Audio Engineering Society

Convention Paper

Presented at the 124th Convention
2008 May 17–20 Amsterdam, The Netherlands

Reproduction of Moving Virtual Sound Sources with Special Attention to the Doppler Effect

Jens Ahrens and Sascha Spors

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

Correspondence should be addressed to Jens Ahrens (jens.ahrens@telekom.de)

ABSTRACT

In this paper, we outline a basic framework for the reproduction of the wave field of moving virtual sound sources. Conventional implementations usually reproduce moving virtual sources as a sequence of stationary positions. This process leads to various artifacts as reported in the literature. On the example of wave field synthesis, we show that the explicit consideration of the physical properties of the wave field of moving sources avoids these artifacts and allows for the accurate reproduction of the Doppler Effect. However, numerical simulations suggest that the artifacts inherent to the reproduction system can lead to a heavy degradation of the reproduction quality.

1. INTRODUCTION

Since several decades, the problem of physically recreating a given wave field has been addressed in the audio community. Independent of the chosen approach, two rendering techniques exist: Data based and model based reproduction [1]. The former case aims at perfectly reproducing a captured sound field. This situation will not be treated in this paper. We concentrate on the latter case where a sound scene is composed of a number of virtual sound sources derived from analytical spatial source models. For stationary virtual scenes accurate reproduction techniques exist. However, the reproduction of dynamic scenes implicates certain peculiarities. This is mostly due to the fact that the speed of sound in air is constant. When a source moves, the propagation speed of the emitted wave field is not affected. However, the emitted wave field differs from that of a static source in various ways. For example, sound waves emitted in the direction of motion experience an increase in frequency, sound waves emitted in opposite direction of motion expe-

rience a decrease in frequency. The whole of these alterations is known as Doppler Effect [2].
Typical implementations of sound field reproduction systems do not take the Doppler Effect into account. Dynamic virtual sound scenes are rather reproduced as a sequence of stationary snapshots. Thus, not only the virtual source also its entire wave field is moved from one time instant to the next.
This concatenation leads to Doppler-like frequency shifts. However, these frequency shifts occur due to warping of the time axis rather than due to the constant speed of sound, a circumstance which introduces artifacts. The artifacts have been recently discussed in the literature in the context of wave field synthesis [3]. We are not aware of an according publication focussing on alternative sound field reproduction methods. See [4] for a treatment of moving virtual sources in binaural (HRTF-based) reproduction. Note that the considerations presented in the present paper are of relevance for sound field reproduction approaches which employ time delays in the procedure of yielding the loudspeaker driving signals.
In this paper, we show that the mentioned artifacts introduced by the process of concatenating stationary virtual source positions can be avoided when the physical properties of the wave field of moving sound sources are a priori taken into account. We derive a mathematical expression for the wave field of a uniformly moving monopole source. On the example of wave field synthesis, we show how such a system can be driven accordingly and discuss the properties of the actual reproduced wave field by means of numerical simulations.

2. WAVE FIELD OF A MOVING SOUND SOURCE

The fundamental prerequisite for model-based sound field reproduction is to have a description of the sound field that is to be recreated. In this section, we derive the wave field emitted by a moving sound source. The presented approach is adopted from [5, 6] and can be regarded as a generalization of the approach described in [4]. The latter is restricted to sound sources moving at speeds slower than the speed of sound. For simplicity, we assume a monopole sound source. However, the presented approach also allows for the treatment of arbitrary source types.

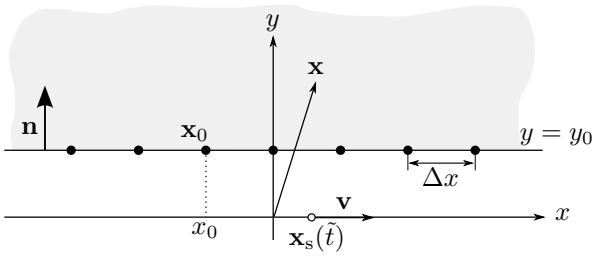


Fig. 1: The coordinate system and geometry used in this paper. The dots • denote the positions of the secondary sources used for wave field synthesis. The grey-shaded area denotes the listening area.

The time-domain free-field Green's function of a stationary sound source at position $\mathbf{x}_\mathrm{s}$, i.e. its spatio-temporal impulse response, is denoted by $g\left(\mathbf{x}-\mathbf{x}_\mathrm{s}, t\right)$. The time-domain Green's function of a moving sound source is then $g\left(\mathbf{x}-\mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t)), t-\tilde{t}(\mathbf{x},t)\right)$, whereby $\tilde{t}(\mathbf{x},t)$ denotes the time instant when the impulse was emitted. Confer to figure 2. $g\left(\mathbf{x}-\mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t)), t-\tilde{t}(\mathbf{x},t)\right)$ is referred to as *retarded Green's function* [5]. $\tilde{t}(\mathbf{x},t)$ is dependent on the location of the receiver $\mathbf{x}$ and the time t that the receiver experiences.
Assume a monochromatic harmonic source oscillating at angular frequency ω_s. Its source function $s_0(\tilde{t})$ reads in complex notation

$$s_0(\tilde{t}) = a_0 \cdot e^{j\omega_s \tilde{t}} \; . \tag{1}$$

In order to yield the wave field produced by a moving source with spatio-temporal impulse response $g\left(\mathbf{x}-\mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t)), t-\tilde{t}(\mathbf{x},t)\right)$ driven by the signal $s_0(\tilde{t})$, we model $s_0(\tilde{t})$ as a dense sequence of Dirac pulses. Each Dirac pulse of the sequence multiplied by $g\left(\mathbf{x}-\mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t)), t-\tilde{t}(\mathbf{x},t)\right)$ yields the wave field created by the respective Dirac pulse. To yield the wave field emitted due to the entire sequence of Dirac pulses, we integrate over $\tilde{t}$ as

$$s(\mathbf{x},t) = \int_{-\infty}^{\infty} s_0(\tilde{t}) \cdot g\left(\mathbf{x}-\mathbf{x}_\mathrm{s}(\tilde{t}), t-\tilde{t}\right) \; d\tilde{t} \; , \tag{2}$$

whereby we temporarily altered the nomenclature for convenience ($\tilde{t} = \tilde{t}(\mathbf{x},t)$).
Assuming a moving monopole sound source, its

