A Geometric Model for Spatial Aliasing in Wave Field Synthesis*

Fiete Winter[1], Jens Ahrens[2], Sascha Spors[1]

[1] *Institute of Communications Engineering, University of Rostock, R.-Wagner-Str. 31, 18119 Rostock, Germany, Email: {fiete.winter; sascha.spors}@uni-rostock.de*

[2] *Division of Applied Acoustics, Chalmers University of Technology, Sven Hultins gata 8a, 41296 Göteborg, Sweden, Email: jens.ahrens@chalmers.se*

1 Introduction

Wave Field Synthesis (WFS) [1] aims at a physically accurate synthesis of a desired sound field inside a target region. Typically, the region is surrounded by a finite number of discrete loudspeakers. For practical loudspeaker setups, this spatial sampling causes spatial aliasing artefacts and does not allow for an accurate synthesis over the entire audible frequency range. In the past, different theoretical treatises of the spatial sampling process for simple loudspeaker geometries, e.g. lines and circles, led to anti-aliasing criteria independent of the listener's position inside a target region. However, no inference about the spatial phenotype of the aliasing artefacts can be made by these models. The present work introduces a geometrical model based on high-frequency approximations of the underlying theory describing the spatial occurrence and the propagation direction of the additional wave fronts caused by spatial aliasing. Based on this model, anti-aliasing criteria for individual loudspeakers and listening positions are derived.

2 Wave Field Synthesis

The fundamental task in WFS is to synthesise a desired (aka. virtual) sound field

$$S(x, y, \omega) = A_S(x, y, \omega)\, \mathrm{e}^{+\mathrm{j}\phi_S(x,y,\omega)} \tag{1}$$

within the defined region Ω (cf. Fig. 1). In $2^1/_2$-dimensional (2.5D) scenarios [2, Sec. 2.3], reproduction is restricted to the horizontal plane, i.e. $z = 0$. The latter variable is omitted, for brevity. The real-valued amplitude and the phase of the virtual sound field are denoted by $A_S(x, y, \omega)$ and $\phi_S(x, y, \omega)$, respectively. The radian frequency $\omega = 2\pi f$ is defined by the temporal frequency f. A distribution of loudspeakers is positioned along the boundary of Ω as so-called secondary sources (loudspeaker symbols). In the traditional theory of WFS [3, Sec. 3.1], a linear Secondary Source Distribution (SSD) along the x-axis is assumed: the secondary sources are located at $\mathbf{x}_0 = [x_0, 0, 0]^{\mathrm{T}}$. Correct synthesis is targeted in the positive y half plane. The sound field synthesised by the SSD is given by the one-dimensional convolution integral

$$P(x, y, \omega) = \int_{-\infty}^{\infty} D(x_0, \omega) G(x - x_0, y, \omega)\, \mathrm{d}x_0\,, y > 0 \tag{2}$$

*This research was supported by grant SP 1295/9-1 of the German Research Foundation (DFG)

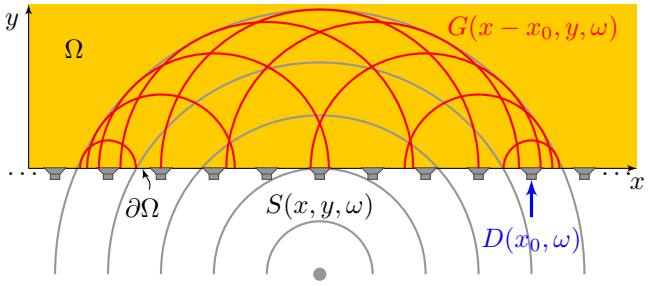

Figure 1: Geometry for Wave Field Synthesis

The sound field emitted by an individual secondary source is commonly modelled as a monopole point source. It is given by the three-dimensional free-field Green's function [4, Eq.(8.41)]

