Identifying direction of reflections with a Soundfield microphone.

POULALION Léo-Polde

February - May 2022

NTNU
Norwegian University of
Science and Technology

UNIVERSITÉ
TOULOUSE III
PAUL SABATIER
Université de Toulouse

Contents

1 Introduction

Identify the position of a sound source is something that may seems common, in fact, our ears can do it pretty well. The goal is to do automatic detection of reflection directions using microphones. So that we will develop in this report how to identify direction of reflections using an array of microphone. First, we will define what is a so called "soundfield microphone", the principle of such a microphone consist in four cardioid microphones placed at the top of each corner of an equilateral tetrahedron, their outputs signals are called the A-format signals. The size of this regular tetrahedron as to be as small as possible so that the distance between each microphone is negligible compare to the distance with the sound source and small compared to the wavelength of the sound. Then, the four signals are combined to simulate a microphone at the center of the tetrahedron with a variable directivity that is changing with the coefficients of the linear combination between the four microphones. These simulated microphones constitute the "B-format" of a soundfield microphone. The most commons directivity functions are the omnidirectionnal directivity and the bidirectionnals function along X,Y and Z axis.These virtuals microphones are mathematical objects, the main directivities are the first orders of atomic orbitals. The omnidirectional microphone, equivalent of an S-orbital, and the "figure-of-eight" microphones along X,Y and Z axis which are the P-orbitals. The Soundfield microphone is often use in musical recordings so that we can, in post-prod, look which directivity sounds the best.
In this project, we will use three omnidirectional microphones placed at the top of an horizontal equilateral triangle. This system is able to locate sources situated on the plan of the triangle, but we will need a fourth microphone to identify sources in the 3-D space.
The purpose of this study is to characterize the efficiency of the soundfield microphone approach to localize a sound source in 2-D and 3-D space.

2 Preconditions

2.1 Impulse response

• *Definition :* The impulse response of a system describes the way this system reacts to a Dirac pulse excitation.
In our case the system is a room, that we will consider linear and causal.

By sending a Dirac pulse through a loudspeaker and measuring how the sound propagates indoor, the impulse response of the room captured by the sensor gives several information on the nature of the room.
But the thing is that a Dirac pulse generated numerically is instantaneous and last only one sample, even if it's possible to generate such a signal, it contains very little information. The power of the signal is low and we have a bad signal to noise ratio. Of course noise is a plague and the signal to noise ratio must be maximised.
So we have to find another way to measure the impulse response. Let's make a quick reminder on impulse:

$$h(n) = DFT^{-1}[H(w_k)] \quad H(w_k) \times DFT(x(n)) = DFT(y(n))$$
$$\iff DFT^{-1}[H(w_k)] \circledast x(n) = h(n) \circledast x(n) = y(n)^{1}$$

Where $h(n)$ is the impulse response of the system and $H(w_k)$ is the transfer function. $x(n)$ and $y(n)$ are respectively the input and output signals. These formulas are true for every $x(n)$ and $y(n)$, in particular, if $x(n) = \delta(n)$ then the output signal is the impulse response.

$$h(n) \circledast \delta(n) = y(n) \iff h(n) = y(n)$$

Using the inverse of the input signal defined such as

$$x(n) \circledast x(n)^{-1} = \delta(n) \tag{1}$$

As long as we can find an inverse to the input signal (usually using Z-transform), we are able to compute the impulse response using the output signal $y(n)$ and the inverted input signal $x(n)^{-1}$. Such that:

$$h(n) \circledast x(n) = y(n) \iff h(n) \circledast x(n) \circledast x(n)^{-1} = y(n) \circledast x(n)^{-1}$$
$$\iff h(n) \circledast \delta(n) = y(n) \circledast x(n)^{-1} = h(n) \tag{2}$$

The fact is that the invert signal does not always exist, or could be very difficult to calculate.

2.2 Sinusoidal sweep signal

In practice we will use as input signal a so called, "chirp signal" or "sweep".

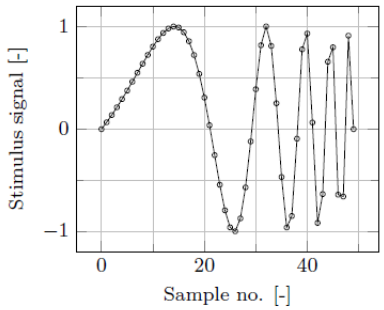

Figure 1 : Time representation of an exponential sine sweep of 50 samples length, the sweep is a sinus whom frequency increase exponentially with time, it contains all frequencies in a chosen bandwidth and we can adjust the time it lasts

We know exactly the expression of the invert of such a signal, which permits using equation (2).

The mathematical expression for a chirp signal which sweeps from f_1 to f_2 in T second is :

$$x_{chirp}(t) = \sin\left[\frac{2\pi f_1 T}{R}\left(e^{\frac{Rt}{Tf_s}} - 1\right)\right] \tag{3}$$

[1] The $\circledast$ symbol is a regular convolution product.

Where $R = \ln \frac{f_2}{f_1}$

The good point with this type of signal is that the impulse response shows a much better signal to noise ratio, and this value becomes better and better with the increasing time T.
Using this type of signal is nowadays a norm in measuring impulse responses, they give good results with few noise, are pretty easy to invert, and have an adjustable bandwidth.
Using a chirp signal is strictly equivalent to use a Dirac pulse, but with a better signal to noise ratio, so we might use the term "Dirac pulse" to talk about the input signal all along this report by language abuse.

2.3 Sound propagation and reflections

The speed of sound in the air is quite sensitive to the temperature : 343,5 m/s at 20°C and 331,5 m/s at 0°C. But it also depends on the air pressure and the humidity. To simplify we will use an approximation of the sound velocity depending only on the temperature.

$$v_s(\Theta) = 331,5 + 0,6 \times \Theta$$

Where Θ is the temperature in Celsius.
When the pulse is emitted by the loudspeaker indoor, the waves propagates in the room and hit the different surfaces such as walls, ground and ceiling, the sound bounce on these surfaces following a sort of Descartes law. Each reflection of the incident wave can be considered as a new source which generates a new sound wave. Supposing that the initial signal is a Dirac pulse, we will see that the microphones capture different pulses which correspond to the different reflections. The first wavefront to reach the microphone is the direct one, it doesn't meet any wall so it's the one that takes the shortest way. Obviously, the second pulse to reach the microphone is the one which takes the second shortest way, the third takes the third shortest way etc.

Of course the material on which the sound bounces will impact the reflection. Absorbing materials are make in the objective of reflecting only a few part of the energy carried by the sound wave, and absorb the majority. On the contrary, hard surfaces will behave like mirrors for light waves, reflecting almost all of the energy of the wave front, limiting the loss. In our experiment, we are interested in reflections and localizing them. Thus it seems logical to use walls in reflective components like concrete to keep powerful pulses even with several bounces in a row. Otherwise the reflections would cause a loss of power and then a worse signal to noise ratio.

Each time the sound wave bounce against a surface, the impact point can be considered as a new sound source emitting a pulse. So it might be useful to analyze each one of them separately. The impulse response measured by the microphone can be cut in slices, one pulse per slice, and then the datas are analysed for each of those windows.

3 Principle of the method

In this chapter, we will describe how the method works, on which points it's based, and define all the theoretical aspects in it.
Before digging in more details, we will explain roughly what is involved in this method. The purpose of this project is to make an algorithm that returns the direction of coming wave fronts. Three omnidirectional microphones placed at the angles of an equilateral triangle, record the same sound source from three different positions. By doing a combination of the Fourier transforms of these three signals, we are able to model virtuals microphones with different directivities, behaving like if the source was recorded by a chosen directivity microphone placed at the center of the equilateral triangle. These virtual microphones constitute the B-format of a Soundfield microphone.
Then an other combination is made, this time between the signals recorded by the B-format microphones to construct a cardioid virtual microphone that can be virtually rotated in space.
Using this cardioid microphone to probe the space, we can analyse the power coming from each direction, the angle that presents the biggest average power describe the case where the virtual microphone is facing the sound source, then we can deduce the position of the source. Alternatively, it could be easier to detect the minimum of intensity, in this case the microphone is pointing 180° away from the source.

3.1 Equipment used

For measuring signals, we need two fundamental instruments, a loudspeaker and a microphone.
For the microphone, we will use an omnidirectional microphone positioned at the angles of an equilateral triangle. An omnidirectional microphone has the same sensitiivity in all directions, whatever the incident angle. The objective here is to identify the position of the source and each one of the reflections, to start we will suppose that the array of microphone and the loudspeaker are both in the same plan Oxy.

The position of the microphones had to be very precise, they are disposed on a circular tool which is remote controlled by a matlab script. We first place the microphone at a small distance from the turntable rotation center. Then we place the sensor at 0°, 120° and 240° to form the perfect equilateral triangle.

Then a excitation signal (Dirac pulse) is sent through the loudspeaker and recorded by the microphones, this is the impulse response of the room.

3.2 Windowing the pulses

As quickly mentioned in the precedent chapter, indoor measurements has to deal with reflection and reverberation.

That means, the microphones record sound coming from every direction, that's why we will focus only on the 20 first millisecond of the impulse response of the room.

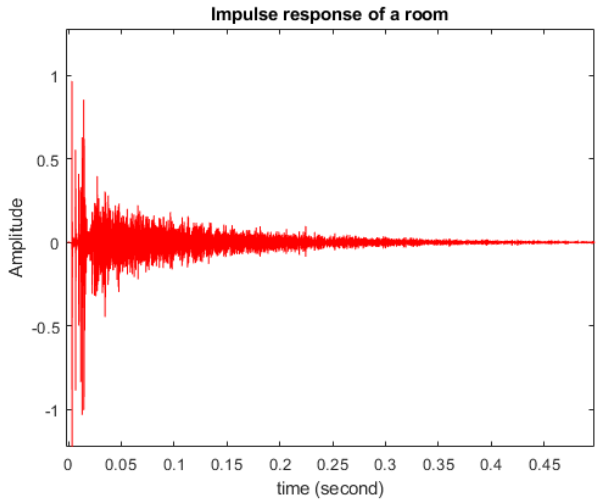

Figure 2 : Impulse response of the room within the measurements will have place. The reverberation time equals more or less 0,5 seconds

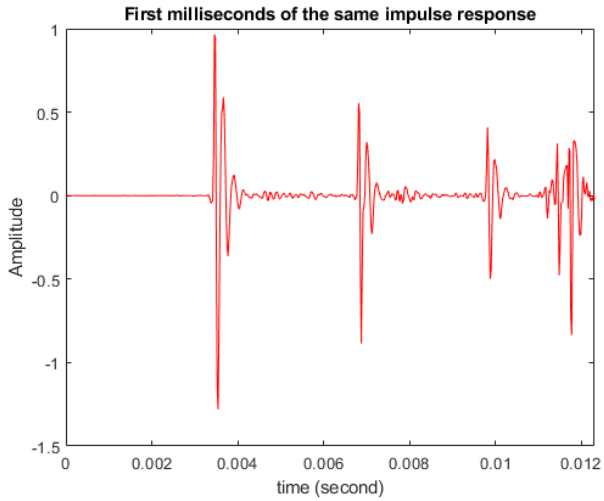

Figure 2 bis : Only the first milliseconds of the impulse response, each pulse correspond to a wave front hitting the microphone.

