POLITECNICO DI MILANO
Scuola di Ingegneria dell'Informazione

POLO TERRITORIALE DI COMO
Master of Science in Computer Engineering

A METHOD FOR THE ESTIMATION OF THE SOUND ABSORPTION COEFFICIENTS OF PLANAR SURFACES

Supervisor: Prof. Fabio ANTONACCI
Assistant Supervisor: Prof. Marco TAGLIASACCHI

Master Graduation Thesis by:
Andrea RIVA
Student Id. number 729903

Accademic Year 2010/2011

POLITECNICO DI MILANO
Scuola di Ingegneria dell'Informazione

POLO TERRITORIALE DI COMO
Corso di Laurea Specialistica in Ingegneria Informatica

METODO PER LA STIMA DEI COEFFICIENTI DI ASSORBIMENTO ACUSTICO DI SUPERFICI PIANE

Relatore: Prof. Fabio ANTONACCI
Correlatore: Prof. Marco TAGLIASACCHI

Tesi di laurea di:
Andrea RIVA
matr. 729903

Anno Accademico 2010/2011

Ai miei genitori,
Beppe e Vanda

Sommario

Il lavoro presentato in questa tesi tratta il problema della stima dei coefficienti di assorbimento acustico di una superficie, attraverso misurazioni *in-situ*. Tali misurazioni sono effettuate direttamente nell'ambiente in cui si trova la superficie da analizzare, senza la necessità di asportare campioni di materiale, per uno studio in laboratorio a posteriori. Il metodo si configura quindi come non distruttivo. Le possibili applicazioni di questa tecnica riguardano la stima delle caratteristiche acustiche di materiali non classificati, o semplicemente sconosciuti. La soluzione proposta si basa sulla misurazione diretta del segnale incidente e di quello riflesso dalla parete. La separazione tra questi due segnali, temporalmente sovrapposti, è realizzata per mezzo di un array di microfoni altamente direzionale. L'utilizzo di questo tipo di array consente inoltre al sistema di essere robusto rispetto a fenomeni di riverbero. I risultati della stima dei coefficienti di assorbimento acustico sono presentati in funzione della frequenza. Non è invece considerata alcuna dipendenza rispetto all'angolo di incidenza, a causa di vincoli di implementazione. È infatti richiesto che il fronte d'onda del segnale incidente sia sempre normale rispetto alla superficie sotto analisi. Il metodo proposto è stato validato attraverso simulazioni ed esperimenti.

Abstract

The work presented by this thesis deals with the problem of the sound absorption coefficient estimation, by means of *in-situ* measurements. These measurements are directly taken in the environment in which the surface under consideration is placed. Thus, there is not the need to remove material samples for *a posteriori* laboratory study. Consequently the method can be classified as non-destructive. The possible applications of this technique regard the estimation of the acoustical properties of unclassified, or simply unknown materials. The proposed solution is based on the direct measurement of the energy related to the signal incident to the surface under test, and to the reflected one. The separation between these two signals, temporally superimposed, is achieved by means of a highly directive microphone array. Moreover, the use of this kind of array allows the system to be robust against reverberation phenomena. The results for the sound absorption coefficient estimation are provided as a function of the frequency. On the contrary, no information about the angle of incidence is considered, due to implementation constraints. In fact, it is required that the incident signal wavefront is always normal respect to the surface under analysis. The proposed method has been validated by means of simulations and experiments.

Ringraziamenti

Desidero innanzitutto ringraziare il Prof. Fabio Antonacci e il Prof. Marco Tagliasacchi, per aver reso possibile questa tesi e per il prezioso tempo che mi hanno dedicato durante questi mesi di lavoro. Un ringraziamento va anche alla Prof.ssa Paola Folegati, per avermi incoraggiato in questa esperienza, per i preziosi consigli, ma soprattutto per la passione che ha saputo trasmettermi.
Un grazie speciale ai miei genitori, a cui sono immensamente riconoscente, per avermi sempre sostenuto durante questo lungo viaggio che ormai volge al termine. Grazie, soprattutto per avermi incoraggiato e spronato nei momenti più difficili.
Un grazie particolare a Veronica per avermi supportato e sopportato, in questi ultimi mesi di lavoro. Infine, parafrasando una canzone a me sì cara, *Oh, thanks for all my friends.*

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Sound absorption operated by walls plays a determinant role in the study of the acoustics of the environments and in the way they interact with propagating wavefronts. An estimate of these parameters can predict the behaviour of a given environment. Many simulation software, starting from the knowledge of the geometry of a room and the absorption coefficients of walls and objects that are contained in, are able to determine how the wavefronts dim during their propagation. Usually these tools are used in the design of new rooms and buildings. The dual case is the one in which, from a preexistent situation, it is necessary to estimate absorption and reflection properties of non-classified or simply unknown materials. The obtained data is a starting point for solving problems related to noise control, for both domestic and industrial purposes. Moreover, an interesting case might be represented by diffusion systems, capable of self-calibration according to the characteristics of the surrounding environment. Such a systems could, for example, pre-equalize the output signal in order to compensate the room response, using the data gathered during the measurement step. The work of this thesis covers the latter case and proposes an *in situ* estimating system of material absorption coefficients. Consequently, measurements can be performed directly in the environments under study, not requiring the removal of material for laboratory analysis. Particular attention was given to the fact that halls where we make measures could be highly reverberant. In fact, the proposed system aims at being robust against wavefronts coming from areas that are not the one under investigation.
Some estimation methods in the literature, such as the Sabine[5] one,

use the information coming from reverberations. But the latter case presents an approximate solution, subject to heavy application limitations. Methods such as the normal incidence ones, are able to gather more valuable results , but obtained only in highly controlled situations. Due to their execution mode, these techniques require small amounts of material to be removed. Therefore, under certain conditions those are classified as destructive methods. Instead the method proposed by Mommertz [3] provides *in situ* measurements, but does not consider the problems coming from reverberation.

Other *in situ* methods have been proposed by the Adrienne [6] research team, but in this case the objective was to identify methodologies to estimate the efficiency of road noise barriers, which goes beyond the scope of this work.

The system presented in this thesis takes the cue from the work proposed by Ducorneau et al [7], another *in situ* method that uses a multipolar weighting linear array [8]. This array allows the system to be robust against reverberations. Our method starts from the same basic idea, but introducing the use of a superdirective array realized by means of a fourth order differential system, which maximizes the front to back ratio [4]. This kind of array makes possible to concentrate its focus on a small portion of the analyzed surface. As the array is very directive, all signals coming from other directions and generated by secondary reflections can be neglected.

Since the reflection coefficient estimation is performed by calculating the ratio between the energy reflected from the material and the incident one, the adopted solution realized by the array described above can achieve significant results. In fact we will show in the experimental chapter that the estimation of the reflected wavefront is almost free of spurious components.

The system uses an MLS source signal (Maximum Length Sequence), which is generated by a speaker oriented in normal direction respect to the plane to be analyzed. Between the plane and the speaker there is the measurement instrument, which is the superdirectional microphone array. As anticipated, this is the feature that allows the system to be robust against reverberation phenomena: the high directivity of the array allows to *hear* only the reflected signal, provided that the plan under study is sufficiently large. Furthermore, the array has the characteristic of being able to isolate the direct signal coming from the

speaker and the reflected one, coming from the wall. So it is possible to *listen to* one direction and to the opposite one in the same moment, but generating two distinct signals.
Due to spatial sampling constrains the system must acquire and process the signals into subbands, in order to avoid aliasing problems. So we require more arrays having the same referencing point, but with different spacings between microphones. The spacing is related to the frequency band that a particular array has to process. Arrays with small spacings are suitable to process high frequency signals, whereas arrays with large spacings are proper for low frequency ones.
When a microphone input is read by the system, it is decomposed into subbands. After that, each subband is transmitted to the proper array processing system, according to its microphone position. The set of the arrays has to ensure the coverage of a frequency range large enough. In order to fulfill the latter requirement we choose to use 4 different arrays of 5 microphones each one. The microphones belonging to the same array are uniformly spaced. With the purpose of optimizing the cost and the manageability of the system, the spacings were chosen so that different arrays share some microphones. Using this option it is possible to realize the latter configuration, using 13 microphones globally and covering a frequency range between 300 and 3000 Hz.
At the completion of subband processing, the signals are reconstructed dually respect to the way they were decomposed. Then we proceed to calculate the impulse responses for each of the reconstructed signals. Finally, the frequency-dependent reflection coefficient is estimated as the ratio of the absolute value of the Fourier transform of the direct and reflected signal.
The overall block diagram of the system is represented in Figure 1.1. There we can identify all the components previously described.
A limitation of this project is that it can't analyze the variation of reflection coefficients as function of the incidence angle of the signal, because all the components (speaker, array and surface under test) must always be aligned and perpendicular to the wall under analysis.
Chapter 2 describes the physics of the sound reflection phenomenon, together with all the basic techniques used in the construction and implementation of the proposed system. Among these, we remember the Wavelet band decomposition and in particular the discrete Wavelet, for its time-invariance characteristic. Information about spectral esti-

