IMPLEMENTING WAVE FIELD SYNTHESIS IN AN ITU SPEC LISTENING ROOM PART 1: KEEPING IT AT EAR LEVEL

R G Oldfield University of Salford, Salford, UK
J A Hargreaves University of Salford, Salford, UK
I A Drumm University of Salford, Salford, UK
A T Moorhouse University of Salford, Salford, UK

1 INTRODUCTION

Listening rooms are designed to be acoustically neutral spaces with acoustic conditions approximating a typical domestic room. As such, they are ideally suited for perceptual testing of audio systems. This paper describes the permanent design, installation and commissioning of a wave field synthesis (WFS) system in the listening room at the University of Salford. The room meets the standards set out in ITU-R BS 1116-1[1], BS 6840-13 and IEC 268-13[2] and it was important that the new installation did not affect the effectiveness of the acoustic treatment or the room's accreditation.

WFS is becoming an increasingly popular multichannel spatial audio technique; providing a spatially accurate sound field made up by the superposition of signals from multiple loudspeakers around the boundary of a listening space. Using such a system in a non-anechoic room presents special challenges, since the inevitable reflections interfere with the rendered sound field; hence this system is intended as a test-bed to prototype and evaluate the effectiveness of new techniques to make WFS more accurate and realistic in real rooms. The installed system presented here consists of 112 axi-concentric loudspeakers powered by fourteen 8-channel power amplifiers. The loudspeakers are mounted on a rectangular frame that can be raised and lowered on a winch such that they can be adjusted to the height of the listener's ears or can be moved out of the way when not in use, thus not affecting the tightly controlled acoustic of the room.

2 WAVEFIELD SYNTHESIS THEORY

Wave field synthesis (WFS) is a multi-channel spatial audio technique that was first formally suggested by Berkhout[3,4] in the late 1980's. WFS uses a large number of loudspeakers to completely (re)create a desired sound field across the entire listening space. By doing this, all of the localization cues are available to the listener as if the sound source itself were actually there. Thus WFS is a volumetric solution for spatial audio which means there is no associated 'sweet spot' in the room corresponding to a position optimized for correct localization as is the case for many standard systems such as stereo, 5.1, ambisonics etc.

The theory of wave field synthesis is essentially an application of the Huygen's Effect which states that each point on a wave font can be considered as the beginning of a new point source. An infinite number of these point sources distributed along the wave front allow it to be completely recreated by the secondary sources. Wave field synthesis, therefore, attempts to reconstruct any sound field using many secondary sources (loudspeakers). Conceptually, if these loudspeakers are fed with different amplitudes and delays then an arbitrary sound field within the enclosed space can be accurately recreated. This theory can be expressed mathematically by the Kirchhoff-Helmholtz equation (1), with the geometry given by Figure 1:

$$P(\mathbf{x},\omega) = -\oint_S \left(G(\mathbf{x}|\mathbf{x}_0,\omega)\frac{\partial}{\partial \mathbf{n}}P(\mathbf{x}_0,\omega) - P(\mathbf{x}_0,\omega)\frac{\partial}{\partial \mathbf{n}}G(\mathbf{x}|\mathbf{x}_0,\omega) \right) dS \tag{1}$$

Where **n** is the inward pointing unit vector normal to the surface, *S* at the position $\mathbf{x}_0$ relative the origin, 0. This equation states that if the pressure and its derivative (the velocity normal to the surface) are known at each point on the surface, then the pressure *P* at any point in the volume, *V* can be determined, that is the wave field in the volume can be determined.

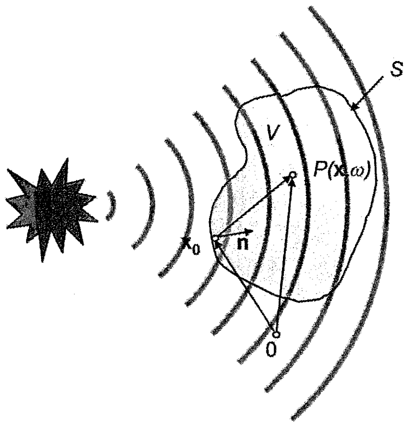

Figure 1. Geometry used for the Kirchhoff-Helmholtz equation

Although the strict application of this equation requires a continuous distribution of an infinite number of monopoles and dipoles on the surface to reproduce the sound field in the volume, it is possible to make some simplifications and to derive driving functions for a limited number of loudspeakers placed on the surface of the volume instead. With these simplifications come some amplitude and spectral errors in the reproduced wave field and a low frequency limit of correct sound field reproduction; however reproduction is still considered perceptually accurate[5]. Further information on these errors and a more detailed overview of the derivation of the WFS driving functions can be found in Spors *et al*[6].

