Combining beamforming methods and antenna rotation for sound sources localization with a spherical microphone array

Jean-Jacques Embrechts[1)]

[1)] Department of Electrical Engineering and Computer Science, Acoustics laboratory, Univ Liège, Quartier Polytech 1, Allée de la découverte, 10, 4000 Liège, Belgium. jjembrechts@ulg.ac.be

Abstract

A 16-microphones spherical array is used to localize and identify reflections in room acoustics. By increasing the number of microphones on the array, the spatial resolution is expected to be improved, but at the expense of cost and complexity. In this study, the number of microphone positions is rather increased by sequentially rotating the spherical array around its vertical axis. We show how to combine the microphone signals measured at different positions of the rotating table to increase the order of beamforming and obtain better spatial resolutions. In particular, four optimized azimuthal positions are defined and tested with three beamforming methods: PWD, DAS and MVDR. The method combining beamforming and antenna rotation is first tested with synthetic plane waves at different frequencies. The results of sound source localization are compared with the ones obtained with the original 16 microphone positions. It is shown that some improvements are obtained in the spatial resolution, depending on the frequency and the beamforming method. The combination method is then tested with real sound fields. Theoretical results are retrieved for 16 and 64 microphone positions.

1 Introduction

The context of this study is the measurement of spatial properties of the sound field in closed spaces with microphone arrays. In particular, we are interested in measuring directional room impulse responses (DRIR) [1].

Early experiments dedicated to these measurements used rotating directional microphones [2, 3]. One of the first application of microphone array for DRIR measurements was realized by Gerzon [4] whose objective was to record concert hall acoustics for posterity. More recently, Merimaa et al [5] used pairs of intensity probes in each x-, y- and z-directions to compute intensity vectors in a spectrogram-like map.

Spherical arrays for room acoustics applications were introduced in 2004 by Gover et al [6]: with a 32-microphones array and MLS excitation signals, they obtained DRIRs in a bandwith of 300 to 3000 Hz. The 16-microphones spherical array developed in our laboratory has been inspired by a similar instrument described by Clapp et al [7], whereas a 32-microphones array has been experimented by Farina and Tronchin [8] to mimic virtual 4th-order cardioid microphones pointing in different directions. With their measuring system, these authors can obtain multichannel room impulse responses and plot a colour map of the sound field spatial distribution, for each 1ms time-window of the measured RIR.

Increasing the number of microphones on the array increases the spatial resolution of DRIR measurements. This would be the main interest of developing 32-microphones systems instead of 16. However, this benefit is obtained at the expense of a greater number of microphones and signals to process. This implies that the sound cards used must be equipped with a greater number of input channels and that the computing facilities (PC or DSP) must be more powerful in order to process the measured signals, sometimes

in real-time. Even if the material can be available today for more and more channel processing, this can be very (too) expensive for some developers.

Therefore comes the idea of increasing the number of microphone positions without increasing the number of microphones and simultaneously processed channels, simply by sequentially rotating the microphone array or antenna. This idea is not new: for example, Rafaely at al [9] have measured DRIRs in an auditorium with an array consisting of 882 microphone positions on an open sphere (in fact, a dual-sphere with two radii). In order to obtain this high-resolution spatial instrument (9 degrees aperture), a single microphone is rotated and sequentially positioned at the 882 locations. The authors' objective was to achieve an accurate identification of early reflections in the room impulse response through an efficient time and spatial separation of these reflections.

The objective of this paper is to study the opportunities and the difficulties of combining the beamforming results obtained with a 16-microphones array positioned at several azimuthal locations. These locations are obtained by sequentially rotating the microphone antenna. Our study is particularly focused on the application of sound source localization or early reflections localization in rooms. Some particular questions that will be addressed are: how to combine the beamforming results obtained at different azimuthal positions ? Is there an optimal set of azimuthal positions ? Is there a significant impact on the spatial resolution in sound source localization tasks ?

Incident plane waves in free field will be used as test sound fields, instead of complex sound fields in enclosures: it is indeed much easier to interpret the impact of such simple sound waves on a spherical array. We guess that the scope of our study will not be restricted by this choice since after all, given the relatively small size of the sphere (20 cm) compared to usual room dimensions, complex sound fields can be viewed as a sum of contributions (reflections and diffractions) which can be locally approximated by incident plane waves.

The paper is organized as follows: in section 2, we recall the basic equations used to reconstruct the sound field on the spherical array, we briefly review the beamforming methods applied in this study and we describe our spherical microphone array. In section 3, we propose a method to combine sets of 16 acoustic pressures measured by the array at several azimuthal positions. An optimized combination is found for four azimuthal positions. In section 4, this method is tested at three different frequencies with synthetic sound fields (theoretical plane waves incident on the array) and three beamforming methods. In section 5, the method is applied to real sound fields and compared with previous theoretical results. Finally, we conclude in section 6.

2 Description of the 16-microphone array and beamforming methods

2.1 Source localization with a spherical array

The following section is inspired by several previous publications [1, 9, 10, 11]. In a first step of sound source localization, the spherical array of microphones is used to approximate the soundfield on its surface, through the spherical harmonics transform (spherical Fourier transform).

Let $P(k, \theta, \phi)$ be the sound pressure on the array surface, at any angular position defined by the usual pair of spherical coordinates (θ, ϕ). k is the wavenumber. The spherical harmonics representation of this pressure field is:

$$P(k, \theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} P_{nm}(k) Y_n^m(\theta, \phi) \qquad (1)$$

Y_n^m are the spherical harmonics functions [11] and P_{nm} are the spherical harmonics coefficients, which are obtained by the inverse transform:

$$P_{nm}(k) = \int_0^{\pi} sin\theta d\theta \int_0^{2\pi} P(k, \theta, \phi) Y_n^{m^*}(\theta, \phi) d\phi \qquad (2)$$

In this formula, the symbol * represents the complex conjugate operation.

The spherical harmonics coefficients can be evaluated if the sound pressure has been measured at the positions of the microphones. Suppose that there are J microphones on the sphere, then, according to [9]:

$$P_{nm}(k) \simeq \sum_{j=1}^{J} \alpha_j P(k, \theta_j, \phi_j) Y_n^{m^*}(\theta_j, \phi_j) \qquad (3)$$

This formula is in fact an approximation of the integral (2) by a linear combination of integrand's values known at specific angular positions (θ_j, ϕ_j). The coefficients of this combination α_j are determined by the choice of these positions on the sphere.

This formula and the choice of microphone positions is a key point in the study of this paper. In this work, we have used the positions defined by Fliege in his paper on cubature formulae for the sphere, as suggested by [7]. The Fliege positions define a nearly-uniform spatial sampling on the sphere.

The spatial sampling of the sound pressure field on the sphere would be exact (equality in equation 3) if the following was true $\forall n, n', m, m'$:

$$\sum_{j=1}^{J} \alpha_j Y_{n'}^{m'}(\theta_j, \phi_j) Y_n^{m^*}(\theta_j, \phi_j) = \delta_{n-n'} \delta_{m-m'} \qquad (4)$$

However, this condition (which is a modified version of the orthogonality property of the spherical harmonics functions) can only be approached in practice for a limited number of values of the indices n and n'. In general, increasing the number of microphone positions will increase this limit.

As a consequence, the coefficients P_{nm} can only be recovered for the values of the index $n \leq N$: this determines the order N of the sound field reconstruction in (1). Typically, we should have $(N+1)^2 \leq J$. Therefore, with 16 microphones located at the Fliege's positions, the order can be extended to $N = 3$.

