ANALYSIS OF CERTAIN CHALLENGES FOR THE USE OF WAVE FIELD SYNTHESIS IN CONCERT-BASED APPLICATIONS

Terence Caulkins, Etienne Corteel and Olivier Warusfel

IRCAM, 1 place Igor Stravinsky, 75004, France
caulkins@ircam.fr

ABSTRACT

Wave Field Synthesis (WFS) provides a means for reproducing 3D sound fields over an extended area. Beyond conventional audio reproduction applications, present research at IRCAM involves augmenting the realism of concert-based applications in which real musicians will be interacting on stage with virtual sources reproduced by WFS. The stake of such a situation is to create virtual sound sources which behave as closely as possible to real sound sources, in order to obtain a natural balance between real and virtual sources. The goal of this article is to point out physical differences between real sound sources and WFS reproduced sources situated at the same position, considering successively the sound field associated to the direct sound of the virtual source and its interaction with the room. Methods for taking into account and compensating these differences are proposed.

1. INTRODUCTION

Wave field synthesis, or holophony, is a sound reproduction technique based on Huygens' principle that can be seen as the equivalent of holography for acoustic waves. For more complete theoretical background regarding this technique, the reader is invited to refer to [1], [2], [3]. In a previous article [4], the consecutive simplifications that must be operated upon the framework described by the Rayleigh integrals in order to achieve practical implementation of Wave Field Synthesis were listed. The stated simplifications include the reduction of the ideal planar distribution of secondary sources to a line, the truncation of the line, and the subsequent spatial sampling of the line.

Section 2 of the present paper aims to show a difference between the direct sound field radiated by a real source and that of a virtual source reproduced by WFS. This difference is seen to be a consequence of the reduction of the plane of secondary sources to a line, and manifests itself as a dispersion of the virtual source's direct sound field over the listening area. The discussion will be initiated by considering the radiative properties of an ideal line source. The impulse response of this type of source is shown to display dispersive-like properties in the nearfield. In the second part of the discussion, the so-called *stationary phase approximation* is restated and its consequences upon the reproduced sound field are examined. This approximation is shown to lead to similar conclusions concerning the dispersive quality of the reproduced wave field. These theoretical considerations are illustrated by simulations for three distinct reproduction situations.

Section 3 of this paper goes on to study the difference between a real source and a WFS reproduced source in terms of interaction with the listening room. An analysis of the power effectively emitted by an ideal monopole array reproducing a virtual dipole source shows a difference with what is expected of true dipole sources. This entails a modification of the ratio of direct/reverberated sound level throughout the listening area for WFS reproduced sources as compared to real sources.

2. SCATTERING OF THE SOUND FIELD EMITTED BY THE SECONDARY SOURCE ARRAY

2.1. Radiation of an ideal line source

A review of the properties inherent to line sources may help to improve the overall comprehension of the physical properties of an ideal WFS line array before truncation and sampling. Let Ω_n represent an n-dimensional infinite, homogeneous, and isotropic space. It is a well documented fact that a line source in Ω_3 is the physical equivalent of a point source in Ω_2 [5]. The propagative behavior of acoustical waves emitted by a line source in Ω_3 can thus be derived from the Green function associated to a point source in Ω_2.

Furthermore, the 2D Green function g_2 can be deduced from the 3D Green function using Hadamard's "method of descent" [6] and yields the following expression:

$\forall t$ and $t_0 > 0,\ \forall \vec{r}$ and $\vec{r_0} \in \Omega_2,$

$$g_2(|\vec{r}-\vec{r_0}|, t-t_0) = \frac{c}{2\pi\sqrt{c^2(t-t_0)^2 - |\vec{r}-\vec{r_0}|^2}} \times U(c(t-t_0) - |\vec{r}-\vec{r_0}|), \quad (1)$$

where U is the step function and c the speed of sound. This expression signifies that if a point source of Ω_2 situated in $\vec{r_0}$ emits an impulse at t_0, the pressure field received in $\vec{r}$ will consist of an impulse at $t(\vec{r}) = \frac{|\vec{r}-\vec{r_0}|}{c} + t_0$ *followed by a residual field of which the amplitude decreases but remains non-zero over an infinite period of time.* In other words, the wavefront emitted by a point source in Ω_2 is followed by a *tail* or *wake* [7]. This is known as diffusive, as opposed to sharp, propagation: the point (resp. line) source in 2D (resp. 3D) behaves as if it were emitting a wave that propagates simultaneously at all velocities between 0 and c [6]. Moreover, the shape of the wake varies as a function of the reception point $\vec{r}$. The Fourier transform of g_2 yields the following expression:

$$G_2 = -\frac{j}{4} H_0^1(k|\vec{r}-\vec{r_0}|) \quad (2)$$

where H_0^1 represents the cylindrical Hankel function of the first kind of order 0 and $k = \frac{2\pi f}{c}$ represents the wave number.

