IMPLEMENTATION OF DIRECTIONAL SOURCES IN WAVE FIELD SYNTHESIS

Jens Ahrens and Sascha Spors

Deutsche Telekom Laboratories, Berlin University of Technology,
Ernst-Reuter-Platz 7, 10587 Berlin, Germany
{jens.ahrens,sascha.spors}@telekom.de

ABSTRACT

Wave field synthesis (WFS) is a spatial audio reproduction technique aiming at physically synthesizing a desired sound field. Typically, virtual sound sources are rendered as emitting spherical or plane waves. In this paper we present an approach to the implementation of sources with arbitrary directivity. The approach is based on the description of the directional properties of a source by a set of circular harmonics. A time domain expression of the loudspeaker driving signals is derived allowing an efficient implementation. Consequences of sampling and truncation of the secondary source distribution as occurring in typical installations of WFS systems are discussed and simulated reproduction results are shown.

1. INTRODUCTION

Wave field synthesis (WFS) is a massive multichannel audio reproduction technique relying on physical descriptions of sound fields rather than on psycho-acoustical principles like conventional stereo and surround sound do. It is suitable to satisfy an extended listening area. Theoretically, there is no such limitation like a sweet-spot. WFS is commonly implemented for two-dimensional reproduction in the horizontal plane. In such a setup the listening area is typically surrounded by (piece-wise) linear arrays of closely spaced loudspeakers termed secondary sources.

For WFS 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 rather concentrate on the latter case where a sound scene is composed of a number of sound sources derived from analytical spatial source models. Typically, these sound sources are rendered emitting plane or spherical waves. This circumstance does not exploit all potentials of WFS since directional properties of sound sources are known to contribute to immersion and presence of a sound scene.

The expansion of a source sound field into spatial harmonics to synthesize source directivity has already been presented in [2] for Ambisonics rendering and in [3] for WFS. However, the focus in [3] lies rather on optimal equalization than on an efficient implementation.

In this paper we present an approach to the implementation of such directional sound sources based on the circular harmonics expansion of a source sound field. It retains the freedom of arbitrarily creating and manipulating directivity. We furthermore provide an efficient time domain implementation scheme.

1.1. Nomenclature

Although arbitrary geometries of the secondary source distribution are possible we will restrict the descriptions in this paper to

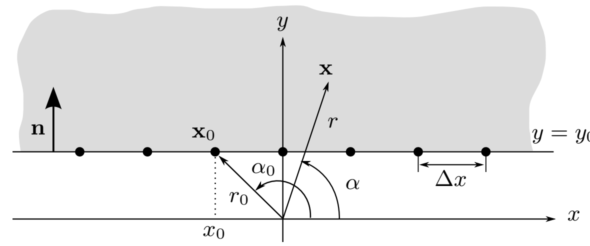


Figure 1: *The coordinate system and geometry used in this paper. The dots • denote the positions of the secondary sources when the distribution is sampled. The grey-shaded area denotes the listening area.*

linear secondary source arrays for two-dimensional reproduction, since this is the case which is implemented most frequently. Two-dimensional in this context means that an observed sound field is independent from one of the spatial coordinates, so that e.g. $P(x, y, z, \omega) = P(x, y, \omega)$.

The following conventions are used in this paper: For scalar variables lower case denotes the time domain, upper case the temporal frequency domain. Vectors are denoted by lower case boldface. The spatial frequency domain is denoted by a tilde placed over the respective symbol. The position vector in Cartesian coordinates is given as $\mathbf{x} = [x\ y]^T$ and is linked to the cylindrical coordinate system via $x = r \cos\alpha$ and $y = r \sin\alpha$ (cf. to figure 1). The vector $\mathbf{k} = [k_x\ k_y]^T$ denotes the wave vector whose absolute value $k = |\mathbf{k}|$ satisfies the dispersion relation $k^2 = \left(\frac{\omega}{c}\right)^2$ with ω being the radial frequency and c the speed of sound.

2. WAVE FIELD SYNTHESIS

The theoretical basis of WFS with linear secondary source arrays is given by the two-dimensional Rayleigh I integral [4]. 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.