Equation (3) of course does not prevent to calculate the spherical harmonics coefficients for higher indices, but the accuracy of the sound field reconstruction will not be guaranteed in this case.

m:	-3/3	-2/2	-1/1	0
n=0	-	-	-	0.000
n=1	-	-	0.001	0.001
n=2	-	0.001	0.001	0.001
n=3	0.336	0.345	0.345	0.162

Table 1: Maximum absolute errors in equation (4) for the 16 Fliege microphone positions, if N=3. The maximum errors are identical for corresponding positive and negative values of the index m

To illustrate this, we have computed the left part of equation (4) and compared it with the right part (1 or 0 according to the values of the indices n, n', m, m'). The difference between left and right parts (we call it an 'error') is a complex number whose maximum absolute value is given in table 1: in this table, the maximum for all n', m' is given, for each pair n, m ($n \leq 3$ and $n' \leq 3$). It can be seen that (4) is nearly perfectly verified if $n \leq 2$, which means that if the sound field is extended in (1) until order 2, then its reconstruction will be close to exact. It is also concluded that we can have significant errors on some $P_{\pm 3m}$ if the sound field is developed until order 3.

The situation is even worse for N=4, since it can be shown in that case that equation (4) is only perfectly verified if $n \leq 1$, some absolute errors being as high as 0.78 for $n \geq 3$.

After reconstruction of the sound field on the sphere, the second step of sound source localization is the beamforming operation: the incident sound field corresponding to a particular 'look-up' direction (θ_L, ϕ_L) is obtained by applying a set of weights W_{nm} to the spherical harmonics coefficients. The output (response) of the array is:

$$Q(k, \theta_L, \phi_L) = \sum_{n=0}^{N} \sum_{m=-n}^{n} P_{nm}(k) W_{nm}^*(k, \theta_L, \phi_L) \qquad (5)$$

These weights depend on the frequency and on the look-up direction. There exists several methods of beamforming. In this paper, we will apply only

three of them: the plane-wave decomposition method (PWD), the delay-and-sum method (DAS) and the minimum-variance distortionless response (MVDR) described in [12].

The PWD method is described in [9, 10, 11]. The weights can be expressed as:

$$W_{nm}^{*}(k, \theta_L, \phi_L) = \frac{d_n}{b_n(k)} Y_n^{m^*}(\theta_L, \phi_L) \qquad (6)$$

with $d_n = 1$ (for PWD) and b_n depends on the sphere boundary. For the rigid sphere with radius r:

$$b_n(k) = 4\pi i^n \left(j_n(kr) - \frac{j_n'(kr)}{h_n'(kr)} h_n(kr) \right) \qquad (7)$$

In (7), j_n and h_n are the spherical Bessel and Hankel functions, j_n' and h_n' their derivatives.

The choice $d_n = 1$ leads to a regular beam pattern in (5), if the sound field is created by a plane wave incident along the direction $\Omega_0 = (\theta_0, \phi_0)$: see an example of a regular beam pattern in figure 4. This beam pattern approaches the Dirac function $\delta(\Omega_0 - \Omega_L)$ as $N \to \infty$, which confirms that a single incident wave is collected by the antenna in this limiting case.

The DAS method consists in applying a different delay (or phase shift) to each output of the individual microphones such that the signals become in-phase for a plane wave coming from the look-up direction. It is shown in [11] that the weights are still defined by (6), with $d_n = |b_n(k)|^2$.

Finally, the MVDR method belongs to the class of optimal beamforming techniques: the weights are such that they minimize the contributions of other directions than the look-up one in the array's response. The weights have a more complex expression, which is also given in [11] and will not be discussed here.

The result of the beamforming will be expressed in this paper by a diagram showing the spatial distribution of $Q(k, \theta_L, \phi_L)$ as a function of both angles. Figure 1 is an example. In this diagram, a simple estimation of the sound source direction is given by the pair of angular values corresponding to the maximum output magnitude $|Q|$.

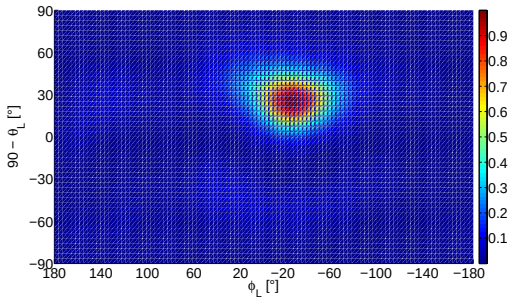

Figure 1: Example of the spatial distribution of the array's output magnitude, for all look-up directions around the sphere: ϕ_L is the azimuth in the horizontal plane and θ_L is the elevation angle. If we imagine the spherical array as being the head of a listener, positive azimuth angles are associated with look-up directions to the left and $\theta_L = 0$ is the look-up direction to the top. The array's output magnitude is normalized such that the maximum value for all look-up directions is equal to 1.

2.2 The spherical microphone array

The spherical array of microphones used in this study has been designed and build to allow for the measurement of directional room impulse responses [1]. Figure 2 shows a picture of the array.

16 low-cost electret omnidirectional microphones are slotted in the rigid sphere surface at the Fliege's position. In fact, the Fliege's positions have been slightly rotated by the same elevation angle of 10°, for practical reasons: to avoid the south pole region in which the support is connected to the sphere and also the north pole which lies on the separation between the two hemispheres (see figure 2).

The sensitivity of the microphones is -42 ± 3 dB (re $1V/Pa$) and their frequency response is flat between 50 Hz and 20 kHz. Their SNR is 58dB. They are connected to the 16 inputs of two soundcards Edirol Firewire FA 101 (eight inputs each) which allow for the recording of the signals at the sampling frequency of 48 kHz.

As this measuring equipment is intended to be used on-site and must be handy and portable, the size of the sphere has been limited. Its radius has been de-

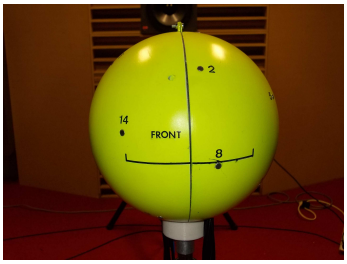

Figure 2: Front face of the microphone array. Some microphones appear on the picture and they are identified by their number. The vertical separation between two hemispheres corresponds to $\phi_L = 0$.

m:	-3/3	-2/2	-1/1	0
n=0	-	-	-	0.064
n=1	-	-	0.064	0.064
n=2	-	0.063	0.068	0.058
n=3	0.307	0.308	0.308	0.152

Table 2: Maximum absolute errors in equation (4) for the 16 Fliege microphone positions, if N=3, and for the optimum α_j's values. The maximum errors are identical for corresponding positive and negative values of the index m

termined after consideration of spatial aliasing errors, which increase with frequency, phase errors (mainly at low frequencies) and the optimization of the WNG (white noise gain). With a radius of 10cm, these errors are acceptable in the useful bandwith of [250-4000] Hz and the WNG is optimum at 1623 Hz for order N=3 (optimum at $kr = N$) [13, 14].

2.3 Optimum values of the α_j coefficients

The coefficients α_j in (3) are essential in the estimation of the spherical harmonics coefficients. Initial values are given by Fliege with more than ten significant digits: in fact, for the 16 Fliege's positions, the 16 α_j 's values are all equal (to 0.7557), except four of them (= 0.8746).

In the following, we propose a method to optimize these values without any reference to a particular sound field (the method does not depend on particular values of the P_{nm} coefficients). The method is based on equation (4). The optimum values of α_j are those which minimize the absolute differences between the left and right parts of the equality sign in (4). The problem is a MSE (mean square error) minimization, the cost to optimize being:

$$C = \sum_{n,n',m,m'}^{orderN} \left| \sum_{j=1}^{J} \alpha_j Y_{n',j}^{m'} Y_{n,j}^{m^*} - \delta_{n-n'}\delta_{m-m'} \right|^2 \quad (8)$$

In this expression, the following simplifications have been adopted: first, the symbol $\sum_{n,n',m,m'}^{orderN}$ represents a quadruple sum, in which the indices n and n' take the integer values from 0 to N, and the index m (resp. m') the integer values from $-n$ to n (resp. $-n'$ to n'). Secondly, the spherical harmonics function $Y_{n',j}^{m'}$ and $Y_{n,j}^{m^*}$ are both evaluated at the position of microphone number j in (θ_j, ϕ_j).

