Implementing spherical microphone array to determine 3D sound propagation in the "Teatro 1763" in Bologna, Italy

Lamberto Tronchin (lamberto.tronchin@unibo.it)
Marco Tubertini (marco.tubertini2@studio.unibo.it)
Department of Industrial Engineering DIN
Inter-University Centre for Acoustics and Music Research CIARM
University of Bologna
Bologna, Italy, I-40136

Andrea Venturi (andrea.venturi78@gmail.com)
Angelo Farina (farina@unipr.it)
Department of Industrial Engineering
University of Parma
Parma, Italy, I-43100

ABSTRACT

A new method for recording the spatial properties of a soundfield, or for generating a synthetic three-dimensional soundfield, is described. The spatial distribution of sound waves passing at a point in space is sampled by means of a number of virtual directive microphones, covering the surface of a sphere. This corresponds to a discretization of the spatial information, which is exactly the spatial equivalent of the PCM sampling of a waveform.

The paper shows how it is possible to obtain information about the direction of early reflections in a theatre, not only by means of a 32-channels microphone, but also by means of traditional B-format microphone (Soundfield). The method is applied to a historical Italian theatre, located in Bologna, i.e. the 1763 theatre in Villa Mazzacorati.

The post processing includes the visualisation of the early reflections on a panoramic picture of the theatre, in which early reflections are reported as different circles on the figure, where the dimension and the intensity are related with the characteristics of the reflections.

Moreover, the influence of the height of the microphone in the calculation of the acoustic parameters was analysed. The measurements were repeated at different height and different position on a transversal line in the theatre, and statistically analysed.

1   INTRODUCTION

The importance of capturing the acoustic characteristics and the sonic behaviour of significant ancient theatres and auditoria was initially proposed by Gerzon [1], and became even more important after the destruction by fire of two theatres in Italy: Teatro Petruzzelli in Bari and

Teatro la Fenice in Venice [2]. Following these catastrophes, many attempts were made to standardise the acoustic measurements in theatre by taking into account several sound source positions, microphone positions and room conditions.

On the other hand, only a few attempts were made to analyse and standardise the effect of test signals employed during the measurements and the types of sound source and microphones [3]. These details become crucial when measured Impulse Responses (IRs) are employed for performing 3D auralisation of the room, rather than for simply obtaining the numerical values of ISO 3382 parameters. Moreover, the full spatial sonic behaviour of a theatre, which includes information about energy, intensity and location of early reflections in the room, is required to determine and solve some acoustic problems that could not be resolved by only considering mono or 2-channel IRs.

2   3D IMPULSE RESPONSE MEASUREMENTS

In 2003 and 2005 a measurement method able to obtain a complete description of spatial sound propagation in the auditorium was described [3, 4]. It incorporates all the previously known measurement techniques in a single, coherent approach. Three different microphone systems were mounted on a rotating beam (a binaural dummy head, a pair of cardioids in ORTF configuration, and a Soundfield microphone) and a set of Impulse Responses were measured at each angular position. The ORTF configuration represents a standard method (adopted by French Radio/Television) for recording dual-channel signals, employing two cardioid microphones spaced 170mm and divergent from each other by 110 degrees. The Soundfield microphone, introduced by Gerzon, enabled the measurement of 4-channel Impulse Responses and, therefore, spatial properties of the sound field. A Soundfield microphone captures a set of four signals known as "B-format" signals: one omnidirectional (sound pressure) and three with a polar pattern called "figure-of-eight", oriented along the three Cartesian axes X, Y, Z (these three channels capture a signal proportional to the Cartesian components of the particle velocity vector). The combination of the three aforementioned different measurement methods provided a general method from which all standard multi-channel playback formats (i.e., 2.0, 5.1, 7.1, 10.2, etc.) could be derived. A further enhancement of this method was recently presented in 2013 [5]. Following this new procedure, the description of spatial sound distribution in a theatre could be retrieved by means of a B-format microphone, without using expensive microphone probes, and also eventually reprocessing previous acoustic measurements in rooms.

