PHASE APPROXIMATION OF LINEAR GEOMETRY DRIVING FUNCTIONS FOR SOUND FIELD SYNTHESIS

Paolo Peretti, Laura Romoli, Stefania Cecchi, Lorenzo Palestini and Francesco Piazza

A3Lab - DIBET - Università Politecnica delle Marche
Via Brecce Bianche 1, 60131 Ancona Italy
p.peretti@univpm.it
web: www.a3lab.dibet.univpm.it

ABSTRACT

Driving functions are particular filters, developed in Wave Field Synthesis context, which allow the reconstruction of a desired sound field produced by a virtual source through a loudspeakers array. In this paper a detailed study of linear array driving functions in terms of time and frequency domains behavior is presented. Subsequently, two phase approximations are explored in order to simplify the temporal response and to achieve shorter filters thus reducing the complexity associated with their implementation. Simulations results are presented in order to demonstrate that the approximations do not introduce artifacts in the reproduced sound fields.

1. INTRODUCTION

Wave Field Synthesis (WFS) is a recently developed technique which allows sound fields reconstruction through arrays composed of a high number of loudspeakers. It is based on the use of driving functions: a driving function describes the sound pressure produced by the virtual (desired) source at the position of each loudspeaker. Therefore, the resulting signal, which each loudspeaker is fed with, is a filtered version of the original virtual source signal [1]. The application of these concepts in real scenarios (cinemas, car environment, home theater, etc) needs a real-time implementation. Unfortunately, this kind of processing requires a large computational complexity, especially when adaptive algorithms are applied to it (e.g., AEC [2]). In [3] an efficient solution based on the time invariant preprocessing of the driving functions has been introduced. Linear array geometry is one of the most used in WFS technology: an application to the problem of digital directivity control of the reproduced sound field was introduced in [4].

In literature, the theoretical formulation of driving functions is derived in the frequency domain. From a practical point of view, they can be regarded as FIR filters, which can be realized either in the time domain or in the frequency domain. For the time domain implementation an approximation of the driving function to a weighted integer delay is usually considered [5]. In general, an integer value for the delay does not result in a good approximation, especially when adaptive algorithms are applied to the overall system because it could heavily weigh on the adaptation stability. In this scenario, time and spatial frequencies-domain implementations have to be considered in order to implement the exact driving function with a considerably reduction of the computational complexity [6]. Implementations in frequency domain can be realized through Overlap and Save (OLS) technique.

It is worth underlying that increasing the distance between the virtual sources to be reproduced and the loudspeakers, these filters become longer with a consequent computational cost increase. If frequency implementations are considered, it may be needed to use a higher FFT order which implies a longer audio frame size (higher latency), unless a partitioned block convolution is adopted (higher computational cost) [7]. In [8] the authors suggest the use of fractional delay (FD) to approximate the time delay component in a better way. However, the magninitude of the driving function and its remaining phase contribution, given by a constant phase-shift, are considered and implemented together in an another filter in order to reproduce the nearly exact driving function. The delay of the global transfer function increases due to the additional delay introduced by the second filter, even though, in the case of multiple sources, the computational complexity decreases. In case of low-order IIR implementation for the second filter the delay can be reduced but a phase distortion is introduced [5].

On the other hand, it is possible to achieve a significative compression of driving functions impulse response by omitting the phase shift or by considering it together with the pure delay in the phase approximation problem. In this paper, after a detailed analysis of the time and frequency (magnitude and phase) domains behavior of the driving functions, we will introduce two possible phase approximations derived from the previous considerations: shorter time domain filters excluding the pure delay component can be obtained, thus reducing the computational cost associated to the filtering operations and preserving the exact acoustic image.

In Section 2, after a review of driving functions theory relative to the line array geometry, a study of their behavior in terms of magnitude and phase will be reported. In Section 3 the phase approximations of driving functions and their practical implementations will be presented and in Section 4 some sound fields simulations using the above approximations will be shown. Finally conclusions will be drawn in Section 5.