This MSE minimization leads to a simple linear system of 16 equations which is solved to give the optimum values of the α_j coefficients. If this is done with the 16 Fliege's positions and $N = 3$, the optimum α_j's are not very different from the initial values proposed by Fliege: all optimum values are equal (to 0.7304), except four of them (= 0.7498). The optimum cost $C = 1.0217$ is not significantly different from the cost calculated with the initial α_j's ($C = 1.1389$).

This result suggests that there's no special interest to replace the initial coefficients of Fliege by the optimum ones. Moreover, if we detail the optimum solution by computing the maximum absolute errors for different values of the indices n, m, we obtain the results in table 2. It is indeed seen that there's a small decrease of the errors for $n = 3$ (compared with table 1), which are the most significant. But this is obtained at the expense of non-zero errors for $n < 3$.

3 Increasing the number of microphone positions by rotating the sphere

3.1 New coefficients α_j

The number of microphone positions is here simply increased by rotating the sphere around its vertical axis $\overrightarrow{oZ}$. The rotation is defined by the azimuthal angle ψ (see figure 3). In this figure, a microphone is moved from its original position P to the new position P'. The new position (θ'_j, ϕ'_j) in the axis system XYZ is such that:

$$\theta'_j = \theta_j \tag{9}$$

$$\phi'_j = \phi_j + \psi \tag{10}$$

The initial number of microphone positions is therefore doubled and, if the measurements of the sound field are repeated at the new microphone positions, the results can be combined with the ones performed at the initial positions to form a new set of pressures $P(k, \theta_j, \phi_j)$ to estimate the spherical harmonics coefficients P_{nm}. However, the new set containing twice the number of microphone positions does not lead to a 'nearly-uniform' spatial sampling, according to Fliege, which means that we must find a new definition for the coefficients α_j in (3).

This definition is the following: the values of α_j will be those which minimize the cost function (8), for the new set of microphone positions (including the initial and new ones). Indeed, if this cost could be reduced to zero, then the reconstruction of the sound field would be exact at the order N.

Of course, it must be checked that the sound field $P(k, \theta, \phi)$ is constant during the two measurements, such that the two sets of results can be combined as if they were obtained at the same time, and not sequentially. This property of the sound field is assumed in the following: it will be discussed later in section 5.3.1.

The new coefficients α_j will of course depend on the rotation angle ψ. Furthermore, we are not restricted to just one additional ψ angle, but we can imagine several rotations of the sphere, repeating the sound field measurements at each angular position: if N_ψ is the number of angular positions of the sphere (including the initial one), this will create at the end $N_\psi.J = J_\psi$ microphone positions for which we have to optimize a similar number of coefficients α_j.

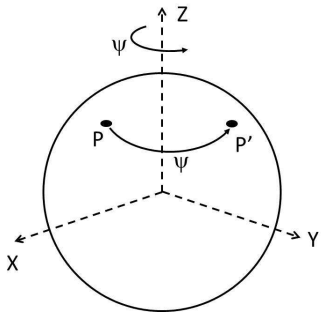

Figure 3: P is the original position of a microphone on the sphere. After azimuthal rotation around the Z axis by an angle ψ, the new position is in P'.

3.2 Results with 64 microphone positions

The objective of multiplying the number of microphone positions is to increase the order N of a correct reconstruction of the sound field by spherical harmonics decomposition. In practice, if order $N = 3$ can be expected with 16 microphones, as explained in the last sections, we would like to increase the order to $N = 4$. An estimation of the required number of microphone positions is $J_\psi \geq (N+1)^2 = 25$ [11, 14]. However, this lower bound could possibly be reached if the new distribution of microphone positions was nearly-uniform, which is not the case if we simply rotate the sphere around its vertical axis.

We therefore created a function which optimizes the cost (8) for a given number N_ψ of azimuthal positions of the sphere. This function returns:

- The optimum set of azimuthal positions ψ_l, for $1 \leq l \leq N_\psi$,
- The corresponding optimum values of the coefficients α_j,
- The optimum value of the cost function,

m:	-3/3	-2/2	-1/1	0
n=0	-	-	-	0.023
n=1	-	-	0.021	0.029
n=2	-	0.041	0.037	0.032
n=3	0.046	0.046	0.041	0.041

Table 3: Maximum absolute errors in equation (4) for the 64 optimum microphone positions (11) and their optimum α_j's values, if the sound field is reconstructed at order $N = 3$. The maximum errors are identical for corresponding positive and negative values of the index m

m:	-4/4	-3/3	-2/2	-1/1	0
n=0	-	-	-	-	0.023
n=1	-	-	-	0.049	0.029
n=2	-	-	0.052	0.037	0.032
n=3	-	0.072	0.070	0.086	0.108
n=4	0.095	0.087	0.078	0.086	0.108

Table 4: Same as table 3, if the sound field is reconstructed at order $N = 4$.

- The maximum absolute errors in equation (4) resulting from this new set of microphone positions associated with their optimum α_j.

Applying this function, it has been shown that a minimum of four azimuthal positions was necessary to obtain a good reconstruction of the sound field at order $N = 4$. The optimum set of four azimuthal positions (in fact, there's an infinity of optimum solutions) can be expressed as the following:

$$[\psi_1, \psi_1 + 30°, \psi_1 + 160°, \psi_1 + 310°] \qquad (11)$$

In this solution, the angle ψ_1 can take any value. The new set of 64 microphone positions is not optimal at order $N = 3$: as shown in table 3, the combination is better that what can be obtained with only 16 microphone positions (compare with table 2), especially for the computation of $P_{\pm 3m}$, but other sets of azimuthal positions can do better and reach practically negligible errors at order 3. These solutions will not be detailed here since we are interested to reach order $N = 4$.

Precisely, table 4 illustrates the maximum absolute errors in (4) for the 64 optimum microphone positions, associated with their optimum α_j's. These errors seem acceptable: approximately less than 0.05 for $P_{nm}(n \leq 2)$ and less than 0.11 if $n \leq 4$. The optimum cost function's value is $C_{opt,N=4} = 0.578$, which can be compared with the corresponding result obtained with the 16 Fliege's position: $C_{Fliege,N=4} = 19.449$.

Another significant difference between the optimized set of 16 versus 64 microphone positions is that the 64 optimum α_j's are no longer positive and approximately equal, as it was the case for 16 positions. Indeed, the coefficients' values are now comprised between -3 and $+3$, which will create some problems that will be discussed later.

3.3 Changing the cost function ?

The cost function (8) attributes the same weight to the errors in equation (4), according to the orders n, m, n', m'. Can this choice be justified ?

If we develop equations (1-4), it can be shown that the approximate value of the spherical harmonics coefficient $\hat{P}_{nm}$ computed by (3) is:

$$\hat{P}_{nm} = P_{nm} + \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} P_{n'm'} \epsilon_{n,n',m,m'} \qquad (12)$$

in which $\epsilon_{n,n',m,m'}$ is the error in (4). It seems therefore that a deeper analysis of the kind of sound field to reconstruct could give some information on the magnitude of the coefficients $P_{n'm'}$ and possible corresponding weights in the cost function.