Principally one can derive 2 major driving functions for the WFS system corresponding to the rendering of point sources and plane waves. A third source type corresponding to a point source in front of the loudspeaker arrays (known as a focused source) can also be derived from the standard point source driving function using time reversal techniques. Examples of the wave field produced by these driving functions are given in Figure 2.

The driving function for a linear array of secondary sources to render a virtual point source can be given as:

$$D_{PS}(\mathbf{x}_0,\omega) = -2\sqrt{\frac{2\pi|\mathbf{x}_{\mathrm{Ref}} - \mathbf{x}_0|}{j\frac{\omega}{c}}}\,\frac{(\mathbf{x}_0 - \mathbf{x}_S)^T\,\mathbf{n}(\mathbf{x}_0)}{|\mathbf{x}_0 - \mathbf{x}_S|^2}A(\omega)\,e^{-j\frac{\omega}{c}|\mathbf{x}_0-\mathbf{x}_S|}\left(\frac{1}{|\mathbf{x}_0 - \mathbf{x}_S|} + j\frac{\omega}{c}\right) \tag{2}$$

Where $A(\omega)$ is the source spectrum, $\mathbf{x}_{Ref}$ is the reference point (usually the centre of the room), where spectral and amplitude errors are minimum, $\mathbf{x}_0$ is the loudspeaker position, $\mathbf{x}_s$ is the primary (virtual) source position, $\mathbf{n}(\mathbf{x}_0)$ is the direction vector of the loudspeaker and all other terms have their usual meanings.

For a plane wave source the driving function can be given as:

$$D_{PW}(\mathbf{x}_0,\omega)=\sqrt{2\pi\left|\mathbf{x}_{\mathrm{Ref}}-\mathbf{x}_0\right|}\times\sqrt{j\frac{\omega}{c}}A(\omega)\,\mathbf{n}_{PW}^{T}\mathbf{n}(\mathbf{x}_0)e^{-jk\mathbf{n}_{PW}^{T}\mathbf{x}_0} \tag{3}$$

In this case $\mathbf{n}_{PW}$ is the direction vector of the plane wave.

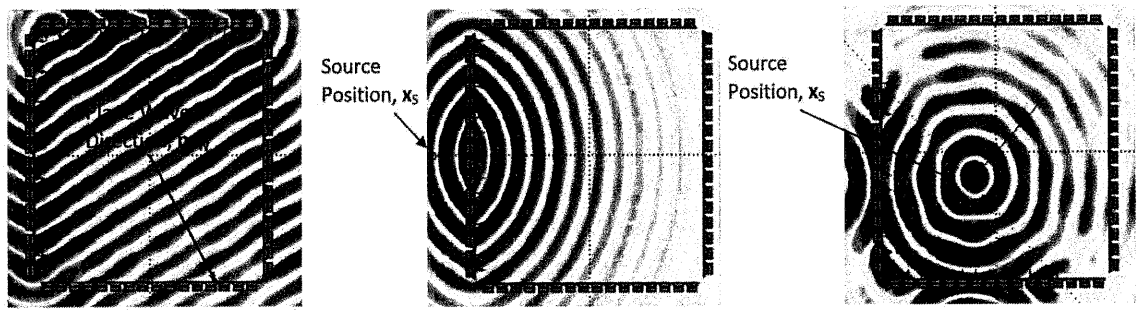

Figure 2. Examples of sources rendered using wave field synthesis: (a) A virtual plane wave, (b) A virtual point source, (c) A focused source

As can be seen from Figure 2, the wave field of each major source type is accurately rendered over the entire listening space so the spatial impression and localisation of sources is almost independent of listener position. This key feature of WFS, along with the accurate creation of plane waves and focused sources, are the major advantages of a WFS system over other spatial audio techniques. It is worth noting here that the Kirchhoff-Helmholtz equation (1) does not restrict sound from travelling out of the listening area and that the approximations used when the distributions of monopoles and dipoles were replaced by an array of loudspeakers cause sound to be radiate backwards as well as forwards from the array. The implications of this will be discussed further in the following sections.

