19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

A THEORETICAL STUDY OF SOUND FIELD RECONSTRUCTION TECHNIQUES

PACS: 43.60.Tj

Fazi, Filippo[1]; Nelson, Philip[1]
[1]Institute of Sound & Vibration Research, Southampton University SO171BJ, United Kingdom;
ff1@isvr.soton.ac.uk ; pan@isvr.soton.ac.uk

ABSTRACT
Three different methods for the physical analysis and synthesis of sound fields are presented and compared: the least squares method, a method based on the Kirchhoff-Helmholtz integral equation and the generalised Fourier transform method (based on spherical harmonics). These three methods constitute the basis for many active control technologies and for some modern multi-channel audio technologies, such as High Order Ambisonics and Wave Field Synthesis. It is analytically demonstrated that the three methods have an equivalent formal background. It is shown that the three techniques give identical reconstruction performance in the target frequency range when using the same number of transducers, regularly arranged over a sphere. The analytical results are confirmed by using numerical simulations.

INTRODUCTION

The reconstruction of a sound field is a problem studied within different branches of acoustics. Particular attention to that problem has been given to the design of multi-channel systems for audio applications. In the first three sections, different methods are presented: the least squares method [1] that is largely used in active control applications, then a method based on the Kirchhoff-Helmholtz integral [2], from which the theory of Wave Field Synthesis is derived, and finally a method based on the generalised Fourier transform [3], which constitutes the basis for the theory of High Order Ambisonics. The theoretical equivalence of these three methods, is demonstrated in the ideal case of a spherical array of an infinite number of transducers used to reconstruct the sound field in the interior region. Under these circumstances, the Kirchhoff-Helmoholtz integral reduces to the simple source formulation that has already been shown [4] to give the same results of the method based on the spherical harmonic decomposition. The simple source formulation and the method based on spherical harmonics are then shown to be a particular case of the least squares method. The proof was inspired by the work of Nelson and Kahana [5] and is based on the singular value decomposition of the propagation matrix. Finally, numerical simulations show that the performance of the three methods is very similar in the interior domain (but not in the exterior domain) even if the number of transducers is finite.

THE LEAST SQUARES METHOD

The acoustic pressure of a generic sound field in a non dispersive medium, defined over a limited region of space Ω where no acoustic sources are located, can be described by a scalar field $p(\mathbf{r},\omega)$ that satisfies the homogeneous Helmholtz equation

$$\nabla^2 p(\mathbf{r},\omega)+k^2 p(\mathbf{r},\omega)=0 \ . \qquad (1)$$

The vector $\mathbf{r}$ identifies a point in the region Ω, and ω is the frequency and k the wave number. The harmonic time dependence $e^{j\omega t}$ is implicit, where $j=\sqrt{-1}$. The target of the sound field reconstruction is to create a sound field $\hat{p}(\mathbf{r},\omega)$ that is as similar as possible to the original sound field $p(\mathbf{r},\omega)$. One way of attempting the reconstruction is to measure the original sound field using an array of *S* arbitrarily arranged microphones, each of them generating a signal $p_s(\omega)$, and then to process these signals in order to obtain the signals which drive an array of *L* arbitrarily arranged loudspeakers, each of them driven by a signal $a_l(\omega)$. Therefore, the *L* loudspeakers generate the reconstructed sound field $\hat{p}(\mathbf{r},\omega)$.

It is possible to measure the reconstructed sound field using the same microphone array that was used to measure the original sound field. In that case, each of the *S* microphones generates a signal $\hat{p}_s(\omega)$. The *S* microphone signals $p_s(\omega)$ and $\hat{p}_s(\omega)$ as well as the *L* loudspeakers signals $a_l(\omega)$ can be represented by vectors $\mathbf{p}(\omega)$, $\hat{\mathbf{p}}(\omega)$ and $\mathbf{a}(\omega)$ respectively. For example:

$$\mathbf{p}(\omega) = \left[\, p_1(\omega), p_2(\omega), p_3(\omega), \ldots\ldots, p_S(\omega) \right]^T . \tag{2}$$