$$G(x - x_0, y, \omega) = \underbrace{\frac{1}{4\pi\sqrt{(x - x_0)^2 + y^2}}}_{A_G(x-x_0,y,\omega)} \underbrace{\mathrm{e}^{-\mathrm{j}\frac{\omega}{\mathrm{c}}\sqrt{(x-x_0)^2+y^2}}}_{\mathrm{e}^{+\mathrm{j}\phi_G(x-x_0,y,\omega)}} \tag{3}$$

with its according amplitude and phase terms. The speed of sound is denoted by c. Each individual secondary source is driven by its according WFS driving signal [5, Eqs. (18) and (21)]

$$D(x_0, \omega) = \tag{4}$$
$$\underbrace{-\sqrt{\frac{\mathrm{j}8\pi\, d(x_0)}{\frac{\omega}{\mathrm{c}}}}\, \phi'_{S,y}(x_0, 0, \omega)\, \overbrace{A_S(x_0, 0, \omega)\mathrm{e}^{+\mathrm{j}\phi_S(x_0,0,\omega)}}^{S(x_0,0,\omega)}}_{A_D(x_0,\omega)}\,.$$

The first derivative of ϕ_S w.r.t. y is denoted by $\phi'_{S,y}$. For details about the 2.5D correction factor $d(x_0)$, see [5]. Note that the phase term of the WFS driving function coincides with the phase of the desired sound field at $\mathbf{x}_0$.

The practical implementation of WFS implies a discretisation of the SSD as the distance between adjacent loudspeakers cannot be chosen arbitrarily small. For a uniform sampling the reproduced sound field $P(x, y, \omega)$ is approximated by

$$P^{\mathrm{S}}(x, y, \omega) = \sum_{n=-\infty}^{\infty} D(n\Delta_x, \omega) G(x - n\Delta_x, y, \omega)\Delta_x\,. \tag{5}$$

with the sampling distance denoted as Δ_x.

3 Modelling of Spatial Aliasing

In the following, a ray approximation for wave field synthesis using a continuous and a discrete SSD will be derived.

3.1 The Local Wavenumber Vector

The concept of the local wavenumber vector was introduced to the context of Sound Field Synthesis (SFS) by Firtha et al. [5, Eq. (15)]. For an arbitrary sound field $P(x, y, \omega)$, it is defined as

$$\mathbf{k}_P(x, y, \omega) := -\nabla \phi_P(x, y, \omega) \overset{\omega \to \infty}{\approx} \frac{\omega}{\mathrm{c}} \hat{\mathbf{k}}_P(x, y, \omega) \quad (6)$$

It describes the local propagation direction of $P(x, y, \omega)$ at a given coordinate $\mathbf{x}$. For simple sound fields such as point/line sources or plane waves, $\mathbf{k}_P(x, y, \omega)$ fulfils the local dispersion relation, i.e. its length is fixed to $\frac{\omega}{\mathrm{c}}$. For arbitrary sound fields, this statement is true for asymptotically high frequencies, see [6, Sec. 5.14]. The normalised vector is denoted as $\hat{\mathbf{k}}_P(x, y, \omega)$. Within the described 2.5D synthesis scenario, it is assumed that the involved sound fields do only propagate in horizontal direction. Hence, the z-component of the involved local wavenumber vectors is zero for $z = 0$. With its length and its z-component fixed to the given values, the local wavenumber vector is determined by one of its remaining components despite an unknown sign of the other component.

