Acoustics seminar - Wave field Synthesis

Juho Luoto
Aalto University
Master's Programme CCIS / AAT

juho.luoto@aalto.fi

Abstract

This article explores the history and current state of wave field synthesis (WFS). WFS has some properties that enable spatial audio without headphones. WFS can be used for telepresence, art installations, live music and alert systems. The widespread application of WFS is hindered by its cost and complexity. Each installation requires a large number of speakers. Even the smallest arrays can have more elements than a state-of-the-art movie theatre surround system. The advances in manufacturing techniques and innovations from the mobile device industry are likely to make WFS a more viable product in the near future.

1 Introduction

This study reviews the state of wave field synthesis (WFS). WFS has its origins in wave field analysis, which has been used for geological applications. In late 1980s, The application of these principles for acoustic synthesis was discovered. WFS can be used to create realistic spherical, planar and focused waves by varying the boundary conditions of a space using an array of sound transducers. The use of WFS is motivated by its ability to create a physically accurate sound field within a listening space. The technique allows some movement within the listening area while preserving the correct directional cues. In some applications, wave field synthesis could be used as a replacement of headphones and binaural rendering, which could be more comfortable than wearing headphones. Two or more people could collaborate within one sound environment and be able to speak to each other directly without relaying their voice through a microphone. The application of WFS is limited by its high cost, limited usable bandwidth and spatial resolution. I cover techniques for synthesis of wave fields and transducer signals, such as the effect of loudspeaker selection, spacing, distribution and truncation artifacts. This study focuses on the use of horizontal arrays, as those are well covered in the literature and most cost effective for commercial applications. Microsoft Copilot LLM was used in the editing phase to rewrite some paragraphs in more fluent language. A generalized theory of wave field synthesis is described in [1]. There are subsets of WFS which allow the use of more simple speaker setups and source audio while removing some of the special features [2].

2 The mathematical basis

The state of all points inside a source-free volume can be solved if the boundary conditions are known, using the Kirchoff-Helmholtz integral.

$$\oint_{\partial V} G(\mathbf{x}|\mathbf{x}_0,\omega)\frac{\partial}{\partial \mathbf{n}}P(\mathbf{x}_0,\omega) - P(\mathbf{x}_0,\omega)\frac{\partial}{\partial \mathbf{n}}G(\mathbf{x}|\mathbf{x}_0,\omega)\ dS_0 \tag{1}$$

$P(x_0,\omega)$ is the pressure field, $G(x|x_0,\omega)$ is a suitable Green's function. [1] In many applications, WFS is implemented through the Rayleigh Integral, which can be derived from the Kirchoff-Helmholtz integral. [3]

$$P(\mathbf{r},\omega) = \iint_S 2G(\mathbf{r}|\mathbf{r}_0,\omega)\frac{P(\mathbf{r}_s,\omega)}{\partial z} \tag{2}$$

2.1 Virtual sources behind the wall

Driving functions for sources beyond the wall are given as

$$E_n(\omega) = \frac{1}{G}A(\omega)\sqrt{\frac{jk}{2\pi}}\sqrt{\frac{z_s^2 z_r}{z_r - z_s}}\frac{e^{-jkr_n}}{r_n^{3/2}}\Delta x \tag{3}$$

Where z points away from the wall and x-axis is along the (horizontal) array. [4] This technique can be used for room acoustics enhancement. The virtual source would be behind the the speakers. An microphone array can be used to record a sound source and it can be played back by the loudspeaker array. This is possible even in real-time. [5]

2.2 Virtual sources in front of the wall

It is also possible to place point-like virtual sources in front of the wall, as presented in Figure 1. A wavefront is focused on a point in space, and as the front propagates through the focal point, it appears like if it radiated from that point. The Kirchoff-Helmholtz integral restricts real sources inside the volume, but virtual sources are allowed, since they can be constructed with sources outside the volume by focusing. This technique is a form of holography and requires solving an inverse problem. The driving functions can be found using Driving functions for sources in front of the wall are given as

$$E_n(\omega) = \frac{1}{G}A(\omega)\sqrt{\frac{k}{2\pi j}}\sqrt{\frac{z_s^2 z_r}{z_r - z_s}}\frac{e^{+jkr_n}}{r_n^{3/2}}\Delta x \tag{4}$$