Without loss of generality, the secondary source array is assumed to be parallel to the x-axis at $y = y_0$ as depicted in figure 1 and to be of infinite length in a first step. The virtual source is situated in the origin of the coordinate system. Any other situation may be treated by an appropriate rotation and/or translation of the coordinate system. The listening area is chosen to be at $y > y_0$. The results derived in this paper hold there only.

The two-dimensional Rayleigh I integral determines the sound pres-

978-1-4244-1619-6/07/$25.00

sure $P(\mathbf{x},\omega)$ created by such a setup reading

$$P(\mathbf{x},\omega) = -\int_{-\infty}^{\infty} \underbrace{2\frac{\partial}{\partial \mathbf{n}} S(\mathbf{x_0},\omega)}_{D(\mathbf{x}_0,\omega)} \cdot \underbrace{\frac{j}{4} H_0^{(2)}(\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|)}_{G_{2D}(\mathbf{x}-\mathbf{x_0},\omega)} \, dx_0 \ , \tag{1}$$

where $S(\mathbf{x},\omega)$ denotes the sound field of a virtual sound source, $G_{2D}(\mathbf{x}-\mathbf{x_0},\omega)$ the 2D free-field Green's function representing the secondary sources, $H_0^{(2)}$ the zeroth order Hankel function of second kind, and $\frac{\partial}{\partial \mathbf{n}}$ the gradient in the direction normal to the secondary source distribution. $\mathbf{x}$ is a point inside the half plane where the sound field is recreated. $D(\mathbf{x}_0,\omega)$ is the driving signal for a secondary source situated at $\mathbf{x_0} = [x_0\ y_0]^T$. The Green's function $G_{2D}(\mathbf{x}-\mathbf{x_0},\omega)$ may be interpreted as the sound field of a line source perpendicular to the x-y-plane intersecting it at $\mathbf{x_0}$.
Typical WFS systems employ loudspeakers with closed cabinets as secondary sources. These approximately have the characteristics of acoustic point sources. This mismatch in source types produces various artifacts that can be compensated for to some extend [4]. We will restrict our descriptions to the artifact free case of secondary line sources. However, our results also hold for secondary point sources within the limitations of the compensation of the mismatch artifacts.

3. DERIVATION OF SECONDARY SOURCE DRIVING FUNCTION

An arbitrary two-dimensional virtual source sound field $S(\mathbf{x},\omega)$ may be decomposed into circular harmonics which is the two-dimensional analogon of the cylindrical harmonics expansion [5]. Its favorable properties in our context are: (a) It provides an orthogonal expansion of the virtual source field, (b) rotations of the virtual source can be simply expressed by phase shifts, and (c) the expansion coefficients can be easily derived from the far-field (plane wave) characteristics of the virtual source via a Fourier transformation [6].
For a virtual source in the origin of the coordinate system this decomposition yields

$$S(\mathbf{x},\omega) = \hat{S}(\omega) \cdot \sum_{\nu=-\infty}^{\infty} \breve{S}^{(2)}(\nu,\omega) H_\nu^{(2)}(\left|\frac{\omega}{c}\right| r) e^{j\nu\alpha} \ , \tag{2}$$

where $\hat{S}(\omega)$ denotes the temporal source spectrum. The coefficients $\breve{S}^{(2)}$ are termed circular harmonics expansion coefficients. Note that $\hat{S}(\omega)$ and $\breve{S}^{(2)}$ may also be combined. But for the sake of an efficient implementation scheme we keep them separated (cf. to section 4).
Taking the directional gradient of equation (2), as indicated in equation (1), and exploiting a recurrence relation of the Hankel function [7] and Euler's identity yields

$$D(\mathbf{x}_0,\omega) = j\frac{\omega}{c} \cdot \hat{S}(\omega) \cdot \sum_{\nu=-\infty}^{\infty} \breve{S}^{(2)}(\nu,\omega) \times \\ \times \left(H_{\nu-1}^{(2)}\left(\left|\frac{\omega}{c}\right| r_0\right) e^{j(\nu-1)\alpha_0} + H_{\nu+1}^{(2)}\left(\left|\frac{\omega}{c}\right| r_0\right) e^{j(\nu+1)\alpha_0} \right) \tag{3}$$

for the driving function of the secondary sources.