3.2 Ray-Approximation of Wave Field Synthesis

Using amplitude-phase notations of the involved quantities, the synthesis integral in Eq. (2) is reformulated to

$$P(x, y, \omega) = \int_{-\infty}^{\infty} A_D(x_0, \omega)\, A_G(x - x_0, y, \omega) \times \mathrm{e}^{+\mathrm{j}(\phi_G(x - x_0, 0, \omega) + \phi_S(x_0, 0, \omega))}\, \mathrm{d}x_0 \,. \quad (7)$$

It is approximated by its so-called Stationary Phase Approximation (SPA) [7, Sec. 4.6]

$$P_{\mathrm{SPA}}(x, y, \omega) = D(x_0^*, \omega) G(x - x_0^*, y, \omega) \times \sqrt{\frac{-\mathrm{j}2\pi}{|\phi''_{S,x}(x_0^*, 0, \omega) + \phi''_{G,x}(x_0^* - x, y, \omega)|}} \,. \quad (8)$$

In general, the approximation assumes a complex-valued function to have a phase term which is rapidly oscillating compared to its slowly changing amplitude. This is fulfilled in the context of WFS for asymptotically high frequencies. The integration over this function yields zero except for the contributions from the vicinity of the stationary point x_0^*. Here, the first-order derivative of the phase term w.r.t. x_0 vanishes and the second-order derivative is non-zero. More figuratively, the reproduced sound field at $\mathbf{x} = [x, y, 0]^{\mathrm{T}}$ is completely determined by a single secondary source located at $\mathbf{x}_0^* = [x_0^*, 0, 0]$. The stationary phase point x_0^* has to fulfil the condition

$$0 \overset{!}{=} \phi'_{S,x}(x_0^*, 0, \omega) + \phi'_{G,x}(x_0^* - x, y, \omega) \quad (9)$$

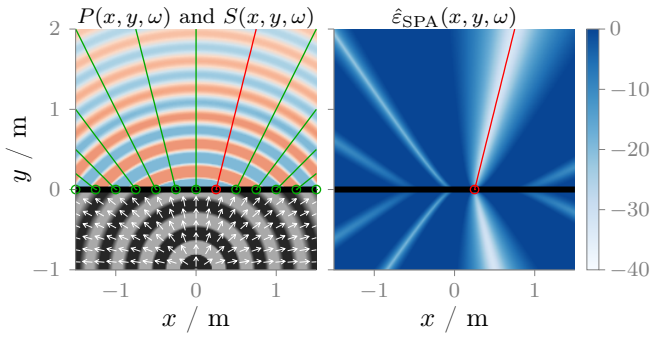

Figure 2: For $y < 0$, the left plot shows the real part of a monochromatic ($f = 1000$ Hz) point source located at $\mathbf{x}_{\mathrm{ps}} = [0, -1, 0]^{\mathrm{T}}$ m as the desired sound field $S(x, y, \omega)$. The white arrows indicate the corresponding normalised local wavenumber vector $\hat{\mathbf{k}}_S(x, y, \omega)$. For $y > 0$, the real part of the sound field $P(x, y, \omega)$ synthesised by a continuous linear SSD along the x-axis (black line) is plotted. Each coloured line indicates the positions $\mathbf{x}$ for which the secondary source at the start of the respective line (circles) is the stationary secondary source. For the secondary source at the red circle, $\hat{\varepsilon}_{\mathrm{SPA}}(x, y, \omega)$, see Eq. (13), is shown in the right plot.

with $\phi'_{\cdot,x}$ being the first-order derivative of the phase w.r.t. to its first argument. Taking the Eq. (6) into account, the derivatives can be replaced by the x-components $\hat{k}_{S,x}$ and $\hat{k}_{G,x}$ of the respective (normalised) local wavenumber vectors. Together with the properties described in (3.1), this allows for the equivalent condition

$$\hat{\mathbf{k}}_S(x_0^*, 0, \omega) \overset{!}{=} \hat{\mathbf{k}}_G(x - x_0^*, y, \omega) \quad \forall y > 0 \,. \quad (10)$$

According to Eqs. (3) and (6), the normalised local wavenumber vector of the three-dimensional free-field Green's function is given by

$$\hat{\mathbf{k}}_G(x - x_0^*, y, \omega) = \frac{1}{\sqrt{(x - x_0^*)^2 + y^2}} \begin{bmatrix} x - x_0^* \\ y \\ 0 \end{bmatrix} . \quad (11)$$

