



HAL
open science

Representation of sound fields for audio recording and reproduction

Filippo Fazi, Markus Noisternig, Olivier Warusfel

► **To cite this version:**

Filippo Fazi, Markus Noisternig, Olivier Warusfel. Representation of sound fields for audio recording and reproduction. Acoustics 2012, Apr 2012, Nantes, France. hal-00811313

HAL Id: hal-00811313

<https://hal.science/hal-00811313v1>

Submitted on 23 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ACOUSTICS 2012

Representation of sound fields for audio recording and reproduction

F. M. Fazi^a, M. Noisternig^b and O. Warusfel^b

^aUniversity of Southampton, Highfield, SO171BJ Southampton, UK

^bInstitut de Recherche et Coordination Acoustique/Musique, 1, place Igor Stravinsky 75004 Paris

ff1@isvr.soton.ac.uk



ACOUSTICS 2012

Representation of sound fields for audio recording and reproduction

F. M. Fazi^a, M. Noisternig^b and O. Warusfel^b

^aUniversity of Southampton, Highfield, SO171BJ Southampton, UK

^bInstitut de Recherche et Coordination Acoustique/Musique, 1, place Igor Stravinsky 75004 Paris

ff1@isvr.soton.ac.uk

Spherical and circular microphone arrays and loudspeaker arrays are often used for the recording and reproduction of a given sound field. A number of approaches and formats are available for the representation of the recorded field, some of which are more popular among the audio engineering community (B-format, virtual microphones, etc.) whilst other representations are widely used in the literature on mathematics (generalized Fourier series, single layer potential, Herglotz wave function, etc.). In this paper, some of the properties of the various approaches and some of their differences are discussed. The analysis includes the representation of near-field sources and the choice of basis functions used for the representation.

1 Introduction

In 3-D audio applications the auditory scene is fully described by the sound field $p(\mathbf{x}, \omega)$. We present several strategies to describe this sound field with respect to the listener, whose position (the centre of his/her head) is the origin of the coordinate system. The position vector \mathbf{x} belongs to the two-dimensional or three-dimensional Cartesian space. The norm of that vector is $x := |\mathbf{x}|$ while its direction is identified by the unitary vector $\hat{\mathbf{x}} := \mathbf{x}/x$. The relation between cartesian and polar/spherical coordinates is given by $\mathbf{x} = [x_1, x_2] = [x \cos \phi_x, x \sin \phi_x]$, $\mathbf{x} \in \mathbb{R}^2$, $\mathbf{x} = [x_1, x_2, x_3] = [x \sin \theta_x \cos \phi_x, x \sin \theta_x \sin \phi_x, x \cos \theta_x]$, $\mathbf{x} \in \mathbb{R}^3$.

2 Representation of sound fields

We want to represent a sound field $p(\mathbf{x})$ that satisfies the homogeneous Helmholtz equation

$$\nabla^2 p(\mathbf{x}, \omega) + k^2 p(\mathbf{x}, \omega) = 0, \quad \mathbf{x} \in V \quad (1)$$

where $k = \omega/c$ is the wave number, ω is the angular frequency and c is the speed of sound, which is assumed to be uniform. V is a bounded subset of \mathbb{R}^2 or \mathbb{R}^3 , depending on whether we are considering a two-dimensional (2-D) or three-dimensional (3-D) problem. Hereafter we will assume that V is a ball (the region of space bounded by a circle - 2D - or by a sphere - 3D) with radius R that contains the listener(s) or the listener's head. The fact that the field p satisfies the homogeneous Helmholtz equation in V implies that the acoustic scattering effect due to the listener's body are not taken into consideration here.

2.1 Integral representation

A simple representation of the sound field is given by a linear superposition of propagating plane waves $\exp(-ik\mathbf{x} \cdot \hat{\mathbf{y}})$ impinging from all possible directions $\hat{\mathbf{y}}$. Such representation is given by the following integral operator

$$p(\mathbf{x}, \omega) = (H\varphi)(\mathbf{x}, \omega) := \int_{\Omega} \exp(-ik\mathbf{x} \cdot \hat{\mathbf{y}}) \varphi(\hat{\mathbf{y}}, \omega) dS(\hat{\mathbf{y}}) \quad (2)$$

