Audio Engineering Society

Convention Paper 7244

Presented at the 123rd Convention
2007 October 5–8 New York, NY, USA

The ill-conditioning problem in Sound Field Reconstruction

Filippo M. Fazi[1] and Philip A.Nelson[1]

[1] Institute of Sound and Vibration Research, University of Southampton, SO171BJ, United Kingdom
ff1@isvr.soton.ac.uk

ABSTRACT

A method for the analysis and reconstruction of a three dimensional sound field using an array of microphones and an array of loudspeakers is presented. The criterion used to process the microphone signals and obtain the loudspeakers signals is based on the minimisation of the least-square error between the reconstructed and the original sound field. This approach requires the formulation of an inverse problem that can lead to unstable solutions due to the ill-conditioning of the propagation matrix. The concepts of generalised Fourier transform and singular value decomposition are introduced and applied to the solution of the inverse problem in order to obtain stable solutions and to provide a clear understanding of the regularisation method.

1. INTRODUCTION

The problem of analyzing and reconstructing a sound field is of basic importance in many branches of acoustics. In the audio industry, the reconstruction of a recorded or synthesized sound field has become a topic of increasing interest because of the diffusion of multi-channel audio systems, which are very indicated for this kind of application. Theories like Wave Field Synthesis [1] and Ambisonics [2] have been applied with success to the realization of large audio systems for multiple listeners, and a considerable effort in the research has been dedicated to further studies related to these theories [3,4,5,6].

The technique for the reconstruction of a sound field presented in this paper is based on the mean squared error minimisation method used in the active control of sound and vibrations [7], and requires a combined use of loudspeakers and microphones. Some applications of this method for audio purposes were proposed for reconstruction of plane waves [8]. The aim of this work is to present an insight of the theory of this method applied to multi-channel audio systems, with particular attention to the problem of ill-conditioning. The electro-acoustics transducers of the multi-channel systems are not considered as independent units but as elements of a complex system, represented by a matrix, called the propagation matrix. The array of loudspeakers and the array of microphones are described as systems with multiple degrees of freedom, represented by independent array modes. It is shown that when attempting the reconstruction of a sound filed, some modes of the microphone array, corresponding to specific combination of the signals of the different microphones, can be reconstructed more "easily" than others by the loudspeaker array. For those modes, the presence of noise and errors in the measurements of the sound field does not affect dramatically the reconstruction. The singular value decomposition is a powerful mathematical tool that allows to analyse the conditioning of the propagation matrix and to identify which array modes are robust and which array modes are more problematic.

The basic theory of the least squares method applied to multi-channel systems is introduced in the first section. Then the concepts of ill-posedness of a problem and ill-conditioning of a matrix are defined in a mathematical sense and their importance for engineering application is highlighted. Then the concept of loudspeaker modes and the singular value decomposition are presented and discussed in the detail. Two examples are finally presented in order to clarify the practical implication of the mathematical results. In the first example the singular value decomposition is applied to simple system of 2 microphones and 2 loudspeakers. The second example involves a system constituted by two large spherical arrays.

THEORY

A monochromatic sound field defined over a closed, bounded and source free region $\Omega \subset R^3$ can be represented by a function

$$p(\mathbf{x},t) = \mathrm{Re}\left[p(\mathbf{x}) e^{j\omega t} \right] \qquad (1)$$

where ω is the angular frequency and $p(\mathbf{x})$ is a complex scalar field that satisfies the homogeneous Helmholtz equation

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

and the wave number k is defined by the dispersion relation $k = \omega / c_0$. The speed of sound c_0 is assumed to be uniform in Ω. Let p_q be the complex value of $p(\mathbf{x})$ at the position $\mathbf{x} = \mathbf{x_q}$. p_q can be obtained by a Fourier transform of the signal captured by an ideal omnidirectional microphone located in that position. $p(\mathbf{x})$ can be sampled using an array of Q omnidirectional microphones located at arbitrary positions $\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_q}, ..., \mathbf{x_Q} \in \Omega$ and can be used to extract Q samples from the sound field, which can be represented by a column vector (the microphone vector)

$$\mathbf{p} = \left[p_1, p_2, ..., p_q, ... p_Q \right]^T \qquad (3)$$

$\mathbf{p}$ is an element of $C^{\mathbf{Q}}$ (the Q-dimensional complex vector space). The elements p_q are, from here on, referred to as microphone coefficients. In general, for a finite number of microphones, the same microphone vector could correspond to different sound fields and therefore vector $\mathbf{p}$ does not usually identify $p(\mathbf{x})$ uniquely. This phenomenon is known as spatial aliasing and is of relevance for many applications, but will not be discussed any further in this paper.

It is possible to attempt to reconstruct $p(\mathbf{x})$ in a different environment using an array of L loudspeakers, located at arbitrary positions $\mathbf{y_1}, \mathbf{y_2}, ..., \mathbf{y}_l, ..., \mathbf{y_L}$ outside Ω and arbitrarily oriented, each of them driven by a signal $a_l(t) = \mathrm{Re}\left[a_l e^{j\omega t} \right]$, where a_l is a complex (frequency dependent) constant, referred to in what follows as a loudspeaker coefficient. As for the microphone vector, it is possible to define the loudspeaker vector

$$\mathbf{a} = \left[a_1, a_2, ..., a_l, ... a_L \right]^T \qquad (4)$$

which is an element of $C^{\mathbf{L}}$. The squared norm of the loudspeaker vector, denoted by $\|\mathbf{a}\|^2$, where

$$\|\mathbf{a}\| = \sqrt{\sum_{l=1}^{L} |a_l|^2} \qquad (5)$$

is representative of the total energy of the signals driving the loudspeakers.

