Audio Engineering Society

Convention Paper

Presented at the 127th Convention
2009 October 9–12 New York NY, USA

Physical and Perceptual Properties of Focused Sources in Wave Field Synthesis

Sascha Spors, Hagen Wierstorf, Matthias Geier and Jens Ahrens

Deutsche Telekom Laboratories, Technische Universität Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany

Correspondence should be addressed to Sascha Spors (Sascha.Spors@telekom.de)

ABSTRACT

Wave Field Synthesis is a well-established high-resolution spatial sound reproduction technique. Its physical basis allows to reproduce almost any desired wave field, even sound sources positioned in between the loudspeakers and the listener. Such sources are known as focused sources. Focused sources have a number of remarkable physical properties, especially in the context of spatial sampling. This paper presents a detailed analysis of the physical properties of focused sources as well as their perceptual impact. Additionally, results of a first informal listening experiment are discussed in order to gather knowledge on the perceptual relevance of the derived artifacts in practical implementations.

1. INTRODUCTION

Wave Field Synthesis (WFS) aims at physically reproducing the sound of complex acoustic scenes as naturally as possible. In theory, WFS creates an almost physically correct reproduction of almost any virtual wave field by using a continuous distribution of acoustic sources (so called secondary sources) placed around the listening area. Amongst other interesting properties, reproduction systems based on physical principles allow to reproduce virtual sources which are positioned in the area between the loudspeakers and the listener [1, 2]. These are known as focused sources, due to their strong relation to acoustic focusing [3].

The physical theory of WFS assumes a spatially continuous distribution of secondary sources. In practical implementations of WFS the secondary source distribution will be realized by a limited number of loudspeakers placed at discrete positions. This implies a spatial sampling process that typically leads to spatial aliasing artifacts. For focused sources these sampling artifacts are of special interest. Some of the interesting properties of focused have already been discussed in a previous study by the authors [4]. This paper extends the previous study in various

ways. The physical properties of focused sources are studied in more detail as well as their perceptual properties. It is shown that focused sources exhibit a number of perceptually relevant artifacts that will be audible for large WFS systems if no further action is taken. This paper is organized as follows: Section 2 introduces the theory of WFS and focused sources. Due to the underlying geometry, a spatio-temporal frequency domain description has shown to be very effective in describing the effects of spatial sampling. Section 3 reviews this spatio-temporal frequency domain description and models the sampling process in this domain. Section 4 discusses various physical properties of focused sources for monochromatic and broadband signals. These results will serve as basis for the discussion of their perceptual properties in Section 5. Here results from the literature will be used to estimate the perceptual artifacts. Furthermore results from a first informal listening test will be presented to validate the findings of this paper.

2. WAVE FIELD SYNTHESIS

The following section introduces the basic theory of WFS, as well as the concept of focused sources.

2.1. Theory of WFS

Typical implementations of WFS systems are restricted to the reproduction in a plane only using (piecewise) linear loudspeaker arrays. The theoretical basis for this setup is given by the first Rayleigh integral [5, 6]. It states that a linear distribution of monopole sources (secondary sources) is capable of reproducing a desired wave field (virtual source) in one of the half planes defined by the linear distribution. The wave field in the other half plane is a mirrored version of the desired wave field.
Without loss of generality the geometry depicted in figure 1 is assumed: A linear secondary source distribution which is located on the x-axis ($y = 0$) of a Cartesian coordinate system. The reproduced wave field is given by

$$P(\mathbf{x},\omega) = -\int_{-\infty}^{\infty} D(\mathbf{x}_0,\omega) G(\mathbf{x}-\mathbf{x}_0,\omega) dx_0 \;, \quad (1)$$

where $\mathbf{x} = [x\, y]^T$ with $y > 0$ and $\mathbf{x}_0 = [x_0\, 0]^T$. The functions $D(\mathbf{x}_0,\omega)$ and $G(\mathbf{x}-\mathbf{x}_0,\omega)$ denote the (secondary source) driving function and the wave field

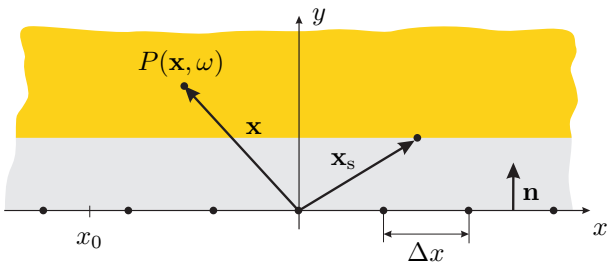


Fig. 1: Geometry used to investigate the physical properties of focused sources. The yellow area denotes the listening area.

emitted by the secondary sources, respectively. In WFS, the secondary source driving function is given by considering the first Rayleigh integral [7], as

$$D(\mathbf{x}_0,\omega) = 2\frac{\partial}{\partial \mathbf{n}} S(\mathbf{x},\omega)\Big|_{\mathbf{x}=\mathbf{x}_0} \;, \quad (2)$$

where $\frac{\partial}{\partial \mathbf{n}}$ denotes the directional gradient with $\mathbf{n} = [0\,1]^T$ and $S(\mathbf{x},\omega)$ the wave field of the virtual source. In theory, the wave field of the secondary sources is given by the two-dimensional free-field Green's function $G_{2\mathrm{D}}(\mathbf{x}|\mathbf{x}_0,\omega)$ for two-dimensional reproduction. $G_{2\mathrm{D}}(\mathbf{x}|\mathbf{x}_0,\omega)$ can be interpreted as the field of a line source intersecting the listening plane at the position $\mathbf{x}_0$. In practical implementations of WFS systems, loudspeakers with closed cabinets are used as secondary sources. These approximately exhibit the characteristics of the three-dimensional free-field Green's function $G_{3\mathrm{D}}(\mathbf{x}|\mathbf{x}_0,\omega)$, hence of acoustic point sources. For $\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0| \gg 1$, both the two- and three-dimensional Green's function can approximately be related to each other as follows [8]

$$\underbrace{\frac{j}{4} H_0^{(2)}\left(\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|\right)}_{G_{2\mathrm{D}}(\mathbf{x}-\mathbf{x}_0,\omega)} \approx -\sqrt{\frac{2\pi|\mathbf{x}-\mathbf{x}_0|}{j\frac{\omega}{c}}}\; \underbrace{\frac{1}{4\pi}\frac{e^{-j\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|}}{|\mathbf{x}-\mathbf{x}_0|}}_{G_{3\mathrm{D}}(\mathbf{x}-\mathbf{x}_0,\omega)} \;, \quad (3)$$

where $H_0^{(2)}(\cdot)$ denotes the Hankel function of the second kind and zeroth-order. Equation (3) states that in principle point sources can be used as secondary sources for two-dimensional reproduction. However, it can also be seen that the amplitude and the temporal spectrum of the reproduced wave field will not

be correct in this case. The latter can be corrected by temporal filtering of the virtual source signal. The amplitude errors can only be corrected to some extent [9].
The reproduction in a plane using secondary point sources is referred to as 2.5D reproduction. This situation will be assumed in the following if nothing else is explicitly mentioned.
According to (3), the driving function for 2.5D reproduction is given by

$$D_{2.5\mathrm{D}}(\mathbf{x}_0,\omega) = \sqrt{\frac{1}{j\frac{\omega}{c}}}\ \underbrace{\sqrt{2\pi\,|\mathbf{x}_{\mathrm{ref}}-\mathbf{x}_0|}}_{g_0}\ D_{2\mathrm{D}}(\mathbf{x}_0,\omega)\ . \tag{4}$$

It can be shown [1, 10] that the amplitude can be chosen as approximately correct on a line parallel to the secondary source distribution (reference line). In this case, the factor g_0 is constant.

2.2. Driving Function for Focused Sources

Acoustic focusing refers to a variety of techniques to focus acoustic wave fields. These have been developed in diverse application areas like e. g. material analysis or medicine [11, 12]. The basic concept underlying most of the techniques is the principle of time-delay law focusing or more generally of time-reversal acoustic focusing [3, 13, 14, 15]. The goal of most applications is a concentration of acoustic energy at the focus point.
For sound reproduction, the goal is to create the illusion of an acoustic source that is situated in front of the loudspeaker array. Note, that this condition implies an important constraint in comparison to the traditional time-reversal principle. Only contributions emerging from the desired focused source should be reproduced at the listener position in order to not confuse the auditory impression by other contributions. Time reversal techniques may result in additional contributions, especially for curved or closed secondary source contours.
Since the secondary sources emit a wave field that travels towards the listener, one can only expect that the desired wave field of a focused source is correct if the focus point is located in between the secondary source distribution and the listener. In the context of WFS, this is a well known limitation of focused sources [1].
We aim at investigating the physical and perceptual artifacts for a focused source with focus point $\mathbf{x}_\mathrm{s}$. One method to derive the desired driving function is to assume an acoustic sink placed within the listening area, as a model for the desired virtual source. For two-dimensional reproduction the virtual source would be given by a line sink. For 2.5D reproduction a reasonable virtual source model is the field of a line sink which has equivalent spectral characteristics as a point source. It is given as [7]

$$S(\mathbf{x},\omega) = \frac{j}{4}\sqrt{j\frac{\omega}{c}}H_0^{(1)}\left(\frac{\omega}{c}\,|\mathbf{x}-\mathbf{x}_\mathrm{s}|\right)\ , \tag{5}$$

where $\mathbf{x}_\mathrm{s} = [x_\mathrm{s}\ y_\mathrm{s}]^T$ denotes the position of the focus point with $y_\mathrm{s} > 0$. Introducing (5) into (4) results in the 2.5D driving function for a focused source

$$D_{2.5\mathrm{D}}(\mathbf{x}_0,\omega) = -\frac{j\frac{\omega}{c}}{2}g_0\frac{y_0-y_\mathrm{s}}{|\mathbf{x}_0-\mathbf{x}_\mathrm{s}|}H_1^{(1)}\left(\frac{\omega}{c}\,|\mathbf{x}_0-\mathbf{x}_\mathrm{s}|\right)\ , \tag{6}$$

