Audio Engineering Society
Convention Paper 7555

Presented at the 125th Convention
2008 October 2–5 San Francisco, CA, USA

Surround system based on three dimensional sound field reconstruction

Filippo M. Fazi[1], Philip A. Nelson[1], Jens E. N. Christensen[1] and Jeongil Seo[2]

[1]Institute of Sound and Vibration Research, University of Southampton
Highfield - S0171BJ, Southampton, U.K.

[2] Broadcasting Media Research Group, Electronics and Telecommunications Research Institute
161 Gajeong-dong, Yuseong-gu, Daejeon, 305-600, KOREA

Correspondence should be addressed to Filippo M. Fazi (ff1@isvr.ston.ac.uk)

ABSTRACT

The theoretical fundamentals and the simulated and experimental performance of an innovative surround sound system are presented. The proposed technology is based on the physical reconstruction of a three dimensional target sound field over a region of the space using an array of loudspeakers surrounding the listening area. The computation of the loudspeaker gains includes the numerical or analytical solution of an integral equation of the first kind. The experimental setup and the measured reconstruction performance of a system prototype constituted by a three dimensional array of 40 loudspeakers are described and discussed.

1. INTRODUCTION

During the last 30 years many technologies have been proposed for the reproduction of a desired sound field using an array of loudspeakers. Wave Field Synthesis [1], [2], [3] is a well known approach, based on the Kirchhoff-Helmoholtz integral (also known as Green formula [4]). Ambisonics is another known method that is widely used in multi-channel audio. Its first formulation was proposed by Gerzon [5] and the further extension of Ambisonics to High Order Ambisonics [6], [7] proved the latter to be an efficient method for the reproduction of a sound field. Other theories have been recently proposed, which are based on the decomposition of the desired sound field in terms of spherical harmonics [8], [9], [10]. Other original theories such as, among others, [11], [12], [13] [14], [15], [16], [17] have been developed and are based on the solution of an acoustic inverse problem or on other principles.

In this paper an alternative sound field reconstruction method is proposed. This approach starts from the ideal assumption that the surround system is constituted of a continuous distribution of secondary sources,

continuously arranged on the boundary of a three dimensional volume. The strength of the secondary sources is calculated by solving an integral equation of the first kind (different from the well known Kirchhoff-Helmholtz integral). Part of the underlying theory was first proposed by the authors in [18] and is here revisited, with the attempt of providing a physical interpretation of some of the mathematical entities introduced.

In the second section of this paper the theoretical fundamentals of the proposed method are presented. The problem of sound field reconstruction is addressed and some relevant mathematical results are reviewed. The reconstructed sound field is represented as a single layer potential, and its unknown density is shown to be the solution of an integral equation of the first kind. The concepts of adjoint operator and singular value decomposition are introduced, and their physical interpretation is proposed for the case under consideration. In Section 3 an explicit solution of the integral equation is illustrated and discussed, for the special case when the boundary of the reconstruction area and of the surface on which the secondary sources are arranged are two concentric spheres. In Section 4 the mathematical analogy of the problem of sound field reconstruction to the problem of acoustic scattering is discussed. It is explained that the strength of the secondary sources that allows the reproduction of a given target sound field is related to the gradient of the same field scattered by a sound-soft object. Simple physical considerations related to the scattering theory are used to derive a simplified formula for the strength of the secondary sources. Section 5 is dedicated to the discussion of the problems related to the practical realization of a surround system based on the method presented. The attention is focused on the effect of the discretisation of the continuous layer of secondary sources. The experiments described in Section 6 show the real performance of the method proposed. A multi-channel system constituted by an array of 40 loudspeakers arranged on a $4m$ diameter sphere was realized. A set of digital filters was designed applying the proposed method and implemented on a PC based system, in order to generate the driving signals for the loudspeakers. The target sound field and the field reproduced by the system were measured with a microphone array and are compared.

Notation

Vectors are represented by lower case bold letters. The convention for the spherical co-ordinates r_x, θ_x, ϕ_x of a given vector $\mathbf{x}$ is illustrated in Figure 1. Matrices are represented by capital bold letters. j is the imaginary unit, that is $j=\sqrt{-1}$. The symbol $[\cdot]^*$ represents the complex conjugate. δ_{nm} is the Kronecker delta ($\delta_{nm}=1$ if $n=m$, $\delta_{nm}=0$ if $n\neq m$).

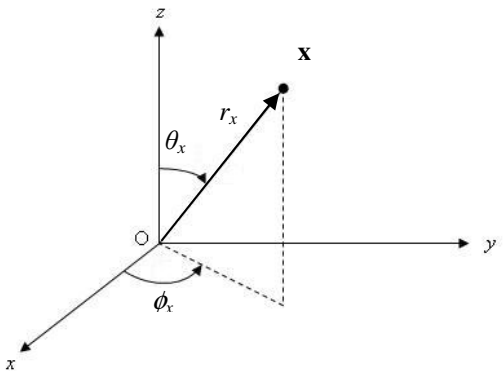

Figure 1: Reference geometry.

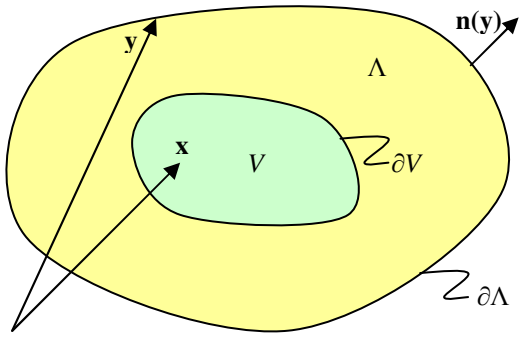

Figure 2: Cross-section of the three dimensional regions V and Λ and other reference geometrical entities.

2. GENERAL THEORY

Let $V \subseteq R^3$ be a bounded and simply connected region of the space, as shown in Figure 2, with boundary ∂V (of class C^2). This region is referred to as the reconstruction volume. The assumption is made that the target sound field $p(\mathbf{x})$ is due to sources of sound that are not contained in V, and that no scattering object is contained within this region. $p(\mathbf{x})$ satisfies the homogeneous wave equation

$$\nabla^2 p(\mathbf{x},t) - \frac{1}{c^2}\frac{\partial^2 p(\mathbf{x},t)}{\partial t^2} = 0 \qquad (1)$$

in V, where c is the speed of sound, considered to be uniform in V. When a single frequency ω is considered and $p(\mathbf{x},t) = \mathrm{Re}\{p(\mathbf{x})e^{-j\omega t}\}$, equation (1) can be reformulated as the Helmholtz equation

$$\nabla^2 p(\mathbf{x}) + k^2 p(\mathbf{x}) = 0 \quad \mathbf{x} \in V \qquad (2)$$

where the $k = \omega/c$ is the wave number and the time dependence $e^{-j\omega t}$ has been omitted. Let now $\Lambda \subseteq R^3$ be a bounded and simply connected region of the space, with boundary $\partial\Lambda$ of class C^2, that fully encloses V. Assume now that a continuous distribution of an infinite number of secondary, monopole-like sources is arranged on the boundary $\partial\Lambda$. The assumption is made that the sound field $p_y(\mathbf{x})$ generated by the secondary source located at $\mathbf{y} \in \partial\Lambda$ can be represented by a Green function $G(\mathbf{x} \mid \mathrm{y})$, thus satisfying the inhomogeneous Helmholtz equation

$$\nabla^2 p_y(\mathbf{x}) + k^2 p_y(\mathbf{x}) = -a(\mathbf{y})\delta(\mathbf{x}-\mathbf{y}) \quad \mathbf{x} \in \overline{V},\ \ y \in \partial\Lambda \qquad (3)$$

where the function $a(\mathbf{y})$ represents the complex strength of the secondary sources. In free field, $p_y(\mathbf{x})$ can be represented by the free field Green function

$$p_y(\mathbf{x}) = a(\mathbf{y})g(\mathbf{x} \mid \mathrm{y}) = a(\mathbf{y})\frac{e^{jk|\mathbf{y}-\mathbf{x}|}}{4\pi|\mathbf{y}-\mathbf{x}|} \quad \mathbf{x} \in R^3,\ \ \mathbf{y} \in \partial\Lambda \qquad (4)$$

In a reverberant environment, the Green function has a more complex expression, which strongly depends on the shape of the reverberant enclosure and on its impedance boundary conditions. The reader can refer to [19] for a detailed discussion on Green functions.

The sound field $\hat{p}(\mathbf{x})$ generated by the infinite number of secondary sources uniformly arranged on $\partial\Lambda$ can be expressed as an integral, representing the linear superposition of the sound fields generated by the single sources:

$$\hat{p}(\mathbf{x}) = (Sa)(\mathbf{x}) = \int_{\partial\Lambda} G(\mathbf{x} \mid \mathbf{y})a(\mathbf{y})dS(\mathbf{y}) \quad \mathbf{x} \in V \qquad (5)$$

In what follows, $\hat{p}(\mathbf{x})$ is also referred to as the reconstructed or reproduced sound field, while the integral introduced is often referred to as a *single layer potential* and $a(\mathbf{y})$ is called *density* of the potential [4].

Two problems are now addressed:

- If the desired sound field $p(\mathbf{x})$ is defined on the boundary ∂V, is it also uniquely defined in the interior volume V ?
- If the desired sound field $p(\mathbf{x})$ is known on the boundary ∂V, is it possible to define a density $a(\mathbf{y})$ such that the reconstructed sound field $\hat{p}(\mathbf{x})$ equals the desired sound field in the reconstruction area V ?

2.1. The interior Dirichlet problem

The rigorous answer to the first question is related to the solvability of the interior Dirichlet problem [4]. The problem can be readdressed in terms of the existence and the uniqueness of the solution of the differential equation

$$\begin{aligned} &\nabla^2 p(\mathbf{x}) + k^2 p(\mathbf{x}) = 0 && \mathbf{x} \in V \\ &p(\mathbf{x}) = f(\mathbf{x}) && \mathbf{x} \in \partial V \end{aligned} \qquad (6)$$

where the function $f(\mathbf{x})$ describes the value field $p(\mathbf{x})$ on the boundary ∂V .This differential equation is known as the Dirichlet problem (and the related boundary condition is called after the same name). As the Kichhoff-Helmholtz integral equation suggests, the

knowledge of both the sound field and its normal derivative on the boundary (the Cauchy boundary condition) uniquely defines a sound field in the interior region V. However, as discussed in [20] and [19], this condition can be relaxed. In fact, with the appropriate regularity assumption on the function $f(\mathbf{x})$, the Dirichlet problem (5) has a unique solution, as proved in [4]. This holds as long as the wave number k is not one of the Dirichlet eigenvalues k_n. The latter are defined as the infinite and countable wave numbers k_n such that the differential equation (5) with homogeneous boundary conditions $f(\mathbf{x}) = 0$ is satisfied.