As said earlier, each pulse can be considered as if they were generated by different audio source.
As you can see on the previous graphs, at the beginning of the IR, it's easy to separate pulses and then analyse from which direction they are coming individually. In the objective of separating the pulses, we will cut the IR in slices using time windows, after that, we will be able to identify the originating direction of each wave front, and then conclude on the position of the sources (the real one and the reflective surfaces).

3.3 The B-format microphones

3.3.1 Preambulum (extracts of a paper written by Peter Svensson)

The particularity of the cardioïd is that it possesses a very profound minimum of sensibility depending on the incident angle of the wave. Once we have built such a microphone, we are able to find the angle for which the power is maximum. When the cardioïd is perfectly aligned with the source, we have rushed the maximum of intensity, that means we have found the angular position of the source.

Before making the variable cardioïd microphone, we have to build 3 virtual microphones with different directivities that will be the base ingredients of our cardioïd mic. These virtual microphones composed the B-format of the Soundfield microphone. The first one is an omnidirectional microphone virtually positioned at the center of the circle the real ones are on (remember the real microphones also have an omnidirectional directivity). The second and the third one are figure of eight along X and Y axis.
What are the linear combination that leads to such microphones ?

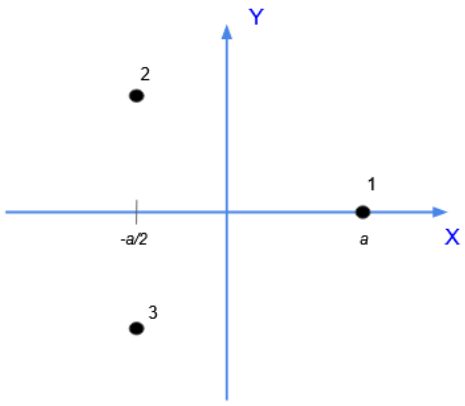

*Figure 3 :*three microphones M_1, M_2, M_3, placed uniformly around a circle of radius a, centered at the origin.

The coordinates of the microphones in a Cartesian system are:

$$\begin{cases} (x_1, y_1) = a(1, 0) \\ (x_2, y_2) = a\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right) \\ (x_3, y_3) = a\left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right) \end{cases}$$

A plane wave is propagating, with an incidence angle θ. Then the sound pressure in this plane wave is described as[2]

$$p(x, y) = p_0 e^{jk\cos\theta \cdot x} e^{jk\sin\theta \cdot y}$$

So, we can find out what the sound pressure is at the three microphones:

$$p_1 = p_0 e^{jka\cos\theta}$$

[2]We assume that the presence of the microphones does not affect the sound pressure. This is true if the microphones are much smaller than the wavelength.

$$p_2 = p_0 e^{-jk\frac{a}{2}\cos\theta} e^{jk\frac{\sqrt{3}a}{2}\sin\theta}$$

$$p_3 = p_0 e^{-jk\frac{a}{2}\cos\theta} e^{-jk\frac{\sqrt{3}a}{2}\sin\theta}$$

We plan to put the microphones so close that $ka << 1$ and then we can express the exponential functions with series expansions[3]:

$$p_1 \approx p_0 + p_0 \cdot jka\cos\theta$$

$$p_2 \approx p_0(1 - jk\frac{a}{2}\cos\theta)(1 + jk\frac{\sqrt{3}}{2}a\sin\theta)$$

$$p_2 \approx p_0(1 - jk\frac{a}{2}\cos\theta + jk\frac{\sqrt{3}}{2}a + (ka)^2\frac{\sqrt{3}}{4}\sin\theta\cos\theta)$$

The order 2 term is negligible

$$p_2 \approx p_0 - p_0 \cdot jk\frac{a}{2}\cos\theta + p_0 \cdot jk\frac{\sqrt{3}a}{2}\sin\theta$$

$$p_3 \approx p_0 - p_0 \cdot jk\frac{a}{2}\cos\theta - p_0 \cdot jk\frac{\sqrt{3}a}{2}\sin\theta$$

We can create a *virtual microphone* by summing the signals of the three microphones with weight factors (that might depend on frequency ω (that is, on the wavenumber, k). Thus we create an output signal s which has a desired directivity function $D(\theta)$:

$$s = w_1 p_1 + w_2 p_2 + w_3 p_3 = p_0 \cdot D(\theta) \qquad (4)$$

3.3.2 The omnidirectional microphone

It seems logical to simply take the mean of the 3 signals to build an omnidirectional microphone. Let's see if our flair is correct. For this microphone, the directivity function is $D(\theta) = 1$, it has the same sensitivity everywhere without depending on θ.

So $$s_W = p_0.D(\theta) = p_0$$

According to equation (4), We have to find w_1, w_2 et w_3 such as:

$$s_W = w_1 p_1 + w_2 p_2 + w_3 p_3 = p_0$$

That's to say

$$w_{W1}(1 + jka\cos\theta) + w_{W2}(1 - jk\frac{a}{2}\cos\theta + jk\frac{\sqrt{3}}{2}a\sin\theta) + w_{W3}(1 - jk\frac{a}{2}\cos\theta - jk\frac{\sqrt{3}}{2}a\sin\theta) = 1$$

[3] $e^x \approx 1 + x$

Identifying sine and cosine, comes out this system:

$$\begin{cases} \left(w_{W1}jka - w_{W2}jk\frac{a}{2} - w_{W3}jk\frac{a}{2}\right)\cos\theta = 0 & \forall\theta \in \mathbb{R} \\ \left(w_{W2}jk\frac{\sqrt{3}}{2}a - w_{W3}jk\frac{\sqrt{3}}{2}a\right)\sin\theta = 0 & \forall\theta \in \mathbb{R} \\ w_{W1} + w_{W2} + w_{W3} = 1 \end{cases}$$

$$\iff \begin{cases} w_{W1} - \frac{w_{W2}}{2} - \frac{w_{W2}}{2} = 0 \\ w_{W2} - w_{W3} = 0 \\ w_{W1} + w_{W2} + w_{W3} = 1 \end{cases}$$

And so the weights factors for the omnidirectional microphone are $w_{W1} = w_{W2} = w_{W3} = \frac{1}{3}$.

3.3.3 The figure of eight microphones

In our coordinate system, the angle $\theta = 0$ correspond to a point belonging to the positive part of the X axis.

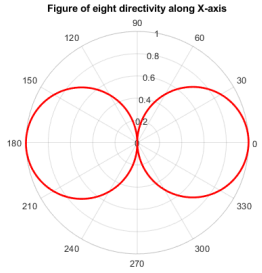

Figure 4

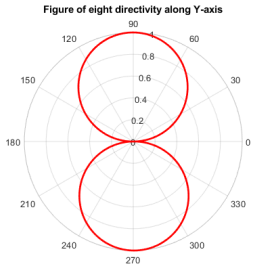

Figure 5

• Bidirectional directivity along the X-axis (figure 4)

For such a microphone the directivity function depending on theta is $D_X(\theta) = \cos\theta$.
The calculation leads to :

$$w_{X1} = -\frac{2j}{3ka}; \qquad w_{X2} = w_{X3} = \frac{j}{3ka}$$

Please refer to the annexes for the detailed calculation.

•Bidirectional directivity along the Y-axis (figure 5)

Along the Y-axis, the directivity function is now $D_Y(\theta) = \cos(\theta + \frac{\pi}{2}) = \sin\theta$ And the weight factors for each real microphones are :

$$w_{Y1} = 0; \qquad w_{Y2} = \frac{-j}{\sqrt{3}ka}; \qquad w_{Y3} = \frac{j}{\sqrt{3}ka} = -w_{Y2}$$

Please refer to the annexes for the detailed calculation.
It's important to notice that for the bidirectional directivity, the weight factors depends on k that's to say depends on

the frequency. It means the combination of the signals have to be done in frequency domain, we will then use frequency response which is the Fourier transform of the impulse response.

3.4 Building a rotating cardioïd microphone

As said in the beginning of this chapter, the objective is to reproduce several times the audio pulse, and moving a bit the direction of the virtual microphone to spot for which angle the power received is maximum. Once this is done, we know that the pit in cardioïd directivity function faces the sources.