$i = \sqrt{-1}$ is the imaginary unit and Ω is the unitary sphere S^2 for the three-dimensional problem or the unitary circle S^1 for the two-dimensional problem. In the mathematical literature, the integral operator H is referred to as the *Herglotz Wave Function* (HWF) [1]. The complex-valued function $\varphi(\hat{\mathbf{y}}, \omega)$ represents the amplitude and phase of the plane waves impinging from each given direction $\hat{\mathbf{y}}$. This function is usually referred to as the Herglotz kernel or Herglotz density (HD). In the simple case of a single plane wave, arriving from direction $\hat{\mathbf{z}}$ we have that $p(\mathbf{x}) = \exp(-ik\mathbf{x} \cdot \hat{\mathbf{z}})$ and the

HD density is the (translated) Dirac delta function, that is $\varphi(\hat{\mathbf{y}}) = \delta(\hat{\mathbf{y}} - \hat{\mathbf{z}})$.

The HWF provides a simple and intuitive representation of the field, but it has some disadvantages: firstly, as will be shown later, only a subset of the sound fields that satisfy (1) can be represented by (2). For instance, the HWF is not adequate for the representation of the near-field component (with respect to the listener) of the sound field. Secondly, the function $\varphi(\mathbf{y}, \omega)$ generally belongs to an infinite dimensional space (typically to $L^2(\Omega)$), which means that an infinite amount of information will be required to describe φ (and an infinite amount of sensor will be required to measure it!). This is however always going to be the case for all representation unless we either make some assumptions on the structure of the field (for example it can be assumed that the density is sparse in some domain) or we choose to provide an approximation of φ .

In order to overcome the first obstacle and represent near-field sources, we can use a different integral operator, the so-called *single-layer potential* (SLP) [1]. This can be interpreted as a continuous distribution of omnidirectional point-sources (three-dimensional problem) or line-sources (two-dimensional problem), continuously distributed on the boundary ∂V of the ball V . Such operator is defined as follows:

$$p(\mathbf{x}, \omega) = (S\psi)(\mathbf{x}, \omega) := \int_{\partial V} G(\mathbf{y}, \mathbf{x}, \omega) \psi(\mathbf{y}, \omega) dS(\mathbf{y}) \quad (3)$$

where the free-field Green's function $G(\mathbf{y}, \mathbf{x}, \omega)$ is given by [1]

$$G(\mathbf{y}, \mathbf{x}, \omega) = \begin{cases} \exp(ik|\mathbf{x} - \mathbf{y}|)(4\pi|\mathbf{x} - \mathbf{y}|)^{-1} & \text{3-D} \\ \frac{i}{4} H_0(k|\mathbf{y} - \mathbf{x}|) & \text{2-D} \end{cases} \quad (4)$$

H_0 is the zero-order Hankel function of the first kind. As for the HWF, the complex-valued function $\psi(\mathbf{y}, \omega)$ is referred to as density of the potential and represents the strength of the point-sources. Since ∂V is a sphere or a circle, the SLP can be simplified as follows:

$$(S\psi)(\mathbf{x}, \omega) = \int_{\Omega} G(R\hat{\mathbf{y}}, \mathbf{x}, \omega) \psi(\hat{\mathbf{y}}, \omega) R^\alpha dS(\hat{\mathbf{y}}) \quad (5)$$

where $\alpha = 1$ for the two dimensional case and $\alpha = 2$ in three dimensions.

As for φ , the description of ψ generally requires an infinite amount of information. The SLP allows for the representation of any sound field that satisfies (2) [2].

2.2 Series representation

A further method to represent the field p is given by the linear superposition of spherical (3-D) or cylindrical (2-D) standing waves $\Phi_n(\mathbf{x})$. The expressions of these standing waves are

$$\begin{aligned} \Phi_n(\mathbf{x}, \omega) &= j_\nu(kx) Y_\nu^\mu(\theta_x, \phi_x), & \text{3D} \\ \Phi_n(\mathbf{x}, \omega) &= J_\nu(kx) \exp(iv\phi_x) (2\pi)^{-1/2}, & \text{2D} \end{aligned} \quad (6)$$

where J_ν is the ν -th order Bessel function, j_ν is the ν -th order spherical Bessel function and Y_ν^μ are spherical harmonics as defined in reference [3]. The relation between n, ν, μ is given by

$$\nu = \begin{cases} (n-1)/2, & n \text{ odd} \\ -n/2, & n \text{ even} \end{cases} \quad (7)$$

for the two-dimensional case, whilst

$$\begin{aligned} \nu &= \lceil \sqrt{n} - 1 \rceil \\ \mu &= n - 1 - \nu - \nu^2 \end{aligned} \quad (8)$$

for the three-dimensional case, where the symbol $\lceil \cdot \rceil$ is the ceiling operator (round argument to the nearest integer towards $+\infty$).