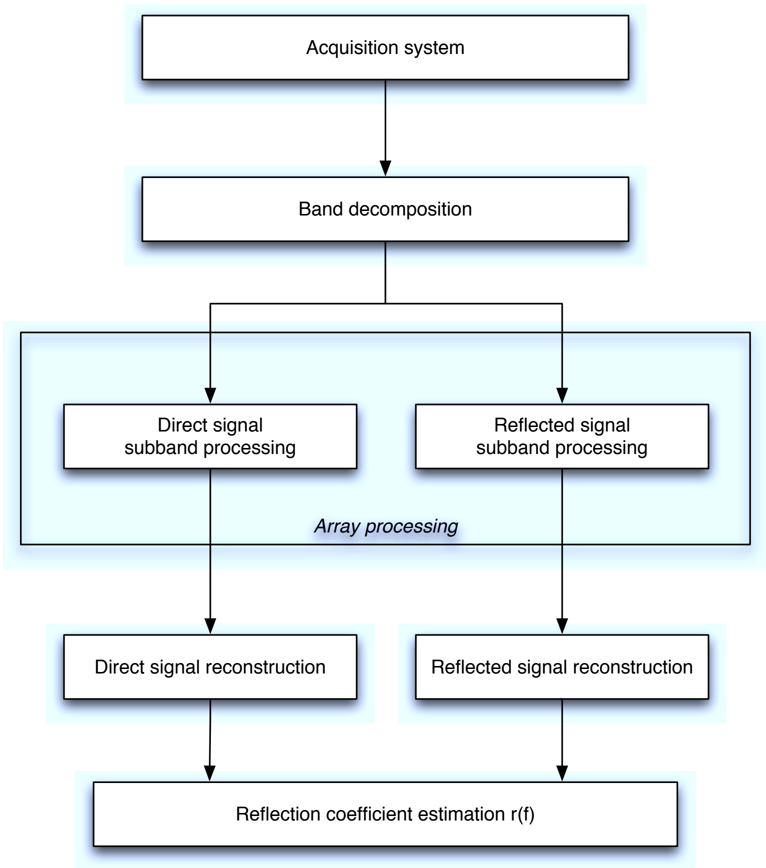


Figure 1.1: Overall system block diagram

mation is reported too. It is used for the system characterization. We can also find the theory for the construction of the superdirectional microphone array. There is also a brief digression on methods already existing in the literature.
Chapter 3 describes the implementation of the techniques outlined in Chapter 2, applied to our real system. Starting from data acquisition we arrive up to the calculation of sound absorption coefficients. Each block represented in Figure 1.1 is analyzed in detail.
Chapter 4 shows the obtained results, first in simulation and then by actual measurements. We extensively talk about the realization of the experimental setup and we show some results on the estimation of the model used in the system.
Chapter 5 concludes this thesis by highlighting the strengths and limitations of this system and laying the foundations for future developments.

Chapter 2

Background

This chapter describes the theory involved in the implementation of our project. Since the estimation of the sound reflection coefficients is the goal of our work, we start from a detailed description of what these values are and how they are obtained. The starting point for the discussion is the general wave equation that describes how the wavefronts propagate into physical mediums. Section 2 briefly describes the state of the art in sound reflection coefficients estimation. After that, the chapter takes into account the description of a superdirectional microphone array, which is the core of our system. The spectral estimation, used in system characterization and the Wavelet decomposition, required for sub-band processing, are presented in Section 4 and in Section 5, respectively.

2.1 Sound reflection

The starting point for this discussion is the study of the acoustic wave equation. For the moment we assume that the medium can be characterized as inviscid and thermally non-conducting. In this case the continuity, momentum and state equations can be respectively written as

$$\frac{\partial \rho}{\partial t} + \mathbf{u} \cdot \nabla \rho + \rho \nabla \cdot \mathbf{u} = 0, \tag{2.1}$$

$$\rho \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \, \mathbf{u} \right] + \nabla P = 0, \tag{2.2}$$

$$P = p_0 + A \, \frac{\rho - \rho_0}{\rho_0} + \frac{B}{2!} \left(\frac{\rho - \rho_0}{\rho_0} \right)^2 + \frac{C}{3!} \left(\frac{\rho - \rho_0}{\rho_0} \right)^3 + \ \ldots , \tag{2.3}$$

where P is the sound pressure, ρ is the average density in the considered volume, $\mathbf{u}$ is the particle velocity, whereas p_0 and ρ_0 are the static values of P and ρ, respectively.
Since the medium is lossless, the energy equation is unnecessary. In absence of an acoustic stimulus the previous equations are respectively satisfied by

$$\rho = \rho_0, \;\; P = p_0, \;\; \mathbf{u} = 0.$$

If an excitation breaks the quiet condition, the previous solutions can be rewritten as

$$\rho = \rho_0 + \delta\rho, \tag{2.4}$$

$$P = p_0 + p, \tag{2.5}$$

$$\mathbf{u} = 0 + \mathbf{u}, \tag{2.6}$$

where $\delta\rho$, p and $\mathbf{u}$ are respectively the variation of density, the pressure and the particle velocity introduced by the stimulus.
The second approximation we introduce is the so-called *small signal* approximation, for which we can neglect non-linear terms. Note that this approximation is valid even for the loudest sounds, since most of the waves disturb the fluid status only in a negligible way. We can formalize the *small signal* approximation assuming

$$|\delta\rho| \ll \rho_0,$$

where $\rho_0 = 1.21\frac{kg}{m^3}$ is the air density measured at 20°C. If the previous inequality is true, it can be proved that

$$|p| \ll \rho_0 c_0^2,$$

where $c_0 = 343\frac{m}{s}$ is the sound propagation speed in the air, at 20°C. Similarly, if one of the previous two inequalities is true it can be proved that

$$|u| \ll c_0^2.$$

Now, if we substitute (2.4), (2.5) and (2.6) in (2.1), (2.2) and (2.3), considering the *small signal* approximation, after some calculations we get

$$\frac{\partial \delta p}{\partial t} + \rho_0 \nabla \cdot \mathbf{u} = 0, \tag{2.7}$$

$$\rho_0 \frac{\partial \mathbf{u}}{\partial t} + \nabla p = 0, \tag{2.8}$$

and

$$p = c_0^2 \delta\rho. \tag{2.9}$$

Combining the latest equations we get

$$\tfrac{\partial p}{\partial t} + \rho_0 c_0^2 \ \nabla \cdot \mathbf{u} = 0,$$

and subtracting its time derivative from (2.8) we obtain

$$\frac{\partial^2 p}{\partial t^2} - c_0^2 \ \nabla^2 p = 0, \tag{2.10}$$

that is the desired wave equation. The general solution for a plane wave is given by

$$p(x, t) = f(x - ct) + g(x + ct), \tag{2.11}$$

where f and g are two arbitrary integrable functions. Respect to the spatial reference point, the first and the second term are an incoming and an outgoing wave, respectively. With few calculations it can be proved that the solution for a spherical wave is given by

$$p(d, t) = f\left(\frac{d - c_0 t}{d}\right) + g\left(\frac{d + c_0 t}{d}\right), \tag{2.12}$$

where d is the radial distance from the source. This solution is similar to the one valid for plane waves. Notice that in this case the sound pressure decreases as the square of the radius. We discuss this phenomenon in Chapter 4, making a comparison respect to the real decaying factor. As the sound source is a conventional speaker, in our work we deal with spherical waves.
A further step required before proceeding to talk about sound reflection phenomenon is the definition of the specific acoustic impedance. It is the ratio between the pressure and the particle velocity at a point

$$Z = \frac{p}{u} = \rho_0 c_0, \tag{2.13}$$

where, in this case, ρ_0 and c_0 are not quantities related to the air, but to a generic medium. The impedance can assume either real or complex values. In the second case it is defined as

$$Z = \tfrac{P}{U} = Re(Z) + jIm(Z) = |Z| \ e^{j\phi}.$$

Knowing the wave equation and the impedance quantity, we own all the elements required to completely understanding the reflection phenomenon. For simplicity, the discussion deals with the plane wave situation, but this approach is still valid even for spherical waves.
Consider the situation reported in Figure 2.1. It represents the reflection phenomenon for a plane wave normally incident respect to the interface between Medium 1 and Medium 2, which in this case are two fluids. Notice that a more interesting case for our purposes is described later, where the case of a fluid-wall interface is taken into account.