The meaning of this theorem is that, in general, the Dirichlet boundary condition implicitly contains the Neumann boundary condition, that is the normal derivative of the sound field on ∂V. In other words, the knowledge of the acoustic pressure on ∂V uniquely defines the normal particle velocity on ∂V, and therefore the sound field in the interior region V. But this does not hold in the case when the wave number is one of the Dirichlet eigenvalues k_n. It is important to emphasize that these values identify the internal resonance frequencies $\omega_n = k_n c$ of an ideal object with the shape of V and pressure release boundaries. The Dirichlet problem (5) then does not have a unique solution, and this can be intuitively explained by the fact that if a field $p_n(\mathbf{x})$, with operating frequency ω_n, is a solution of (5), then also any other sound field $ap_n(\mathbf{x})$, $a \in R$ is a solution of (5). It can actually be argued that, under these circumstances, it is not possible to obtain the normal derivative of the field on ∂V from the boundary value of the field $f(\mathbf{x}) = 0$.

2.2. The integral equation

The second problem that has been addressed is related to the feasibility of the reconstruction of the desired sound field using a continuous layer of monopole-like secondary sources, when the target field is known on ∂V. The target is to find a density function $a(\mathbf{y})$, if the latter exists, such that the reconstructed sound field $\hat{p}(\mathbf{x})$ expressed by (5) equals the target sound field $p(\mathbf{x})$. Following what has been said about the solvability of the Dirichlet problem, if the operating frequency is not one of those identified by the Dirichlet eigenvalues, when the sound field is accurately reconstructed on ∂V then it is also reconstructed in V. The reconstruction problem can be formulated as

$$p(\mathbf{x}) = (Sa)(\mathbf{x}) = \int_{\partial\Lambda} G(\mathbf{x} \mid \mathbf{y}) a(\mathbf{y}) dS(\mathbf{y}) \qquad \mathbf{x} \in \partial V \tag{7}$$

where $p(\mathbf{x})$ is given and the density $a(\mathbf{y})$ is the unknown of the problem. Note that the integral operator $(Sa)(\mathbf{x})$ has been defined here as the restriction of the single layer potential (5) to the boundary ∂V. Equation (7) is an integral equation of the first kind and the determination of $a(\mathbf{y})$ from the knowledge of $p(\mathbf{x})$ on the boundary ∂V represents an inverse problem. A thorough insight into the solution of integral equations is beyond the scope of this paper, and the interested reader is referred to [21] or any other text book on functional analysis. It is important to highlight that equation (7) represents a problem that is, in general, ill-posed [22]. This implies that a solution $a(\mathbf{y})$ might not exist, and even if it exists it might be non-unique or not continuously dependent on the data $p(\mathbf{x})$. The latter concept implies that small variations or errors on $p(\mathbf{x})$ can result in very large errors in the solution $a(\mathbf{y})$, which is therefore said to be unstable [23].

2.3. Adjoint operator and Singular value decomposition

Even if the exact solution to equation (7) does not always exist or is not stable, it is always possible to calculate an approximate and stable solution. In this section the singular value decomposition of the integral operator S is proposed as one of the possible methods for solving (7). Attention is focused in this paper on the physical interpretation of this technique, while the reader can refer to [18] and to [4] for the mathematical detail.

To begin with, the scalar product of two square integrable functions $f(\mathbf{y})$ and $g(\mathbf{y})$, with the same domain D, is here defined as

$$\langle f \mid g \rangle = \int_D f(\mathbf{x}) * g(\mathbf{x}) dS(\mathbf{x}) \tag{8}$$

The adjoint operator S^+ of S is defined to be such that

$$\langle g \mid S\ a \rangle = \langle S^+ g \mid a \rangle \tag{9}$$

It can be easily verified that, for the case under consideration, the adjoint operator S^+ is given by

$$(S^+ g)(\mathbf{y}) = \int_{\partial V} G(\mathbf{y} \mid \mathbf{x}) * g(\mathbf{x}) dS(\mathbf{x}) \qquad \mathbf{y} \in \partial\Lambda \tag{10}$$

The physical meaning of the single layer potential $(Sa)(\mathbf{x})$ is represented by the sound field generated by the continuous distribution of secondary sources on $\partial\Lambda$, evaluated at $\mathbf{x} \in \partial V$. Similarly, its adjoint operator $(S^+ g)(\mathbf{y})$ can be regarded as the time reversed version of the sound field generated by a continuous distribution of monopole-like secondary sources on ∂V, evaluated at $\mathbf{y} \in \partial\Lambda$. The time reversal is due to the fact that the kernel of the integral (10) is the complex conjugate of the Green function $G(\mathbf{y} \mid \mathbf{x})$. Considering, as an example, the case of the free field Green function $g(\mathbf{x} \mid \mathbf{y})$ introduced in equation (4), it can be easily verified that $g(\mathbf{y} \mid \mathbf{x})$ and $g(\mathbf{y} \mid \mathbf{x})^*$ differ because of the sign of the exponential. This implies that while $g(\cdot \mid \mathbf{x})$ can be considered as the representation of a sound field generated by a source of outgoing (diverging) spherical waves located at $\mathbf{x}$, $g(\cdot \mid \mathbf{x})^*$ could be understood as the case of a source of incoming (converging) spherical waves located at the same position $\mathbf{x}$. Alternatively, the sound field represented by $g(\cdot \mid \mathbf{x})^*$ could be regarded as the time reversed version of $g(\cdot \mid \mathbf{x})$, as

$$\mathrm{Re}\left\{e^{-j\omega t} g(\mathbf{x} \mid \mathbf{y})^*\right\} = \mathrm{Re}\left\{\frac{e^{-j(k|\mathbf{x}-\mathbf{y}|+\omega(-t))}}{4\pi|\mathbf{x}-\mathbf{y}|}\right\} = \mathrm{Re}\left\{e^{-j\omega(-t)} g(\mathbf{x} \mid \mathbf{y})\right\} \tag{11}$$

For both interpretations, the generated spherical wave fronts are converging towards $\mathbf{x}$. As a consequence of what has been said, the adjoint operator S^+ could be interpreted as a continuous distribution of "sources of incoming waves" on ∂V. As explained in [18], the operator S is compact and therefore its adjoint operator is compact too and the composite operator S^+S is compact and self-adjoint (the reader can refer to [21] or [24] for the definition and mathematical explanation of these concepts). It is therefore possible to apply the properties of compact self adjoint operators [21], [24] and to perform a spectral decomposition of S^+S. As a first step, the set of functions $a_n(\mathbf{y})$ is computed, these functions being solutions of the eigenvalue problem

$$(S^+S\ a_n)(\mathbf{y}) = \lambda_n a_n(\mathbf{y}) \qquad \mathbf{y} \in \partial\Lambda,\ \ n = 1,2,3...\infty \tag{12}$$

where the eigenvalues λ_n are real, positive numbers. Considering the interpretation of S and S^+ introduced above, the effect of the composite operator $(S^+Sa)(\mathbf{y})$ could be understood as follows: a sound field is generated by the continuous distribution of monopole-like sources on $\partial\Lambda$ with strength $a(\mathbf{y})$, and this field, on the boundary ∂V, is described by the function $\hat{p}(\mathbf{x})$. Mathematically this implies that $\hat{p}(\mathbf{x}) = (Sa)(\mathbf{x})$. A different sound field is then generated by a continuous distribution of monopole-like sources on ∂V (the operator S^+), and the strength of the sources is determined by the function $\hat{p}(\mathbf{x})$. The sound field is then time reversed (or alternatively the secondary sources on ∂V can be replaced by "sources of incoming waves"), and the function $\hat{a}(\mathbf{y})$ describes the value of this field on $\partial\Lambda$. In mathematical terms, this corresponds to $\hat{a}(\mathbf{y}) = (S^+\hat{p})(\mathbf{y})$. Summarizing, the effect of the operator S^+S can be understood as if the field generated by the secondary sources on $\partial\Lambda$ was propagated from $\partial\Lambda$ to ∂V and then propagated back to $\partial\Lambda$. Attention is now focused on the relation between $\hat{a}(\mathbf{y})$ and $a(\mathbf{y})$: if these two functions are such that $\hat{a}(\mathbf{y}) = \lambda_n a(\mathbf{y})$, $\lambda_n \in R^+$, then the function $a(\mathbf{y})$ is one of the eigenfunctions $a_n(\mathbf{y})$ of S^+S, and is a solution of (12). This means that the action of S^+S on $a_n(\mathbf{y})$, is simply an amplification or attenuation of the function, corresponding to positive real number λ_n. When, in practical cases, the layer of secondary sources is substituted by an array of loudspeakers, the functions $a_n(\mathbf{y})$ could be interpreted as *array modes* [23].

The set of functions $a_n(\mathbf{y})$ constitutes an orthogonal set of functions for $\partial\Lambda$ [21], [24], meaning that any square integrable function $a(\mathbf{y})$ defined on $\partial\Lambda$ can be expressed as

$$a(\mathbf{y}) = \sum_{n=1}^{N} \langle a_n \mid a \rangle a_n(\mathbf{y}) + (Qa)(\mathbf{y}) \tag{13}$$

The integer N depends on the dimension of the range of S and might be infinite. $(Qa)(\mathbf{y})$ is the orthogonal projection of $a(\mathbf{y})$ on the null-space of S [18], [24]. The latter is defined as the set of functions $\tilde{a}(\mathbf{y})$ such that

$$N(S) = \left\{\tilde{a}(\mathbf{y}) : (S\tilde{a})(\mathbf{x}) = 0\right\} \tag{14}$$

If this set is empty, then the set of functions $a_n(\mathbf{y})$ is complete and equation (13) can be regarded as a

generalized Fourier series. We can also generate a set of orthogonal functions $p_n(\mathbf{x})$ on ∂V by letting the operator S act on the functions $a_n(\mathbf{y})$ such that

$$\mu_n p_n(\mathbf{x}) = (Sa_n) \qquad n = 1,2,3,...\infty \tag{15}$$

where the positive real numbers $\sigma_n = \sqrt{\lambda_n}$ are the singular values of S. It can be also proved [24] that

$$\sigma_n a_n(\mathbf{x}) = (S^+ p_n) \qquad n = 1,2,3,...\infty \tag{16}$$

It is possible to express any square integrable function $p(\mathbf{x})$ on ∂V as the infinite series

$$p(\mathbf{x}) = \sum_{n=1}^{N} \langle p_n \,|\, p \rangle p_n(\mathbf{x}) + (Rp)(\mathbf{x}) \tag{17}$$