4. IMPLEMENTATION

Although equation (3) gives an exact frequency domain description of the driving function for a directional source it is expensive to implement since it involves the evaluation of Hankel functions and a sum with a potentially great number of addends.
Towards a more efficient description in the time domain we employ the large argument approximation[1] of the Hankel function $H_\nu^{(2)}(z) = \sqrt{\frac{2}{\pi z}} e^{-j(z-\nu\frac{\pi}{2}-\frac{\pi}{4})}$ [7] to equation (3) yielding

$$D(\mathbf{x}_0,\omega) \approx 2\sqrt{\frac{2}{\pi r_0^3}} y_0 \cdot e^{-jkr_0} \cdot \hat{S}(\omega) \times \\ \times \underbrace{\sqrt{\frac{\omega}{jc}}}_{=G(\omega)} \underbrace{\sum_{\nu=-\infty}^{\infty} j^\nu \cdot \breve{S}^{(2)}(\nu,\omega) \cdot e^{j\nu\alpha_0}}_{=\bar{S}^{(2)}(\alpha_0,\omega)} \ . \tag{4}$$

In equation (4) the fact that r_0 and α_0 are not independent from each other in our specialized geometry has been exploited.
Note that the sum corresponds to the plane wave decomposition (PWD) $\bar{S}^{(2)}(\alpha_0,\omega)$ of the virtual source's wave field [6].
An inverse temporal Fourier transform [5] of (4) leads then to

$$d(\mathbf{x}_0,t) \approx \underbrace{2\sqrt{\frac{2}{\pi r_0^3}} y_0}_{f(\mathbf{x}_0)} \cdot \underbrace{\delta(t-\frac{r_0}{c}) * \hat{s}(t)}_{=\hat{s}(t-\frac{r_0}{c})} * g(t) * \bar{s}^{(2)}(\alpha_0,t) \ , \tag{5}$$

where $f(\mathbf{x}_0)$ is a weighting factor depending on the location of the respective secondary source, $\hat{s}(t)$ is the time domain source signal, $g(t)$ is the impulse response of a filter having a frequency response of $G(\omega)$, and $\bar{s}^{(2)}(\alpha_0,t)$ is the time domain correspondence of the PWD $\bar{S}^{(2)}(\alpha_0,\omega)$ of the virtual source directivity. Thus, the time domain driving signal for a given secondary sound source can be approximately yielded by a weighting and delaying of the time domain source signal and two filtering operations.
The filter $G(\omega)$ exhibits a high pass character with a slope of 3 dB per octave which is typical for virtual sources in WFS [1]. It is independent from the source/loudspeaker position. The second filter, capturing the spatio-temporal source characteristics, can be realized efficiently in the time domain by noting that: (a) Its impulse response will be short for typical source models, since only the source characteristics and not the room characteristics are captured and (b) the PWD coefficients can be computed from the circular harmonics coefficients by an inverse Fourier transformation with respect to the angle α_0. This Fourier transformation can be realized efficiently by the Fast Fourier Transformation (FFT). The evaluation of the potentially infinite sum in equation (3) is thus facilitated. However, for linear arrays the angles α_0 are generally not sampled equidistantly (see figure 1). This problem can be overcome by applying an inverse FFT for equidistant angles α_0 to compute the PWD coefficients and an interpolation to the intermediate angles in a second step to derive the required values of $\bar{s}^{(2)}(\alpha_0,t)$ for the actual loudspeaker/source geometry. Rotations of the virtual source can be efficiently realized by rotating

[1] In this case it is a far-field/high frequency approximation, i.e. the virtual source has to be sufficiently far away from the secondary sources, resp. the emitted frequency has to be high enough. Note that this approximation essentially corresponds to the stationary phase approximation commonly applied in WFS theory [4] for point source reproduction.