3 SYSTEM DESIGN

When designing a WFS system it is important that an appropriate number of secondary sources are used. Theoretically there is no limitation on the number of loudspeakers, but practical factors (such as the number of audio channels and space requirements) limit how many can be used in reality. If too few loudspeakers are used the reproduced wave field will exhibit spatial aliasing artifacts[7], which can cause source colouration and localization errors at higher frequencies. The frequency at which spatial aliasing occurs is a function of the distance between secondary sources (loudspeakers) and is traditionally formulated as *(4)*.

$$f_{al}=\frac{c}{2d} \tag{4}$$

Where d is the distance between the loudspeakers in the array and c is the velocity of sound in air. It should be noted that Equation *(4)* describes the worst case scenario, where the direction of source propagation is normal to the loudspeaker array; as the source angle decreases the spatial aliasing frequency will increase. Typical WFS systems have a loudspeaker spacing of ~10-30cm (f_{al}

= 600Hz – 1.7kHz) and it is widely considered that within these limits the artifacts are perceptually inconsequential as the human auditory system is not too sensitive to spatial aliasing. However, to increase the frequency at which these artifacts occur in this system, the loudspeaker spacing was made as small as possible, such that the rims of the loudspeakers were almost touching one another. Using smaller drivers would reduce the loudspeaker spacing but result a system with poor low frequency performance, so a compromise was met by using loudspeakers with a 100mm driver as described in the next sub-section.

3.1 Loudspeaker Hardware

The loudspeaker system was designed to be modular in nature. Lightweight 6mm plywood enclosures were constructed which each contained eight loudspeakers and these were clamped directly onto the supporting frame, which was constructed from 48.3mm steel pipe. Fourteen enclosures were arranged in a rectangle four enclosures deep and three wide, using 112 loudspeakers in total (as shown in Figure 4). The loudspeakers themselves are KEF CR100QR, which comprise a 100mm woofer and a coaxial 19mm tweeter in a sealer 0.7 litre enclosure which also contains the passive crossover (at 3kHz). Originally designed for in-ceiling installation, they are ideally suited to creating WFS arrays since they include mounting hardware and their individual enclosures eliminate coupling through the enclosure volume. Individually they reproduce 110Hz to 27kHz ±6dB and can be fitted with separation of just 12.95cm, giving an aliasing frequency of approximately 1.33kHz.

The design brief required that the system could be used at a range of different listening heights, from seated to standing, and should not compromise the acoustics of the room for other uses. In response to this the innovative solution of suspending the loudspeaker on a winch system was adopted. This was straightforward because the room had been constructed with a concrete slab ceiling for acoustic isolation purposes and a flying system of the type most commonly seen in theatre lighting grid systems was designed and installed by a certified contractor according to LOLER regulations. The steel pipe was supported at four hanging points which were diverted by a system of pulleys back to a single ratcheted hand winch. The loudspeakers can be positioned at any height from 0.7m upwards and winched to within a few inches of the ceiling so that they are roughly flush with the existing acoustic treatment and both visually and acoustically unobtrusive. Cabling was routed around the steel ring and slung across to a strain relief bar (located so as to avoid the cable's weight pulling the flying frame askew) and finally through glands on a plate on the wall furthest from the door. Photos and a plan view of the system are shown in Figure 3 and Figure 4 respectively.