It is possible to express the relation between the vectors $\hat{\mathbf{p}}(\omega)$ and $\mathbf{a}(\omega)$ using the matrix product

$$\hat{\mathbf{p}}(\omega) = \mathbf{H}(\omega)\mathbf{a}(\omega). \tag{3}$$

$\mathbf{H}(\omega)$ is a *SxL* frequency dependent matrix, called the propagation matrix, and its (*s,l*)th element represents the linear electro-acoustic transfer function between the *l*-th loudspeaker and the *s*-th microphone. This transfer function depends on the characteristics and arrangement of the transducers and can include, in general cases, the effect of the reflections that occur when the reconstruction is attempted in a reverberant environment. In the ideal case of free field conditions, omnidirectional ideal microphones and ideal monopole-like loudspeakers the elements of matrix $\mathbf{H}(\omega)$ can be reduced to free field Green functions $G\left(\mathbf{r_l} \,|\, \mathbf{r_s}, \omega\right)$ [2].

It is possible to choose a set of loudspeakers signals $\mathbf{a}(\omega)$ that minimize, in a least squares sense, the difference between vectors $\hat{\mathbf{p}}(\omega)$ and $\mathbf{p}(\omega)$. The solution to the presented problem is well known [1]:

$$\mathbf{a}(\omega) = \mathbf{H}^{+}(\omega)\mathbf{p}(\omega). \tag{4}$$

$\mathbf{H}^{+}(\omega)$ represents the Moore-Penrose pseudo-inverse of matrix $\mathbf{H}(\omega)$. The solution of this inverse problem is strongly sensitive to the conditioning of $\mathbf{H}(\omega)$, expressed by the spread of its singular values [1]. This problem can be largely simplified if the sound field reconstruction is attempted in an anechoic environment. If this is the case, the conditioning of $\mathbf{H}(\omega)$ depends only on the transducer characteristics and arrangement.

Matrix $\mathbf{H}(\omega)$ can be decomposed by using the singular value decomposition [5]

$$\mathbf{H}(\omega) = \mathbf{U}(\omega)\mathbf{\Sigma}(\omega)\mathbf{V}^{\mathbf{H}}(\omega) \tag{5}$$

where the superscript $\mathbf{H}$ represents the Hermitian transposed matrix and $\mathbf{U}(\omega)$ and $\mathbf{V}(\omega)$ are unitary matrices, representing an algebraic basis change. Their columns are the singular vectors of $\mathbf{H}(\omega)$. The matrix $\mathbf{\Sigma}(\omega)$ is diagonal and real valued matrix of the singular values of $\mathbf{H}(\omega)$. In the special case of $\mathbf{H}(\omega)$ being a normal matrix, the singular values correspond to the absolute values of the eigenvalues of $\mathbf{H}(\omega)$. This matrix decomposition is useful because the pseudoinverse matrix $\mathbf{H}^{+}(\omega)$ can be expressed as

$$\mathbf{H}^{+}(\omega) = \mathbf{V}(\omega)\mathbf{\Sigma}^{-1}(\omega)\mathbf{U}^{\mathbf{H}}(\omega). \tag{6}$$

It is also possible to regularise the singular values $\mathbf{\Sigma}(\omega)$ in order to obtain a more stable solution to the inverse problem.

THE KIRCHHOFF-HELMHOLTZ INTEGRAL

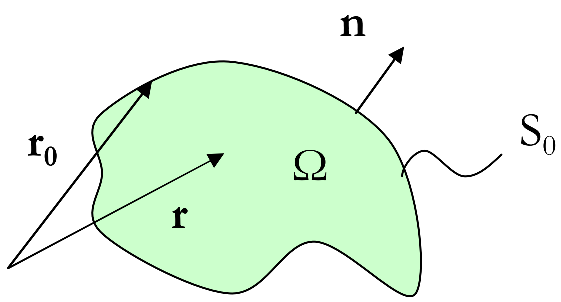

Figure 1.- Parameter used for the Kirchhoff-Helmholtz integral

A well known solution of the Helmholtz equation (1) is given by the Kirchhoff-Helmholtz integral [2]. Referring to Figure 1 the Kirchhoff-Helmholtz integral can be written as