3   TEST SIGNAL AND SOUND SOURCE

The techniques employed for the measurement of the Impulse Response have been significantly improved in recent years. The ESS (Exponential Sine Sweep) method allows for the calculation of the Impulse Response of an electro-acoustical system by avoiding contamination due to distortions, which usually occur in the loudspeaker and provides a large value of the S/N ratio [7], and that could be eventually emulated [8]. The choice of the sound source represents an important issue during the measurements and especially during the subsequent 3D auralisation. The ISO 3382/2009 standard [9] requires employing an omnidirectional sound source for measurements of room Impulse Responses. The main advantage of the ESS signal method is that it provides a huge dynamic range (often in excess of 100dB) and is completely immune from nonlinear effects. The main problem, however, is that the signal must be radiated by a loudspeaker system. Even by employing a state-of-the-art, measurement grade dodecahedron loudspeaker, the source is never really an "omnidirectional point source". The

polar pattern at medium-high frequencies deviates significantly from omnidirectionality, albeit the directivity limits of ISO 3382 are met. Furthermore, the radiated spectrum significantly changes with the direction and even with the distance from the sound source. Thus, for the analysis of spatial properties of rooms, it is better to employ sound sources which provide better omnidirectionality and a spatially-invariant spectrum such as specially-modified starter pistols equipped with an omnidirectional diffuser, or industrial-grade firecrackers, even if these sound sources provide smaller dynamic ranges (although usually better than 80 dB) and a not-perfectly-flat spectrum [10].

Even though an omnidirectional source does not correspond to the effective directivity pattern of real-world sound sources, it is preferable when the purpose of the measurements is to precisely determine the spatial sound distribution in a room. It avoids exploiting room effects (abnormal concentration of energy and focusing for selected orientations of a directive source), as can happen when employing highly directive loudspeakers. However, when the purpose of the measurements is to determine the acoustic response of a room when a particular kind of source is used (a particular musical instrument or the human voice), a directive sound source could be added to the omnidirectional during the measurements. When employing an electrical test signal, the source is usually a dodecahedron or another type of spherical source.

For wide-band Impulse Response measurements to be used for auralisation it is necessary to use a specially-built dodecahedron with wide frequency response (30Hz to 16kHz, minimum) and incorporating a perfectly-flat digital equalisation system. In many cases the need to employ a portable and cheap system makes it preferable to employ "truly impulsive" sound sources such as starter pistols, balloons or firecrackers. Recent studies [11, 12] analysed the properties of balloons and revealed that they are, in general, unsatisfactory sound sources.

4   MICROPHONES

So far, the ISO 3382/2009 standard [9] requires omnidirectional, monoaural microphones to be utilised in the measurement of monophonic acoustic parameters and only specifies the dimension of the microphones (preferably less than 13mm). Moreover, the ISO 3382/2009 standard describes the characteristics of binaural microphones (real heads or dummy heads), which could be used to measure binaural Impulse Responses and IACC. The standard also considers using figure-of-eight microphones to measure some lateral-energy parameters, such as LF and LFC, but does not provide technical specifications for these types of directive microphones. However, it is evident that monoaural or even figure-of-eight microphones cannot provide complete information about spatial sound distribution in the theatre. For this purpose, a multi-microphone system is necessary to capture the complete spatial information.

4.1   B-Format Microphone

Leaving aside binaural measurements, required only for binaural parameters, a B-format microphone (such as the Soundfield[TM]) has been considered for many years to be the optimal transducer for performing 3D Impulse Response measurements in theatres and auditoriums. The W channel is good for the monoaural parameters (omnidirectional), the Y channel provides the figure-of-eight signal required for computing LF, and the other two directive channels (X and Z) can be used for recreating the entire 3D soundscape inside a playback environment using the well-known 1st-order Ambisonics technology.