Eqs. (10) and (11) are solved for $\mathbf{x}$ yielding

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} x_0^* \\ 0 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} \hat{k}_{S,x}(x_0^*, 0, \omega) \\ \hat{k}_{S,y}(x_0^*, 0, \omega) \\ 0 \end{bmatrix} , 0 \le \gamma \le \infty \quad (12)$$

which is the parametric definition of a ray starting at $\mathbf{x}_0^*$ with the direction $\hat{\mathbf{k}}_S(x_0^*, 0, \omega)$. The reproduced sound field along the given ray is fully determined by the secondary source located at $\mathbf{x}_0^*$. As shown in the left plot of Fig. 2, the rays (coloured lines) are perpendicular to the wave front of the synthesised sound fields. The right plot shows the normalised error

$$\hat{\varepsilon}_{\mathrm{SPA}}(x, y, \omega) = 20 \log_{10} \left| \frac{P_{\mathrm{SPA}}(x, y, \omega) - P(x, y, \omega)}{P(x, y, \omega)} \right| \quad (13)$$

of the reproduced sound field and its SPA for a distinct $\mathbf{x}_0^*$ (red circle). Along the corresponding ray (red line), the error is significantly reduced. The approximation can be regarded as valid.

3.3 Ray Approximation of Aliasing Components

An elegant and commonly used model to describe the sampling process is the multiplication of the continuous quantity by a dirac impulse comb [8, Sec. 11.3.1]. The sampled driving function reads

$$D^{\mathrm{S}}(x_0,\omega) = D(x_0,\omega)\,\Delta_x \sum_{n=-\infty}^{\infty} \delta(x_0 - n\Delta_x) \tag{14}$$

with $\delta(x_0 - n\Delta_x)$ being a dirac impulse located at $n\Delta_x$. Note, that $D^{\mathrm{S}}(x_0,\omega)$ is still a continuous function, however only non-zero at integer multiples of Δ_x. With the Fourier series of the impulse comb [8, Eq. (11.12)], Eq. (14) is rearranged to

$$D^{\mathrm{S}}(x_0,\omega) = \sum_{\eta=-\infty}^{\infty} \underbrace{D(x_0,\omega)\mathrm{e}^{-\mathrm{j}2\pi\eta\frac{x_0}{\Delta_x}}}_{D^{\mathrm{S}}_\eta(x_0,\omega)} . \tag{15}$$

The η-th aliasing component of the discrete driving function is denoted by $D^{\mathrm{S}}_\eta(x_0,\omega)$, whereas the 0-th component is the original continuous driving function. The η-th component of the sound field $P^{\mathrm{S}}(x,y,\omega)$ synthesised by the discrete SSD is given by

$$P^{\mathrm{S}}_\eta(x,y,\omega) = \int_{-\infty}^{\infty} D^{\mathrm{S}}_\eta(x_0,\omega) G(x - x_0, y, \omega)\,\mathrm{d}x_0 . \tag{16}$$

Summing over all components, i.e. over all η, will result in $P^{\mathrm{S}}(x,y,\omega)$ given by Eq. (5), again. The individual aliasing components are approximated via their SPA

$$P^{\mathrm{S}}_{\eta,\mathrm{SPA}}(x,y,\omega) = D^{\mathrm{S}}_\eta(x^*_0,\omega) G(x - x^*_0, y, \omega) \times \sqrt{\frac{-\mathrm{j}2\pi}{|\phi''_{S,x}(x^*_0,0,\omega) + \phi''_{G,x}(x^*_0 - x, y, \omega)|}} . \tag{17}$$

Compared to Eq. (9), the condition for the stationary phase point x^*_0 is extended by an additional phase term caused by the complex exponential of $D^{\mathrm{S}}_\eta(x_0,\omega)$, see Eq. (15). The equivalent condition for the x-components of the normalised local wavenumber vectors is given by

$$\hat{k}_{S,x}(x^*_0,0,\omega) + \frac{\eta\mathrm{c}}{\Delta_x f} \overset{!}{=} \hat{k}_{G,x}(x - x^*_0, y, \omega) \tag{18}$$