$$\alpha p(\mathbf{r},\omega)=\int_{S_0} p(\mathbf{r_0},\omega)\frac{\partial G(\mathbf{r_0}\,|\,\mathbf{r},\omega)}{\partial \mathbf{n}}-\frac{\partial p(\mathbf{r_0},\omega)}{\partial \mathbf{n}}G(\mathbf{r_0}\,|\,\mathbf{r},\omega)\,dS_0$$
$$\alpha=\begin{cases}1 \text{ if } \mathbf{r}\in\Omega\\ 1/2 \text{ if } \mathbf{r}\in S_0\\ 0 \text{ if } \mathbf{r}\notin\Omega\end{cases} \tag{7}$$

where $G(\mathbf{r_0}\,|\,\mathbf{r},\omega)$ is the Green function, which is chosen depending on the boundary conditions of the studied acoustic problem. In the special case of free field conditions, $G(\mathbf{r_0}\,|\,\mathbf{r},\omega)$ is the free field Green function and the two terms in the integrand of equation (7) can be regarded as representing surface distributions of dipole and monopole sources respectively. This equation has an important direct implication: once the acoustic pressure $p(\mathbf{r_0},\omega)$ of the original sound field and its derivative in respect to the $\mathbf{n}$ are known on the surface S_0, then the sound field can be reconstructed in free field conditions using a continuous distribution of monopole- and dipole-like secondary sources (loudspeakers) on the surface S_0. This approach also implies that the reconstructed sound field is zero at all locations outside the reconstruction area Ω.

THE GENERALISED FOURIER TRANSFORM

The scalar field describing the sound field inside Ω can be analysed by using the generalised Fourier transform. In mathematical terms, this means that the scalar field is described by means of a (generally infinite) set of orthogonal functions, which represent a complete basis for that space. There are many possible sets of basis functions and one of the possible and convenient choices is represented by the spherical harmonic series. As shown by Williams [2], the sound field defined over a limited, source free, region of space Ω can be expressed as

$$p(\mathbf{r},\omega)=\sum_{n=0}^{\infty}\sum_{m=-n}^{n} b_{nm}\, j_n(kr)Y_n^m(\hat{\mathbf{r}})$$
$$b_{nm}\, j_n(kr)=\int_{S_0} p(\mathbf{r_0},\omega)Y_n^m(\hat{\mathbf{r}}_\mathbf{0})^*\,dS_0 \tag{8}$$

The symbol $*$ represents the complex conjugate, $r=\|\mathbf{r}\|$, $\hat{\mathbf{r}}$ is the unit vector pointing in the same direction as $\mathbf{r}$, defined as $\hat{\mathbf{r}}=\mathbf{r}/r$, $j_n(kr)$ is n-th order spherical Bessel function of the first kind and $Y_n^m(\hat{\mathbf{r}})$ is the spherical harmonic of order n and mode m. The term S_0 represents the surface of a sphere of radius r that is centred on the origin and contained in Ω, and b_{nm} are the coefficients of the spherical spectrum.

When attempting to reconstruct a sound field using an array of L loudspeakers, it is possible to analyse the original sound field $p(\mathbf{r},\omega)$ and the sound field generated by each loudspeaker using a generalised Fourier transform, and then define the loudspeaker signals a_l using a mode matching method. If the spherical harmonics are chosen as basis functions and if the loudspeakers are assumed to be monopole-like sources arranged on the surface of a sphere containing Ω, the loudspeaker signals $a_l(\omega)$ can be computed from [3, 4]

$$a_l(\omega) = \frac{4\pi}{L}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{b_{nm}}{(-jk)h_n^{(2)}(kr_l)}Y_n^m(\hat{\mathbf{r}}_\mathbf{l}), \tag{9}$$

where $h_n^{(2)}(kr_0)$ is a spherical Hankel function of the second kind [2]. The factor $4\pi/L$ arises from the Gaussian quadrature of the orthogonality relation for the spherical harmonics [2,3]. The coefficients b_{nm} can be computed, for example, if the acoustic pressure of the original sound field is known on the surface of a sphere S_s with radius $r_s \le r_l$ and centred on the origin, under the condition that the terms $j_n(kr_m) \neq 0$:

$$b_{nm} = \frac{1}{j_n(kr_m)r_s^2}\int_{S_s} p(\mathbf{r_s},\omega)Y_n^m(\hat{\mathbf{r}}_\mathbf{s})^* dS_s . \tag{10}$$

