An analytic method for transforming spatial filtering of WFS driving functions into temporal filtering

Gergely Firtha[1], Péter Fiala[1]

[1] *Department of Networked Systems and Services, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary, Email: firtha/fiala@hit.bme.hu*

Introduction

The aim of sound field synthesis is to reproduce a virtual target sound field over an extended listening area using a densely spaced loudspeaker arrangement, known as the secondary source distribution (SSD). By feeding the loudspeakers with specific driving functions, the superposition of sound fields from each SSD element should ideally match the target sound field in the intended receiving area. One prominent sound field synthesis method is Wave Field Synthesis (WFS) [1, 2].

The spatial filtering of the time-space dependent driving functions is a frequently emerging question in practical sound field synthesis applications. As an important example: basic WFS theory assumes a continuous secondary source distribution for reproduction, while practical WFS applications use a discrete loudspeaker array, hence spatially sampling the loudspeaker driving functions. This discretization results in aliasing wavefronts emerging from the individual loudspeaker elements and following the intended virtual wavefront. To avoid these aliasing waves spatial bandlimitation of the driving signals is required leading to a spatial filtering problem. The direct spatial bandlimitation of the WFS driving functions is, however, not straightforward in real-time applications, being a computationally complex problem requiring more specialized hardware architectures. Analytical approaches for the problem are usually termed as Local Wave Field Synthesis (LWFS) techniques [3, 4].

For simple virtual sound fields, WFS driving functions exhibit a simple space-time interconnection resulting from the characteristics of wave propagation. This interconnection allows one to derive unique temporal filters for each SSD element, so that the resulting temporal filtering is equivalent to analytical spatial filtering. This paper discusses this spatial-to-temporal filter transformation, focusing on antialiasing as a direct application of the theory.

Theoretical basics

Local propagation vector: Consider an arbitrary steady-state sound field at an angular frequency ω, described by the general polar form:

$$P(\mathbf{x},\omega) = A^P(\mathbf{x})\mathrm{e}^{-\mathrm{j}\frac{\omega}{c}\phi^P(\mathbf{x})}, \tag{1}$$

where $A^P(\mathbf{x},\omega)$ and $\phi^P(\mathbf{x})$ are real-valued functions, and c is the speed of sound. This formulation applies to both plane waves and (3D) point sources. The propagation dynamics of the sound field are governed by its phase function $\phi^P(\mathbf{x})$, which can be used to define the local propagation vector $\hat{\mathbf{k}}^P(\mathbf{x})$:

$$\hat{\mathbf{k}}^P(\mathbf{x}) = [\hat{k}^P_x(\mathbf{x}),\ \hat{k}^P_y(\mathbf{x}),\ \hat{k}^P_z(\mathbf{x})]^{\mathrm{T}} = \nabla_{\mathbf{x}}\phi^P(\mathbf{x},\omega). \tag{2}$$

The local propagation vector is a unit vector perpendicular to the wavefront, pointing towards the local propagation direction [5].

In the temporal domain, the sound field can be obtained by taking the inverse Fourier transform of (1), leading to the expression:

$$p(\mathbf{x},t) = A^P(\mathbf{x})\,\delta\left(t - \frac{1}{c}\phi^P(\mathbf{x})\right). \tag{3}$$

2.5D Wave Field Synthesis: Consider a smooth convex SSD located at $\mathbf{x}_0 = [x_0,\ y_0,\ 0]^{\mathrm{T}}$ consisting of a continuous distribution of 3D point sources, described by the 3D Green's function. In this geometry the synthesized field at a receiver position $\mathbf{x} = [x,\ y,\ 0]^{\mathrm{T}}$ inside the area bounded by the SSD is described by the Kirchhoff approximation of 2.5D Kirchhoff-Helmholtz integral, from which the 2.5D driving functions can be extracted as [5]

$$D(\mathbf{x}_0,\omega) = \underbrace{\sqrt{8\pi\mathrm{j}k}}_{H_{\mathrm{pre}}(\omega)}\underbrace{w(\mathbf{x}_0)\sqrt{d_{\mathrm{ref}}(\mathbf{x}_0)}\hat{k}^P_{\mathrm{n}}(\mathbf{x}_0)}_{A(\mathbf{x}_0)}P(\mathbf{x}_0,\omega), \tag{4}$$

where $k^P_{\mathrm{n}}(\mathbf{x}_0)$ is the normal component of the local wavenumber vector. The driving function consists of a frequency dependent prefilter $H_{\mathrm{pre}}(\omega)$, a secondary source selection window $w(\mathbf{x}_0)$, a gain factor allowing amplitude correction along a reference curve (c.f. [5]) and the virtual field measured on the SSD.