However, a 1st order Ambisonics playback system is currently considered incapable of providing accurate spatial cues to the listeners, as this technology does not synthesise sound fields exhibiting significant polarisation and consequently the sound is perceived to be coming from almost anywhere. A possible solution is to employ high-order Ambisonics (HOA) systems, which require capturing a multichannel signal corresponding to the spherical harmonics expansion up to 3rd or 4th order. While it was found that HOA works very well with synthetic signals (where the high-order spherical harmonic signals are computer-generated), the recording of HOA signals is problematic, even when employing microphone arrays composed of dozens of elements: when the directivity of the harmonic patterns is high, the S/N ratio is poor at low frequencies and the spatial accuracy of the pattern is disrupted at high frequencies, resulting in the reduction of the useful bandwidth to less than one octave band.

Another viable approach is to employ "advanced" decoding methods applied to the 1st order B-format signal, which perform a "spatial analysis" of the signal, and therefore "steer" the sound just in the very precise directions from where it arrives from, at every instant and at every frequency. Two of these methods are currently being employed, i.e. Sirr/Dirac and Harpex [12, 13]. The former is based on the sound intensity theory while the latter is based on plane wave decomposition. Higher order B-Format microphone can be achieved employing Eingemike [5].

5   SPATIAL ANALYSIS OF 3D IMPULSE RESPONSES WITH B-FORMAT IRS

The techniques employed to analyse 3D impulse response is based on the work of Farina and Tronchin [5]. It exploits the capabilities of the B-Format signal of detecting the direction-of-arrival of each impinging wavefront by computing the "instantaneous" sound intensity vector I and the instantaneous value of the energy ratio r_E and is based on the same vector decomposition scheme initially proposed in [16], also related to the later SIRR method [12]. The three components of the sound intensity vector can be simply obtained from the B-Format Impulse Response by means of the following equations:

$$I_x = w \cdot x; \quad I_y = w \cdot y; \quad I_z = w \cdot z; \tag{1}$$

Where w, x, y and z represent the four signals of the B-Format IR. Keeping in mind these signals are proportional to sound pressure and particle velocity, the total energy density can be computed by means of the following equation:

$$E_D = \frac{w^2 + x^2 + y^2 + z^2}{c} \tag{2}$$

The modulus of the sound intensity vector is:

$$|I| = \sqrt{I_x^2 + I_y^2 + I_z^2} \tag{3}$$

The ratio between the active intensity and energy density is computed as:

$$r_E = \frac{|I|}{E_D \cdot c} = \frac{\sqrt{(w \cdot x)^2 + (w \cdot y)^2 + (w \cdot z)^2}}{w^2 + x^2 + y^2 + z^2} \tag{4}$$

Finally, the *azimuth* (horizontal) *"a"* and *elevation* (vertical) *"e"* angles are simply obtained from trigonometric equations, i.e.:

$$a = \arctan\frac{I_y}{I_x}; \qquad e = \arcsin\frac{I_z}{|I|} \tag{5}$$

All these quantities are averaged over 1ms time slices, creating a "time history" of the above-defined descriptors along the whole length of the measured Impulse Response. The results can be visualised dynamically by means of a properly developed post-processing tool, plotting at every "frame" a circle, located at position *(a,e)*, having a diameter proportional to the sound intensity modulus *|I|* and opacity proportional to r_E. The moving circle is plotted over a panoramic 360°x180° photographic image taken from the microphone position, while a synchronised marker moves over the Impulse Response graph so that it is easy to see the arrival direction of each reflection and how much the corresponding wavefront is "polarised".

The meaning of r_E is related to the fact that the sound energy is significantly oriented along one direction (r_E approaching 1, travelling wave), or instead is diffuse (r_E approaching 0, standing wave).The chart does not display the "Impulse Response of r_E", but rather the superposition of the "energetic" Impulse Response (that is E_D in dB) and the "intensimetric" Impulse Response (that is, *|I|* in dB), having aligned both dB scales so that, for a perfectly plane, progressive wave (when r_E=1) the two values in dB are the same both for E_D and *|I|*.

