Audio Engineering Society

Convention Paper

Presented at the 125th Convention
2008 October 2–5 San Francisco, CA, USA

Reproduction of Virtual Sound Sources Moving at Supersonic Speeds in Wave Field Synthesis

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 conventional implementations of wave field synthesis, moving sources are reproduced as sequences of stationary positions. As reported in the literature, this process introduces various artifacts. It has been shown recently that these artifacts can be reduced when the physical properties of the wave field of moving virtual sources are explicitly considered. However, the findings were only applied to virtual sources moving at subsonic speeds. In this paper we extend the published approach to the reproduction of virtual sound sources moving at supersonics speeds. The properties of the actually reproduced sound field are investigated via numerical simulations.

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, in sources moving slower than the speed of sound, the sound waves emitted in the direction of motion experience an increase in frequency. Sound

waves emitted in opposite direction of motion experience 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 but 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. Furthermore, this approach is limited to the reproduction of virtual sources moving slower than the speed of sound. 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, 5] for a treatment of moving virtual sources in binaural (HRTF-based) reproduction.
Various alternative implementations of the conventional approach of concatenating stationary source positions as outlined above are being applied both frame-based as well as in a sample-by-sample fashion. Most notably, in [3] it is proposed to incorporate the retarded time of a moving source (see section 2) into the driving function of a stationary source. Results presented ibidem show that this strategy still leaves prominent artifacts.
As shown by the authors in [6], the mentioned artifacts occurring in conventional implementations can be avoided when the physical properties of the wave field of moving sound sources are a priori taken into account. However, the approach in [6] was exclusively applied to virtual sources moving slower than the speed of sound. In this paper, we extend this approach to the reproduction of virtual sources moving at supersonic speeds. Our work can also be regarded as an extension of the approach presented in [7] which focuses on the reproduction of the frequency content present in supersonic booms of aircrafts but does not physically reproduce the actual wave front.
Note that the considerations presented in this paper are of relevance only for sound field reproduction approaches which employ time delays in the procedure of yielding the loudspeaker driving signals.

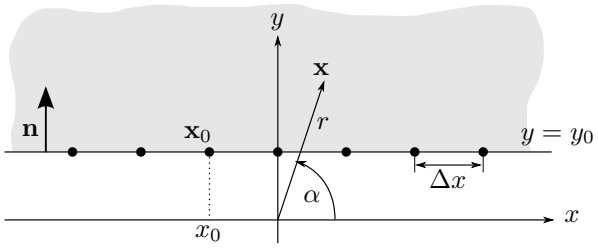


Fig. 1: The coordinate system and geometry used in this paper. The dots • denote the positions of the secondary sources used for wave field synthesis. The grey-shaded area denotes the listening area.

2. THE WAVE FIELD OF A MOVING SOURCE

The fundamental prerequisite for model-based sound field reproduction is the knowledge of the sound field that is to be recreated. In this section, we derive analytical expressions of the sound field of a moving sound source. For simplicity, we assume a monopole source. However, the presented approach also allows for the treatment of arbitrary source types. The derivation below follows [8, 9].
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 $g\left(\mathbf{x}-\mathbf{x}_\mathrm{s}, t\right)$. See figure 1 for a sketch of the coordinate system. 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* [8]. $\tilde{t}(\mathbf{x},t)$ 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 weighted 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