where g_0 denotes an amplitude normalization factor. Note, that the spectral correction due to the secondary source type mismatch cancels out.
Equation (6) can be related to the traditional formulation of the driving function used in WFS [1, eq.(2.30)] by replacing the Hankel function in (6) by its large-argument approximation [8]

$$D_{2.5\mathrm{D}}(\mathbf{x}_0,\omega) \approx -g_0\sqrt{\frac{\frac{\omega}{c}}{2\pi j}}\frac{y_0-y_\mathrm{s}}{|\mathbf{x}_0-\mathbf{x}_\mathrm{s}|}\frac{e^{j\frac{\omega}{c}|\mathbf{x}_0-\mathbf{x}_\mathrm{s}|}}{\sqrt{|\mathbf{x}_0-\mathbf{x}_\mathrm{s}|}}\ , \tag{7}$$

where g_0 is explicitly given in [1]. Equation (7) is equal to the traditional formulation for a suitable choice of the normalization factor g_0. For the chosen geometry and a reference line parallel to the x-axis this factor does not depend on the actual secondary source position x_0.
As already mentioned above, the driving function for a focused source will not result in the reproduction of an acoustic sink for the entire listening area. The secondary sources are driven such that they create a wave field that converges towards the focus point $\mathbf{x}_\mathrm{s}$ (gray area in figure 1). After the focus point the field diverges again like the field of a point source located at the focus point (yellow area in figure 1). Note, that the reproduction of a focused source with WFS can be regarded as the 2.5D analogon to time-delay law focusing, as can be concluded from the exponential term in (7).

Figure 4(a) illustrates the reproduced wave field for a virtual focused source. The driving function given by (6) was used for the simulation.
Note, that the theory presented so far can be extended straightforward to curved or closed-contour secondary source distributions. However, in this case the secondary sources driven for a particular focused source have to be selected sensibly in order to maintain the desired impression. A suitable scheme can be found e. g. in [16].

3. SPATIO TEMPORAL FREQUENCY DOMAIN DESCRIPTION

The following section will derive a spatio temporal frequency domain formulation of the respective wave fields. For purely two-dimensional reproduction of focused sources this already has been derived in [4]. Here we focus on 2.5D reproduction.

3.1. Reproduced Wave Field

The calculation of the reproduced wave field, as given by (1), together with the shift-invariant field of the secondary point sources, as given by (3), can be interpreted as a convolution with respect to the x-coordinate. Applying a spatial Fourier transformation to (1) with respect to the x-coordinate results consequently in

$$\tilde{P}(k_x, y, \omega) = -\tilde{D}(k_x, \omega)\, \tilde{G}(k_x, y, \omega) \;, \qquad (8)$$

where k_x denotes the spatial frequency (wave number). Spatial frequency domain quantities are denoted by a tilde over the respective variable. Equation (8) states that the spatio temporal spectrum of the reproduced wave field is given by multiplying the spatial temporal spectra of the driving function and the secondary sources, respectively. The derived frequency domain description is very powerful for the analysis of spatial sampling and near-field effects as will be shown in the following section. However, the spatial Fourier transform of the driving function and the secondary sources are required for the analysis.

3.2. Spectra of Secondary Sources and Driving Function

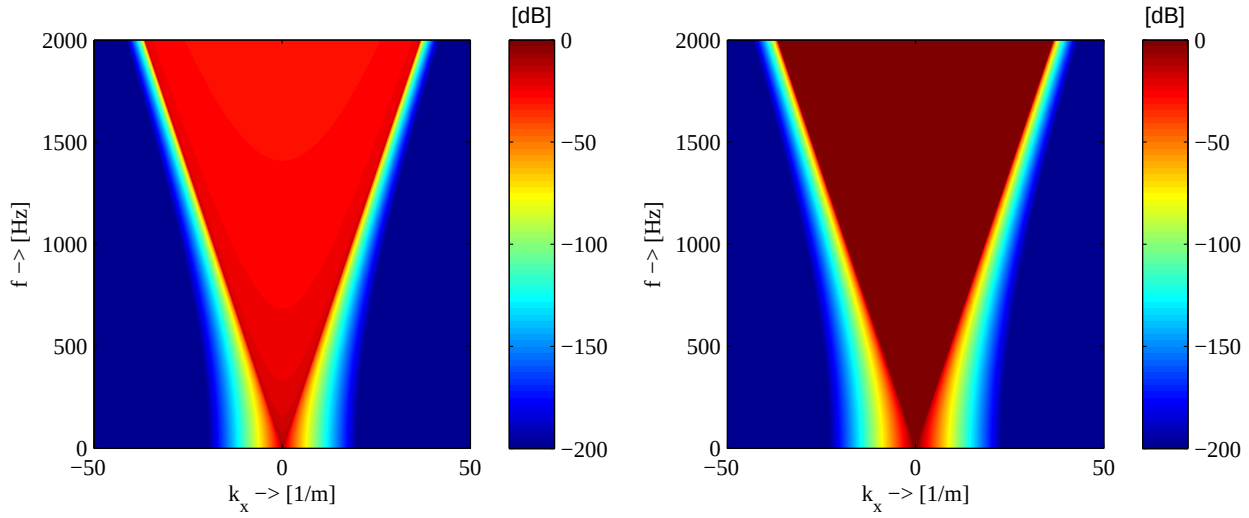
The spatial Fourier transformation of the secondary point sources $\tilde{G}_{3\mathrm{D}}(k_x, y, \omega)$ is is given by [17, 3.876-1/2]

$$\tilde{G}_{3\mathrm{D}}(k_x, y, \omega) = \begin{cases} -\frac{j}{4} H_0^{(2)}(\sqrt{(\frac{\omega}{c})^2 - k_x^2}\, y) & \text{for } |k_x| < \left|\frac{\omega}{c}\right| \;, \\ \frac{1}{2\pi} K_0(\sqrt{k_x^2 - (\frac{\omega}{c})^2}\, y) & \text{for } \left|\frac{\omega}{c}\right| < |k_x| \;, \end{cases} \qquad (9)$$

which is valid for $y > 0$. $K_0(\cdot)$ denotes zeroth-order modified Bessel function of second kind. The spectrum $\tilde{G}_{3\mathrm{D}}(k_x, y, \omega)$ consists of two contributions: a traveling wave contribution for $|k_x| < \left|\frac{\omega}{c}\right|$ and an evanescent contribution for $\left|\frac{\omega}{c}\right| < |k_x|$. Figure 2(a) shows the absolute value of $\tilde{G}_{3\mathrm{D}}(k_x, y, \omega)$ for $y = 1$ m. The wedge $\left|\frac{\omega}{c}\right| < |k_x|$ containing the traveling wave contributions can be seen clearly. The evanescent contributions are only well visible for the low frequencies due to their rapid decay. Although these contributions decay rapidly, $\tilde{G}(k_x, y, \omega)$ is not strictly bandlimited with respect to the spatial frequency k_x. This holds especially for low frequencies and/or short distances y to the secondary source distribution.
The spectrum of the driving function for two-dimensional reproduction of a focused source was already derived in [4]. The driving function for 2.5D reproduction (6) differs only in terms of the spectral correction. However, since both the secondary sources and the desired virtual source are corrected this correction effectively cancels out in the equations. Hence, the spatial spectrum of the driving function is given as [4]

$$\tilde{D}_{2.5\mathrm{D}}(k_x, \omega) = e^{j k_x x_\mathrm{s}} \begin{cases} e^{j\sqrt{(\frac{\omega}{c})^2 - k_x^2}\, y_\mathrm{s}} & \text{for } |k_x| < \left|\frac{\omega}{c}\right| \;, \\ e^{-\sqrt{k_x^2 - (\frac{\omega}{c})^2}\, y_\mathrm{s}} & \text{for } \left|\frac{\omega}{c}\right| < |k_x| \;, \end{cases} \qquad (10)$$

which is valid for $y_\mathrm{s} > 0$.
As for the spectrum of the secondary sources (9), the spectrum of the driving function (10) consists of a propagating and an evanescent part. Figure 2(b) shows the absolute value of $\tilde{D}_{2.5\mathrm{D}}(k_x, \omega)$ for $\mathbf{x}_\mathrm{s} = (0, 1)$ m. As for the spectrum of the secondary sources, the evanescent contributions for $\left|\frac{\omega}{c}\right| < |k_x|$ decay rapidly.
The reproduced wave field for a focused source is given by introducing (9) and (10) into (8).
Practical implementations of WFS will always be based on spatially discrete secondary sources. This



(a) secondary source spectrum

(b) driving function spectrum

Fig. 2: Spectrum (absolute value) of the secondary sources $\tilde{G}_{3\mathrm{D}}(k_x, y, \omega)$ for $y = 1$ m and driving function $\tilde{D}_{2.5\mathrm{D}}(k_x, \omega)$ for a virtual source at position $\mathbf{x}_\mathrm{s} = (0, 1)$ m.

constitutes a spatial sampling of the continuous secondary source distribution.