Assuming a simple virtual sound field as given by (1) the driving functions can be written in the time domain as

$$d(\mathbf{x}_0,t) = h_{\mathrm{pre}}(t) *_t A^D(\mathbf{x}_0)\delta\left(t - \frac{1}{c}\phi^P(\mathbf{x}_0)\right), \tag{5}$$

with $A^D(\mathbf{x}_0) = A(\mathbf{x}_0)\cdot A^P(\mathbf{x}_0)$ being the real valued gain factor of the driving function, $*_t$ denoting temporal convolution and $h_{\mathrm{pre}}(t)$ is the temporal WFS prefilter impulse response [6]. In the following due to the associativity of convolution this prefiltering is excluded from the discussion.

Spatial to temporal filter transformation

This section presents an analytical transformation to express equivalent temporal filters for arbitrary spatial filters based on formulation (5). The idea behind the transformation is the following: For a given SSD position, the spatially filtered driving function represents the spatial average of driving functions in the neighboring SSD elements, weighted by the spatial impulse response. The driving functions are Dirac impulses in each array position, differing only in time shift and gain. Therefore, their weighted sum is a rescaled image of the spatial filter's impulse response, with its amplitude modulated by the driving function gains. Our aim is to express this weighted sum at each SSD position as the result of temporal moving-averaging of the corresponding Dirac impulse.

Assume an arbitrary filter impulse response defined in the spatial domain, denoted by $h_x(s)$. By choosing a suitable parametrization of the driving functions $d(s,t)$ (e.g. by polar angle in case of a circular SSD, or linear position on a linear SSD) the spatially filtered driving function is written as

$$d'_x(s,t) = h_x(s) *_x d(s,t) = \int h_x(s-s_0)\, d(s_0,t)\, \mathrm{d}s_0, \quad (6)$$

where $*_x$ denotes a circular convolution for a convex SSD contour or a linear convolution in case of an infinite long linear SSD. Our goal is to find a temporal filter impulse response for each SSD element $h_t(s,t)$, so that the temporal convolution of the driving function at each SSD element with the corresponding temporal filter results in the spatially filtered driving function. Mathematically this requirement is written as

$$\begin{aligned} d'_x(s,t) = h_x(s) *_x d(s,t) = h_t(s,t) *_t d(s,t) = \\ = \int_{-\infty}^{\infty} h_t(s,t-t_0)\, d(s,t_0)\, \mathrm{d}t_0 = d'_t. \end{aligned} \quad (7)$$

The temporal convolution in (7) can be evaluated by substituting equation (5) and exploiting the sifting property of the Dirac-delta

$$d'_t(s,t) = A^D(s)\, h_t\left(s, t - \frac{1}{c}\phi^P(s)\right). \quad (8)$$

With applying a time shift of $\frac{1}{c}\phi^P(s)$ to both sides the temporal filter impulse response is connected with the spatial convolution as

$$\begin{aligned} h_t(s,t) = \frac{1}{A^D(s)} \int h_x(s-s_0)\, d\left(s_0, t + \frac{1}{c}\phi^P(s)\right) \mathrm{d}s_0 = \\ \frac{1}{A^D(s)} \int h_x(s-s_0)\, A^D(s_0) \delta\left(t + \frac{1}{c}\left(\phi^P(s) - \phi^P(s_0)\right)\right) \mathrm{d}s_0. \end{aligned} \quad (9)$$

To evaluate the spatial convolution the generalized sifting property of the Dirac-delta may be applied [7], which states that

$$\int f(s_0)\delta(g(s_0))\mathrm{d}s_0 = \sum_i \frac{f(s_i)}{\left|\frac{\partial}{\partial s_0} g(s_0)\right|_{s=s_i}}, \quad g(s_i) = 0. \quad (10)$$