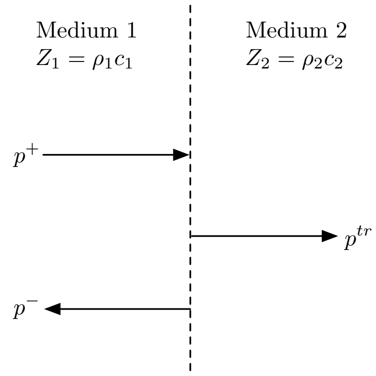


Figure 2.1: Reflection and transmission of a plane wave, normally incident to a fluid-fluid interface

From the figure, we can see that a portion of the incoming sound pressure (p^+) is reflected (p^-) in the opposite direction respect to the arriving one. The remaining part (p^{tr}) continues travelling through Medium 2. Consequently p^+ is related to the incoming wave, p^- to the reflected wave and p^{tr} to the transmitted one.
If we set our reference point ($x = 0$) at the interface between the two mediums, using the wave equation solution for plane waves (2.11) we can write

$$p^+ = p^+ \left(t - \frac{x}{c_1} \right), \tag{2.14}$$

$$p^- = p^- \left(t + \frac{x}{c_1} \right), \tag{2.15}$$

$$p^{tr} = p^{tr} \left(t - \frac{x}{c_2} \right), \tag{2.16}$$

where c_1 and c_2 are the speed of propagation in the Medium 1 and in the Medium 2, respectively. The reflection coefficient R and the transmission coefficient T are defined by

$$R = \frac{p^-}{p^+}, \tag{2.17}$$

and

$$T = \frac{p^{tr}}{p^+}. \tag{2.18}$$

As the pressure at the interface must be the same on either side, we can write

$$p^+ + p^- = p^{tr}. \tag{2.19}$$

Dividing (2.19) by p^+ and using (2.17) and (2.18), we obtain

$$1 + R = T. \tag{2.20}$$

Similarly as we have seen for the sound pressure, we must guarantee that the particle velocity is continuous at the interface. Consequently

$$u^+ + u^- = u^{tr}. \tag{2.21}$$

Rewriting (2.21) and using the impedance definition (2.13), after some calculations we get

$$R = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \tag{2.22}$$

and

$$T = \frac{2Z_2}{Z_2 + Z_1}, \tag{2.23}$$

that are the so-called *pressure* reflection and transmission coefficients. The *power* reflection and transmission coefficients are more common in use and are the ones we estimate in this work. The *power* coefficients are defined as the ratio of the reflected power W^- to the incident power W^+ in one case, and as the ratio of the transmitted power W^{tr} to the incident power, in the other one. The calculations are straightforward and after few passages we obtain

$$r \equiv \frac{W^-}{W^+} = \frac{(p^-_{rms})^2/Z_1}{(p^+_{rms})^2/Z_1} = \frac{(p^-_{rms})^2}{(p^+_{rms})^2} = R^2, \tag{2.24}$$

for the *power* reflection coefficient and

$$\tau \equiv \frac{W^{tr}}{W^+} = \frac{(p^{tr}_{rms})^2/Z_2}{(p^+_{rms})^2/Z_1} = T^2\frac{Z_1}{Z_2}, \tag{2.25}$$

for the *power* transmission coefficient. We can show that the conservation of energy principle is fulfilled, in fact

$$\tau + r = 1. \tag{2.26}$$

Another common quantity used in practice is the transmission loss TL, expressed in decibels and defined as

$$TL \equiv -10 \log_{10} \tau. \tag{2.27}$$

In the latter part we have analyzed the reflection phenomenon for a plane wave normally incident to the interface between two fluids. A more interesting case for our purposes is represented by the one in which we are in the same incident condition, but the interface is between a fluid and a wall. The situation is sketched if Figure 2.2.

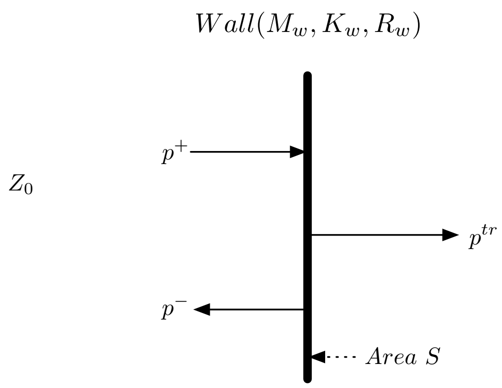


Figure 2.2: Reflection and transmission of a plane wave, normally incident to a fluid-wall interface

This discussion is valid, for example, for a wall fixed at the edges surrounded by the air. The wall is characterized by its mechanical mass M_w, stiffness K_w and damping resistance R_w. A good starting point is the Newton's law that describes the motion of the wall

$$M_w \dot{u}_w = -R_w u_w - K_w \xi_w + S(p^+ + p^- - p^{tr}), \tag{2.28}$$

where $\xi_w = \int u_w dt$ is the wall displacement. Introducing the wall characteristics expressed *per-unit-area*

$$m_w = M_w/S, \;\; r_w = R_W/S, \;\; k_w = K_w/S, \tag{2.29}$$

and substituting them in (2.28), we get

$$m_w \dot{u}_w + r_w u_w + k_w \int u_w dt = p^+ + p^- - p^{tr}. \tag{2.30}$$

Notice that if the wall has no mass, resistance and stiffness, equation (2.29) is the same of the (2.19) that is valid in the case of a fluid-fluid interface. Proceeding as we have done for the latter case, the equation (2.19) that describes the continuity of particle velocity is still valid. In addition to this, also the wall particle velocity u_w is equal to u^{tr}. Now, assuming the incident wave to be time harmonic, we can rearrange equation (2.29) as

$$(j\omega m_w + r_w + k_w/j\omega)\ U_w = P^+ + P^- + P^{tr}, \tag{2.31}$$

where the upper case for u and p represents their time harmonic component. Since

$$U_w = U^{tr} = \frac{P^{tr}}{Z_0}, \tag{2.32}$$

equation (2.30), after a division by P^+, can be rewritten as

$$1 + R = (1 + Z_w/Z_0)\ T, \tag{2.33}$$

where $Z_w = (j\omega m_w + r_w + k_w j\omega)$ is the specific acoustic impedance of the wall, and Z_0 is the characteristic impedance of the fluid that surrounds the wall. Since Z_0 is the same on both sides of the wall, equation (2.19) can be reduced to

$$1 - R = T. \tag{2.34}$$

Combining equations (2.32) and (2.33) we get the solution for

$$T = \frac{2Z_0}{(2Z_0 + r_w) + j(\omega m_w - k_w/\omega)} \tag{2.35}$$

and

$$R = \frac{r_w/2Z_0 + (j\omega m_w/2Z_0)[1 - (\omega_0/\omega)^2]}{(1 + r_w/2Z_0) + (j\omega m_w/2Z_0)[1 - (\omega_0/\omega)^2]}, \tag{2.36}$$

where $\omega_0 = \sqrt{k_w/m_w}$ is the resonance frequency of the wall.
The last case we take into account is the one in which the wavefront doesn't travel in a normally incident path respect to the wall, but in an oblique one. Notice that this case is included only for completeness, as our system works in a source-wall perpendicular situation only.
When the angle of incidence is not normal we could encounter the refraction acoustic phenomenon, which is described as the variation of a wave travelling path, when the wave passes through two different mediums. It is formalized by the Snell's law. Assuming that the incident

ray lies in Medium 1, and the transmitted ray in Medium 2, we can write

$$\frac{\sin\theta_t}{\sin\theta_i} = \frac{c_2}{c_1}, \tag{2.37}$$

where c_2 and c_1 are the speeds of sound propagation in each of the mediums, whereas θ_t and θ_i are the angles of the transmitted and the incident rays.
The panel is assumed to be flexible but not stiff, and the signals are supposed to be time harmonic. The situation is the one presented in Figure 2.3.