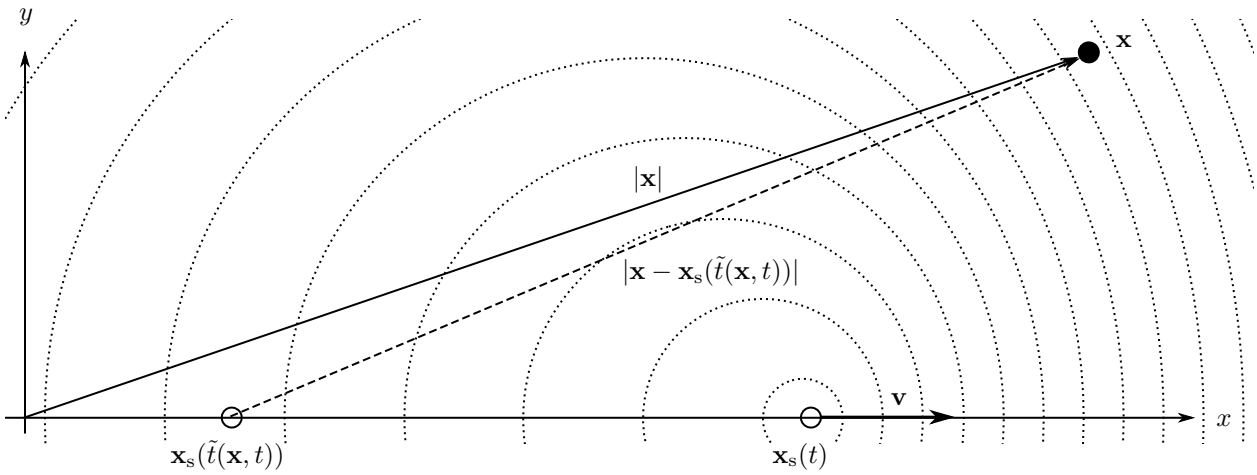


Fig. 2: Derivation of the Green's function of a moving sound source.

yield the wave field emitted due to the entire sequence of Dirac pulses, we integrate over $\tilde{t}$ as

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

whereby we temporarily altered the nomenclature for convenience ($\tilde{t} = \tilde{t}(\mathbf{x},t)$).
Assuming a moving monopole sound source, its Green's function explicitly reads

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

Note that

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

is referred to as *retarded time* [8]. 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*, with v being the speed of the sound source.
For convenience, we assume the virtual source to move uniformly along the x-axis in positive x-direction (cf. to figure 2). As outlined in [6], arbitrary trajectories can be approximated by assuming a piece-wise uniform motion and an appropriate translation and rotation of the coordinate system. At time $t = 0$ 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

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

and the exploitation of the sifting property of the delta function [10]. It turns out that the integral has different solutions for $M < 1$, $M = 1$, and $M > 1$. In the following sections, we present solutions to the integral in (2) for subsonic ($M < 1$) as well as supersonic ($M > 1$) sound sources and briefly comment on the case of sources moving at the speed of sound ($M = 1$).

2.1. Sound sources moving at subsonic speeds

For $M < 1$, the integral boundaries in (2) can be kept and the solution, i.e. the sound field $s_{M<1}(\mathbf{x},t)$ of a source moving at a speed $v < c$ reads then

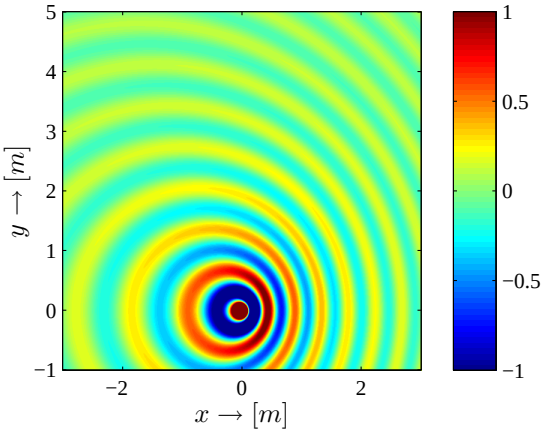
$$s_{M<1}(\mathbf{x},t) = \frac{1}{4\pi} \cdot \frac{s_0(\tilde{t}(\mathbf{x},t))}{\Psi(\mathbf{x},t)}\ , \qquad (6)$$

whereby

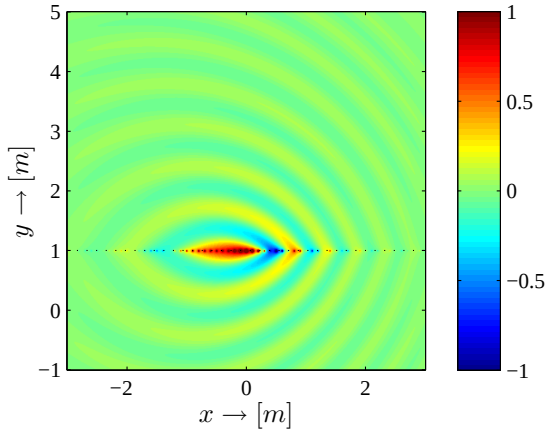
$$\tilde{t}(\mathbf{x},t) = t - \frac{M\Phi(x,t) + \Psi(\mathbf{x},t)}{c(1-M^2)}\ ,$$

$$\Psi(\mathbf{x},t) = \sqrt{\Phi^2(x,t) + y^2(1-M^2)}\ ,$$

$$\Phi(x,t) = x - vt - x_\mathrm{s}(0)\ .$$



(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.