As the microphone vector represents all of the available information about the original sound field $p(\mathbf{x})$, it is reasonable to try to generate with the loudspeakers array a reconstructed sound field $\hat{p}(\mathbf{x})$ such that its corresponding microphone vector $\hat{\mathbf{p}}$, obtained by sampling $\hat{p}(\mathbf{x})$ using the previously defined microphone array, is as close as possible to $\mathbf{p}$. In other words, the target is to minimise the distance between $\hat{\mathbf{p}}$ and $\mathbf{p}$. For the metric of $C^{\mathbf{Q}}$, the distance is defined as

$$d(\mathbf{p},\hat{\mathbf{p}}) := \|\mathbf{p} - \hat{\mathbf{p}}\| = \sqrt{\sum_{q=1}^{Q} \left| p_q - \hat{p}_q \right|^2} \qquad (6)$$

and represents the root mean squared error between the Q samples of $p(\mathbf{x})$ and the Q samples of $\hat{p}(\mathbf{x})$. In order to solve this minimization problem, the microphone vector $\hat{\mathbf{p}}$ must be expressed in terms of the loudspeaker vector $\mathbf{a}$.

If the loudspeakers are assumed to be linear systems, it is possible to define the linear electro-acoustical transfer function $H_{ql}(\omega)$ between the input of the *l*-th loudspeaker and the *q*-th microphone of the array positioned in the environment where the reconstruction is performed. The set of all the transfer functions between each loudspeaker can be represented by the propagation matrix $\mathbf{H}$ (later on also called propagation operator), of dimension $Q \times L$. Hence the microphone vector of the reconstructed sound field can be expressed as

$$\hat{\mathbf{p}} = \mathbf{H}\mathbf{a} \qquad (7)$$

If this expression is substituted in (6), then the minimization problem becomes

$$d(\mathbf{p},\hat{\mathbf{p}}) := \|\mathbf{p} - \mathbf{H}\mathbf{a}\| \qquad (8)$$

It can be proved [8] that, if the propagation matrix $\mathbf{H}$ is full rank, then the choice of the loudspeaker vector $\mathbf{a}_0$ that minimises the distance $d(\mathbf{p},\hat{\mathbf{p}})$ (or alternatively the mean squared error) is unique and is given by

$$\mathbf{a}_0 = \left(\mathbf{H}^{\mathbf{H}}\mathbf{H}\right)^{-1} \mathbf{H}^{\mathbf{H}}\mathbf{p} = \mathbf{H}^{+}\mathbf{p} \qquad (9)$$

The symbol $[\cdot]^{\mathbf{H}}$ denotes the Hermitian transpose of a matrix and the symbol $[\cdot]^{-1}$ denotes the inverse matrix. Matrix $\mathbf{H}^{\mathbf{H}}$ is called the adjoint matrix (or adjoint operator) of the propagation matrix. $\mathbf{H}^{+}$ is called the Moore-Penrose pseudo-inverse matrix of $\mathbf{H}$.

2. ILL-CONDITIONING PROBLEM

Even if equation (9) gives a unique solution to the least squares problem (8), it does not explicitly shows that there are subsets of microphone vectors $\mathbf{p}$ that can not be reconstructed and subsets of vectors that can be reconstructed perfectly. Furthermore, the reconstruction of different vectors $\mathbf{p}$ having the same norm could imply the usage of a considerably different amount of energy by the loudspeaker array. The knowledge of the robustness of a solution is also of great importance for engineering applications, as it is important to know what the effects of noise and measurement errors are on the sound field reconstruction. It is useful to define, in a mathematical sense, the concept ill-posedness of a problem. According to the definition of well-posed equation given by Hadamard [9], if an operator $\mathbf{H}$ from a subset U of a normed space X into a normed space Y is defined, a problem of the form $\mathbf{p} = \mathbf{H}\mathbf{a}$ is said to be ill posed if [10] one of the following three conditions is verified:

- the solution $\mathbf{a}$ does not exist for all $\mathbf{p} \in Y$;
- the solution $\mathbf{a}$ is not unique;
- the solution $\mathbf{a}$ does not depend continuously on the data $\mathbf{p}$ (instability).

The instability is of primary importance for engineering applications, and

In the case of the sound field reconstruction problem the least squares solution $\mathbf{a}_0$ always exist, but for some arrangements of microphone and loudspeaker arrays it is often possible that either the distance $\|\mathbf{p} - \hat{\mathbf{p}}\| > 0$, that means the reconstruction of $\mathbf{p}$ is not perfect, or that small errors in the sampling $\mathbf{p}$ or small variations in the estimation of matrix $\mathbf{H}$ result in large loudspeaker gains and in large errors $\|\mathbf{p} - \hat{\mathbf{p}}\|$. In this case, the propagation matrix is said to be ill-conditioned. The concept of array modes and Singular Value Decomposition are now introduced as these concepts are very useful for the study and physical understanding of the ill-conditioning of matrix $\mathbf{H}$.