$(Rp)(\mathbf{x})$ being the orthogonal projection of $p(\mathbf{x})$ on the null-space of S^+. Combining equations (13) and (15) and keeping in mind that $(S(Qa))(\mathbf{x}) = 0$ because of the definition of the null-space (14), it is possible to express the action of S on $a(\mathbf{y})$ as [24]

$$(Sa)(\mathbf{x}) = \sum_{n=1}^{N} \sigma_n \langle a_n \,|\, a \rangle p_n(\mathbf{x}) = \hat{p}(\mathbf{x}) \qquad \mathbf{x} \in \partial V \tag{18}$$

It is now possible to calculate an approximate solution of the integral equation (7) as

$$a(\mathbf{y}) = \sum_{n=1}^{N} \frac{1}{\sigma_n} \langle p_n \,|\, p \rangle a_n(\mathbf{y}) \tag{19}$$

This equation provides a very powerful method for solving the problem of sound field reproduction with a continuous layer of monopole-like secondary sources. In fact, considering equations (15), (17) and (19), the single layer potential with the density computed in equation (18) is given by

$$\begin{aligned} &\int_{\partial\Lambda} G(\mathbf{x} \,|\, \mathbf{y}) a(\mathbf{y}) dS(\mathbf{y}) = \\ &= \int_{\partial\Lambda} G(\mathbf{x} \,|\, \mathbf{y}) \left(\sum_{n=1}^{N} \frac{1}{\sigma_n} \langle p_n \,|\, p \rangle a_n(\mathbf{y}) \right) dS(\mathbf{y}) = \\ &= \sum_{n=1}^{N} \frac{1}{\sigma_n} \langle p_n \,|\, p \rangle (Sa_n)(\mathbf{x}) = \sum_{n=1}^{N} \langle p_n \,|\, p \rangle p_n(\mathbf{x}) = \\ &= p(\mathbf{x}) - (Rp)(\mathbf{x}) = \hat{p}(\mathbf{x}) \end{aligned} \tag{20}$$

This implies that the reconstructed sound field $\hat{p}(\mathbf{x})$ is that part of the target field that does not belong to the null-space of the adjoint operator S^+. This result is consistent with the theorem reported in [24], stating that the closure of the range of S, that is the sub-space identified by all fields that can be generated by the continuous distribution of secondary sources on $\partial\Lambda$, is the orthogonal complement to the null-space of S^+. In other words, with the usual condition on the Dirichlet eigenvalues, if the target field has a pressure profile $p(\mathbf{x})$ on ∂V that can be expressed as a linear combination of the orthogonal functions $p_n(\mathbf{x})$, then it is ideally possible to determine a density $a(\mathbf{y})$ such that $\hat{p}(\mathbf{x}) = p(\mathbf{x})$ in V.

Considering equation (16), it can be noticed that even if some of the functions $p_n(\mathbf{x})$ do not rigorously belong to the null-space of S^+, their related singular value σ_n can be so small that it is possible to consider these $p_n(\mathbf{x})$ as if $(S^+ p_n)(\mathbf{x}) \approx 0$. This sheds some light on the ill-conditioning of the inverse problem represented by the integral equation (7). In fact, if for a given n the corresponding singular value σ_n is very small, then its reciprocal is very large and the norm of the density $a(\mathbf{y})$ computed with the series (19) might become unreasonably large. Furthermore, as explained in [18], the factor $1/\sigma_n$ is related to the amplification of errors contained in the data $p(\mathbf{x})$ and therefore to the stability of the system. In order to prevent this ill-conditioning problem, it is possible to regularize the solution (19), applying a cut-off to the spectrum of the operator, using the Tikhonov regularization or any other regularization technique [4], [24].

It has been shown that the possibility of calculating a density $a(\mathbf{y})$ for reconstructing the target sound field $p(\mathbf{x})$ with the single layer potential (5) from the knowledge of the boundary values of $p(\mathbf{x})$ on ∂V is strongly related to the null-space of the adjoint operator

S^+. It is important to emphasize that if it is not possible to exactly calculate the density $a(\mathbf{y})$, this does not imply that the latter does not exist and that the target sound field can not be perfectly reconstructed by the continuous distribution of secondary sources. An analogous if not identical problem arises in the field of Acoustic Holography: suppose that an attempt is made to determine the surface normal velocity of a radiating plate from the knowledge of the radiated sound field on a given region. It turns out that the vibration modes that generate evanescent waves cannot in practice be determined from far field measurements. However, this does not mean that these low-efficiency vibro-acoustic modes do not exist. This highlights the fact that the feasibility of the reconstruction of the target sound field not only depends on the arrangement $\partial\Lambda$ of the secondary sources, but also is very much related to the surface ∂V on which the sound field has been defined or measured.

3. THE EXAMPLE OF THE CONCENTRIC SPHERES

Even if equation (19) provides a very powerful method for solving the integral equation (7), the determination of the functions $a_n(\mathbf{y})$ and $p_n(\mathbf{x})$ is in general not trivial. In the vast majority of the cases, the calculation is undertaken numerically, while for some special geometries of Λ and V, an analytical calculation is possible. A special case is now considered, where the boundaries of the two volumes V and Λ are two concentric spheres, with radius R_V and R_Λ, respectively. The assumption of a free field is also made here. It can be shown that for the case under consideration the functions $a_n(\mathbf{y})$ and $p_n(\mathbf{x})$ and the singular values σ_n can be expressed analytically as

$$\begin{aligned} a_n(\mathbf{y}) &= \frac{1}{R_\Lambda} Y_n^m(\theta_y, \phi_y) \\ p_n(\mathbf{x}) &= \frac{\gamma_n}{R_V} Y_n^m(\theta_x, \phi_x) \\ \sigma_n &= R_\Lambda R_V k \left| h_n^{(1)}(kR_\Lambda) j_n(kR_V) \right| \end{aligned} \tag{21}$$

$j_n(\cdot)$ is the spherical Bessel function of order n, $h_n^{(1)}(\cdot)$ is the spherical Hankel function of the first kind and order n. The spherical harmonics $Y_n^m(\theta,\phi)$ are defined as [20]

$$Y_n^m(\theta,\phi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos\theta) e^{im\phi} \tag{22}$$

where $P_n^m(\cdot)$ are associated Legendre functions. The factor γ_n, having unitary norm, is given by

$$\gamma_n = j \frac{h_n^{(1)}(kR_\Lambda) j_n(kR_V)}{\left| h_n^{(1)}(kR_\Lambda) j_n(kR_V) \right|} \tag{23}$$

and represents a phase shift applied to each spherical harmonic of order n due to the action of S. It can be noticed that the spherical harmonics $Y_n^m(\theta,\phi)$ have two indices, while the singular values σ_n have only one index. This is due to the degeneracy of the singular values, and it implies that one eigenspace of dimension $(2n+1)$ is associated with the singular values σ_n. Hence, for each order n, it is possible to generate a set of $(2n+1)$ orthogonal spherical harmonics which span that subspace. In other words, all the spherical harmonics of order n and degree m are associated with the same singular values σ_n. This degeneracy is typical of symmetrical geometries (such as the sphere), and arises in many other fields of physics (a well known example in quantum physics is the degeneracy of two electronic configurations, which have the same energy level).

As a result of the orthogonality of the spherical harmonics (A5), it is easy to verify the mutual orthogonality of the set of functions $p_n(\mathbf{x})$ and $a_n(\mathbf{y})$ defined by equation (21). Using the expansion of the free field Green function (A1), reported in Appendix 1, it possible to show that the functions $a_n(\mathbf{y})$ and $p_n(\mathbf{x})$ and the singular values σ_n satisfy equations (12), (15) and (16).

The spherical wave spectrum $S_{nm}(r)$ of the target sound field $p(\mathbf{x})$, calculated at $r = R_V$, is defined as [20]

$$S_{nm}(R_V) = \int_0^{2\pi} d\phi \int_0^{\pi} p(R_V, \theta_x, \phi_x) Y_n^m(\theta,\phi)^* \sin(\theta) d\theta \tag{24}$$

It is now possible to calculate analytically the density $a(\mathbf{y})$ as

$$a(\mathbf{y}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{Y_n^m(\theta_y, \phi_y)}{{R_\Lambda}^2 jkh_n^{(1)}(kR_\Lambda) j_n(kR_V)} S_{nm}(R_V) \quad (25)$$

It can be observed that the denominator of (22) equals zero for those wave numbers k_n such that $j_n(k_n R_V) = 0$. These wave numbers correspond to the Dirichlet eigenvalues introduced in section 2.1, and identify the frequencies of resonance of a spherical cavity with radius R_V and pressure release boundaries (sometimes called the Dirichlet sphere). Under these circumstances, it is not possible to calculate uniquely the density $a(\mathbf{y})$ with (25) because of the non uniqueness of the solution of the Dirichlet problem (6). Considering equation (16), it is easy to notice that the function $Y_n^m(\theta_x, \phi_x)$ belongs to the null-space of S^+ and therefore is not in the range of the single layer potential S. Once again, it is important to emphasise that this problem in the reconstruction is not due to the arrangement of the layer of secondary sources on $\partial\Lambda$, but it is due to the boundary ∂V where the target sound field has been defined.

3.1. Some simplified formulations

The determination of the spherical spectrum of the target sound field involves the calculation of the integral (24). For some special cases, this integral can be solved analytically. Considering the expansion of the free field Green function (A1) it is possible to derive the following relation for the spherical spectrum of an omnidirectional point source located at $\mathbf{y_0} > R_V$ (formally a monopole with volume velocity $q_{VOL} = -(j\rho ck)^{-1}$):

$$S_{nm}^{ps}(R_V) = jkh_n^{(1)}(k|\mathbf{y_0}|) j_n(kR_V) Y_n^m(\theta_{y_0}, \phi_{y_0})^* \quad (26)$$

It important to emphasize that the source location $\mathbf{y}$ should not be in V but can be within Λ. Similarly, considering the expansion (A2) of a plane wave with unitary amplitude and wave vector $\mathbf{k_0}$, it is possible to derive its spherical spectrum as

$$S_{nm}^{pw}(R_V) = 4\pi j^n j_n(kR_V) Y_n^m(\theta_{k_0}, \phi_{k_0})^* \quad (27)$$

Finally, considering the spherical harmonic summation formula (A3) and the trigonometric relations (A4), two simplified formulae can be derived for the density for the reconstruction of the sound field due to a monopole source or a plane wave, respectively:

$$a^{ps}(\mathbf{y}) = \sum_{n=0}^{\infty} \frac{h_n^{(1)}(k|\mathbf{y_0}|)}{{R_\Lambda}^2 h_n^{(1)}(kR_\Lambda)} \frac{2n+1}{4\pi} P_n(\cos(\zeta)) \quad (28)$$

$$a^{pw}(\mathbf{y}) = \sum_{n=0}^{\infty} \frac{4\pi j^{n-1}}{k{R_\Lambda}^2 h_n^{(1)}(kR_\Lambda)} \frac{2n+1}{4\pi} P_n(\cos(\zeta)) \quad (29)$$

where $P_n(\cdot)$ is the Legendre polynomial of degree n and (see relation (A4))

$$\cos(\zeta) = \sin(\theta_y)\sin(\theta')\cos(\phi_y - \phi') + \cos(\theta_y)\cos(\theta') \quad (30)$$

and $\theta', \phi' = \theta_{y_0}, \phi_{y_0}$ for the case of a point source and $\theta', \phi' = \theta_{k_0}, \phi_{k_0}$ for a plane wave. It can be noticed that the summation over the different degrees m has been reduced to the computation of a single Legendre polynomial.

