Time-Domain Behaviour of Spherical Microphone Arrays at High Orders

Till Rettberg and Sascha Spors
Institut für Nachrichtentechnik, Universität Rostock, 18119 Rostock
Email: {till.rettberg, sascha.spors}@uni-rostock.de

Introduction

Spherical microphone arrays have been widely studied for three-dimensional sound field analysis. Expansion into spherical modes provides an elegant framework for beamforming and plane wave decomposition. In theory, constant directivity over the full frequency range is achieved at the expense of robustness for lower frequencies, which manifests as a low white noise gain (WNG) [1]. To prevent excessive noise amplification, in practice usually some form of gain limiting is applied to the radial filters. Equivalently, this can be considered as regularization of an ill-conditioned inverse problem. This contribution focuses on the impact of regularization in the time domain. It is argued that regularization necessarily smears the beamformer's impulse response, which should be taken into account in radial filter design.

Modal Beamforming

Only a brief problem statement is given here, a thorough presentation of spherical array processing can be found e.g. in [2]. The sound field captured by a spherical array of radius r is represented by the spherical wave spectrum $S_n^m(\omega)$, band-limited to a maximum order N. Then $S_n^m(\omega)$ can be expanded into plane waves incident from (ϕ, θ) by division with the radial function $B_n(\omega)$ and a subsequent inverse spherical Fourier transform

$$P(\phi, \theta, \omega) = \sum_{n=0}^{N} \underbrace{\frac{1}{B_n(\omega)}}_{=:D_n(\omega)} \sum_{m=-n}^{n} S_n^m(\omega) Y_n^m(\phi, \theta) \,, \qquad (1)$$

where the radial function $B_n(\omega)$ depends on the array design. For an open sphere of radius r with cardioid microphones $B_n(\omega) = 4\pi i^n (j_n(\frac{\omega}{c} r) - i j_n'(\frac{\omega}{c} r))$ may be chosen, with the spherical Bessel function $j_n(\cdot)$ and speed of sound c. The radial filters denoted by $D_n(\omega)$ in (1) correspond to the reciprocal of $B_n(\omega)$ and therefore grow arbitrarily large as $B_n(\omega)$ decays towards zero for $\omega \to 0$ and $n > 0$. Consequently, some kind of gain-limited radial filters $\hat{D}_n(\omega)$ have to be used in practice.

This can be stated from a numerical viewpoint: Assume the sound field can be modeled as a superposition of a finite number of plane waves. Their complex amplitudes are gathered in a vector $\mathbf{p}$ for each frequency. Matrix $\mathbf{M}$ models a measurement with a spherical sensor, which yields the spherical spectrum $\mathbf{s} = \mathbf{M}\mathbf{p}$. Plane wave decomposition as in (1) can then be regarded as the inverse problem $\mathbf{p} = \mathbf{M}^{-1}\mathbf{s}$, which is ill-conditioned for low frequencies. This corresponds to physical intuition that it is difficult to achieve directional resolution at large wave lengths with a finite aperture sensor. To obtain a meaningful result, regularization has to be applied; probably the most common choice would be the least squares solution using zero'th order Tikhonov regularization [3]:

$$\begin{aligned} \mathbf{p} &= \min \left\{ \|\mathbf{M}\mathbf{p} - \mathbf{s}\|_2^2 + \|\alpha \mathbf{I}\|_2^2 \right\} \\ &= \left(\mathbf{M}^H \mathbf{M} + \alpha^2 \mathbf{I} \right)^{-1} \mathbf{M}\mathbf{s} \,, \end{aligned} \qquad (2)$$

with the regularization parameter α. In beamforming literature (e.g. [4]) (2) is often given as the classical solution to a minimization problem with a constraint on total power, using Lagrange multipliers.

In order to assess different regularization strategies, the radial filters $D_n(\omega)$ in (1) are replaced by their gain-limited counterparts $\hat{D}_n(\omega)$, introducing a set of regularization filters $H_n(\omega)$:

$$\hat{D}_n(\omega) := H_n(\omega) \cdot D_n(\omega) \,, \qquad (3)$$

such that $|\hat{D}_n(\omega)| \le g_{\max}$ for a maximum radial filter gain $g_{\max}$. A similar formulation was used in [5], [6] in the context of Higher Order Ambisonics. $H_n(\omega)$ are order-dependent and real-valued in range $[0, 1]$. The realization of $H_n(\omega)$ depends on the employed regularization method, which is discussed next.