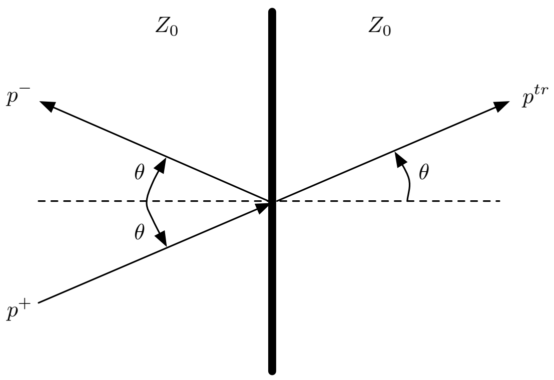


Figure 2.3: Reflection and transmission of a plane wave, travelling with oblique incidence respect to a fluid-wall interface

The angle of the transmitted wave is the same of the incident one, due to the Snell's law and to the fact that the fluid is the same on both sides of the wall. The starting point for the discussion is again the Newton's law. The passages to get the results are the same of the latter cases, so we report only the final equations for

$$R = \frac{j\omega m_w \cos\theta/2Z_0}{1 + j\omega m_w \cos\theta/2Z_0}, \tag{2.38}$$

and

$$T = \frac{1}{1 + j\omega m_w \cos\theta/2Z_0}. \tag{2.39}$$

Notice that all the cases presented up to now deal with thermally non-conducting mediums. Instead, if a medium is thermally conducting part of the kinetical energy of the wave is converted into heat, due to impedance mismatch. This phenomenon, together with the absorption within the material and the wavefront damping, is the way in which a material absorbs energy. Up to now we have talked about sound energy reflected and transmitted by a wall, but nothing about the energy absorbed in the wall structure. Introducing this new quantity and applying the energy conservation law, we can write

$$W^+ = W^- + W^{tr} + W^{abs}, \tag{2.40}$$

where $W^+, W^-$and W^{tr} are well known quantities, whereas W^{abs} is the energy absorbed by the wall. The situation is depicted in Figure 2.4.

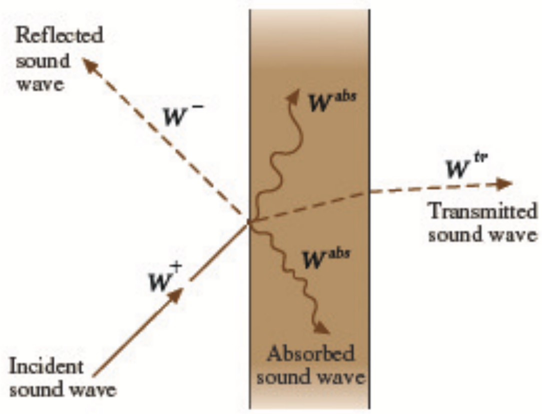


Figure 2.4: Interaction of the sound wave energy with a surface [1]

In real world applications we are interested in the portion of the energy that a surface doesn't reflect. In equation (2.40) this quantity is represented by W^{tr} and W^{abs}, that are quantities difficult to estimate both empirically and theoretically. Consequently, what is done in the reality is to measure the energy associated to the incident and the reflected signal, obtaining the others with some straightforward calculations. Dividing equation (2.40) by W^+, we get

$$1 - \frac{W^-}{W^+} = \frac{W^{tr}}{W^+} + \frac{W^{abs}}{W^+}, \tag{2.41}$$

where the right part of the equation is the quantity we want to know. It is the portion of energy the wall doesn't reflect. If we introduce the *power* absorption coefficient, defined by

$$\alpha = \frac{W^{tr}}{W^+} + \frac{W^{abs}}{W^+}, \tag{2.42}$$

recalling the definition of *power* reflection coefficient of equation (2.24), and rearranging equation (2.41), we can write

$$\alpha = 1 - r, \tag{2.43}$$

or in an equivalent way

$$\alpha = 1 - R^2. \tag{2.44}$$

The values of α and r are limited in the [0,1] range. If a material has an absorption coefficient $\alpha = 1$ it can be referred as a totally absorbing, whereas if $\alpha = 0$, it can be referred as a totally reflecting. Appendix A reports the absorption coefficients for some common materials at the standard octave frequency bands.
The absorption coefficient varies with the frequency of the incident ray and the angle of incidence, as we have seen before. As anticipated, this work estimates α only as function of the frequency. The angle dependency is not taken into account, since the source has to be normal respect to the plane. Moreover, in this way we can neglect the refraction phenomenon.

2.2 State of the art

This section covers some of the most significant methods already present in the literature, for the sound absorption coefficients estimation.
The first method we take into account is the Sabine's one. Let's start from the Sabine equation, defined by

$$T = \frac{0.161V}{\sum_i S_i \alpha_i}, \tag{2.45}$$

where T represents the number of seconds that the intensity of the sound takes to drop by 60 dB, V(expressed in m^3) is the volume of the room in which the experiment is taken, S_i (expressed in m^2) is the area of the i^{th} surface of the room, and α_i is the power absorption coefficient for the i^{th} surface. If we explicit the previous equation as function of $A = \sum_i S_i \alpha_i$, we get

$$A = \frac{0.161V}{T}. \tag{2.46}$$

Defining

$$\bar{\alpha} = \frac{\alpha_1 S_1 + \alpha_2 S_2 + \ \ldots \ + \alpha_i S_i}{S_1 + S_2 + \ \ldots \ + S_i} = \frac{A}{S} \tag{2.47}$$

as the average absorption coefficient, knowing the value of the reverberation time T and the geometry of room, we can estimate this value. The equation (2.45) is purely empirical; nevertheless, in some highly controlled situations it can be useful to estimate the absorption coefficient for a surface. Assume that a room has a regular and known

geometry and that all the surfaces have an absorption coefficient that is known, as well. If we cover an entire surface of the room (i.e. the floor or a wall) with the material under test, and if we are able to measure the reverberation time T, we can estimate the absorption coefficient for that material, since it is the only unknown in equation (2.46). In the case that we know the geometry of the room, but nothing about the absorption coefficients of the surfaces, we can at least estimate $\bar{\alpha}$. The pro of this method is that it is very simple. The cons are that it can be used only in highly controlled situations and in rooms with a regular geometry. Moreover, the equation from which it is derived from neglects all the most important sound propagation phenomena. Even in the cases for which the method provides a result, it doesn't carry information about frequency or incident angle dependency.
A more interesting case is represented by the normal incidence methods. As the name suggests, these methods are able to estimate the absorption coefficients only for normally incident sound waves. These techniques are all based on the Kundt tube method. The device adopted to realize the measurements is reported in Figure 2.5 and it is referred as Kundt tube, from the name of the inventor August Kundt.

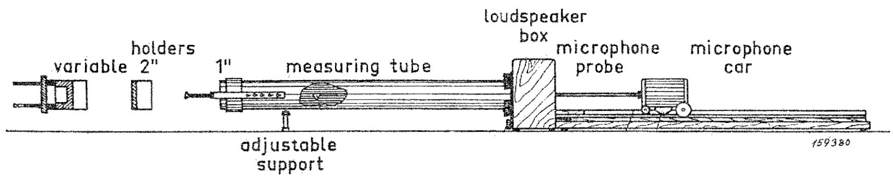


Figure 2.5: Kundt tube [2]

At one end of the tube there is a portion of the material under test, positioned on an appropriate holder, if required. At the opposite end there are the speaker and the microphone support. This last one is made up of a *microphone car* that by means of a rigid support allows the microphone probe to be placed everywhere inside the tube. The operating principle is based on the standing waves that the speaker leads inside the tube.
We encounter the standing waves phenomenon when two waves with the same frequency travels in opposite directions in an enclosure or on a fixed-edges support (i.e. strings and membranes). Under this condition the wave pattern does not advance in the space. The nodes are points of zero amplitude in every time instant, and they occur

spaced half-wavelengths apart. Instead, antinodes are points of maximum amplitude, spaced halfway between the nodes. Some standing wave patterns are reported in Figure 2.6. Notice that the x-axis refers to the space. We define the standing wave ratio as

$$SWR \equiv \frac{|P_{max}|}{|P_{min}|}, \tag{2.48}$$

where $|P_{max}|$ and $|P_{min}|$ are respectively the maximum and the minimum value of $|P|$ in the tube. The maximum value of $|P|$ occurs when the incident and the reflected wave are perfectly in phase. Similarly, the minimum occurs when the two phasors are 180° out of phase. These values can be measured moving the microphone probe inside the tube, while the source is emitting the sinusoidal wave at the desired frequency.

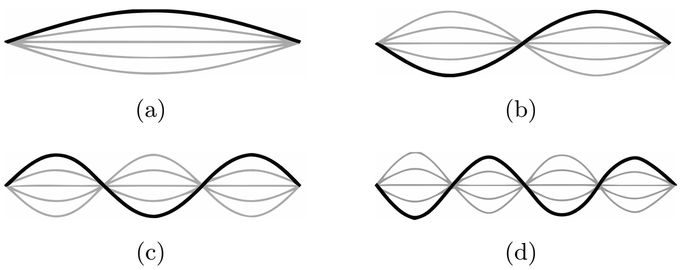


Figure 2.6: Standing wave patterns for signals at different frequencies.