[4]

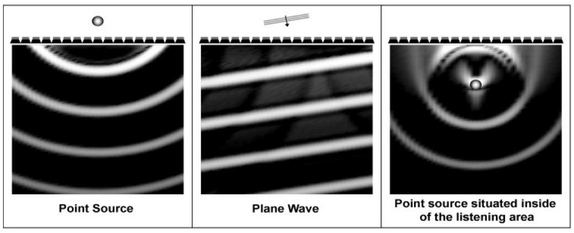

Figure 1: Sources that can be created with WFS [6]

3 Wave field synthesis system components

Various methods for WFS are covered in [2]. The choice of method depends on the desired features of the wavefield. The most complete implementation of WFS is achieved with Huygens arrays. In this case, the listening area is limited to the near-field, while PBAP works in the far-field [2]. Aside from audio applications, Huygens arrays are covered in ultrasound and radio engineering literature. The computational models used in those fields can be useful for acoustic simulations. Ideally, a WFS system can produce virtual sources between the speaker array and the listener [2], [4]. If we relax this requirement, we get planewave-based angle panning (PBAP), also known as far-field WFS. With this method, we are still able to produce virtual sources at and behind the speaker array. This could be useful for stereophonic audio when the listener is far from the speakers compared to the speakers mutual distance. [2]

3.1 Characteristics of the virtual source

The virtual source might be moving. The Kirchoff-Helmholtz integral assumes time-invariant source position. Moving the source causes spectral spreading due to the doppler effect and multipath propagation. This is most noticeable with sinusoidal signals and less noticeable on broadband noises. [7] Fractional delays must be employed even for slowly moving sources, otherwise audible artifacts will result [7]. The source emits a range of frequencies. It might be behind or in front of the loudspeaker array. It can be situated in different positions laterally and vertically. Many WFS implementations use a horizontal array, where rendering vertical position is not possible to do physically, but psychoacoustic methods might be implemented. This is not necessarily a problem because human source localization capability is less accurate in elevation. In the case of source behind the array, channel based audio from a microphone array could be used. [7] WFS sound reproduction requires accurate spatial information of the sound source. This requires applying object-based audio methods for the sound source. An advantage of WFS is its ability to reproduce virtual source radiation patterns [8].

3.2 Transducer array

Ideally, a WFS array would extend to infinity, have infinitesimally small speakers distributed in continuum and have infinite bandwidth. It is common for WFS systems to only extend horizontally, as humans do not perceive sound very well in the vertical axis. 2D WFS theory assumes line sources, but in practice the sources are realized as point sources. This results in faster sound decay over distance [3]. A wave field listening system requires a very large amount of loudspeakers. The highest presentable frequency is limited by the spacing of the array. Small speakers are required for high resolution, but good bass response requires large speakers.
Due to wide human hearing range, WFS is challenging to use in audio applications [1]. WFS-like techniques are used in ultrasound imaging. In that application domain, the bandwidth limitation is not as problematic as in audio applications. Less than one one octave is sufficient, and the frequency range can be selected so that the wave field is rendered accurately.
In audio application, accurate sound reproduction is required over the whole human hearing range, which spans several octaves. It could be beneficial to integrate wave field synthesis with other sound production techniques to simultaneously achieve an impressive soundstage and frequency response. However, this is a complex task, as it would involve combining different types of speakers with crossover filters, potentially leading to phase and amplitude distortions. This approach might be suitable for certain applications. I believe this is very non-trivial, since many speakers would have to be combined with crossover filters, creating phase and amplitude distortions.

In some applications, WFS could be used for select virtual sources, while using conventional sound recreation methods as foundation. There are methods to improve the frequency response. Missing fundamental effect can be used for advantage [9]. This requires some nonlinear processing, but it can be implemented in current DSPs. This technique is used in smart phones. A disadvantage of psychoacoustic bass enhancement methods is those do not work well on all people.