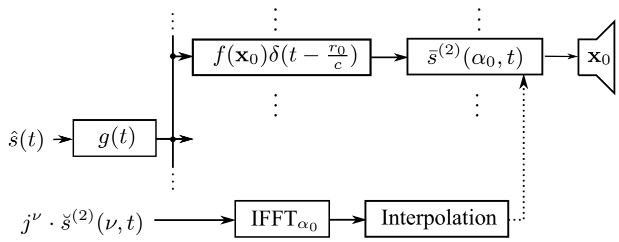


Figure 2: *Block diagram of the implementation scheme of the approximated time domain driving function $d(\mathbf{x}_0,t)$ for a secondary source at $\mathbf{x}_0$.*

the PWD coefficients before the interpolation step. A block diagram of the entire implementation scheme is depicted in figure 2. Note that the usual amplitude and spectral corrections necessary due to the application of two-dimensional theory in a three-dimensional world should also be applied in practical realizations [1, 4].

5. TRUNCATION AND SAMPLING

Practical implementations of linear secondary source arrays will always be of finite length and employ a finite number of loudspeakers. The consequences of this truncation and sampling are treated in this section. A thorough quantitative investigation is beyond the scope of this paper and is subject to ongoing research.
The truncation is modeled by multiplying the secondary source driving function $D(\mathbf{x}_0,\omega)$ with a suitable window function $w(x_0)$ [4]. Incorporating $w(x_0)$ into equation (1) yields the wave field $P_{tr}(\mathbf{x}_0,\omega)$ of a truncated linear source distribution as

$$P_{tr}(\mathbf{x},\omega) = -\int_{-\infty}^{\infty} w(x_0)D(\mathbf{x}_0,\omega)G_{2D}(\mathbf{x}-\mathbf{x}_0,\omega)dx_0 \ . \quad (6)$$

A spatial Fourier transform of $P_{tr}(\mathbf{x},\omega)$ yields the spatial spectrum $\tilde{P}_{tr}(\mathbf{k},\omega)$ of the reproduced wave field as

$$\tilde{P}_{tr}(\mathbf{k},\omega) = \underbrace{-\frac{1}{2\pi}\left(\tilde{w}(k_x) *_{k_x} \tilde{D}(k_x,\omega)\right)}_{\tilde{D}_{tr}(k_x,\omega)} \tilde{G}_{2D}(\mathbf{k},\omega) \ , \quad (7)$$

whereby the asterisk $*_{k_x}$ denotes convolution with respect to the spatial frequency variable k_x. The finite extension of a secondary source distribution of length L centered around $x = 0$ can be modeled by a rectangular window $w_R(x_0)$ as

$$w_R(x_0) = \text{rect}\left(\frac{x_0}{L}\right) = \begin{cases} 1 & \text{for } |x_0| \leq \frac{L}{2} \\ 0 & \text{otherwise} \ . \end{cases} \quad (8)$$

The Fourier transformation of $w_R(x_0)$ with respect to x_0 is [5]

$$\tilde{w}_R(k_x) = |L|\frac{\sin\frac{k_x}{2}}{\frac{k_x}{2}} \ . \quad (9)$$

The convolution of $\tilde{D}_{tr}(k_x,\omega)$ with $\tilde{w}_R(k_x)$ adds high frequency components (i.e. artefacts) to $\tilde{D}_{tr}(k_x,\omega)$ and smears it. This results in a limitation of the spatial fine structure of the reproduced sound field. Note that there are indications that this demonstrates a fundamental difference of the capability of WFS to render directional sources compared to Ambisonics [2].

Of course, other window functions may be applied some of which provide potential to reduce truncation artefacts [4]. Nevertheless, they all will exhibit similar qualitative properties leading to the above mentioned consequences.
Concerning spatial sampling we will derive an anti-aliasing condition. One of the authors derived such a condition in [8] for the reproduction of plane waves reading