As will be discussed later on, in practical situations it can be useful (but not mandatory!) to truncate the order of the series (25), (28), (29) to a given order N. Under this assumption, for high operating frequencies such that $kR_\Lambda, k|\mathbf{y_0}| >> N$, then the large argument limits or far field approximation (A7) of the spherical Hankel functions can be used in (28), which can be rewritten as

$$a^{hf}(\mathbf{y}) = \frac{e^{jk(|\mathbf{y_0}| - R_\Lambda)}}{R_\Lambda |\mathbf{y_0}|} \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(\cos(\zeta)) \quad (31)$$

This is a very powerful formula, which is extremely useful for the real time implementation of a panning function. In fact, the only frequency dependent part of (28) is the complex exponential $e^{jk(|\mathbf{y_0}| - R_\Lambda)}$, and its Fourier transform corresponds in the time domain to a simple delay equal to the distance $|\mathbf{y_0}| - R_\Lambda$. The denominator of (28) is a simple attenuation due to the distance of the point source and the term represented by the series of Legendre polynomials is a gain factor depending on the relative angle between $\mathbf{y_0}$ and $\mathbf{y}$.

In the very special case when the distance of the virtual source $\mathbf{y_0}$ equals the sphere radius R_Λ, equation (28) can be rewritten in the very simple formulation

$$a^{real}(\mathbf{y}) = \frac{1}{{R_\Lambda}^2}\sum_{n=0}^{\infty}\frac{2n+1}{4\pi}P_n\left(\cos(\zeta)\right) \qquad (32)$$

This equation confirms the intuitive consideration that when the attempt is made to reconstruct the sound field due to a monopole source located on the boundary $\partial\Lambda$ with a continuous layer of monopole-like secondary sources, then only one monopole located at $\mathbf{y_0}$ will be active. This can be simply proved invoking the completeness relation of the spherical harmonics (A6): re-expanding the Legendre polynomial in equation (32) in terms of spherical harmonics (refer to equation (A3)) and applying relation (A6), equation (32) reduces to the Dirac delta functions

$$a(\mathbf{y}) = \delta(\phi_y - \phi_{y_0})\delta(\cos(\theta_y) - \cos(\theta_{y_0})) \qquad (33)$$

It should be pointed out that the completeness relation (A6) holds only if the series (32) is infinite. As will be shown later, in practical situations this series is truncated to a given order N and the simple panning functions (32) can in this case be considered as a spherical *sinc* function.

4. ANALOGY WITH SCATTERING THEORY

The study of the singular values in equation (21) has shown that in the case of the concentric spheres the ill-conditioning of the inverse problem (7) is directly related to the ratio between the two radii R_V and R_Λ, which governs the roll-off of the singular values σ_n. The question therefore arises of what happens when the two regions V and Λ coincide. In this case it is possible to express the density function $a(\mathbf{y})$ in the following way that results in an interesting physical interpretation. Assume that the region of the space $\Lambda = V$ identifies an object with pressure release boundaries (often called a sound soft object [4]), and assume that the target sound field $p(\mathbf{x})$ is due to sources of sound not contained in Λ. Then let $p_S(\mathbf{x})$ be the scattered sound field due to the scattering of $p(\mathbf{x})$ by the sound soft object Λ, and let the total sound field $p_T(\mathbf{x}) = p(\mathbf{x}) + p_S(\mathbf{x})$. Then the density $a(\mathbf{y})$ that allows the perfect reconstruction of the target sound field within Λ with a continuous distribution of monopole-like secondary sources on its boundary is given by

$$a(\mathbf{y}) = \nabla p_T(\mathbf{y})\cdot\mathbf{n}(\mathbf{y}) \qquad (34)$$

where $\mathbf{n}(\mathbf{y})$ is the unit vector normal to $\partial\Lambda$ at $\mathbf{y}$ and pointing towards the exterior of Λ, as shown in Figure 2. This result means that, as a result of the Euler equation, the strength of the secondary sources is proportional to the normal particle velocity of the total field (target plus scattered field) measured on the boundary $\partial\Lambda$.

This result is not proved in this paper, but the special case is analyzed of the target sound field due to a plane wave (as in equation (27)) reconstructed by a continuous distribution of secondary sources on a spherical surface $\partial\Lambda$ with radius $R = R_\Lambda = R_V$. The radial derivative of a sound field scattered by a sound soft sphere is derived in Appendix 2 and is given by equation (B5), and its value on the boundary $\partial\Lambda$ is given by

$$\begin{aligned} p_T(\mathbf{x}) &= \sum_{n=0}^{\infty}\frac{4\pi j^n k}{h_n^{(1)}(kR)}\left[h_n^{(1)}(kR)\frac{\partial j_n(kR)}{\partial(kR)} - j_n(kR)\frac{\partial h_n^{(1)}(kR)}{\partial(kR)}\right] \\ &\quad\cdot\sum_{m=-n}^{n}Y_n^m(\theta_x,\phi_x)Y_n^m(\theta_{k_0},\phi_{k_0})^* = \\ &= \sum_{n=0}^{\infty}\frac{4\pi j^{n-1}}{kR^2 h_n^{(1)}(kR)}\sum_{m=-n}^{n}Y_n^m(\theta_x,\phi_x)Y_n^m(\theta_{k_0},\phi_{k_0})^* \end{aligned} \qquad (35)$$

where the Wronskian relation [20]

$$\frac{\partial h_n^{(1)}(x)}{\partial(x)}j_n(x) - h_n^{(1)}(x)\frac{\partial j_n(x)}{\partial(x)} = \frac{j}{x^2} \qquad (36)$$

has been used. Considering the addition formula for spherical harmonics (A3), it can be easily seen that equation (29) and equation (35) coincide, as expected.

It is now possible to derive a simplified formulation for the density (34) by simple consideration of the scattering theory described in [4]. Assume that the incident field is a plane wave with wave vector $\mathbf{k_0}$. As a first step, it can be easily proved that the normal derivative of the field $\nabla p_S(\mathbf{y})\cdot\mathbf{n}(\mathbf{y})$ scattered by a pressure release infinite plane equals the normal derivative of the incident field $\nabla p(\mathbf{y})\cdot\mathbf{n}(\mathbf{y})$. Hence, for the case of an infinite plane,

$$\nabla p_T(\mathbf{y})\cdot\mathbf{n}(\mathbf{y}) = 2\nabla p(\mathbf{y})\cdot\mathbf{n}(\mathbf{y}) \qquad (37)$$

As a second step, it can be argued that, at high frequencies the scattering object Λ can be locally considered at each point $\mathbf{y} \in \partial\Lambda$ as an infinite plane tangent to $\partial\Lambda$ on $\mathbf{y}$. Finally the boundary $\partial\Lambda$ of the scattering object can be divided, as shown in Figure 3, into the *illuminated region* $\partial\Lambda_-$ and the *shadow region* $\partial\Lambda_+$, such that $\partial\Lambda_+ : \{\mathbf{y} \in \Lambda : \mathbf{n}(\mathbf{y}) \cdot \mathbf{k_0} \geq 0\}$ and $\partial\Lambda_- = \partial\Lambda \setminus \partial\Lambda_+$. It can be assumed that, at high frequencies, the total sound field $p_T(\mathbf{y})$ vanishes in the shadow region $\partial\Lambda_+$. The latter two approximations are quite accurate when the wavelength of interest is much smaller than the dimensions of the scattering object Λ. As a consequence of what has been discussed, the following approximate expression for the density $a(\mathbf{y})$ can be derived:

$$a(\mathbf{y}) \approx \begin{cases} 2\nabla p(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y}) & \mathbf{y} \in \partial\Lambda_- \\ 0 & \mathbf{y} \in \partial\Lambda_+ \end{cases} \tag{38}$$

This very powerful formulation provides a very simple (although not exact) expression for the density $a(\mathbf{y})$ for the case when the normal derivative $\nabla p(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y})$ of the target sound field is known *and* when the direction of the incoming plane wave is also known.

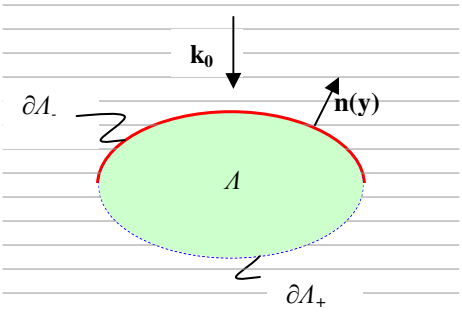

Figure 3: illuminated region $\partial\Lambda_-$ and shadow region $\partial\Lambda_+$ of the boundary of the scattering object Λ.

5. PRACTICAL REALIZATION OF A SURROUND SYSTEM

The theoretical discussion which has been presented starts with some ideal assumptions which might not hold in practical cases. If a surround system based on sound field reproduction is realized, some of these assumptions must be removed. In the first place, the continuous distribution of monopole-like secondary sources on $\partial\Lambda$ is replaced by an array of loudspeakers (or analogous electroacoustic transducers). This implies that the secondary sources can no longer be considered as monopole-like sources. Furthermore, their number is obviously finite, with the consequence that the ideal continuous distribution of secondary sources is replaced by a discrete distribution of secondary sources.

The description of the target sound field represents another practical issue. It is, in general, possible to proceed in two different ways: it is either possible to assume that the target sound field is generated by a given source with known directivity and located at a known position, or it can be stated that no *a priori* information on the target field is given. In this latter case, some pieces of information on the target field can be obtained by measuring it with a microphone array. This process implies that the target field is spatially sampled, and as a consequence the description of the field can be affected by errors like spatial aliasing. The reader can refer to the literature dedicated to microphone arrays, as for example [25] and [26], for a thorough discussion of this subject.