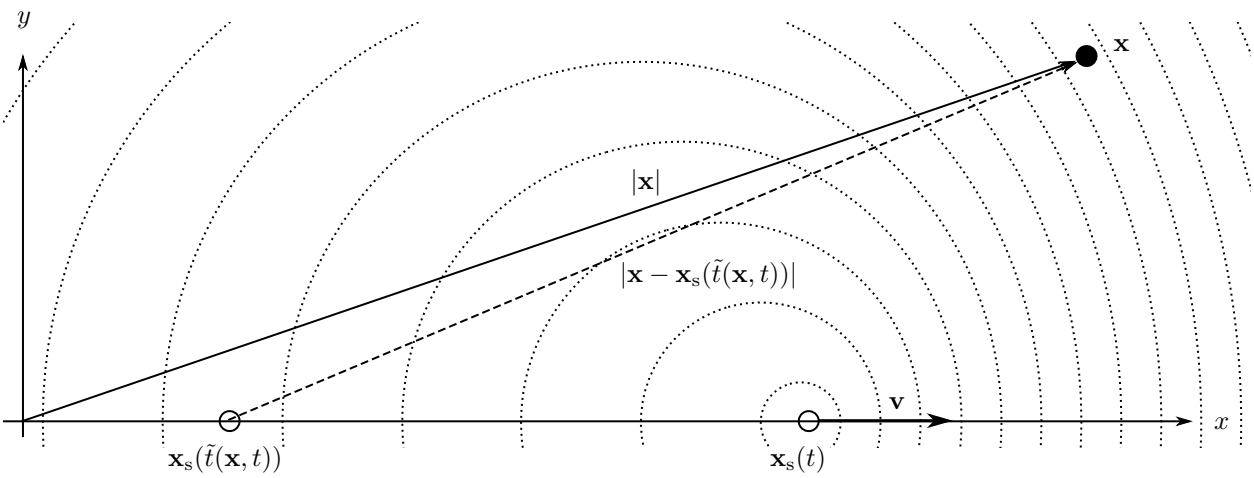


Fig. 2: Derivation of the Green's function of a moving sound source.

Green's function explicitly reads

$$g\left(\mathbf{x} - \mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t)), t - \tilde{t}(\mathbf{x},t)\right) = \\ = \frac{1}{4\pi} \frac{\delta\left(t - \tilde{t}(\mathbf{x},t) - \frac{|\mathbf{x}-\mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t))|}{c}\right)}{|\mathbf{x} - \mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t))|} \; . \quad (3)$$

Note that

$$\tau(\mathbf{x},t) = \frac{|\mathbf{x} - \mathbf{x}_\mathrm{s}(\tilde{t}(\mathbf{x},t))|}{c} \quad (4)$$

is referred to as *retarded time* [5]. It denotes the duration of sound propagation from the source to the receiver. In the remainder of this paper, $M = \frac{v}{c}$ denotes the *Mach number*.

For convenience, we assume the virtual source to move uniformly along the x-axis in positive x-direction (cf. also to section 5 and figure 2). At time $\tilde{t}_0 = 0$ the source is located at position $x_s(0)$. For this particular source trajectory, the integral in equation (2) can be solved via the substitution

$$u = \tilde{t}(\mathbf{x},t) + \tau(\mathbf{x},t) \quad (5)$$

and the exploitation of the sifting property of the delta function [7]. The solution is then valid for all points in space except for the instantaneous position of the sound source where it exhibits a pole. It turns out that the integral has different solutions for $M < 1$, $M = 1$, and $M > 1$. In the remainder of this paper, we exclusively treat the case of $M < 1$, i.e. virtual sources moving at subsonic speeds. Note however that the presented approach also allows for the treatment of supersonic speeds.

For $M < 1$, the integral boundaries can be kept and the solution, i.e. the sound field $s(\mathbf{x},t)$ of a source moving at a speed $v < c$ reads then

$$s(\mathbf{x},t) = \frac{1}{4\pi} \cdot \frac{s_0(\tilde{t}(\mathbf{x},t))}{\Psi(\mathbf{x},t)} \; , \quad (6)$$

whereby

$$\tilde{t}(\mathbf{x},t) = t - \frac{M\Phi(x,t) + \Psi(\mathbf{x},t)}{c(1-M^2)} \; , \\ \Psi(\mathbf{x},t) = \sqrt{\Phi^2(x,t) + y^2} \; , \\ \Phi(x,t) = x - vt - x_\mathrm{s}(0) \; .$$

A snapshot of the wave field of a moving sound source described by equation (6) is depicted in figure 3(a).

For $M = 0$, i.e. a static source, equation (6) reads

$$s_{M=0}(\mathbf{x},t) = \frac{1}{4\pi} \cdot \frac{s_0(t-\tau)}{|\mathbf{x} - \mathbf{x}_\mathrm{s}|} \quad (7)$$

which corresponds to the familiar expression for the sound field of a static harmonic monopole sound source (confer to equation (10)).

3. WAVE FIELD SYNTHESIS

In this section, we demonstrate how a moving virtual sound source can be reproduced using the find-

ings derived in section 2. Exemplarily, we use wave-field synthesis (WFS) employing a linear array of secondary sources (loudspeakers).

The theoretical basis of WFS employing linear secondary source arrays is given by the two-dimensional Rayleigh I integral [8]. It states that a linear distribution of monopole line sources is capable of reproducing a desired wave field (a virtual source) in one of the half planes defined by the secondary source distribution. The wave field in the other half (where the virtual source is situated) is a mirrored copy of the desired wave field. For convenience, the secondary source array is assumed to be parallel to the x-axis at $y = y_0$ as depicted in figures 1 and 3(b). The listening area is chosen to be at $y > y_0$.