Hence, the dynamic display of the spatial-temporal distribution of sound along the duration of the Impulse Response does not only carry the information of the trajectory of the reflected sound, but also about the degree of diffusion. The sound is fully diffused when the level of E_D is much larger than the level of *|I|* (and hence r_E approaches 0). When, instead, the two levels are almost equal (meaning that r_E=1), the sound is strongly "polarised", a propagative wave traveling in a precise direction. This method allows for processing of a large amount of B-format IRs previously measured in the scientific community by means of Soundfield microphones or other tetrahedral probes, obtaining much more information than traditional acoustic parameters.

6   POST PROCESSING

After measuring in each position a 3D impulse response (either B-format or 32-channels), it is possible to post-process the results in two ways:
- A graphical analysis can be performed, showing the spatial distribution of the incoming energy along the running time – this allows to "see" from where the room's reflections are coming;
- An audible rendering can be presented to a group of listeners, inside a special room equipped with a suitable array of loudspeaker, surrounding completely the listening area around a sphere
The graphical analysis is performed thanks to two post-processing software tools:
- the first shows the "moving circle", which corresponds to the instantaneous direction of arrival of the sound intensity, based on the analysis of a B-format impulse response performed according with the algorithm described in par. 5.
-     the second creates an animated colour video rendering of the sound map, over-plotted over the panoramic image. In this case no graphical algorithm is required, as a standard graphic library is employed for obtaining the colour map, based on the 32 "instantaneous" values of the sound pressure level captured by the 32 virtual microphones. The sound map technique revealed useful to discover the main sources of noise inside car cockpit too [14]. These tools create an animated video rendering of the instantaneous sound intensity vector, in one case,

and of a colour map of sound distribution, in the second case, plotted over the panoramic image. A frame of such video renderings is shown in Figure 2. The audible rendering is obtained by reprocessing the original impulse response recording: a new set of virtual microphones is extracted, one feeding each loudspeaker of the playback array. Again, the processing is slightly different for B-format impulse response, and for 32-channels impulse responses, although a same methodological approach is employed.

The set of filters employed for deriving the "playback" virtual microphones is obtained by solving a linear equation system, imposing that the signals re-recorded placing the probe (either the B-format microphone or the Eigenmike™) at the centre of the playback system are maximally similar to the original signals recorded in the theatre. This approach, which is not Ambisonics-based, also corrects inherently for deviations from ideality of the loudspeakers employed, both in terms of magnitude/phase response, and in terms of placement/aiming/shielding.

7 THE CASE STUDY – VILLA MAZZACORATI

The case study here reported was conducted inside the "Teatro 1763" in Villa Mazzacorati, Bologna, Italy. In this experiment the Impulse Responses have been measured to analyze also the variability of sound parameters (particularly clarity, definition and reverberation time) on lateral and vertical moving inside the theatre. The Impulse Responses have been measured with a digitally equalised dodecahedron sound source fed by a SineSweep signal and three different microphones for the measurements.

7.1 Results on Villa Mazzacorati

7.1.1 Spherical array measurements

The analysis of the acoustic measurements in the Teatro 1763 was initially focused on the directions of early reflection.
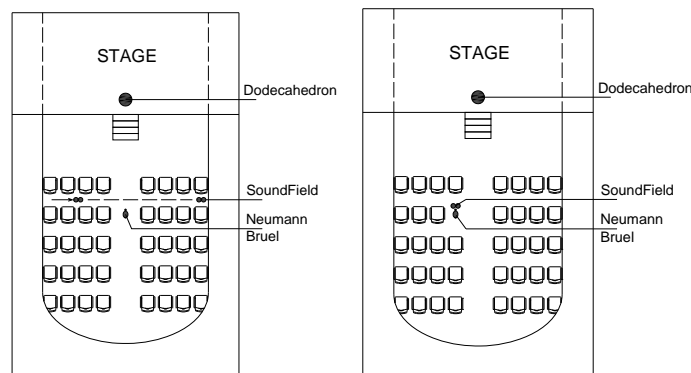