2.1. Array modes

It has been shown that all the vectors $\mathbf{a}$, obtained with all the possible combinations of loudspeaker coefficients a_l, belong to the complex vector spaces of dimension L. This means that each coefficient a_l identifies a one dimensional subspace of $C^{\mathbf{L}}$ that can be identified by a unitary complex vector $\mathbf{a}_l$. In other words, it is possible to express $\mathbf{a}$ as

$$\mathbf{a} = \sum_{l=1}^{L} a_l \mathbf{a}_l \qquad (10)$$

The vectors $\mathbf{a}_l$ are mutually orthogonal and represent the canonical orthonormal base for $C^{\mathbf{L}}$. It is possible to identify another set of respectively L non zero, orthogonal vectors $\left\{\mathbf{v_n} \in C^L\right\}$, which are an orthonormal basis for the vector space to which they belong. It is possible to call the vectors $\mathbf{v_n}$ loudspeaker array modes, for a reason that will become clearer later on. It holds that:

$$\begin{aligned}
&\mathbf{v_n} = [v_{n1}, v_{n2}, \ldots, v_{nL}] \\
&\|\mathbf{v_n}\| = 1 \\
&\langle \mathbf{v_n} | \mathbf{v_m} \rangle = \sum_{l=1}^{L} \overline{v_{nl}} v_{ml} = \delta_{nm} \\
&n, m = 1, 2, \ldots, L
\end{aligned} \qquad (11)$$

where δ_{nm} is the Kronecker delta ($\delta_{nm} = 1$ if $n = m$ and $\delta_{nm} = 0$ otherwise). It follows that any vector $\mathbf{a}$ can be expressed as

$$\begin{aligned}
&\mathbf{a} = \sum_{n=1}^{L} \langle \mathbf{v_n} | \mathbf{a} \rangle \mathbf{v_n} \\
&\langle \mathbf{v_n} | \mathbf{a} \rangle = \sum_{l=1}^{L} \overline{v_{nl}} a_l
\end{aligned} \qquad (12)$$

This basic result implies that it is possible to describe the sound field generated by the loudspeaker array $\hat{p}(\mathbf{x})$ in terms of the sound fields generated by L independent combination of loudspeaker array modes $\mathbf{v_n}$.

What has been said for the loudspeaker array can also be proved for the microphone array. It is possible to define an orthogonal set of Q vectors $\left\{\mathbf{u_m} \in C^Q\right\}$, called microphone array modes, such that each microphone vector $\mathbf{p}$ can be expressed as

$$\begin{aligned}
&\mathbf{p} = \sum_{m=1}^{Q} \langle \mathbf{u_m} | \mathbf{p} \rangle \mathbf{u_m} \\
&\langle \mathbf{u_m} | \mathbf{p} \rangle = \sum_{q=1}^{Q} \overline{u_{mq}} p_q
\end{aligned} \qquad (13)$$

This result is less intuitive than the loudspeaker case: the sampled version of the sound field $p(\mathbf{x})$ can be expressed as the linear combination of Q independent microphone array modes.

If two matrices $\mathbf{V}$ and $\mathbf{U}$ are created by putting side by side all the column vectors $\mathbf{v_n}$ and $\mathbf{u_n}$ respectively, then it is possible to rewrite equation (12) and (13) in a matrix formulation

$$\begin{aligned}
&\mathbf{a} = \mathbf{V}\mathbf{V}^{\mathbf{H}}\mathbf{a} \qquad V_{ln} = v_{ln} \\
&\mathbf{p} = \mathbf{U}\mathbf{U}^{\mathbf{H}}\mathbf{p} \qquad U_{qm} = u_{qm}
\end{aligned} \qquad (14)$$

It is easy to see that matrix $\mathbf{V}$ and $\mathbf{U}$ are unitary, that is

$$\begin{aligned}
&\mathbf{V}\mathbf{V}^{\mathbf{H}} = \mathbf{I_L} \\
&\mathbf{U}\mathbf{U}^{\mathbf{H}} = \mathbf{I_Q}
\end{aligned} \qquad (15)$$

where $\mathbf{I_L}$ is the $L \times L$ identity matrix. As a simple example of array modes, the very basic case of the system composed of two loudspeakers and two omnidirectional microphones arranged as in Figure 1 is presented. The canonical bases for the loudspeaker and microphone array are respectively

$$\begin{aligned} \mathbf{a_1} &= [1,0] \quad \mathbf{a_2} = [0,1] \\ \mathbf{p_1} &= [1,0] \quad \mathbf{p_2} = [0,1] \end{aligned} \tag{16}$$

It is possible to define one loudspeaker array mode $\mathbf{v_1}$ as the combination of the two loudspeakers operating in phase and another mode $\mathbf{v_2}$, orthogonal to $\mathbf{v_1}$, as the combination of the two loudspeakers operating in opposition of phase, that is

$$\mathbf{v_1} = \left[\sqrt{1/2}, \sqrt{1/2}\right] \quad \mathbf{v_2} = \left[\sqrt{1/2}, -\sqrt{1/2}\right] \tag{17}$$

It is easy to see the orthonormality relation $\langle \mathbf{v_1} | \mathbf{v_1} \rangle = \langle \mathbf{v_2} | \mathbf{v_2} \rangle = 1$ and $\langle \mathbf{v_1} | \mathbf{v_2} \rangle = \langle \mathbf{v_2} | \mathbf{v_1} \rangle = 0$. In a similar way one can define two microphone array modes as the microphone coefficients being in quadrature, that is

$$\mathbf{u_1} = \left[\sqrt{1/2}, j\sqrt{1/2}\right] \quad \mathbf{u_2} = \left[j\sqrt{1/2}, \sqrt{1/2}\right] \tag{18}$$

Again, the orthonormality relation (11) is verified.