The two-dimensional Rayleigh I integral determines the sound pressure $p_{\mathrm{WFS}}(\mathbf{x}, t)$ created by such a setup reading

$$p_{\mathrm{WFS}}(\mathbf{x}, t) = \int_{-\infty}^{\infty} \underbrace{-\frac{\partial}{\partial \mathbf{n}} s(\mathbf{x}, t)|_{\mathbf{x}=\mathbf{x}_0}}_{d(\mathbf{x}_0, t)} *_t g(\mathbf{x}, t)\; dx_0 \; . \tag{8}$$

$s(\mathbf{x}, t)$ denotes the sound field of the virtual source and $\frac{\partial}{\partial \mathbf{n}}$ the gradient in the direction normal to the secondary source distribution (confer also to figure 1). The asterisk $*_t$ denotes convolution with respect to time.

The driving function $d(\mathbf{x}_0, t)$ for a loudspeaker at position $\mathbf{x}_0$ is thus yielded by evaluating the gradient of the desired virtual sound field in direction normal to the loudspeaker distribution at the position of the respective loudspeaker.

Note however that the virtual source's wave field is not perfectly reproduced in the receiver's half-space. This is due to the fact that the physical requirements can not be perfectly fulfilled in practical implementations. Equation (8) requires an infinitely long continuous distribution of secondary sources, practical implementations can only employ a finite number of discrete loudspeakers. The array has thus a finite length. Furthermore, equation (8) requires secondary line sources which are positioned perpendicular to the receiver plane. Practical implementations typically employ loudspeakers with closed cabinets as secondary sources. These are more accurately described by point sources rather than line sources. This fact is known as secondary source mismatch and has to be compensated for as

$$d_{\mathrm{corr}}(\mathbf{x}, t) = f(t) *_t d(\mathbf{x}, t) \; . \tag{9}$$

$f(t)$ is a filter with frequency response $F(\omega) = 2\sqrt{2\pi j k d_{\mathrm{ref}}}$, the asterisk $*_t$ denotes convolution with respect to time, and d_{ref} denotes the reference distance from the secondary source array, to which the amplitude of the reproduced wave field is referenced. See [9] for a thorough treatment of the properties of WFS.

For convenience, we do not explicitly compensate for the secondary source mismatch in the expressions for the driving functions. However, in the simulations this compensation is performed.

3.1. Conventional driving functions

In this section, we briefly outline the two most common conventional approaches of implementing moving sound sources in WFS. What both approaches have in common is the fact that they do not explicitly consider the physical properties of the wave field of a moving source. The source motion is rather modeled as a sequence of stationary positions. Thus, not only the virtual source moves but also its entire wave field. The difference between the approaches is the consideration of the retarded time τ defined by equation (4) as discussed below.

The starting point of all implementations is the wave field of a monochromatic static monopole source reading [10]

$$s(\mathbf{x}, t) = \frac{1}{4\pi} \frac{s_0\left(t - \frac{|\mathbf{x}-\mathbf{x}_{\mathrm{s}}|}{c}\right)}{|\mathbf{x}-\mathbf{x}_{\mathrm{s}}|} = \frac{1}{4\pi} \frac{s_0\left(t-\tau\right)}{|\mathbf{x}-\mathbf{x}_{\mathrm{s}}|} \; . \tag{10}$$

Calculation the driving function as outlined in section 3 yields

$$d_\tau(\mathbf{x}, t) = \frac{y}{|\mathbf{x}-\mathbf{x}_{\mathrm{s}}|} \left(\frac{1}{|\mathbf{x}-\mathbf{x}_{\mathrm{s}}|} + \frac{j\omega_{\mathrm{s}}}{c} \right) \cdot s(\mathbf{x}, t) \; . \tag{11}$$

In equation (11), the retarded time is implicitly correctly considered when static conditions are assumed. This implies that the driving signal for a given secondary source is yielded by a frequency dependent scaling of the source signal and by delaying the source input signal by $t_{\mathrm{delay}} = \tau$. Note that t_{delay} features the retarded time τ as experienced by the secondary source under consideration. For virtual sources located far away from the secondary source

array this implies relatively long delays.
The alternative to the approach given by equation (11) is to shorten t_{delay} by the travel time $\tau_{\min}$ of the sound wave from the virtual source to the closest secondary source, i.e.

$$d_{\tau_{\text{short}}}(\mathbf{x}, t) = d_{\tau}(\mathbf{x}, t + \tau_{\min}) \; . \tag{12}$$

This procedure maintains the relative delays between the secondary sources and therefore ensures the correct curvature of the reproduced sound wave. It saves processing resources since the resulting delays to be implemented are generally significantly shorter than in equation (11).
As far as we are aware, equation (12) is the most common implementation approach.

3.2. Proposed approach

Contrary to the common approach, we propose to derive the secondary source driving function from the wave field of the moving source. For a virtual harmonic monopole sound source of angular frequency ω_s moving uniformly along the x-axis as described in section 2, the driving function $d(\mathbf{x}, t)$ reads

$$d(\mathbf{x}, t) = \frac{y}{\Psi(\mathbf{x}, t)} \left(\frac{1}{\Psi(\mathbf{x}, t)} + \frac{j\omega_s}{c(1 - M^2)} \right) \times \\ \times \, s(\mathbf{x}, \tilde{t}(\mathbf{x}, t)) \; . \tag{13}$$

Note that $d(\mathbf{x}, t)$ in equation (13) implicitly includes static virtual sources.
The wave field reproduced by a linear WFS array driven by equation (13) is depicted in figure 3(b). The overall length of the loudspeaker array is 8 m. The virtual source moves at a speed $v = 120 \frac{\text{m}}{\text{s}}$ along the x-axis in positive x-direction ($M \approx \frac{1}{3}$).