COMPARISON BETWEEN THE DIFFERENT METHODS

When the reconstruction of the sound field is limited to the source free region Ω, that is, when the interior problem is considered, the knowledge of both the acoustic pressure $p(\mathbf{r_0},\omega)$ and its derivative, as required by equation (7), is redundant, and the exact sound field reconstruction can be achieved using a continuous distribution of either monopole-like or dipole-like secondary sources [2]. Under the assumption of free field conditions, if monopole-like secondary sources are used, the sound field inside Ω can be described using the simple source formulation [2]

$$p(\mathbf{r},\omega) = \int_{S_l}\mu(\mathbf{r_l},\omega)G(\mathbf{r_l}\,|\,\mathbf{r},\omega)\,dS_l . \tag{11}$$

The function $\mu(\mathbf{r_l},\omega)$ represents the "driving signal" of the monopole-like source located at $\mathbf{r_l}$, and can be analytically expressed as described by Williams [2]. This result implies that, under the above-mentioned conditions, it is possible to achieve an exact reconstruction of the sound field inside Ω using a continuous distribution of one type only of secondary source. Poletti [4] has proved that, in the special case of S_0 being the surface of a sphere with radius r_l and centred on the origin, the computation of the function $\mu(\mathbf{r_l},\omega)$ is equivalent to (9) and $a_l(\omega) \xrightarrow{L=\infty} \mu(\mathbf{r_l},\omega)r_l^2 4\pi/L = \mu(\mathbf{r_l},\omega)dS_l$. However, it is important to point out that the use of only one type of secondary source implies that the sound field outside Ω is not necessarily zero. For this reason, the simple source formulation and the Kirchhoff-Helmholtz equation are equivalent only for the interior domain Ω, but not outside that region.

Referring now to the least squares method and to equation (3), it is possible to consider the special and ideal case of L monopole-like loudspeakers regularly arranged over the surface of a sphere with radius r_l and S ideal omnidirectional microphones regularly arranged over the surface of a sphere with radius $r_s \le r_l$. The elements of matrix $\mathbf{H}(\omega)$ can be expressed as [2]

$$\mathbf{H}_{\mathbf{s},\mathbf{l}}(\omega) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}(-jk)h_n^{(2)}(kr_l)\,j_n(kr_s)Y_n^m(\hat{\mathbf{r}}_\mathbf{s})Y_n^m(\hat{\mathbf{r}}_\mathbf{l})^* . \tag{12}$$

It is also possible to decompose the matrix $\mathbf{H}(\omega)$ as the product of three matrices defined as follows

$$\mathbf{H}(\omega) = \mathbf{\Xi}\mathbf{\Lambda}(\omega)\mathbf{\Psi}^{\mathbf{H}}, \quad \mathbf{\Xi}_{\mathbf{s,nm}} = \sqrt{\frac{4\pi}{S}} Y_n^m(\hat{\mathbf{r}}_\mathbf{s}), \quad \mathbf{\Psi}_{\mathbf{l,nm}} = \sqrt{\frac{4\pi}{L}} Y_n^m(\hat{\mathbf{r}}_\mathbf{l})$$
$$\mathbf{\Lambda}(\omega) = \frac{\sqrt{SL}}{4\pi}(-jk)diag\left[h_0(kr_l)j_0(kr_s), h_1(kr_l)j_1(kr_s), h_1(kr_l)j_1(kr_s), \ldots\right] \tag{13}$$

If the number of loudspeakers L equals the number of microphones S and both tend to infinity, it can be demonstrated that matrices $\mathbf{\Xi}$ and $\mathbf{\Psi}$ are frequency independent and unitary (e.g. $\mathbf{\Xi}^{\mathbf{H}}\mathbf{\Xi} = \mathbf{I}$) due to the orthogonality of the spherical harmonics [2]:

$$\left(\mathbf{\Xi}^{\mathbf{H}}\mathbf{\Xi}\right)_{nm,\nu\mu} = \sum_{s=1}^{S} Y_n^m(\hat{\mathbf{r}}_\mathbf{s})^* Y_\nu^\mu(\hat{\mathbf{r}}_\mathbf{s}) \frac{4\pi}{S} \xrightarrow{S=\infty} \int_{\hat{S}} Y_n^m(\hat{\mathbf{r}}_\mathbf{s})^* Y_\nu^\mu(\hat{\mathbf{r}}_\mathbf{s}) d\hat{S} = \delta_{n\nu}\delta_{m\mu} \ . \tag{14}$$