As before, we have to make linear combination between the omni-microphone and the two figures in eight to build the cardioid. Remember that the directivity functions of the B-format microphones are :

$$D_W(\theta) = 1 \qquad D_X(\theta) = \cos\theta \qquad D_Y(\theta) = \sin\theta$$

The directivity function for a cardioïd that has maximum sensibility in $\theta = 0$ is :

$$D_{card}(\theta) = \frac{1}{2}(1 + \cos\theta)$$

And so the weight factors are;

$$w_{Wcard} = w_{Xcard} = \frac{1}{2} \qquad and \qquad w_{Ycard} = 0$$

$$\implies D_{card} = w_{Wcard}.D_W(\theta) + w_{Xcard}.D_X(\theta) + w_{Ycard}.D_Y(\theta)$$

Which is logical if we think of it graphically.
Now for a cardioïd that has maximum sensibilty in $\theta = \theta_{aim}$, the directivity function can be expressed as follow :

$$D_{varcard}(\theta) = \frac{1}{2}(1 + \cos(\theta - \theta_{aim})) \tag{5}$$

However; $\cos(\theta - (\theta_{aim})) = \cos\theta\cos(\theta_{aim}) + \sin\theta\sin(\theta_{aim})$

$$\text{Then, } D_{vc}(\theta) = \frac{1}{2}\left(1 + \cos\theta\cos(\theta_{aim}) + \sin\theta\sin(\theta_{aim})\right)$$

$$\implies D_{vc}(\theta) = \frac{1}{2} + \frac{1}{2}\cos(\theta_{aim})\cos\theta + \frac{1}{2}\sin(\theta_{aim})\sin\theta$$

$$\implies D_{vc}(\theta) = \frac{1}{2}D_W(\theta) + \frac{1}{2}\cos(\theta_{aim})D_X(\theta) + \frac{1}{2}\sin(\theta_{aim})D_Y(\theta))$$

Identifying with the equation (6) right below

$$D_{vc}(\theta) = w_{Wvc}D_W(\theta) + w_{Xvc}D_X(\theta) + w_{Yvc}D_Y(\theta) \tag{6}$$

We conclude :

$$w_{Wvc} = \frac{1}{2} \qquad w_{Xvc} = \frac{1}{2}\cos\theta_{aim} \qquad w_{Yvc} = \frac{1}{2}\sin\theta_{aim} \qquad \overline{P}_{varcard}(\theta_{aim}) = A^2\frac{1}{2}(1 + \cos(\theta - \theta_{aim}))^2$$ [5]

3.5 Finding the source probing the power coming from each direction

Now that the cardioïd microphone is built, we will measure the power received by the microphone for each values of θ_{aim} in $[0; 2\pi[$.
The power contained in a signal is the square of his root mean square value (V^2_{RMS} expressed in Volts).
Let $x(n)$ be the time representation of our discrete signal, where $n \in \mathbb{N}$ is the index of sample.
Then the V_{RMS} of $x(n)$ is defined as :

$$V_{RMS}(x(n)) = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x(n)^2} \qquad V_{RMS}(x(t)) = \sqrt{\int_{-\infty}^{+\infty} x(t)^2}$$

Upwards, both expressions for discrete and continuous signals.

We are interested in the frequency representation because remember the weight factors depends on the frequency.
In a perfect world, we should compute the average power over all frequencies which is proportional to the square root of the spectral density of power ($S(f)$) of the signal. However because we did the serial expansion of the exponential function as mentioned in footnote (3), we have to limit our study to relatively small frequencies.
The principle of the algorithm is a for loop with the criterion "for θ_{aim} in $\vec{\Theta}$" where $\vec{\Theta}$ is a vector that contains discretes values between 0 and 2π radians with a chosen step value. Such a vector is built using the matlab function linspace [4]. The value of θ_{aim} traduce the main direction of the cardioïd.
For each iteration of the loop the programm returns the power received by the microphone in an array. Plotting the average power depending on θ_{aim}, a curve like figure (6) appears, we decide to call this function $\overline{P}(\theta_{aim})$
Then, a curve fitting is done to find the angle from which the wave sound is coming.
The function $\overline{P}(\theta_{aim})$ can be approximated, it can be done using for example Lagrange interpolation (with polynomials), or Fourier series. The last is the one method we chose, the Fourier series present a great advantage : making the development until only second order is enough to reproduce perfectly the curve. Indeed, as said a bit earlier, $\overline{P}(\theta_{aim})$ represent the amplitude of the signal over frequency multiplied by the directivity function:

[4] $\vec{\Theta} = linspace(0, 359, 360)$ create an array from 0 to 359 with 360 values, in this case it corresponds to every degrees : $\vec{\Theta} = [0, 1, 2, ..., 358, 359]$
[5] The formula shows a $\cos(x)^2$ which can be linearised in $\alpha\cos(2x) + \beta$ using linearisation formulas. So second order fourier series, that is to say with $\cos(2\theta)$ and $\sin(2\theta)$ terms is enough. The detail of the calculus is described in annexes

The general expression for a fourier decomposition until second order is :

$$F(\theta) = a_0 + a_1 \cos\theta + b_1 \sin\theta + a_2 \cos(2\theta) + b_2 \sin(2\theta)$$

Where a_0, a_1, a_2, b_1 and $b_2 \in \mathbb{R}$ are called the coefficients of the decomposition.
The detailed calculus in annexes permits to find out :

$$a_0 = \frac{3A^2}{4} \qquad a_1 = A^2 \cos(\theta_{aim}) \qquad b_1 = A^2 \sin(\theta_{aim})$$

$$a_2 = \frac{A^2}{4} \cos(2\theta_{aim}) \qquad b_2 = \frac{A^2}{2} \sin(2\theta_{aim})$$

The value of θ_{aim} can now be deduced :

$$\frac{\sin(\theta_{aim})}{\cos(\theta_{aim})} = \frac{b_1}{a_1} = \tan(\theta_{aim}) \implies \theta_{aim} = \tan^{-1}\left(\frac{b_1}{a_1}\right)$$

The Matlab add-on "curve fitting tool" allows us to return the values of the coefficients of the Fourier decomposition in an array, in particular the values for a_1 and b_1. And then print the value of θ_{aim} using the native function atan2.[6]

3.5.1 Building a variable bidirectional microphone

The directivity function for a figure in eight microphone is written as follow :

$$\text{along X-axis:} \qquad D_{X\,figin8}(\theta) = \cos\theta$$

$$\text{along Y-axis:} \qquad D_{Y\,figin8}(\theta) = \sin\theta$$

The derectivity function of a bidirectional microphone with a maximum sensitivity in the direction of θ_{aim} is :

$$D_{figin8} = \cos(\theta - \theta_{aim}) \tag{7}$$

4 Simulations

4.1 Interest of simulations

Before doing any measurements, using simulated data permit to develop a program and check is proper functioning before putting the blame on setup measurements. Moreover, using different types of data such as impulses responses, frequency response, spherical wave fronts or plane waves is a good way to spot any failure of the algorithm and permit to identify where it could be improved.

4.2 Using digital pulses

A pulse generated numerically has the advantage of being really sharp in the sens that it contains only one or two samples different from zero, a more realistic representation of a Dirac pulse is something with a oscillating tail. In the first case it is pretty easy to find a window that doesn't cut a part of the pulse, whereas in the second, it can be difficult to find the end of the pulse, making the choice of the best window much more difficult.

By taking the Fourier transform of these digital pulses, the frequency response of each A-format microphone is perfectly flat all over the frequency range.

For these data sets: digital pulses and perfect waves fronts, the algorithm seems to work really well, even adding quite strong noise, we still have a 0.1 degree precision overall. This is encouraging for the following even if this simulations are not the more realistic ones.

4.3 Average power coming from each direction

As said earlier, the principle of this approach is to probe the space with a virtual cardioïd microphone. The cardioïd directivity has a sensitivity that varies depending on the orientation of the sound wave. Here is two representations of the $\overline{P}(\theta_{aim})$ function. Using the fourier decomposition, the goal is to have the analytical expression of this curve.

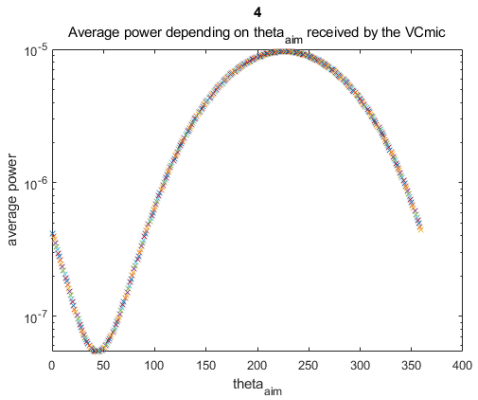

Figure 6 : Average power coming from each direction (logarithmic scale along y axis), the plot shows a unique minimum

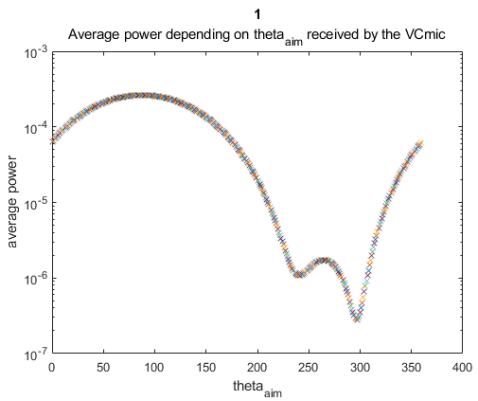

Figure 7 : Same plot as previous figure but for a different wavefront, the curve now possesses two minimums instead of one.

[6] atan2(Y,X) returns the value of the arcus tangent of the argument of the complex number X + iY, so that we are sure the programm compute the good sign

The figure (6) is characteristic from a cardioïd microphone that has been rotated. The minimum and the maximum are separated by 180° which is coherent with our model.
However, the figure (7) seems a bit strange, the sounding has been done by the same cardioïd virtual microphone, but this time, the curve presents 2 minima and 2 maxima. Both maxima, the big and the small one, are separated by 180°.
In conclusion this curve seems like it was computed using an *supercardioïd microphone*, which directivity function is showed below.

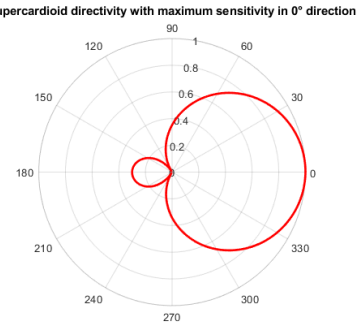

Figure 8 : Supercardioïd directivity, it's equation is $D_{supcard}(\theta) = 0,35 + 0,65cos(\theta)$

The polar equation of such a function can be thought like a flawed cardioïd.

$$D_{supercard} = abs(0,5 - \delta + (0,5 + \delta)\cos(\theta)) \qquad (8)$$

With $\delta \in [0; 0,5]$.
Therefore, this shows that our cardioid microphone isn't as perfect as it might be.

4.4 Wavefronts are not planwaves

In our experiment, the wavefront are considered as spherical wavefront, indeed the loudspeaker is very small and so the waves it generates can be considered as spherical wavefronts. Moreover, the distances between the three microphones is negligible before the distance with the source, then the spherical waves can theoretically be considered as plan waves propagation.
However, in the simulations, the wave received by the microphones were not so flat, in particular, the furthest microphone receive less sound energy than the closest. And this difference even if it's no big deal, deform the virtual microphones because is like if some microphones has a bigger weight than others. The linear combination between the three microphones were calculated in the hypothesis that each signal had the same average power.
A solution to counter this problem is to correct the amplitudes received by the microphones so that the average amplitude is barely the same for all three microphones.