2. WFS DRIVING FUNCTIONS THEORY

WFS is a recently developed technique which guarantees an optimal acoustic field auralization. It is based on the *Kirchhoff Principle* which is directly derived from *Huygens Principle* and states that *the sound field inside a volume can be calculated if the pressure field and normal particle velocity due to the primary sources on the enclosing surface are known*. The mathematical form of this principle is given by the Kirchhoff-Helmholtz Integral. The idea of WFS is to sample the surface, enclosing the listening area, with an appropriate number of loudspeakers in order to control the acoustic sound field inside this area [1]. A particular choice

 ISSN 2076-1465

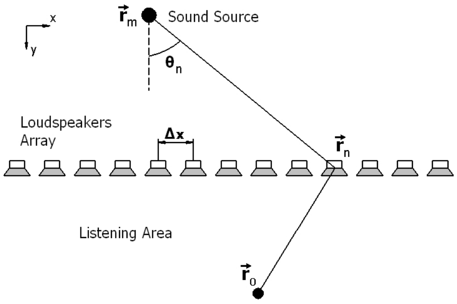
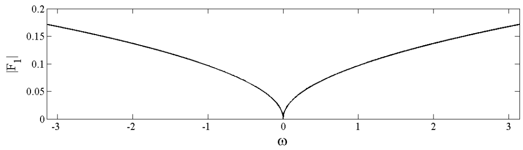
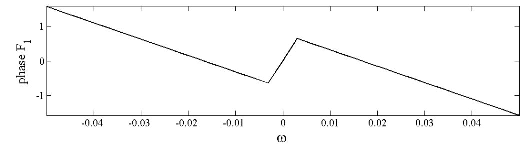


Figure 1: Linear array geometry used for driving functions calculation.



(a)



(b)

Figure 2: Magnitude (a) and a detail of phase responses (b) for a driving function.

of surface allows to define the formulas which permit the reproduction of a monopole source with a linear array of loudspeakers. This surface is constituted by a plane and a hemisphere of infinite radius. In order to find the desired functions, which can be used in a real environment, three approximations have been applied:

- the surface is reduced to a plane curve;
- the line is considered to be of finite length;
- the continuous line is sampled at specific positions.

Starting from the Kirchhoff-Helmoltz integral, the analytical form of each driving function, for a virtual source positioned behind the line array, is obtained in [1] (Fig. 1):

$$Q_n(\omega) = S(\omega)\frac{\cos\theta_n}{D_n(\theta_n,\omega)}\sqrt{\frac{jk}{2\pi}}\sqrt{\frac{|y_0-y_n|}{|y_0-y_m|}}\frac{e^{-jk|\vec{r}_n-\vec{r}_m|}}{|\vec{r}_n-\vec{r}_m|}, \quad (1)$$

where $S(\omega)$ is the sound pressure level of the virtual source to be reproduced, $\vec{r}_m = [x_m\ y_m]$ is its position in the plane, $\vec{r}_n = [x_n\ y_n]$ is the position of the n-th loudspeaker in the plane and θ_n is the angle between the line joining the n-th loudspeaker with the source and the y axis, $\vec{r}_0 = [x_0\ y_0]$ is the center of the listening area, k is the wave number, ω is the angular frequency and c is the sound velocity. It is worth underlying that in (1) the loudspeaker directivity function D_n is considered but, in the following loudspeakers will be viewed as omnidirectional sources. Therefore D_n will be assumed equal to 1. Equation (1) consists in two components: one depends on time while the other one is time independent. In case of non-moving source the first component is just given by $S(\omega)$ [3].

Starting from (1) the sound pressure level at each point of the plane, is obtained:

$$P(\vec{r},\omega) = \sum_{n=1}^{N} Q_n(\omega)\frac{e^{-jk|\vec{r}_n-\vec{r}|}}{|\vec{r}_n-\vec{r}|}\,\Delta x, \quad (2)$$

where $\vec{r} = [x\ y]$ is a point within the listening area, N is the array loudspeakers number and Δx defines the distance between two adjacent loudspeakers.