Assuming the signals time harmonic and sinusoidal, it can be proved that the *pressure* reflection coefficient is given by

$$R = \frac{SWR - 1}{SWR + 1}\, e^{j2(kd)_{max}}, \tag{2.49}$$

or

$$R = -\frac{SWR - 1}{SWR + 1}\, e^{j2(kd)_{min}}, \tag{2.50}$$

where $k = \omega/c_0$ is the wavenumber and d is the distance from the material termination. Knowing R, the calculation of α is straightforward. Both the previous equations are frequency dependent, due to the presence of k in their respective solutions. The first equation has to be used when the extremum closest to the material termination of the tube is a maximum, whereas the second one has to be used in the opposite situation. The pros of the Kundt method are that it provides appreciable results and that it is relatively simple to implement.

The cons are that it can only provide results for a single frequency at time and that it constitutes a laboratory method. Since it requires a small portion of material, in certain condition it can be classified as a destructive method. Nonetheless some more sophisticated techniques take the cue from it, such as the so-called *two-microphone* method. In this case the pressure signals are measured by two fixed microphones, instead of only one microphone probe. Knowing the distances between the microphones and the surface, it is possible to estimate the reflection coefficient. A modified version of this technique, using the transfer function between the two microphones, yields the reflection coefficient as a function of frequency.

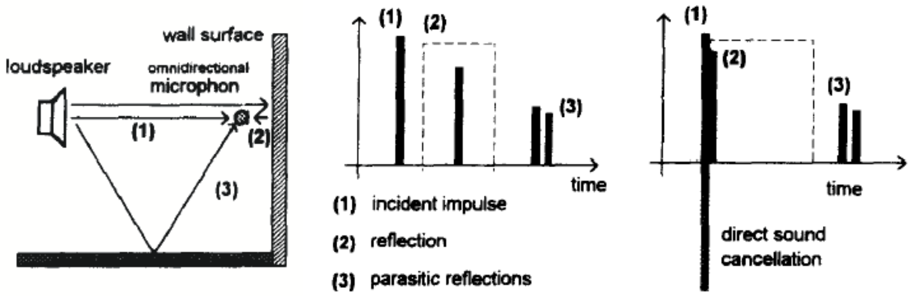
We have also to mention the standard methods ISO 10534-1 [9] and ISO 10534-2 [10]. The first one is the standardized version of what we have proposed in this discussion, except some minor differences. The second one is similar too, but it uses wideband signals as sources and the *two-microphone* method.

Instead, the Mommertz [3] method is specifically designed to measure the sound absorption of surfaces under real conditions. The functioning principle of such a system is the so-called subtraction technique, combined with the pre-equalization of the source signal. This procedure allows to compensate the linear distortions mainly caused by the loudspeaker, but also by the measurement instrument cascade. A microphone-loudspeaker system is positioned in front of the surface (Figure 2.7(a)). The environment is stimulated by a maximum length sequence and the impulse response is measured. The first step is to cancel the direct sound and the reverberations, windowing them out (Figure 2.7(b)). The second step is to cancel the portion of the direct signal that is still present in the reflected one (Figure 2.7(c)). This operation can be done only if the source signal is known by shape, amplitude and time delay. Lastly, the FT (Fourier Transform) of the *cleaned* reverberation signal and of the incident one are calculated. The ratio between the absolute value of these quantities gives us the *pressure* reflection coefficient $R(f, \Theta)$, where f if the frequency and Θ is the incidence angle.

Also in this case the calculation of α is straightforward, in fact we get

$$\alpha(f, \Theta) = 1 - |R(f, \Theta)|^2. \tag{2.51}$$

The strength of this system is that it can estimate the reflection coefficients as function of the incident angle. In fact, using this method



(a) (b) (c)

Figure 2.7: Functioning principle of the Mommertz method [3]

we have no loudspeaker-surface positioning limitations. Since the subtraction technique is a tricky operation, the system limits the problems that could arise, introducing a pre-equalization of the source signal. It is a pre-equalization of the MLS source signal, based on the overall system frequency response. The pros of the Mommertz method are that it can estimate the absorption coefficients as function of the frequency and the angle of incidence, and that it is an *in situ* method. The cons are that it is not really robust against the reverberations, and that sometimes the subtraction technique could fail. In fact, the windowing operation described in Figure 2.7(b) works only when the reverberations are distant enough in time from the reflected signal, but under real conditions usually it happens they are mixed together. Moreover the subtraction technique works only if the information gathered for the incident wavefront is coherent with the one we have. This method is, however, the base for more sophisticated solutions including the work proposed by the Adrienne [6] research team. The solutions of this work have been standardized in the ISO 13472-1 [11] normative.
The last method we are going to present is the one from which the work of this thesis takes the cue from. It is the solution for the coefficients estimation problem, proposed by Ducorneau et al [7]. For obvious reasons the functioning principle is the same of our work. It proposes to estimate the absorption coefficients directly measuring the energy of the incident and the reflected signals. For this reason it is also classified as an *intensimentric* method. Taking the frequency-by-frequency ratio of the absolute values of the quantities cited above, we get the desired result. It is an estimation of the absorption coefficients as function of the frequency.

The strength of this solution is the idea to acquire the direct and the reflected signals using a high directivity microphone array that allows to be robust against the reverberation phenomena. Consequently it is the ideal solution for *in situ* measurements.
Since the overall system is presented in Chapter 3 we focus only on the main difference between the method presented in our work and the one introduced by Ducorneau et al [7]: the acquiring device. The Ducorneau's solution uses a multipolar weighting microphone array. Such a device allows to obtain a narrow beamform, but the one implemented in our system gets a narrower beamform, providing at the same time the maximum front-to-back ratio, as showed in the next section.

2.3 Differential microphone arrays

The core of the work presented in this thesis is the use of a superdirectional microphone array. It is a particular type of differential array. The discussion presented in this section describes the theory required to realize such a device and it explains the way to obtain a maximum front-to-back ratio, during the signal acquisition. Instead, the array implementation is discussed in Chapter 3.
A first order differential microphone array is a device whose response is proportional to a zero-order acoustic pressure signal and to the first-order spatial derivative of the acoustic pressure field. Generalizing, an n^{th} order differential microphone array is a device whose response is proportional to a zero-order acoustic pressure and to all the spatial derivatives of the pressure field, up to order n.
This microphone array class has a higher directivity respect to the one that uses the *wighted delay-and-sum* technique.
The spatial derivatives are calculated by means of the finite-differences approximation. Later we show the error introduced by this approximation. The starting point for the discussion is the wave equation for a plane wave travelling to the array. We assume the signal time harmonic and given as function of the wavelength k, distance l and time t. We get

$$p(k, l, t) = P_0 e^{j(\omega t - kl\cos\theta)}, \tag{2.52}$$

where P_0 is the plane wave amplitude, ω is the angular frequency and θ is the wave angle of arrival. Dropping the time dependence and taking the n^{th} order spatial derivatives, we get

$$\frac{\partial^n p(k,l)}{\partial l^n} = P_0(-jk\cos\theta)^n \; e^{-jkl\cos\theta}. \tag{2.53}$$

The output of a differential microphone array is the linear combination of spatial derivatives, such as the one reported in the previous equation and the overall response is a weighted sum of terms of the form $\cos^n\theta$. Analyzing the first order spatial derivative and the one obtained by the finite differences, we want to estimate the error introduced by this approximation. The first equation required for the comparison is extracted from (2.53), dropping the n power, since we are considering the first order case. Instead the second equation, which is the one that describes the finite-difference approximation, is given by

$$\begin{aligned}\frac{\Delta p(k,l,\theta)}{\Delta l} &= \frac{p(k,l+d/2,\theta) - p(k,l-d/2,\theta)}{d} \\ &= \frac{-j2P_0\sin\left(kd/2\cos\theta\right)e^{-jkl\cos\theta}}{d},\end{aligned} \tag{2.54}$$

where d is the spacing between the microphones. Dividing the equation obtained from (2.53) (with $n = 1$) and equation (2.54), we get the amplitude bias error introduced by the approximation. It is defined as

$$\epsilon = \frac{\sin kd/2}{kd/2} = \frac{\sin \pi d/\lambda}{\pi d/\lambda}. \tag{2.55}$$

The result is plotted in Figure 2.8. We can infer that the microphone spacing has to be less than 1/4 of the wavelength, in order to have a bias error less than 1 dB.