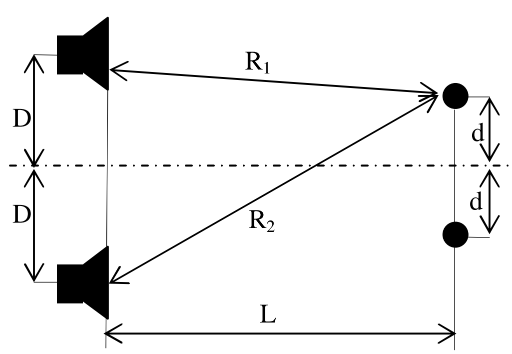

Figure 1: diagrammatic representation of the two-channels system

2.2. Singular Value Decomposition

It is possible to define the matrix $\mathbf{H^H H}$, that is the product of the adjoint matrix of the propagation matrix times the propagation matrix. This matrix is self-adjoint, meaning that it is equal to its Hermitian transpose as $\left(\mathbf{H^H H}\right)^{\mathbf{H}} = \mathbf{H^H H}$. For that reason, $\mathbf{H^H H}$ is a normal matrix and the spectral theorem can be applied to it [11]. This implies that L orthonormal eigenvectors $\mathbf{v_n}$ and L (not necessarily distinct) real eigenvalues $\sigma_n^{\ 2} \geq 0$ exist such that

$$\mathbf{H^H H v_n} = \sigma_n^{\ 2} \mathbf{v_n} \tag{19}$$

This means that the spectral decomposition of the self-adjoint operator $\mathbf{H^H H}$ identifies L loudspeaker array modes $\mathbf{v_n}$, as defined in (11), which represent a complete orthonormal basis for $C^{\mathbf{L}}$. It is important to emphasize that, even if the operator $\mathbf{H^H H}$ transform a loudspeaker vector into another loudspeaker vector, the loudspeaker array modes $\mathbf{v_n}$ and the eigenvalues $\sigma_n^{\ 2}$ depend not only on the loudspeaker arrangement and characteristic but also on the microphone array. It is now possible to apply the propagation matrix to each array mode corresponding to non-zero $\sigma_n^{\ 2}$ obtaining

$$\mathbf{H v_n} = \tilde{\mathbf{u}}_\mathbf{n} = \sigma_n \mathbf{u_n} \tag{20}$$

The expression of the vector $\tilde{\mathbf{u}}_\mathbf{n}$ as a product of a vector $\mathbf{u_n}$ times a scalar σ_n is due to the fact that the vectors $\mathbf{u_n}$ are mutually orthogonal and of unitary norm. This can be proved considering the propriety of the adjoint operator [11]

$$\langle \mathbf{H v_n} | \mathbf{v_m} \rangle = \langle \mathbf{v_n} | \mathbf{H^H v_m} \rangle \tag{21}$$

From this and considering (11) and (20) it follows that

$$\langle \mathbf{u_n} | \mathbf{u_m} \rangle = \frac{1}{\sigma_n \sigma_m} \langle \mathbf{Hv_n} | \mathbf{Hv_m} \rangle =$$
$$= \frac{1}{\sigma_n \sigma_m} \langle \mathbf{v_n} | \mathbf{H^H H v_m} \rangle = \quad (22)$$
$$\frac{\sigma_m^{\ 2}}{\sigma_n \sigma_m} \langle \mathbf{v_n} | \mathbf{v_m} \rangle = \delta_{nm}$$

This means that the loudspeaker array modes $\{\mathbf{v_n}\}$ can generate a set of orthogonal microphone array modes $\{\mathbf{u_n}\}$. In general $\{\mathbf{u_n}\}$ is not a complete basis for $C^{\mathbf{Q}}$, as the complex subspace of $C^{\mathbf{Q}}$ spanned by the microphones array modes generated by (20), $span\{\mathbf{u_n}\}$, has dimension $R \le Q$, corresponding to the rank of the propagation matrix. However, it is possible to create other $Q-R$ microphone array modes $\left[\mathbf{u_{R+1}},....,\mathbf{u_Q}\right]$ mutually orthogonal and orthogonal to $span\{\mathbf{u_n}\}$ such that the set $\{\mathbf{u_n}\} \oplus \left[\mathbf{u_{R+1}},....,\mathbf{u_Q}\right]$ is a complete base of $C^{\mathbf{Q}}$. Because of the described results, each microphone vector and loudspeaker vector can be represented as a combination of microphone array modes and loudspeaker array modes as in (14). The values σ_n are called singular values of $\mathbf{H}$ and the vectors $\mathbf{u_n}$ and $\mathbf{v_n}$ are respectively the left and right singular vectors of $\mathbf{H}$. Letting $T = \min(Q,L)$. It is possible to create a $Q \times L$ matrix $\mathbf{\Sigma}$ consisting of a diagonal matrix of dimension $T \times T$,whose elements are the singular values σ_n ordered in decreasing order, and $|Q-L|$ extra columns or rows, which contain only zeros. It is then easy to show that

$$\mathbf{HV} = \mathbf{U\Sigma} \quad (23)$$

and, as a consequence of (15)

$$\mathbf{H} = \mathbf{U\Sigma V^H} \quad (24)$$

This important result represents the Singular Value Decomposition of the propagation matrix. Furthermore, it can be shown that the pseudo-inverse matrix $\mathbf{H^+}$ can be simply expressed as

$$\mathbf{H^+} = \mathbf{V\Sigma^+U^H} \quad (25)$$

where the matrix $\mathbf{\Sigma^+}$ is computed by transposing $\mathbf{\Sigma}$ and substituting each non zero element with its reciprocal.