The same approach as of Sec. 3.2 is taken to solve the equation for $\mathbf{x}$. The corresponding ray equation for the aliasing components reads

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} x^*_0 \\ 0 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} \hat{k}_{S,x}(x^*_0,0,\omega) + \frac{\eta\mathrm{c}}{\Delta_x f} \\ \sqrt{1 - \left(\hat{k}_{S,x}(x^*_0,0,\omega) + \frac{\eta\mathrm{c}}{\Delta_x f}\right)^2} \\ 0 \end{bmatrix} \tag{19}$$

It is evident from the y-component of the ray's direction vector, that the condition

$$\left|\hat{k}_{S,x}(x_0,0,\omega) + \frac{\eta\mathrm{c}}{\Delta_x f}\right| \leq 1 \tag{20}$$

has to be fulfilled in order to have a real-valued solution. Otherwise the η-th aliasing component is not excited by the secondary source located at x_0. Fig. 3 shows an example for the aliasing components and their corresponding ray approximation. As in Fig. 1, the rays are locally perpendicular to the wave fronts of the sound fields.

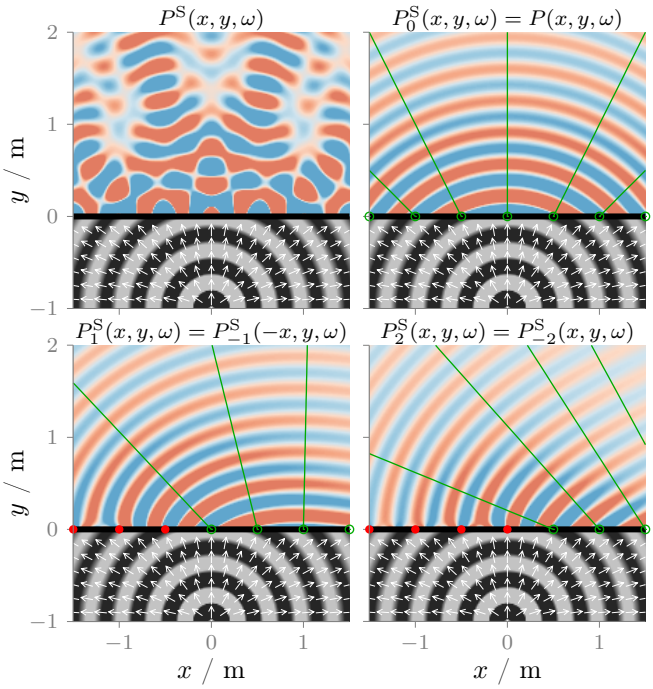

Figure 3: For $y < 0$, the same quantities as in the left plot of Fig. 2 are shown in all four plots. For $y > 0$, the sound field $P^{\mathrm{S}}(x,y,\omega)$ synthesised by a discrete ($\Delta_x = 0.5$ m) linear SSD along the x-axis is plotted in the upper left. The remaining plots show the aliasing components $P^{\mathrm{S}}_\eta(x,y,\omega)$ for $\eta \in [-2,2]$. Due to the symmetry, the plots for negative η can be produced by inverting the x-axis. Each green line indicates the positions $\mathbf{x}$ for which the secondary source its start (green circles) is the stationary secondary source. For the secondary sources marked by the red solid circles, the condition of (20) is not fulfilled.