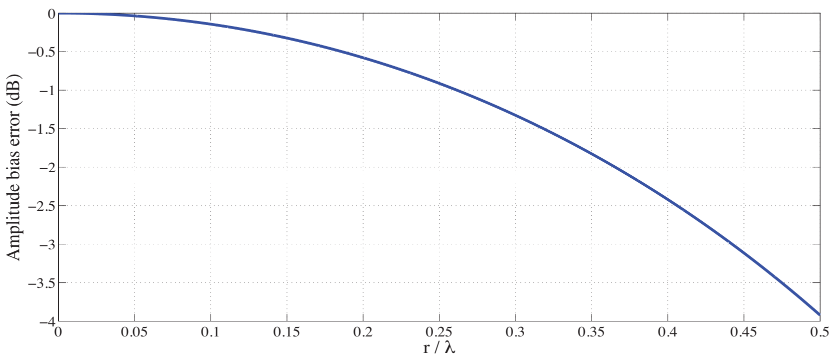


Figure 2.8: Amplitude bias error due to finite differences approximation

We start considering the first order differential array sketched in Figure 2.9. The plane wave arriving to the array produces an output that can be written as

$$E(k,\theta) = P_0 \left(1 - e^{-jkd\cos\theta}\right), \tag{2.56}$$

where the time dependence is neglected for simplicity. Introducing an arbitrary time delay between the two subtracted zero-order microphone responses, we get

$$E(\omega,\theta) = P_0 \left(1 - e^{-j\omega(\tau + d\cos\theta/c)}\right), \tag{2.57}$$

where τ is the applied time delay. For small microphone spacing condition, $kd \ll \pi$ and $\omega\tau \ll \pi$, equation (2.57) can be approximated by

$$E(\omega,\theta) \approx P_0\omega \left(\tau + d/c\cos\theta\right). \tag{2.58}$$

Defining a_0, $a_1$and ζ_1 such that

$$\zeta_1 = a_0 = \frac{\tau}{\tau + d/c}, \tag{2.59}$$

$$1 - \zeta_1 = a_1 = \frac{d/c}{\tau + d/c}, \tag{2.60}$$

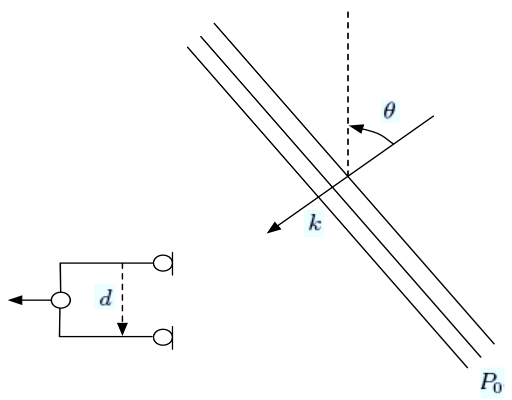


Figure 2.9: First order microphone array

and substituting the previous quantities in equation (2.58) we get

$$E_N(\theta) = a_0 + a_1\cos\theta = \zeta_1 + (1 - \zeta_1)\cos\theta, \tag{2.61}$$

where E_N, is the normalized response of the array. The most important thing to notice about this result is that the response is frequency independent. Moreover, properly tuning τ and so ζ_1, we get the desired directivity response. The value of τ is also proportional to the propagation time of a wave that travels axially between zero-order microphones. For a first order differential microphone array, the definition of τ is given by

$$\tau = \frac{da_0}{ca_1} = \frac{d\zeta_1}{c(1-\zeta_1)}. \tag{2.62}$$

The directivity response for an n^{th} order array, starting from equation (2.57) can be written as

$$E(\omega, \theta) = P_0 \prod_{i=1}^{n} \left[1 - e^{-j\omega(\tau_i + d/c\cos\theta)}\right], \tag{2.63}$$

where τ_i is the i^{th} chosen time delay. Proceeding as we have done for the first order case, assuming small spacings between microphones, we can write

$$E(\omega, \theta) \approx P_0 \omega^n \prod_{i=1}^{n} [\tau_i + d/c\cos\theta]. \tag{2.64}$$

Defining the i^{th} tunable coefficient ζ_i as

$$\zeta_i = \frac{\tau_i}{\tau_i + d/c}, \tag{2.65}$$

equation (2.64) becomes

$$E(\omega, \theta) \approx P_0 \omega^n \prod_{i=1}^{n} [\zeta_i + (1-\zeta_i)\cos\theta]. \tag{2.66}$$

Finally, expanding the previous equation we get

$$E(\omega, \theta) = P_0 \omega^n (a_0 + a_1 \cos\theta + a_2 \cos^2\theta + \ \dots \ + a_n \cos^n\theta). \tag{2.67}$$

We can notice that the directivity response is proportional to τ_i, d and ω^n. The frequency dependence to the ω^n term can be easily compensated by a low-pass filter whose frequency response is proportional to ω^{-n}. This is an appreciable result since the directivity response, after the previous compensation, is frequency independent. Normalizing equation (2.67), neglecting all the constant terms, and assuming frequency independence, we get

$$E_N(\theta) = a_0 + a_1 \cos\theta + a_2 \cos^2\theta + \ \dots \ + a_n \cos^n\theta, \tag{2.68}$$

or equivalently

$$E_N(\theta) = \prod_{i=1}^{n} [\zeta_i + (1 - \zeta_i) \cos\theta]. \tag{2.69}$$

This is the normalized directivity response for a generic n^{th} order differential microphone array.

Up to now we have seen what is the directivity response for a differential microphone array, but nothing about the way to choose the tunable parameters τ_i, or indirectly ζ_i. As shown in equation (2.69) ζ_i are the roots of equation (2.68).

We now introduce a method to find the values of such tunable parameters, which realizes a maximum front-to-back ratio that is similar to the one implemented in our work. In this situation, the directional gain for the signals propagating in front of the microphone array is maximum respect to the one for the signals in the rear. Let's introduce the front-to-back ratio, defined as

$$F(\omega) = \frac{\int_0^{2\pi} \int_0^{\pi/2} |E(\omega, \theta, \phi)|^2 \sin\theta \; d\theta d\phi}{\int_0^{2\pi} \int_{\pi/2}^{\pi} |E(\omega, \theta, \phi)|^2 \sin\theta \; d\theta d\phi}, \tag{2.70}$$

where θ and ϕ are the spherical coordinate angles and $E(\omega, \theta, \phi)$ is the pressure response. If the microphones used in the implementation are axisymmetric, (2.70) can be rewritten as

$$F(\omega) = \frac{\int_0^{\pi/2} |E(\omega, \theta, \phi)|^2 \sin\theta \; d\theta}{\int_{\pi/2}^{\pi} |E(\omega, \theta, \phi)|^2 \sin\theta \; d\theta}. \tag{2.71}$$

Converting the integrals with finite summations and substituting $E(\omega, \theta, \phi)$ with the one provided in equation (2.68), we get

$$F(a_0, \; \dots \; , a_n) = \left[\sum_{i=0}^{n} \sum_{j=0}^{n} \frac{a_i a_j}{1 + i + j} \right] \left[\sum_{i=0}^{n} \sum_{j=0}^{n} \frac{(-1)^{i+j} a_i a_j}{1 + i + j} \right]^{-1} \tag{2.72}$$

or in a more compact way

$$F(\mathbf{a}) = \frac{\mathbf{a}^T \mathbf{B} \mathbf{a}}{\mathbf{a}^T \mathbf{H} \mathbf{a}}, \tag{2.73}$$

where $\mathbf{a}$ is the weighting vector, $\mathbf{H}_{i,j} = \frac{(-1)^{i+j}}{1+i+j}$ is a Hankel matrix and $\mathbf{B}_{i,j} = \frac{1}{1+i+j}$ is a special case of a Hankel matrix, built as a Hilbert matrix. The eigenvector corresponding to the maximum eigenvalue is

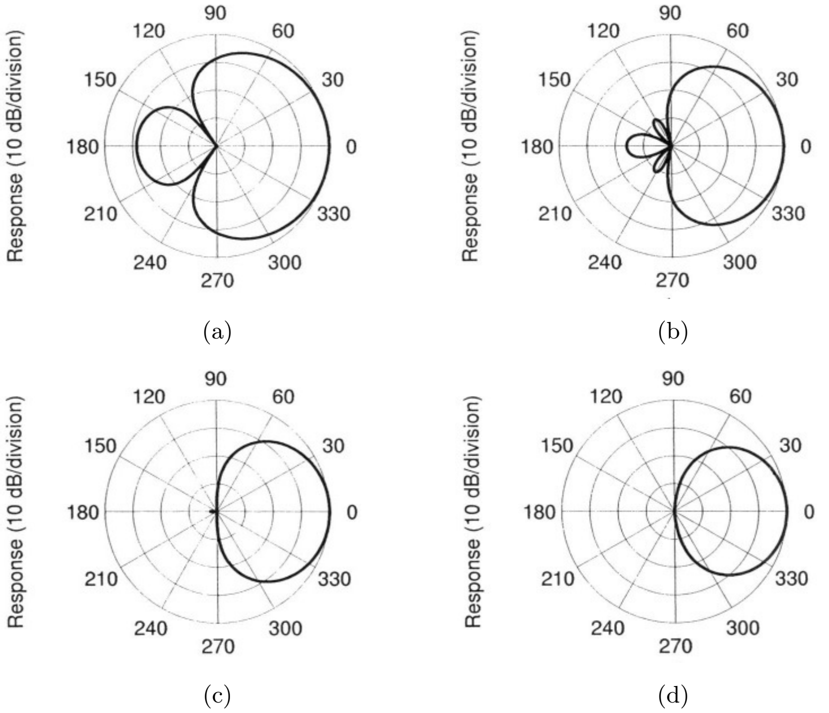


Figure 2.10: Directivity response for maximum front-to-back ratio for (a) first, (b) second, (c)third, (d) fourth order differential microphone arrays [4]

the solution of the discussed problem. As the matrices **H** and **B** are real and positive definite, the eigenvalue and the eigenvector are real, too. Figure 2.10 shows the directivity response for differential arrays up to the fourth order, with a maximum front-to-back ratio.
Table 2.3 summarizes the performances of these differential arrays. As we can see, the beamwidth becomes narrower as the order of the array increases. In the same way, the front-to-back ratio increases as the differential order gets higher.