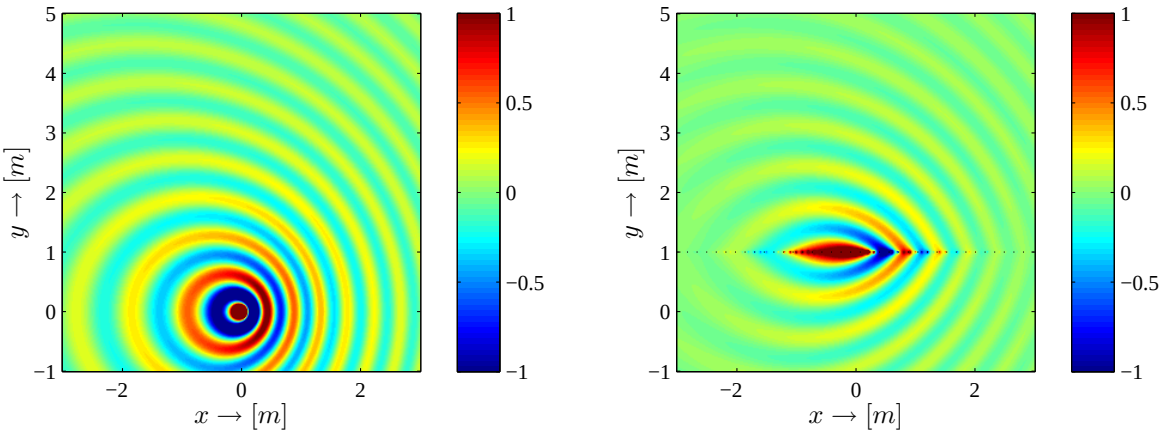
4. RESULTS

In this section, we present a number of simulations in order to analyze the properties of the different implementation approaches. We assume a linear array of secondary monopole sources. The secondary sources are placed at an interval of $\Delta x = 0.1$ m throughout the simulations. The loudspeaker array is situated parallel to the x-axis and symmetrically around the y-axis at $y_0 = 1$ m except where stated explicitly.
As inherent to WFS, the reproduced wave field only approximates the desired one for $y > y_0$. Due to the fact that we assume secondary monopole sources, the reproduced wave field on the other side of the loudspeaker array (where $y < y_0$) is a mirrored version.
The different parameters of the simulation where chosen such that the respective property of the reproduced wave field under consideration is strongly pronounced. The low energy artifacts apparent in the spectrograms in figures 4, 5, and 7 arise due to the properties of the short term Fourier analysis of the signal.
All simulations were carried out at a temporal sampling frequency of 44.1 kHz and the secondary source driving signals were calculated sample-by-sample. When we speak of the conventional implementation approach, we refer to $d_{\tau}(\mathbf{x}, t)$ (confer to section 3.1), except where stated explicitly.
We emphasize that all conclusions that we draw in this section are yielded from the results of numerical simulations. We have not performed analytical analyses to confirm the theses.

4.1. Artifacts reported in the literature

The artifacts arising in conventional WFS implementations of moving virtual sound sources are discussed in [3]. However, our own investigations suggest some reinterpretations. We therefore review the compilation given in [3] and comment on it considering the findings derived in this paper. Confer also to section 4.2 for further interpretations of the results.

Doppler shift deviation The Doppler shift deviation as referred to in [3] describes the discrepancy between the measured/perceived frequency shift due to the source motion in the conventional WFS implementations and the frequency shift due to the Doppler Effect. The frequency shift in the conventional WFS implementations arises as a consequence of a warping of the time axis. The frequency emitted by a moving virtual source measured by a stationary receiver when the source moves towards the receiver increases linearly with the speed of the source. It will not go to infinity as long as the source's speed does not do so. When a natural source moves towards a receiver, the measured frequency increases faster than linear with the source's speed and approaches infinity as the source's speed approaches the speed of sound.
This Doppler shift deviation is rather subtle for mod-



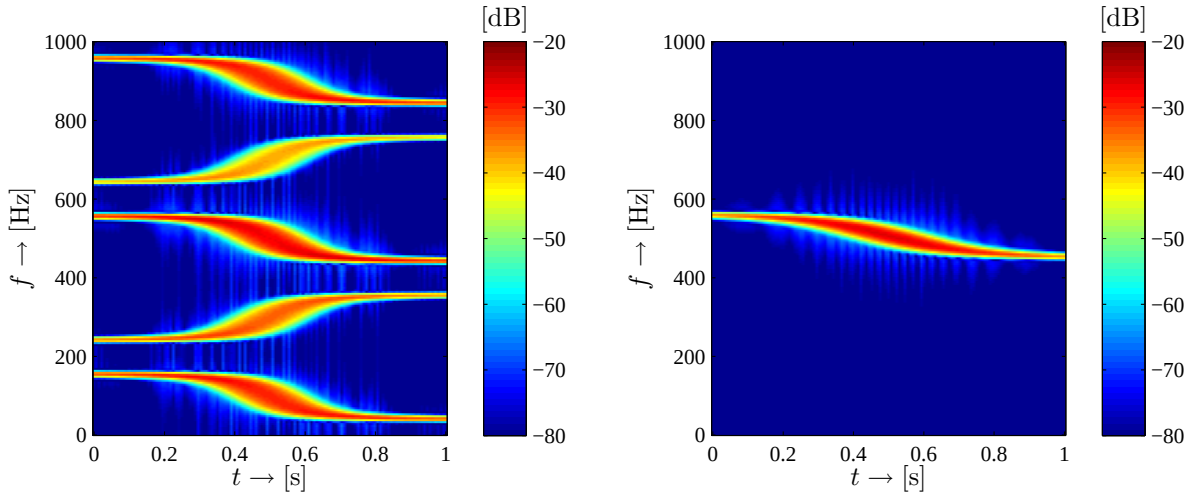
(a) $\Re\{s(\mathbf{x}, t_0)\}$

(b) $\Re\{p_{\mathrm{WFS}}(\mathbf{x}, t_0)\}$. The loudspeaker array indicated by the dotted line is situated symmetrically around the y-axis at $y_0 = 1$ m and its overall length is 8 m. The loudspeakers are positioned at intervals of $\Delta x = 0.1$ m. Tapering is applied.