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

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

which corresponds to the familiar expression for the sound field of a static harmonic monopole sound source [6].

2.2. Sound sources moving at supersonic speeds

For sound sources moving at supersonic speeds, the integral in (2) has to be split into a sum of two integrals after the substitution (5) reading

$$s_{M>1}(\mathbf{x}, t) = \int_{u_1}^{\infty} (\cdot)\ du + \int_{u_2}^{\infty} (\cdot)\ du, \tag{8}$$

whereby

$$u_{1,2} = \frac{1}{v}\left(\pm(x_\mathrm{s}(0) - x) + y\sqrt{M^2 - 1}\right)\ .$$

$(\cdot)$ denotes the argument of the integral in (2). The solution yields the wave field $s_{M>1}(\mathbf{x}, t)$ of a monopole sound source moving at a supersonic speed v reading

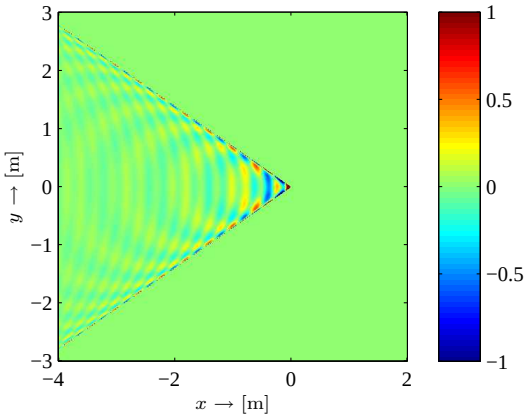
$$s_{M>1}(\mathbf{x}, t) = \begin{cases} s_1(\mathbf{x}, t) + s_2(\mathbf{x}, t) & \text{for } \Phi(x,t)^2 + y^2(1-M^2) \geq 0 \\ & \text{and } x_\mathrm{s}(0) + vt \geq x \\ 0 & \text{elsewhere}\ , \end{cases} \tag{9}$$

with

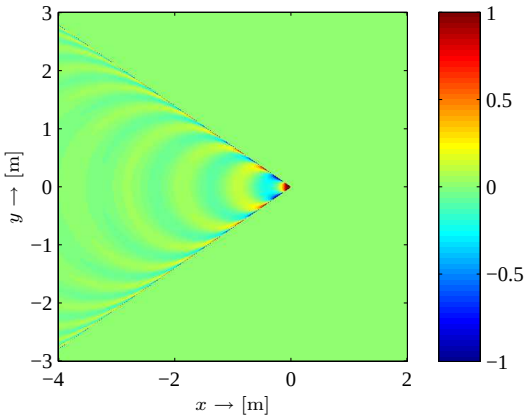
$$s_{1,2}(\mathbf{x}, t) = \frac{1}{4\pi} \frac{s_0(\tilde{t}_{1,2}(\mathbf{x}, t))}{\Psi(\mathbf{x}, t)}\ ,$$

$$\tilde{t}_{1,2}(\mathbf{x}, t) = t - \frac{M\Phi(x,t) \pm \Psi(\mathbf{x}, t)}{c(1-M^2)}\ ,$$

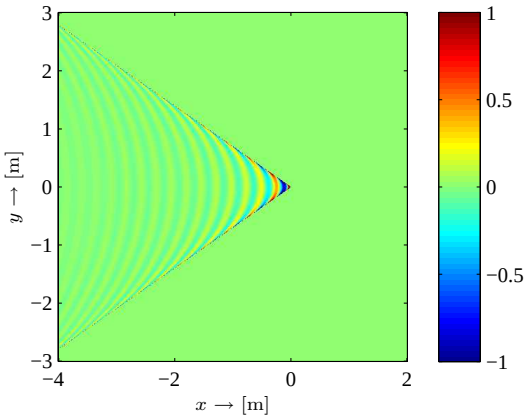
The most prominent property of the wave field of a supersonic source is the formation of the so-called *Mach cone*, a conical sound pressure front following the moving source. See figure 4(a). Note that the



(a) Wave field $s_{M>1}(\mathbf{x}, t)$ of a supersonic source.