In room acoustics, we can consider that the sound field at a microphone position is locally represented by a sum of plane waves arriving from different directions, amplitudes and delays. The contribution of any early reflection can therefore be assimilated with the contribution of a plane wave. The spherical harmonics coefficients of a plane wave of unit amplitude, incident upon a rigid sphere with a diameter of 20cm, along the direction Ω_0, are the following [11]:

$$P_{nm} = b_n(k) Y_n^{m^*}(\Omega_0) \qquad (13)$$

with $b_n(k)$ given by (7). Computing these coefficients for several frequencies and several directions of incidence has shown that:

- at 250 Hz, the P_{nm} are negligible (less than 0.02, compared to $P_{00} = 3.22$) if $n > 2$,
- at 1623 Hz, the P_{nm} are comprised between 1 and 3 if $n \leq 3$, start to decrease from $n = 4$ and reach values less than 0.1 if $n > 5$,
- at 3000 Hz, the P_{nm} magnitudes only decrease from $n = 10$.

This information about the spherical harmonics coefficients of a plane wave sound field shows that there's no particular reason to weight the contributions of the errors $\epsilon_{n,n',m,m'}$ in (8) differently, if we want to reconstruct the sound field at different frequencies until order $N = 4$.

4 Testing the optimized 64 microphone positions with synthetic sound fields

4.1 At the frequency 1623 Hz

As explained earlier, this frequency has first been chosen because it corresponds to the optimum value of the WNG for the reconstruction of the sound field at $N = 3$.

A great number of simulations have been done, applying the three beamforming methods presented in section 2.1 to theoretic sound fields created by plane waves incident on the spherical antenna. The spherical harmonics coefficients P_{nm} are determined at the order 3 or 4, using the 16 Fliege microphone positions or the 64 optimized ones. Then, beamforming is applied through equation (5) for several look-up directions around the sphere. Also, several directions of plane wave incidence have been analysed.

This big amount of results is summarized in the following. As the general conclusions are not really influenced by the direction of incidence, only the direction $(\theta_0, \phi_0) = (90°, 0°)$ has been presented.

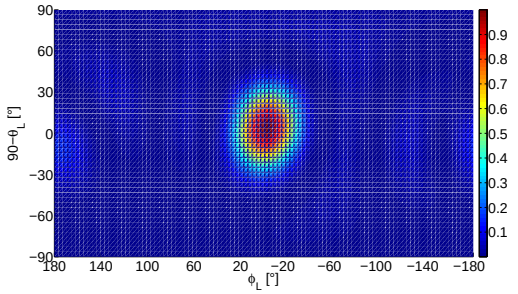

Figure 4: Spatial distribution of the array's output magnitude, for all look-up directions around the sphere, obtained by the PWD beamforming with $N = 3$ and 16 microphone positions.

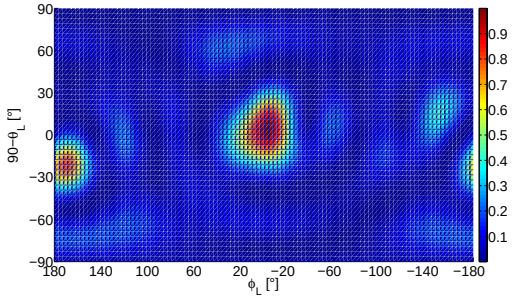

Figure 5: Same as figure 4 with $N = 4$.

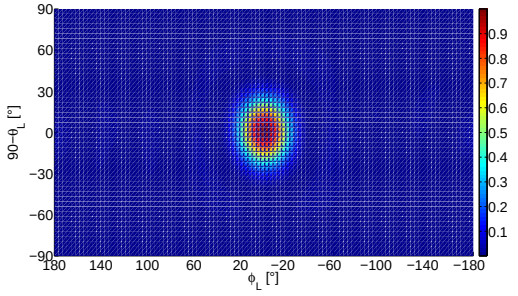

Figure 6: Same as figure 4 with $N = 4$ and 64 optimized microphone positions.

4.1.1 PWD beamforming

Figures 4 and 5 illustrate the results of the PWD method for an incident plane wave along the direction $(\theta_0, \phi_0) = (90°, 0°)$, for a reconstruction of the sound field at respectively $N = 3$ and $N = 4$, with 16 Fliege microphone positions. For $N = 3$, the source position is identified at the maximum output value in $(\theta_L, \phi_L) = (86.4°, 0°)$, just above the exact value, and the central lobe of sensitivity (at $-10dB$ from maximum) is approximately $\pm 30°$ wide. For $N = 4$, the quality of the results significantly deteriorates, due to the errors in the P_{nm} calculations with 16 microphones, as shown in figure 5.

On the other hand, the quality is improved with 64 optimized microphone positions. For $N = 3$, we obtain approximately the same results as in figure 4, except that the source position is now correctly identified in $(\theta_L, \phi_L) = (90°, 0°)$ and the main lobe of sensitivity is perfectly centred and symmetric. This is also the case for $N = 4$ in figure 6 (which can be compared with figure 5), the source position is still correct and the main lobe width has been slightly reduced to $\pm 25°$.

4.1.2 DAS beamforming

The DAS beamforming is not subject to the same sensitivity to P_{nm} errors than PWD. This is shown in figure 7 for $N = 4$ and 16 microphones at the Fliege's positions. For the DAS beamforming method, using 64 instead of 16 microphone positions does not bring significant improvements. Therefore, the results will not be illustrated here.

4.1.3 MVDR beamforming

Figures 8 and 9 illustrate the results of the MVDR method with 16 microphone positions. For $N = 3$, the maximum output value is in $(\theta_L, \phi_L) = (86.4°, 0°)$ and the main sensitivity lobe is thinner than with PWD (approximately $\pm 10°$ at $-10db$). For $N = 4$ and 16 microphones, the main lobe is now perfectly centred, but somewhat wider than with $N = 3$.

The quality is once again improved with 64 optimized microphone positions. For $N = 3$, the main lobe of sensitivity is perfectly centred and thinner than $\pm 1°$ at $-10db$. It is also the case for $N = 4$ (see figure 10), though the main lobe is still somewhat wider than with $N = 3$. However, the improvement compared with 16 microphones is outstanding (compare figures 9 and 10).

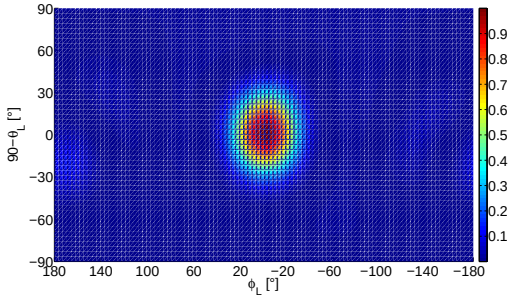

Figure 7: Spatial distribution of the array's output magnitude, for all look-up directions around the sphere, obtained by the DAS beamforming with $N = 4$ and 16 microphone positions.

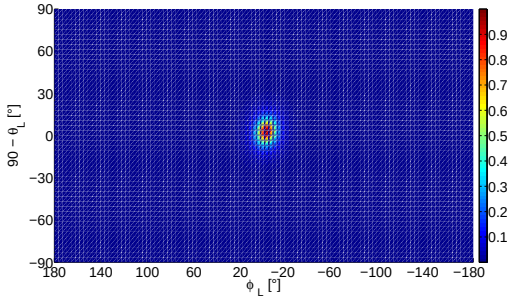

Figure 8: Spatial distribution of the array's output magnitude, for all look-up directions around the sphere, obtained by the MVDR beamforming with $N = 3$ and 16 microphone positions.

4.2 At the frequency 250 Hz

A similar analysis has been done at the low frequency bound of the microphone array's bandwith. The main conclusions are the following:

- the results of the PWD beamforming already deteriorates from $N = 3$ with 16 microphones. At

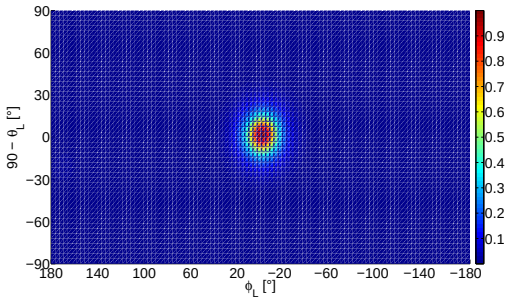

Figure 9: Same as figure 8 with $N = 4$.