5 Measurements

5.1 Equipment description

• Microphone :
This microphone is an omni microphone using a 48V power amplification. Reference : NTi Audio M4261
• Loudspeaker :
The loudspeaker used is a SEAS 1-inch tweeter, model H615, it is small enough to generate almost perfect spherical wave fronts.
• Turntable : The turntable is a NORSONIC NOR265. It is controlled through a Matlab script and permit the rotation of a bar on which can be attached a microphone or a loudspeaker.
• characteristics of the room :
The room is almost empty with concrete walls and floor that reflects a lot sound waves. That implies a long reverberation time, around 6 seconds. It is interesting for the measurements to have very reflective walls so that we don't lose too much power at each reflection and it's better for the results. The tricky thing is that we are able to analyse only the primary reflections, indeed it becomes diffuse sound quite fast and the microphone captures lots of wave fronts we are not interested in.
• Soundcard :
The soundcard play the role of DAC and ADC converter. For all the measurements, we will use a sampling frequency $f_s = 48000Hz$.

The data sheets of the devices can be found on the bibliography.

5.2 Setup

Here is a representation of the position of the sound source, the microphone and the walls.

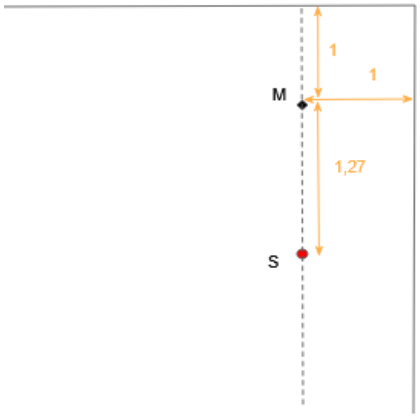

Figure 9 : Setup used for the measurements, distances are in meters, S is the loudspeaker and M the microphone.

The microphone needs to be far enough the microphone so that the sound waves hitting the microphone can be considered as plane waves. The loudspeaker and microphones are

placed at a 2.1 meters height. We placed them the furthest from the ground to delay as much as possible the primary reflection coming from the floor. We recall that we are interested only in waves propagating in the horizontal plan, that's to say in reflections that involves only the walls. In particular, reflections coming from the roof and from the floor are something to avoid. Some absorbing material was put on the ground to limit reflection.
The A-format microphones are placed on the peaks of a triangle inscribed in a circle of radius $a = 0.072m$.

5.3 Calibrate the equipment

• Microphone and loudspeaker
It is important to well adjust the gain of the loudspeaker and the microphone to have a good signal to noise ratio. Indeed if the loudspeaker is not loud enough or if the microphone possesses not enough gain, we will have a lot of noise in our measured signals. On the other hand, if the gain of the loudspeaker or the microphone is too high, we could have some saturation in the signals, or even worse, equipment may be damaged !
• Input signal
As explained in the preconditions chapter, a exponential sweep is used for these measurements. The signal starts at 250 Hz until 24 kHz which is $\frac{f_s}{2}$, and lasts around 7 seconds.
• Turntable
To be used properly, the turntable has to be calibrated, we have to tell to the turntable which angle correspond to our zero.

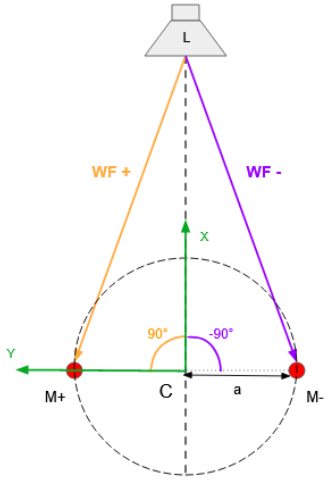

Figure 10 : Method used to find the position of the 0° position in our own coordinates system. M+ and M- represents the two positions of the microphones separated by 180° degrees

The goal is here to find a couple of points M+ and M- separated by an angle of 180° in polar coordinates, such as the time of arrival of the direct sound is the same for both microphones M+ and M-. Then we know the position of our zero degree (our x-axis)

5.4 Theoretical values

Knowing the distances between each wall, the loudspeaker and the microphone, theoretical values such as angles and distances of the "phantom sources" can be deduced using geometry. We can compare this values with the ones estimated by the algorithm to conclude on the efficiency of this approach.

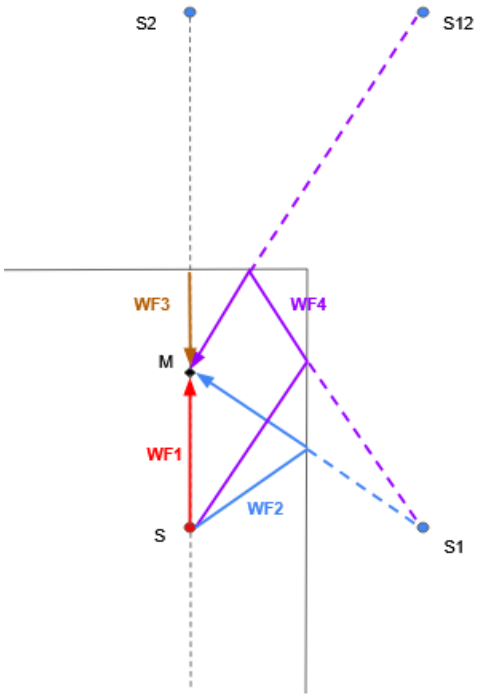

Figure 11 : Schema illustrating the phantom source method used to find the real values of the angles

The theoretical values are:

$$d_1 = 1,27m, \quad d_2 = 2,37m, \quad d_3 = 3,27m, \quad d_4 = 3,84m$$
$$\theta_1 = 0°, \quad \theta_2 = 57.6°, \quad \theta_3 = 180°, \quad \theta_4 = 148.6°$$

The index numbers represents the wave front number.
With an height of 2.1 m, the wave reflection coming from the floor travels a distance of 4.39 m. Which is right after the fourth wave front (about 75 samples time difference).

6 Results

6.1 Impulse responses

Here is the time representation of the firsts samples of the impulse responses captured by the 3 microphones. It seems to have a good signal to noise ratio, (almost 100/1).

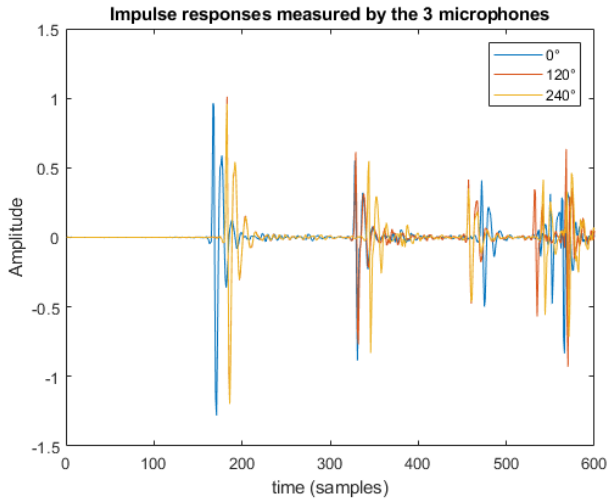

Figure 12 : 600 First samples of the impulse response captured by 3 microphones placed on a circle of radius a=0.072 cm

Each, pulse is coming from a different direction, the position of the walls, and of the measuring tools are thought such as each pulse is delayed enough from each other to permit a proper analysis. Indeed, as shown in figure (12), each pulse is separated from the previous by at least 100 samples (2ms). The last wave front is a bit overlapped - at sample 560 - by the pulse coming from the floor reflection.

6.1.1 The choice of the window

As said earlier, each pulse reaching the microphone can be considered as coming from a new virtual source. To identify the propagation direction of each wavefronts, we cut the impulse response in different slices in order to isolate each pulse. These slices are what we call windows. Despite the work that has been done on the measurements techniques, the distance between each microphones, and all our goodwill, finding the right size for the windows is something hard. In fact, the choice of the windows that is to say their lengths and their positions, seems to have too much impact on the results. As you can see on the table (1) in annexes, only a few samples difference in length or in position leads to strong inaccuracy on the estimate angle.

The thing is that we don't want such instability depending on the initial conditions, finding the best window for each pulse seems to be a case-by-case choice, and nothing let guess, a priori, which window could work better than an other. We have to figure out that 1 sample is a really really short time ($1/48000s$).

Thus we decided to automatise the windows. Basically what we did is we found the time position of the peak of each pulse (the time when the wave front hits the microphone) and asked the algorithm to start the window 4 samples before the peak, so that the window start right before the pulse. And then we could try different windows lengths to see which ones give the best results. The result of this empirical study is that shorter windows seems to give stronger accuracy. However with square windows, the results were not enough satisfying, that is why we tried to improve our windows as detailed in the next part.

6.1.2 Windowed pulses

Here is shown the windowing of the 2 first pulses before they get into the fft algorithm.

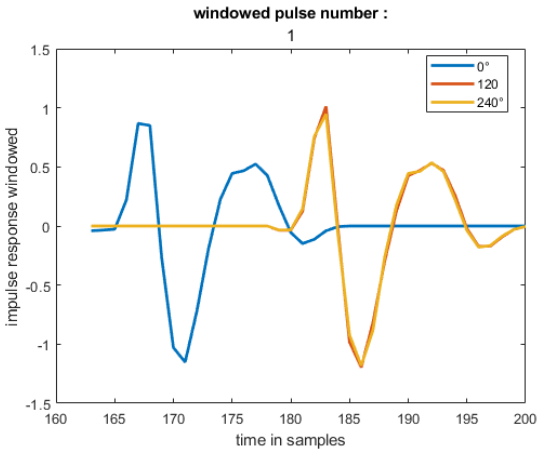

Figure 13

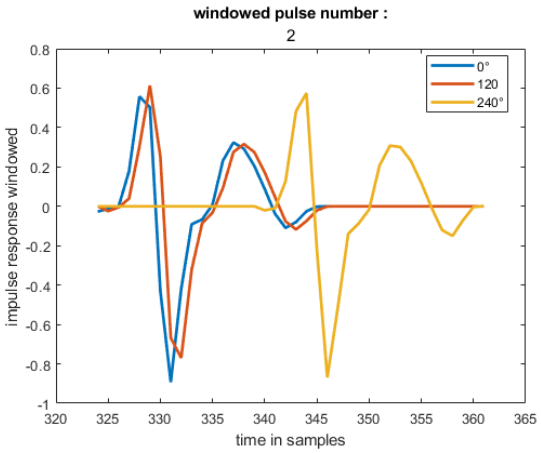

Figure 14

- A remark on these graphs :

We have built windows such as it begins like a regular square window but it ends like the second half of a Hanning window, so the values of the end of the pulse are smoothly getting down to zero.