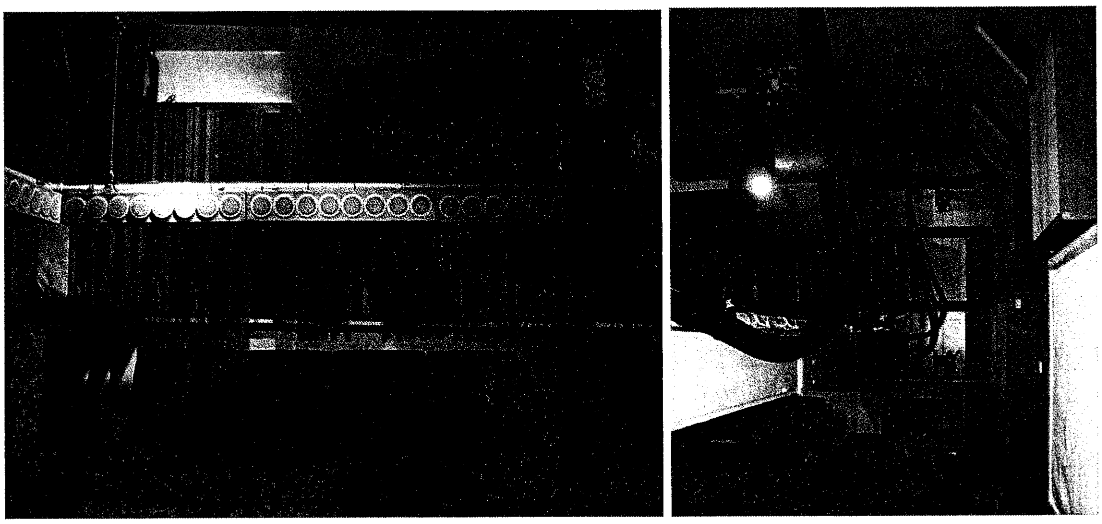

Figure 3. Loudspeakers suspended on winch system in listening room

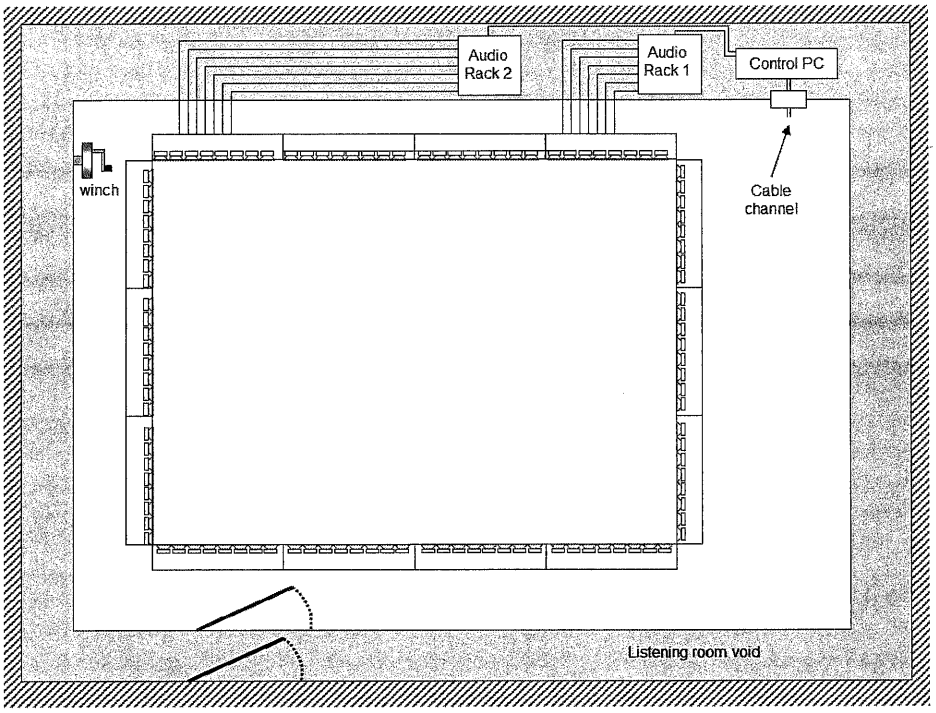

Figure 4. Plan view of listening room with WFS install

3.2 Audio hardware and amplification

The control PC, audio D to A hardware and amplification are all located in the void around the side of the room as shown in Figure 4. The control PC is fitted with 2 RME HDSPe MADI soundcards each of which is capable of outputting 64 channels of audio in MADI (Multi-channel Audio Digital Interface) format. Each MADI signal is then fed via coax into two daisy-chained RME M-32DA converters which each convert 32 channels of the MADI signal into 32 audio outputs. These signals in turn feed Cloud CX A850 8-channel power amplifiers. The loudspeaker cables (each one carrying 8 channels) are fed through the wall to the loudspeakers. Overall system level is set by calibrating the amplifier gain using a dummy load.