Figure 1 displays the magnitude of G_2 expressed in dB as a function of the distance $\vec{r}$ to the considered source. This value is normalized by a factor $\sqrt{r}$ to compensate for the $3dB$ attenuation per distance doubling, r being the radius (i.e. distance to the line source). A global $\sqrt{k}$ factor of frequency correction is also included to account for the -3dB per octave magnitude attenuation that occurs in the far field of line sources. One can observe that the corrected frequency response displays very little energy for low frequencies in the near field ($kr \ll 1$), meaning that the "true" (uncorrected) frequency response tends to be flat in the near field. In the far field, the corrected pressure field exhibits a flat frequency response, meaning that the magnitude of the "true" pressure field is attenuated by $-3dB$ per octave. This shows that the global level of bass frequencies increases with the distance to the line source in Ω_3 until it reaches a stable ratio as compared to high frequencies (i.e. $-3dB$ per octave in the far field). The near field of the line source exhibits dispersive-like qualities, even though there is no real variation in sound celerity due to the medium (which is homogenous and isotropic).

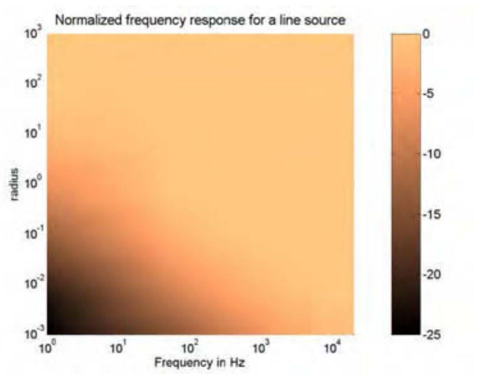

Figure 1: *Frequency response in dB of an infinite line source given at increasing radii and corrected by $\sqrt{kr}$ factor*

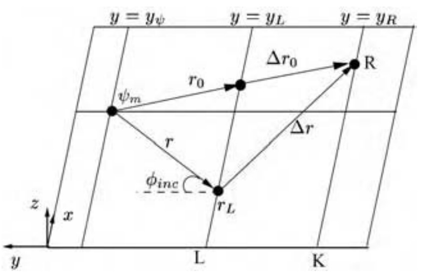

Figure 2: *Diagram describing the $2\frac{1}{2}D$ monopole operator. The wave field of a notional source ψ_m is being reproduced at receiving a position R by an infinite line L of monopole sources.*

2.2. Radiation of an ideal line source as predicted by the stationary phase approximation

The stationary phase approximation is a well-known asymptotic evaluation of the integrals arising in the solution to the wave equation. In the configuration displayed in Figure 2, the notional source ψ_m is separated from the reception point R by an infinite line L of secondary monopole sources. By applying the stationary phase approximation to the Rayleigh I integral [3], the so-called $2\frac{1}{2}D$ monopole operator is derived [8], and yields a measure of the pressure field at the receiving point R:

$$\begin{aligned} P(r_R,k) &= \sqrt{\frac{|y_R - y_L|}{|y_R - y_\psi|}}\sqrt{\frac{jk}{2\pi}}\int_{-\infty}^{+\infty} S(\omega)\cos\phi_{inc} \\ &\times \frac{\exp^{-jkr}}{\sqrt{r}}\frac{\exp^{-jk\Delta r}}{\Delta r}dx_L \end{aligned} \quad (3)$$

where $S(\omega)$ represents the signal fed to the notional source.

This approximation is based on the oscillatory nature of the exponential function. For the needs of this article, let it suffice to say that the accuracy of the approximation increases for large values of k, r_0 and Δr_0.