In the following section a mathematical analysis is presented for the case when the continuous distribution of secondary sources on $\partial\Lambda$ is replaced by an array of loudspeakers. The latter are here assumed to be omnidirectional point sources. It is worth mentioning that the extension of the theory presented in Section 4 to the case of loudspeaker with more complex directivity is in principle feasible and is the object of ongoing research.

5.1. Sampling of the density function

Let L be the number of these transducers and no assumption is made regarding their arrangement, apart from the fact that each loudspeaker position $\mathbf{y}_l$ is on the boundary $\partial\Lambda$. Let the delta function $\delta(\mathbf{y} - \mathbf{y_0})$ be such that, for any square integrable function $f(\mathbf{y})$ defined on $\partial\Lambda$,

$$\int_{\partial\Lambda} f(\mathbf{y})\delta(\mathbf{y} - \mathbf{y_0})dS(\mathbf{y}) = f(\mathbf{y_0}) \qquad \mathbf{y_0} \in \partial\Lambda \tag{39}$$

The sampled density $a_S(\mathbf{y})$ can be expressed as

$$a_S(\mathbf{y}) = \sum_{l=1}^{L} a(\mathbf{y})\delta(\mathbf{y}-\mathbf{y}_l)\Delta S_l \qquad (40)$$

where the real and positive coefficients ΔS_l have the dimension of an area and depend on the sampling scheme adopted. In the special case of regular sampling, ΔS_l is a constant that does not depend on the index l. The reconstructed sound field is expressed as

$$\hat{p}(\mathbf{x}) = \int_{\partial\Lambda} G(\mathbf{y}\,|\,\mathbf{x}) a_S(\mathbf{y}) dS(\mathbf{y}) \quad , \ \mathbf{x}\in \mathrm{V} \qquad (41)$$

It is now possible to study the effect of discretisation on the reconstruction performance. As shown in Appendix 3, the reconstructed sound field on the boundary ∂V is given by

$$\hat{p}(\mathbf{x}) = p(\mathbf{x}) + e(\mathbf{x}) \quad \mathbf{x}\in \partial\mathrm{V} \qquad (42)$$

where the reconstruction error $e(\mathbf{x})$ is defined as

$$e(\mathbf{x}) = \sum_{n=0}^{\infty}\sum_{\nu=0}^{\infty} e_{n\nu} \frac{\sigma_n}{\sigma_\nu} \langle p_\nu \,|\, p\rangle p_n(\mathbf{x}) \qquad (43)$$

$$e_{n\nu} = \sum_{l=1}^{L} a_n(\mathbf{y}_l)^* a_\nu(\mathbf{y}_l)\Delta S_l - \delta_{n\nu} \qquad (44)$$

The orthogonality error $e_{n\nu}$ is strongly related to the loudspeaker arrangement and is due to the fact that the sampled versions of the two functions $a_n(\mathbf{y})$ and $a_\nu(\mathbf{y})$ are not orthogonal.

Regular arrangement of secondary sources on a spherical surface

Once again, the case when $\partial\Lambda$ and ∂V are two concentric spheres is studied as an example. It is possible to substitute $p_n(\mathbf{x}), a_n(\mathbf{y}), \sigma_n$ with their expressions reported in equation (21) and then to use the definition (24) of the spherical spectrum to obtain

$$e_{n\nu}^{m\mu} = \frac{4\pi}{L}\sum_{l=1}^{L} Y_n^m(\theta_l,\phi_l)^* Y_\nu^\mu(\theta_l,\phi_l) - \delta_{n\nu}\delta_{m\mu} \qquad (45)$$

$$e(\mathbf{x}) = \sum_{n,m}\sum_{\nu,\mu} e_{n\nu}^{m\mu} \frac{h_n^{(1)}(kR_\Lambda) j_n(kR_V)}{h_\nu^{(1)}(kR_\Lambda) j_\nu(kR_V)} Y_n^m(\theta_x,\phi_x) S_{\nu\mu}(R_V) \quad \mathbf{x}\in\partial\mathrm{V} \qquad (46)$$

where the relation $\Delta S_l = 4\pi R_\Lambda^2 / L$ holds for an L-points ideal regular sampling of a sphere with radius R_Λ. A compact notation for the usual double summations over the different n,ν and m,μ has been used. Because of the separability of the Helmholtz equation in spherical geometry [20], equation (46) can be used also for the reconstructed sound field in the interior region V by substituting the argument of the spherical Bessel function $j_n(kR_V)$ with $j_n(k|\mathbf{x}|), \mathbf{x}\in V$. It might be useful to point out that while $e_{n\nu}^{m\mu}$ depends on the arrangement of the secondary sources, the factor $S_{\nu\mu}(R_V)/j_n(kR_V)$ depends only on the target sound field.

It has been demonstrated [7] that for a regular or almost regular arrangement of loudspeakers, the error $e_{n\nu}^{m\mu}$ is small when $n,\nu < (L+1)^2$. That is if the number of loudspeakers L is larger than the total number of spherical harmonics up to a given order n or ν. This assertion is consistent with analogous results described in previous works [7], [8], [9], and corresponds to the case of regular sampling of a function defined over a sphere. Similarly, the error $e_{n\nu}^{m\mu}$ for large orders $n,\nu > (L+1)^2$ can be interpreted as the effect of spatial aliasing. In order to attempt to reduce the reconstruction error, it is possible to discard the component of the spherical spectrum $S_{\nu\mu}(R_V)$ such that $\nu > (L+1)^2$. But if this process partially reduces the effect of aliasing due to the spatial properties of the target sound field, it introduces a new error due to the truncation of the spherical spectrum, that is

$$e_T(\mathbf{x}) = \sum_{n>N_T}^{\infty} \frac{j_n(k|\mathbf{x}|)}{j_n(kR_V)} \sum_{m=-n}^{n} Y_n^m(\theta_x,\phi_x) S_{nm}(R_V) \quad \mathbf{x}\in\mathrm{V} \qquad (47)$$

where N_T is the order of the truncation. As shown previously [7], [8], [9], the behavior of the spherical Bessel functions ensures that this error is small for small values of $k|\mathbf{x}|$.

The fact that the secondary sources are point sources is a further cause of aliasing. In fact, even if the original spherical spectrum of the target sound field is limited to a given order N and there is zero contribution to the reconstruction error (46) for $\nu > N$, there is a non-zero contribution for $\nu < N$ and $n > N$. This is due to the fact that the spherical spectrum of the sound field generated by each secondary source is not limited, as shown by the expansion of the free field Green function (A1). An analogous phenomenon happens in the usual case of a time dependent function that is regularly sampled. When the latter is reconstructed using delta functions instead of their low-pass filtered version (*sinc* functions), the reconstruction is affected by aliasing error. In the one dimensional case, this error appears as a periodic repetition of the spectrum of the reconstructed function. The extension of this spectral repetition to the case of spherical spectra is not discussed here and the reader is referred to [25].

It is useful to point out that the roll-off of the spherical spectrum of the target field is related to the distance from the origin of the virtual sources. This is demonstrated by considering equation (26) and observing the behavior of the spherical Hankel functions shown in Figure 4. It can be easily deduced that the farther is the location of the point source $\mathbf{y_0}$ from the origin, the steeper is the roll-off of the spherical spectrum S_{nm} and the smaller is the consequent contribution of high orders ν to the reconstruction error (46).

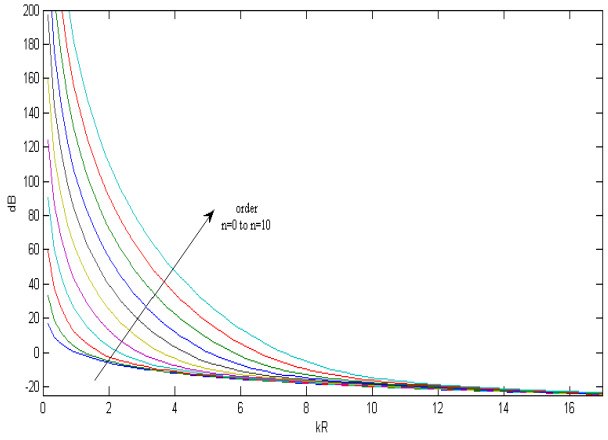

Figure 4: Absolute value of spherical Hankel functions.

5.2. Numerical solution of the integral equation

An alternative approach consists of a direct solution of the discretised version of the integral equation (7). It is possible to sample both functions $a(\mathbf{y})$ and $p(\mathbf{x})$ and then equation (7) can be rewritten in a matrix form as

$$\mathbf{p} = \mathbf{S}\mathbf{a} \tag{48}$$

where the vector $\mathbf{p}$ and $\mathbf{a}$ and the matrix $\mathbf{S}$ are defined as

$$\begin{aligned} \mathbf{p} &= [p(\mathbf{x}_1)\Delta S'_1, p(\mathbf{x}_2)\Delta S'_2, \ldots, p(\mathbf{x}_K)\Delta S'_K]^T \\ \mathbf{a} &= [a(\mathbf{y}_1)\Delta S_1, a(\mathbf{y}_2)\Delta S_2, \ldots, a(\mathbf{y}_L)\Delta S_L]^T \\ \mathbf{S} &= \begin{vmatrix} G(\mathbf{x}_1 \mid \mathbf{y}_1) & G(\mathbf{x}_1 \mid \mathbf{y}_2) & \ldots & G(\mathbf{x}_1 \mid \mathbf{y}_L) \\ G(\mathbf{x}_2 \mid \mathbf{y}_1) & & & \vdots \\ \vdots & & \ddots & \\ G(\mathbf{x}_K \mid \mathbf{y}_1) & \ldots & & G(\mathbf{x}_K \mid \mathbf{y}_L) \end{vmatrix} \end{aligned} \tag{49}$$

K is the number of sampling points of $p(\mathbf{x})$ while L is the number of sampling point of $a(\mathbf{y})$, corresponding to the number of secondary sources. Once again, the real and positive coefficients ΔS_l and $\Delta S'_k$ depend on the sampling schemes adopted. The strength of the secondary sources, represented now by vector $\mathbf{a}$, can be calculated after having computed the pseudo inverse of matrix $\mathbf{S}$ as [11]

$$\mathbf{S}^+ = (\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H \tag{50}$$

where the symbol $[\cdot]^H$ represents the Hermitian transpose of a matrix. The vector $\mathbf{a}$ can now be computed from

$$\mathbf{a} = \mathbf{S}^+\mathbf{p} \tag{51}$$

This process is analogous to the solution of the integral equation (7) described in this paper earlier on (the matrix $\mathbf{S}$ is a compact operator acting between two Hilbert spaces of finite dimension). As in the previous case, it is possible to perform a Singular Value Decomposition of the matrix $\mathbf{S}$ and to apply regularization schemes to obtain stable solutions. For the usual case of the two concentric spheres and regular sampling, it has been shown [28] that the singular values and singular vectors of matrix $\mathbf{S}$ coincide with

the expression given in (21) (up to the constant coefficients ΔS_l and $\Delta S'_k$) when both Z and L tend to infinity.