Any sound field that satisfies (1) can be represented by the following series [3]

$$p(\mathbf{x}, \omega) = \sum_{n=1}^{\infty} \Phi_n(\mathbf{x}, \omega) c_n(\omega) \quad (9)$$

The sound field in V is fully described by the set of coefficients c_n . This set of coefficients is infinite but countable. We can define a synthesis operator W , acting on the set $\{c_n\}$ of coefficients, such that

$$W\{c_n\} = \sum_{n=1}^{\infty} \Phi_n(\mathbf{x}, \omega) c_n(\omega) \quad (10)$$

2.3 Series representation of densities

As explained above, the densities φ and ψ are functions defined over a circle or a sphere. They can therefore be represented by a series as follows

$$\varphi(\hat{\mathbf{y}}, \omega) = \sum_{n=1}^{\infty} \tilde{\Psi}_n(\hat{\mathbf{y}}) \varphi_n(\omega) \quad (11)$$

The series coefficients φ_n are computed from the inner product of φ with the functions Ψ_n . In symbols

$$\varphi_n(\omega) = \langle \Psi_n | \varphi \rangle := \int_{\Omega} \Psi_n(\hat{\mathbf{y}})^* \varphi(\hat{\mathbf{y}}, \omega) dS(\hat{\mathbf{y}}) \quad (12)$$

Clearly, the same representation can be used for ψ . The set of functions Ψ_n defines a frame for the space of function to which the density belongs and span that space. The functions $\tilde{\Psi}_n$ constitute the dual frame. The reader can refer to [4] for an introduction to the frame theory.

A special case for a frame is when the functions have unitary norm and are mutually orthogonal, that is when $\langle \Psi_n | \Psi_m \rangle = \delta_{n,m}$ (the symbol on the right-hand side is the Kronecker delta), and form an orthonormal basis for the space they span. In this case we also have that $\tilde{\Psi}_n = \Psi_n$. In our case, since the densities functions we are interested in are defined on the sphere or on the circle, the most intuitive choice is to use spherical harmonics and complex exponentials as basis functions, respectively. Thus we write the following (generalized) Fourier series

$$\varphi(\hat{\mathbf{y}}, \omega) = \sum_{n=1}^{\infty} \Psi_n(\hat{\mathbf{y}}) \varphi_n(\omega) \quad (13)$$

$$\varphi_n(\omega) = \langle \Psi_n | \varphi(\cdot, \omega) \rangle \quad (14)$$

where

$$\begin{aligned} \Psi_n(\hat{\mathbf{y}}) &= Y_\nu^\mu(\theta_y, \phi_y), & 3\text{D} \\ \Psi_n(\hat{\mathbf{y}}) &= \exp(i\nu\phi_y)(2\pi)^{-1/2}, & 2\text{D} \end{aligned} \quad (15)$$

and the relations between n, ν and μ are given by formulae (7) and (8).

The use of orthonormal bases for the representation of the density has many advantages, but it is by no means the only possibility. The use of non-orthogonal and/or overcomplete frames (like Gabor frames [5]) may be a better choice for some applications.

2.4 Relation between different representations

We have seen that a given sound field can be represented by means of the following operators

$$p = \begin{cases} H\varphi, & \text{Herglotz wave function} \\ S\psi, & \text{single layer potential} \\ W\{c_n\}, & \text{Standing waves superposition} \end{cases}$$

We will refer to the densities φ , ψ and to the set of coefficients $\{c_n\}$ as *formats*. We will present now some simple relations that allow to obtain one format from the other.

We start with the well-know Jacobi-Anger expansion that allows for the representation of a plane wave by superposition of spherical waves. This is given by [1]

$$\exp(-ik\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}) = \sum_{n=1}^{\infty} 2\pi\alpha r_n(kx) \Psi_n(\hat{\mathbf{x}}) \Psi_n(\hat{\mathbf{y}})^* \quad (16)$$

where $r_n(kx) = i^{-\nu(n)} J_{\nu(n)}(kx)$, $\alpha = 1$ for the two dimensional case and $r_n(kx) = i^{-\nu(n)} j_{\nu(n)}(kx)$, $\alpha = 2$ in three dimensions. We have omitted the dependency on ω to simplify the notation, and we will do so also hereafter. Substituting the complex exponential in the Herglotz wave function (2) with the Jacobi-Anger expansion above leads to

$$\begin{aligned} p(\mathbf{x}) &= \int_{\Omega} \sum_{m=1}^{\infty} 2\pi\alpha r_m(kx) \Psi_m(\hat{\mathbf{x}}) \Psi_m(\hat{\mathbf{y}})^* \\ &\times \sum_{n=1}^{\infty} \Psi_n(\hat{\mathbf{y}}) \langle \Psi_n | \varphi \rangle dS(\hat{\mathbf{y}}) \end{aligned} \quad (17)$$