In the present problem the zeros of the Dirac's argument are found where

$$\frac{\phi^P(s_i)}{c} = t + \frac{\phi^P(s)}{c} \quad (11)$$

is satisfied. In the following it is assumed that a single zero s_i exists, satisfying (11). This assumption means that the wavefront arrives at each SSD element at a unique time instant, being strictly true for a virtual plane wave and a linear SSD. With this assumption and by utilizing that

$$\frac{\partial}{\partial s_0} g(s_0) = -\frac{1}{c}\frac{\partial}{\partial s_0}\phi^P(s_0) = -\frac{\hat{k}_t^P(s_0)}{c}, \quad (12)$$

i.e. the derivative of the phase function is the tangential component of the local propagation vector, the integral can be evaluated. The resulting general transformation relation between the spatial and temporal filters are given as

$$h_t(s,t) = \frac{A^D(s_i)}{A^D(s)} \frac{c}{|\hat{k}_t^P(s_i)|} h_x(s - s_i). \quad (13)$$

The above formulation already allows one to transform an arbitrary spatial impulse response into an equivalent temporal filter for each SSD element, as long as the virtual field model is known and (11) can be solved. In order to give a more general solution it is assumed that the virtual field is locally plane, being a usual high frequency assumption in WFS theory. As a further general WFS assumption it is supposed that the SSD is locally plane. These requirements inherently ensure that a single solution exists for (11)[1]. With these assumptions the phase function is given as

$$\phi^P(\mathbf{x}) = \hat{k}_x^P(\mathbf{x})x + \hat{k}_y^P(\mathbf{x})y \quad (14)$$

and (11) is satisfied where

$$s_i = s + \frac{c \cdot t}{\hat{k}_t^P(s)} \quad (15)$$

Finally, as a crucial approximation it is assumed that both the amplitude of the driving function and the propagation vector varies slowly along the SSD, i.e $A^D(s_i) \approx A^D(s)$ and $\hat{k}_t^P(s_i) \approx \hat{k}_t^P(s)$ holds. With these assumptions the corresponding filter transform reads as

$$h_t(s,t) = \frac{c}{|\hat{k}_t^P(s)|} h_x\left(-t\, \frac{c}{\hat{k}_t^P(s)}\right). \quad (16)$$

Equation (16) can be directly expressed in the spectral domain by taking the Fourier transform of both sides, relating the frequency response of the spatial and the temporal filters. By denoting $\mathcal{F}(h_x(s)) = H_x(k_s)$ the corresponding transform is given by

$$H_t(s,\omega) = H_x\left(-\frac{\omega}{c}\hat{k}_t^P(s)\right) = H_x\left(-k_t^P(s)\right), \quad (17)$$

where $k_t^P(s)$ is the tangential component of the local wavenumber vector, being a vector in a steady state sound field, pointing in the local propagation direction with the length being ω/c [7]. Hence, as the main result of the present discussion, the temporal filter transfer can be obtained from the wavenumber content of the spatial filter by simple rescaling in terms of the local wavenumber vector.

[1] Except for a plane wave arriving normally to the SSD, at which case the temporal filter is transformed into a Dirac-delta.

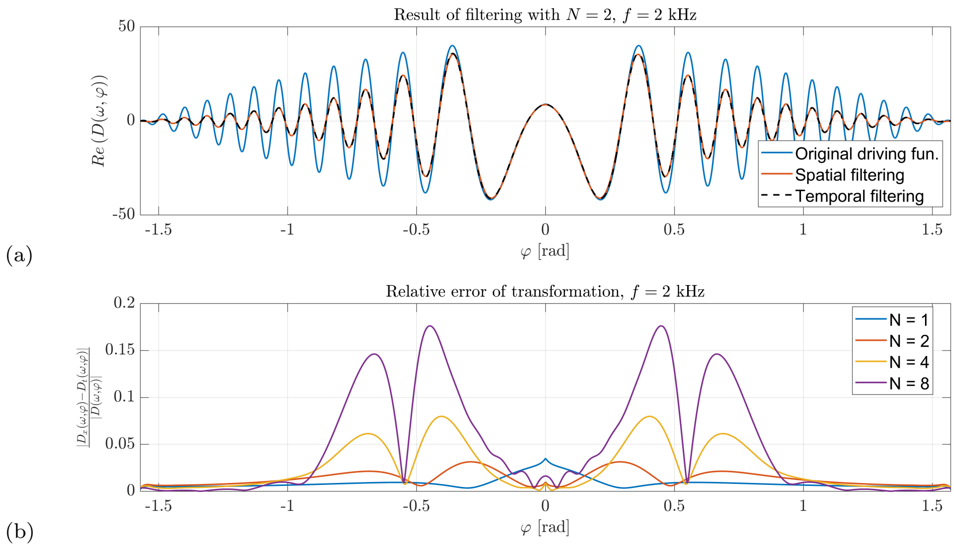

Figure 1: Comparison of spatially filtered quasi-continuous WFS driving functions and the result of equivalent temporal filtering (a) and the introduced error as the function of spatial filter impulse response length (b)