Fig. 3: Simulated wave fields of a source oscillating monochromatically at $f_\mathrm{s} = 500$ Hz and moving along the x-axis in positive x-direction at $v = 120\frac{\mathrm{m}}{\mathrm{s}}$. Due to the employment of the complex notation for time domain signals (see equation (1)), only the real part $\Re\{\cdot\}$ of the considered wave field is depicted. The wave fields have been scaled to have comparable levels. The values of the sound pressure are clipped as indicated by the colorbars.

erate source speeds. It is therefore hardly visible in the spectrograms presented in this paper since spectrograms in general suffer from the trade off between time and frequency resolution. By comparing figures 4(a) and 4(b), it can be seen that the measured/perceived frequency is slightly lower in the conventional WFS implementation in figure 4(b) than with the natural source in figure 4(a) when the virtual source moves towards the receiver.
When the source approaches the receiver, the measured frequency is substantially lower in the conventional implementation for high source speeds. When the source moves away from the receiver, the discrepancy is less obvious.
See section 4.2 for comment on the Doppler shift deviations present in the proposed approach.

Spectral broadening Besides the systematic Doppler shift deviation, the second artifact in conventional implementations reported in [3] which can be exclusively attributed to a moving virtual source is a broadening of the source's temporal spectrum. This artifact is considered to be the most disturbing one. In [3] it is claimed that the spectral broadening arises due to the fact that each secondary source experiences a different time warping of its input signal. Thus, for monochromatic signals each secondary source also emits a monochromatic signal but with a frequency individual to each secondary source. The superposition of the wave fields of the different secondary sources then exhibits are broader temporal spectrum than the virtual source.
However, although this explanation sounds reasonable, our simulations do not confirm that this spectral broadening does indeed per se occur. The implementation of $d_{\tau_{\mathrm{short}}}(\mathbf{x}, t)$ (confer to section 3.1) does introduce artifacts similar to a spectral broadening. However, these artifacts turn out to be much more subtle than intuitively expected, especially for long secondary source arrays and virtual sources at far distances from the secondary source array. When $d_\tau(\mathbf{x}, t)$ is implemented, we do not find a considerable broadening at all in our simulations when the virtual source is sufficiently far away from the sec-



(a) Conventional implementation. Note that more transpositions occur at frequencies higher than depicted.

(b) Proposed approach.

Fig. 6: Spectrograms of a virtual source travelling with $v = 40$ m/s and $f_\mathrm{s} = 500$ Hz observed at $\mathbf{x}_\mathrm{mic} = [0 \quad 4]^T$ m. The loudspeaker array has an overall length of 50 m and is located at $y_0 = 0.01$ m. Values are clipped as indicated by the colorbars.

ondary sources (i.e. a few meters).

Additionally to the conventional driving functions discussed in section 3.1, the traditional WFS literature, e.g. [11], features alternative formulations. We have not investigated their properties with respect to moving virtual sources since they pose restrictions even on stationary virtual sources. The application of approximations in their derivation requires that the virtual source be positioned sufficiently far away from the secondary sources. The driving functions given in section 3.1 themselves do not per se carry along restrictions.

Our investigations revealed that severe artifacts arise in the conventional WFS approaches ($d_\tau(\mathbf{x}, t)$ and $d_{\tau_\mathrm{short}}(\mathbf{x}, t)$, section 3.1) when the virtual sound source moves at very close distances to the secondary sources (at a few centimeters distance). Confer to figure 6(a). It depicts the reproduced wave field when the virtual source moves at a distance of 1 cm behind the loudspeaker array. This circumstance is not considered in [3].

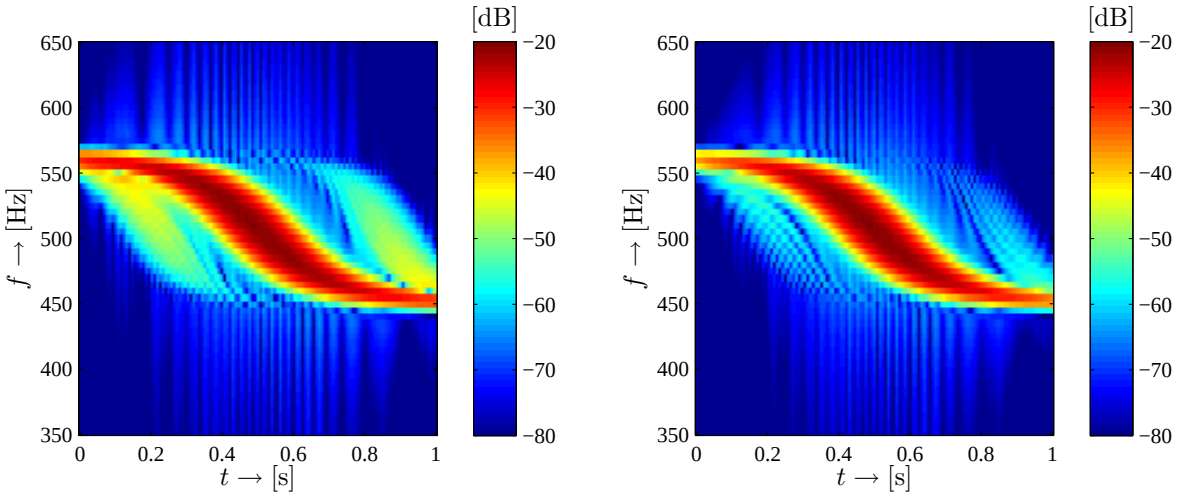
Unfortunately, the signals in the spectrograms which are used to illustrate the spectral broadening in [3] seem to suffer from prominent truncation artifacts (confer to section 4.2) which make their interpretation hard. Furthermore, the virtual source is relatively far away from the secondary sources. No further numerical evidence of the spectral broadening is given in [3].