Noting that $\sqrt{j} = \exp j\frac{\pi}{4}$, a set of driving functions $Q_\psi(x_L, k)$ for the secondary monopole sources situated along line L (represented by the term $\frac{\exp^{-jk\Delta r}}{\Delta r}$ in equation (3)) can now be extracted. This is done for an average listening depth $y_{R_{av}}$, introducing only an amplitude error in the synthesized wave field for receivers on lines $y_R \neq y_{R_{av}}$:

$$\begin{aligned} Q_\psi(x_L,k) &= \sqrt{k}\exp^{-j(kr-\frac{\pi}{4})} \\ &\times \left[\sqrt{\frac{|y_{R_{av}} - y_L|}{|y_{R_{av}} - y_\psi|}}\frac{S(\omega)\cos\phi_{inc}}{\sqrt{2\pi r}}\right] \end{aligned} \quad (4)$$

The term $\sqrt{k}\exp^{-j(kr-\frac{\pi}{4})}$ situated outside of the brackets in equation (4) makes it clear that the driving functions applied to the secondary sources are *frequency dependent* in regard to both phase and magnitude. The term $\exp j\frac{\pi}{4}$ (constant in the frequency domain) can be translated as a "negative delay" of an eighth of a period to be applied to all the frequency components of the source signal in the time domain. Moreso, the $\sqrt{k}$ term implies a magnitude dependence of $+3dB$ per octave. In other words, the stationary phase suggests that low frequencies be emitted with lower levels and in advance in comparison to higher frequencies.

An important specificity of WFS systems is the capacity to reproduce sound sources within the listening area, i.e. "focused" sources. These sources are generated simply by applying a time reversal on the delays suggested by the stationary phase equations. The driving function for such sources is thus equal to:

$$\begin{aligned} Q_\psi^{foc}(x_L,k) &= \sqrt{k}\exp^{-j(-kr+\frac{\pi}{4})} \\ &\times \left[\sqrt{\frac{|y_{R_{av}} - y_L|}{|y_{R_{av}} - y_\psi|}}\frac{S(\omega)\cos\phi_{inc}}{\sqrt{2\pi r}}\right] \end{aligned} \quad (5)$$

Note that for focused source reproduction low frequencies must be emitted *after* high frequencies because of the phase inversion appearing in equation (5).

Aside from virtual point sources, WFS systems allow the reproduction of "plane waves". These correspond to point sources situated at very large distances as compared to the size of the listening area. It is to be remarked that the situation described in Section 2.1, i.e. a line source emitting an impulse, is identical to the emission by an ideal linear WFS array of a "plane wave" propagating perpendicularly to the array. It is therefore clear that dispersive effects will also appear in the sound field of WFS reproduced plane waves.

In any case, the stationary phase approximation proposes a correction of dispersive effects (frequency dependent delays and $\sqrt{k}$ filtering) independently of virtual source and listening area positioning. Although it is clear that dispersive qualities appear in sound fields emitted by infinite continuous line arrays, nothing guarantees that this is also the case for the finite and discrete linear monopole arrays used in Wave Field Synthesis. Simulations in the following Section will allow to decide whether dispersive effects exist in ideal WFS setups.

2.3. Simulations

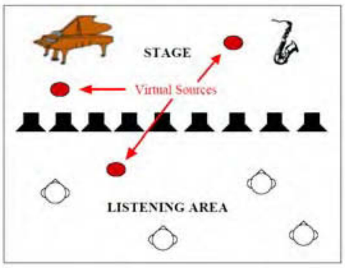

Figure 3: *Typical concert situation in which a WFS system is being used to render virtual sources alongside real instruments.*

The configuration for the following simulations consists of a linear array of 32 ideal monopole sources with regular 16.5cm spacing reproducing sound sources situated on stage as well as inside of the listening area. This is the simulation of a real situation in which the loudspeaker array would be placed between the stage and the audience so as to generate virtual sources to accompany real instruments (Figure 3).

Three WFS sound reproduction situations are chosen to illustrate dispersive effects in the reproduced soundfield.

- **Situation 1:** Reproduction of a virtual source at various distances behind the monopole array (i.e. on stage).
- **Situation 2**: Reproduction of a virtual source 1 meter in front of the monopole array (i.e. in the listening area).
- **Situation 3**: Reproduction of a plane wave travelling perpendicularly away from the array into the listening area.