Figure 1: plans of Teatro 1763 Villa Mazzacorati, with positions of microphones for two type of measurements

Since "Teatro 1763" is a little theatre, as reported in figure 1, the reverberant sound field is very small, and the energy analysis of the impulse responses could give proper information also in the central part of the tail of the impulse response. Figure 2 shows a frame from the movie

obtained following the theory described in par. 5. From the picture, it appears as a fairly strong reflection was caused by the lateral wall on the left side, due to the characteristics of the wall (a mixed wood and plastered wall).
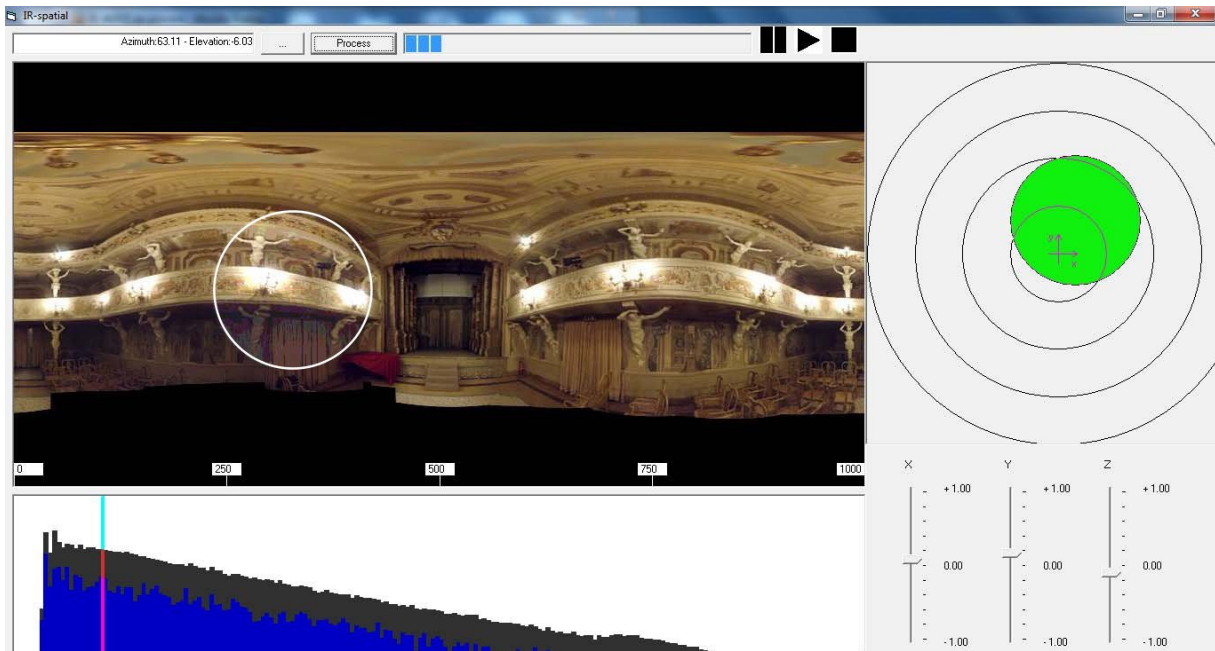