Different Limiting Filters

Literature (e.g. [1]) has suggested to impose frequency-dependent maximum order N by simply discarding contributions where the radial filter magnitude exceeds $g_{\max}$. This corresponds to regularization filters

$$H_n(\omega) = \begin{cases} 1 & |D_n(\omega)| \le g_{\max} \\ 0 & \text{else} \end{cases} . \qquad (4)$$

Note that this also is a common approach for solving bin-wise inverse problems such as (2) using singular value decomposition. Alternatively, instead of dropping excessive contributions they may be hard-clipped to $g_{\max}$:

$$H_n(\omega) = \begin{cases} 1 & |D_n(\omega)| \le g_{\max} \\ \frac{g_{\max}}{|D_n(\omega)|} & \text{else} \end{cases} \qquad (5)$$

In [7] it was argued that a hard threshold is detrimental for the spatial response and the following soft-knee characteristic was proposed:

$$H_n(\omega) = \frac{2}{\pi} \arctan \left(\gamma \frac{|D_n(\omega)|}{g_{\max}} \right) \qquad (6)$$

with the scaling factor $\gamma = \frac{\pi}{2}$ as used in [7]. Finally, solving the constrained optimization problem (2) leads to the regularization filters

$$H_n(\omega) = \frac{|B_n(\omega)|^2}{|B_n(\omega)|^2 + \alpha^2}\,, \tag{7}$$

where α is related to $g_{\max}$ by $\alpha = (1-\sqrt{1-1/g_{\max}^2}\,)/(1+\sqrt{1-1/g_{\max}^2}\,)$ [5].

Time Domain Implications

Plugging the gain-limited radial filters $\hat{D}_n(\omega)$ (3) and the spherical wave spectrum of a single incident plane wave $S_n^m(\omega) = Y_n^{m*}(\phi_0, \theta_0)$ into (1) yields [8]

$$\hat{P}(\phi, \theta, \omega) := \sum_{n=0}^{N} H_n(\omega) \frac{2n+1}{4\pi} L_n(\cos\Theta)\,, \tag{8}$$

where $L_n(\cos\Theta)$ denote the Legendre polynomials and Θ is the angle difference between look direction and incident plane wave. Eq.(8) constitutes the beam pattern of a regularized but otherwise ideal modal beamformer of finite order. Without modal gain limiting (i.e. $H_n(\omega) = 1$) the beampattern is frequency independent. Conversely, its time domain response is a perfectly localized impulse at $t = 0$, while the angular spread (beam width) depends on the maximum order N. All regularization filters $H_n(\omega)$ presented here are real-valued and therefore zero-phase. This implies impulse responses with even symmetry and thus non-causal contributions. These may be perceptually critical in auralization applications, e.g. data-based wave field synthesis.

Figure 1 depicts simulated impulse responses of a modal beamformer in azimuth under different regularization schemes. Assuming a maximum order $N = 23$, an open spherical sensor of radius 0.5m with cardioid capsules and a maximum filter gain of $g_{\max} = 10^3 \mathrel{\hat{=}} 40$dB. To focus on the regularization effects no equalization has been applied. The subfigures correspond to different regularization filters $H_n(\omega)$: *top left*: excessive orders are discarded (4), *top right*: hard limited to $g_{\max}$ (5), *bottom left*: soft limited as by (6). The regularization filters can also be interpreted as order-dependent spectral windows, causing side lobes in time domain. In this sense, simply discarding the excessive orders shares the undesirable properties of a rectangular window. This corresponds to a well known result from Fourier series that relates the number of continuous derivatives of a function to the decay of its Fourier coefficients [9]. A smooth regularization filter as in (6) is obviously beneficial. Binwise Tikhonov regularization (7), *bottom right*, achieves similarly good localization in time. Interestingly, larger side lobes appear at the rear direction.