Situation 1 exhibits a source/array positioning that enters into the theoretical framework set by Huyghens' principle (i.e. primary source situated outside of the listening area). Virtual sources are placed on the perpendicular line running through the center of the monopole array so as to limit windowing effects as much as possible. The resulting soundfield is recorded for different source positions (1m, 5m, 10m and 50m behind the array) on a virtual omnidirectional microphone situated in the listening area. This recording simulation is carried out at 1m and 10m in front of the monopole array. The results are represented in Figure 5. It can be seen that between 100 and 1000 Hz the frequency responses simulated 1m in front of the monopole array tend to be flat ($\pm$ 1dB). In the same frequency band but 10m away, all sources (except for the one 1m behind the array), exhibit frequency responses that become more and more disturbed as the virtual source moves away from the monopole array. Below 100Hz none of the virtual sources exhibit flat frequency responses and the magnitude differences between the two recording positions are maximal.

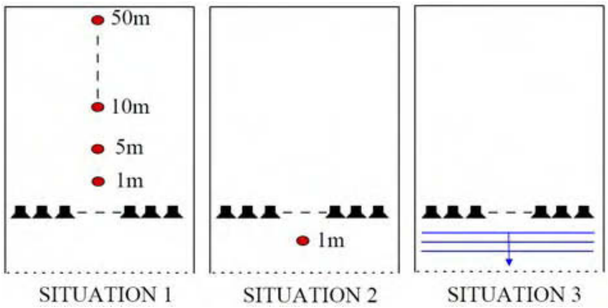

Figure 4: *Situations chosen to illustrate dispersive qualities in WFS sound fields.* **Situation 1** *involves reproducing virtual point sources behind the monopole array;* **Situation 2** *involves reproducing a point source in front of the array;* **Situation 3** *involves reproducing a plane wave.*

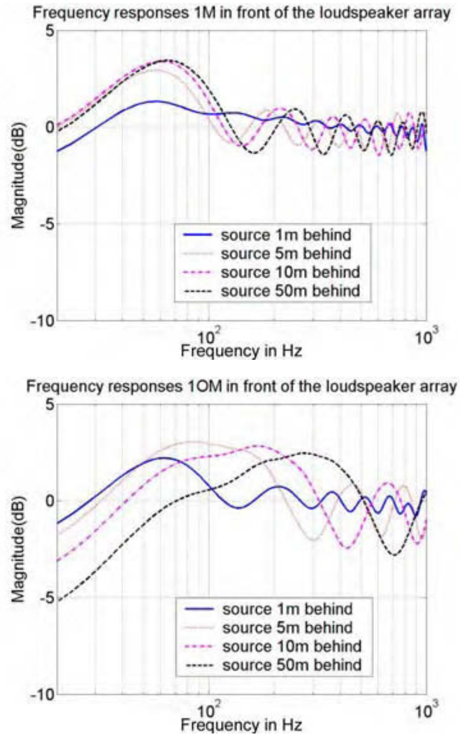

Figure 5: *Frequency responses for various virtual source positions behind the monopole array recorded on a single microphone placed 1m and 10m in front of the monopole array.*

Situation 2 displays the reproduction of a focused source situated 1m inside of the listening area. In order to center the description on the virtual source, the sound field is recorded upon concentric microphone arrays of increasing radii (r = 0.1m, 0.5m and 50m). These arrays are reduced to arcs situated in the listening area as shown in Figure 6 so as to be contained in the visibility window of the focused source [4]. Microphone positions are numbered from 1 to 28 starting from the left extremity of the array. Figure 7 shows the results for this simulation. The first observation that can be made is that in the immediate vicinity of the source