Confer to figure 4 for our respective simulations. All spectrograms ibidem suffer from a spectral broadening at first sight. However, this seeming spectral broadening is a consequence of the applied short time analysis and should not be misinterpreted as inherent to the signal [7].

4.2. Artifacts present in the proposed approach

Inspection of the simulations presented in this paper suggests that the reproduced wave field does suffer from a number of artifacts even if the presented approach is applied. However, there are indications that these artifacts are not directly related to the source motion. It is rather such that the artifacts of which WFS inherently suffers from get a more prominent quality due to the time-variant property of the reproduced wave field [12]. This circumstance is discussed in this section.

Note that the artifacts presented in this section are also apparent in conventional implementations.



(a) No tapering applied.

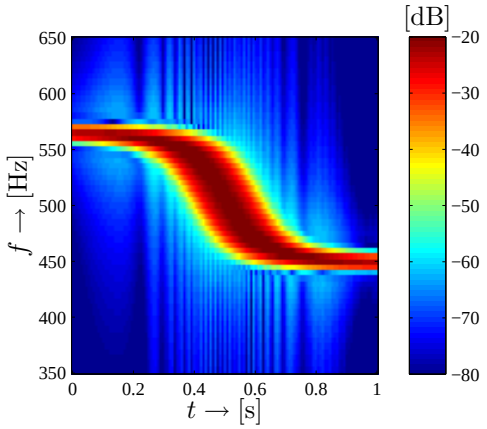
(b) Cosine tapering.

Fig. 7: Spectrograms of a virtual source travelling with $v = 40$ m/s and $f_\mathrm{s} = 500$ Hz observed at $\mathbf{x}_\mathrm{mic} = [0 \quad 4]^T$ m. The loudspeaker array has an overall length of 20 m and is located at $y_0 = 1$ m. Values are clipped as indicated by the colorbars.

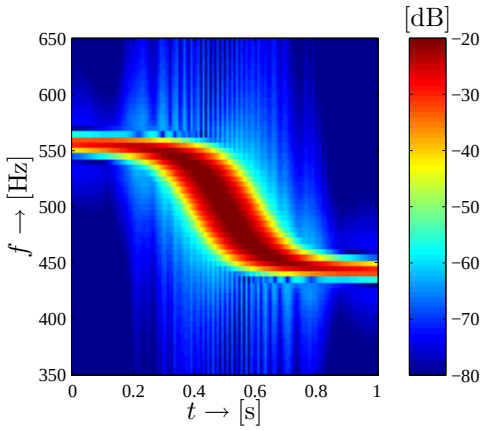
Truncation artifacts Practical implementations of WFS systems always employ secondary source arrays of finite length although the physical theory requires an infinite array. As consequence, the reproduced wave field suffers from artifacts which appear as the wave field of point sources situated at the ends of the secondary source array [11]. This is a very subtle disturbance with stationary virtual sources. However, for moving virtual sources, the truncation artifacts appear as delayed respectively anticipated echoes of the moving source (confer to figure 7(a)). When the secondary source distribution is long, the echoes are audible as such. When the secondary source distribution is only a few meters long, the echoes appear close in time to the virtual source and their combination results in strongly disturbing beats which can be perceived as spectral broadening.
In order to minimize truncation artifacts, tapering can be applied [11]. Tapering is an attenuation of the secondary source driving signal towards the ends of the array (confer to figure 7(b)). Different weighting functions can be applied having different side effects. In figure 7(b), a raised-cosine window was chosen.
Note that as a consequence of tapering, moving sources become audible later and disappear faster.

Spatial aliasing Spatial aliasing artifacts constitute a distortion of the spatial structure of the reproduced wave field [13]. When a sound source is moving, even for monochromatic sound sources, the frequency of the reproduced wave field is both time and position dependent. The spatial distortion of such a wave field introduces strongly audible artifacts. This can be observed in figure 5. The source frequency ω_s is successively doubled from figure 5(a) to figure 5(c). The higher ω_s the more transposed copies of the source arise. Their combination results in strongly audible beats and can be perceived as spectral broadening.
Note that the spatial aliasing frequency of the loudspeaker array employed in the simulations, i.e. the frequency above which spatial aliasing occurs, is $f_\mathrm{alias} \approx 1700$ Hz [13].

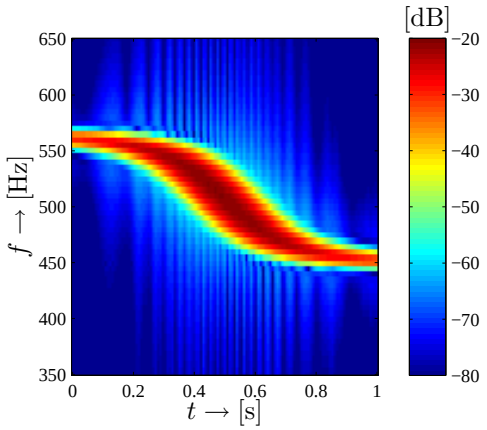
Amplitude errors Our investigations do not allow to draw conclusions about the consequences of the WFS-inherent amplitude errors [12].



(a) Sound source.

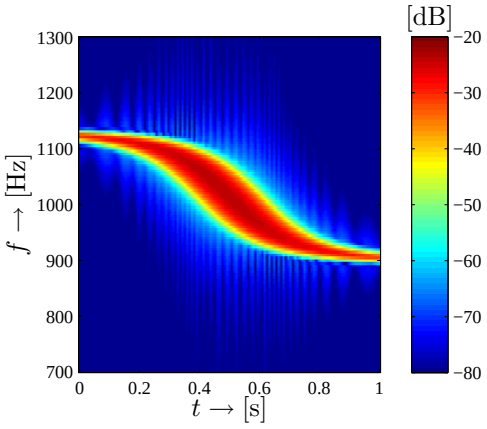


(b) WFS conventional implementation.