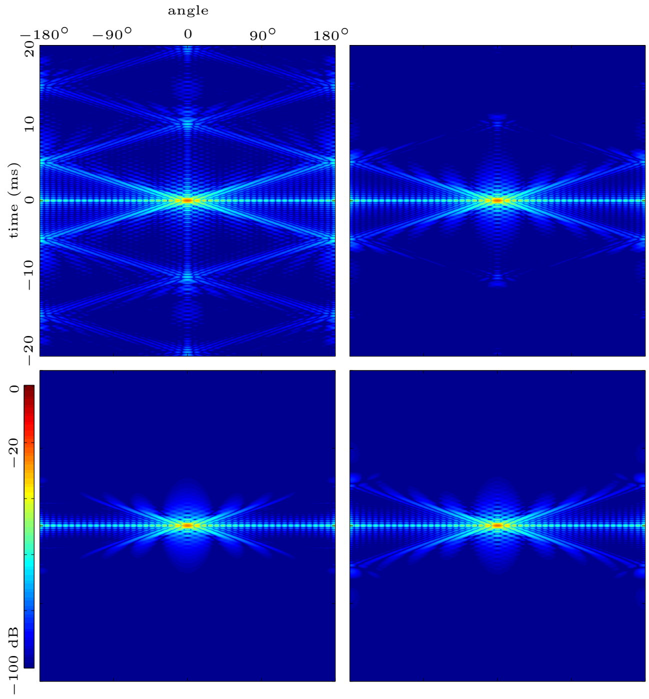

Figure 1: Impulse responses of modal beamformers using different regularization filters $H_n(\omega)$. Top left: equation (4), top right: (5), bottom left: (6), bottom right: (7). Axes and coloring are the same for all subplots.

Conclusion

This paper briefly discussed the impact of regularization on the impulse response of a modal beamformer, which is necessary in practice to prevent an excessive WNG. For auralization applications, the perceptual influence has to be evalutated by listening tests. Regularization schemes with nonlinear phase may be employed to reduce noncausal contributions. If linear phase is desired, results from time window design are applicable. Time domain implications of practical equalization schemes have to be investigated. From a theoretical point of view, links to mathematical regularization theory should be discussed further.

References

[1] Boaz Rafaely. Analysis and design of spherical microphone arrays. *IEEE Transactions on Speech and Audio Processing*, 13(1):135–143, January 2005.

[2] Boaz Rafaely. Plane-wave decomposition of the sound field on a sphere by spherical convolution. *The Journal of the Acoustical Society of America*, 116(4):2149–2157, 2004.

[3] Georges A. Deschamps and Hrair S. Cabayan. Antenna synthesis and solution of inverse problems by regularization methods. *IEEE Transactions on Antennas and Propagation*, 20(3):268–274, May 1972.

[4] Otis L. Frost III. An algorithm for linearly constrained adaptive array processing. *Proceedings of the IEEE*, 60(8):926–935, 1972.

[5] Sébastien Moreau, Jérôme Daniel, and Stéphanie Bertet. 3D sound field recording with higher order ambisonics-objective measurements and validation of spherical microphone. In *Proc. of 120th Audio Engineering Society Convention*, Paris, France, May 2006.

[6] Robert Baumgartner, Hannes Pomberger, and Matthias Frank. Practical implementation of radial filters for ambisonic recordings. In *International Confernece on Spatial Audio*, Detmold, Germany, November 2011.

[7] Benjamin Bernschütz, Christoph Pörschmann, Sascha Spors, and Stefan Weinzierl. Soft-Limiting der modalen Amplitudenverstärkung bei sphärischen Mikrofonarrays im Plane Wave Decomposition Verfahren. In *Proc. of German Annual Conference on Acoustics (DAGA)*, Düsseldorf, Germany, March 2011.

[8] Boaz Rafaely. Phase-mode versus delay-and-sum spherical microphone array processing. *Signal Processing Letters, IEEE*, 12(10):713–716, October 2005.

[9] Wladimir I. Smirnow. *Lehrgang der höheren Mathematik, Bd II, VI.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1955.