The orthogonality relation of the basis functions

$$\langle \Psi_m | \Psi_n \rangle = \int_{\Omega} \Psi_m(\hat{\mathbf{y}})^* \Psi_n(\hat{\mathbf{y}}) dS(\hat{\mathbf{y}}) = \delta_{n,m} \quad (18)$$

yields

$$p(\mathbf{x}) = \sum_{n=1}^{\infty} 2\pi\alpha r_n(kx) \Psi_n(\hat{\mathbf{x}}) \langle \Psi_n | \varphi \rangle \quad (19)$$

The comparison of this result with equation (6) leads to the following relation between the Herglotz density and the standing wave coefficients:

$$c_n = 2\pi\alpha i^{-\nu(n)} \langle \Psi_n | \varphi \rangle \quad (20)$$

$$\varphi(\hat{\mathbf{y}}) = \sum_{n=1}^{\infty} (2\pi\alpha)^{-1} i^{\nu(n)} c_n \Psi_n(\hat{\mathbf{y}}) \quad (21)$$

We derive a similar relation for the single layer potential density ψ . We start from the expansion of the free-field Green's function (sometimes referred to as addition theorem for the fundamental solution)

$$G(\mathbf{y}, \mathbf{x}) = \sum_{n=1}^{\infty} i^{\nu(n)} \rho_n(ky) r_n(kx) \Psi_n(\hat{\mathbf{x}}) \Psi_n(\hat{\mathbf{y}})^* \quad (22)$$

where

$$\rho_n(ky) = \begin{cases} ikh_{\nu(n)}(ky) & 3\text{-D} \\ i\pi H_{\nu(n)}(ky)/2 & 2\text{-D} \end{cases} \quad (23)$$

H_ν and h_ν are Hankel functions and spherical Hankel functions of the first kind, respectively. Substituting this formula in the expression of the single layer potential (5) and with the same mathematical manipulations as those we applied above we obtain

$$c_n = R^\alpha \rho_n(kR) \langle \Psi_n | \psi \rangle \quad (24)$$

$$\psi(\hat{\mathbf{y}}) = \sum_{n=1}^{\infty} \frac{c_n}{R^\alpha \rho_n(kR)} \Psi_n(\hat{\mathbf{y}}) \quad (25)$$

The factor R^α originates from the domain of integration being ∂V instead of Ω .

Combining the results above we obtain the following formulae to compute the Herglotz density from the SLP density and vice-versa:

$$\psi(\hat{\mathbf{y}}) = \sum_{n=1}^{\infty} \frac{2\pi\alpha}{i^{\nu(n)} R^\alpha \rho_n(kR)} \langle \Psi_n | \varphi \rangle \Psi_n(\hat{\mathbf{y}}) \quad (26)$$

$$\varphi(\hat{\mathbf{y}}) = \sum_{n=1}^{\infty} \frac{i^{\nu(n)} R^\alpha \rho_n(kR)}{2\pi\alpha} \langle \Psi_n | \psi \rangle \Psi_n(\hat{\mathbf{y}}) \quad (27)$$

We consider now two special cases: a plane wave impinging from direction $\hat{\mathbf{z}}$ and a point source at position \mathbf{z} , where $z > R$. From the equations (16) and (22) we obtain

$$c_n^{PW} = 2\pi\alpha i^{-\nu(n)} \Psi_n(\hat{\mathbf{z}}) \quad (28)$$

$$c_n^{SW} = \rho_n(kz) \Psi_n(\hat{\mathbf{z}}) \quad (29)$$

In the light equations (21) and (25) we obtain the following expressions for the Herglotz and SLP densities

$$\varphi(\hat{\mathbf{y}})^{PW} = \sum_{n=1}^{\infty} \Psi_n(\hat{\mathbf{z}})^* \Psi_n(\hat{\mathbf{y}}) = \delta(\hat{\mathbf{y}} - \hat{\mathbf{z}}) \quad (30)$$