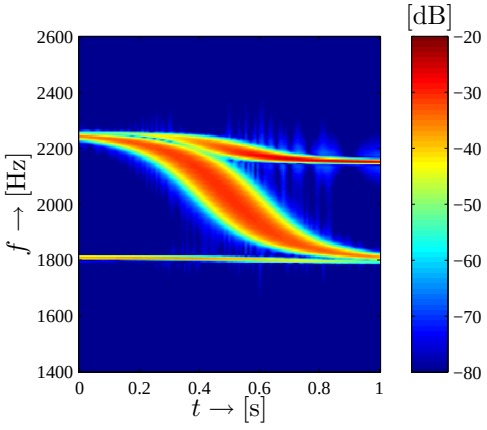


(c) Proposed approach.

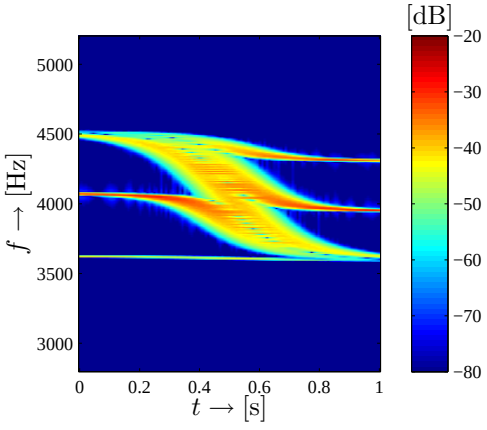
Fig. 4: Spectrograms of a virtual source travelling with $v = 40$ m/s and $f_\mathrm{s} = 500$ Hz observed at $\mathbf{x}_\mathrm{mic} = [0 \;\; 4]^T$ m. The WFS array has an overall length of 50 m and is located at $y_0 = 1$ m. Values are clipped as indicated by the colorbars.



(a) $f_s = 1000$ Hz



(b) $f_s = 2000$ Hz



(c) $f_s = 4000$ Hz

Fig. 5: Spectrograms of a virtual source travelling with $v = 40$ m/s emitting different frequencies f_s observed at $\mathbf{x}_\mathrm{mic} = [0 \;\; 4]^T$ m. The WFS array has an overall length of 50 m and is located at $y_0 = 1$ m. Values are clipped as indicated by the colorbars.

Deviation of the Doppler shift Comparison of figures 4(a) and 4(c) reveals that also the present approach suffers from a deviation of the frequency shift caused by the source motion compared to the Doppler shift. However, the extremals of the preceived/measured frequency correspond closely to those created by a natural source. This holds true even for very high source speeds. The transition of the preceived/measured frequency from one extremal to the other - when the virtual source passes the receiver - shows some deviations.
The simulations presented in this paper do not give hints on the cause for this deviation. Informal listening suggests that the deviation is hardly perceptible, even in an A-B-comparison.

Other As mentioned in section 4.1, severe artifacts arise in the conventional WFS approaches when the virtual sound source moves at very close distances to the secondary sources (at a few centimeters distance). According to our simulations, the presented approach is robust towards this. Confer to figure 6(b). It depicts the reproduced wave field when the virtual source moves at a distance of 1 cm behind the loudspeaker array.

5. ARBITRARY TRAJECTORIES AND IMPLEMENTATION

In order to keep the equations simple, we assumed in the above derivation that the virtual source moves uniformly along the x-axis. In order to enable arbitrary source trajectories, the formulation has to be modified. The instantaneous position $\mathbf{x}_s(\tilde{t}(\mathbf{x},t))$ of the sound source in equation (3) has to be replaced by

$$\mathbf{x}_s(\tilde{t}(\mathbf{x},t)) = \mathbf{x}_s(\tilde{t}_0) + \int_{\tilde{t}_0}^{\tilde{t}(\mathbf{x},t)} \mathbf{v}(t)\; dt \;, \qquad (14)$$

whereby the instantaneous speed $\mathbf{v}(t)$ is a vector. $\mathbf{v}(t)$ itself is the integration over the instantaneous acceleration $\mathbf{a}(t)$ with respect to time.
However, the solution to the integral in equation (2) can be unknown. As a work around, the source trajectory can be assumed to be piece-wise linear with constant speed. Then, the approach presented in this paper can be directly applied.
It might be favorable to turn and translate the coordinate system such that the virtual source always moves on the x-axis. In this case, care has to be taken that the gradient in equation (8) is taken properly with respect to the normal vector of the secondary source distribution. This operation can be comfortably implemented. The driving function expressed in dependency of the direction of the normal vector $\mathbf{n} = [\cos\alpha_\mathbf{n}\;\; \sin\alpha_\mathbf{n}]^T$ reads

$$\begin{aligned} d(\mathbf{x},t) = \\ = \Bigg[\cos\alpha_\mathbf{n}\left(\frac{\Phi(x,t)}{\Psi^2(\mathbf{x},t)} + \frac{j\omega_\mathrm{s}}{c(1-M^2)}\left(M + \frac{\Phi(x,t)}{\Psi(\mathbf{x},t)}\right)\right) + \\ + \sin\alpha_\mathbf{n}\,\frac{y}{\Psi(\mathbf{x},t)}\left(\frac{1}{\Psi(\mathbf{x},t)} + \frac{j\omega_s}{c(1-M^2)}\right)\Bigg] \times \\ \times\, s(\mathbf{x},\tilde{t}(\mathbf{x},t)) \;. \qquad (15) \end{aligned}$$

Confer to figure 8.
Note that equation (15) describes the driving function for monochromatic signals. The factor ω_s constitutes a frequency dependent weight. For broadband signals, this frequency dependent weighting is essentially a filtering operation. Comparable filtering also occurs in the reproduction of static virtual sources [9].
The treatment of arbitrarily shaped secondary source arrays is straightforward. In that case, a modified version of the Kirchhoff-Helmholtz integral replaces the Rayleigh integral in equation (8) [9]. Since the secondary sources are mutually independent, the directive for the calculation of the secondary source driving signal stays essentially the same and equation (15) can be applied. Note that for arbitrarily shaped secondary source arrays, a secondary source selection has to be additionally performed [14].
The implementation requires furthermore the ability to evaluate the source input signal continuously. For this operation, numerous approaches exist. The most important of which can be found in [15].