4 The Spatial Aliasing Frequency

The highest temporal frequency f up to which no spatial aliasing artefacts occur is generally termed spatial aliasing frequency. Exceeding this frequency can be regarded as an violation of the anti-aliasing criterion.
Solving Eq. (20) for f yields the frequency $f^{\mathrm{S}}_\eta(x_0)$ above which the secondary source at x_0 excites $P^{\mathrm{S}}_\eta(x,y,\omega)$. No spatial aliasing is contributed, if f does not exceed this threshold for any η. Hence, the aliasing frequency for an individual secondary source reads

$$f^{\mathrm{S}}(x_0) = \min_\eta f^{\mathrm{S}}_\eta(x_0) = \frac{\mathrm{c}}{\Delta_x \left(1 + \left|\hat{k}_{S,x}(x_0,0,\omega)\right|\right)} . \tag{21}$$

Solving the Eq. (18) for f yields the frequency $f^{\mathrm{S}}_\eta(x,y,x_0)$ at which the secondary source located at x_0 contributes the η-th aliasing component P^{S}_η to a distinct position $\mathbf{x}$ inside the target region. The aliasing frequency $f^{\mathrm{S}}(x,y)$ for this position is defined as the frequency up to which no secondary source reproduces any aliasing at $\mathbf{x}$. Hence, the minimum of $f^{\mathrm{S}}_\eta(x,y,x_0)$ over all loudspeakers and aliasing components η defines this frequency as

$$f^{\mathrm{S}}(x,y) = \min_{x_0} \frac{\mathrm{c}}{\Delta_x \left|\hat{k}_{S,x}(x^*_0,0,\omega) - \hat{k}_{G,x}(x - x^*_0, y, \omega)\right|} . \tag{22}$$