(b) Backward travelling component $s_1(\mathbf{x}, t)$.



(c) Forward travelling component $s_2(\mathbf{x}, t)$.

Fig. 4: Wave field of a source traveling at 600 $^{\text{m}}/_{\text{s}}$ ($M \approx 1.7$).

Mach cone is a direct consequence of causality.

For the receiver this has two implications: (1) He/She does not receive any sound wave before the arrival of the Mach cone, (2) after the arrival of the Mach cone the receiver is exposed to a superposition of the wave field which the source radiates into backward direction $s_1(\mathbf{x}, t)$ and the wave field $s_2(\mathbf{x}, t)$ which the source had radiated into forward direction before the arrival of the Mach cone. $s_1(\mathbf{x}, t)$ carries a frequency shifted version of the emitted signal propagating in opposite direction to the source motion (figure 4(b)), $s_2(\mathbf{x}, t)$ carries a time-reversed version of the emitted signal following the source (figure 4(c)). The latter is generally also shifted in frequency.

2.3. Sound sources moving at the speed of sound

The integral in (2) can also be solved for $M = 1$. In that case, the lower integral boundary is finite, the upper boundary is infinite. The result then resembles the circumstances for $M > 1$, i.e the receiver is not exposed to the source's wave field at all times. It is rather such that the source moves at the leading edge of the sound waves it emits. The wave field can not surpass the source. The leading edge of the wave field is termed *sound barrier*.

Unlike for $M > 1$, the resulting wave field is not composed of two different components. It contains only one single component carrying the frequency shifted input signal.

Informal listening suggests that it can not be assumed that the human ear is aware of the details of the properties of the wave field of a transonic source (a source moving exactly at the speed of sound). We therefore do not present an explicit treatment here. For convenience, we propose to assume that the wave field of a transonic source is perceptually indistinguishable from the wave field $s_1(\mathbf{x}, t)$ of a source moving at a speed slightly faster than the speed of sound c.

3. WAVE FIELD SYNTHESIS

In this section, we demonstrate how a moving virtual sound source can be reproduced using the findings 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 [11, 12]. 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

$$p_{\mathrm{WFS}}(\mathbf{x}, t) = \int_{-\infty}^{\infty} \underbrace{-\frac{\partial}{\partial \mathbf{n}} s(\mathbf{x}, t)|_{\mathbf{x}=\mathbf{x}_0}}_{d(\mathbf{x}_0, t)} *_t \; g(\mathbf{x}, t) \; dx_0 \; . \tag{10}$$

$s(\mathbf{x}, t)$ 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.
Due to the fact that the physical requirements can not be perfectly fulfilled in practical implementations, the virtual source's wave field is not perfectly reproduced in the receiver's half-space. Equation (10) 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 (10) requires secondary line sources which are positioned perpendicular to the receiver plane [12]. 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

$$d_{\mathrm{corr}}(\mathbf{x}, t) = f(t) *_t d(\mathbf{x}, t) \; . \tag{11}$$

$f(t)$ 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 [12] for a thorough treatment of the properties of WFS.
For convenience, we do not explicitly compensate for the secondary source mismatch in the analytical expressions for the driving functions. However, in the simulations this compensation is performed.

3.1. Driving function for subsonic sources

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 $d(\mathbf{x}, t)$ derived from (6) and (10) reads [6]

$$d_{\mathrm{sub}}(\mathbf{x}, t) = \frac{y(1-M^2)}{\Psi(\mathbf{x}, t)} \left(\frac{1}{\Psi(\mathbf{x}, t)} + \frac{j\omega_s}{c(1-M^2)} \right) \times \\ \times \, s(\mathbf{x}, t) \; . \tag{12}$$

Note that $d_{\mathrm{sub}}(\mathbf{x}, t)$ in equation (12) implicitly includes static virtual sources.
The wave field reproduced by a linear WFS array driven by equation (12) is depicted in figure 3(b). The overall length of the loudspeaker array is 8 m. The virtual source moves at a speed $v = 120\frac{\mathrm{m}}{\mathrm{s}}$ along the x-axis in positive x-direction ($M \approx \frac{1}{3}$).