Figure 2: image of IR-spatial, software to analyze indoor reflections

7.1.2 Lateral and vertical movements

Two different movements have been done. In the first example, the SoundField (B-Format), the BK (omnidirectional) and the Neumann (dummy head) microphones were fixed in only one position of the theatre.
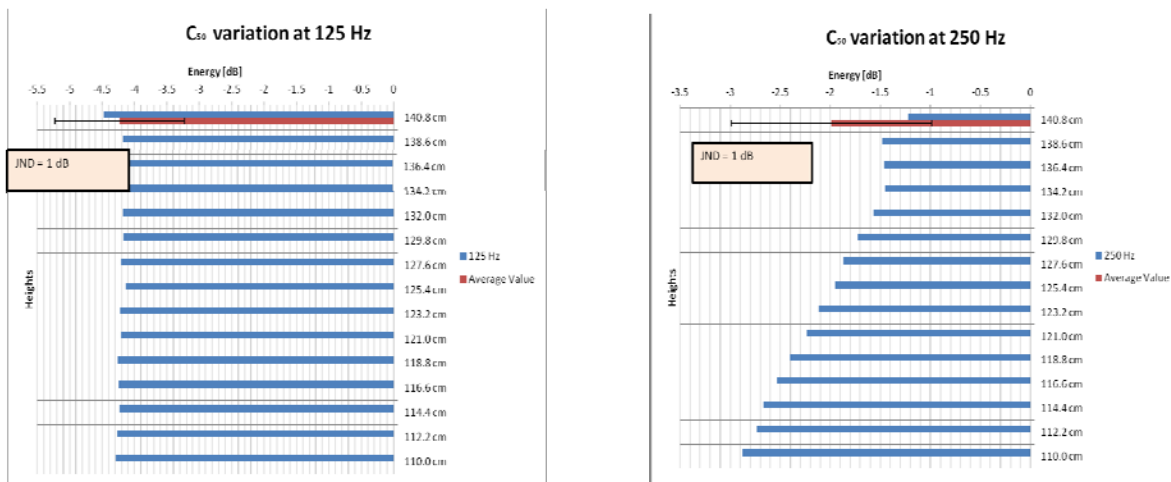

Figure 3: variability of parameter C50 in vertical moving at frequencies 125Hz and 250 Hz for SoundField microphone

Only the SoundField microphone have been lifted for 15 different heights from 1.10m to 1.40m, while Neumann and BK microphones were used like a reference measurement at the same height. In the second example, the SoundField microphone was moved from right to left side of the theatre in 10 different positions in front of the stage, whereas the Neumann and the BK microphones were fixed in the middle of the room.

The measurements have been processed with Aurora to extract from every sound signal its Impulse Response and calculate its acoustic parameters. For each measurement, the acoustic parameters have been calculated for each octave-band frequency (from 125 Hz to 8 kHz) and they have been compared with the corresponding JND to analyze the variability.
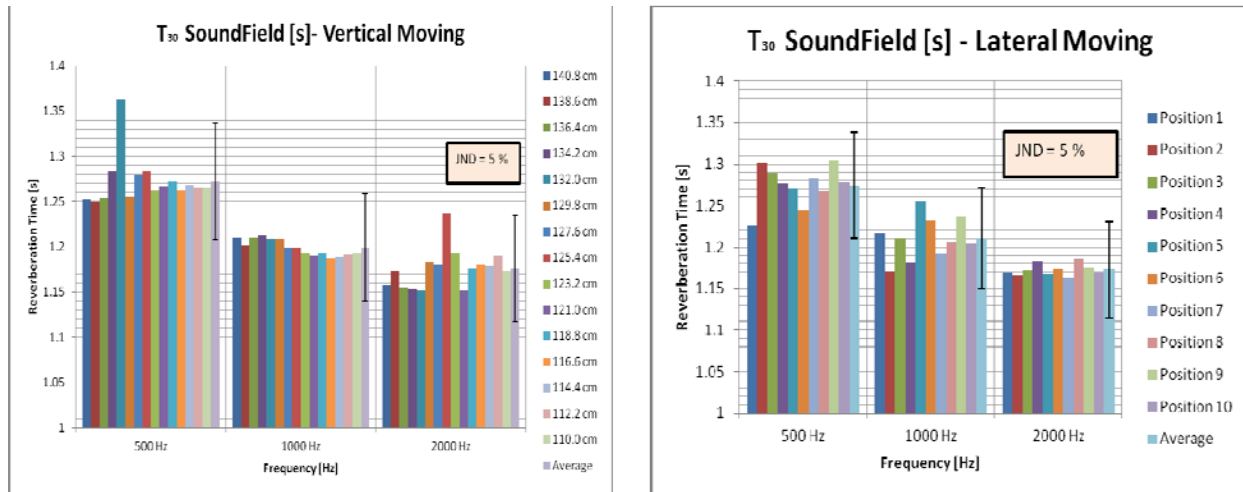

Figure 4: graphics about variability of parameter T30 for lateral and vertical moving at frequencies 500 Hz, 1000 Hz and 2000 Hz.