$$\varphi(\hat{\mathbf{y}})^{SW} = \sum_{n=1}^{\infty} \frac{i^{\nu(n)} \rho_n(kz)}{2\pi\alpha} \Psi_n(\hat{\mathbf{z}})^* \Psi_n(\hat{\mathbf{y}}) \quad (31)$$

$$\psi(\hat{\mathbf{y}})^{PW} = \sum_{n=1}^{\infty} \frac{2\pi\alpha}{i^{\nu(n)} R^\alpha \rho_n(kR)} \Psi_n(\hat{\mathbf{z}})^* \Psi_n(\hat{\mathbf{y}}) \quad (32)$$

$$\psi(\hat{\mathbf{y}})^{SW} = \sum_{n=1}^{\infty} \frac{\rho_n(kz)}{R^\alpha \rho_n(kR)} \Psi_n(\hat{\mathbf{z}})^* \Psi_n(\hat{\mathbf{y}}) \quad (33)$$

The fact that the absolute values of the Hankel functions and spherical Hankel functions diverge close to the origin suggests that the series (31) diverges. This can be shown for spherical geometry and small kz by considering the norm of φ^{SW}

$$\|\varphi^{SW}\| = \left(\sum_{\nu=0}^{\infty} \sum_{\mu=\nu}^{\nu} \left| \frac{i^{\nu+1}}{4\pi} h_\nu(kz) Y_\nu^m u \right|^2 \right)^{\frac{1}{2}} \quad (34)$$

and observing that, in view of the high-order approximation of the spherical Hankel functions [3]

$$h_\nu(kz) \approx -i \frac{(2\nu)!}{2^{\nu} \nu! (kz)^{\nu+1}}, \nu \gg kz \quad (35)$$

the density φ is not bounded (with respect to L^2 metrics). This implies that the field due to a point source cannot be represented accurately by the HWF with bounded density $\|\varphi\| < \infty$.

3 Approximate representation

All the formats for sound field representations presented above require an infinite amount of information, as the densities φ and ψ are functions that belong in general to an infinite dimensional space and the set of series coefficients $\{c_n\}$ is infinite. It is however possible to provide an approximate representation by expressing φ, ψ and c_n with a finite amount of information. The most intuitive way for this is to remove infinitely many elements from the series (9) and (11) and to keep only a finite set of coefficients.

Obviously we want to measure and store only that piece of information that is most *important* for the given metric that we choose to use. It may be argued that the ideal compression algorithm for spatial audio should act as an operator \mathcal{A} that *projects* one of the sound field descriptors, say for example φ , onto a finite dimensional subspace $\mathcal{H} \subseteq L^2(\Omega)$, such that the the norm $\|\mathcal{H}\mathcal{A}\varphi - \mathcal{H}\varphi\|_{\mathcal{L}}$ with respect to a psychoacoustically-motivated metric \mathcal{L} is minimal. If we know a basis $\{\Psi_n\}_{n \in \{1, \dots, N\}}$ for that space, we can write

$$\mathcal{A}\varphi(\mathbf{y}) = \sum_{n=1}^N \tilde{\Psi}_n(\mathbf{y}) \langle \Psi_n | \varphi \rangle \quad (36)$$

$$\varphi(\mathbf{y}) = \mathcal{A}\varphi(\mathbf{y}) + \tilde{\mathcal{A}}\varphi(\mathbf{y}) \quad (37)$$

where the orthogonal functions Ψ_n span \mathcal{H} and the operator $\tilde{\mathcal{A}}$ projects φ onto the orthogonal complement of \mathcal{H} .

4 Microphone techniques

The complex directivity pattern of a microphone $D(\hat{\mathbf{y}}, \omega)$ evaluated for a given direction $\hat{\mathbf{y}}$ can be regarded as the response of the microphone with respect to a plane wave impinging on the transducer from that direction $\hat{\mathbf{y}}$. When sensing a sound field $p = H\mathbf{a}$ with a microphone, the microphone signal $s(\omega)$ can be regarded as the inner product of the density φ with the directivity function D , that is

$$s(\omega) = \langle D(\cdot, \omega) | \varphi(\cdot, \omega) \rangle \quad (38)$$