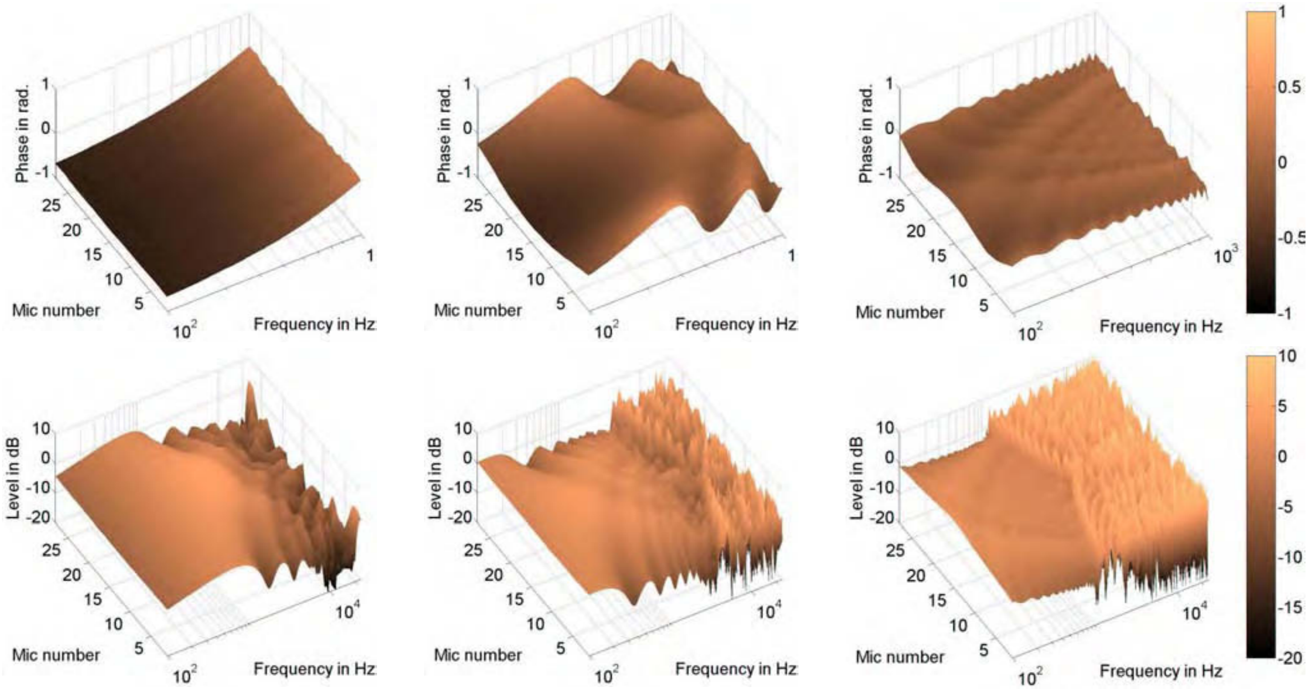

Figure 7: *Phase* (**top figures**) *and frequency* (**bottom figures**) *evolution of a WFS synthesized focused source recorded on concentric microphone arrays situated at different distances* (**left :** *0.1m*, **middle :** *0.5m*, **right :** *10m*) *from the source*)

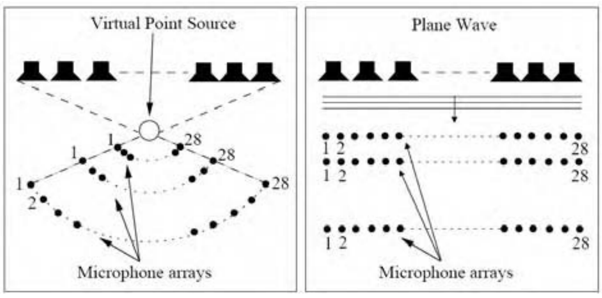

Figure 6: **Left** *: Mic arrays (radii resp. 0.1m, 0.5m, and 10m) recording a focused source situated 1m in front.* **Right** *: Mic arrays recording a plane wave (0.1m, 10m and 50m from array)*

and underneath the aliasing frequency ($\simeq$ 1200 Hz for this setup), the wave field exhibits a +3dB/octave frequency response as well as phase delay for lower frequencies. At this point, the contributions of the entire array arrive simultaneously (by design) and are summed independently of frequency, which means that the $\sqrt{-jk}$ filtering suggested by the stationary phase approximation locally distorts the sound field. As the distance to the source increases, the frequency response flattens out, as does the phase diagram. These observations are well in agreement with the theoretical considerations of Section 2.2. Indeed, the $\sqrt{-jk}$ filtering suggested by the stationary phase approximation ensures a realistic sound field in the far field of the monopole array (r = 10m, f $\geq$ 100).

Situation 3 describes the reproduction of a plane wave propagating perpendicularly to the monopole array into the listening area. The wave field is recorded on linear microphone arrays running parallel to the monopole array as shown in Figure 6. Figure 8 shows the results for this simulation. Phase and frequency responses in the nearfield of the loudspeaker array and below the aliasing frequency are seen to be flat. They tend towards a +3dB/octave frequency response in the far field as well a phase advance for low frequency components. To explain this, one may turn towards linear array radiation prediction techniques. It is a classical approximation to consider that the contributions of all the sources composing a linear array become coherent in the farfield around the perpendicular bisector of the array. The same situation arises at the focal point of virtual sources located within the listening area, causing an inaccuracy in the corrections suggested by the stationary phase approximation (cf **Situation 2**).

2.4. Consequences on practical implementation

Practical implementation of WFS involves the application of a multiequalization scheme to compensate for the complex directivity patterns exhibited by real loudspeakers. For detailed knowledge on this subject the reader is invited to refer to [9].