Therefore, the WFS technique, in the case of line array geometry, becomes the application of N filters, one for each loudspeaker, to the source stream $S(\omega)$ [3]. The filters frequency response $F_n(\omega)$ is given by the time invariant part of (1), weighted by Δx and it can be seen as a product of two functions:

$$F_n(\omega) = F_n^{|\cdot|}(\omega)F_n^{\phi}(\omega), \quad (3)$$

where $F_n^{|\cdot|}$ and F_n^{ϕ} weigh on the module and phase behavior, respectively. The first part is given by

$$F_n^{|\cdot|}(\omega) = \frac{|\cos\theta_n|}{|\vec{r}_n-\vec{r}_m|}\sqrt{\frac{|y_0-y_n|}{|y_0-y_m|}}\sqrt{\frac{|\omega|}{2\pi c}}\,\Delta x \quad (4)$$

because $k = \omega/c$ and $|\sqrt{j}| = 1$. The frequency dependence is given only by $\sqrt{|\omega|}$ so that the module has a symmetric radix behavior (Fig. 2(a)). As regards F_n^{ϕ}, considering both sources behind and in front of the line array, it is given by

$$F_n^{\phi}(\omega) = ae^{-jak|\vec{r}_n-\vec{r}_m|}\sqrt{ja\mathrm{sgn}(\omega)} \quad (5)$$

where $\mathrm{sgn}(\cdot)$ represents the sign function defined as

$$\mathrm{sgn}(x) = \begin{cases} -1 & x<0 \\ 0 & x=0 \\ 1 & x>0 \end{cases} \quad (6)$$

and a is a constant which is equal to 1 or -1 in case of virtual sources behind and in front of the array, respectively. This extension is given by the time reversal approach [9]. It can be seen that F_n^{ϕ} is the product of two terms: the first one is $e^{-jak|\vec{r}_n-\vec{r}_m|}$ which represents a delay of $a|\vec{r}_n-\vec{r}_m|/c$ samples. In case of negative delays the implementation derives from the following complex number property:

$$e^{jx} = e^{-j(2\pi-x)}. \quad (7)$$

Regarding the second term of (5), it can assume three values, depending on a and ω: $\sqrt{j}$, 0 and $\sqrt{-j}$. By using complex number formulas, considering $x \in \{-1,1\}$, it can be found that

$$\sqrt{xj} = \sqrt{e^{x\frac{\pi}{2}j}} = \pm e^{x\frac{\pi}{4}j}. \quad (8)$$

Considering the positive solution, the term $\sqrt{ja\mathrm{sgn}(\omega)}$ represents a constant phase shift of $\pi/4$ or $-\pi/4$ radians, depending on source position. The negative solution leads to the same results shifted by π. Furthermore, it is worth underlying that $F_n^{\phi}(\omega)$ (and consequently F_n) is a symmetric complex conjugate function. The phase-shift could be clearly seen between the first and the second positive frequency bins of Fig. 2(b).

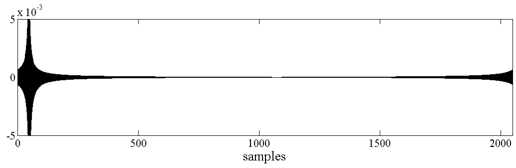


Figure 3: Detail of the driving function in the time domain.

Finally, in Fig. 3 the behaviour of the real part of the time domain driving function is shown. It should be noted its oscillating behaviuor due to not purely real inverse Fourier Transform of the driving function (except for particular values of delay). In the next section this problem will be analyzed in detail.

3. PHASE APPROXIMATION

The exact inverse Fourier transform of $F_n^{\phi}(\omega)$ does not generate a pure real time-domain filter. In fact, this characteristic could be guaranteed forcing to 0 the imaginary part of frequency response at $\omega = 0$ and $\omega = \pi$ [10]. While this constraint is always valid at $\omega = 0$, it is generally not valid at $\omega = \pi$. Forcing the filter phase at $\omega = \pi$ to 0 or π is not the best solution because it involves a phase discontinuity at this frequency. The consequence of this operation is a filter which does not decay smoothly and presents an oscillating behavior in the time domain (Fig. 3). This filter could not be truncated without loss of information.