3.3. Spatial Sampling of Secondary Source Distribution

The discretization of the secondary source distribution is modeled by spatial sampling of the driving function. This is performed by multiplying $D(x, \omega)$ with a series of spatial Dirac functions at the positions of the loudspeakers. For an equidistant spacing this reads

$$D_S(x, \omega) = D(x, \omega) \cdot \frac{1}{\Delta x} \sum_{\mu=-\infty}^{\infty} \delta(x - \Delta x \mu) \ , \quad (11)$$

where $D_S(x, \omega)$ denotes the sampled driving function and Δx the distance (sampling period) between the sampling positions (indicated by the dots • in figure 1). Applying a spatial Fourier transformation to (11) results in

$$\tilde{D}_S(k_x, \omega) = 2\pi \sum_{\eta=-\infty}^{\infty} \tilde{D}\left(k_x - \frac{2\pi}{\Delta x}\eta, \omega\right) \ . \quad (12)$$

Equation (12) states that the spectrum $\tilde{D}_S(k_x, \omega)$ of the sampled driving function is given as a superposition of the shifted continuous spectra $\tilde{D}(k_x - \frac{2\pi}{\Delta x}\eta, \omega)$ of the driving function. Introducing the spectrum of the sampled driving function $\tilde{D}_S(k_x, \omega)$ into (8) results in the spectrum $\tilde{P}_S(k_x, y, \omega)$ of the wave field reproduced by a spatially discrete secondary source distribution. Figure 3 illustrates on a qualitative level the calculation of the reproduced wave field. The blue areas denote the propagating parts of the driving function $\tilde{D}_S(k_x, \omega)$ and the secondary sources $\tilde{G}(k_x, y, \omega)$, respectively, the light gray areas their evanescent contributions. As for virtual point sources [18], four different types of overlaps between the spectrum of the sampled driving function and the secondary sources can be identified. Two propagating contributions emerge from the overlap of the propagating and evanescent contributions of the driving function with the propagating part of the secondary sources. Two evanescent contributions emerge from the overlap of the propagating and evanescent contributions of the driving function with the evanescent part of the secondary sources. These four cases are not discussed in detail here, please refer to [18] for a detailed treatment in the context of virtual point sources. However, the results derived there can be transferred straightforwardly to focused sources.

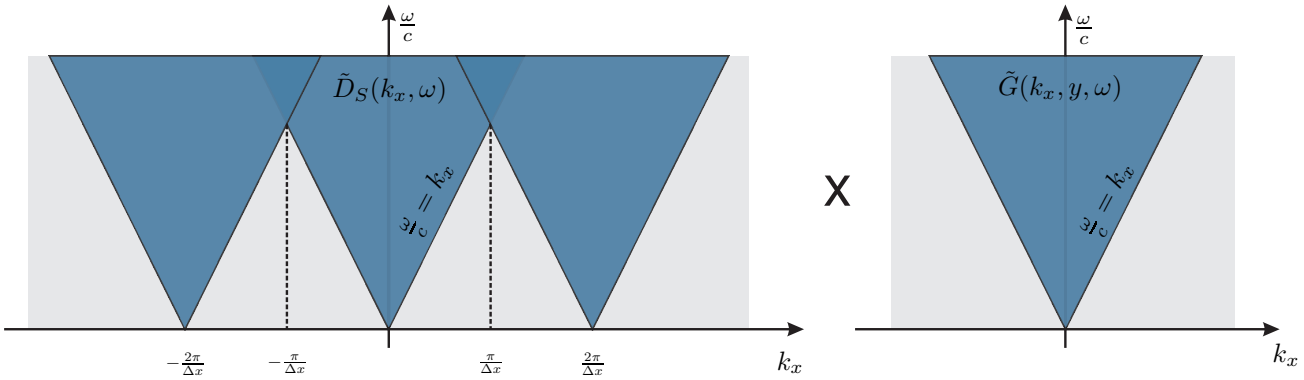


Fig. 3: Qualitative illustration of the computation of the spectrum of the reproduced wave field $\tilde{P}_S(k_x, y, \omega)$ for a sampled secondary source distribution.

3.4. Aliasing Artifacts in the Driving Function

The reproduction with spatially discrete secondary sources can be interpreted as a sampling and interpolation process. The driving function is sampled at the spatially discrete secondary source positions and the secondary sources propagate their field into the continuous space. Consequently, artifacts due to the secondary source sampling can be expected when (1) the spectrum of the driving function is not band-limited, and (2) the spectrum of the secondary sources is not band-limited.

The first condition ensures that there exists a sampling interval Δx where no spectral overlaps occur in the sampled driving function, the second condition ensures that the spectral repetitions in the sampled driving function will be filtered out by the characteristics of the secondary sources.

Inspection of (9) and (10) reveals that both the spectrum of the driving function and the secondary sources are not strictly bandlimited for a given frequency ω due to their evanescent contributions. These contributions decay rapidly, except for low frequencies and/or short distances y to the secondary source distribution. When considering only the propagating part of the driving function the following anti-aliasing condition can be derived

$$f_{\mathrm{al}} \leq \frac{c}{2\Delta x} \,. \tag{13}$$

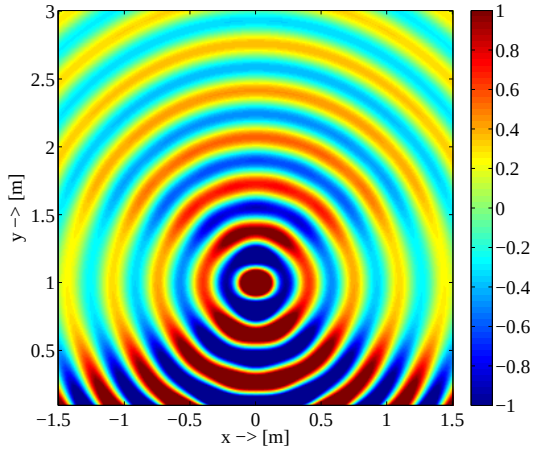
It was already shown in [4] that focused sources exhibit interesting properties with respect to their sampling artifacts. Aliasing which is present in the driving function may not lead to sampling artifacts in the entire listening area. Positions close to the focus point show less aliasing artifacts than expected when analyzing the driving function only. The next section will discuss this and other interesting properties of focused sources.

4. PHYSICAL PROPERTIES OF FOCUSED SOURCES

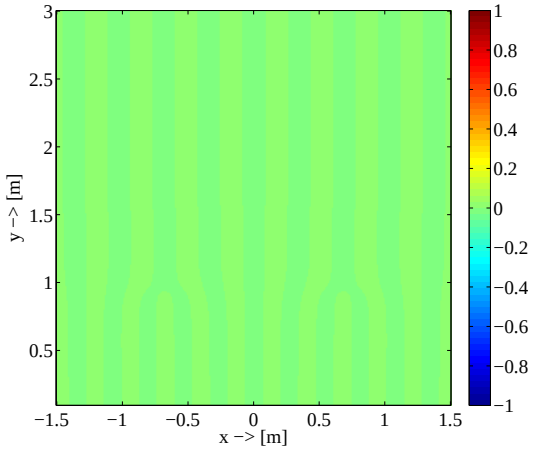
The following section discusses the physical properties of focused sources. Note, that a discussion of the basic properties of focused sources in the context of two-dimensional reproduction has already been presented in [4]. This paper focuses on the properties of 2.5D reproduction. For completeness, the same properties that have been discussed for two-dimensional reproduction are also discussed here, even when the results are similar.

4.1. Monochromatic Signals

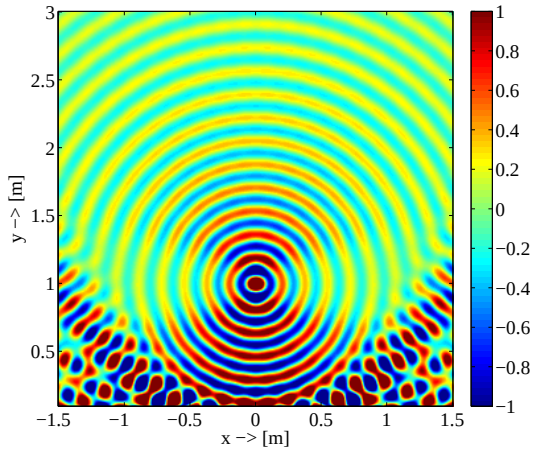
In the following, the influence of spatial sampling of secondary sources in the context of focused sources will be investigated. The representation of the driving function in the spatio-temporal frequency domain is especially useful to investigate spatial sampling artifacts. This is due to the fact that spatial sampling results in repetitions of the spatial spectrum. The spectrum of the reproduced wave field for a sampled secondary source distribution is given by introducing the spectrum of the driving function (10) together with (12) and (9) into (8). The reproduced wave field is given by evaluating



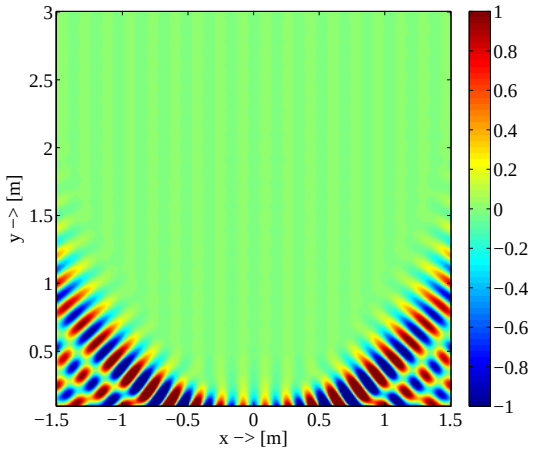
(a) reproduced wave field $f = 1000$ Hz



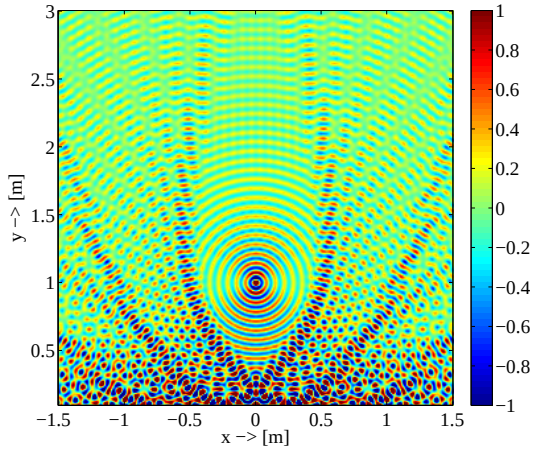
(b) aliasing contributions $f = 1000$ Hz