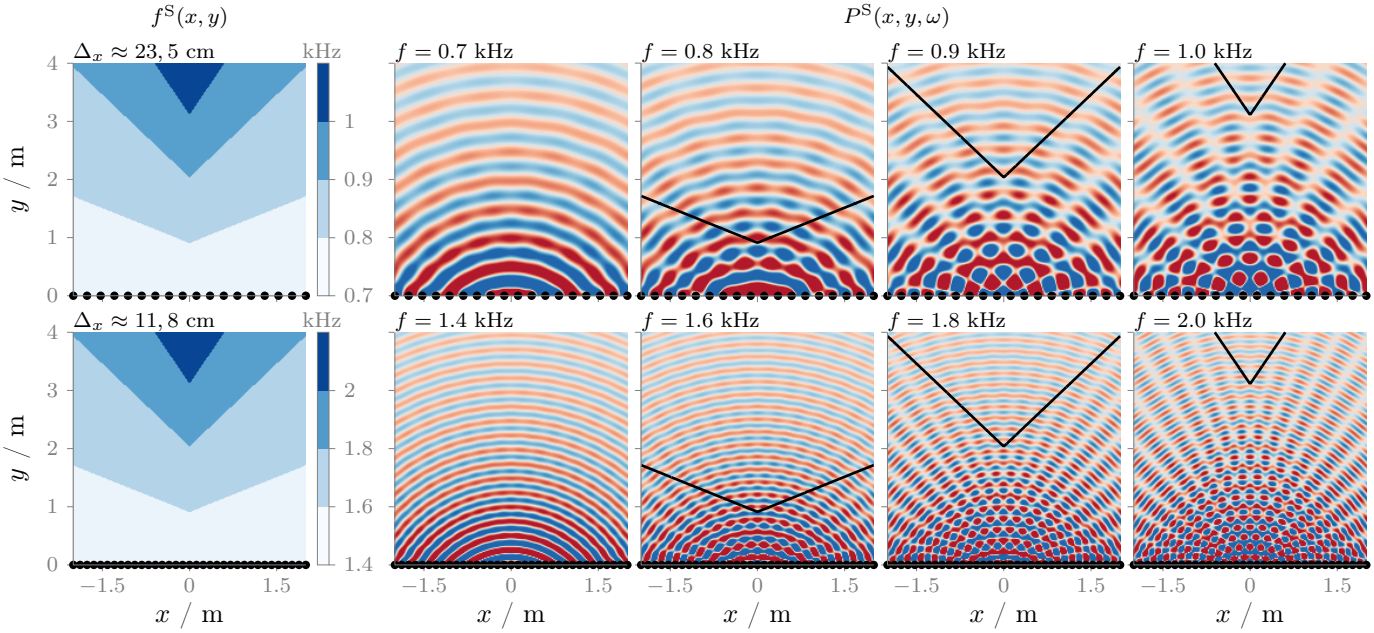

Figure 4: The left plots show the estimated aliasing frequency $f^{\mathrm{S}}(x, y)$, see Eq. (22), for a point source located at $\mathbf{x}_{\mathrm{ps}} = [0, -1, 0]^{\mathrm{T}}$ m and a linear SSD with two different sampling distances. The frequency was computed via exhaustive search over a dense grid of x_0. The SSD was truncated to $[-2, 2]$ m. The right plots show the real part of the synthesised sound fields $P^{\mathrm{S}}(x, y, \omega)$, see Eq. (5), for different frequencies f. For the positions below the solid black lines, the anti-aliasing criterion involving Eq. (22) is violated.

An analytical solution to the minimisation problem for simple desired sound fields $S(x, y, \omega)$, e.g. point sources and plane waves, is possible. However, iterative optimisation algorithms may be feasible for arbitrary sound fields. The left plots in Fig. (4) show the estimated spatial aliasing frequency for a virtual point source for two different sampling distances Δ_x. For validation, the reproduced sound field $P^{\mathrm{S}}(x, y, \omega)$ is plotted in the remaining plots: The predicted spatial phenotype of aliasing with stronger artefacts at positions closer to the SSD is in agreement with the plotted sound fields. As the presented ray model is only an approximation of underlying synthesis problem, the strict separation between aliasing-free and aliased regions does not reflect the nature of the artefacts gradually reducing with increasing distance to the SSD.

5 Conclusion

This paper presented a geometrical model for the spatial aliasing artefacts occurring in practical implementations of WFS. This allowed for the formulation of local anti-aliasing criteria depending on the position of the loudspeakers and/or the listener. The model correctly predicting stronger aliasing artefacts for positions near the SSD. An extension of the model to arbitrary SSD geometries including non-uniform sampling scheme is possible by treating latter as locally linear with a local sampling distance. Other SFS techniques such as Near-Field-Compensated Higher Order Ambisonics (NFC-HOA) [9] may be included into the model. Future work may also connect the predicted aliasing frequency with results from listening tests investigating the perceived impairment in SFS, e.g. [10].

References

[1] A. J. Berkhout. "A Holographic Approach to Acoustic Control". In: *J. Aud. Eng. Soc.* 36.12 (1988), pp. 977–995.

[2] E. N. G. Verheijen. "Sound Reproduction by Wave Field Synthesis". PhD thesis. Delft University of Technology, 1997.

[3] E. W. Start. "Direct Sound Enhancement by Wave Field Synthesis". PhD thesis. Delft University of Technology, 1997.

[4] E. G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. London: Academic Press, 1999.

[5] G. Firtha, P. Fiala, F. Schultz, and S. Spors. "Improved Referencing Schemes for 2.5D Wave Field Synthesis Driving Functions". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 25.5 (2017), pp. 1117–1127.

[6] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders. *Fundamentals of Acoustics*. 4th Ed. Dec. 1999.

[7] P. M. Morse and H. Feshbach. *Methods of theoretical physics*. New York: McGraw-Hill, 1953.

[8] B. Girod, R. Rabenstein, and A. Stenger. *Signal and Systems*. Wiley, 2001.

[9] J. Daniel. "Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format". In: *Proc. of 23rd Intl. Aud. Eng. Soc. Conf. on Signal Processing in Audio Recording and Reproduction*. Copenhagen, Denmark, 2003.

[10] H. Wierstorf, C. Hohnerlein, S. Spors, and A. Raake. "Coloration in Wave Field Synthesis". In: *Proc. of 55th Intl. Aud. Eng. Soc. Conf. on Spatial Audio*. Helsinki, Finland, Aug. 2014.