Z}\} = \mathbf{1}, \operatorname{rank}\{\mathbf{Z}\} = 1. \end{aligned} \tag{5}$$

By omitting the rank constraint, the degrees of freedom increase from Λ to Λ^2, an approach known as the Semi Definite Relaxation Problem (SDR) of (5) [9, 10, 11]

$$\begin{aligned} \underset{\mathbf{Z} \in \mathcal{H}^{\succcurlyeq}}{\text{minimize}} \quad & \operatorname{trace}\{\mathbf{C}\mathbf{Z}\} \\ \text{subject to} \quad & \operatorname{diag}\{\mathbf{Z}\} = \mathbf{1}. \end{aligned} \tag{6}$$

Practically, this problem can be solved by some freely available programs such as `CVX` [12, 13].

Due the rank constraint relaxation, the vector $\mathbf{z}$ containing the polarity information for each microphone must be extracted from the optimal matrix $\mathbf{Z}$, now. For our sign-retrieval problem, the simplest way is to take the signs of the eigenvector associated to the highest eigenvalue of $\mathbf{Z}$. Since $\mathbf{C}$ is symmetric and real-valued, the optimal matrix $\mathbf{Z}$, its eigenvalues and eigenvectors must be real-valued.

Simulation

In order to evaluate the quality of the reconstruction, a simulation study was conducted using the 64pt sampling scheme. 50 radiation patterns were generated as real-valued random SH coefficients (according to the standard normal distribution) up to the order $N = 3$. Five different strategies were compared:

- zero phase: no sign retrieval assuming zero phase
- SDR: SDR-based sign retrieval of all signs
- SDR rgn+dbl: SDR-based sign retrieval with common-sign regions according to Algo. 1
- SDR rgn-dbl: SDR-based sign retrieval with a variation of Algo. 1: lines of doublons are removed from retrieval equation and their implicit phase is taken
- try rgn-dbl: the permutation-based sign retrieval with non-doublons approach as in SDR-rgn-dbl.

In Fig. 2, the Mean Square Error of each of the simulations, their median value and interquartiles are depicted. The MSE is defined as $MSE = \frac{J}{\Lambda}$ with the cost function from Eq. (3), using a non-zero $\mathring{\mathbf{w}}$ for orders up to $N = 3$. The

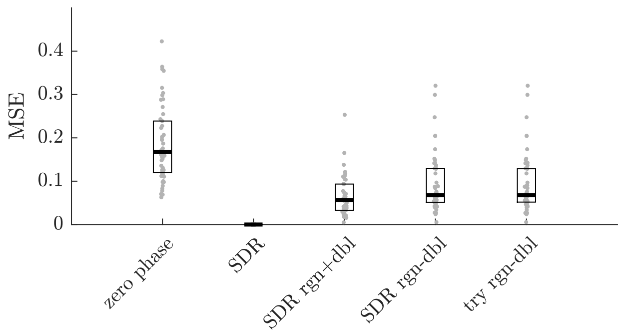

Figure 2: Median and interquartiles of MSE for reconstructed random directivity pattern from magnitude

statistics over the computation time for reconstruction time is depicted in Fig. 3 using MATLAB on a MacBook Pro with Intel Core i5 Processor (2.5 GHz) and 16GB RAM. On average, Algo. 1 detected 6 regions and 34 doublons.

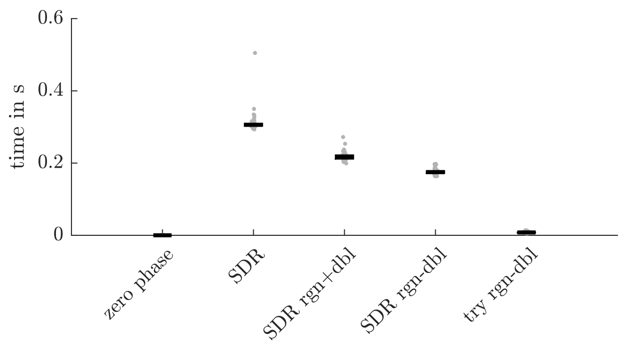

Figure 3: Median and interquartiles of reconstruction time of random directivity pattern from magnitude

Fig. 2 shows that the MSE of zero-phase approximation is outperformed by any other polarity-retrieving approach. Be it with SDR or by trying all polarity combinations, the use of region decomposition generally leads to a greater MSE than the SDR method without regions. In fact, the region decomposition algorithm is not conservative enough and may create larger regions than it should. The regions decomposition without inclusion of the doublons in "single microphone regions" appears to lead to smaller MSE than when ignoring them. Furthermore, both *SDR rgn-dbl* and *try rgn-dbl* lead to the same MSE, as they both reach the global minimum of the exact same cost function. Regarding the computation time, the use of region decomposition appears to be a significant improvement in the case of SDR. The exclusion of the doublons in the optimization process also reduce the computational time, as it drastically reduces the number of regions. Nonetheless, this enables a very rapid determination of the phase by trying all polarity combinations, which is much faster than when using the SDR method.