Therefore, two phase approximation methods for $F_n^{\phi}(\omega)$ will be described. In order to overcome the mentioned problem, both of them aim at obtaining a pure real time-domain filter. Furthermore, it is needed to maintain the magnitude response of the new filter equal to the driving function magnitude for all frequencies and the phase response as close as possible to the phase response of $F_n^{\phi}(\omega)$.

3.1 First Method

The first solution is given by the approximation of $F_n^{\phi}(\omega)$ with a linear phase symmetric FIR filter. Therefore, it is necessary to omit the radix term of (5), which introduces a constant phase shift of $\pi/4$ for all frequencies, as previously seen. Since this phase shift is applied to all signals to be reproduced through the loudspeakers, its exclusion can be considered not to be too significant. Therefore, taking into account a sampling frequency f_s, $F_n^{\phi}(f)$ is approximated by

$$F_n^{\phi}(f) = ae^{-ja\frac{2\pi f}{c}|\vec{r}_n - \vec{r}_m|} \tag{9}$$

since $\omega = 2\pi f$. $F_n^{\phi}(f)$ becomes a simple delay of $d = f_s|\vec{r}_n - \vec{r}_m|/c$ samples for a source behind the line array and a delay of $d = P - f_s|\vec{r}_n - \vec{r}_m|/c$ samples for a source in front of the array, depending on the points P of Inverse Fourier Transform (7). In both case P is necessary greater than $\lceil|\vec{r}_n - \vec{r}_m|/c\rceil$.

In order to obtain a pure real time-domain filter the delay is constrained to be integer. The obtained time domain filter equals 0 except for the d-th sample and the implementation becomes really simple. In all other cases, the time domain filter has also an imaginary part different from 0 and the real part is not so simple as the previous case. It is possible to approximate the delay to the nearest integer, but it is not a good solution because the consequent error is not negligible [5] (see section 4). Therefore, a possible solution is to realize (9) using a FD as in [8]. However, this approach requires an additional filter in order to take into account both $F_n^{|\cdot|}(f)$ and the phase shift, whose frequency responce is

$$H_n(\omega) = h_n\sqrt{j\omega}, \tag{10}$$

where h_n is a constant depending on the driver position (4). This filter can be implemented as a Fractional Hilbert Transformer (FHT) multiplied by $\sqrt{\omega}$ but a further delay is introduced with this implementation [11].

On the contrary the proposed technique is based on the calculation of the whole driving function $F_n(f)$ (considering (9) for $F_n^{\phi}(f)$) at higher sampling frequency Lf_s in order to have a better resolution for the delay, where L is the upsample factor. Therefore, an anti-aliasing filter is needed and a downsampling operation by a factor L is performed to come back to the initial sampling frequency f_s [12]. In this way no additional filter is needed.

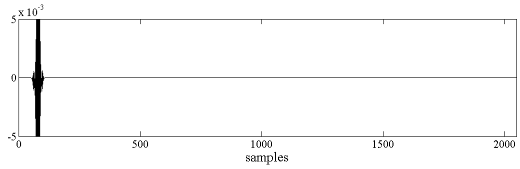
3.2 Second Method

The second type of phase approximation takes also into account the radix term of $F_n^{\phi}(\omega)$. In fact, the first solution could present some issues: in case of multiple sources reproduction it could result in a psychoacoustic sensation different from the target one. Furthermore, sources behind and in front of the array are reproduced with different phase shifts: they have a phase shift of $\pi/4$ and $-\pi/4$ radians, respectively. Moreover, the idea of WFS is to replicate the sound field generated by reference sources with an array of loudspeakers and the previous approximation does not fulfill this constraint.