2.3. Meaning of the Singular Value Decomposition

It is interesting at this point to understand the meaning of the singular value decomposition of the propagation operator in terms of the array modes, and to show how it is related to the study of the ill-conditioning of the problem. It is useful to split both $C^{\mathbf{L}}$ and $C^{\mathbf{Q}}$ in two subspaces. The subspace spanned by the loudspeaker array modes $\mathbf{v_n}$ corresponding to non-zero singular values σ_n contains all the loudspeaker modes that are observable by the microphone array. On the other hand, the subspace spanned by the loudspeaker modes corresponding to zero singular values is the *null space* of the propagation matrix, $\mathcal{N}(\mathbf{H})$, and represents the loudspeaker array modes that are not observable by the microphone array. The subspace of $C^{\mathbf{Q}}$ spanned by the microphone array modes generated by (20) is the *range* of the propagation matrix, $\mathcal{R}(\mathbf{H})$, and represents the microphone array modes that can be reconstructed by the loudspeaker array. Its orthogonal complement, spanned by $\left[\mathbf{u_{R+1}},....,\mathbf{u_Q}\right]$, represents the microphone array modes that can not be reconstructed by the loudspeaker array.

It is possible to analyse separately two cases: $R = Q < L$ and $R = L < Q$. In the first case, corresponding to a number of loudspeakers larger then the number of microphones, matrix $\mathbf{\Sigma}$ has the form

$$\mathbf{\Sigma_{Q\times L}} = \begin{vmatrix} \sigma_1 & & \mathbf{0} & \mathbf{0} & \cdots & 0 \\ & \ddots & & \vdots & \ddots & \vdots \\ \mathbf{0} & & \sigma_{\mathbf{Q}} & \mathbf{0} & \cdots & 0 \end{vmatrix} \quad (26)$$

It can be seen that $\mathcal{R}(\mathbf{H}) = C^{\mathbf{Q}}$ and all microphone array modes can be reconstructed. $\mathcal{N}(\mathbf{H})$ has dimension $L-Q$. This means that there are some loudspeaker array modes that do not actively contribute to the minimization problem (8), and as a consequence even if a large amount of energy is generated by the loudspeaker array operating in one mode that belongs to the null space of the propagation matrix, the microphone array will not detect any energy. It is easy to verify that, in this case, the solution (9) implies that the loudspeaker vector $\mathbf{a_0}$ is such that it does not activate any of these modes and hence minimise the amount of energy required to reconstruct the desired microphone array modes. From a geometrical point of view, the least squares method (9) always defines a solution $\mathbf{a_0}$ that is perpendicular to $\mathscr{N}(\mathbf{H})$ and therefore lies on the orthogonal complement of $\mathscr{N}(\mathbf{H})$ to $C^{\mathbf{L}}$. Any other solution $\mathbf{a}':\mathbf{p}=\mathbf{Ha}'$ having a projection on $\mathscr{N}(\mathbf{H})$ would have a squared norm $\|\mathbf{a}'\|^2 > \|\mathbf{a_0}\|^2$

In the case that $R=L<Q$, when there are more microphones than loudspeakers, matrix $\mathbf{\Sigma}$ has the form

$$\mathbf{\Sigma_{Q\times L}} = \begin{vmatrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_{\mathbf{L}} \\ \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} \end{vmatrix} \tag{27}$$

It can be observed that $\mathscr{N}(\mathbf{H})=\{0\}$ and that there are some microphone array modes that can not be reconstructed by the loudspeaker array. This problem is ill posed, as the exact solution can not always be found. In this case the least squares solution (9) generates an orthogonal projection $\hat{\mathbf{p}}$ of $\mathbf{p}$ on $\mathscr{R}(\mathbf{H})$, and then finds an loudspeaker vector $\mathbf{a_0}$ that reconstructs that projection. The root mean square error $\|\mathbf{p}-\hat{\mathbf{p}}\|$ is the norm of the component of $\mathbf{p}$ perpendicular to $\mathscr{R}(\mathbf{H})$. In other words, the least squares solution automatically finds a solution that reconstructs the array modes that can be reconstructed by the loudspeaker array and neglects the other modes.

It is really important to point out that in both the cases described the singular value decomposition defines a one to one relation between each microphone modes $\mathbf{u_n}\in\mathscr{R}(\mathbf{H})$ and the loudspeaker array mode $\mathbf{v_n}\notin\mathscr{N}(\mathbf{H})$ from which it is generated (if the multiplicity of the corresponding singular value is unitary). More intuitively, the reconstruction of the microphone array mode $\mathbf{u_n}\in\mathscr{R}(\mathbf{H})$ depends on one and only one loudspeaker array mode. Furthermore, the corresponding modal amplification is given by the corresponding singular value, which is to say that if the loudspeaker array is operating in the mode $\mathbf{u_n}$ and $\|\mathbf{a}\|=1$, then it can be easily shown that the microphone array is detecting only the mode $\mathbf{v_n}$ and $\|\mathbf{p}\|=\sigma_n$. This important relation derives directly from (25) and implies that some loudspeaker array modes are seen to be more efficient than others by the microphone array, and the measure of the efficiency is expressed by the corresponding singular values. It is important to make clear that the loudspeaker array modes are considered to be efficient or inefficient in respect to what can be detected by the microphone array, but nothing is argued about the effect of each loudspeaker array mode on the reconstructed sound field $\hat{p}(\mathbf{x})$. Considering the inverse problem (9), it can be argued that some microphone array modes require more effort to be reconstructed than others. In more detail, if a microphone array mode is related to an inefficient loudspeaker array mode, then its reconstruction requires a large amount of energy, proportional to $1/\sigma_n{}^2$. In the limit that σ_n is very close to 0, then an almost infinite amount of energy is required by the loudspeaker array. The ratio between the largest and the smallest singular value of $\mathbf{H}$ is the condition number of the propagation matrix and is a good index of the conditioning of the matrix. In a more intuitive sense, if some loudspeaker array modes are severely less efficient than others, then the problem is ill-posed and the presence of a little noise in the microphone signals or little errors in the estimation of the propagation matrix can have devastating effects when attempting the reconstruction due to the “explosion” of the modes with low efficiency.