Even though is not formally proved in this paper, the sampling process can be actually considered as a kind of regularization method (the total number of the singular values of the operator is limited to $\min(Z, L)$). It could be argued that increasing both Z and L has in general the effect of achieving a more accurate reconstruction in the absence of errors (the dimension of the range of $\mathbf{S}$ is larger) but the solution is in general less stable (a larger number of singular values can lead to a higher condition number) [22].

If on one side no analytical calculation is required by this method, its implementation is, on the other side, more computationally expensive, as the numerical inversion of a matrix is involved. Furthermore, the sampling of the target sound field on the boundary ∂V leads to potentially large reconstruction errors at high frequencies due to spatial aliasing. A thorough discussion of this topic is not presented in this paper, but it is again possible to refer to the representative example of the spherical geometry - in this case the problem of aliasing due to sound field sampling as well as the problem arising from the Dirichlet eigenvalues are extensively studied in the literature dedicated to spherical microphone arrays [25], [26], [27].

6. EXPERIMENTS

The results presented in this paper have been validated experimentally. For this purpose, a system of 40 loudspeakers was assembled in the anechoic chamber of the Institute of Sound and Vibration Research (ISVR). The loudspeakers were arranged almost regularly on the surface of a sphere having a radius of $2m$. The acoustic centers of the loudspeakers were assumed to be located at $0.2m$ from their connection to the supporting structure of the system, resulting in an effective radius of the spherical array of $1.8m$. A diagram of the arrangement of the 40 loudspeakers is shown in Figure 5, while Figure 6 shows a picture of the loudspeaker array in the anechoic chamber of the ISVR.

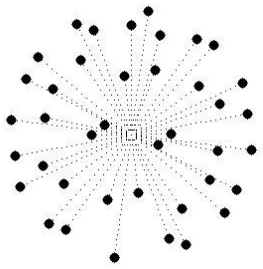

Figure 5: Diagram of the loudspeaker arrangement.

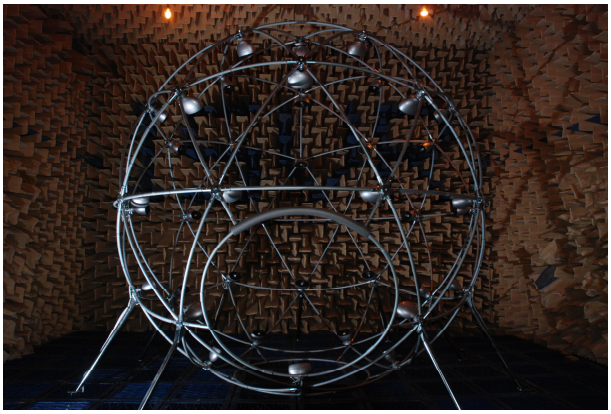

Figure 6: Picture of the loudspeaker array in the anechoic chamber of the ISVR.

The loudspeakers used are KEF® HTS3001. The measurement of their three dimensional radiation pattern was the object of a previous experiment, described in the detail in [29]. The results suggested that the device measured could approximate well an omnidirectional point source on a wide frequency range. The system was constituted also by an RME multi-channel system with MADI technology for the digital to analog conversion of the signals and by custom made power amplifiers. The sampling frequency adopted was $48kHz$. The test signal was a sine sweep with

exponentially varying frequency from *40Hz* to 19kHz. The latter was processed by a PC based system in order to generate the driving signals for the loudspeakers. The reproduced sound field was measured with a linear array of 29 omnidirectional microphones, shown in Figure 7, with a spacing $0.07m$ between each microphone. This spacing imposed the upper limit of the measurable frequency limit to be constrained below $4.9kHz$ (the measured temperature was $20°C$ and the calculated sound speed was ~$343m/s$). The phase and amplitude response of each microphone was measured and corrected (over the frequency band of interest) by applying a dedicated inverse filter. The array was then translated and the measurement repeated, in order to obtain a 29x29 measurement grid, spanning an area of approximately $2m$ x $2m$ of the horizontal plane ($\theta = 90^\circ$).

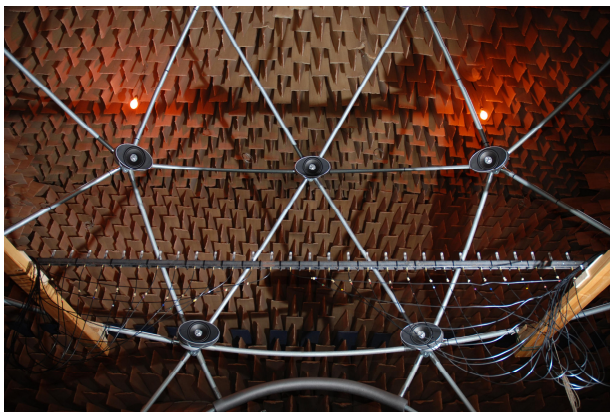

Figure 7: Picture of the microphone array.

The target of the experiment was to achieve the reconstruction of the sound field due to an omnidirectional virtual source, located in two different positions. The spherical co-ordinates describing the location of the first virtual source are $r_{PS1} = 3.23m$,$\theta_{PS1} = 77.3^\circ$, $\phi_{PS1} = 72^\circ$, and the co-ordinates of the second location are $r_{PS2} = 1.18m$,$\theta_{PS2} = 90^\circ$, $\phi_{PS2} = 0^\circ$. The measurement of the target sound field was carried out after having played back the test signal through an extra loudspeaker (KEF® HTC3001) positioned at one of the two abovementioned locations.

6.1. Computation of the digital filters

The loudspeaker signals were generated by filtering the test signal, with a bank of digital filters (one FIR filter corresponding to each loudspeaker). The filters were designed from their frequency domain formulation, which was applied to a discrete number of frequencies on the band of interest, with a frequency resolution of $11.72Hz$. This implied that the total length of the filters was 4096 points.

Two different methods were used for the computation of the filters. The first one was derived from the analytical solution of the integral equation (7) for the case of a sound field due to an omnidirectional point sources with order-limited spherical spectrum. The frequency domain expression of these filters is reported in equation (28), with the only difference that the series was truncated to the order $N = 5$. Figure 8 and Figure 9 report the frequency domain and the time domain representation, respectively, of one of the digital filters computed with this method for the case of the virtual point source located at $r_{PS1} = 3.23m$,$\theta_{PS1} = 77.3^\circ$, $\phi_{PS1} = 72^\circ$. The plotted filter is related to the loudspeaker located at $r_\Lambda = 1.8m$, $\theta_9 = 63.4^\circ$, $\phi_9 = 72^\circ$.

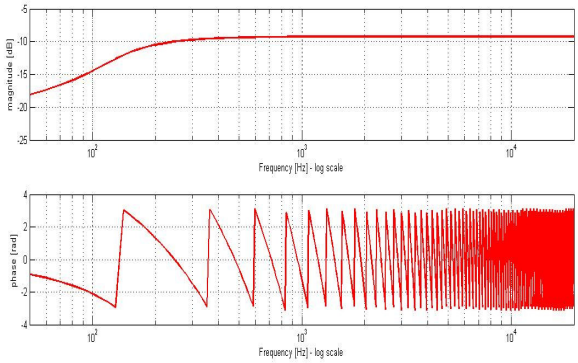

Figure 8: Frequency domain representation (magnitude and phase) of one of the digital filters calculated with the first method.

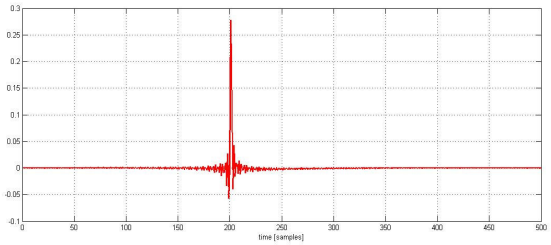

Figure 9: Time domain representation of one of the digital filters calculated with the first method.

The second method implies the numerical solution of the integral equation (7), described by equation (49), (50) and (51). The latter was calculated, as in the previous case, for a discrete number of frequency values with the resolution of $11.72Hz$. In order to avoid problems in the filter calculation arising from spatial aliasing, ill-conditioning or from the Dirichlet eigenvalues, the vector target sound field $\mathbf{p}$ was defined as the simulated sound field generated by the given virtual point source, scattered by a rigid sphere of a $0.1m$ radius centered in the origin and sampled on the surface of two concentric spheres, with radius $0.1m$ and $0.3m$, respectively. The rigid sphere was introduced in order to avoid problems related to the Dirichlet eigenvalues [7], [9]. 120 sampling positions were defined on each spherical surface, following an approximately regular spherical sampling scheme. In the frequency range below $1.5kHz$, all 240 sampling points were used, while for the upper frequency range only the 120 sampling points on the smaller sphere were considered, in order to reduce the effect of spatial aliasing. As a consequence of the presence of the scattering sphere, the Green function appearing in equation (49) corresponds to the Neumann Green function representing the solution of the inhomogeneous Helmholtz equation in a free field in the presence of the considered scattering sphere [20].

Figure 10 and Figure 11 report the representation in the frequency and the time domain of one of the digital filters computed with the second method, for the same loudspeaker and virtual source location as in Figure 8 and 9.

Observing Figure 8 and Figure 10, it can be noticed that the filters designed by applying the two different methods are in fact very similar in the frequency range below $4kHz$, apart from the minor discontinuity at $1.5kHz$ in Figure 10 due to the previously mentioned change of dimension of vector $\mathbf{p}$ (from 240 to 120 sampling points). Above $4kHz$, the filters computed with the two different methods are noticeably different. The analysis of the other 39 filters highlights the erratic behavior at high frequencies of the filters computed with the second method. This can be interpreted as an artifact of spatial aliasing arising from the sampling of the target sound field, and causes errors in the reproduction.

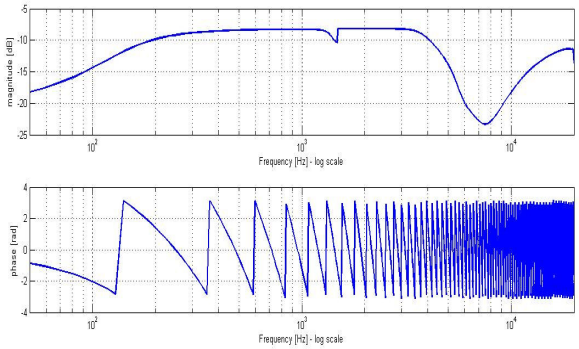

Figure 10: Frequency domain representation (magnitude and phase) of one of the digital filters calculated with the second method.