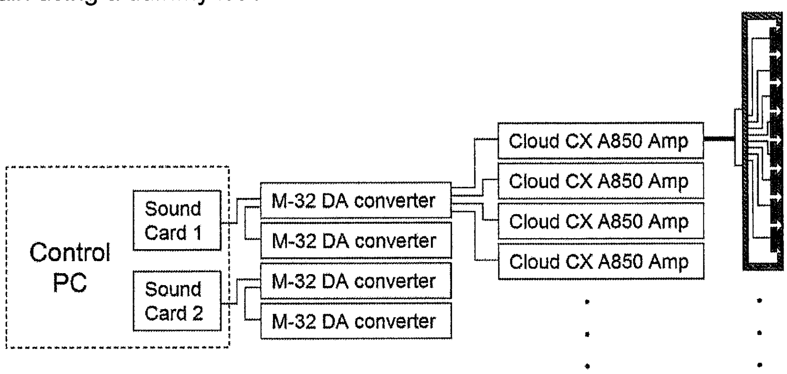

Figure 5.Audio hardware diagram

3.3 Operation of the system

The system is controlled by the PC in the listening room void as per Figure 4. The control software consists of a series of multi-channel audio APIs as libraries that can be accessed easily from higher level programming languages making it easy to try out different techniques and algorithms. A Matlab implementation of the WFS render engine has also been written to enable easy algorithm development. A variety of different source types either static or moving can be implemented simultaneously to enable modelling of complex audio scenes.

4 COMISSIONING THE SYSTEM

The WFS system was commissioned using a linear array of 16 Rode NT55 condenser microphones with an inter-microphone spacing of 9cm, positioned in the middle of the room. With a 4ms Gaussian pulse as the test signal a series of spherical and plane wave sources were outputted though the system and the impulse response measured at each of the 16 microphone positions. Looking at the multiple impulse response traces in the time domain shows the measured propagating wave front (see Figure 7). This measured response can then be compared to the analytical solution of point sources and plane waves as shown in Figure 6.