Application of the multiequalization technique ensures that the reproduced sound field is correct along a certain microphone control line. The multiequalization scheme "automatically" takes into account the scattering effects described and simulated in Sections 2.2 and 2.3 and compensates them. However, for positions situated before and after the control line (in regard to the natural progression of the wavefront) nothing is known a priori about the validity of the wave field. The simulations carried out in the previous part point to the fact that the wave fields for the three types of WFS sources display frequency and phase characteristics that vary during propagation.

This knowledge may prove to be useful when installing WFS setups in large concert halls (where the reproduction zone must be

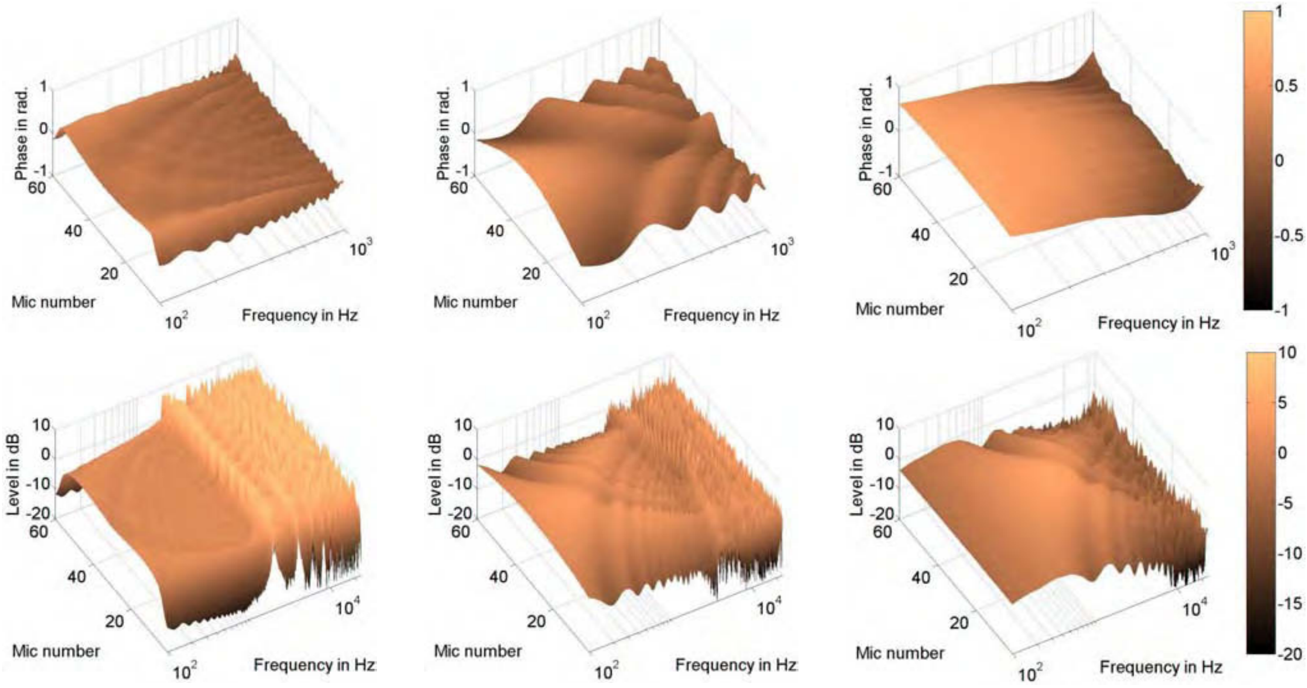

Figure 8: *Phase* (**top figures**) *and frequency* (**bottom figures**) *evolution of a WFS synthesized plane wave recorded on linear microphone arrays situated at different distances* (**left :** *0.1m*, **middle :** *10m*, **right :** *50m*) *from the monopole array*

as large as possible but far away from the loudspeaker system) or, oppositely, when dealing with WFS setups in small rooms (where the reproduction zone is situated near the reproduction system and is reduced in size). For small reproduction rooms, the microphone array upon which the desired wave field is specified may be placed close to the loudspeaker array since the reproduced wavefield will have very little space to disperse over. For larger rooms, the control array of microphones must be placed further away to account for dispersion effects. This suggests the use of different filter banks according to the size of the reproduction room to ensure a realistic sound field over the targeted listening area.