$$\omega \leq \frac{2\pi c}{\Delta x(1+|\cos\theta_{pw}|)} \ , \quad (10)$$

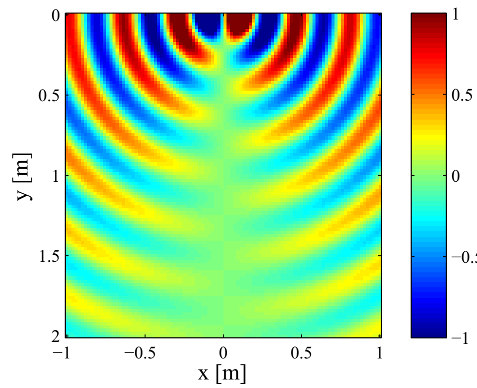
whereby θ_{pw} denotes the angle between the secondary source array and the propagation direction of the plane wave and Δx the distance between the secondary sources. From equation (10) conclusions can be drawn for arbitrary sound fields since similar to the circular harmonics expansion any sound field may be decomposed into plane waves [5]. Note that both expansions can be yielded from one another [6].
For a given virtual source field the anti-aliasing frequency can be found by inserting its plane wave components into equation (10) and looking for the worst case, i.e. the plane wave components where θ_{pw} approaches closest to 0 respectively to π.
Note that although not apparent from equation (10) the anti-aliasing frequency also depends on the location of the listener because for truncated arrays the listener position determines the reproducible range of θ_{pw} [8].
Furthermore, there are indications that also the sampling reduces the reproducible spatial fine structure. In [9] it is shown that a bandlimited sound field has a limited complexity in any circular region in two-dimensional space. Thus, it can be resynthesized by a limited number of secondary sources. Inversely, a limited number of secondary sources (e.g. a truncated sampled array) is then only capable of reproducing a sound field with limited complexity.

6. RESULTS

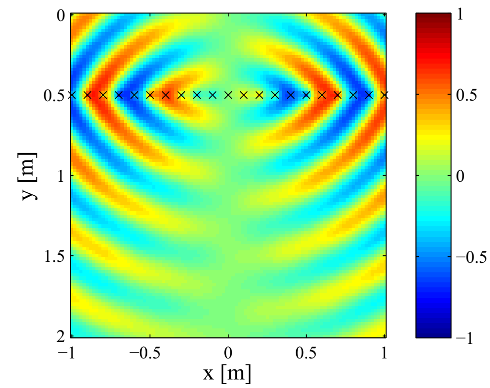
In this section we simulate a sample implementation of a directional source to illustrate the above derived observations.
One of the elementary types of directional sources is a dipole. In this case $\breve{S}^{(2)}(\nu,\omega)$ is given as

$$\breve{S}^{(2)}_{\text{dipole}}(\nu,\omega) = \begin{cases} -\sqrt{k} & \text{for } \nu = 1 \\ \sqrt{k} & \text{for } \nu = -1 \\ 0 & \text{otherwise} \ . \end{cases} \quad (11)$$

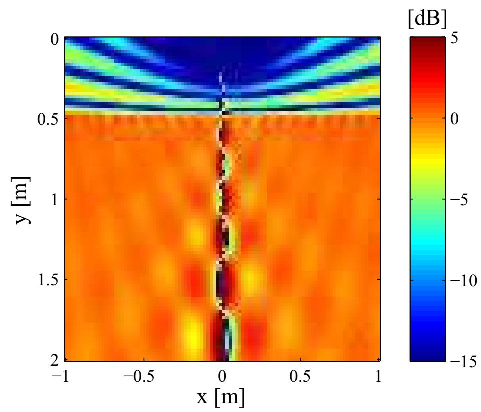
The sound field of such a dipole described by (11) and the sound field of a WFS system rendering it are shown in figure 3(a) respectively 3(b). Here, an array of 40 monopole sources with a spacing of $\Delta x = 10$ cm is modeled. Thus, the array has an overall length of 4 m. The emitted signal is monochromatic with a frequency of 1000 Hz. The marks in figure 3(b) indicate the secondary source positions. Of course, behind the array where our reproduction equations are not valid the two wave fields do not coincide. We find a mirrored version of the sound field of the listening area there. In front of the array a good concordance can be seen. Figure 3(c) further illustrates this. It shows the ratio of the two sound fields in logarithmic scale.
Figure 3(c) reveals that the zero in the dipole's directivity can not be properly reproduced. Note the irregularities around $x = 0$. As described in section 5 these irregularities are consequences of sampling and truncation of the secondary source distribution. To separately identify the contributions of sampling figure 3(d) shows