In a geometric sense, the singular value decomposition means that a hyper-sphere of unitary radius in $C^{\mathbf{L}}$ is mapped to a hyper-ellipse in $C^{\mathbf{Q}}$, having its principal axis of length σ_n directed as the eigenvectors $\mathbf{u_n}$. The more the hyper-ellipse looks like a sphere, the more robust is the system.

In order to regularise an ill-posed problem, one could deliberately decide not to attempt the reconstruction of the microphone array modes which relate to non observable loudspeaker array modes. This technique is known as truncated singular value decomposition [9,10] and implies that the smaller singular values are converted into zero and hence not inverted in the computation of $\mathbf{\Sigma}^+$. Alternatively, it is possible to limit the total energy by restricting the allowed value of $\mathbf{a}$ to the interior of a hyper-sphere in $C^{\mathbf{L}}$ of certain radius [10], with the direct implication that the range of the propagator is restricted to the interior of a hyper-ellipse defined in $C^{\mathbf{Q}}$. A further possibility is represented by the Tikhonov regularization. This method consists in adding a small quantity β to all singular values σ_n. It can be shown [9,10] that the Tikhonov regularization method implies the trade off between the correct reconstruction of the sound field and a limit on the total energy of the signals driving the loudspeakers.

3. EXAMPLES

3.1. Two channel system

In order to give a better understanding of the presented results, it is useful to study the reconstruction problem in the simple case represented by Figure 1. In this example, the electro acoustic transfer functions between each loudspeaker and each microphone constituting the propagation matrix are assumed to be free field Green functions of the form

$$H_{ql} = \frac{\exp\left(-jk\left\|\mathbf{x_q} - \mathbf{y_l}\right\|\right)}{4\pi\left\|\mathbf{x_q} - \mathbf{y_l}\right\|} \tag{28}$$

Observing Figure 1, it can be easily shown that the propagation matrix is

$$\mathbf{H} = \begin{vmatrix} \dfrac{\exp(-jkR_1)}{4\pi R_1} & \dfrac{\exp(-jkR_2)}{4\pi R_2} \\ \dfrac{\exp(-jkR_2)}{4\pi R_2} & \dfrac{\exp(-jkR_1)}{4\pi R_1} \end{vmatrix} \tag{29}$$

The Singular Value Decomposition of $\mathbf{H}$ defines two loudspeaker array modes $\mathbf{v_1}$ and $\mathbf{v_2}$ of the form expressed by (17), that is in one mode the loudspeakers are operating in-phase and in the other mode in opposition of phase. Numerical simulations show that the microphone array modes are such that the microphone coefficients defining $\mathbf{u_1}$ are in phase, while those defining $\mathbf{u_2}$ are in opposition of phase.

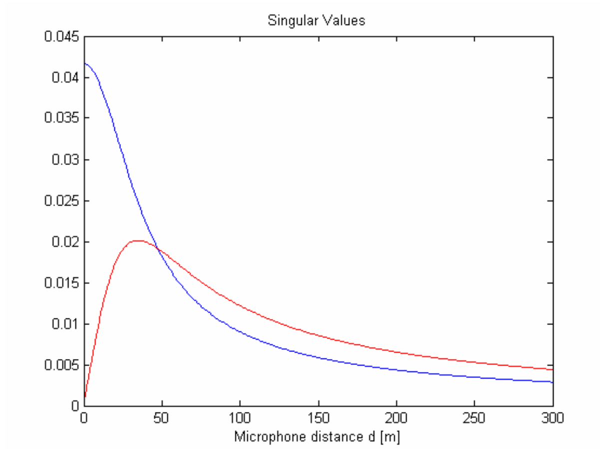

Figure 2 : singular values of the two-channel system

Figure 2 shows how the singular values associated with $\mathbf{v_1}$ (blue line) and $\mathbf{v_2}$ (red line) vary as a function of the distance between the two microphones and for a fixed arrangement of the loudspeakers. It can be observed that when the two microphones are closely spaced, the first loudspeaker array mode $\mathbf{v_1}$ can be well observed by the microphone array, while the second mode is not. This implies that the reconstruction of the microphone array mode $\mathbf{u_2}$ is more "problematic" than the reconstruction of $\mathbf{u_1}$ when the microphones are close together. In these circumstances, matrix $\mathbf{H}$ is ill-conditioned. In the extreme case when the two microphones are coincident, the singular value associated with $\mathbf{v_2}$ approaches 0, and the propagation matrix becomes singular. This means that the loudspeaker array mode $\mathbf{v_2} \in \mathcal{N}(\mathbf{H})$, and therefore can not be observed by the microphone array. This has the intuitive consequence that it is impossible to create two microphone coefficients in opposition of phase when the two microphones are in the same position. When the distance between the microphones is increased, the singular value associated with $\mathbf{v_2}$ grows while the other singular value decreases. At a certain microphone distance the two singular values coincide, and if the distance between the microphones is increased, the array mode $\mathbf{v_1}$ becomes less observable than $\mathbf{v_2}$ and, as a consequence, the reconstruction of the microphone array mode $\mathbf{u_1}$ requires more energy than the reconstruction of $\mathbf{u_2}$.