3. DIRECT/REVERBERATED SOUND RATIO FOR DIRECTIVE SOURCES

Rendering realistic spatial impression involves reproducing, beside the direct sound of the virtual source, a coherent room effect, especially when real and virtual sources are mixed together on stage. Contrarily to a classical audio situation where synthetic room effect is rendered in addition to direct sound, the aim of this article is to explore the 'natural' room effect emanating from the interaction between the WFS virtual source itself and the listening room. A priori if the direct sound field reconstructed by WFS were entirely accurate (which is not the case, as was shown in Section 2), the resulting room effect would automatically be entirely accurate. This Section aims to give a measure of the accuracy of the reproduced room effect for WFS. This can be done simply by characterizing the ratio of direct/reverberated sound in the listening room. Section 2 dealt with describing the direct sound field, which can also be piloted using directive sources [4]. The energy density of the reverberated sound field is for its part linked to volume and absorption of the listening room, as well as the power emitted by the source itself. The power effectively emitted by the WFS array when reproducing the ideal properties of sources such as monopoles and dipoles will therefore be calculated so as to compare its behavior in terms of dir/rev ratio with that of real sources. The proposed configuration is a WFS reproduction system made of a linear array of 32 ideal monopole transducers with 16.5cm spacing surrounded by a circular array composed of 64 evenly distributed omnidirectional microphones. The array is used to reproduce a virtual source situated at different positions (in front or behind the WFS array), and associated to various directivity patterns and/or orientations (cf Figure 9).

The first analysis deals with the synthesis of a source situated 3m in back of the array, which represents the case of a source situated on stage. As described in [4], WFS allows for the synthesis of a directivity pattern associated to the virtual sound source. We consider here the case of a dipole pattern simulated with different orientations. Results are shown in Figure 9. It appears that the power emitted by the array for different dipole orientations *varies*. This obviously would not be the case for a real source of which the power level is independent of its orientation in space. Moreso, the right side of Figure 9 shows that the expected ratio of -4.7dB for dipole/monopole power is attained only for certain dipole orientations ($\sim 40^o$ for the source synthesized behind the array and $\sim 45^o$ for the source generated 1m in front of the array).

One reason for which the direct/reverberated sound level in the listening room varies for different dipole orientations that the window through which the source feeds the listening room is limited by the size of the loudspeaker array. A proposition for compensating this windowing effect was proposed in [4]. It involves injecting artificial image sources using extra loudspeaker arrays along the lateral walls of the listening room. This would then naturally modify the ratio of direct/reverberated sound for the different possible