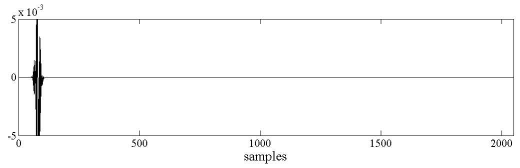
In order to consider the phase shift in addition to the time delay it is necessary to approximate $F_n^{\phi}(\omega)$ with the function $e^{j(-2\pi fd+\varphi)}$, where d and φ are real variables representing the time delay and the phase shift, respectively. In the first method, φ is considered equal to 0. Also in this case a phase approximation is performed. The value of d is the same of the first case, while φ is equal to $a\pi/4$. The approximation for d could be performed at higher sampling frequency, obtaining a single filter as in the first method. After the downsample operation, the frequency-domain driving function is multiplied by $\sqrt{ja\mathrm{sgn}(\omega)}$ in order to perform a phase-shift of φ. However, adding φ at all frequencies produces a filter which has not a null imaginary part in the time domain. Therefore, in order to obtain a pure real time-domain filter, the imaginary part of $F_n^{\phi}(f)$ at frequency bins relative to 0 Hz and $f_s/2$ Hz has to be set to 0. In this case, the last bin could be forced to 0 as the first bin because the driving function has a low-pass behaviour due to the anti-aliasing filter application. It is worth underlying that the second method generates non symmetric filters in the time domain.

3.3 Filter Truncation

Fig. 4 shows the time domain behavior of the driving function in Fig. 3 after applying the two phase approximation methods. The great avantage of this implementation arises from the fact that, in each cases, the power of the filter impulse response is bounded in a small time range. Taking into account that the delay value is well known, an application

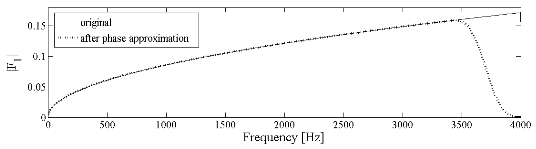


(a)

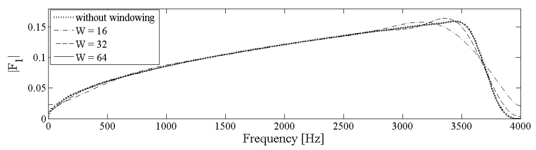


(b)

Figure 4: Detail of the driving function in the time domain after the first (a) and second (b) phase approximation procedures ($L = 16$).



(a)



(b)

Figure 5: Magnitude response at positive frequencies for F_n with and without phase approximation (a) and considering different window lengths (b) ($f_s = 8$ kHz, $L = 16$).

of a window centered in d should be considered. Rectangular window is not a good solution. Extremes smoothing with Hanning or Blackman window leads to better results [10].

After the windowing operation, the driving function could be implemented in a more optimized way. In fact, it is composed of two parts. The first one is a pure delay, which becomes a simple memory shift in a practical implementation, and the other one is a FIR filter. Especially in the case of sources far from the array, the first part of the driving function is typically much longer than the second one. It should be noted that, except for a constant gain depending on the distance between the source and the loudspeaker positions, there are L FIR filters, one for each fractional part, that could be pre-calculated in order to reduce the computational load in case of a real-time application.

4. SIMULATIONS

Some simulations have been performed to evaluate the effectiveness of the proposed approach. Since the differences between the described methods can be found only in the phase term of the driving function, in the following, for the sake of brevity, only the second method will be considered. Driving function filters of length $N_f = 1024$ are considered.

L	MSE [dB]
1	-12.8
4	-25.8
16	-39.3
64	-51.6

Table 1: Mean Square Error (MSE) evaluation by considering different values for L ($f_s = 48$ kHz) with respect to the no-approximation approach. MSE is evaluated in the range between the aliasing frequency of the array (loudspeakers spacing 0.08 m, aliasing frequency 4.2 kHz).