The statistical analysis performed on these measurements, showed some remarkable results, especially with relation to Clarity. Whilst at 125 Hz Clarity was substantially constant with the movement of the Soundfield Mic, at 250 Hz, the variation of Clarity C50 in the vertical plane is systematic, even though within the JND value. All the reverberation times (i.e. EDT, T20 and T30) slightly vary both during lateral and vertical movement of the microphone. However, their variation at mid frequencies (500 Hz, 1kHz and 2kHz) is always within the JND range (figure 3). These results show that the intrinsic characteristics of theatre are respected in every position and at all heights.

8   CONCLUSIONS

The new method based on spherical array microphone has been applied to the Teatro 1763 in Bologna; moreover, in this theatre the variation of some acoustic parameters (as clarity and RT) has been considered. It was shown that, despite the spherical array could give very précis information about the directions of early reflections, some parameters (as Clarity) vary considerably moving the microphone vertically or horizontally. Nevertheless, these changes are within the JND of these parameters, but further experiments are necessary to determine the variation of these acoustic parameters (and others) in different (more reverberant) rooms.

REFERENCES

1 M. Gerzon - "Recording Concert Hall Acoustics for Posterity", JAES Vol. 23, Number 7 pages 569-571 (1975).

2 L. Tronchin, A. Farina - "The acoustics of the former Teatro "La Fenice", Venice", JAES Vol. 45, Number 12 p. 1051-1062 (1997).

3 A. Farina, R. Ayalon – "Recording concert hall acoustics for posterity" - 24th AES Conference on Multichannel Audio, Banff, Canada, 26-28 June 2003.

4 A. Farina, L. Tronchin – "Measurement and reproduction of spatial sound characteristics of auditoria", Acoustical Science and Technology, 26(2), pag 193-199, 2005.

5 A. Farina, L. Tronchin – "3D sound characterisation in theatres employing microphone arrays", Acta Acustica United with Acustica, 99(1), Pages 118-125, 2013.

6 A. Farina, A. Amendola, A. Capra, C. Varani – "Spatial analysis of room impulse responses captured with a 32-capsules microphone array", 130th AES Conference, London, 13-16 May 2011.

7 A. Farina – "Simultaneous measurement of impulse response and distortion with a swept-sine technique", 110th AES Convention, Paris 18-22 February 2000.

8 L. Tronchin – "The emulation of nonlinear time-invariant audio systems with memory by means of Volterra series", Journal of the Audio Engineering Society, 60(12), Pages 984-996, 2013.

9 ISO 3382-1 – "Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces", 2009.

10 A. Farina - "Impulse Response Measurements" - 23rd Nordic Sound Symposium, Bolkesjø (Norway) 27 – 30 September 2007.

11 J Pätynen, BF Katz, T Lokki – "Investigations on the balloon as an impulse source" – J. Acoust Soc Am. 2011 Jan, 129(1) pp. 27-33.

12 V. Pulkki - "Spatial Sound Reproduction with Directional Audio Coding.", J. Audio Eng. Soc., 55(6):503–516, 2007.

13 S. Berge, N. Barrett, - "High Angular Resolution Planewave Expansion", Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics May 6-7, 2010, Paris, France.

14 M. Binelli, A. Venturi, A. Amendola, A. Farina – "Experimental analysis of spatial properties of the sound field inside a car employing a spherical microphone array", 130th AES Conference, London, 13-16 May 2011