3.2. Multi-channel system

The reconstruction of a sound field can be in general attempted using a large number of loudspeakers and microphones. In the following example, the numerical simulation of the reconstruction of a sound field using a system constituted by 81 loudspeakers and 81 omnidirectional microphones is illustrated. The loudspeakers and the microphones are arranged almost regularly (sphere packing [12]) over the surface of two concentric spheres of radius R_L and R_M respectively, and $R_L > R_M$. As in the previous example, the elements of the propagation matrix $\mathbf{H}$ are free field Green functions. The singular value decomposition is applied to the propagation matrix $\mathbf{H}$ in order to analyse the stability of the system.

If the sound field $p(\mathbf{x})$ is square-integrable in Ω, then it possible to express $p(\mathbf{x})$ by the generalized Fourier series [13]

$$p(\mathbf{x}) = \sum_{n=-\infty}^{+\infty} b_n \psi_n(\mathbf{x}) \tag{30}$$

Where the set of $\psi_n(\mathbf{x})$ represents a complete set of orthonormal functions for Ω, that is [11]

$$\langle \psi_n | \psi_m \rangle = \delta_{nm} \tag{31}$$

and the series coefficients are computed by

$$b_n = \langle \psi_n \,|\, p \rangle = \int_\Omega \overline{\psi_n(\mathbf{x})} p(\mathbf{x}) d\Omega(\mathbf{x}) \tag{32}$$

A useful example of generalized Fourier series is represented by the spherical harmonics expansion, sometimes also called the Fourier-Bessel series. It can be proved [14] that the sound field in an interior, source free region Ω can be represented using the following expression

$$p(\mathbf{x}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} j_n(k\|\mathbf{x}\|) Y_n^m(\hat{\mathbf{x}}) b_{nm} \tag{33}$$

The terms b_{nm} are the coefficients of the series, $j_n(\cdot)$ is the spherical Bessel function of order n, $\hat{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|$ and $Y_n^m(\cdot)$ is a spherical harmonic as defined in [14].

It has been shown [15] that the singular value decomposition of the propagation matrix can be computed analytically in the case of spherical loudspeaker and microphone arrays, like the ones described, and when the number of transducers tends to infinite. This result holds with good approximation when the arrays are constituted by a large but finite number of transducers. In more detail, both the microphone array modes and the loudspeaker array modes correspond to sampled spherical harmonics, while the singular values are such that

$$\sigma_n = k \left| j_n(kR_M) h_n^{(2)}(kR_L) \right| \tag{34}$$

where $h_n^{(2)}(\cdot)$ is the Hankel function of the second kind of order n [14]. This important result gives a very good insight into the analysis of the conditioning of the propagation matrix, as it explicitly expresses relates the radius of the transducer arrays to the singular values and consequently the condition number of the propagation matrix.

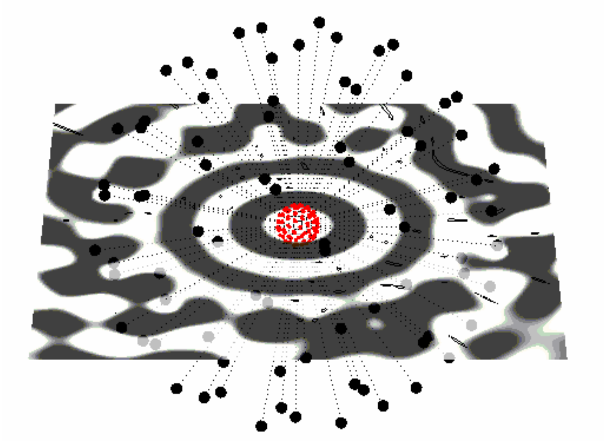

Figure 3: multi-channel system - sound field generated by the first loudspeaker array mode

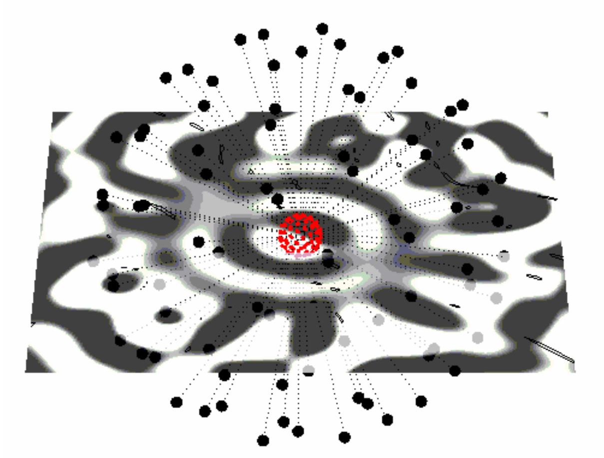

Figure 4: multi-channel system - sound field generated by the first loudspeaker array mode

Figure 3 and Figure 4 show the horizontal cross-section of the sound field generated by the first two loudspeaker array modes $\mathbf{v_1}$ and $\mathbf{v_2}$ obtained by a numerical computation of the singular value decomposition of the propagation matrix. The black dots represent the loudspeakers, while the red dots represent the microphones. Considering equation (34), it can be observed that in the interior of the loudspeaker array the sound fields in Figure 3 and Figure 4 can be well represented by the product of one spherical harmonic and one Bessel function ($n = 0$ in Figure 3 and $n = 1$ in Figure 4).