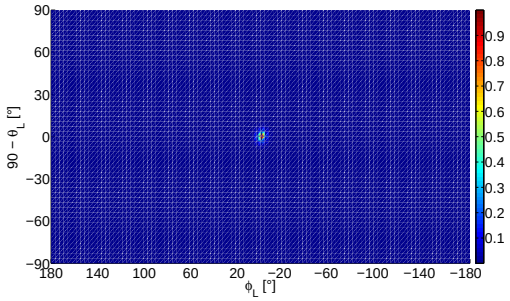

Figure 10: Same as figure 8 with $N = 4$ and 64 optimized microphone positions.

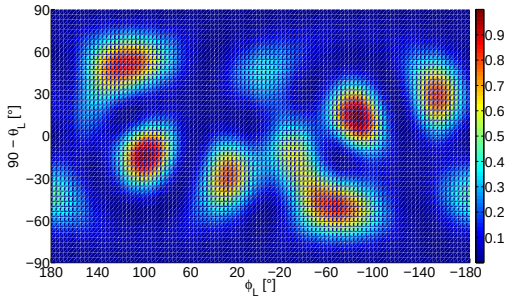

Figure 11: Spatial distribution of the array's output magnitude at 250 Hz, obtained by the PWD beamforming with $N = 3$ and 64 microphone positions. The plane wave is still incident along the angular coordinates $(\theta_0, \phi_0) = (90°, 0°)$.

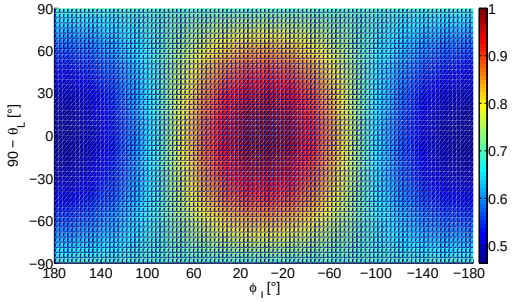

Figure 12: Spatial distribution of the array's output magnitude at 250 Hz, obtained by the DAS beamforming with $N = 4$ and 64 microphone positions.

this frequency, the 64 optimized microphone positions don't bring any improvement: the results are even worse, as shown in figure 11. This case will be analysed in the following.

- the DAS beamforming results provide a correct detection of the source angular position, but with a low spatial resolution as can be seen in figure 12. With this beamforming method, the results obtained with 16 or 64 microphone positions are similar, as for 1623 Hz.
- with the MVDR beamforming, the direction of incidence is correctly detected at 250 Hz, with a very good resolution (not shown), similar to the one obtained at 1623 Hz. The resolution obtained with 16 microphones is even somewhat better than with 64 microphones in this case, but the differences are not significant.

Now, what is the reason of the bad performance of the PWD method at low frequency ? As explained in section 3.3, the plane wave can be reconstructed at 250 Hz with only low order P_{nm}'s, so we would expect that errors on these coefficients for $n > 3$ would have little impact on the accuracy of the array's output. This is obviously not the case as shown in figure 11. The reason clearly appears if we look at equation (6) which gives the weights multiplying the P_{nm}'s in the expression of the array's output (5). For PWD beamforming, these weights are proportional to $\frac{1}{b_n(k)}$,

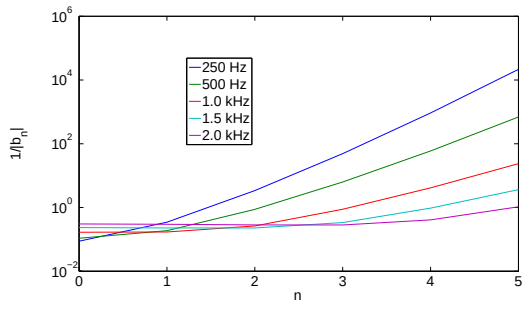

Figure 13: Inverse of $b_n(k)$ (absolute value) for several orders n and several frequencies;

a ratio that is illustrated in figure 13: this ratio significantly increases with the order n at low frequencies, resulting in a great amplification of the errors on P_{nm}. For the 64 optimized microphone positions, these errors are small for $n \leq 3$, but not zero as seen in tables 3 and 4, and these errors are so amplified that they completely deteriorate the accuracy of the results in figure 11.

This problem is greatly attenuated at 1623 Hz, since the weights W_{nm} are much smaller at this frequency, as shown in figure 13.

4.3 At the frequency 2500 Hz

Another problem appears at high frequencies, i.e. spatial aliasing. This is illustrated in figure 14 which shows the array's output computed by PWD beamforming with 16 microphones ($N = 3$). Secondary lobes appear, in particular around $(\theta_L, \phi_L) = (120°, 170°)$. In the following, the improvements brought by the optimized 64 microphone positions are described:

- with the PWD beamforming and 64 microphones positions, the main lobe is perfectly centred and the aliasing disappears at $N = 3$ and $N = 4$: see figure 15.
- the same improvement is observed with the DAS beamforming. With 64 microphone positions, the main lobe is perfectly centred at $N = 3$ and $N = 4$, the aliasing disappears and the width of the main lobe decreases ($\pm 22°$ for $N = 4$), which is new

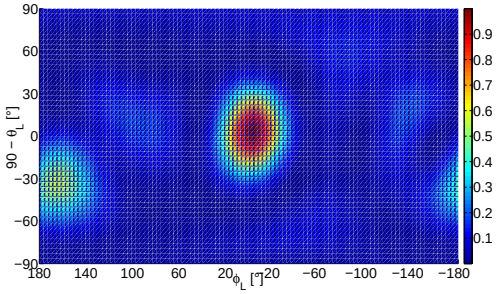

Figure 14: Spatial distribution of the array's output magnitude at 2500 Hz, obtained by the PWD beamforming with $N = 3$ and 16 microphone positions. The plane wave is still incident along the angular coordinates $(\theta_0, \phi_0) = (90°, 0°)$.

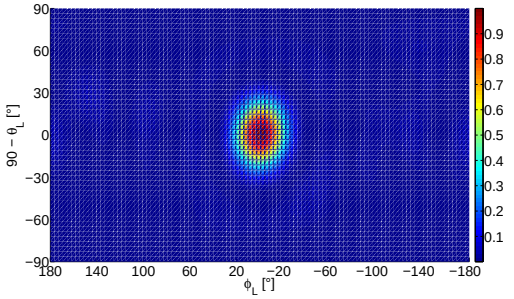

Figure 15: Same as figure 14 with $N = 4$ and 64 microphone positions.

compared with the frequency 1623 Hz.

- the same improvement is observed with MVDR beamforming, but what is more significant here is the centering and the thinning of the main lobe as a result of combining 64 microphone positions instead of 16 (this was already the case at 1623 Hz).

4.4 Summary

The improvements brought by the 64 optimized microphone positions, compared with the initial 16 Fliege's positions, are listed below:

- the PWD beamforming method gives improved source localization above 1 kHz (centering, symmetry of the main lobe and aliasing suppression). However, the inherent small errors in the calculation of the coefficients P_{nm} are greatly amplified at low frequencies, which lead to deteriorated results.
- the DAS method gives similar results with 16 and 64 microphone positions until 1623 Hz. At higher frequencies, the aliasing observed with 16 microphones disappears if 64 microphone positions are used.
- the MVDR method gives similar results with 16 and 64 microphone positions at low frequencies, but significant improvements have been observed at 1623 Hz and 2500 Hz (in particular, the thinning of the main lobe which leads to a very high spatial resolution in the detection of the source direction).