The first test refers to the evaluation of the window length W. Fig. 5(b) shows the magnitude behavior of the driving function without filter truncation and by using a window length of 16, 32 and 64 samples. It can be seen that $W = 64$ is a good approximation for the driving function. Furthermore, 99.9% of the initial power is preserved after the phase approximation with $W = 64$. Using windows of 16 and 32 samples the ratios between the powers of truncated and non truncated filters are 99.7% and 98.6%, respectively. Sound field reconstruction tests confirm that also these window lengths result in a good approximation for the driving function.

The second test refers to the relevance of upsample factor L. Fig. 5(a) shows the introduction of the low pass behavior due to the analysis filter. The appropriate L becomes higher increasing the ratio between the frequency to be reproduced and the sampling frequency f_s: $L = 16$ is good for all frequencies. Fig. 6 shows some examples of sound field due to a sinusoidal monopole source reproduced by a line array. To evaluate the upsampling factor importance, two extremely different values for L are considered: 1 and 16. $L = 1$ represents an approximation by a simple integer delay that can be found in many implementations of WFS [5]. In this case there is a considerable difference between sound fields reproduced with and without phase approximation and a phase distortion can be easily viewed. On the other hand, a good approximation for the driving functions is obtained with $L = 16$. This mistake becomes more visible by increasing the reproduced frequency. It should be noted that in simulations of Fig. 6 a low sampling frequency f_s is chosen in order to emphasize the sound field distortion in case of $L = 1$. However, also in the case of high audio quality sampling frequency, considerable improvements, in terms of Mean Square Error with respect to the no-approximation case, are obtained (Tab. 1). Listening tests will be carried out for a better tuning of the algorithm parameters L and W.

The last test refers to the computational saving achievable with the proposed approach. It is evident that using only one filter, whose length is equal to FD, instead of two filters allows to obtain a lower computational cost [8]. Moreover, in order to analyze the computational load with respect to the frequency implementation, an input signal divided into frame of length F_l is considered. In case of no phase approximation, if $F_l > N_f$ OLS is performed, while partitioned OLS has to be used in the other cases [7]. Otherwise, in case of phase approximation with $W = 64$, an OLS implementation followed by a memory shift is performed. It is worth to underline that the proposed approach permits a considerable decrease of the computational load, especially in case of short frames which are necessary to obtain low latency (Table 2).

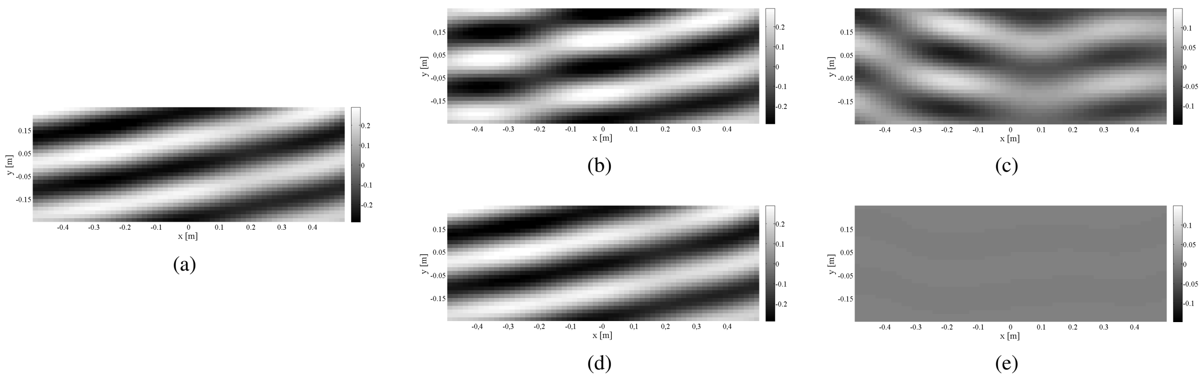


Figure 6: Sound fields of a monopole source ($f = 1.5$ kHz, $\vec{r}_m = [-1,4]$) reproduced with an array ($N = 21$, along x, centered in $\vec{r}_c = [0,2]$). In (a) no phase approximation has been considered. In (b) phase approximation type 2 with $L = 1$ has been considered and (c) shows the difference between (b) and (a). (d) and (e) are obtained with the same procedure of (b) and (c) considering $L = 16$.