For example we can sense the sound field $p = H\varphi$ with a cardioid microphone, whose directivity function (ideally frequency independent) is given by

$$D(\hat{\mathbf{y}}) = \frac{1}{2} (1 + \cos \theta_y) = \sqrt{\pi} Y_0^0(\hat{\mathbf{y}}) + \sqrt{\frac{\pi}{3}} Y_1^0(\hat{\mathbf{y}}) \quad (39)$$

where the axis of the cardioid is orientated towards the north pole ($\theta = 0$). The microphone signal is given by

$$\begin{aligned} s &= \int_{\Omega} \left[\sum_{\nu, \mu} Y_\nu^\mu(\hat{\mathbf{y}}) \varphi_{\nu, \mu} \right] D(\hat{\mathbf{y}})^* dS(\hat{\mathbf{y}}) \\ &= \sqrt{\pi} (\varphi_{0,0} + 3^{-1/2} \varphi_{1,0}) \end{aligned} \quad (40)$$

where

$$\varphi(\hat{\mathbf{y}}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} Y_\nu^\mu(\hat{\mathbf{y}}) \varphi_{\nu, \mu} \quad (41)$$

In fact, given two plane waves impinging from direction $\theta_{z1} = 0$ and $\theta_{z2} = \pi$, respectively, we see that (see equation (30))

$$s_{z1} = \sqrt{\pi} (Y_0^0(0,0) + 3^{-1/2} Y_1^0(0,0)^*) = 1 \quad (42)$$

$$s_{z1} = \sqrt{\pi} (Y_0^0(\pi,0) + 3^{-1/2} Y_1^0(\pi,0)^*) = 0 \quad (43)$$

4.1 Blumlein pair

It is possible to sense the sound field with an array M of ideally coincident directional microphones, directed towards different directions, thus obtaining a set of signals $\{s_m\}_{m \in \{1, \dots, M\}}$. A widely used microphone array configuration is the so-called *Blumlein pair*, wherein two ideally coincident figure-of-eight microphones are used, arranged in such a way that their axes lie on the horizontal plane and are perpendicular. The directivity of the two microphones is therefore given by

$$\begin{aligned} D_1(\hat{\mathbf{y}}) &= \cos \phi_y \sin \theta_y \\ &= \sqrt{\frac{2\pi}{3}} \left[-Y_1^1(\theta_y, \phi_y) + Y_1^{-1}(\theta_y, \phi_y) \right] \end{aligned} \quad (44)$$

$$\begin{aligned} D_2(\hat{\mathbf{y}}) &= \sin \phi_y \sin \theta_y \\ &= i \sqrt{\frac{2\pi}{3}} \left[+Y_1^1(\theta_y, \phi_y) + Y_1^{-1}(\theta_y, \phi_y) \right] \end{aligned} \quad (45)$$

The microphone signals can be simply derived as

$$s_1(\omega) = \sqrt{\frac{2\pi}{3}} \left[-\varphi_{1,1}(\omega) + \varphi_{1,-1}(\omega) \right] \quad (46)$$

$$s_2(\omega) = i \sqrt{\frac{2\pi}{3}} \left[\varphi_{1,1}(\omega) + \varphi_{1,-1}(\omega) \right] \quad (47)$$

This basic microphone array allows to reconstruct the projection of the Herglotz density φ onto the subspace \mathcal{H} spanned by $Y_1^1(\theta_y, \phi_y)$ and $Y_1^{-1}(\theta_y, \phi_y)$. The microphone directivity functions represent an orthogonal basis for that space (although not orthonormal), thus the reconstructed Herglotz density is given by

$$\varphi_{\mathcal{H}}(\hat{\mathbf{y}}, \omega) = \frac{3}{4\pi} \left[s_1(\omega) D_1(\hat{\mathbf{y}}) + s_2(\omega) D_2(\hat{\mathbf{y}}) \right] \quad (48)$$

The factor $3(4\pi)^{-1}$ has been introduced since the basis is orthogonal but not orthonormal, but can be removed. The functions $3(4\pi)^{-1} D_m(\hat{\mathbf{y}})$ define the dual basis for the array [4].

The Herglotz density for the case of a single plane wave impinging from direction $\hat{\mathbf{z}}$ is given by

$$\begin{aligned} \varphi_{\mathcal{H}}(\hat{\mathbf{y}}) &= \frac{3}{4\pi} \sin(\theta_y) \sin(\theta_z) \cos(\phi_y - \phi_z) \\ &= Y_1^{-1}(\hat{\mathbf{y}}) Y_1^{-1}(\hat{\mathbf{z}})^* + Y_1^1(\hat{\mathbf{y}}) Y_1^1(\hat{\mathbf{z}})^* \end{aligned} \quad (49)$$

which resembles the figure-of-eight directivity function, with its axis rotated towards the direction identified by $\hat{\mathbf{z}}$.