5 Testing the optimized 64 microphone positions with real sound fields

5.1 Measurement of the sound field

The objective of this part of the study is to check if similar improvements in the source localization are observed with 4x16 microphone positions and real sound fields. In order to meet this objective, the very simple sound field created by a unique source in an anechoic room has been tested. The experimental procedure is the following: the spherical microphone stand is fixed on a rotating table and positioned at 4m from a loudspeaker (see figure 16). At this distance, it is supposed that the wavefronts reaching the microphones can be locally approximated by plane waves. This assumption is not absolutely necessary, except if we want to compare with the results of theoretical plane waves. Finally, the centres of the sphere and the loudspeaker are at 1m50 from the floor, which is also absorbing (full anechoic room).

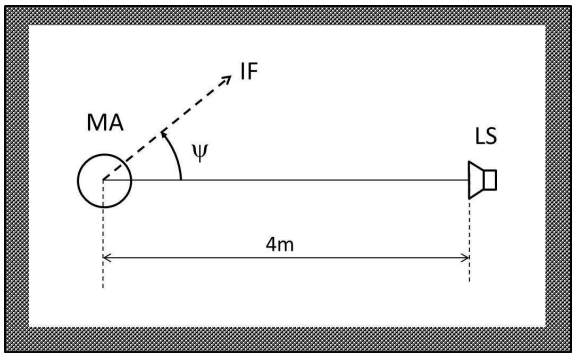

Figure 16: Plane view of the experimental setup in the anechoic chamber. The microphone array (MA) is situated at 4m from the loudspeaker (LS). ψ measures the azimuthal rotation of the sphere. (IF) is the initial front direction of the microphone array (when $\psi = 0$).

In this experiment, pure tones are emitted by the loudspeaker. Only the frequency 1623 Hz has been analysed in the following.

The rotating table is first positioned at a given azimuth angle ψ. Then, the loudspeaker is fed with the pure tone signal. After approximately ten to twenty seconds, the recording starts simultaneously on each microphone channels. The signals captured by the microphones are sampled at 48 kHz, digitized and stored as *.wav files (each file contains approximately 10 seconds of the signal).

Beamforming can be applied to 16 simultaneously recorded files (corresponding to any ψ angle) or several (usually four) measurements at different ψ angles can be combined as explained in previous sections. During post-processing, the sound files are read and

the samples are stored in 16 or 64 vectors $P(k, \theta_j, \phi_j)$. Then, the appropriate α_j are applied to obtain the spherical harmonics coefficients P_{nm}, following (3). Finally, the beamforming method (PWD, DAS or MVDR) is applied.

5.2 Results with 16 microphones

Figures 17 and 18 illustrate the results of the PWD method for an incident plane wave along the direction $(\theta_0, \phi_0) = (90°, 0°)$, the coefficients P_{nm} being calculated with 16 Fliege microphone positions, until respectively order $N = 3$ and order $N = 4$. These figures are very similar to figures 4 and 5 obtained with theoretical plane waves. For $N = 3$, the source position is identified at the maximum output value in $(\theta_L, \phi_L) = (90°, 0°)$ and the central lobe of sensitivity (at $-10dB$ from maximum) is approximately $\pm 30°$ wide. For $N = 4$, the PWD method leads to deteriorated results.

The DAS method (not shown) gives nearly correctly centred main lobes for $N = 3$ and $N = 4$ (maximum output value in $(\theta_L, \phi_L) = (93.6°, 0°)$). The central lobe of sensitivity (at $-10dB$ from maximum) is approximately $\pm 30°$ wide.

Finally, the MVDR method also gives very similar output results to those presented in figures 8 and 9, obtained with theoretic plane waves. The main lobe width is identical and the maximum output value is in $(\theta_L, \phi_L) = (93.6°, 0°)$.

5.3 Results with 64 microphones

Measurements with 16 microphone positions have therefore confirmed the results obtained with theoretical sound fields. Now, the next step is to analyse the results obtained by combining the measurements operated at several azimuths ψ. In this subsection, we will use the combination of pressure signals measured at the four optimized azimuth angles defined in (11) with $\psi_1 = 0$, in order to identify the sound source location at $(\theta_0, \phi_0) = (90°, 0°)$.

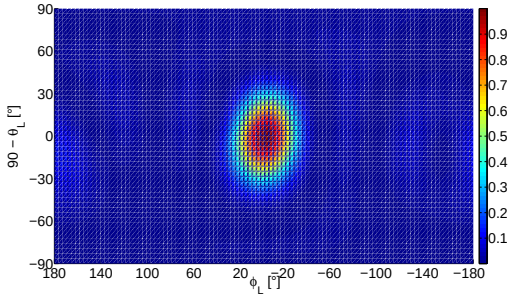

Figure 17: Spatial distribution of the array's output magnitude, obtained from measured acoustic pressures by the PWD beamforming with $N = 3$ and 16 microphone positions.

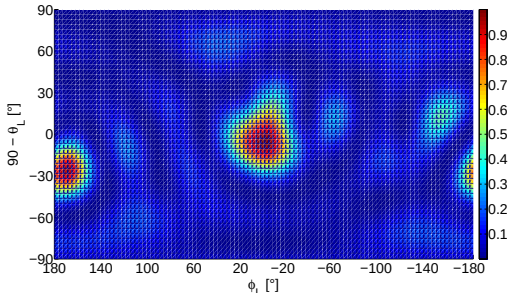

Figure 18: Same as figure 17 with $N = 4$.

5.3.1 Preprocessing operations

Even if the same signal is sent to the loudspeaker at each run, the amplitude and the phase of the waves reaching the spherical array are not exactly the same for two different azimuthal positions of the rotating table. This is because the two sets of 16 recordings are initiated at different times, at which the initial phase of the wave is not necessarily the same. Also, the wave amplitude can change, due to slightly different mechanical behaviours of the loudspeaker.

Preprocessing operations are necessary to unify the amplitudes and phases of the incident sound waves among different measurements. One solution would be to have a microphone recording the incident wave at a fixed position, but we did not provide such a microphone during our experiments. However, the microphone number 1 is situated approximately at the zenith of the spherical array, and therefore its position is 'nearly' fixed, i.e. its displacements are small compared to the wavelength. Figure 19 shows the uncorrected signals recorded at microphone nb.1 at the four optimized azimuthal positions of the rotating table: clearly, if the amplitude is approximately constant, the phase differences are significant and should be corrected. At 1623 Hz, the maximum distance between two positions of microphone 1 during rotation is $\lambda/6$, which introduces maximum phase differences of $\pi/3$. Some phase differences shown in figure 19 are significantly greater and cannot be simply explained by the displacements of microphone nb.1.

The method we have used to find the amplitude and phase corrections is therefore based on the amplitude and phase variations measured at microphone nb.1. This method is certainly not the best accurate. Fortunately, after having tested several methods, we have shown that the final result of the beamforming methods and source localization was not really sensitive to these corrections. The method we have used is the following:

- calculate the amplitude $A_1(\psi)$ by autocorrelation of the signal $s_1(\psi)$ recorded at microphone nb.1 for the azimuth ψ of the rotating table. The amplitude correction is $\frac{A_1(\psi_1)}{A_1(\psi)}$, with $\psi_1 = 0$;
- compute the cross-correlation between $s_1(\psi)$ and $s_1(\psi_1)$ for several delays and find the delay D_ψ that maximizes the cross-correlation;
- for all microphones (including nb.1) and all $\psi \neq \psi_1$, apply the corresponding amplitude correction and the delay D_ψ.