6. CONCLUSIONS

A basic theoretical framework for the reproduction of moving virtual sound sources was presented. On the example of wave field synthesis it was demonstrated that the explicit consideration of the physical properties of the wave field of a moving source avoids artifacts particular to conventional implementations. Doppler Effects both for moving sources and inherently also for moving listeners are accurately

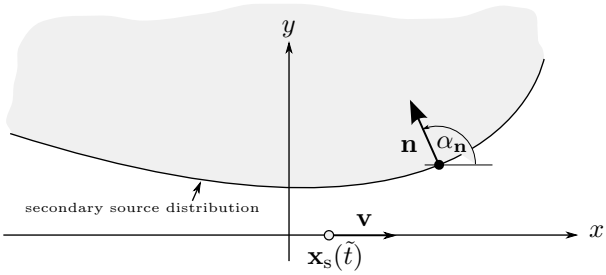


Fig. 8: Example of a curved secondary source array. The dot • denotes the position of the considered secondary source. The grey-shaded area denotes the listening area.

reproduced. The frequency shift observed by a receiver does not match perfectly the one occurring for real sources. However, informal listening suggests that this deviation is perceptually not significant. A thorough analytical and perceptual analysis of this circumstance is beyond the scope of this paper.
It turned out that what was referred to in the literature as artifacts of the reproduction system can not be assigned to the reproduction method. It is rather such the description of the wave field of the moving virtual source as a sequence of stationary positions is not appropriate. This description represents a physically non-meaningful wave field and therefore any sound field reproduction technique driven by this approach will reproduce a physically non-meaningful wave field. It can be assumed that the findings derived in this paper also hold true for any type of sound field reproduction method that employs time delays in the derivation of the secondary source driving signals.
Real-world implementations of WFS suffer from a number of unavoidable artifacts due to the fact that the physical basis of WFS can not be perfectly implemented. For static virtual sound sources, these inherent artifacts are reported to be of minor perceptual significance. Simulations presented in this paper suggest that some of these inherent artifacts are strongly audible with moving virtual sources. Most notably, spatial aliasing is suspected to cause a major degradation of the reproduction accuracy by widening the virtual source's temporal spectrum. Artifacts due to the employment of finite length loudspeaker arrays (truncation artifacts) can be strongly attenuated by applying appropriate tapering. The consequence of the WFS-inherent amplitude error can not be assessed by means of the presented results. Contrary to conventional implementations, our approach does not introduce additional artifacts, when a virtual source moves in the vicinity of the secondary sources.
Further research work on this topic includes the treatment of virtual sources moving at supersonic speeds, compensation for the inherent frequency shifts which are unwanted in certain applications, and focussing of moving virtual sources.
Finally, a thorough analysis of the remaining artifacts both analytically and perceptually has to be conducted.

ACKNOWLEDGEMENTS

We thank Holger Waubke of Österreichische Akademie der Wissenschaften for providing us with the notes of his lecture on theoretical acoustics [6].

7. REFERENCES

[1] R. Rabenstein and S. Spors. Sound field reproduction. In Benesty, J., Sondhi, M., Huang, Y, (Eds.), Springer Handbook on Speech Processing and Speech Communication, Springer Verlag, 2007.

[2] C. Doppler. Über das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels. In *Abhandlungen der königlichen böhmischen Gesellschaft der Wissenschaften, 2, pp. 465–482*, 1842.

[3] A. Franck, A. Gräfe, T. Korn, and M. Strauß. Reproduction of moving virtual sound sources by wave field synthesis: An analysis of artifacts. *32nd Int. Conference of the AES, Hillerød, Denmark*, Sept. 2007.

[4] H. Strauss. Simulation instationärer Schallfelder für virtuelle auditive Umgebungen. Fortschrittberichte VDI 10/652, VDI Verlag, Düsseldorf, 2000.

[5] J.D. Jackson. *Classical Electrodynamics*. Wiley, New York, 1975.

[6] H. Waubke. Aufgabenstellung zur Seminararbeit zur Vorlesung "Theoretische Akustik". IEM Graz, 2003.

[7] B. Girod, R. Rabenstein, and A. Stenger. *Signals and Systems*. J.Wiley & Sons, 2001.

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

[9] S. Spors, R. Rabenstein, and J. Ahrens. The theory of wave field synthesis revisited. In *124th Convention of the AES*, Amsterdam, The Netherlands, May 17–20 2008.

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

[11] E.N.G. Verheijen. Sound reproduction by wave field synthesis. PhD thesis, Delft University of Technology, 1997.

[12] S. Spors, M. Renk, and R. Rabenstein. Limiting effects of active room compensation using wave field synthesis. In *118th AES Convention*, Barcelona, Spain, May 2005. Audio Engineering Society (AES).

[13] S. Spors and R. Rabenstein. Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis. *120th Convention of the AES, Paris, France*, May 2006.

[14] S. Spors. Extension of an analytic secondary source selection criterion for wave field synthesis. *123rd AES Conv., NY, NY, USA, 5–8 Oct.*, 2007.

[15] T.I. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine. Splitting the unit delay: Tools for fractional delay filter design. *IEEE Signal Processing Magazine*, 13:30–60, January 1996.