4.2 Soundfield microphone

A more complex example is given by the so-called soundfield microphone array, used for Ambisonics recordings. This array is constituted by four coincident cardioid (or subcardioid) capsules, whose axes point towards the vertices of a regular tetrahedron. In this case the directivity functions of the four microphones are given by the following equations:

$$D_1(\hat{\mathbf{y}}) = \frac{1 + \cos \theta_y}{2} \quad (50)$$

$$D_2(\hat{\mathbf{y}}) = \frac{1 + \cos \gamma \cos \theta_y + \sin \gamma \sin \theta_y \cos \phi_y}{2} \quad (51)$$

$$\begin{aligned} D_3(\hat{\mathbf{y}}) &= \frac{1 + \cos \gamma \cos \theta_y + \sin \gamma \sin \theta_y \cos(\phi_y - \frac{2}{3}\pi)}{2} \\ D_4(\hat{\mathbf{y}}) &= \frac{1 + \cos \gamma \cos \theta_y + \sin \gamma \sin \theta_y \cos(\phi_y + \frac{2}{3}\pi)}{2} \\ \gamma &= 2 \arctan(\sqrt{2}) \end{aligned}$$

The microphone signals can be calculated as in the previous cases and are expressed by a linear combination of the zero and first order spherical-harmonics. The Herglotz density reconstructed when sensing the field with a soundfield microphone is the orthogonal projection of φ onto the subspace spanned by $Y_0^0, Y_1^0, Y_1^1, Y_1^{-1}$.

The reproduced Herglotz density is computed by the sum of the product of the microphone signals with the dual basis for that array. Note that in this case the microphone directivity functions are not mutually orthogonal. The dual basis functions or dual directivity functions \tilde{D}_m are given by

$$\begin{aligned} \tilde{D}_1(\hat{\mathbf{y}}) &= \frac{1 + 9 \cos \theta_y}{8\pi} \\ \tilde{D}_2(\hat{\mathbf{y}}) &= \frac{1 + 9 \left[\cos \gamma \cos \theta_y + \sin \gamma \sin \theta_y \cos \phi_y \right]}{8\pi} \\ \tilde{D}_3(\hat{\mathbf{y}}) &= \frac{1 + 9 \left[\cos \gamma \cos \theta_y + \sin \gamma \sin \theta_y \cos(\phi_y - \frac{2}{3}\pi) \right]}{8\pi} \\ \tilde{D}_4(\hat{\mathbf{y}}) &= \frac{1 + 9 \left[\cos \gamma \cos \theta_y + \sin \gamma \sin \theta_y \cos(\phi_y + \frac{2}{3}\pi) \right]}{8\pi} \end{aligned}$$

It can be shown that $\langle D_m | \tilde{D}_n \rangle = \delta_{n,m}$, which defines the biorthogonality relation between the basis and the dual basis [4]. The projected density is therefore

$$\varphi_{\mathcal{H}}(\hat{\mathbf{y}}, \omega) = \sum_{m=1}^4 s_m(\omega) \tilde{D}_m(\hat{\mathbf{y}}) \quad (52)$$

It is of course possible to extract the coefficients $\varphi_{0,0}, \varphi_{1,0}, \varphi_{1,-1}, \varphi_{1,1}$ from the microphone signals. This operation is known as the *Ambisonics encoding* process [6], and the coefficients $\varphi_{0,0}, \varphi_{1,0}, \varphi_{1,-1}, \varphi_{1,1}$ (expressed as time-domain functions and for real-valued spherical harmonics) are known as *B-format* signals.