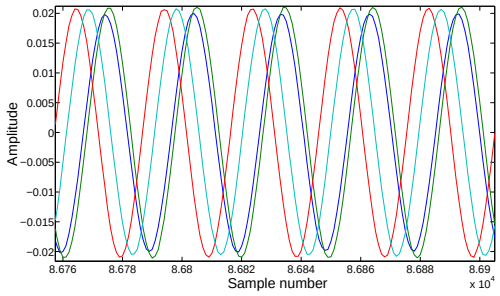

Figure 19: Amplitude (relative units) of four pressure measurements at microphone number 1 situated close to the north pole of the sphere ($\theta_1 = 0$).

5.3.2 Beamforming results

The application of the beamforming methods to the measurements at 64 optimized microphone positions, associated with the optimized values of the α_j, don't give reliable results, as shown in figures 20 and 21 for PWD and DAS ($N = 4$). Wrong source locations, spatial aliasing and artifacts can be seen in these figures and they are also present for the MVDR beamforming method (not shown).

In this case, contrary to the measurements at the 16 Fliege's positions, the results obtained for theoretic plane waves are not retrieved.

What may be the reasons of these discrepancies between theory and measurements ? In addition to inaccuracies related to the spatial discretization in the calculation of the P_{nm} in equation (3), measurements are affected by other errors, such as:

- the microphone are not exactly at the Fliege's positions on the sphere;
- the sphere is not perfectly rigid, therefore the measured pressures deviate from their theoretical values;
- instrumental errors are created by all the devices

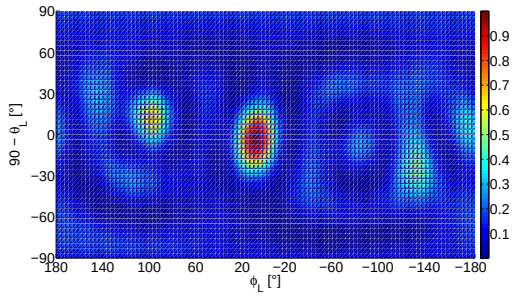

Figure 20: Spatial distribution of the array's output magnitude, obtained from measured acoustic pressures by the PWD beamforming with $N = 4$ and 64 optimized microphone positions.

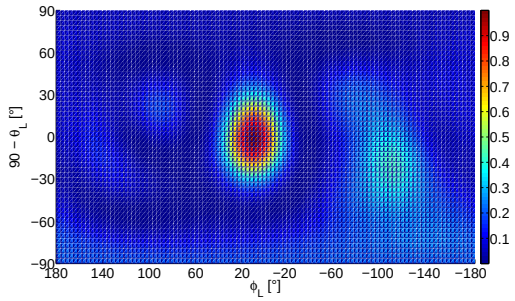

Figure 21: Same as figure 20 for the DAS method.

in the recording channels.

All these additional inaccuracies are acceptable with the measurements at 16 Fliege's positions, but their effect is amplified with the 64 optimized positions. Not only because the number of measurements increases, but also because of the greater absolute values of the α_j coefficients which amplify the errors. Indeed, a significant difference between the 16 versus 64 microphone positions is that the 64 optimum α_j's are no longer positive and approximately equal, as it is the case for the 16 Fliege's positions. As explained in section 3.2, the coefficients' values are comprised between -3 and $+3$, instead of approximately 0.8.

5.4 Results with 64 microphones and modified α_j coefficients

The idea to solve the problem of amplified errors in the calculation of P_{nm} from 64 measured acoustic pressures $P(k, \theta_j, \phi_j)$ is simply to extend the solution offered by the 16 Fliege's positions which was shown to be satisfying with measured results in section 5.2.

We have defined earlier $\epsilon_{n,n',m,m'}$ as:

$$\epsilon_{n,n',m,m'} = \sum_{j=1}^{J} \alpha_j Y_{n'}^{m'}(\theta_j, \phi_j) Y_n^{m^*}(\theta_j, \phi_j) - \delta_{n-n'}\delta_{m-m'} \tag{14}$$

A noticeable property of this 'error' is that its absolute value is not affected by a rotation ψ of the spherical array around its vertical axis, if we keep the same value of the coefficients α_j:

$$\epsilon^{\psi}_{n,n',m,m'} = \left(e^{i(m'-m)\psi}\right) \epsilon^{\psi=0}_{n,n',m,m'} \tag{15}$$

Therefore, if we combine the measurements at several azimuth angles ψ_p $(p = 1, N_\psi)$ and choose the same coefficients $\frac{\alpha_j}{N_\psi}$ at each rotation, then:

$$\epsilon^{N_\psi}_{n,n',m,m'} = \left(\frac{1}{N_\psi} \sum_{p=1}^{N_\psi} e^{i(m'-m)\psi_p}\right) \epsilon^{\psi=0}_{n,n',m,m'} \tag{16}$$

Equation (16) shows that the absolute value of $\epsilon^{N_\psi}_{n,n',m,m'}$ is less with N_ψ positions than with just one

m:	-4/4	-3/3	-2/2	-1/1	0
n=0	-	-	-	-	0.000
n=1	-	-	-	0.000	0.000
n=2	-	-	0.094	0.134	0.095
n=3	-	0.371	0.2380	0.296	0.383
n=4	0.371	0.238	0.191	0.296	0.383

Table 5: Maximum absolute errors in equation (4) for the 64 optimum microphone positions (11) and $\alpha_j = \alpha_j^{Fliege}/4$, if the sound field is reconstructed at order $N = 4$. As usual, the maximum errors are identical for corresponding positive and negative values of the index m

azimuthal position, if this particular choice of coefficients α_j is adopted. As a consequence, the cost function (8) is also decreased if we increase the number of azimuthal positions.

As the Fliege's α_j values are particularly interesting since they are approximately equal, positive and bounded in absolute value, we decide in the following to test the combination of four azimuthal positions with $\alpha_j = \frac{\alpha_j^{Fliege}}{4}$ at each rotation. As a first trial, we keep the four azimuthal positions defined in (11). We therefore obtain a cost function $C_{N=4} = 3.544$ which is comprised between the values presented in section 3.2 for two other choices of coefficients, i.e. 0.578 (64 optimized $\alpha_j, N = 4$) and 19.449 (16 Fliege's $\alpha_j, N = 4$).

We consider this result as a fairly good value for the cost function, especially since it has been obtained without optimization. Now, if the errors $\epsilon_{n,n',m,m'}$ are detailed in table 5 for a reconstruction of the sound field at the order 4, we can see that the P_{nm} are exactly reconstructed for $n \leq 1$, which is a characteristic of the Fliege's coefficients. The other maximum errors for $n > 1$ are greater than in table 4, as expected, but they are still acceptable, as will be seen in the following. As a comparison, the errors $|\epsilon_{n,n',m,m'}|$ can be as high as $0.38 (n = 2)$ and $0.78 (n > 2)$ for a reconstruction with the 16 Fliege's positions and coefficients at order $N = 4$.

The figures 22, 23 and 24 show the array's output magnitude for $N = 4$ and this new choice of

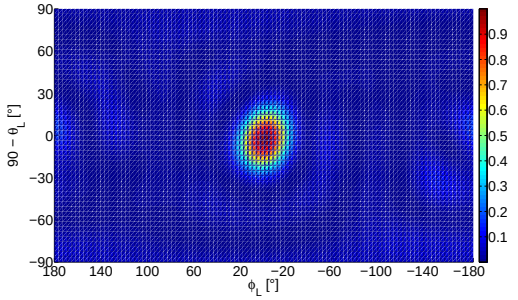

Figure 22: Spatial distribution of the array's output magnitude, obtained from measured acoustic pressures by the PWD beamforming with $N = 4$, 64 microphone positions and $\alpha_j = \alpha_j^{Fliege}/4$. The source incidence is still $(\theta_0, \phi_0) = (90°, 0°)$.