3.3 Array spacing, spatial aliasing and truncation issues

Wavelength along the array is critical for complete reconstruction of the wave field. For large radiation angles, the array elements need to be more densely spaced. [10][11]. With high source frequency, the array needs to be densely spaced. If the density is inadequate, the sound source is not perceived in the correct location. In one study, 46 speakers were placed with an interval of 0.112 meters. This resulted in spatial aliasing frequency of 1500 Hz. [4]
The linear array allows for the easiest analysis, but a logarithmic array is better suited for covering a large range of frequencies. With logarithmic array, the array gain pattern can be made similar at different frequencies. The directivity of a fixed distance array increases with frequency. [12] This causes sound coloration. At low frequencies, arrays suffer from diffraction and near field effects. It is important to

have omnidirectional elements, otherwise the wave gets distorted at large incident angles. [13] The listener is positioned in a source free volume. The observed sound emissions emitted by the secondary sources (the speakers) appear to come from the virtual source(s). [1] The width of the WFS array affects the size of the listening area and the useful range of wave incidence angles that can be synthesized. The listener has to be at a correct distance away from the loudspeaker array. There is a minimum and a maximum distance for spatial sound reproduction. Outside this range artifacts break the listening experience. Minimum distance is set by the lowest frequency due to diffraction and large angle from the speaker movement direction required to focus the beam in front off the listener. The finite width of realizable transducer arrays results in truncation artifacts near the edges of the listening area. Phase distortion occurs where there are no further elements to support the wavefront. The effect is similar to diffraction from an aperture.

3.4 Loudspeaker elements

There are multiple different loudspeaker options to choose from: electrodynamic cone speakers, piezoelectric, Distributed mode loudspeakers (DML) and multiactuator panels (MAP). Electrodynamic cone loudspeaker is a very common transducer most people are likely recognize. It consist of a magnet, a coil and a cone that acts as the radiator. Either the magnet or the coil can be attached to the moving cone. These variations are called moving magnet and moving coil. Attaching the coil to the cone results in lesser moving mass and thus broad frequency response.
Piezoelectric transducers are often used as buzzers and in ultrasonic applications. Constructing broadband loudspeakers from piezoelements is impractical due to their high capacitive reactance. Impedance matching network is required to deliver power efficiently to them.
A DML consists of one or more flat plates that are excited with multiple drivers [9] [14]. The flat shape of these speakers allows them to be used as a video projection surface, which can be desirable in a virtual reality application. [14] The problem with DML is is that their element radiation pattern is very inconsistent over frequency [15].
Multiactuator panels are a type of DML where a large number of transducers is used to excite a single large radiating surface. [13]
It is also possible to utilize multiple DML panels as individual array elements. In this case the elements might be tall and the array elements laid out side-by-side.
Cheap mass manufacturing and miniturization is possible using MEMS technology. MEMS loudspeakers (micro electromechanical systems) are very small speakers manufactured with methods used in semiconductor technology.

3.5 Signal distribution, control and networking

The large number of channels mandates some form of efficient signal distribution system. There are two main approaches: an audio interface with large amount of channels [16], with analog speakers or digital speakers with their own dedicated DACs. In integrated arrays, integrating the DAC into the array assembly seems to be the most practical as the DAC can be optimized for the array application and the amount of external wiring is kept minimum. Class D amplifiers could be useful here, because they can be controlled with pulse-width modulated signal and thus act as DACs.

Dante protocol has been used to send audio over ethernet to a WFS array with integrated audio interfaces [17]. Ethernet cable can carry data tens of meters without interference, which allows freedom in the placement of the control computer. With network based audio, the array can be transported and installed with minimal reconfiguration. The risk of wrong channel order is small.

Wireless speakers could be used as an array. This would require accurate control over the transmission latency between the controlling device and the receivers on the array elements.

4 Psychoacoustics and perceptual considerations

In free-field stereo listening, the illusion of sound angle of arrival is created by using time and amplitude differences for placing the sound sources within the stereo field [18]. The sound source direction is best perceived when both speakers are playing sound. Extreme panning causes the sound to be localized to one of the speakers. The listening area is laterally in the middle of the speakers. Large lateral movement away from the sweet spot causes the sound to be heard from the nearest speaker. The virtual sound sources must always be behind the speakers. Room acoustics affect the stereo listening experience significantly.