Validity of the approximations:

The presented results are valid under certain assumptions, requiring that the amplitude and propagation direction of the virtual field must change slightly in the proximity of each SSD element. This approximation is strictly valid for linear SSDs with virtual plane waves. For other geometries error is introduced, with the magnitude varying as the function of the spatial impulse response length.

The next section investigates the validity of the presented results by examining a typical WFS application where circular SSD synthesis is used to recreate a virtual plane wave. The amplitude of synthesis is referenced to the center of the array. The driving functions for the synthesis scenario are given e.g. by (4.31) in [7].

In the present example the circular driving functions (being the function of the polar angle φ) are filtered with a spatial low-pass filter. The transfer of the low pass filter is defined in the wavenumber domain, chosen to be an N-th order Butterworth design, given as

$$H_x(k_s) = \frac{1}{\sqrt{1+(k_s/k_c)^{2N}}}. \tag{18}$$

The investigation was carried out on a single angular frequency ω and the cut-off wavenumber was arbitrarily chosen to be $k_c = \frac{k}{2} = \frac{\omega}{2c}$. With applying the transform given by (17) the corresponding equivalent filter array is obtained by substituting $k_s = -\frac{\omega}{c}\hat{k}_t^P(s)$, yielding

$$H_t(\omega, \mathbf{x}_0) = \frac{1}{\sqrt{1+\left(2\hat{k}_t^P(\mathbf{x}_0)\right)^{2N}}}. \tag{19}$$

Figure 1 (a) depicts the result of temporal filtering of the driving functions in comparison with the result of direct spatial filtering in case of a second order Butterworth filter. It is verified that at low filter orders the equivalent temporal filter bank approximates fairly the result of spatial filtering. Figure 1 (b) illustrates the relative error of temporal filtering. It is shown that with increasing filter order—i.e. with increasing spatial impulse response length—the error of approximation also increases, since local approximations discussed in the previous section do not longer hold.

Application for antialiasing filtering:

The practical application of the theoretical results is discussed in relation to the spatial antialiasing filtering of WFS driving functions. The physical presence of aliasing in the synthesis is due to non-zero loudspeaker spacing, which violates theoretical requirements. This results in the presence of aliasing wavefronts in the synthesized field, which follow the intended virtual wavefront and can cause perceivable coloration in the listening position. The mathematical model of aliasing involves the spatial sampling of the driving function distribution, sampled at the actual loudspeaker positions, resulting in the overlapping of the spatial driving function spectra. Figure 2 (a) illustrates the result of aliasing in the case of a circular WFS geometry.

In this circular geometry, the SSD is discretized with a sampling arc length of $\mathrm{d}s = \frac{R}{N}$ (with R and N being the SSD radius and the number of loudspeakers respectively), leading to a sampling wavenumber of $k_s = \frac{2\pi N}{R}$. To avoid aliasing, spatially bandlimited driving functions must be evaluated, bandlimited to the Nyquist wavenumber, being half of the sampling wavenumber $k_{\text{Nyquist}} = \frac{\pi N}{R}$ [7]. However, achieving spatial bandlimitation before the evaluation of the driving functions is computationally extensive, as direct implementation would require spatial oversampling, antialiasing filtering, and downsampling. The presented equivalent temporal filtering approach provides a simple solution for antialiasing filtering, as it inherently provides analytically spatial-bandlimited loudspeaker signals.