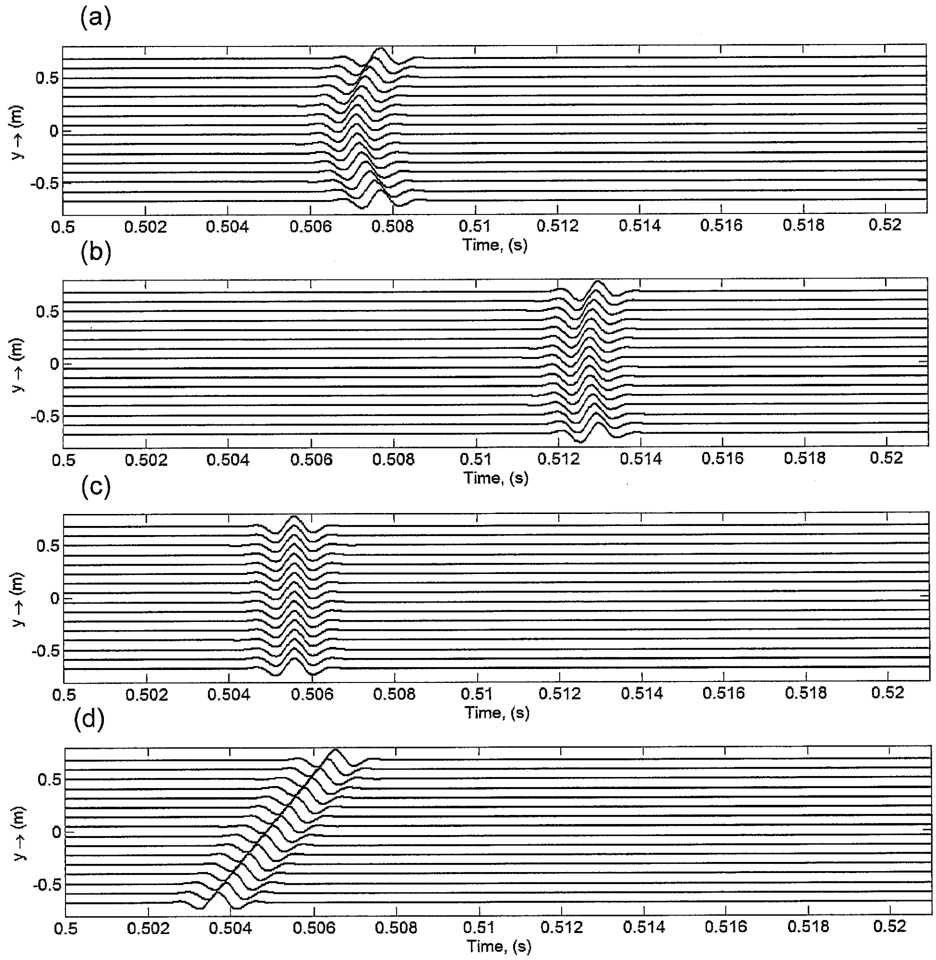

Figure 6. Multiple impulse response plots for the analytical solution of: (a) Point source at [-0.5m, 0m], (b) Point source at [-2.5m, 0m], (c) Plane wave at 0° and (d) Plane wave at 45°