4.3 High-order spherical microphone arrays

A technique analogous to that described above can be applied to the so-called high-order spherical microphone arrays. These are constituted by many microphones, typically omnidirectional and flush-mounted on the surface of a rigid sphere. The transducers can be therefore regarded as directional microphones, but their directivity pattern $D_m(\hat{\mathbf{y}}, \omega)$ is typically represented by an infinite series of spherical harmonics. This implies that it is not possible to extract the series coefficients $\varphi_{0,0}, \varphi_{1,0}, \dots, \varphi_{N,N}$ with a finite number of microphones without contamination from higher order coefficients. This problem is usually referred as *spatial aliasing* [7]. Nevertheless, sensing the sound field with this kind of arrays should allow us to project the Herglotz density φ onto a finite dimensional subspace \mathcal{H} , whose dimension is given by the number M of microphones used. A set of functions $\Psi_n(\hat{\mathbf{y}}, \omega)$ (typically frequency dependent) that span this space may not be easy to derive analytically but may be computed numerically by discretization of the HWF and using techniques such as the singular value

decomposition [1, 2]. Assuming these functions and their dual frame $\tilde{\Psi}_n(\hat{\mathbf{y}}, \omega)$ are known, we can then express the projected kernel as usual:

$$\varphi_{\mathcal{H}}(\hat{\mathbf{y}}, \omega) = \sum_{n=1}^M \tilde{\Psi}_n(\hat{\mathbf{y}}, \omega) \langle \Psi_n | a \rangle \quad (53)$$

and the residual density $a - a_{\mathcal{H}}$ is orthogonal to the subspace \mathcal{H} spanned by the functions Ψ_n . These functions Ψ_n may be obtained as linear combinations of the directivity of the real microphones constituting the array:

$$\Psi_n(\hat{\mathbf{y}}, \omega) = \sum_{m=1}^M q_{n,m}(\omega) D_m(\hat{\mathbf{y}}, \omega) \quad (54)$$

Finally it is possible to combine the microphone signals in order to create a set of *virtual microphones*. The number J of virtual microphones may also be larger than the number of real microphones of the array, but they can at most define a (redundant) frame still for the same M -dimensional subspace \mathcal{H} . The virtual microphone signals v_j and their directivity functions \mathcal{D}_m are given by linear combinations of the signals and directivity functions of the real microphones. Thus

$$v_j(\omega) = \sum_{m=1}^M d_{j,m}(\omega) s_m(\omega) \quad (55)$$

$$\begin{aligned} &= \sum_{m=1}^M d_{j,m}(\omega) \langle D_m(\cdot, \omega) | a(\cdot, \omega) \rangle \\ &= \langle \mathcal{D}_j(\cdot, \omega) | a(\cdot, \omega) \rangle \end{aligned}$$

$$\mathcal{D}_j(\hat{\mathbf{y}}, \omega) = \sum_{m=1}^M d_{j,m}(\omega) D_m(\hat{\mathbf{y}}, \omega) \quad (56)$$

As usual, the projected Herglotz density is given by the sum of the dual frame/virtual directivity functions $\tilde{\mathcal{D}}$ weighted by the virtual microphone signals:

$$\varphi_{\mathcal{H}}(\hat{\mathbf{y}}, \omega) = \sum_{j=1}^J \tilde{\mathcal{D}}_j(\hat{\mathbf{y}}, \omega) v_j(\omega) \quad (57)$$

Acknowledgments

The authors would like to thank Dr Marco Liuni for his precious scientific contribution to this work.

This work has been partially funded by the ANR Contenus et Interactions (CONTINT 2010) project ‘‘Sample Orchestrator 2’’, by the U.K. Royal Academy of Engineering and by the Engineering and Physical Sciences Research Council.

References

[1] D. L. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, ser. Applied Mathematical Sciences. Berlin: Springer, 1992.

[2] F. M. Fazi, ‘‘Sound field reproduction,’’ Ph.D. dissertation, University of Southampton, 2010.

[3] E. G. Williams, *Fourier Acoustics : Sound Radiation and Nearfield Acoustical Holography*. San Diego: Academic Press, 1999.

[4] S. G. Mallat, *A wavelet tour of signal processing*, 2nd ed. San Diego, Ca.: Academic, 1999.

[5] P. Balazs, M. Dorfler, F. Jaillet, N. Holighaus, and G. Velasco, ‘‘Theory, implementation and applications of nonstationary gabor frames,’’ *Journal of Computational and Applied Mathematics*, vol. 236, no. 6, pp. 1481–1496, 2011.

[6] J. Daniel, J.-B. Rault, and J.-D. Polack, ‘‘Ambisonics encoding of other audio formats for multiple listening conditions,’’ in *Proc. 105th Conv. Audio Eng. Soc.*, San Francisco, California, 1998.

[7] B. Rafaely, B. Weiss, and E. Bachmat, ‘‘Spatial aliasing in spherical microphone arrays,’’ *IEEE Trans. on Signal Process.*, vol. 55, no. 3, pp. 1003–1010, 2007.