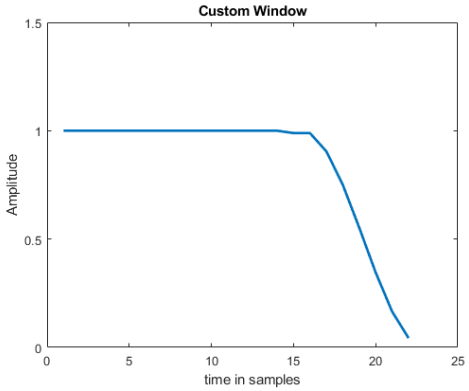

Figure 15 : shape of the window used.

Then an other tricky thing is that we cannot just take the figure (12) and cut it in windows, otherwise the window does not contain the same power for each microphone. For example, for the first wave front, the pulse arrive at the microphone placed at 0° at sample 167 and rush the two others at sample number 183. If we end the window at sample 262 then it contains more power from the 0° microphone, the edge of the window is in this case 95 samples after the peak for the 0° microphone, but only 79 samples after the peak for 120° and 240° microphones, so the windows contains more energy coming from the 0° sensor. This is not fair, while constructing the B-format microphones, the "weight" of the first response will be greater than the two others and this will distort the results.
To compensate this effect, we took windows of same sizes (70 samples after the peak) for the 3 microphones, and then to respect the delay that is between each one, we add some zeros before and after the pulses to slide the position of the peaks, while doing this way we did not add any phase shift effects.

6.2 Frequency response

6.2.1 A-format microphones

Below, it's the frequency response for two different wave fronts (the digit in the title indicates the number of the wave front)

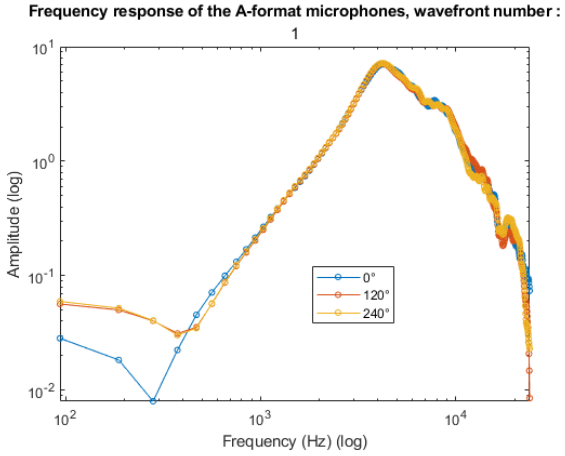

Figure 16 : Fourier transform of the first pulse captured by the 3 omni microphones

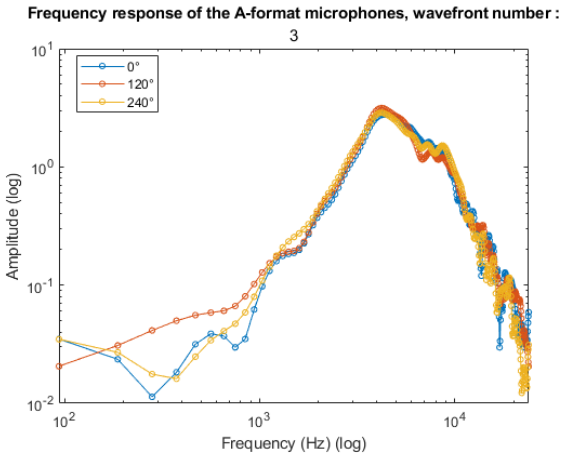

Figure 17 : Fourier transform of the third pulse captured by the 3 omni microphones

As we can see on the figures upwards, all frequencies below around 500 Hz behaves strangely, our sweep starts at 250 Hz so all frequencies below this value are noise. Then there is this slope that the three microphones seems to follow, and after 5000 Hz the frequency response seems unstable, and is different for the three microphones, this high frequencies are also random noise. To lead a proper study we will make sure that the noisy part of the spectrum will not be taken into account while computing the power coming from each direction. In conclusion we have to limit ourselves to a small bandwidth.

6.2.2 Frequency range

Like described in the preamble by Peter Svensson, we use Taylor expansion of the exponential function to make the linear combination. And so :

$$e^{-jka\cos(\theta)} \approx 1 - jka\cos\theta$$

As long as $ka << 1$. With a the radius of the circle and k the wavenumber such as $2\pi = \lambda k$.
Supposing that the linearization works well until $ka = 0,5$ then:

$$ka < 0,5 \iff \frac{2\pi}{\lambda}a < 0,5 \iff \frac{2\pi f}{c}a < 0,5$$
$$\iff f < 0,5\frac{c}{2\pi a}$$

With $a = 0.072$ m, we have to limit us to a very small bandwidth otherwise the serial expansion becomes too false:

$$f < \frac{0,5 \times 342}{2\pi \times 0,072} = 378Hz$$

Knowing that our sweep start at 250Hz, we will restrict to a 250-378 Hz bandwidth, which is ridiculous. For using a larger frequency range, there is only one parameter we can change : the radius a.
However, we cannot choose the radius a as small as possible, remember that for finding the zero degree of our system,

we need to study time delay between two points, (see figure (10)). And if a is too small we aren't capable anymore of finding correctly where the zero is placed.
In conclusion, the choice of the value of a is crucial, we have to balance between a bandwidth large enough and a precise position of our A-format microphones (which is linked with finding the position of the 0°).

6.3 Results with a smaller radius

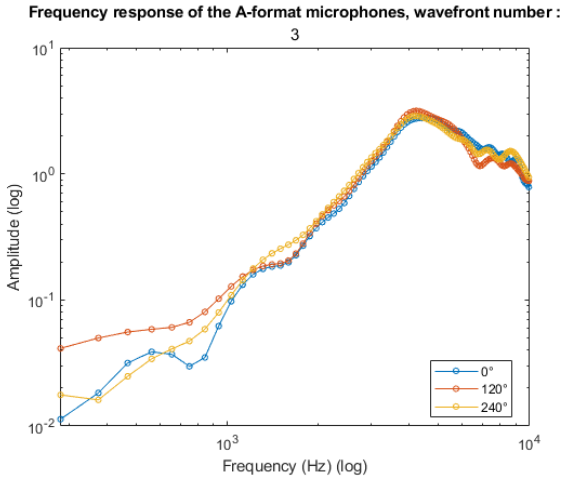

Figure 18 : fft of the signal recorded by the three A-format microphones with a = 7.2cm, third wave front.

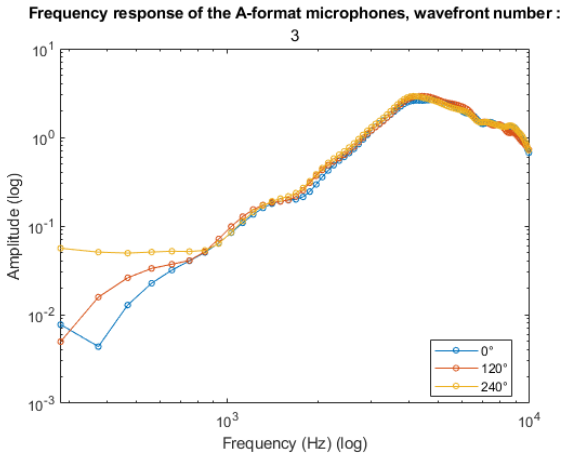

Figure 19 : fft of the signal recorded by the three A-format microphones with a = 2.8cm, third wave front.

Considering the case a=7,2 cm, there is a other effect that is uncomfortable, it generates some non-linearities in the frequency response, these ones are particularly visible in figure (18), between 1000Hz and 2000Hz. This might be a problem because again, while combining the ffts to build the B-format microphones, the weight of each might not be the same.
Hopefully, looking at the figure (19), we see that the three frequency responses are much more stacked one on each other, and so the linear combination between them is more balanced.

Measuring the power coming from each direction can now be done and a larger frequency range, $f < \frac{0,5\times342}{2\pi\times0.028} = 972Hz$. The average power is computed using 250-970 Hz bandwidth.

6.4 Identifying the position of the source using the time-delay method

6.4.1 Quick presentation of the method

The time delay method is quite an ancient and well-tried approach, it can be found in the literature since 1981. Basically the principle here is to record the same sound source with several microphones placed in a chosen form. Then the analysis of the data is only focused on the time domain. The algorithm find the exact time when a pulse hits each microphone and then, using the time delays between them and a bit of geometry, the script is able to return the direction from where the wave front is coming. A detailed explanation can be found in the annexes.

6.4.2 Comparing the results from the 2 different approaches

In the tables (2) (3) and (4) in annexes, it is mentioned the theoretical values, and the estimated values by the two algorithms for different placement of the microphones.

As showed in table (2), the precision for the time delay method is rather good, the sound field spot well the source for the first pulse but the precision drops greatly for the others wave fronts. Moreover, the time delay approach has the advantage to be very few sensitive to the choice of the windows, in fact as long as the peaks are situated in the window, the results will be okay.

The table (3) presents the results for closer microphones placement. Because of the phenomenon described in part 6.3, the results of the soundfield microphone approach are much more satisfying. However, even if both methods give quite good results, it seems that the time delay approach stay the best.

To finish, in the table (4), we see that the variable cardïoid approach give really good results, because the radius a is small, so we can measure average power in a larger bandwidth and thus be more precise. On the other hand, the time-delay analysis is for the first time less effective than the soundfield microphone method. These results are logical, the time-delay method is more efficient if the distance between the sensors is large, and if the sampling frequency is high. The all algorithm is based on detecting maxima, and even if we use oversampling in it to precisely return the position of the maximum, if the distance is too small between the microphones, the error is bigger when the the pulses are separated by 2,37 samples for example. Because 2,37 isn't close enough to an entire number of sample.
However to palliate, the short distances, we can use and higher oversampling frequency, such results are shown in ta-

ble (5). As you can see by increasing p by a factor 10, we can reach a more satisfying estimation.