$\hat{S}$ is the surface of a sphere with unit radius and $\delta_{n\nu} = 1$ if $n = \nu$ and $\delta_{n\nu} = 0$ otherwise. The diagonal complex valued matrix $\mathbf{\Lambda}(\omega)$ can be transformed into a real valued diagonal matrix if multiplied by a unitary diagonal matrix $\mathbf{\Gamma}(\omega)$ such that $\mathbf{\Lambda}_{\mathbf{nm}}(\omega)\mathbf{\Gamma}_{\mathbf{nm}}(\omega) = \left|\mathbf{\Lambda}_{\mathbf{nm}}(\omega)\right|$. Comparing this result with equation (5), it is possible to observe that equation (13) represents a possible singular value decomposition of matrix $\mathbf{H}(\omega)$: the singular values of the matrix are the absolute values of the elements of the diagonal matrix $\mathbf{\Lambda}(\omega)$, while the spherical harmonics are the singular functions, up to a phase rotation due to the degeneracy of the singular values (the singular value corresponding to the n-th order has a $2n+1$ multiplicity). The pseudo-inverse matrix $\mathbf{H}^+(\omega)$ can be easily computed from

$$\mathbf{H}^+(\omega) = \mathbf{\Psi}\mathbf{\Lambda}^{-1}(\omega)\mathbf{\Xi}^{\mathbf{H}} \ . \tag{15}$$

The important consequence of this result is that the conditioning of matrix $\mathbf{H}(\omega)$ depends only on the radial coordinates of the transducer positions defined by r_l and r_s. Substituting equation (13) and (15) into equation (4) and using equation (10), the loudspeaker signals can be analytically expressed as

$$a_l(\omega) = \frac{16\pi^2}{LS} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{\sum_{s=1}^{S} p_s Y_n^m(\hat{\mathbf{r}}_\mathbf{s})^*}{j_n(kr_s)(-jk)h_n^{(2)}(kr_l)} Y_n^m(\hat{\mathbf{r}}_\mathbf{l}) \xrightarrow{S=\infty} \frac{4\pi}{L} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{b_{nm}}{(-jk)h_n^{(2)}(kr_l)} Y_n^m(\hat{\mathbf{r}}_\mathbf{l})$$

This result corresponds to equation (9). This implies that, under the conditions previously described, the simple source formulation and the spherical harmonic decomposition can be seen as a special case of the least squares method, corresponding to the particular choice of monopole-like loudspeakers regularly arranged over the surface of a sphere. However, the LSM is more general, as no assumption is made about the characteristics and arrangement of the transducers and of the reconstruction environment. It is useful to note that the generalised Fourier transform is a powerful tool for computing the pseudo inverse of the matrix $\mathbf{H}(\omega)$.

NUMERICAL SIMULATIONS

Since the number of loudspeakers and microphones is finite in practical cases the reconstruction of the sound field is never exact over an arbitrarily extended area. When dealing with a finite number of transducers, the summation in equations (8), (9) etc. must be truncated at the *N*-th order, with the consequence of a truncation error in the reconstruction. However, numerical simulations show that the presented results hold to a first approximation when using a finite number of transducers. Figure 2 represents the simulation of the reconstruction on the horizontal plane $z = 0$ of the sound field due a 200 Hz harmonic spherical wave using 81 monopole-like loudspeakers (+ 81 dipoles for the KHI) and 81 omni-directional microphones (81 ideal intensity probes for the KHI), almost regularly arranged on the surface of two spheres with radius $r_l = 2$ m and $r_s = 0.2$ m respectively ($r_s = r_l = 2$ m for the KHI). The black dots represent the positions of the loudspeakers and the red sphere represents the sphere on which the microphones are placed. The reconstruction is attempted using the three methods described. The continuous line represents the region of space inside which the reconstruction normalised square error is less than 5%. It can be seen that the reconstruction is achieved in a similar way in the three cases. In the case of the Kirchhoff-Helmholtz integral, the sound field is almost zero outside the reconstruction area. In general the three methods give very different results in the exterior domain.