Conclusion

We discussed productive ways to find a suitable real-valued sign to an absolute-valued radiation pattern measurement. This is done because radiation patterns phase is not robust except at low frequencies. As a consequence, interpolation of complex-valued patterns in spherical harmonics may often turn out overly complicated or inaccurate, raising the need for optimum phase solutions.

We showed that taking only the positive sign can be sub-optimal, and we defined a cost function to measure simplicity and match of a radiation pattern due to its sign pattern.

For the sign-retrieving algorithm, we demonstrated that naive iteration through all thinkable sign combinations is unproductive for a large number of signs. However, if only a few signs are searched, the effort is low.

Consequently, we presented an algorithm decomposing the convex-hull grid of sound pressure samples into peak regions supposed to be of uniform sign, each, which effectively reduces the number of signs to iterate.

For more complex and sophisticated sign retrieval, we demonstrated semidefinite relaxation methods on the magnitude-least-squares problem, and we compared its accuracy and calculation time when operating on all signs individually, or aggregated signs of the region algorithm, compared to the all-positive sign solution.

While the semidefinite relaxation approach for the magnitude-least-squares problem outperforms the region-based and all-positive solutions in accuracy and simplicity of the directivity, it imposes still a substantial computational load. We could see that the region-based solution with semidefinite relaxation delivers less accurate/simple representation, but is more efficient.

For arbitrary, complex-valued directivities, the retrieval of polarity only is sub-optimal, and more continuous phase values are desirable. However, this renders both the naive search approach and the semidefinite relaxation approach more difficult. We hope that our contribution on the decomposition in peak regions is useful to optimally initialize local optimizers.

References

[1] F. Hohl, "Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten," Master's thesis, Institute for Electronic Music and Acoustics (IEM), November 2009.

[2] M. Pollow, G. Behler, and B. Masiro, "Measuring directivities of natural sound sources with a spherical microphone array," in *Ambisonics Symposium 2009*, June 2009.

[3] I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely, "Acoustic centering of sources measured by surrounding spherical microphone arrays," *The Journal of the Acoustical Society of America*, vol. 130, pp. 2003–2015, Oct 2011.

[4] D. Deboy and F. Zotter, "Acoustic center and orientation analysis of sound-radiation recorded with a surrounding spherical microphone array," in *Proceedings of the Int. Symp. on Ambisonics and Spherical Acoustics (Proceedings of the Int. Symp. on Ambisonics and Spherical Acoustics)* (M. Noisternig, ed.), (Paris (Frankreich)), 09 2010. procedure: peer-reviewed.

[5] D. Deboy and F. Zotter, "Tangential intensity algorithm for acoustic centering," in *Fortschritte der Akustik - DAGA 2011 (DAGA 2011)*, (Düsseldorf (Deutschland)), 03 2011. procedure: without peer reviewing.

[6] N. R. Shabtai and M. Vorländer, "Acoustic centering of sources with high-order radiation patterns," *The Journal of the Acoustical Society of America*, vol. 137, pp. 1947–1961, Apr 2015.

[7] J. Zaar, "Phase unwrapping for spherical interpolation of head-related transfer functions," Master's thesis, Institute for Electronic Music and Acoustics (IEM), December 2011.

[8] F. Zagala, "Optimum-phase primal signal and radiation-filter modelling of musical instruments," Master's thesis, Institute for Electronic Music and Acoustics (IEM), January 2019.

[9] P. W. Kassakian, *Convex Approximation and Optimization with Applications in Magnitude Filter Design and Radiation Pattern Synthesis.* PhD thesis, Electrical Engineering and Computer Sciences University of California at Berkeley, May 2006.

[10] Z.-q. Luo, W.-k. Ma, A. M.-c. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, pp. 20–34, May 2010.

[11] T. Bendory, R. Beinert, and Y. C. Eldar, "Fourier phase retrieval: Uniqueness and algorithms," 2017.

[12] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control* (V. Blondel, S. Boyd, and H. Kimura, eds.), Lecture Notes in Control and Information Sciences, pp. 95–110, Springer-Verlag Limited, 2008.

[13] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1." http://cvxr.com/cvx, March 2014.