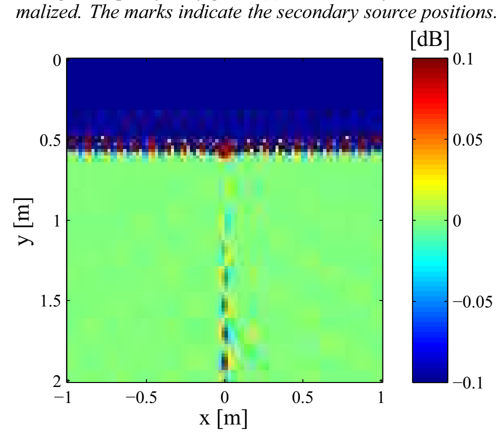
(a) *Wave field $P_{dipole}(\mathbf{x}, \omega)$ of a dipole.*



(b) *Simulated wave field $P_{WFS}(\mathbf{x}, \omega)$ of a WFS array rendering the dipole from figure 3(a). The sound fields are normalized. The marks indicate the secondary source positions.*



(c) *Ratio of $|P_{WFS}(\mathbf{x}, \omega)|$ and $|P_{dipole}(\mathbf{x}, \omega)|$ for an array of 40 loudspeakers. Values are clipped between -15 and 5 dB.*



(d) *Ratio of $|P_{WFS}(\mathbf{x}, \omega)|$ and $|P_{dipole}(\mathbf{x}, \omega)|$ for an array of 1000 loudspeakers. Values are clipped between -0.1 and 0.1 dB.*

Figure 3: Simulation results (linear array at $y = y_0$ rendering a dipole, $\Delta x = 10$ cm, 1kHz).

the ratio of the sound field of an array of 1000 monopole sources with $\Delta x = 10$ cm rendering the dipole from figure 3(a) and the sound field of the dipole itself. Here it can be assumed that the influence of truncation is negligible. The clipping of the color scale in figure 3(d) has been changed compared to figure 3(c) to better allow identification of the artefacts inside the listening area.

7. CONCLUSIONS

An efficient time domain scheme for the implementation of virtual sound sources of arbitrary directivity was presented for two-dimensional wave field synthesis. The approach is based on an expansion of the directional properties of a virtual source into circular harmonics. It was shown that the driving signals for the secondary sound sources can be derived by weighting, delaying, and filtering of the input audio signal.
Truncation and sampling of the secondary source distribution, as occurring in typical installations of WFS systems, limit the reproducible fine structure of a rendered sound field such that e.g. zeros in the directivity are not reproduced exactly. There are indications that this demonstrates a fundamental difference to the ability of Ambisonics to render directional sources [2].

8. REFERENCES

[1] S. Spors, H. Teutsch, A. Kuntz, and R. Rabenstein, "Sound field synthesis," In Y. Huang and J. Benesty, Audio Signal Processing for Next-Generation Multimedia Communication Systems, Kluwer Academic Publishers, 2004.

[2] J. Ahrens and S. Spors, "Rendering of virtual sound sources with arbitrary directivity in higher order ambisonics," *123rd AES Convention, New York, NY, USA, 5th-8th October*, 2007.

[3] E. Corteel, "Synthesis of directional sources using wave field synthesis, possibilities, and limitations," *EURASIP Journal on Advances in Signal Proc.*, vol. 2007, Article ID 90509, 2007.

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

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

[6] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis," in *110th AES Convention*, Amsterdam, Netherlands, May 2001.

[7] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*. New York: Dover Publications Inc., 1968.

[8] S. Spors, "Spatial aliasing artifacts produced by linear loudspeaker arrays used for wave field synthesis," *Second IEEE-EURASIP Int. Symposium on Control, Communications, and Signal Processing, Marrakech, Morocco*, March 2006.

[9] H. Jones, R. Kennedy, and T. Abhayapala, "On dimensionality of multipath fields: Spatial extent and richness," *IEEE ICASSP, Orlando, Florida, USA*, 2002.