Mic. Order	F (dB)	Beamwidth (degs)	Null(s) (degs)
1^{st}	11.4	115	125
2^{nd}	24.0	80	104, 114
3^{rd}	37.7	65	97, 122, 154
4^{th}	51.8	57	94, 111, 133, 159

Table 2.1: Maximum front-to-back power ratio for differential arrays

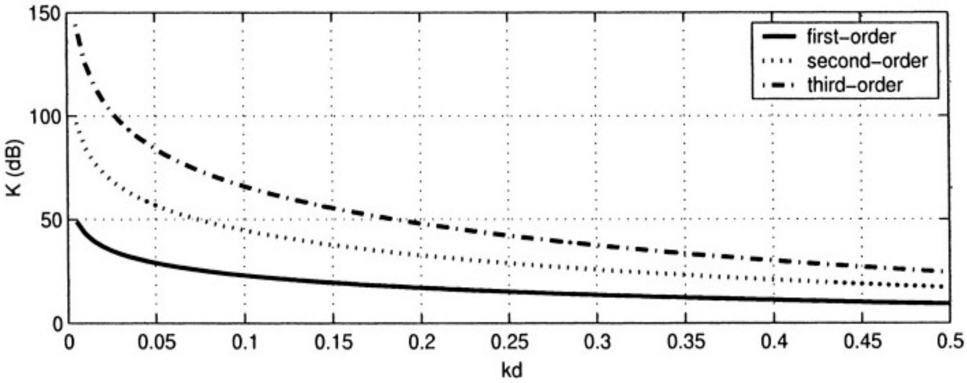


Figure 2.11: Array sensitivity error [4]

Appendix B reports the numerical solutions for the above cases. Increasing the order of the differential arrays we achieve better results, but we could encounter problems related to the sensitivity to electronic noise, and to microphone mismatching. To quantify this issue, we introduce a quantity that characterizes the sensitivity of the array to random amplitude and position errors. It is the so-called sensitivity function and it is defined as

$$K = \frac{\sum_{m=1}^{n} |b_m|^2}{|\sum_{m=1}^{n} b_m e^{-jk(r_m + c\tau_m)}|^2}, \tag{2.74}$$

where m is the microphone index, τ_m is the delay associated with the microphone m and b_m are the amplitude shading coefficients. In the case of differential arrays, assuming $kd \ll 1$, the previous equation reduces to

$$K \approx \frac{n+1}{[\prod_{m=1}^{n} k(d + c\tau_m)/2]^2}. \tag{2.75}$$

Notice that the sensitivity error is inversely proportional to the frequency response of the array. Figure 2.11 shows the values of K as function of the product of the wavelength by the microphone spacing. As we expect, the array sensitivity error is greater for higher order differential microphone arrays.

2.4 Spectral estimation

In the implementation of the work presented by this thesis, we have to consider the linear distortion introduced by the loudspeaker and by

the cascade of the measuring instruments. These undesired phenomena could affect in a significant way the quality of the impulse responses we obtain from the measured signals. Having good impulse responses means to have them shorter as possible. In order to compensate the effects of the linear distortion we have to pre-equalize the source signal using the inverse frequency response of the system. In such a way we aim to obtain a real frequency response as flat as possible.
The system characterization is realized by means of the so-called spectral estimation process. In the literature there are mainly two types of spectral estimation: non-parametric and parametric. The non-parametric approach makes no assumption on the model of the system under analysis. Instead, the parametric method assumes that the signal under test is generated by a system with a known functional form. In this discussion we present the latter approach, since it is the one chosen for the implementation of our work.
The parametric spectral estimation allows to approximate any power spectral density (PSD), provided that it is continuous and that we have enough information about the studied signal. The PSD is the quantity that we want to get. We define the PSD estimator as

$$\phi(\omega) = \left| \frac{B(\omega)}{A(\omega)} \right|^2 \sigma^2, \tag{2.76}$$

where

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ \dots \ + a_n z^{-n}, \tag{2.77}$$

$$B(z) = 1 + b_1 z^{-1} + b_2 z^{-2} + \ \dots \ + b_n z^{-n}. \tag{2.78}$$

and

$$A(\omega) = A(z)|_{z=e^{j\omega}}. \tag{2.79}$$

The PSD estimator can be associated with a signal obtained by filtering a white noise $e(n)$ of power σ^2, through a filter with transfer function $H(\omega) = B(\omega)/A(\omega)$. Consequently, equation (2.76) can be rewritten as

$$Y(z) = H(z)E(z) = \frac{B(z)}{A(z)}E(z). \tag{2.80}$$

Rearranging equation (2.80), we get

$$A(z)Y(z) = B(z)E(z). \tag{2.81}$$

Notice that the values associated to m and n are *tunable* parameters of the estimation process. Their choose is quite difficult and it depends

on the application we are interested in. Moreover, from the values of these parameters depends the model chosen for the estimation.
For $m = 0$ and $n \neq 0$ we have an AR (autoregressive) model, described by

$$A(z)Y(z) = E(z). \tag{2.82}$$

For $m \neq 0$ and $n = 0$ we have a MA (moving average) model, described by

$$Y(z) = B(z)E(z). \tag{2.83}$$

Lastly, for $m \neq 0$ and $n \neq 0$ we have an ARMA (autoregressive moving average) model, described by the generic equation (2.81). The AR model is the most frequently used in applications, and it allows to describe well also smoothly varying signals. For this reason, it is the one we choose for our implementation. Notice that in this case, for stability, $A(z)$ must have all its zeros inside the unit circle.
Starting from the generic ARMA model equation (2.81), and rewriting it as function of the parameters $\{a_i\}_{i=1}^n$ and $\{b_j\}_{j=1}^n$, we get

$$y(t) + \sum_{i=1}^{n} a_i y(t-i) = \sum_{j=0}^{m} b_j e(t-j). \tag{2.84}$$

Multiplying the previous equation by $y^*(t-k)$ and taking the expectation we get

$$\begin{aligned} r(k) + \sum_{i=1}^{n} a_i r(k-i) = \sum_{j=0}^{m} E\{e(t-j)y^*(t-k)\} &= \sigma^2 \sum_{j=0}^{m} b_j h_{j-k}^* \\ &= 0, \end{aligned} \tag{2.85}$$

for $k > m$. Notice that $r(k) = E\{y(t)y^*(t-k)\}$ is the autocovariance sequence (ACS). For an AR model m is equal to 0, so combining the set of equations obtained from the (2.85) for $k > 0$, and the one obtained from the (2.85) for $k = 0$, we get a system of linear equations, called Yule-Walker equations. After some rearrangements we get

$$\begin{bmatrix} r(1) \\ \vdots \\ r(n) \end{bmatrix} + \begin{bmatrix} r(0) & \cdots & r(-n-1) \\ \vdots & \ddots & \vdots \\ r(n-1) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{2.86}$$

or in a more compact way

$$\mathbf{r}_n + \mathbf{R}_n \boldsymbol{\phi} = \mathbf{0}, \tag{2.87}$$

where the vector $\boldsymbol{\phi} = [a_1 \ \dots \ a_n] = \mathbf{R}_n^{-1}\mathbf{r}_n$ contains the desired solution. Once we get the FIR (Finite Impulse Response) filter coefficients that approximate the PSD, we are able to obtain the inverse filter, referred as *whitening filter*. Filtering the source signal by the inverse filter we pre-equalize the signal in order to compensate the linear distortion phenomenon.