F_l	A	B	B/A	F_l	A	B	B/A
64	69	2103	30.58	1024	84	164	1.96
128	66	1037	15.81	2048	91	90	0.99
256	69	567	8.27	4096	117	114	0.97
512	73	298	4.09	8192	109	106	0.97

Table 2: Time elapsed (ms) to filter a 10^6-samples signal with a 1024-taps driving function with (A) and without (B) phase approximation.

5. CONCLUSIONS

In this paper, after a review of driving functions theory, its frequency response has been analyzed. In particular, it has been shown that the phase term is composed by a pure delay and a constant phase shift. Two phase approximations, with the aim of obtaining pure real time-domain filters, have been described: the former ignores the constant phase shift and approximates the group delay to the nearest integer up to a suitably selected upsampling factor; the latter takes into account the phase shift by a proper multiplication in the frequency domain in order to have a pure real filter in the time domain. In this way, extremely shorter time domain filters are obtained, excluding the pure delay components due to the propagation distance, in order to reduce the computational cost associated to their implementation. However, the differences between the approximated and the exact sound images are negligible and this fact becomes important especially in the case of application of adaptive algorithms.

Future works will be oriented toward the extension of the proposed approach to moving sources where driving functions change over time, also considering different loudspeakers array geometries (e.g. circular).

REFERENCES

[1] A. J. Berkhout, D. De Vries, and P. Vogel, "Acoustic Control by Wave Field Synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[2] H. Buchner, S. Spors, and W. Kellermann, "Wave-Domain Adaptive Filtering: Acoustic Echo Cancellation for Full-Duplex Systems based on Wave-Field Synthesis," in *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing. ICASSP 2004*, Montreal, Canada, May 2004, vol. 4, pp. 117–120.

[3] L. Romoli, P. Peretti, S. Cecchi, L. Palestini, and F. Piazza, "Real-Time Implementation of Wave Field Synthesis for Sound Reproduction Systems," in *Proc. of IEEE Asia Pacific Conf. on Circuits and Systems. APCCAS 2008*, Macao, China, Nov. 2008.

[4] L. Romoli, P. Peretti, L. Palestini, S. Cecchi, and F. Piazza, "A New Approach to Digital Directivity Control of Loudspeakers Line Arrays using Wave Field Synthesis Theory," in *Proc. of Int. Work. on Acoustic Echo and Noise Control. IWAENC 2008*, Seattle,WA,USA, Sep. 2008.

[5] E.N.G. Verheijen, *Sound Reproduction by Wave Field Synthesis*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1997.

[6] H. Buchner and S. Spors, "A general derivation of Wave-Domain Adaptive Filtering and application to Acoustic Echo Cancellation," in *Proc. of IEEE Asilomar Conf. on Signals, Systems, and Computers.*, Pacific Grove, CA, USA, Oct. 2008.

[7] W. G. Gardner, "Efficient Convolution without Input-Output Delay," *J. Audio Eng. Society*, vol. 43, no. 3, pp. 127–136, Mar. 1995.

[8] A. Franck, K. Brandenburg, and U. Richter, "Efficient Delay Interpolation for Wave Field Synthesis," in *Proc. of 125th AES Conv.*, New York, NY, USA, Oct. 2008.

[9] S. Yon, M. Tanter, and M. Fink, "Sound Focusing in Rooms: The Time-Reversal Approach," *J. Acoust. Soc. Am.*, vol. 113, no. 3, pp. 1533–1543, Mar. 2003.

[10] A.V. Oppenheim, R.V. Schafer, and J.R. Buck, *Discrete Time Signal Processing*, Pranctice-Hall, 1999.

[11] C.C. Tseng and S.C. Pei, "Design of Discrete-Time Fractional Hilbert Transformer," in *Proc. of IEEE Int. Symp. on Circuits and Systems. ISCAS 2000*, Geneve, Switzerland, May 2000, vol. 5, pp. 525–528.

[12] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, Inc., 1983.