Binaural listening with headphones reduces the room acoustic effects to minimum. With closed headphones, the surrounding real environment has no effect. Using HRTFs, virtual sources can be placed around the listener. Listening to a 3D environment requires some form of control over the virtual head orientation. In first-person video games, mouse is used to control the head orientation. This can be done with the assistance of inertial or optical headtracking.

Wavefield synthesis could be used to free the listener from wearing headphones while retaining the envelopment of the spatial audio. With WFS, it is possible to produce an authentic sound field that should, in theory, indistinguishable from a natural soundscape. WFS also allows multiple users to experience the same 3D soundscape side-by-side, as long as the users do not occlude the array [1][19]. This capability has been combined with 3D displays to create a multi-viewer virtual reality room [19]. Stereo cues are intrinsic to the sound field, which means that moving or rotating the head does not require additional processing.

Wave field synthesis suffers from sound coloration, which can be mitigated by some methods proposed by Wittek in [11]. Directional imaging, image focus, locatedness, sound color, listening area, robustness, depth, distance and immersion are some of the psychoacoustic quantities to consider [11].

5 Use cases and applications

WFS has been tested for telepresence applications [20] [21]. The presented linear "Omniwall" uses facial recognition and beam steering to capture the users speech, and this speech is rendered using a panoramic WFS array on the receiving end [20]. Another study used a small array of 8 loudspeakers with array width of 1.6 meters. Kinect was used to capture the location of the person speaking. [21] This kind of system could be helpful to distinguish the speaker in group video meetings where multiple people are colocated.

The most prominent application of WFS are in art installations and telepresence. Researching commercial products is cumbersome, because a different term or trademark might be used to refer to WFS as a concept. Examples of art installations are readily available by use of search engines. WFS can be used in the living room [22] and movie theatres [23]. WFS with moving sources can be applied to electronic music shows [16]. One experiment conducted in 2003 used 24 speakers controlled by a RME Hammerfall audio interface. The array control was implemented using FIR filter software running on a Linux PC. This system was capable of playing nine moving sources in real-time. Part of this project was a software for creating sound source paths for the WFS system. The user could define the virtual source positions in space, and the software computed the required filters for controlling the sound field. [16] Considering that the article was released 20 years ago, it can be assumed that computing power is not a relevant bottleneck for PC based installations. This application has potential for interesting abstract spatial effects. The source media can be crafted to use the strengths of WFS systems and while circumventing its flaws. For instance, WFS could be used for mid frequency effects, while bass can be produced with subwoofers as usual.

WFS has been used for sound art installations. Miles Scharff has created an installation called 'Sound Hologram'. The Spat 5 MAX/MSP sound spatialization library from IRCAM has been used for this purpose [17]. The library designs the filters based on desired source positions. The array consists of 64 individually addressable speakers controlled over ethernet with Dante protocol [17]. The installation is located in an indoor garden.

6 Discussion

WFS setups can be expensive due to the large amount of loudspeakers and playback channels required, but the emergence of MEMS speakers will make loudspeaker arrays more affordable. Advances in Bluetooth audio could make it possible to create arrays where all array elements are independent wireless speakers.

In near-field applications where the source signal can be limited in bandwidth, such as alarm sounds, WFS can be an useful technique. In a vehicle, the alarm sound could direct attention to the right direction. A desktop workstation display could have a attached sound bar for wave field synthesized sound.
Currently there mobile phones with two speakers. More MEMS speakers could be added to a phone for more complex sound experience via WFS. The reverse process, wave field analysis, can be used for sound localization and spatial source separation. The possibility of using a DML and integrating the WFS sound system into a wall could be used in a entertainment facility.

7 Conclusion

Wave field synthesis is a method for recreating a realistic soundfield inside a listening room. It has potential to increase immersion in virtual reality environments. The method is based on the premise of a continuous boundary condition. In practice, the boundary must be discretized. The layout of the transducer array must be considered carefully. The choice of loudspeaker drivers affects sound coloration, localization and frequency response.
WFS is not commonly used, so there is a lack of standards and formats. In many applications, more traditional spatial audio solutions, like binaural playback and 7.1 surround sound are more practical. WFS has found use in sound installations. In this application, the composition can be created around the sound system.
This article was written based on a literature study. Proper understanding of the capabilities and limitations of WFS would require computer simulations and experiments with a real WFS array and associated control software. For research and educational purposes, knowledge and practice of use of array speakers would be insightful.