3.2. Driving function for supersonic sources

The driving function for supersonic sources derived from (9) and (10) reads

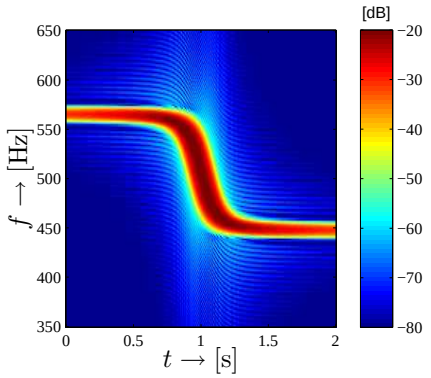
$$\begin{aligned} d_{\mathrm{sup}}(\mathbf{x}, t) \; &= \; d_1(\mathbf{x}, t) + d_2(\mathbf{x}, t) = \\ &= \; \frac{y(1-M^2)}{\Psi(\mathbf{x}, t)} \left(\frac{1}{\Psi(\mathbf{x}, t)} + \frac{j\omega_s}{c(1-M^2)} \right) \times \\ &\quad \times \, s_1(\mathbf{x}, t) \; + \\ &\quad + \; \frac{y(1-M^2)}{\Psi(\mathbf{x}, t)} \left(\frac{1}{\Psi(\mathbf{x}, t)} - \frac{j\omega_s}{c(1-M^2)} \right) \times \\ &\quad \times \, s_2(\mathbf{x}, t) \; . \end{aligned} \tag{13}$$

3.3. Driving function for transonic sources

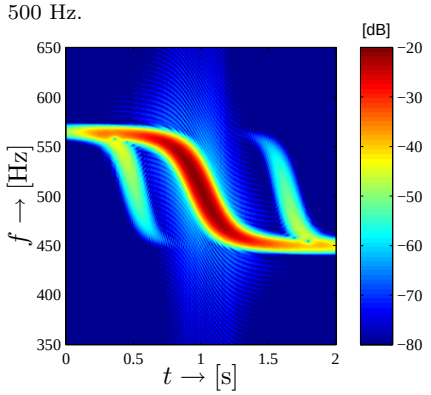
As outlined in section 2.3, we propose to reproduce $s_1(\mathbf{x}, t)$ of a virtual source moving slightly faster than the speed of sound in order to approximate a transonic source. The appropriate driving function is then $d_1(\mathbf{x}, t)$.

4. RESULTS

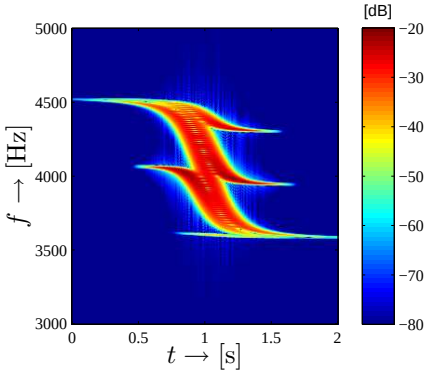
In this section, we present a number of simulations in order to analyze the properties of the proposed



(a) Real source. The emitted frequency is 500 Hz.



(b) Truncation artifacts. The length of the array is 40 m. The emitted frequency is 500 Hz.



(c) Spatial aliasing. The desired signal is the S-shaped one in the middle. The emitted frequency is 4000 Hz.

Fig. 5: Spectrograms illustrating artifacts apparent in the reproduced wave field of a subsonic source ($v = 40$ m/s). The virtual source passes the receiver at $t \approx 1$ s.

approach with focus on the case of $M > 1$. The case of $M < 1$ is thoroughly treated in [6].

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. Its overall length is 14 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.