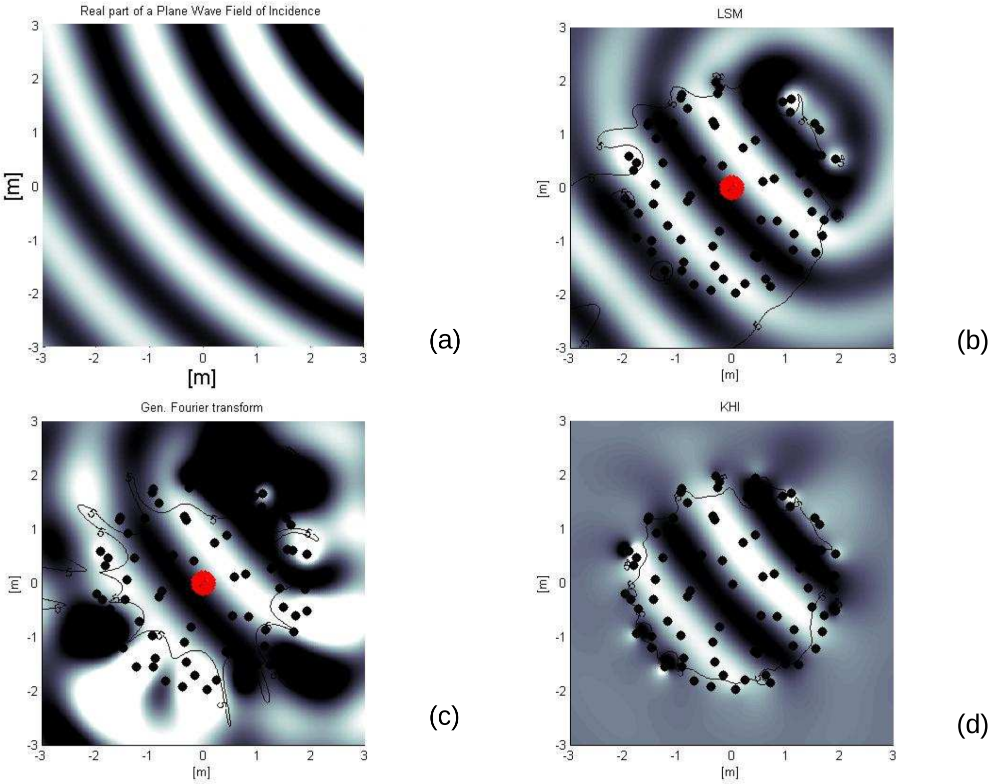

Figure 2.-Numerical simulation of sound field reconstruction. (a) original sound field, (b) least squares method, (c) generalised Fourier transform, (d) Kirchhoff-Helmholtz integral

CONCLUSIONS AND FURTHER STUDIES

The theories of three different methods for the reconstruction of the sound field have been presented and analytically compared for the special case of a spherical array of transducers. It has been shown that both the simple source formulation and the generalised Fourier method can be seen as a particular case of the least squares method. The generalised Fourier transform allows a powerful method for the computation of the pseudo-inverse of the propagation matrix $\mathbf{H}(\omega)$ and for the study of its conditioning. For the case under study and under the ideal assumption of an infinite number of transducers, the singular values of $\mathbf{H}(\omega)$ can be computed analytically and depend only on the frequency and on the radial coordinates of the positions of the transducers. The application of this approach to arrays with different geometries and different transducer directivity could be the object of further studies.

References:

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

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

[3] F.M. Fazi and P.A. Nelson: Analisi comparativa tra due tecniche di ricostruzione del campo acustico. Proceedings of the 34° Convegno Nazionale Associazione Italiana Acustica (2007).

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

[5] P.A. Nelson and Y. Kahana: Spherical harmonics, singular value decomposition and the head related transfer function. Journal of Sound and Vibration 239, 607-638 (2001).