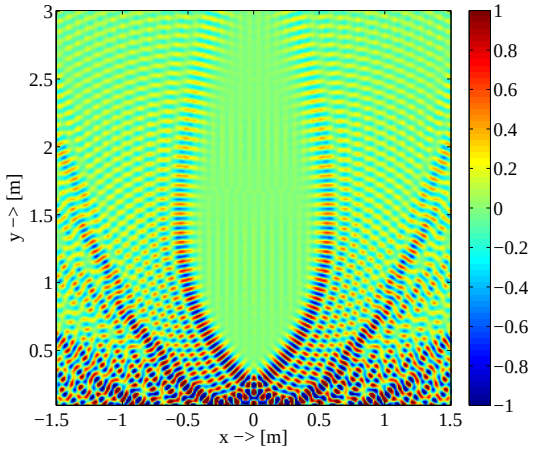
(c) reproduced wave field $f = 2000$ Hz



(d) aliasing contributions $f = 2000$ Hz



(e) reproduced wave field $f = 5000$ Hz



(f) aliasing contributions $f = 5000$ Hz

Fig. 4: Reproduced wave field and its sampling artifacts for the reproduction of a monochromatic focused source ($\mathbf{x}_\mathrm{s} = (0, 1)$ m, $\Delta x = 0.15$ m, 2.5D WFS).

the spectral repetitions of the driving function for all η. The contributions emerging from the secondary source sampling by evaluation only the contributions for $\eta \neq 0$. In order to illustrate the properties of focused sources both cases are evaluated separately and shown in figure 4. For instance, figures 4(a) and 4(b) show the reproduced wave field and its aliasing contributions for a focused source at position $\mathbf{x}_\mathrm{s} = (0, 1)$ m emitting a monochromatic signal with frequency $f = 1$ kHz. The anti-aliasing frequency (13) for the simulated scenario is $f_\mathrm{al} \approx 1140$ Hz.

The focus point is clearly visible at the desired position, no aliasing artifacts are visible since the condition (13) is fulfilled. The contributions that can be seen in figure 4(b) are near-field artifacts that will be discussed in detail later. The situation changes when increasing the frequency of the focused source. Figures 4(c) to 4(f) show the reproduced wave field and its aliasing contributions for $f = 2$ kHz and $f = 5$ kHz. Prominent aliasing artifacts become visible. However, in the vicinity of the focus point these aliasing artifacts decrease leaving the focus point (almost) free of spatial aliasing. This is a very remarkable property of focused sources, which is based on the fact that at the focus point the wave fields emitted by the individual secondary sources superimpose with equal phase.

This is also illustrated in figure 5, where the phase of the wave field of figure 4(a) is shown. Furthermore, a phase jump can be observed along the line parallel to the x-axis with $y = 1$ m. This is due to the fact that the wave field converges towards the focus point and diverges after the focus point. This behavior is also known from time-reversal focusing [19].

4.2. Evanescent Contributions

It has been derived in [18, 10] that spatial sampling of the secondary source distribution leads to evanescent contributions in the reproduced wave field when reproducing virtual plane waves or point sources. The formulation of the reproduced wave field in the spatio-temporal frequency domain allows to explicitly split the reproduced wave field into its propagating and evanescent contributions. For this purpose, the case differentiation in the spectra of the driving function (10) and the secondary sources (9) are evaluated separately for the propagating and evanescent parts. Figure 6 shows these two contributions for

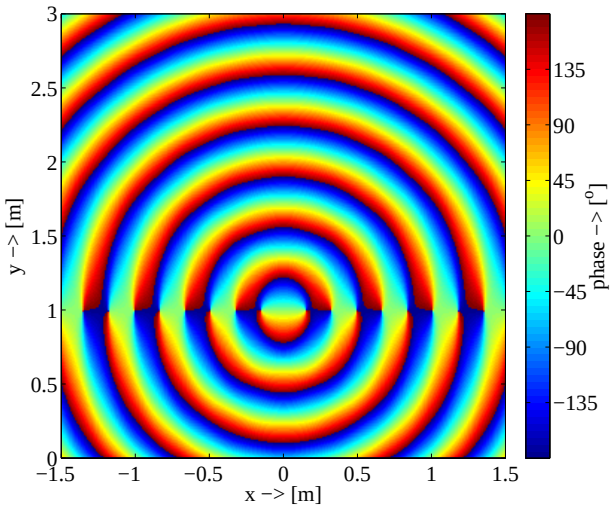
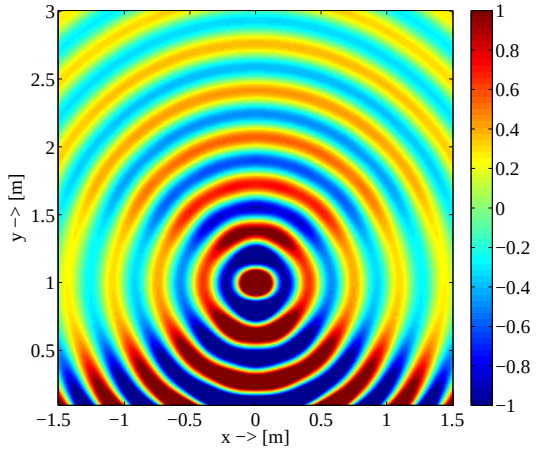


Fig. 5: Phase of reproduced wave field for the reproduction of a monochromatic focused source ($f = 1$ kHz, $\mathbf{x}_\mathrm{s} = (0, 1)$ m, $\Delta x = 0.15$ m, 2.5D WFS)

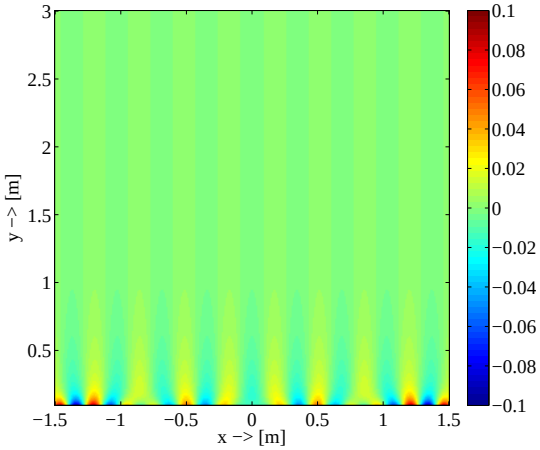
different frequencies of the monochromatic focused source. Note the different color scales that have been used for the propagating and evanescent contributions. The evanescent contributions are present near the secondary source distribution and decay quite rapid with increasing distance. It can be observed further, that no specific evanescent contributions are present near the focus point. This implies that these contributions of the virtual source may not be reproduced as if a real source would be present at the focus point. This effect is also known from time-reversal focusing techniques [3]. However, its hard to draw conclusions from this finding since the role of evanescent contributions in human perception does not seem to be clear.

4.3. Amplitude Distribution

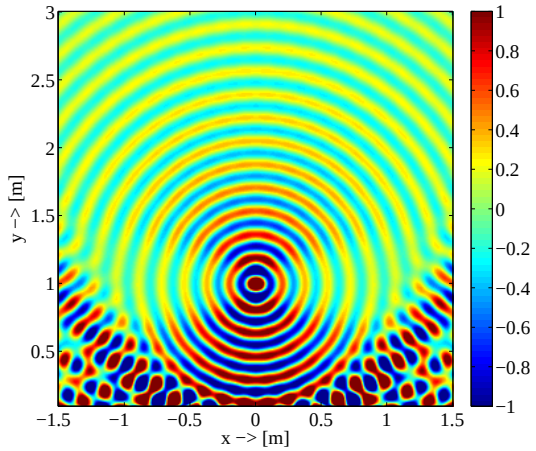
It is well known that the reproduction in a plane using secondary point sources (2.5D reproduction) may lead to an amplitude mismatch in comparison with the desired virtual source [9, 1]. In WFS this holds for the reproduction of virtual point sources and plane waves. Figure 7 shows the amplitude of the reproduced wave field of figure 4(a) along two axes. Figure 7(a) shows the normalized amplitude of the focused source along the y-axis ($x = 0$ m). The amplitude of a real source placed at the po-



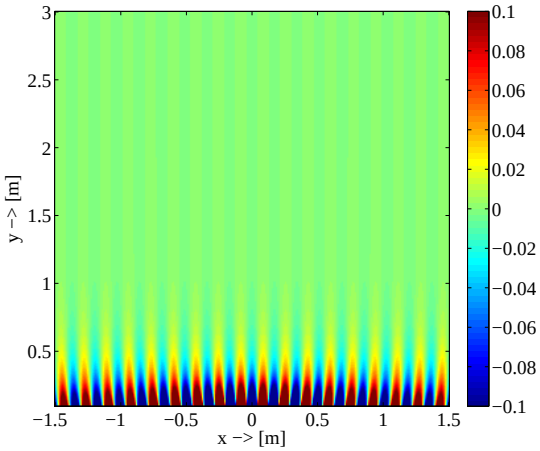
(a) propagating wave field $f = 1000$ Hz



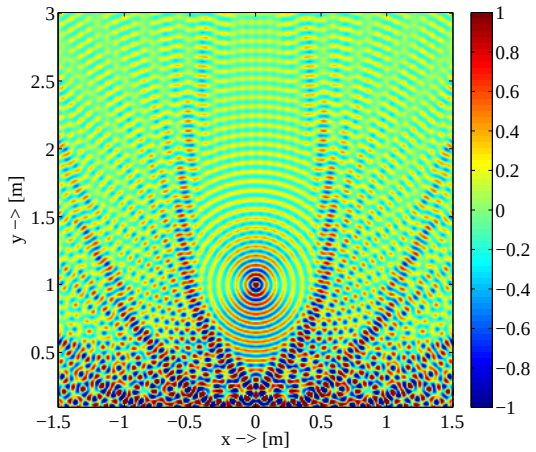
(b) evanescent wave field $f = 1000$ Hz



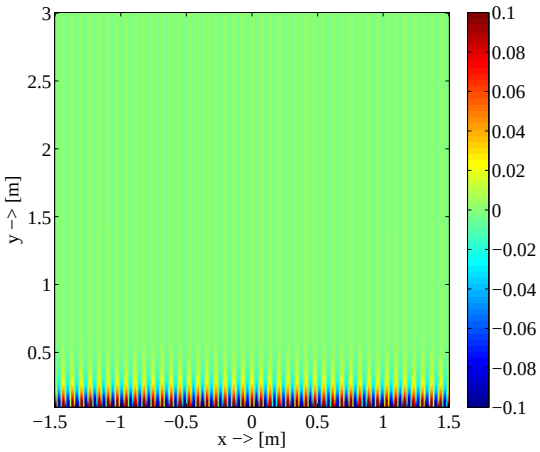
(c) propagating wave field $f = 2000$ Hz