References

[1] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. of the 124th AES convention*, (Amsterdam, The Netherlands), May 2008.

[2] J. Smith, "A spatial sampling approach to wave field synthesis: PBAP and huygens arrays," Nov. 2019.

[3] K. Furuya, Y. Haneda, Y. Hiwasaki, and S. Koyama, "Inverse wave propagation in wave field synthesis," in *Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space*, Oct 2010.

[4] M. Boone, E. Verheijen, and G. Jansen, "Virtual reality by sound reproduction based on wave field synthesis," May 2000.

[5] D. de Vries, "Sound reinforcement by wavefield synthesis: Adaptation of the synthesis operator to the loudspeaker directivity characteristics," *JAES*, vol. 44, no. 12, 1996.

[6] "WFS." http://recherche.ircam.fr/equipes/salles/WFS_WEBSITE/Index_wfs_site.htm. Accessed: 2024-05-9.

[7] A. Franck, A. Graefe, T. Korn, and M. Strauss, "Reproduction of moving sound sources by wave field synthesis: An analysis of artifacts," in *Audio Engineering Society Conference: 32nd International Conference: DSP For Loudspeakers*, Sep 2007.

[8] T. Ziemer, "A psychoacoustic approach to wave field synthesis," in *Audio Engineering Society Conference: 42nd International Conference: Semantic Audio*, Jul 2011.

[9] J. J. L. Basilio Pueo, German Ramos, "Strategies for bass enhancement in multiactuator panels for wave field synthesis," *Applied Acoustics*, vol. 71, pp. 722 – 730, March 2010.

[10] E. Start, *Direct sound enhancement by wave field synthesis*. PhD thesis, Technical University of Delft, 1997.

[11] H. Wittek, *Perceptual differences between wavefield synthesis and stereophony*. PhD thesis, Department of Music and sound recording; School of Arts, Communications and Humanities. University of Surrey, 2007.

[12] M. van der Wal, E. W. Start, and D. de Vries, "Design of logarithmically spaced constant-directivity transducer arrays," *J. Audio Eng. Soc*, vol. 44, no. 6, pp. 497–507, 1996.

[13] E. Corteel, K.-V. NGuyen, O. Warusfel, T. Caulkins, and R. Pellegrini, "Objective and subjective comparison of electrodynamic and map loudspeakers for wave field synthesis," 2007.

[14] M. Boone and W. de Bruijn, "On the applicability of distributed mode loudspeaker panels for wave field synthesis-based sound reproduction," May 2000.

[15] E. Corteel, "Synthesis of directional sources using wave field synthesis, possibilities, and limitations," *EURASIP J. Adv. Signal Process*, vol. 2007, p. 188, jan 2007.

[16] M. Baalman, "Application of wave field synthesis in the composition of electronic music," 01 2003.

[17] "Sound hologram." https://www.milescharff.com/installations/soundholo-j3kpj. Accessed: 2024-05-9.

[18] V. Pulkki, *Communication Acoustics: an introduction to Speech, Audio and Psychoacoustics*. Wiley, 2015.

[19] J. Springer, C. Sladeczek, M. Scheffler, J. Hochstrate, F. Melchior, and B. Frohlich, "Combining wave field synthesis and multi-viewer stereo displays," in *IEEE Virtual Reality Conference (VR 2006)*, pp. 237–240, 2006.

[20] G. Heinrich, C. Jung, V. Hahn, and M. Leitner, "A platform for audiovisual telepresence using model- and data-based wave-field synthesis," in *Audio Engineering Society Convention 125*, Oct 2008.

[21] M. Gasparini, S. Cecchi, L. Romoli, A. Primavera, P. Peretti, and F. Piazza, "Kinect application for a wave field synthesis-based reproduction system," in *Audio Engineering Society Convention 133*, Oct 2012.

[22] B. Klehs and T. Sporer, "Wave field synthesis in the real world: Part 1 - in the living room," in *Audio Engineering Society Convention 114*, Mar 2003.

[23] T. Sporer and B. Klehs, "Wave field synthesis in the real world: Part 2 - in the movie theatre," in *Audio Engineering Society Convention 116*, May 2004.