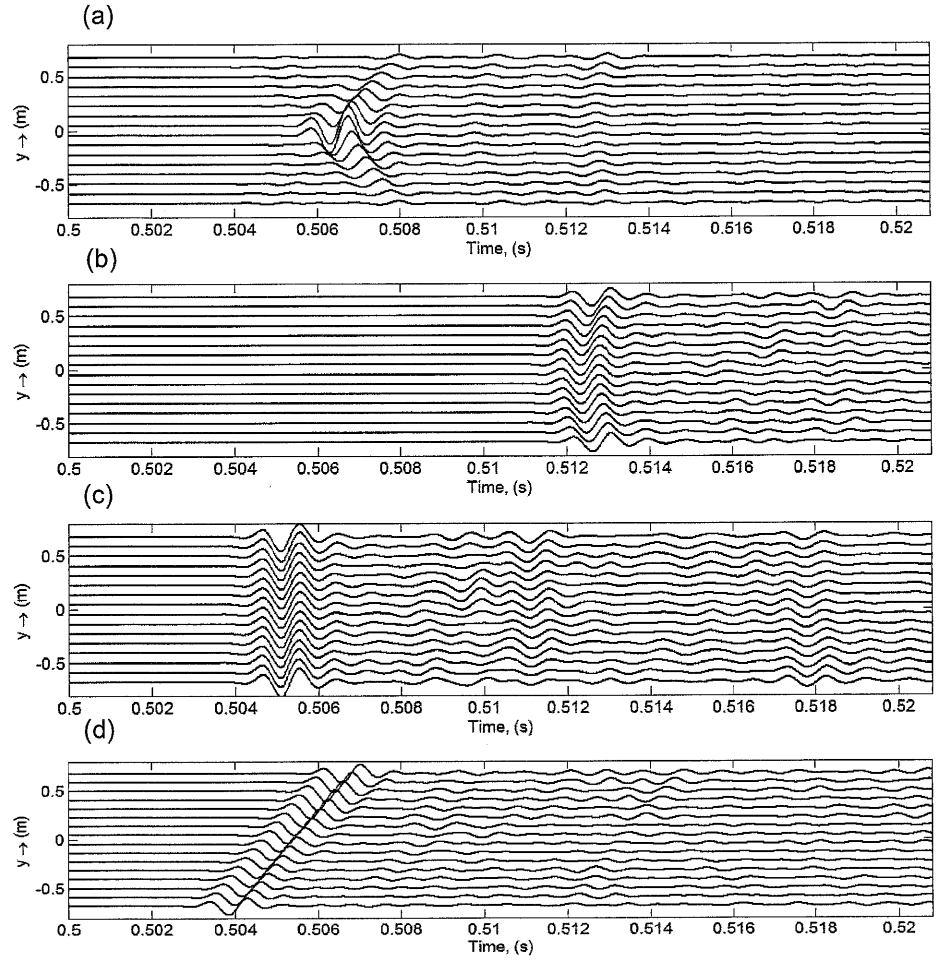

Figure 7. Multiple impulse response plots for measured WFS reproduction of: (a) Focused point source at [-0.5m, 0m], (b) Point source at [-2.5m, 0m], (c) Plane wave at 0° and (d) Plane wave at 45°. In each case [0, 0] was taken as the middle of the room

The above plots demonstrate that the system is capable of accurately rendering plane wave, point sources and even focused sources in front of the loudspeaker array with a high degree of accuracy. However after the main wavefront passes reflections from the walls of the room can clearly be seen as ripples in the impulse responses. Quantifying the significance of these reflections and testing mitigation strategies are two of the key applications of this installation, as it will allow assessment of how successfully WFS may be used in real room environments and how subjective perception is affected.

5 APPLICATIONS

The WFS system presented here has been deliberately installed in a 'real room' environment specifically to test and improve upon how WFS performs under typical listening conditions. One of the major areas of research in WFS is how to overcome the problem of reflections in the reproduction room and having a permanent installation in a tightly controlled acoustic environment such as is presented here allows the quantification of how this affects both objective and subjective responses and consequently what the ramifications are for real world installations. It also allows the development, implementation and testing of algorithms used to cancel out the unwanted reflected energy from the loudspeaker signals which is a key area of WFS research. One possible approach might be to introduce cancelling image sources, akin to the low frequency method discussed in part two of this paper[8].

The system also provides an ideal test bed for auralising soundscapes for research into how sound affects our sense of place and emotions. Rather than recording a soundscape which has fixed and immovable sources, WFS allows a fully customisable and dynamic soundscape to be synthesised that can even be interacted with in real-time to assess people's preferences.

Once the physical accuracy of the system has been verified it is also proposed to use it to simulate other lower resolution loudspeaker systems. For example the system could render sources decoded according to ambisonics, vector-based amplitude panning (VBAP) and 5.1, allowing immediate subjective comparison between different rendering methodologies with no physical change of hardware and even combine aspects of different systems to result in a perceptually optimal hybridised system.

6 CONCLUSIONS

A Wave Field Synthesis (WFS) system has been implemented in an ITU spec listening room. Innovative features include the ability to adjust the height of the system on a winch mechanism so that it may be adjusted to the ideal listening height or lifted in amongst the ceiling treatment when not in use. The system will initially be used as a test-bed for evaluating and improving the performance of WFS in real rooms and once fully commissioned be available for subjective listening tests.

7 REFERENCES

1. ITU-R BS.1116-1, "Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems", International Telecommunications Union, Geneva, Switzerland (1997)

2. IEC.60268-13, BS.6840-13, "Sound System Equipment – Part 13: Listening Tests on Loudspeakers", International Electrotechnical Commission, Geneva, Switzerland (1998)

3. A. J.Berkhout "A Holographic Approach to Acoustic Control", J. Audio Eng. Soc., 36(12), December 1988, pp977-995.

4. A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic Control by Wave Field Synthesis", *J. Acoust Soc. Am, 93(5)*, May 1993.

5. E. N. G Verheijen. *Sound Reproduction by Wave Field Synthesis*. PhD thesis, Technical University of Delft, (1997).

6. S. Spors, R. Rabenstein, and J. Ahrens, "The Theory of Wave Field Synthesis Revisited," *Proc. 124th Conv. Audio Eng. Soc.,* Amsterdam, The Netherlands (2008)

7. S. Spors and R. Rabenstein. "Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis". *Conv. Audio Eng. Soc.*, Paris, France, (2006).

8. J. Hargreaves and M. Wankling, "Implementing Wave Field Synthesis in an ITU Spec Listening Room Part 2: Bass without modes," *Proc. IoA Reproduced Sound*, Brighton, UK, 2011