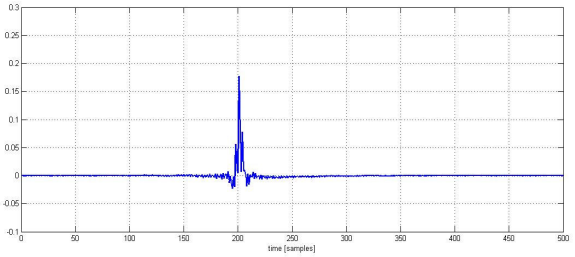

Figure 11: Time domain representation of one of the digital filters calculated with the second method.

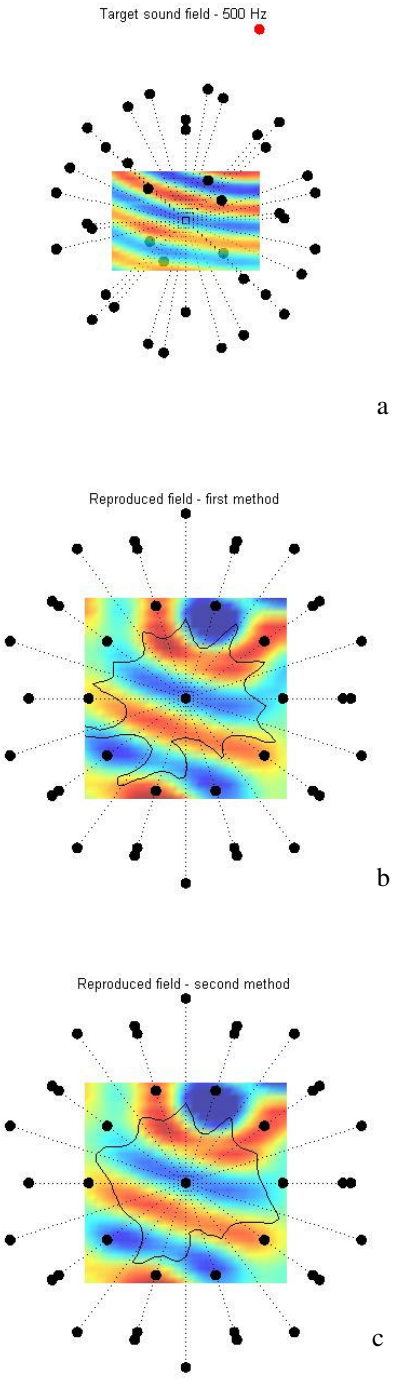

Figure 12: Measurement of the target (a) and reconstructed (b-c) sound field for the first location of the virtual source and for a frequency of 500Hz.

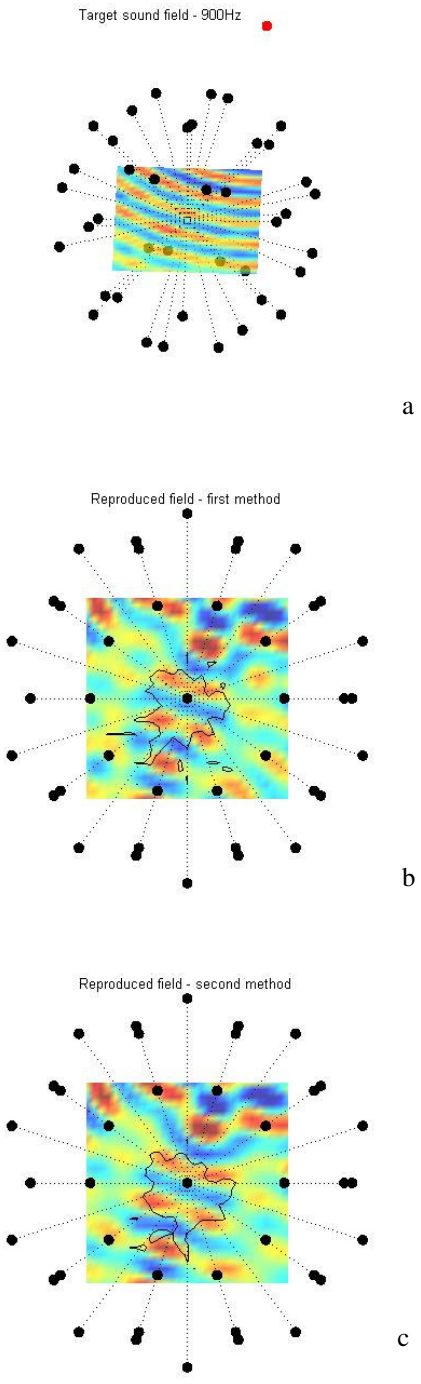

Figure 13: Measurement of the target (a) and reconstructed (b-c) sound field for the first location of the virtual source and for a frequency of 900Hz.

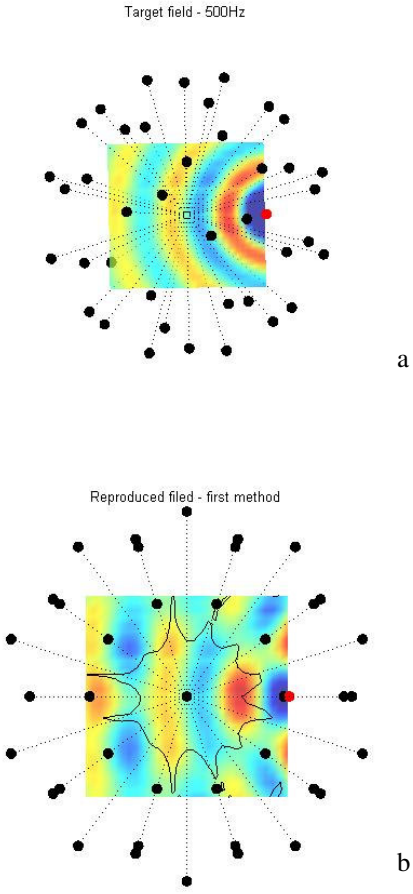

Figure 14: Measurement of the target (a) and reconstructed (b) sound field for the second location of the virtual source and for a frequency of 900Hz.

6.2. Experimental results

Figure 12a represents the first target sound field, generated by a source placed at $r_{PS1} = 3.23m$, $\theta_{PS1} = 77.3^\circ$, $\phi_{PS1} = 72^\circ$, for a frequency of 500Hz. The red circle represents the position of the virtual source. The measurements were performed using the microphone array described on the horizontal plane ($\theta = 90^\circ$). Figure 12b and Figure 12c represent the same field reconstructed by the loudspeaker array, implementing the digital filters designed with the first and second method respectively. The continuous contour line defines the area within which the normalized reconstruction error, defined as

$$NRE(\mathbf{x}) = \frac{\left|\hat{p}(\mathbf{x}) - p(\mathbf{x})\right|^2}{\left|p(\mathbf{x})\right|^2} 100 \qquad (52)$$

is less than 20%.

Analogously, Figure 13a, Figure 13b and Figure 13c represent the target and reconstructed field, for the same virtual source location but for a frequency of 900 Hz. Finally, Figure 14a represents the target field for the case of a point source located at $r_{PS2} = 1.18m$, $\theta_{PS2} = 90^\circ$, $\phi_{PS2} = 0^\circ$ (inside the loudspeaker array) for a frequency of 500Hz, while Figure 14b represents its reconstruction with the first method.

It can be observed that the reconstruction performance of the two methods is similar for the two considered frequencies. These experimental results demonstrate the feasibility of the sound field reconstruction using the presented method. The size of the reconstruction area depends, as expected, on the frequency of the sound field. A detailed analysis of the measured reconstruction error and its comparison with its theoretical formulation is the topic of ongoing work.

7. CONCLUSIONS

The theoretical fundamentals of a new method for the reproduction of a target sound filed has been introduced and the practical realization of a multi-channel system based on this technology has been presented. The experimental results have demonstrated that a good reconstruction of the target field is achievable over an area which varies as a function of the frequency.

It is important to emphasize that, even though the experiments were carried out using a system constituted by a spherical loudspeaker array, the theoretical approach presented in this paper can be extended to system with different geometries, which will exhibit different reconstruction performance. As it has been said, analytical expression of the density of the single layer potential can be derived for special geometries.

Finally, it is possible to argue that the approach proposed can be considered as a general theoretical

framework for the study of sound field reconstruction. Even if this subject is not formally discussed in this paper, it is possible to observe the correspondence of some of the theoretical results presented here with the formulations of Wave Field Synthesis, High Order Ambisonics and some other methods, which could be therefore interpreted as special cases of the general theory presented in this paper.

8. ACKNOWLEDGEMENTS

This work was supported by the IT R&D program of MIC/IITA. [2007-S004-01,Development of Glassless Single-User 3D Broadcasting Technologies].

9. REFERENCES

[1] A.J. Berkhout, "A holographic approach to acoustic control." *Journal of the Audio Engineering Society*, vol. 36, pp. 977–995, (1988).

[2] D. de Vries, E.W. Start, V.G. Valstar, "The Wave Field Synthesis concept applied to sound reinforcement: Restrictions and solutions," in *96th Convention of the Audio Engineering Society*, (1994).

[3] S. Spors, H. Buchner, R. Rabenstein, "Efficient active listening room compensation for Wave Field Synthesis," in 116th *Convention of the Audio Engineering Society* (2004).

[4] D. Colton and R. Kress, "Inverse Acoustic and Electromagnetic Scattering Theory" Springer-Verlag (1992).

[5] M.A. Gerzon "Periphony : With-height sound reproduction*." Journal of the Audio Engineering Society* 21 (1), 2-10 (1973).

[6] J. S. Bamford, J. Vanderkooy, "Ambisonic Sound for Us," in *99th Convention of the Audio Engineering Society*, (1995).

[7] J. Daniel, R. Nicol, S. Moureau, "Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging," in *114th Convention of the Audio Engineering Society*, (2003).

[8] D. B. Ward, T. D. Abhayapala "Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers." *IEEE Transaction on Speech and Audio Processing, 9*, 697-707 (2001).

[9] M. Poletti, "Three-Dimensional Surround System Based on Spherical Harmonics". *Journal of the Audio Engineering Society*, Vol. 53, No. 11 (2005).

[10] Y. Jennifer WU and T.D. Abhayapala, ``Soundfield reproduction using theoretical continuous loudspeaker", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. X, pp., Las vegas, USA (2008).

[11] O.Kirkeby and P.A. Nelson "Reproduction of plane wave sound fields." *Journal of the Acoustical Society of America* 94, 2992-3000 (1993).