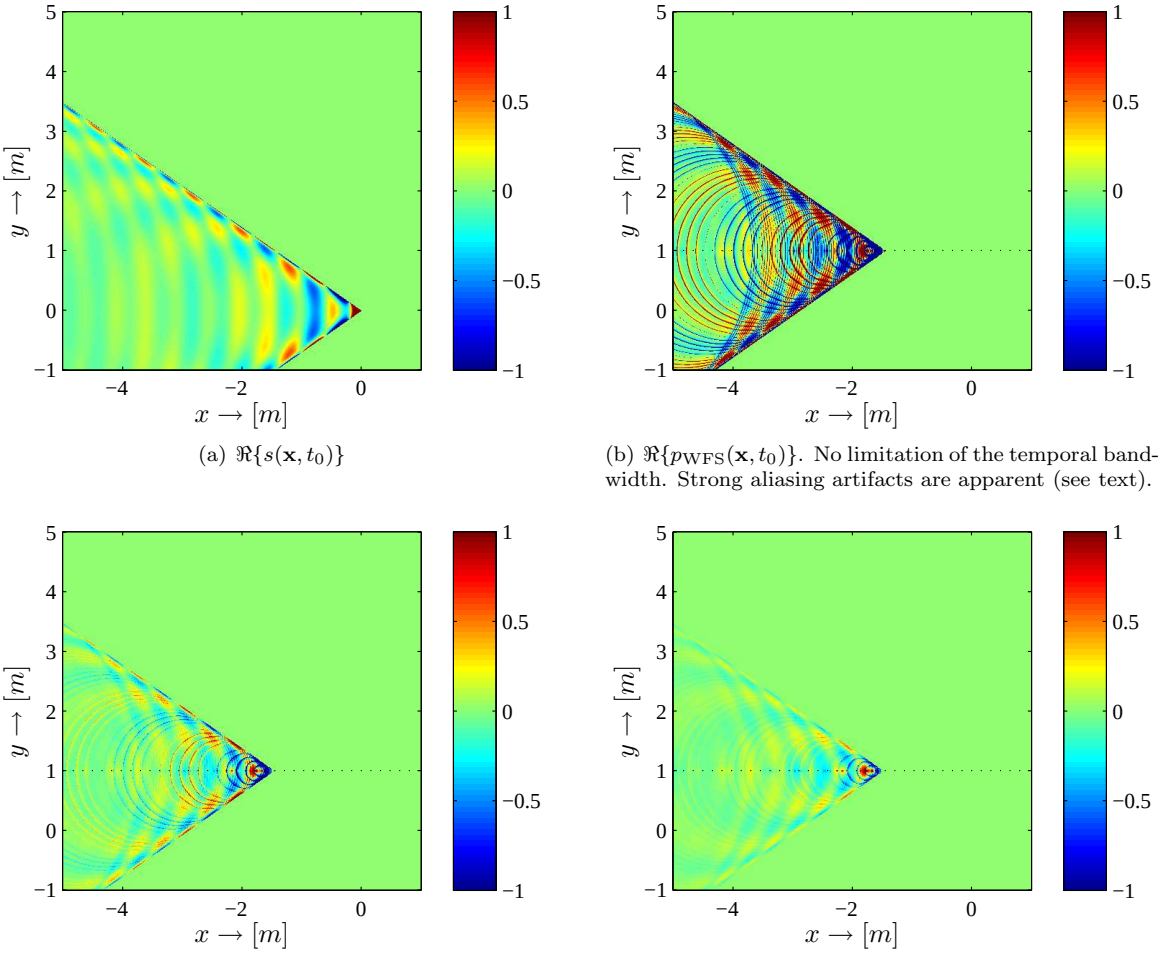
4.1. Artifacts apparent in the reproduced wave field

As outlined in [6], the reproduced wave field suffers from two major artifacts: (1) echo-like artifacts due to spatial truncation of the secondary source array, and (2) spatial aliasing when the frequency content of the reproduced wave field is above the spatial aliasing frequency. Figure 5 shows spectrograms of the reproduced wave field observed at $\mathbf{x}_R = [0\ -4]^T$. The loudspeaker array similar to the one used in the simulations in figure 3, i.e. the loudspeaker array 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

In figure 5(b) a pre- and a post-echo additional to the desired signal are apparent. The shorter the array the closer in time to the desired signal the echoes occur. These truncation artifacts can be significantly reduced by the application of tapering (i.e. an attenuation of the secondary sources towards the very ends of the array) [11, 6].

Figure 5(c) depicts the spectrogram of a virtual source reproduced above the spatial aliasing frequency. For the given array with a loudspeaker spacing of $\Delta = 0.1$ m the spatial aliasing frequency is approximately 1700 Hz [13].

Finally, another artifact resulting from spatial truncation of the secondary source distribution is an incorrect amplitude envelope of the receiver signal. This circumstance can be observed when comparing e.g. figures 5(a) and 5(b). At the very ends of the depicted time window, the receiver signal due to the real source is significantly higher in amplitude than



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

(b) $\Re\{p_{\mathrm{WFS}}(\mathbf{x}, t_0)\}$. No limitation of the temporal bandwidth. Strong aliasing artifacts are apparent (see text).