(d) evanescent wave field $f = 2000$ Hz

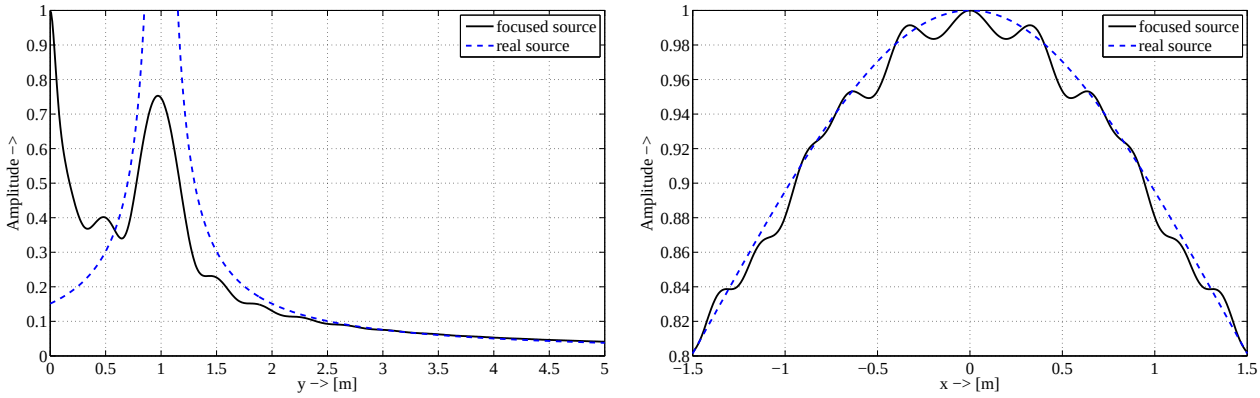


(e) propagating wave field $f = 5000$ Hz



(f) evanescent wave field $f = 5000$ Hz

Fig. 6: Reproduced wave field split into propagating and evanescent contributions for the reproduction of a monochromatic focused source ($\mathbf{x}_\mathrm{s} = (0, 1)$ m, $\Delta x = 0.15$ m, 2.5D WFS).



(a) $x = 0$ m (b) $y = 3$ m

Fig. 7: Amplitude distribution of the reproduced wave field shown in figure 4(a) along the y-axis ($x = 0$ m) and on a line parallel to the x-axis for $y = 3$ m.

sition of the focused source is shown for reference. It can be observed that, after some distance to the focus point, the amplitude decay is almost correct. This is remarkable, since the amplitude decay, for a (non-focused) virtual point source, is not reproduced correctly along this axis. The traditional driving function used for virtual point sources results in an amplitude error which is inversely proportional to the square root of the distance to the virtual point source. This is due to the secondary source dimensionality mismatch in 2.5D reproduction. However, for focused sources there is no amplitude error.

Figure 7(b) shows the normalized amplitude of the focused source along a line parallel to the x-axis for $y = 3$ m. Also here, the amplitude of the reproduced wave field is almost correct. The deviations in figure 7(a) and 7(b) result from the far-field approximations used in the derivation of the driving function and truncation artifacts.

4.4. Truncated Arrays

Up to now, the linear secondary source distribution was assumed to be of infinite length in the x-direction. However, practical implementations of linear loudspeaker arrays will always be of finite length. The impact of this truncation on the reproduced wave field and the spatial sampling artifacts for virtual point sources and plane waves has been investigated in detail [20, 18, 21]. Mathematically, truncation can be modeled by multiplying the secondary source driving function $D(x_0, \omega)$ with a suitable window function $w(x_0)$ [20]. Incorporating $w(x_0)$ into (1) yields the wave field $P_{\text{tr}}(\mathbf{x}, \omega)$ reproduced by a truncated linear array. We will not perform a quantitative analysis in the course of this paper and will rather stick to qualitative findings, and results that have been derived in the context of time-reversal focusing.

As described in [18] for virtual point sources, truncation leads to a limited listening area and other artifacts. The same argumentation to derive the listening area, as used for virtual point sources, holds also for focused sources when considering the time-delay law focusing analogon. Figure 8 shows an geometric approximation of the listening area for reproduction of a focused source with a truncated linear array.

The limited listening area is not the only artifact emerging from truncated secondary source contours. Truncation artifacts, due to the diffraction from the ends of the secondary source distribution, are also present. It was shown for virtual point sources that these can be modeled as additional sources at the ends [21]. Truncation artifacts can be improved by applying a smooth spatial window (tapering) to the driving function. Note, that truncation artifacts may be quite critical in the context of broadband signals since they may account to the pre-echo artifacts that will be discussed in the next section.

From the theory of time-delay law focusing it is also known that truncation increases the size of the focus point [3]. In physics, this phenomena is also known

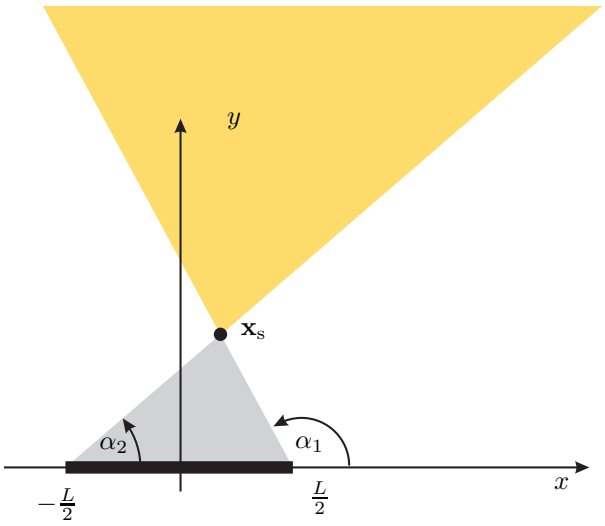


Fig. 8: Approximation of listening area (shown by yellow wedge) for reproduction of a focused source with a truncated linear array.

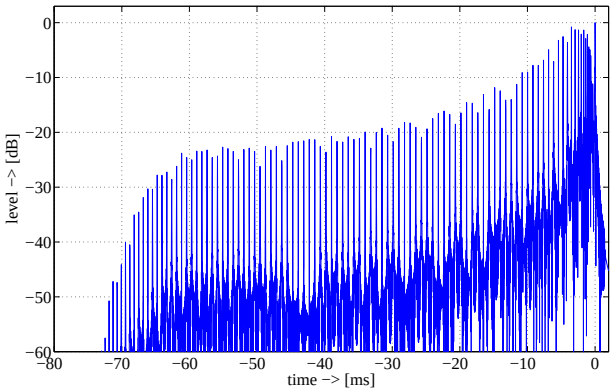


Fig. 10: Impulse response at position $\mathbf{x} = (15, 4)$ m ($\mathbf{x}_s = (0, 1)$ m, $\Delta x = 0.15$ m, $N = 200$, 2.5D WFS).

as diffraction limit since it determines the resolution of e. g. optic systems.

4.5. Broadband Signals

Of special interest are the properties of focused sources in the context of broadband signals, since these allow to draw conclusions on their perceptual properties. Some fundamental findings for 2.5D reproduction have already been derived in [4]. These will be summarized and extended. Their potential perceptual influence is discussed in Section 5.