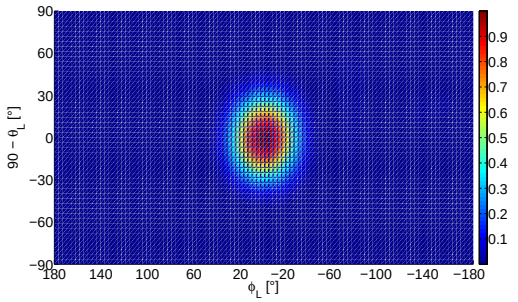

Figure 23: Same as figure 22 for the DAS method.

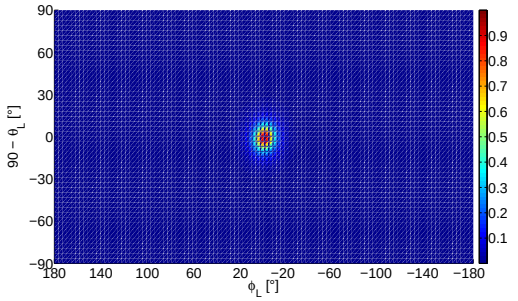

Figure 24: Same as figure 22 for the MVDR method.

Method	16 mic. ms.	64 mic. ms.	64 mic. th.
PWD	??	$\pm 23°$	$\pm 25°$
DAS	$\pm 30°$	$\pm 30°$	$\pm 30°$
MVDR	$\pm 16°$	$\pm 9°$	$\pm 3°$

Table 6: Central lobe width [°] at $-10dB$ around the maximum value in the distribution of array's output values. The sound field is reconstructed at order $N = 4$ with 16 or 64 microphones positions, from measured ('ms') or theoretical ('th') acoustic pressures at the microphones.

coefficients. The source direction is exactly identified for DAS and MVDR (for PWD, maximum in $(\theta_L, \phi_L) = (93.6°, 0°)$). The widths of the central lobes of sensitivity (at $-10dB$ from maximum) are given in table 6 and compared with previous methods. PWD and MVDR beamforming methods result in an improvement of the spatial resolution at $N = 4$, if 64 microphone positions are used instead of 16. This was predicted by theoretical results, unless the improvement brought by MVDR is less than expected. Furthermore, PWD with 64 microphone positions and $N = 4$ is better than with 16 microphones and $N = 3$. Finally, as predicted by theory, the DAS method doesn't increase the spatial resolution.

5.5 Further developments

The set of four azimuth angles (11) is not optimal with the constraint $\alpha_j = \alpha_j^{Fliege}/4$. We have found another combination of angles that leads to a cost function value $C_{N=4} = 2.732$ instead of 3.544. This combination has been tested, but it does not give significant improvements in the spatial resolution.

We have also tried to increase the number of azimuth angles. First, with $\alpha_j = \alpha_j^{Fliege}/6$. The combination of 6 azimuth angles has been arbitrarily set to:

$$[0°, 60°, 120°, 180°, 240°, 300°] \qquad (17)$$

No significant improvement has been obtained with these 96 microphone positions at order $N = 4$. If the 96 optimal α_j's are used (those which minimize the cost function (8) instead of the previous ones), there's no significant improvement at all, even if the cost function in this case is equal to zero, which means that the reconstruction of the sound field is exact at order $N = 4$.

Furthermore, if the combination (17) is used with 96 α_j values optimized until order $N = 5$, this does not provide better results.

Finally, we have tested intermediate α_j's between the optimized ones and the solution $\alpha_j = \alpha_j^{Fliege}/N_\psi$, with $N_\psi = 4$ or 6. Without success.

All these further developments show that the results presented in the previous section are probably the best we can obtain with measurements operated with 4 or 6 azimuth positions of the rotating table.

6 General conclusion

Increasing the number of microphones on the array generally provides a better spatial resolution of the antenna while detecting the location of sound sources, but at the expense of greater cost and complexity. In this paper, we have tested the solution of increasing the number of microphone positions by rotating the sphere around its vertical axis.

We have shown how to combine the sets of 16 pressure measurements to obtain an estimation of the spherical harmonics coefficients of the sound field. In particular, four optimized azimuthal positions of the rotating table have been defined for the 16 Fliege's microphone positions on the sphere, together with corresponding optimized α_j coefficients in equation (3).

The first series of tests consisted in applying three beamforming methods to theoretic pressures at the 64 microphone positions. We have shown that the quality of detection and spatial resolution in the source localization tasks is improved for the PWD beamforming method above 1 kHz, compared with 16 microphone positions. For PWD, results are deteriorated at low frequencies. The MVDR method gives similar results with 16 or 64 microphone positions at low frequencies, but significant improvements

have been observed at 1623 Hz and 2500 Hz. Finally, the DAS method provides improved results only at 2500 Hz.

During the second series of tests, the 64 optimized microphone positions were tested with real sound fields. Acoustic pressures generated by a loudspeaker have been measured in anechoic conditions. Beamforming results obtained with the 16 Fliege's positions correspond to theoretic sound fields' results. However, the improvements theoretically predicted for 64 microphone positions are only retrieved if other (non-optimized) sets of α_j coefficients are used.

Further developments will test this method on more complex sound fields, such as those existing in closed spaces. In particular, the room impulse responses could be analysed with these beamforming methods and the quality of early reflections' localization could be investigated by combining several azimuthal positions of the spherical array.

References

[1] J.J. Embrechts: Measurement of 3D room impulse responses with a spherical microphone array. Proceedings of Euronoise 2015 conference, Maastricht, The Netherlands.

[2] R. Thiele: Richtungsverteilung und zeitfolge der schallrückwürfe in räumen. Acustica **3 Suppl. 2** (1953) 291–302.

[3] S. Strom, H. Dahl, A. Krokstad, E. Eknes: Acoustical design of the Grieg Memorial Hall in Bergen. Applied acoustics **18** (1985) 127–142.

[4] M. Gerzon: Recording concert hall acoustics for posterity. J. Audio Eng. Soc. **23** (1975) 569–571.

[5] J. Merimaa, T. Lokki, T. Peltonen, M. Karjalainen: Measurement, analysis and visualization of directional room responses. AES 111th Convention paper 5449, Audio Engineering Society (2001).

[6] B.N. Gover, J.G. Ryan, M.R. Stinson: Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array. J. Acoust. Soc. Am. **116(4) Pt.1** (2004) 2138–2148.

[7] S.W. Clapp, A.E. Guthrie, J. Braasch, N. Xiang: Investigations of room acoustics with a spherical microphone array. AES 131st Convention paper 8459, Audio Engineering Society (2011).

[8] A. Farina, L. Tronchin: 3D sound characterization in theatres employing microphone arrays. Acta Acust united Ac **99** (2013) 118–125.

[9] B. Rafaely, I. Balmages, L. Eger: High-resolution plane-wave decomposition in an auditorium using a dual-radius scanning spherical microphone array. J. Acoust. Soc. Am. **122(5)** (2007) 2661–2668.

[10] J. Meyer, G. W. Elko: A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. ICASSP Proceedings II, IEEE (2002) 1781–1784.

[11] B. Rafaely, Y. Peled, M. Agmon, D. Khaydin, E. Fisher: Spherical microphone array beamforming. In Speech processing in modern communication, challenges and perspectives (I. Cohen, J. Benesty, S. Gannot, editors), chapter 11. Springer, Berlin, 2010.

[12] H. Sun, E. Mabande, K. Kowalczyk, W. Kellermann: Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing. J. Acoust. Soc. Am. **131** (2012) 2828–2840.

[13] M.Park, B. Rafaely: Sound-field analysis by plane-wave decomposition using spherical microphone array. J. Acoust. Soc. Am. **118(5)** (2005) 3094–3103.

[14] B. rafaely: Analysis and design of spherical microphone arrays. IEEE Trans. Speech Audio Process. **13** (2005) 135–143.