7 Conclusion

This new approach for detecting position of audio sources is establishing itself as an alternative to time delay method. If we use small distances between the A-format microphone, the results are very satisfying. Also, it could be interesting to develop a virtual cardioïd that can rotate along the two angles of a spherical coordinate system θ and ϕ, so that it could detect source's position in 3-D. However, unless it gets improved, this soundfield microphone method has several drawbacks. First, the results are really sensitive on the choice of the window, even if a short windows length seems to be stable, we are not sure if it will work in different situations. Then, the complexity of the algorithm is quite heavy and it takes a certain time for the computer to return the results, there is lots of derivation for each wave fronts, so has definitely a higher complexity than the time-delay algorithm. To finish, the biggest problem of this elegant method, is the fact that if several wave fronts coming from different places of the room hit the microphone at the same time, the program will not be able to detect the position of the source, in fact it will think that there is only 1 pulse instead of two so it will be lost. Of course, it's also true for the time-delay method, but it will have less impact on the results if several waves hit the microphones at the same time.

Concerning a more general aspect of this internship, I personally really enjoyed it. It is the best way for students to see the "real world", and it was actually very surprising to see that the research field is far away from what we see at university. I had the chance to learn using equipment and form myself to programming (Matlab), it feels like french students are not used to python, Matlab, Latex, and this is in my opinion something that should change. About, moving abroad, it's a very good opportunity to practice English and discover new cultures, and will recommend future students to leave France, at least several months, to see the rest of the world. At last, I would like to thank Peter Svensson for his time and his dedication, he really transmitted the taste of making research and I learned many things along his side.

8 Bibliography

About the EASERA Software :
https://scout.univ-toulouse.fr/sw?type=gwtmail&state=30&oidMessage=674884361&AttachmentId=1

http://www.afmg-support.de/SoftwareDownloadBase/AFMG/EASERA12/EASERATutorial_EN.pdf

About the MATLab Software :
http://perso.unifr.ch/ales.janka/analnum/intro_matlab_fr.pdf

https://www.math.u-bordeaux.fr/~ayger/MATLABSignal/Tutorials/TutMatlab.pdf
https://fr.mathworks.com/

Various documents on the subject :
https://scout.univ-toulouse.fr/sw?type=gwtmail&state=30&oidMessage=674884361&AttachmentId=0

Lectures notes from "audio technology" class:
https://ntnu.blackboard.com/webapps/blackboard/execute/content/file?cmd=view&content_id=_1671018_1&course_id=_31962_1&launch_in_new=true

https://ntnu.blackboard.com/webapps/blackboard/execute/content/file?cmd=view&content_id=_1680675_1&course_id=_31962_1&launch_in_new=true

Data sheets:
https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjs8aaRwdn3AhW9QvEDHW9E\BBoQFnoECAMQAQ&url=https%3A%2F%2Fwww.nti-audio.com%2FPortals%2F0%2Fdata%2Fen%2FMeasurement-Microphones-Sp
pdf&usg=AOvVaw1OuvlpiGrJeACZtSvLAv9O

http://www.technibook.com/common/confrm.php?query=a3j4v383f574g484a4m4z5w5n4g4k51416n4i524c494i4\x5q215b3j4n392r4w4e4m5r474y5w5r2u4w23403p205a3e3

9 Annexes

9.1 Linear combination

9.1.1 Making the fig-8 directivity along X-axis

In this part we will show how the weight factors of the linear combination leads us to a bidirectional directivity $S_X = \cos\theta$.
According to equation (4), we have:

$$w_{W1}(1 + jka\cos\theta) + w_{W2}(1 - jk\frac{a}{2}\cos\theta + jk\frac{\sqrt{3}}{2}a\sin\theta) + w_{W3}(1 - jk\frac{a}{2}\cos\theta - jk\frac{\sqrt{3}}{2}a\sin\theta) = \cos\theta$$

Identifying sine and cosine parts, it comes:

$$\begin{cases} \left(w_{W1}jka - w_{W2}jk\frac{a}{2} - w_{W3}jk\frac{a}{2}\right)\cos\theta = \cos\theta & \forall\theta \in \mathbb{R} \\ \left(w_{W2}jk\frac{\sqrt{3}}{2}a - w_{W3}jk\frac{\sqrt{3}}{2}a\right)\sin\theta = 0 & \forall\theta \in \mathbb{R} \\ w_{W1} + w_{W2} + w_{W3} = 0 \end{cases}$$

$$\iff \begin{cases} \left(w_{W1} - \frac{w_{W2}}{2} - \frac{w_{W2}}{2}\right)jka\cos\theta = \cos\theta \\ (w_{W2} - w_{W3})jk\frac{\sqrt{3}}{2}a\sin\theta = 0 \\ w_{W1} + w_{W2} + w_{W3} = 0 \end{cases}$$

$$\iff \begin{cases} w_{W1} - w_{W2} = \frac{1}{jka} \\ w_{W2} = w_{W3} \\ w_{W1} + 2w_{W2} = 0 \end{cases}$$

$$\iff \begin{cases} w_{W2} = w_{W3} \\ w_{W2} = \frac{j}{3ka} \\ w_{W1} = -2w_{W2} = \frac{-2j}{3ka} \end{cases}$$

$$\iff \begin{cases} w_{W1} = \frac{-2j}{3ka} \\ w_{W2} = \frac{j}{3ka} \\ w_{W2} = \frac{j}{3ka} \end{cases}$$

9.1.2 Making the fig-8 directivity along X-axis

In this part we will show how the weight factors of the linear combination leads us to a bidirectional directivity $S_Y = \sin\theta$.
According to equation (4), we have:

$$w_{W1}(1 + jka\cos\theta) + w_{W2}(1 - jk\frac{a}{2}\cos\theta + jk\frac{\sqrt{3}}{2}a\sin\theta) + w_{W3}(1 - jk\frac{a}{2}\cos\theta - jk\frac{\sqrt{3}}{2}a\sin\theta) = \sin\theta$$

Identifying sine and cosine parts, it comes:

$$\begin{cases} \left(w_{W1}jka - w_{W2}jk\frac{a}{2} - w_{W3}jk\frac{a}{2}\right)\cos\theta = 0 & \forall\theta \in \mathbb{R} \\ \left(w_{W2}jk\frac{\sqrt{3}}{2}a - w_{W3}jk\frac{\sqrt{3}}{2}a\right)\sin\theta = \sin\theta & \forall\theta \in \mathbb{R} \\ w_{W1} + w_{W2} + w_{W3} = 0 \end{cases}$$

$$\iff \begin{cases} \left(w_{W1} - \frac{w_{W2}}{2} - \frac{w_{W2}}{2}\right)jka\cos\theta = 0 \\ (w_{W2} - w_{W3})jk\frac{\sqrt{3}}{2}a = 1 \\ w_{W1} + w_{W2} + w_{W3} = 0 \end{cases}$$

$$\iff \begin{cases} w_{W1} - \frac{w_{W2}}{2} - \frac{w_{W2}}{2} = 0 \\ w_{W2} - w_{W3} = \frac{2}{jk\sqrt{3}a} \\ w_{W1} + w_{W2} + w_{W3} = 0 \end{cases}$$

$$\iff \begin{cases} w_{W1} - \frac{w_{W2}}{2} - \frac{w_{W2}}{2} = 0 \\ \frac{3}{2}w_{W2}\frac{3}{2} = 0 \implies w_{W2} = -w_{W3} \\ -2w_{W3} = \frac{2}{jk\sqrt{3}a} \implies w_{W3} = \frac{j}{k\sqrt{3}a} \end{cases} \quad (3-1)$$

$$\iff \begin{cases} w_{W1} = 0 \\ w_{W2} = \frac{-j}{\sqrt{3}ka} \\ w_{W2} = \frac{j}{\sqrt{3}ka} \end{cases}$$

9.2 Fourier series decomposition

9.2.1 Cardioïd directivity

$$
\begin{aligned}
\overline{P}_{varcard}(\theta) &= A^2 \times D_{varcard}(\theta) \\
&= A^2[0,5 + 0,5\cos(\theta - \theta_{aim})]^2 \\
&= A^2[0,5 + 0,5(\cos\theta\cos(\theta_{aim}) + \sin\theta\sin(\theta_{aim})]^2 \\
&= \frac{A^2}{2}[1 + \cos^2\theta\cos^2(\theta_{aim}) + \sin^2\theta\sin^2(\theta_{aim}) + 2\cos\theta\cos_{aim} + 2sin\theta\sin(\theta_{aim}) + 2\cos\theta\sin\theta\cos(\theta_{aim})\sin(\theta_{aim})] \\
&= \frac{A^2}{2}\left[1 + \left(\frac{1+\cos(2\theta)}{2}\right)\cos^2(\theta_{aim}) + \left(\frac{1-\cos(2\theta)}{2}\right)\sin^2(\theta_{aim})\right] \\
&+ A^2(\cos\theta\cos(\theta_{aim}) + \sin\theta\sin(\theta_{aim}) + \cos\theta\sin\theta\cos(\theta_{aim})\sin(\theta_{aim})) \\
&= \frac{A^2}{2}\left[1 \times 1 + \left(\frac{\cos^2(\theta_{aim})}{2}\right) + \frac{\cos^2(\theta_{aim})}{2}\cos(2\theta) + \frac{\sin^2(\theta_{aim})}{2} - \left(\frac{\sin^2(\theta_{aim})}{2}\right)\cos(2\theta)\right] \\
&+ A^2[\cos(\theta_{aim})\cos\theta + \sin(\theta_{aim})\sin\theta + \cos(\theta_{aim})\sin(\theta_{aim})\cos\theta\sin\theta] \\
&= \frac{A^2}{2}\left(\frac{3}{2} \times 1 + \left(\frac{\cos^2(\theta_{aim})}{2} - \frac{\sin^2(\theta_{aim})}{2}\right)\cos(2\theta)\right) + A^2(\cos(\theta_{aim})\cos\theta + \sin(\theta_{aim})\sin\theta) \\
&+ \frac{A^2}{2}(2\cos(\theta_{aim})\sin(\theta_{aim})\sin(2\theta))
\end{aligned}
$$

Knowing that $\cos^2 x - \sin^2 x = \cos(2x)$

$$
= \frac{3A^2}{4} \times 1 + A^2\cos(\theta_{aim})\cos\theta + A^2\sin(\theta_{aim})\sin\theta + \frac{A^2}{4}\cos(2\theta_{aim})\cos(2\theta) + \frac{A^2}{2}\sin(2\theta_{aim})\sin(2\theta)
$$