[12] S. Ise, "A Principle of Sound Field Control Based on the Kirchhoff–Helmholtz Integral Equation and the Theory on Inverse Systems," *Acustica—Acta Acustica,* vol. 85, pp. 78–87 (1999).

[13] T. Betlehem, T. Abhayapala, "Theory and Design of Sound Field Reproduction in Reverberant Rooms," *Journal of the Acoustical Society of America,* vol. 117, pt. 1, pp. 2100–2111 (2005).

[14] J.Ahrens and S. Spors, "Reproduction of a plane-wave sound field using planar and linear arrays of loudspeakers.", *IEEE Int. Symposium on Communications Control and Signal Processing* (ISCCSP) (2008)

[15] W.H. Cho, J.G. Ih and M.M. Boone, "Holographic Design of Source Array for Achieving a Desired Sound Field" in *19th Convention of the Audio Engineering Society* (2001).

[16] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, (1997).

[17] A. Farina et al. "Ambiophonic Principles for the Recording and Reproduction of Surround Sound for Music", in *19th Convention of the Audio Engineering Society* (2001).

[18] F.M. Fazi and P.A. Nelson, "Application of functional analysis to the sound field reconstruction", *proceedings of the Institute of Acoustics, 23rd conference on Reproduced Sound* (2007)

[19] Morse, P.M., and H. Feshbach, "Methods of theoretical physics", Part I : chapters 1 to 8: McGraw Hill Book Co, (1953).

[20] E. Williams, "Fourier acoustics. Sound Radiation and Nearfield Acoustical Holography". New York: Academic (1999).

[21] E. Kreyszig, "Introductory Functional Analysis with Applications", Wiley Classics Library (1978)

[22] G.A. Deschamps and H.S. Cabayas, "Antenna Synthesis and Solution of Inverse Problems by Regularisation Methods" *IEEE transactions on antennas and propagation*, Vol. AP-20, no.8 (1972)

[23] F.M. Fazi and P.A. Nelson, "The ill-conditioning problem in sound field reconstruction", *presented at the 123th Convention of the Audio Engineering Society*, convention paper 7244 (2007).

[24] R.Potthast, "Introduction to Functional Analysis and Inverse Problems", Lecture Notes (2007)

[25] B. Rafaely, "Spatial Aliasing in Spherical Microphone Arrays", *IEEE transactions on signal processing,* Vol. 55, no.3 (2007)

[26] J. Meyer and G. W. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)* vol. II, pp. 1781–1784 (2002)

[27] M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposition using spherical microphone array", *Journal of the Acoustical Society of America* Vol. 118, no.15, 3094-3103 (2005).

[28] F.M. Fazi and P.A. Nelson " A Theoretical study of sound field reconstruction techniques" *Proceedings of the 19th International Congress on Acoustics* (2007)

[29] F.M. Fazi, V. Brunel, P.A. Nelson and J. Seo "Measurement and Fourier-Bessel analysis of loudspeaker radiation patterns using a spherical array of microphones", *presented at the 124th Convention of the Audio Engineering Society*, (2008).

10. APPENDICES

Appendix 1. Useful formulae

Spherical harmonic expansion of the free field Green function [20]

$$\frac{e^{jk|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} = \sum_{n=0}^{\infty} jkh_n^{(1)}(k|\mathbf{y}|)j_n(k|\mathbf{x}|)\sum_{m=-n}^{n} Y_n^m(\theta_x,\phi_x)Y_n^m(\theta_y,\phi_y)^* \tag{A1}$$

Spherical harmonic expansion of a plane wave [20]

$$e^{j\mathbf{k_0}\cdot\mathbf{x}} = 4\pi\sum_{n=0}^{\infty} j^n j_n(k|\mathbf{x}|)\sum_{m=-n}^{n} Y_n^m(\theta_x,\phi_x)Y_n^m(\theta_{k_0},\phi_{k_0})^* \tag{A2}$$

Summation formula of the spherical harmonics [4]

$$\sum_{m=-n}^{n} Y_n^m(\theta,\phi) Y_n^m(\theta',\phi')^* = \frac{2n+1}{4\pi} P_n\left(\cos(\zeta)\right) \tag{A3}$$

where $P_n(\cdot)$ is the Legendre polynomial of degree n and ζ is the angle between the directions identified by θ,ϕ and θ',ϕ'. It holds that

$$\begin{aligned}\cos(\zeta) &= \cos(\phi)\sin(\theta)\cos(\phi')\sin(\theta') + \sin(\phi)\sin(\theta)\sin(\phi')\sin(\theta') + \cos(\theta)\cos(\theta') = \\ &= \sin(\theta)\sin(\theta')\cos(\phi-\phi') + \cos(\theta)\cos(\theta')\end{aligned} \tag{A4}$$

Orthogonality of the spherical harmonics [20]

$$\int_0^{2\pi} d\phi \int_0^{\pi} Y_n^m(\theta,\phi) Y_{n'}^{m'}(\theta,\phi)^* \sin(\theta) d\theta = \delta_{nn'}\delta_{mm'} \tag{A5}$$

Completeness relation of the spherical harmonics [20]

$$\sum_{n=0}^{\infty}\sum_{m=-n}^{n} Y_n^m(\theta,\phi) Y_n^m(\theta',\phi')^* = \delta(\phi-\phi')\delta(\cos(\theta)-\cos(\theta')) \tag{A6}$$

Large argument approximation of spherical Hankel functions ($x \to \infty$) [20]

$$h_n^{(1)}(x) \approx (-j)^{n+1} \frac{e^{jx}}{x} \tag{A7}$$

Appendix 2. Scattering by a sound soft sphere

The sound field $p_S(\mathbf{x})$ scattered by an impenetrable object is a radiating solution of the Helmholtz equation, which can be expressed as [4]

$$p_S(\mathbf{x}) = \sum_{n=0}^{\infty} B_{nm} h_n^{(1)}(k|\mathbf{x}|) \sum_{m=-n}^{n} Y_n^m(\theta_x,\phi_x) \tag{B1}$$

If the incident field is a plane wave, the expansion of which is given by (A2), and the scattering object is a sound soft ball of radius R, then the total sound field must vanish on its boundary $\partial\Lambda$ and the following condition must be satisfied

$$0 = p(\mathbf{x}) + p_S(\mathbf{x}) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left[4\pi \quad j^n j_n(k|\mathbf{x}|) Y_n^m(\theta_{k_0},\phi_{k_0})^* + B_{nm} h_n^{(1)}(k|\mathbf{x}|)\right] \; Y_n^m(\theta_x,\phi_x) \tag{B2}$$

$\mathbf{x} \in \partial\Lambda$

This imposed condition implies that

$$B_{nm} = -4\pi j^n Y_n^m(\theta_{k_0},\phi_{k_0})^* \frac{j_n(kR)}{h_n^{(1)}(kR)} \tag{B3}$$

Therefore the total sound field is

$$p_T(\mathbf{x}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} 4\pi \quad j^n \left[j_n(k|\mathbf{x}|) - \frac{j_n(kR)}{h_n^{(1)}(kR)} h_n^{(1)}(k|\mathbf{x}|) \right] \quad Y_n^m(\theta_x, \phi_x) Y_n^m(\theta_{k_0}, \phi_{k_0})^* \tag{B4}$$

and it radial derivative is

$$p_T(\mathbf{x}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} 4\pi \quad j^n k \left[\frac{\partial j_n(k|\mathbf{x}|)}{\partial(k|\mathbf{x}|)} - \frac{j_n(kR)}{h_n^{(1)}(kR)} \frac{\partial h_n^{(1)}(k|\mathbf{x}|)}{\partial(k|\mathbf{x}|)} \right] \quad Y_n^m(\theta_x, \phi_x) Y_n^m(\theta_{k_0}, \phi_{k_0})^* \tag{B5}$$

Appendix 3. Single layer potential with sampled density

It is assumed that the target pressure profile $p(\mathbf{x}), \mathbf{x} \in \partial V$ has zero orthogonal projection on the null-space of the adjoint operator S^+ (which means that the target pressure profile is defined on a finite dimensional Hilbert space which is a subset of the closure of the range of the operator S) and it is also assumed that the operating wave number is not one of the Dirichlet eigenvalues k_n. Then the target sound field can be perfectly reconstructed in V by the single layer potential (5) with the continuous density $a(\mathbf{y})$ expressed by equation (19).

Using this result, the action of the single layer potential S on the sampled density function $a_S(\mathbf{y})$ can be expressed in terms of its eigenvectors and eigenvalues by applying equation (18). Hence

$$\begin{aligned}
(Sa_S)(\mathbf{x}) &= \sum_{n=1}^{N} \sigma_n \langle a_n | a_S \rangle p_n(\mathbf{x}) = \sum_{n=1}^{N} \sigma_n p_n(\mathbf{x}) \int_{\partial\Lambda} a_n(\mathbf{y})^* \sum_{l=1}^{L} \delta(\mathbf{y} - \mathbf{y}_l) a(\mathbf{y}) \Delta S_l dS(\mathbf{y}) = \\
&= \sum_{n=1}^{N} \sigma_n p_n(\mathbf{x}) \sum_{l=1}^{L} \Delta S_l \int_{\partial\Lambda} a_n(\mathbf{y})^* \delta(\mathbf{y} - \mathbf{y}_l) \sum_{\nu=1}^{N} \frac{1}{\sigma_\nu} \langle p_\nu | p \rangle a_\nu(\mathbf{y}) dS(\mathbf{y}) = \\
&= \sum_{n=1}^{N} \sigma_n p_n(\mathbf{x}) \sum_{\nu=1}^{N} \frac{1}{\sigma_\nu} \langle p_\nu | p \rangle \sum_{l=1}^{L} a_n(\mathbf{y}_l)^* a_\nu(\mathbf{y}_l) \Delta S_l
\end{aligned} \tag{C1}$$

Using the definition (43) and (44) of $e(\mathbf{x})$ and $e_{n\nu}$ respectively, it is possible to write

$$\begin{aligned}
(Sa_S)(\mathbf{x}) &= \sum_{n=1}^{N} \sigma_n p_n(\mathbf{x}) \sum_{\nu=1}^{N} \frac{1}{\sigma_\nu} \langle p_\nu | p \rangle (\delta_{n\nu} + e_{n\nu}) = \sum_{n=1}^{N} \langle p_n | p \rangle p_n(\mathbf{x}) + e_{n\nu} \sum_{n=1}^{N} \sum_{\nu=1}^{N} \frac{\sigma_n}{\sigma_\nu} \langle p_\nu | p \rangle p_n(\mathbf{x}) = \\
&= p(\mathbf{x}) + e(\mathbf{x})
\end{aligned} \tag{C2}$$