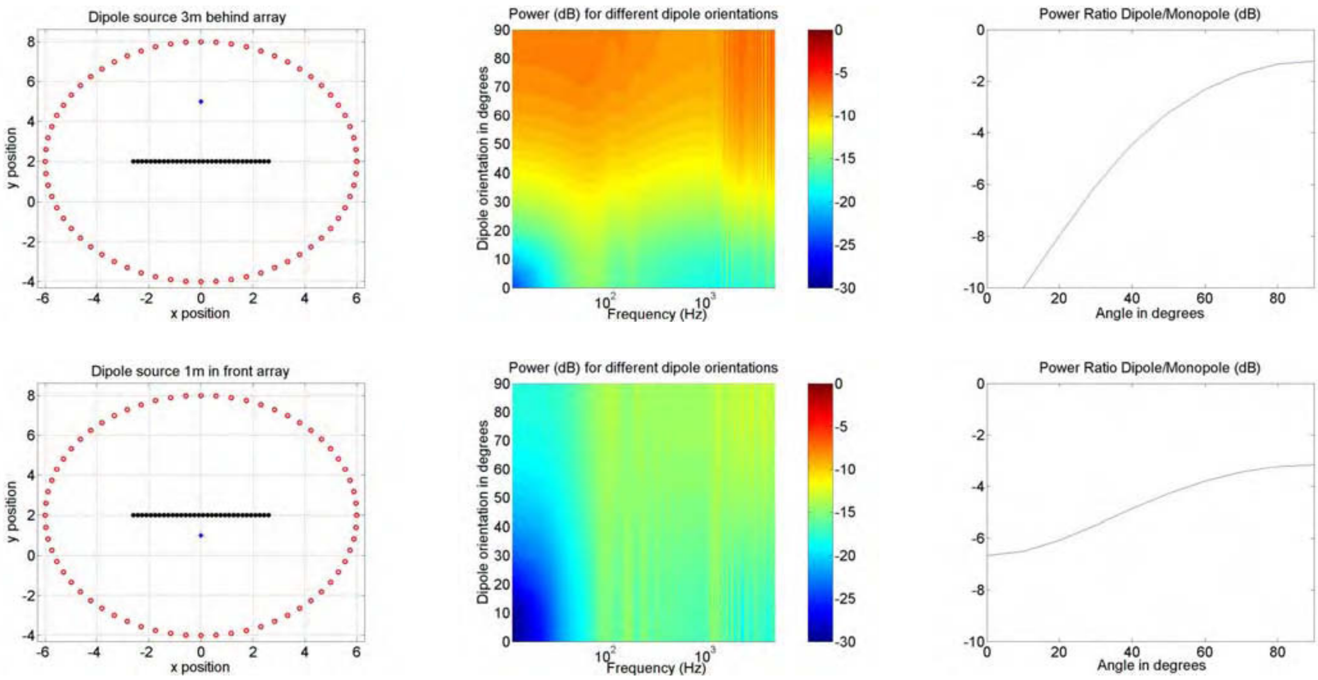

Figure 9: **Left** : *Power measurement configurations, i.e. source 3m behind (top) and 1m in front (bottom). Red dots = microphones, blue dot = virtual source, black dots = monopoles.* **Middle :** *Power emitted by the array reproducing a virtual source situated 3m behind (top) and 1m in front (bottom) as a function of dipole orientation. O^o corresponds to a dipole that displays lobes running parallel to the array; 90^o corresponds to a dipole that is perpendicular to the array.* **Right :** *Ratio of dipole/monopole power integrated from 20-2000 Hz for different dipole orientations, considering a virtual source 3m behind the array (top) and 1m in front (bottom).*

dipole orientations. A less costly solution in terms of calculation power would be to add extra room effect using dedicated room effect channels [10], provided that the power error is negative.

The other reason for which the direct/reverberated level varies lies in the linear nature of the WFS array. The emitted soundfield manifests a symmetry around the axis of the array ("cylindrical propagation"). Directivity patterns or orientations that do not exhibit such cylindrical symmetry will not be accurately rendered in terms of associated power.

4. CONCLUSION

After examining the consequences of the reduction of the plane of secondary sources to a line, simulations were carried out showing dispersive qualities in sound fields emitted by WFS monopole arrays when synthesizing sources situated in front and behind the array, as well as plane waves. This observation led to the conclusion that synthesizing realistic sources in Wave Field Synthesis entails adapting the multiequalization scheme to the reproduction room size, depending on whether we are considering a large hall for concert applications or a small room for virtual reality applications. The next part of the article was committed to studying the ratio of direct/reverberated sound generated by an ideal loudspeaker array when synthesizing basic directivity figures (monopole, dipole). It was shown that this ratio cannot be fulfilled by WFS and depends on the orientation of the virtual sources. Solutions for restoring the perceptual effect linked to room interaction were proposed involving the injection of artificial image sources or reverberation in order to compensate for the observed differences.

5. REFERENCES

[1] A. J. Berkhout, "A holographic approach to acoustic control," *J. Audio Eng. Soc.*, pp. 977–995, 1988.

[2] A. J. Berkhout, D. De Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, pp. 2764–2778, 1993.

[3] E. W. Start, *Direct Sound Enhancement by Wave Field Synthesis*, Ph.D. thesis, TU Delft, 1997.

[4] T. Caulkins, E. Corteel, and O. Warusfel, "Wave field synthesis interaction with the listening environment, improvements in the reproduction of virtual sources situated inside the listening room," in *Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03)*, 2003.

[5] M. Bruneau, *Manuel d´acoustique fondamentale*, Hermes, 1998.

[6] R. Nicol, C. Depollier, and J. Hardy, "Lending an ear to the dimensionality of space," in *ICA*, 1998.

[7] L. Bombelli and S. Sonego, "Relationships between various characterisations of wave tails," 1994.

[8] P. Vogel, *Application of Wave Field Synthesis in Room Acoustics*, Ph.D. thesis, TU Delft, 1993.

[9] E. Corteel, U. Horbach, and R. S. Pellegrini, "Multichannel inverse filtering of multiexciter distributed mode loudspeakers for wave field synthesis," in *AES 112th Convention*, 2002.

[10] E. Corteel and R. Nicol, "Listening room compensation for wave field synthesis: What can be done?," in *Proc. AES 23rd Conference*, 2003.