(c) $\Re\{p_{\mathrm{WFS}}(\mathbf{x}, t_0)\}$, $f_{\max} = 3000$ Hz.

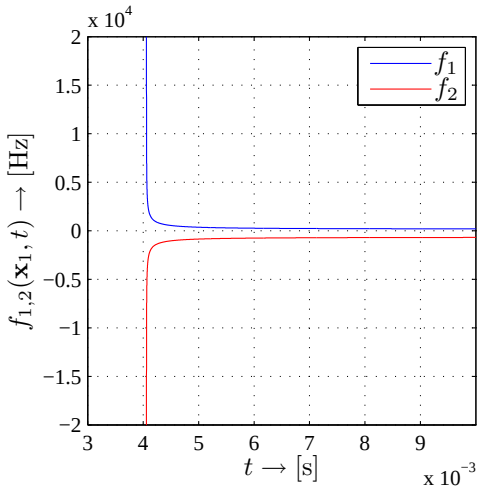
(d) $\Re\{p_{\mathrm{WFS}}(\mathbf{x}, t_0)\}$, $f_{\max} = 2000$ Hz.

Fig. 6: 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 $v = 600$ m/s ($M \approx 1.7$). 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. The loudspeaker array in figures 6(b)-6(d) is indicated by the dotted line. It is situated symmetrically around the y-axis at $y_0 = 1$ m and its overall length is 14 m. The loudspeakers are positioned at intervals of $\Delta x = 0.1$ m.

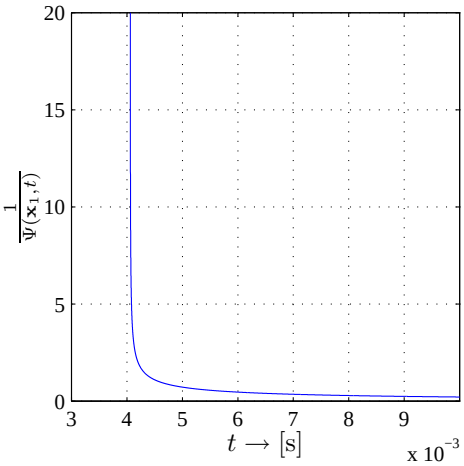
the receiver signal due to the virtual source. In the center of the plot, i.e. when the source is behind the secondary sources from the receivers point of view, the amplitude due to the virtual source is similar to that due to the real source.

4.2. Direct application of the driving function for $M > 1$

Figure 6(b) shows a simulation of a WFS system reproducing the wave field depicted in figure 6(a). The virtual source moves at $v = 600$ m/s, i.e. $M \approx 1.7$. Due to the omnidirectionality of the secondary sources, the reproduced wave field in figure 6(b) is



(a) $f_{1,2}(\mathbf{x}_1, t)$. Negative frequencies indicate time reversal of the input signal.



(b) $\frac{1}{\Psi(\mathbf{x}_1, t)}$.

Fig. 7: Details of the wave field of a source of $v = 600$ m/s ($M \approx 1.7$) oscillating at $f_s = 500$ Hz observed at $\mathbf{x}_1 = [1\ 1]^T$. The Mach cone arrives at $t \approx 4 \cdot 10^{-3}$ s.

symmetric with respect to the secondary source contour. Note that strong artifacts are apparent. It can be shown that these artifacts occur due to temporal as well as spatial aliasing.

This can by verified by analyzing the instantaneous frequencies $f_1(t)$ and $f_2(t)$ of the reproduced wave field components $s_1(\mathbf{x}, t)$ and $s_2(\mathbf{x}, t)$. Confer to figure 7(a). It can be seen that $f_1(t)$ and $f_2(t)$ are infinite at the singularity of the Mach cone, i.e. at the moment of the arrival of the Mach cone. After the arrival they decrease quickly to moderate values. The former means that $f_1(t)$ and $f_2(t)$ will exceed any limit imposed on a reproduction system due to discrete treatment of time and discretization of the secondary source distribution.

4.3. Modified driving function

In order to prevent *temporal* aliasing in digital systems due to discretization of the time, it is desirable to limit the bandwidth of the temporal spectrum of the driving function. Typical bandwidths in digital systems are 22050 Hz for systems using a temporal sampling frequency of 44100 Hz and 24000 Hz for systems using a temporal sampling frequency of 48000 Hz.