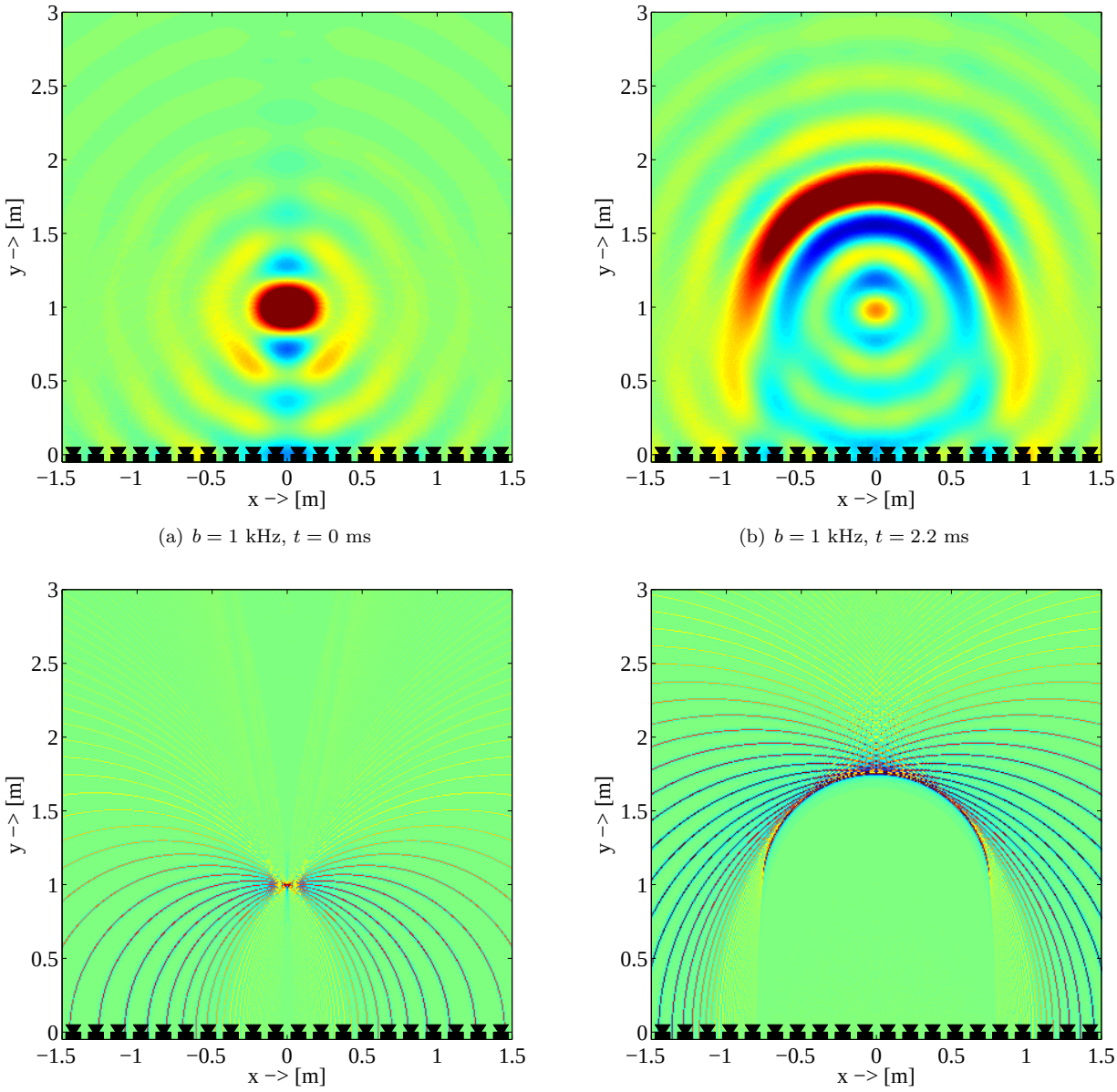
Numerical simulations of the reproduced wave field for different temporal bandwidths of the virtual source signal have been performed to illustrate the effect of spatial aliasing. Figure 9 shows two temporal snapshots of the resulting wave field (spatio-temporal impulse response) for two different bandwidths. Figures 9(a) and 9(b) show the wave field for a bandwidth below the spatial aliasing frequency of $b = 1$ kHz of the simulated setup. The focusing of the wave field at the focus point is clearly visible in figure 9(a). It can be observed in figure 9(b) that the wave front exhibits the desired properties of a point source placed at the focus point. No sampling artifacts are visible, as expected. The artifacts that can be observed are due to the truncation and the temporal band-limitation.

The situation changes dramatically when the bandwidth is increased. Figures 9(c) and 9(d) show the reproduced wave field for a typical audio bandwidth of $b = 20$ kHz. The same time instants as for the previous examples have been used. Aliasing artifacts in the form of additional wave fronts are clearly visible. Interestingly, these artifacts appear before the first wavefront of the focused source. The incidence angles of the additional wavefronts may differ from the desired direction of the virtual source for a given listener position.

A comparison of figures 9(a) and 9(c) reveals that the focus point is much smaller for a higher bandwidth. This is a fundamental result from time-reversal techniques [3]. It is also known that the size of the focus point increases when the aperture (total length) of the secondary source distribution decreases.

The temporal structure of the sampling artifacts is due to the time-reversal nature of acoustic focusing. Figure 10 shows the resulting impulse response for one fixed listener position. It can be concluded from figure 9 and 10 that the spatial aliasing artifacts for broadband signals result in pre-echos that may have different incidence angles than the desired focused source. The temporal extension of these depends on the position of the virtual source and listener, the loudspeaker distance and the total length of the loudspeaker array. Figure 10 shows a rather extreme situation.

In order to create the focus point of the source at $\mathbf{x}_s$ the second source with the greatest distant has to emit sound in the first place. This fact can be used



(a) $b = 1$ kHz, $t = 0$ ms

(b) $b = 1$ kHz, $t = 2.2$ ms

(c) $b = 20$ kHz, $t = 0$ ms

(d) $b = 20$ kHz, $t = 2.2$ ms

Fig. 9: Reproduced wave field for the reproduction of a temporally bandlimited focused source ($\mathbf{x}_\mathrm{s} = (0, 1)$ m, $\Delta x = 0.15$ m, $N = 200$, 2.5D WFS)

to estimate the temporal extent of the pre-echos

$$t_{\max}(\mathbf{x}) = \arg\max_{\mathbf{x}_0} \left\{ \frac{|\mathbf{x}_0 - \mathbf{x}_\mathrm{s}| + |\mathbf{x}_\mathrm{s} - \mathbf{x}|}{c} - \frac{|\mathbf{x} - \mathbf{x}_0|}{c} \right\} . \quad (14)$$

Equation (14) was used to compute figure 13.

4.6. Pre-Equalization

The formulation of the driving function, as given by (6) or (7), includes a $\sqrt{\omega/c}$ pre-filter that is independent of the virtual source position. This filter is well known in WFS and typically termed as pre-equalization filter. It can be applied to the source signal before the calculation of the loudspeaker driving signals is split into the different loudspeaker channels [1]. The pre-equalization filter is typically applied only till the spatial aliasing frequency. It is evident from figure 3, that the spatial aliasing adds energy to the frequency regions where the spectral overlaps due to sampling occur.

A particular problem with pre-equalization, in the context of focused sources, is that their spatial sampling artifacts are very position-dependent. This can be seen in figure 4. Close to the position of the virtual source sampling artifacts are only present for relatively high frequencies. However, this also implies that the pre-equalization filter would have to be adapted to the listener position. Practical implementations use fixed pre-equalization filters. For moving listeners and/or moving focused sources this may lead to spectral distortion of the source signal. Figure 11 shows the frequency response without the pre-equalization filter at two different listener positions. It is evident that the frequency when the spatial aliasing becomes prominent (aliasing frequency) is very different for these two listener positions. While for the listener position close to the focused source with $\mathbf{x} = (0, 3)$ m the aliasing frequency is about 5 kHz, it is as low as 1 kHz for the distant position $\mathbf{x} = (15, 4)$ m. The 3 dB per Octave slope continues much further in the former case.

5. PERCEPTUAL PROPERTIES OF FOCUSED SOURCES

The presented physical properties have potential influence on the perception of focused sources. The perception of pre-echo artifacts is discussed in this section. Also coloration artifacts due to pre-equalization are considered. Thereafter, an informal listening test is carried out to evaluate the perception.

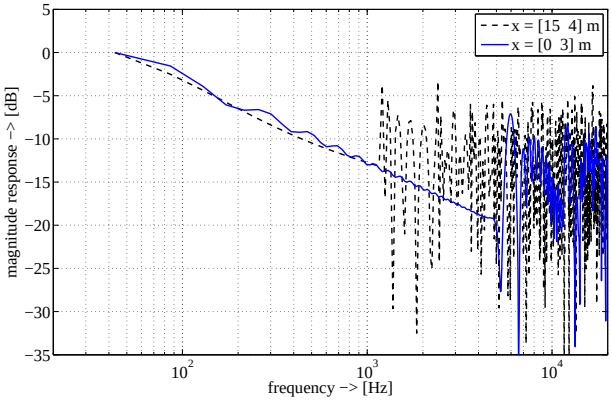
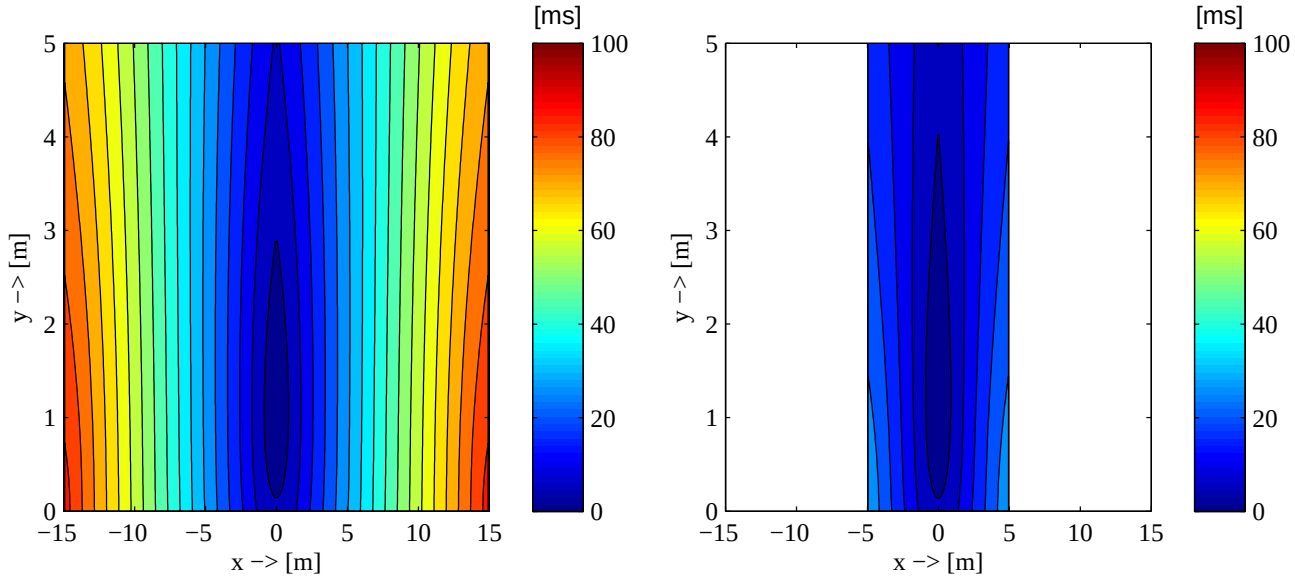


Fig. 11: Frequency response (normalized magnitude) for two listener positions without pre-equalization ($\mathbf{x}_\mathrm{s} = (0, 1)$ m, $\Delta x = 0.15$ m, $N = 200$).