Identifying with the fourier decomposition :

$$
a_0 \times 1 + a_1\cos\theta + b_1\sin\theta + a_2\cos(2\theta) + b_2\sin(2\theta)
$$

We can conclude on the expression of the coefficients :

$$
a_0 = \frac{3A^2}{4} \qquad a_1 = A^2\cos(\theta_{aim}) \qquad b_1 = A^2\sin(\theta_{aim}) \qquad a_2 = \frac{A^2}{4}\cos(2\theta_{aim}) \qquad b_2 = \frac{A^2}{2}\sin(2\theta_{aim})
$$

9.2.2 Figure of eight directivity

$$
D_{varfig8} = A^2\cos(\theta - \theta_{aim})
$$

The average power of the signal depending on θ_{aim} is then:

$$
\begin{aligned}
\overline{P}_{varfig8}(\theta_{aim}) &= A^2(\cos(\theta - \theta_{aim})^2) \\
&= A^2(\cos\theta\cos(\theta_{aim}) + \sin\theta\sin(\theta_{aim}))^2 \\
&= A^2\big[\cos^2\theta\cos^2(\theta_{aim}) + \sin^2\theta\sin^2(\theta_{aim} + 2\cos\theta\cos(\theta_{aim})\sin\theta\sin(\theta_{aim}) \\
&= A^2\left[(1+\cos(2\theta))\frac{\cos^2(\theta_{aim})}{2} + (1-\cos(2\theta))\frac{\sin^2(\theta_{aim})}{2} + 2\cos(\theta_{aim})\sin(\theta_{aim})\cos\theta\sin\theta\right] \\
&= \frac{A^2}{2}\cos^2\theta_{aim} \times 1 + \frac{A^2}{2}\cos\theta_{aim}\cos(2\theta) + \frac{A^2}{2}\sin^2(\theta_{aim}) \times 1 - \frac{A^2}{2}\sin^2(\theta_{aim})\cos(2\theta) \\
&+ A^2\cos(\theta_{aim})\sin(\theta_{aim})\sin(2\theta) \\
&= \frac{A^2}{2} + \frac{A^2}{2}(\cos^2(\theta_{aim}) - \sin^2(\theta_{aim}))\cos(2\theta) + A^2\cos(\theta_{aim})\sin(\theta_{aim})\sin(2\theta) \\
&= \frac{A^2}{2} + \frac{A^2}{2}\cos(2\theta_{aim})\cos(2\theta) + \frac{A^2}{2}\sin(2\theta_{aim})\sin(2\theta)
\end{aligned}
$$

Identifying with the fourier decomposition at first order :

$$F(\theta) = a_0 \times 1 + a_1 \times \cos\theta + b_1 \times \sin\theta$$ [7]

We can conclude on the expression of the coefficients :

$$a_0 = \frac{A^2}{2} \qquad a_1 = \frac{A^2}{2}\cos(2\theta_{aim}) \qquad b_1 = \frac{A^2}{2}\sin(2\theta_{aim})$$

9.3 Tables of results

Wavefront n°	Start	End	Length	Real value (°)	Estimate value (°)
1	10	300	290	0	-0.8
2	319	366	47	57.6	56.1
3	447	530	83	180	179.6
4	530	560	30	148.6	131.6
1	10	300	290	0	-0.8
2	319	368	49	57.6	53.3
3	448	530	82	180	-166.3
4	530	561	31	148.6	138.9

Table 1 : The column start defines the sample where the window start, end is the sample number where the windows end, and length is the length of the window. Here we chose windows arbitrarily to have the closest results to the real values.

Wavefront n°	Real value (°)	Time delay (°)	Error time-delay	Soundfield(°)	Error Soundfield
1	0	0,9	0,9	-1,5	1,5
2	57,6	50,4	7,1	58,3	0,7
3	180	177,2	2,8	173,1	6,9
4	148,6	144,5	5,1	162,4	13,8

Table 2 : Comparison of the results between two different methods to localise a sound source, data set of microphone on a 7.2cm radius circle.

[7] Notice that we have here a period doubling but it's useless to go until second order decomposition otherwise $a_1 = b_1 = 0$

Wavefront n°	Real value (°)	Time delay (°)	Error time-delay	Soundfield(°)	Error Soundfield
1	0	0,5	0,5	-0,8	0,8
2	57,6	56,0	1,6	57,3	0,3
3	180	178,9	1,1	175,4	4,6
4	148.6	150,9	2,3	146,0	2,6

Table 3 : Same table as above using a different data set, microphone are now placed on a 2.8 cm radius circle.

Wavefront n°	Real value (°)	Time delay (°)	Error time-delay	Soundfield(°)	Error Soundfield
1	0	1,5	1,5	-0,4	0,4
2	57,6	56,2	1,4	57,0	0,6
3	180	177,5	2,5	177,3	2,7
4	148.6	152,8	4,2	151,1	2,5

Table 4 : Same table as above using a different data set, microphone are now placed on a 1.7 cm radius circle.

Wavefront n°	Real value (°)	p=8	Error for p=8	p=80	Error for p=80
1	0	1,5	1,5	0,2	0,2
2	57,6	56,2	1,4	57,5	0,1
3	180	177,5	2,5	179,6	0,4
4	148.6	152,8	4,2	155,2	6,6

Table 5 : Effect of the oversampling rate p on the precision of the time-delay method.p is the oversampling rate such as the oversampling frequency $f_{os} = p \times f_s$

9.4 Time-delay method (Extract of lecture notes from Peter Svensson)

It exists an other way to spot the angular position of a sound source using an array of microphone. Such a method consist in analyse the time delay between the different microphones placed in triangle. Let's call t_1, t_2 and t_3 the time of arrival of the acoustic Dirac pulse for the microphones 1,2 and 3. Even if the microphones are really close from each other, the lag between t_1 and t_2 can be in the order of a few samples.
Recall that the microphones are on a circle or radius $r = a \approx 0,02$ m. So the distance between each microphone is $\sqrt{3}a \approx 0,03$ m. Such a distance is run in $\tau_s = \frac{0,03}{343} \approx 8,8e^{-5}s$ Where $v_s = 343$ m/s is the sound velocity in the air at 20°C.
To get the time delay in samples, which is more significant, we multiply τ_s by the sampling frequency which is usually $f_s = 48000Hz$.
And So :

$$\tau_{sp} = \tau_s \times f_s \approx 4,2 \text{ samples}$$

Depending on the position of the source, the delay between the three microphone will change, and we are able to find the angular position using this data.

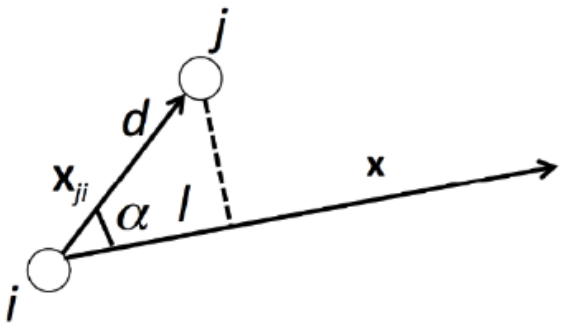

Figure 20 : Schematic diagram of the time-delay method (with the kind permission of P.Svensson)

On the figure above, i and j are two microphones, x_{ji} is the distance between the two microphones that we know, x is a vector pointing to the source, (-x is the propagation vector) and α is the angle that we are searching for.

Detailing a bit of the calculations: If t_i and t_j are respectively the times when the pulse hit the microphones i and j, τ_{ji} is the time delay between i and j such as $\tau_{ji} = t_j - t_i$
Thus we can write l depending on τ_{ji}:

$$l = -v_s.\tau_{ji}^{8} \qquad \text{and also,} \qquad l = ||x_{ji}||.\cos\alpha = d.\cos\alpha$$

That leads to

$$\alpha = \cos^{-1}\left(-\frac{v_s.\tau_{ji}}{d}\right) = -\alpha$$

The angle of the source cannot be determined with only 2 microphones !

Again we suppose that the source is located in the same plan than the triangle formed by the microphones.

$$\vec{x_{ji}}.\vec{x} = |\vec{x_{ji}}|\,|\vec{x}|\cos\alpha \implies |\vec{x_{ji}}|\,.\cos\alpha = \frac{\vec{x_{ji}}.\vec{x}}{|x|}^{9}$$

Also, $l = |x_{ji}|.\cos\alpha = -v_s.\tau_{ji}$
Then, we conclude

$$v_s.\tau_{ji} + \vec{x_{ji}}.\vec{x} = 0 \tag{9}$$

[8] There is a minus sign because the the wave is moving along $-\vec{x}$
[9] The norm of $\vec{x}$ can be defined such that |x|=1

If we settle $|x| = 1$, then we can express $\vec{x}$ in function of θ such that:

$$\vec{x} = (\cos\theta, \sin\theta)$$

Where θ is an angle in the cartesian coordinates system and thus independent of the position of the sensors.
The equation () gives the following system.

$$\begin{cases} v_s.\tau_{21} + \vec{x_{21}}.\vec{x} = 0 \\ v_s.\tau_{31} + \vec{x_{31}}.\vec{x} = 0 \\ v_s.\tau_{32} + \vec{x_{32}}.\vec{x} = 0 \end{cases} \tag{10}$$

Where $\vec{x} = (x, y)$ and $\vec{x_{ji}} = (X_{ji}, Y_{ji})$ Which can be written in a matrix equation form:

$$\begin{bmatrix} X_{21} & Y_{21} \\ X_{31} & Y_{31} \\ X_{32} & Y_{32} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \tau_{21} \\ \tau_{31} \\ \tau_{32} \end{bmatrix} \iff \mathcal{P}.\mathcal{M} = \Delta \tag{11}$$

As we can see, this system is over determined, there is 2 unknowns and 3 equations, to solve the system in the more precise way we can use the pseudo-inverse matrix. In this case we will use the all three equations instead of two to reduce indeterminations that may be caused by noise during the measurements.