In order to prevent respectively reduce *spatial* aliasing of the WFS system under consideration, it is desirable to further limit the bandwidth of the temporal spectrum of the driving function to values in the order of the spatial aliasing frequency which is typically a few thousand Hertz. Recall that the critical frequency above which spatial aliasing occurs in the given secondary source array is approximately 1700 Hz (confer to section 4.1).

A simple means to limit the bandwidth is to simply fade-in the driving signal from a moment on when its temporal frequency has dropped below a given threshold. This strategy also avoids the circumstance that the amplitude of the driving signal is infinite at the moment of arrival of the Mach cone. Real-world implementations of WFS systems can not reproduce arbitrarily high amplitudes.

Confer to figure 7(b). It depicts the factor $\Psi(\mathbf{x}, t)^{-1}$ which determines the amplitude of the wave field around the Mach cone.

The simulations in figures 6(c) and 6(d) show the reproduced wave field when the driving function is faded-in after the instantaneous frequency of the driving function has dropped below 3000 Hz (figure 6(c)) respectively 2000 Hz (figure 6(d)). The aliasing artifacts are significantly reduced.

Note that the shorter the fade-in of the driving function is the better the impulsive property of the Mach cone is preserved. However, shorter fade-ins result in stronger spatial aliasing since they impose more high frequency content onto a signal.

Finally, it has to be considered that spatial aliasing is not necessarily audible under all circumstances.

5. PERCEPTUAL ASPECTS

Informal listening suggests that the human auditory system is not aware of all the properties of the wave field of supersonic sources. Especially the fact that the wave field contains a component carrying a time-reversed version of the source's input signal is confusing. Depending on the specific situation, it might be preferable to exclusively reproduce $s_1(\mathbf{x}, t)$, i.e. the component of the wave field carrying the non-reversed input signal.
Furthermore, only the localization when exposed to $s_1(\mathbf{x}, t)$ is plausible since $s_1(\mathbf{x}, t)$ assures localization of the source in its appropriate location (however with some bias due to the retarded time τ). Exposure of the receiver to $s_2(\mathbf{x}, t)$ suggests localization of the source in the direction where it "comes from". This also seems unnatural. Finally, the exposure of the receiver to a superposition of $s_1(\mathbf{x}, t)$ and $s_2(\mathbf{x}, t)$ suggests the localization of two individual sources.

6. CONCLUSIONS

An approach to the reproduction of the wave field of virtual sound sources moving at supersonic speeds was presented. The approach constitutes an extension to a treatment of the reproduction of the wave field of virtual sound sources moving at subsonic speeds previously published by the authors. It was shown that the reproduced wave field suffers from spatial aliasing artifacts due to the fact that the instantaneous frequency of the virtual sound field is infinite at the moment of arrival of the Mach cone. As workaround, it was proposed to fade-in the driving signal for a given secondary source right after the instantaneous frequency of the driving signal has dropped below a desired threshold. A short fade-in preserves the impulsive quality of the Mach cone.
In order to optimize the reproduction of the sound field of supersonic virtual sources, it is necessary to perform preceptive experiments investigating which properties of the virtual wave field have to be reproduced in order to evoke a plausible perception both in terms of frequency content and localization.

ACKNOWLEDGEMENTS

We thank Holger Waubke of Austrian Academy of Sciences for providing us with the notes of his lecture on theoretical acoustics [9].

7. REFERENCES

[1] R. Rabenstein and S. Spors. Multichannel 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] Y. Iwaya and Y. Suzuki. Rendering moving sound with the doppler effect in sound space. *Applied Acoustics, Technical note*, 68:916–922, 2007.

[6] J. Ahrens and S. Spors. Reproduction of moving virtual sound sources with special attention to the doppler effect. In *124th Convention of the AES*, Amsterdam, The Netherlands, May 17–20 2008.

[7] N. Epain and E. Friot. Indoor sonic boom reproduction using ANC. In *Proceedings of Active*, Williamsburg, Virginia, Sep. 20–22 2004.

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

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

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

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

[12] 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.

[13] S. Spors and R. Rabenstein. Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis. In *120th Convention of the AES, Paris, France*, May 20–23 2006.