5.1. Pre-Echos

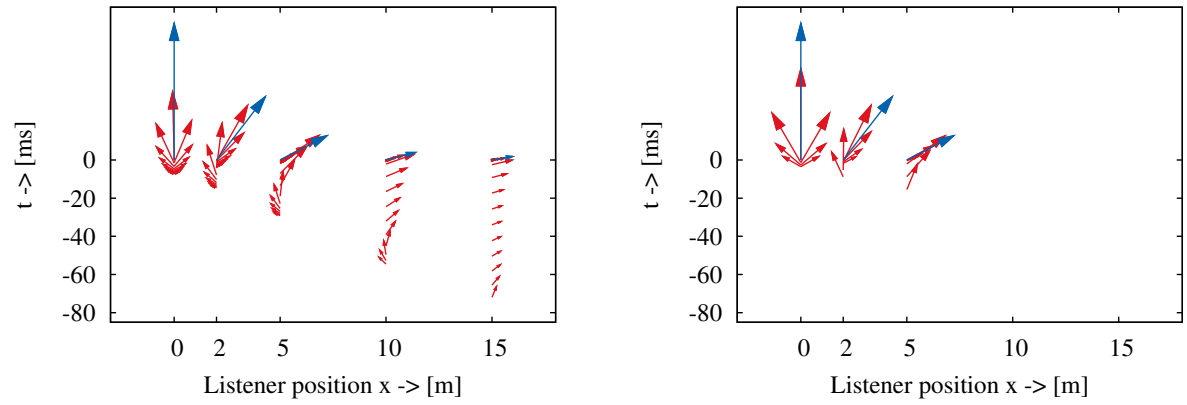
Due to the time-reversal nature of focused sources, spatial aliasing can lead to very serious perception artifacts as it produces pre-echos (as shown in figure 9 and 10). As a spatial aliasing artifact these pre-echos will only exist for frequencies above the aliasing frequency. Pre-echos do not occur naturally, so they may be perceived as a sort of distortion. Furthermore, they arrive from different directions than the focused source and therefore can have an influence on its spatial perception. The pre-echos are shown in figure 9(c) and 9(d) for a broadband impulse. These figures can be interpreted as two temporal snapshots of the spatio-temporal impulse response of a focused source. The reproduced wave field for arbitrary source signals is given by temporal convolution. As a result, the pre-echos will be more dominant for transient signals than for stationary ones when using real-world signals like speech or music. Because stationary signals can mask pre-echos and have no steep flanks in their time signal, which could be smeared by pre-echos.

The influence of echos on the spatial perception of a source has been analyzed for a long time. One of the major findings was the precedence effect [22, 23, 24].



(a) $L = 30$ m (b) $L = 10$ m

Fig. 13: Length of pre-echo artifacts for a focused source at position $\mathbf{x}_s = (0, 1)$ m, $\Delta x = 0.15$ m, $N = 200$, 2.5D WFS).



(a) $L = 30$ m loudspeaker array (b) $L = 10$ m loudspeaker array

Fig. 14: Pre-echos (red) and the focused source (blue) for different listener positions $\mathbf{x}$ (see Fig. 12) depending on time. On the left side for the 30 m and on the right side for the 10 m loudspeaker array. The length of the vectors are proportional to the amplitudes of the arriving sounds and the incidence angles of the vectors are the same as the ones for the sounds.

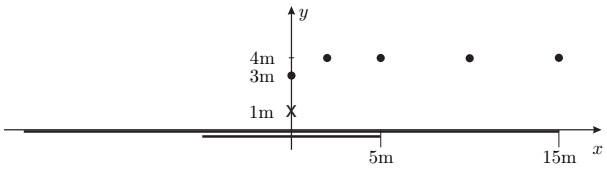


Fig. 12: Setup used for the experiments. x denotes the position of the focused source $\mathbf{x}_\mathrm{s} = (0, 1)$ m, • marks the listener positions (L=10/30 m, $\Delta x = 0.15$ m).

The precedence effect describes the phenomenon, that the direction of a perceived sound is not altered by echos of this sound, arriving from different directions and occurring in a time window of 1–40 ms after the leading wave front. In the case of focused sources the possibility hence exists, that the perceived direction of the focused source is determined by the direction of the first pre-echo. On the other hand, the precedence effect only occurs if the level of the repetition occurring after the leading wave front is not higher than 10–15 dB. So if the amplitude of the wave front from the focused source is much higher than the amplitudes of the pre-echos, the focused source will be perceived from the right location. Furthermore this can lead to the perception of a second source, if enough pre-echos arrive from another dominant direction than the wave front of the focused source.

For further investigation on this topic, a simulation of the amplitude and incidence angles of the pre-echos has been performed for two arrays with a total length of $L = 30$ m and $L = 10$ m. A focused source was positioned at $\mathbf{x}_\mathrm{s} = (0, 1)$ m and listeners were placed at $\mathbf{x} = (0, 3)$, $(2, 4)$, $(5, 4)$ m for both arrays and in addition at $\mathbf{x} = (10, 4)$, $(15, 4)$ m for the larger array as shown in figure 12. In order to evaluate the results, an informal listening test has been carried out as discussed in section 5.3 and 5.4. In figure 10, the impulse response of a broadband focused source at the listener position $\mathbf{x} = (15, 4)$ m is shown. There are clearly visible pre-echos starting at a time of $t = -72$ ms. Due to the geometry of the given positions, this is the position with the longest pre-echo time and the smallest amplitude differences between pre-echos and focused source signal, as can be concluded from figure 13 and 14. The amplitude of the first pre-echos is more than 20 dB lower than the amplitude of the focused source signal, which implies that the perceived direction of the focused source itself can only be influenced by the last, more intense pre-echos, arriving directly before the wave front of the focused source. The incidence angles of the last pre-echos, however, are nearly the same as the incidence angle of the focused source. Therefore, the perceived direction of the focused source is not influenced by the pre-echos. The incidence angles of the first pre-echos can be quite different from the one of the focused source depending on the listener position. For pre-echos at the listener position $\mathbf{x} = (5, 4)$ m or $\mathbf{x} = (2, 4)$ m, the direction of the pre-echos changes over time from nearly the opposite direction to the direction of the focused source (see figure 14). So in these cases it is liable to assume, that there will be more than one or one very wide source audible. Due to the occurrence of the pre-echos only for frequencies above the aliasing frequency additional sources will be perceived as a highpass filtered version of the focused source.

At the listener position $\mathbf{x} = (0, 3)$ m, the pre-echos arrive from different directions symmetrically, so there will be an effect like phantom sources [25] in stereo reproduction and the focused source and the pre-echos will be perceived as arriving from the front of the listener. On the other hand for the positions where no direction errors occur, there will be some sort of distortion audible due to the pre-echos arriving from the same direction as the focused source signal.

The smaller array ($L = 10$ m) is faced by the same sort of artifacts, but the pre-echo time is much smaller. For this reason it can be assumed that the distortion has a smaller impact on the perception of the focused source. But changes of the incidence angle of pre-echos for the position at the end of the array take place in a shorter time frame than for the larger array. Thus in the case of the smaller array, there may be direction errors audible for all positions with $x \neq 0$.

5.2. Coloration

As mentioned in section 4.6, the spatial aliasing frequency for focused sources depends heavily on the position of the listener. Nonetheless fixed pre-equalization filters are used in practical implementations, since WFS is designed inherently as multi-listener system.

The spatial aliasing frequency in the vicinity of the focused source is e. g. more than four times higher than the one at the end of the array, as can be concluded from figure 11. In this case the pre-equalization will yield a bass boost for listener positions near the focused source signal, because only the low frequencies are equalized and the frequency spectrum after this region is descending until spatial aliasing occurs. Therefore a coloration of the focused source depending on the distance of the listener may be observable. This problem is especially prominent for moving focused sources or listeners. The listener position dependent abrupt change from an (almost) spatial aliasing free zone to a zone with spatial aliasing artifacts will result in severe coloration in conjunction with the fixed pre-equalization.

5.3. Experimental Setup

The two setups used (see figure 12) for the discussion of the perceptual properties and the subjective experiments account for the fact that the pre-echo artifacts are prominent for large WFS systems. For a systematic evaluation of their properties, an technical approach has been chosen where the arrays have been recreated virtually for the given listener positions using dynamic binaural resynthesis [26]. This is due to the lack of a real loudspeaker array with hundreds of high-quality loudspeakers placed in an environment with adequate acoustical properties. On the other hand, the virtual recreation allows the seamless switching between different listener positions.

For this purpose a set of binaural room impulse responses (BRIRs) were generated that captured the acoustics of the focused source as reproduced by the WFS systems for the different listener positions. The BRIRs are computed by summing up the individual contributions of the loudspeakers, including the driving function and the pre-equalization filter. An aliasing frequency of 1 kHz was assumed for the pre-equalization. The acoustic path from the loudspeaker to the listener was modeled by taking the respective head-related impulse responses (HRIR) for the angle under which the user sees the loudspeaker for a given head-rotation. The HRIRs were measured with the FABIAN mannequin [27] under free-field conditions, interpolated and weighted according to the loudspeaker distance. The angular resolution of the BRIR dataset was one degree. Headphone compensation filters were applied to the final dataset.

Sets of BRIRs were computed for the five positions shown in figure 12 using the array of $L = 30$ m length. For the $L = 10$ m array, only the three positions $\mathbf{x} = (0, 3)$, $(2, 4)$ and $(5, 4)$ m were used, because the other two are outside of the listening area in this case. The actual sound reproduction was realized with the SoundScape Renderer [28] running in binaural room scanning (BRS) mode using the pre-computed BRIR sets. STAX SR Lambda Professional headphones in combination with a Polhemus Fastrak head-tracker were used for the experiment.