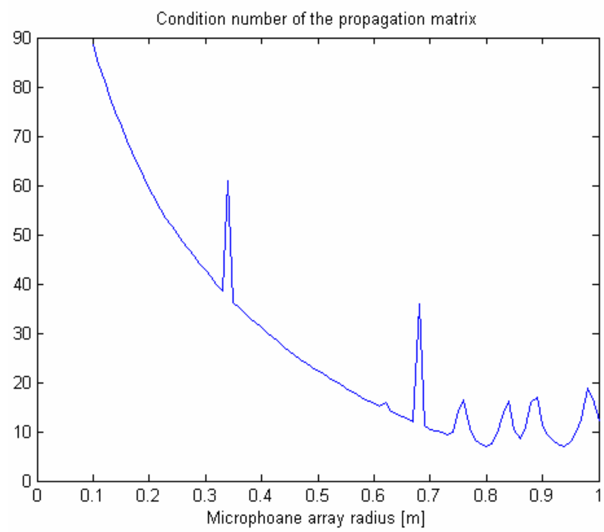

Figure 5: multi-channel system - condition number of the propagation matrix

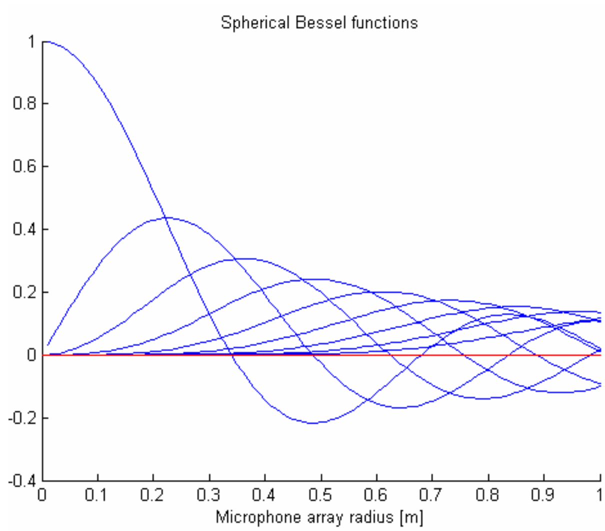

Figure 6: spherical Bessel Functions

Figure 5 shows how the condition number of the propagation matrix varies as a function of the radius of the microphone array and for fixed wave number $k = 9.24$ (corresponding to $500\ Hz$) and loudspeaker array radius $R_L = 2\ m$. Figure 6 plots the spherical Bessel functions of the first 8 orders. It is easy to observe the relation between the condition number and the behavior of the Bessel functions: the condition number is very large close to the origin, as all the Bessel functions of order $n > 0$ are very small close to the origin while the Bessel function of order $n = 0$ is unitary at the origin. This implies that only the first loudspeaker array mode $\mathbf{v_1}$ is observable if all the microphones are located at the origin, and as a consequence only the first microphone array vector $\mathbf{u_1}$, corresponding to all microphone coefficients being in-phase, can be reconstructed. A similar result was obtained in the previous 2-channel system example, and shows that the propagation matrix becomes ill-conditioned when the microphones are closely spaced. This result holds if the microphones are considered to be omnidirectional, but this might not be the case for microphones with different directivity, as shown by [3].

Figure 5 also shows the presence of some peaks in the condition number, the first of which occurs at $R_M = 0.34\ m$ and the second at $R_M = 0.68\ m$. Observing Figure 6, it can be noticed that, for a given wave number k, the spherical Bessel functions of order $n = 0$ is zero at $R_M = 0.34\ m$ and at $R_M = 0.68\ m$. These distances correspond to the radius of the nodal circles represented in Figure 3. As a consequence, if the microphone array radius is one of those two values, that is all microphones are located on a nodal surface of the sound field generated by the first loudspeaker array mode $\mathbf{v_1}$, the singular value associated to $\mathbf{u_1}$ is zero and the associated loudspeaker array mode belongs to the null space of the propagation matrix.

4. CONCLUSIONS

The theory of the sound field reconstruction based on the minimization of the least squared error between the original and reconstructed sound field has been presented and discussed. The concept of array modes has been introduced and the singular value decomposition has been first defined in a mathematical sense and its application to a multi-channel system has been described. The practical implication derived from ill-conditioning of the propagation matrix have been discussed in the detail and clarified with two examples. In the case of the presented multi-channel system, the condition number has been shown to be related to the microphone array radius. It has been demonstrated that ill-conditioning of the propagation matrix can occur if the microphones are closely spaced or if they all lie in the nodal points of the sound field generated by one loudspeaker array mode.

5. REFERENCES

[1] A.J. Berkhout, D. De Vries and P. Vogel (1993) “Acoustic control by wave field synthesis.” Journal of the Acoustical Society of America 93, 2764-2778.

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

[3] J. Daniel (2000) “Acoustic field representation, application to the transmission and the reproduction of complex sound environments in the multimedia context.” PhD. Paris: Université Paris 6.

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

[5] A. Farina et al. (2001). “Ambiophonic Principles for the Recording and Reproduction of Surround Sound for Music”, AES 19th International conference, Schloss Elmau, Germany.

[6] M. Poletti (2005) “Three-Dimensional Surround System Based on Spherical Harmonics”. J. Audio Eng. Soc., Vol. 53, No. 11

[7] P.A. Nelson and S.J. Elliott (1992) “Active Control of Sound”. Academic, London.

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

[9] D. Colton and R. Kress (1992) “Inverse Acoustic and Electromagnetic Scattering Theory” Springer-Verlag

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

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

[12] N. J. A. Sloane, R. H. Hardin and W. D. Smith *et al.*, " Tables of spherical codes", [Online document] Available at http://www.research.att.com/~njas/packings/

[13] L. Nowinski (1981) "Applications of Functional Analysis in Engineering", Hardcover

[14] 43. E.G. Williams (1999) "Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography". New York: Academic.

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