Figure 2 (b) depicts the result of antialiasing filtering. The equivalent temporal antialiasing filter is chosen to

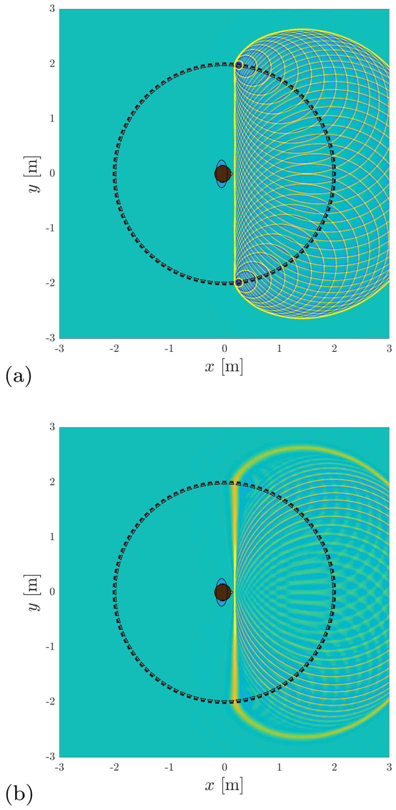

Figure 2: Application of the filter transform strategy to antialiasing filtering.

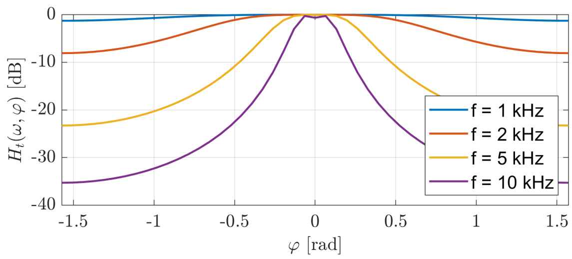

Figure 3: Illustration of equivalent frequency dependent window functions

be a 2nd order Butterworth design, with the cut-off frequency given by

$$\omega_c(\mathbf{x}_0) = k_{\text{Nyquist}} \frac{c}{\hat{k}_t^P(\mathbf{x}_0)} = \frac{\pi N}{R} \frac{c}{\hat{k}_t^P(\mathbf{x}_0)}. \qquad (20)$$

It is verified that aliasing wavefront are highly attenuated behind the virtual wavefront at a the receiver position. A simple strategy for changing the position of antialiased synthesis is further discussed in [5].

It should be noted that the presented spatial-to-temporal filter transform strategy can be also interpreted as a frequency dependent spatial windowing approach: in the aspect of antialiasing filtering SSD positions where the local wanumber vector is higher than the Nyquist wanumber are highly attenuated, with the window width decreasing with increasing angular frequency. The actual shape of the spatial windows are given by the wavenumber spectrum of the spatial lowpass filter $H_x(k_s)$, rescaled in terms of the local wavenumber vector. The set of window functions applied in the present simulation scenario are illustrated in Figure 3.

Conclusion

The present paper discussed an analytical approach for performing spatial filtering of wave field synthesis driving signals as equivalent temporal filtering. The transformation relies on rescaling the spatial filter transfer spectrum in terms of the local wavenumber of the target field, measured along the SSD. The proposed method offers the advantage of easy implementation in existing hardware architectures. The effectiveness of the method is demonstrated by applying it to spatial antialiasing of WFS driving functions. The examples provided confirm that the proposed transformation yields feasible results, as long as relatively short spatial filters are applied.

Acknowledgement

This work was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Supported by the ÚNKP-22-5-BME-318 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

This work was supported by the OTKA PD-143129 grant.

This work was supported by the OTKA K-143436 grant.

References

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

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

[3] Fiete Winter, Frank Schultz, Gergely Firtha, and Sascha Spors. A geometric model for prediction of spatial aliasing in 2.5d sound field synthesis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(6):1031–1046, 2019.

[4] Nara Hahn, Frank Schultz, and Sascha Spors. Cylindrical radial filter design with application to local wave field synthesis. *J. Audio Eng. Soc*, 70(6):510–525, 2022.

[5] Gergely Firtha, Péter Fiala, Frank Schultz, and Sascha Spors. Improved referencing schemes for 2.5d wave field synthesis driving functions. *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, 25(5):1117–1127, May 2017.

[6] Frank Schultz. *Sound Field Synthesis for Line Source Array Applications in Large-Scale Sound Reinforcement*. PhD thesis, University of Rostock, 2016.

[7] G. Firtha. *A Generalized Wave Field Synthesis Framework with Application for Moving Virtual Sources*. PhD thesis, Budapest University of Technologies and Economics, 2019.