9.5 Pseudo inverse matrix method

When a linear system of several is overdetermined (when we have more equation than unknowns) it exists a way to optimise the resolution of this system. In our case the two unknows x and y do not fulfill the system (10), indeed, none of the three equations are exact, the noise prevent the exactitude. However the more equation we have the more precise we can be and hopefully there is a way to minimize the error using the all three equations. The least squares method consist in define an error ϵ and minimize it, this error can be the sum of the squared values of the three equations.
The pseudo inverse method has the same objective as the previous one, that's to say get the more precise result. If a linear system of equation can be written in a matrix equation, then we can solve it with the pseudo inverse method. The pseudo inverse is nothing more a generalisation of inverse for non squares matrices. The tip is to remember that if $\mathcal{P}$ is a matrix square or not, then, using the transpose matrix, $\mathcal{P}^T.\mathcal{P}$ is a square matrix and it's inverse is defined. Let's take our matrix equation (11).

$$\mathcal{P}\mathcal{M} = \Delta \iff \mathcal{P}^T\mathcal{P}\mathcal{M} = \mathcal{P}^T\Delta \implies \mathcal{M} = \left(\mathcal{P}^T\mathcal{P}\right)^{-1}\mathcal{P}^T\Delta$$

And $\mathcal{M}$ contains the coordinates of the source calculated in the most accurate way, that's to say using all the equations of our system.

9.6 Using cardioïd microphone instead of omni

We will treat a other approach of our variable cardioid virtual microphone. We will this time use cardioïd microphones centered on the rotation point of the turntable and with different orientations (0°,120° and 240°). Using such directivity function permit to combine the signals directly in the objective to form the virtual variable cardioïd microphone. Without making the B-format signals, we make less steps and so less approximations to build the virtual microphone.

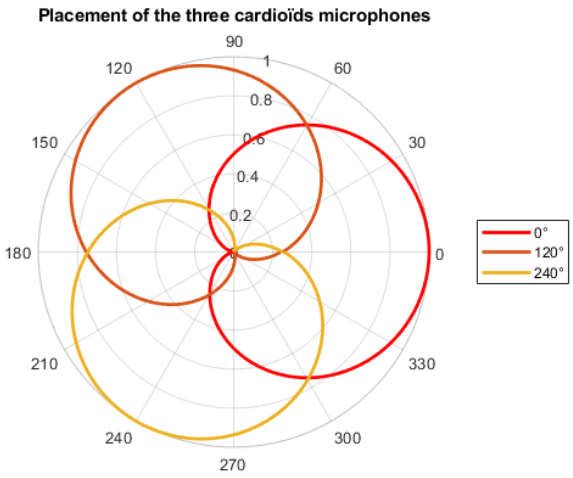

Figure 21 : Orientation of the three microphones, actually we don't use 3 microphones but only one with it we measure the impulse response in three different orientations.

As follows the derivations that lead us to the right weight factors of the linear combinations.

Our starting point is the three directivity functions :

$$D_1(\theta) = \frac{1}{2}(1+\cos(\theta)) \qquad D_2(\theta) = \frac{1}{2}(1+\cos(\theta-120)) \qquad D_3(\theta) = \frac{1}{2}(1+\cos(\theta-240))$$

Using the formula $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$ and equation (4) it comes:

$$\begin{aligned}
&w_1 D_1(\theta) + w_2 D_2(\theta) + w_3 D_3(\theta) = D_{vc}(\theta) = \frac{1}{2}(1+\cos(\theta-\theta_{aim}) \\
&\iff w_1\left(\frac{1}{2}(1+\cos(\theta))\right) + w_2\left(\frac{1}{2}(1+\cos(\theta-120))\right) + w_3\left(\frac{1}{2}(1+\cos(\theta-240))\right) \\
&= \frac{1}{2}(1+\cos(\theta)\cos(\theta_{aim}) + \sin(\theta)\sin(\theta_{aim}) \\
&\iff \frac{1}{2}(w_1+w_2+w_3) + \cos(\theta)\left(\frac{w_1}{2} + \frac{w_2}{2}\cos(120) + \frac{w_3}{2}\cos(240)\right) + \sin(\theta)\left(\frac{w_2}{2}\sin(120) + \frac{w_3}{2}\sin(240)\right) \\
&= \frac{1}{2}(1+\cos(\theta)\cos(\theta_{aim}) + \sin(\theta)\sin(\theta_{aim}))
\end{aligned}$$

And then identifying sine and cosine parts, the following system appears.

$$\begin{cases} 1 = w_1 + w_2 + w_3 \\ \cos(\theta_{aim}) = w_1 - \frac{1}{2}w_2 - \frac{1}{2}w_3 \\ \sin(\theta_{aim} = \frac{\sqrt{3}}{2}w_2 - \frac{\sqrt{3}}{2}w_3 \end{cases} \iff \begin{cases} \cos(\theta_{aim} = 1 - \frac{3}{2}(w_2+w_3) \\ \sin(\theta_{aim} = \frac{\sqrt{3}}{2}(w_2 - w_3) \end{cases}$$

Which can expressed as a matrix equation:

$$\begin{bmatrix} -\frac{3}{2} & -\frac{3}{1} \\ \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}\begin{bmatrix} w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} \cos(\theta_{aim}) - 1 \\ \sin(\theta_{aim}) \end{bmatrix} \iff \begin{bmatrix} w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & -\frac{3}{1} \\ \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}^{-1}\begin{bmatrix} \cos(\theta_{aim}) - 1 \\ \sin(\theta_{aim}) \end{bmatrix}$$

$$\iff \begin{bmatrix} w_2 \\ w_3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -1 & \sqrt{3} \\ -1 & -\sqrt{3} \end{bmatrix}^{-1}\begin{bmatrix} \cos(\theta_{aim}) - 1 \\ \sin(\theta_{aim}) \end{bmatrix} \iff \begin{bmatrix} w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3}\cos(\theta_{aim}) + \frac{1}{3} + \frac{1}{\sqrt{3}}\sin(\theta_{aim}) \\ -\frac{1}{3}\cos(\theta_{aim}) + \frac{1}{3} - \frac{1}{\sqrt{3}}\sin(\theta_{aim}) \end{bmatrix}$$

And if we inject the results in the previous system we have:

$$\begin{cases} 1 = w_1 + w_2 + w_3 \\ w_2 = -\frac{1}{3}\cos(\theta_{aim}) + \frac{1}{3} + \frac{1}{\sqrt{3}}\sin(\theta_{aim}) \\ w_3 = -\frac{1}{3}\cos(\theta_{aim}) + \frac{1}{3} - \frac{1}{\sqrt{3}}\sin(\theta_{aim}) \end{cases} \iff \begin{cases} w_1 = 1 + \frac{2}{3}\cos(\theta_{aim}) \\ w_2 = -\frac{1}{3}\cos(\theta_{aim}) + \frac{1}{3} + \frac{1}{\sqrt{3}}\sin(\theta_{aim}) \\ w_3 = -\frac{1}{3}\cos(\theta_{aim}) + \frac{1}{3} - \frac{1}{\sqrt{3}}\sin(\theta_{aim}) \end{cases}$$

Now that the weight factors are known, the linear combination in order to form the variable cardioïd can be done.

9.6.1 Find the Zero

An important thing that have to be done is placing the three real microphones in a precis way. As before, the principle is to set a 0° angle that faces the loudspeaker. Again there is a technique, using the omnidirectional we used a time delay method, however for the cardioïds microphones it will not be possible. Indeed the three microphones are supposed to be at the same place (just the orientation changes) thus, there is no time delay between each microphones. Nevertheless, we can use the cardioïd characteristic, comparing the energy received by the microphone at +120° and -120°. We will compare the two signals in frequency domain because it's more precise. When the two frequency responses are perfectly overlapping each other we know that we are at +120° and -120° and we can then deduce the position of the 0°.
Yet, it seems to be a little inconsistency while calibrating the equipment, even if our frequency responses look the same (which means we are well at 120° and -120°) there is still a time delay of one sample between the two signals. Weird

9.6.2 Results

The measurements were done using a cardioïd microphone Rode NT5 and a sampling rate of 48000Hz.
We will presents results coming from three different measurements sets. We used the same windows as described part 6.1.2, however the accuracy may be improved, just have not figured yet how to do it.

It seems important to mention that as shown in the previous part, the weight factors don't depend on the frequency for this setup. That means we can either evaluate the average power in time or in frequency domain. We propose tables that compare the two cases (while studying time or frequency domain).

Furthermore, as documented in the data sheet of the microphone, the directivity function isn't a perfect cardioïd for high frequencies. Thus we low pass the signals before we analyse the average power over frequency. For the following results, we used a bandwidth from 350 Hz to 8250 Hz. After several tries it seems that this bandwidth gives the best results.

Wavefront n°	Real value (°)	Time domain (°)	Error t domain	Frequency domain (°)	Error fr domain
1	0	0,9	1,4	2,7	2,7
2	57,6	47,8	9,8	97,5	39,9
3	180	-174,4	6,6	179,9	0,1
4	148,6	149	0,4	147,8	0,8

Table 6 : Results of the variable cardioïds made of real cardioïds method, comparison of integrate average power over time and frequency

As you can see on the results, it is pretty hard to say which one works the best, both have not that good results. The frequency domain study seems more accurate but there is this very big error about the second wavefront, nothing really satisfying here.

Next two tables represents the same experiment as before, however we tried to be a bit more precise on the placement of the microphone. Moreover, the following two measurements were done one right after the other with the exact same conditions, we can then think about the reliability of the equipment.

Wavefront n°	Real value (°)	Time domain (°)	Error t domain	Frequency domain (°)	Error fr domain
1	0	8,2	8,2	7,4	7,4
2	57,6	33,3	23,3	66,7	9,1
3	180	-170,8	9,2	-176,7	3,3
4	148,6	140,2	8,4	152,7	4,1

Table 7 : Results of the variable cardioïds made of real cardioïds method, session 1.

Wavefront n°	Real value (°)	Time domain (°)	Error t domain	Frequency domain (°)	Error fr domain
1	0	8,6	8,6	8,7	8,7
2	57,6	34,2	23,4	67,1	9,5
3	180	-171,3	9,7	-176,6	3,4
4	148,6	140,8	7,8	151,3	2,7

Table 8 : Results of the variable cardioïds made of real cardioïds method, comparison of integrate average power over time and frequency, session 2.

As you can see on the previous tables (6) and (7), our attempt to raise precision is nothing but satisfying. Moreover, in the two same conditions, the results of the method are shifted about 1° for example for the fist and fourth wave fronts using integration along frequency domain.