Three different broadband recordings with a sampling rate of 44.1 kHz were used as source material: female speech, a short melody played on a cello and a short rhythmic pattern played with castanets. For each of the five positions, a reference stimulus was created using the respective HRIR from the focused source position to the listeners ears. Also here BRIRs sets were computed. The loudness levels of reference and test stimuli were adjusted subjectively to match, which was especially critical in the presence of massive artifacts.

The $5 + 3$ pairs of reference and test stimulus were presented to a panel of 5 expert listeners which could freely switch position, array length, source material, test stimulus and reference.

The test subjects were asked to freely describe the perceived differences between each focused source stimulus and the corresponding reference. They were asked to pay attention if there are differences in coloration and spatial impression, and if they hear additional artifacts apart from the original sound.

5.4. Results

The analysis of the statements from the test subjects revealed the following perceptual properties:

There are audible artifacts as soon as the listener moves away from the central position ($x = 0$ m). These artifacts are strongly dependent on and correlated with the input signal. They were perceived as a high-pass filtered and sometimes distorted version of the original source signal. Audible distortions were comb filtering, smearing of transients, even chirping and whistling sounds (especially with castanets as source signal). The artifacts are in many cases perceived as arriving from different directions than the desired focused source, thereby for the majority of

test subjects acting as a separately perceived source and for a few subjects as a contribution to the room impression.
Head movement plays an important role in the perception of the artifacts as sometimes their position and their intensity is changing depending on the head orientation. This shows that a head-tracked dynamic resynthesis, as used here, is quite important to analyze the effects. The results of the short array and the long array are very similar for their common listener positions except the perceived positions of the artifacts are a little narrower on the short array.
At listener position $\mathbf{x} = (0, 3)$ m most subjects perceive a slight low-pass which is audible more clearly on the castanets sound. This is most likely due to the pre-equalization filter mismatch discussed in section 5.2. No further artifacts were perceived. Some subjects did not hear a difference in coloration between reference and virtual WFS stimulus.
At position $\mathbf{x} = (2, 4)$ m, there are already massive artifacts arising which are described as annoying by all subjects. Most subjects perceive the artifacts as arriving from the left at about the same angle as the original sound arrives from on the right side. Two subjects perceive an additional negative elevation. The sound of the artifacts is described as metallic and chirping for all source material. The artifacts of the speech signal additionally generate the impression of a whispering, smoky, high-pass filtered voice. The ones of the cello sound raspy.
At position $\mathbf{x} = (5, 4)$ m, some subjects reported artifacts from the left, others a continuous distribution in the frontal area. The artifacts sound like a comb filter and like chirping (especially the castanets). The artifacts caused by the voice input signal sound breathy and hoarse, the ones caused by the cello sound smeared, whistling and squeaky.
At position $\mathbf{x} = (10, 4)$ m the perceived characteristics of the artifacts are similar: squeaking, chirping, comb filtered. The voice artifact sounds breathy and whispering. Some subjects perceive the artifacts as arriving from the left, others hear them from the left *and* the right (together with the original signal). The remaining subjects perceive the artifacts as distributed about the whole frontal area.
At position $\mathbf{x} = (15, 4)$ m all subjects agree that the artifacts arrive from the same direction as the original sound. The clicking noise of the castanets is hardly audible anymore because it is nearly completely masked by hideously smeared artifacts. The voice signal obtains a vocoder-like quality.

5.5. Discussion

The results of the informal listening test confirm the theoretical considerations on the perception of artifacts caused by the spatial sampling artifacts of focused sources. The pre-echos are clearly audible as some sort of high-pass distortions. Also the different incidence angles of the pre-echos can lead to the perception of multiple sources. But the differences in perception between the $L = 10$ m loudspeaker array and the $L = 30$ m array are smaller than predicted. One reason for this could be the larger amplitude of late pre-echos (which are present for both arrays) in contrast to the smaller amplitude for very early pre-echos, which are only present for the larger array. In further tests the different types of perceived distortions for focused sources should be investigated in more detail.

6. CONCLUSIONS

This paper presents a detailed analysis of the physical and perceptual properties of focused sources as used in the context of WFS. Focused sources exhibit a number of remarkable physical properties. The sampling artifacts, inherent to massive multichannel sound reproduction, almost disappear for positions close to the focused sources. The amplitude decay of a focused source is reproduced correctly in a 2.5D scenario, furthermore. Besides these positive properties also a number of potential artifacts exist. For broadband signals, additional wave fronts arriving before the first wave front of the desired focused source may be present. This spatial sampling artifact is caused by the spatial undersampling of the secondary sources in typical implementations. These pre-echos may further exhibit different incidence angles than the first wave front arriving from the desired focused source. The analysis of the evanescent contributions implies that the near-field contributions of the focused source are not reproduced correctly. However, further analysis is required here. The pre-echo artifacts of focused sources have major impact on their perception as has been concluded from the literature on the precedence effect. These findings have been confirmed by the subjects in the informal listening test. The listening test unveiled

a number of interesting properties in the context of perception of pre-echos. Depending on the listener position spatial aliasing of focused sources can lead to the perception of different types of distortion and multiple sources.

The analysis, as presented in this paper, shows that focused sources exhibit a number of artifacts that are not desirable for high-quality reproduction of sound. Focused sources are a key feature of massive multichannel sound reproduction systems. They allow new types of auditory scenes, e. g. containing sources placed between listeners. Therefore, an improvement of the found artifacts is very desirable. Note, that the standard driving function has been used in this paper. The authors are not aware of published extensions that improved the fidelity of focused sources. One way to overcome the artifacts of focused sources could be not to use all loudspeakers for a particular focused source position. However, this would also limit the listening area which is not desirable in all situations.

In the future more formal subjective experiments will be carried out in order to investigate the perceptual properties of focused sources and to understand the psychoacoustic mechanisms. This could open up the potential to improve their artifacts in the future.

7. REFERENCES

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

[2] J. Ahrens and S. Spors. Focusing of virtual sound sources in higher order ambisonics. In *124th AES Convention*, Amsterdam, The Netherlands, May 2008. Audio Engineering Society (AES).

[3] S. Yon, M. Tanter, and M. Fink. Sound focusing in rooms: the time-reversal approach. *Journal of the Acoustical Society of America*, 113(3):1533–1543, March 2003.

[4] S. Spors and J. Ahrens. Spatial sampling artifacts of focused sources in wave field synthesis. In *NAG–DAGA International Conference on Acoustics*, Rotterdam, The Netherlands, March 2009.

[5] A.J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *Journal of the Acoustic Society of America*, 93(5):2764–2778, May 1993.

[6] S. Spors, R. Rabenstein, and J. Ahrens. The theory of wave field synthesis revisited. In *124th AES Convention*, Amsterdam, The Netherlands, May 2008. Audio Engineering Society (AES).

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

[8] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, 1972.

[9] J.-J. Sonke, D. de Vries, and J. Labeeuw. Variable acoustics by wave field synthesis: A closer look at amplitude effects. In *104th AES Convention*, Amsterdam, Netherlands, May 1998. Audio Engineering Society (AES).

[10] J. Ahrens and S. Spors. Sound field reproduction using planar and linear arrays of loudspeakers. *IEEE Transactions on Audio, Speech and Signal Processing*, 2009. Submitted.

[11] M. Fink, G. Montaldo, and M. Tanter. Time reversal acoustics. In *IEEE Ultrasonics Symposium*, pages 850–859, 2004.

[12] M. Fink. Time-reversed acoustics. *Scientific American*, pages 91–97, Nov. 1999.

[13] S. Yon, M. Tanter, and M. Fink. Sound focusing in rooms. II. The spatio-temporal inverse filter. *Journal of the Acoustic Society of America*, 114(6):3044–3052, Dec. 2003.

[14] M. Fink. Time reversal of ultrasonic fields – Part I: Basic principles. *IEEE Transactions on Ultrasocics, Ferroelectrics, and Frequency Control*, 39(5):555–566, Sept. 1992.

[15] D. de Vries and A.J. Berkhout. Wave theoretical approach to acoustic focusing. *Journal of the Acoustic Society of America*, 70, Sept. 1981.

[16] S. Spors. Extension of an analytic secondary source selection criterion for wave field synthesis. In *123th AES Convention*, New York, USA, October 2007. Audio Engineering Society (AES).

[17] I.S. Gradshteyn and I.M. Ryzhik. *Tables of Integrals, Series, and Products*. Academic Press, 2000.

[18] S. Spors and J. Ahrens. Spatial aliasing artifacts of wave field synthesis for the reproduction of virtual point sources. In *126th AES Convention*, Munich, Germany, May 2009. Audio Engineering Society (AES).

[19] B.G. Lucas and T.G. Muir. The field of a focusing source. *Journal of the Acoustic Society of America*, 72, Oct. 1982.

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

[21] D. de Vries, E.W. Start, and V.G. Valstar. The Wave Field Synthesis concept applied to sound reinforcement: Restrictions and solutions. In *96th AES Convention*, Amsterdam, Netherlands, February 1994. Audio Engineering Society (AES).

[22] Hans Wallach, Edwin B. Newman, and Mark R. Rosenzweig. The precedence effect in sound localization. *American Journal of Psychology*, 57:315–336, 1949.

[23] Wolfgang Haas. The influence of a single echo on the audibility of speech. *Acustica*, 1:49–58, 1951.

[24] J. Blauert and J. Braasch. Acoustic communication: The precendence effect. In *Forum Acousticum*, 2005.

[25] J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, 1996.

[26] A. Lindau, T. Hohn, and S. Weinzierl. Binaural resynthesis for comparative studies of acoustical environments. In *122th AES Convention*, Vienna, Austria, 2007. Audio Engineering Society (AES).

[27] A. Lindau and S. Weinzierl. FABIAN - An instrument for software-based measurement of binaural room impulse responses in multiple degrees of freedom. In *Proc. of the Tonmeistertagung*, Nov. 2008.

[28] M. Geier, J. Ahrens, and S. Spors. The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods. In *124th AES Convention*, Amsterdam, The Netherlands, May 2008. Audio Engineering Society (AES).