2.5 Wavelet transform

When we are going to process a signal coming from the real world, it is necessary to find a representation that models the characteristics we are interested in. After the sampling and quantization process, the acquired time-discrete signal is requested to be represented also in the frequency domain. The mathematical tool that allows to convert a time-domain representation into the frequency one, is the so-called Fourier Transform (FT) or, in the particular case of discrete signals, the Discrete Fourier Transform (DFT). The basic idea behind the DFT is to decompose the input signal and to represent it as a function of complex harmonics. The obtained representation is totally in the frequency domain, in the sense that we lose any temporal information related to a certain frequency. Obviously, the DFT is invertible by means of the Inverse Discrete Fourier Transform (IDFT). If we are interested in the time or frequency domain in two separate instances, the DTF is the choice that suits for us. Instead, if we are interested in a representation that is a compromise between the time domain and the frequency one, the DFT is not enough since, after the transform, the signal is totally represented in the frequency domain. This strictly domain-related representation is the great limitation of the DFT. Since our system needs this kind of *mixed* representation, as described in Chapter 3, we need to find an efficient solution that meets our requirements. Such a solution is given by the Wavelet Transform (WT), specifically designed for a mixed-approach of the signals analysis. As we are dealing with discrete signals, also in this case we introduce the discrete version of the WT, the DWT.

Using the WT, the time-frequency localization is realized by a set of base functions, defined as

$$\frac{1}{\sqrt{s}}\,\psi\left(\frac{t-u}{s}\right), \tag{2.88}$$

where $u \in \mathbb{R}$ is the time shifting parameter, $s \in \mathbb{R}^+$ is the scaling parameter, and ψ is the wavelet base. The wavelet base must satisfy the following requirements:

$$\int_{-\infty}^{+\infty} |\psi(t)| \; dt < \infty, \tag{2.89}$$

and

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 \; dt < \infty, \tag{2.90}$$

that means it must be absolutely and square integrable. Moreover it must be zero average and normalized. In mathematical terms, we get

$$\int_{-\infty}^{+\infty} \psi(t) \; dt = 0, \tag{2.91}$$

and

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 \; dt = 1. \tag{2.92}$$

There exist a large number of wavelet bases, and their choice depends on the application requirements. The WT for a continuous signal $x(t)$ is defined by

$$W_x(s,u) = \int_{\mathbb{R}} x(t) \; \frac{1}{\sqrt{s}} \; \psi\left(\frac{t-u}{s}\right) dt. \tag{2.93}$$

In the discrete case, the input signal is denoted by $x[n]$, and the scaling and time-shifting parameters are usually defined as

$$s = 2^{-j} \tag{2.94}$$

and

$$u = ka = k2^{-j}, \tag{2.95}$$

where $j, k \in \mathbb{Z}$. This configuration is the so-called dyadic discrete wavelet transform. We obtain a mixed time-frequency signal representation with multiresolution, such as the one reported in Figure 2.12. Each time-frequency rectangle has a variable area of $2^j \times 2^{-j}$. This means that we can choose with which degree of detail we want to describe each portion of the signal. Applying some wavelet bases to the dyadic configuration we obtain a set of scaled and shifted bases that are orthogonal, avoiding the redundancy in the time-frequency description of the signal.

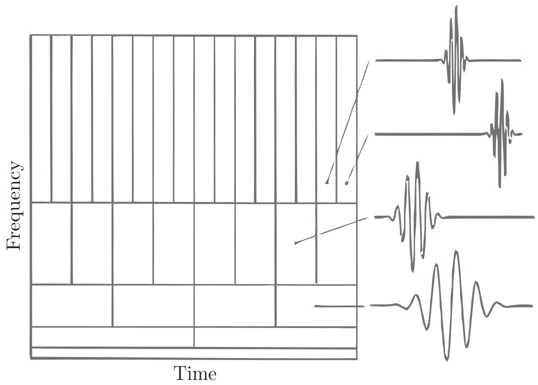


Figure 2.12: Time-frequency multiresolution, achieved by DWT

It can be proved that the calculation of the WT in the discrete case, with a dyadic configuration, can be viewed as the decomposition of the input signal into low-pass and hi-pass components, subsampled by a factor 2. The Inverse Discrete Wavelet Transform (IDWT) performs the reconstruction in a dual way. The DWT and the IDWT are respectively represented in Figure 2.13(a) and Figure 2.13(b).

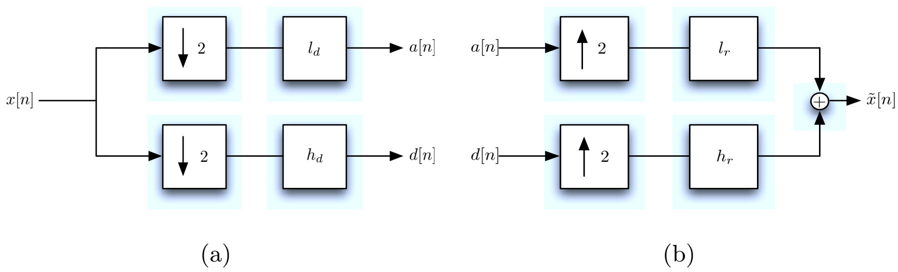


Figure 2.13: One level DWT(a) and IDWT(b) process

The low-pass components are usually referred as approximation coefficients $a[n]$, whereas the high-pass ones as detail coefficients $d[n]$. The filters used in the decomposition are the wavelet filters, directly obtained from the mother wavelet. We achieve prefect reconstruction if at the end of the IDWT we get $x[n] = \tilde{x}_n$. In order to get this result the wavelet filters must fulfill the Quadrature Mirror Filter (QMF) conditions, defined as

$$L_d^*(\omega + \pi)L_r(\omega) + H_d^*(\omega + \pi)H_r(\omega) = 0, \tag{2.96}$$

and

$$L_d^*(\omega)L_r(\omega) + H_d^*(\omega)H_r(\omega) = 2, \tag{2.97}$$

where * denotes the complex conjugate, whereas L_d, H_d, L_r and H_r represent respectively the FT of the decomposition wavelet low-pass and high-pass filters, and the FT of the reconstruction wavelet low-pass and high-pass filters.
The DWT algorithm can be applied recursively N times, obtaining an N-level decomposition. At each level, the approximation coefficients of the previous level are the input for the new one. In such a way the desired level of detail is obtained. Since the wavelet filters have the cut-off frequency in the middle of the band of the input signal, at each level, the approximation and detail coefficients represent respectively the first and the second half of the frequency components. The situation for a 3-level decomposition is reported in Figure 2.14.

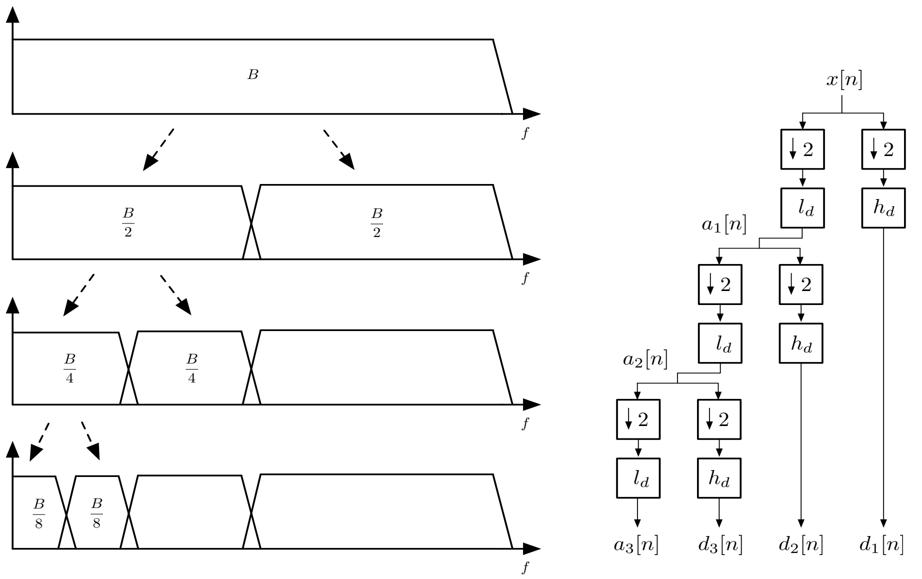


Figure 2.14: 3-level DWT decomposition

The DWT suffers the drawback of not being a time-invariant transform. A time invariant system is a system for which a time shift of the input sequence causes a corresponding shift in the output. In mathematical terms, we get

$$x(t) \rightarrow y(t) \Rightarrow x(t + \delta t) \rightarrow y(t + \delta t), \tag{2.98}$$

where $x(t)$ and $y(t)$ are respectively the input and the output signal of the system, while δt is the time delay applied to the input signal.

For the work presented in this thesis the fact that the DWT is not time-invariant represents a serious problem. In fact, after the wavelet decomposition we need to apply some time delays to the obtained signals, as described in Chapter 3. The solution is given by the Discrete Stationary Wavelet Transform (SWT) and its inverse ISWT. The functioning principle is the same of the DWT and IDWT. It can be proved that removing the signal downsampling and upsampling, and introducing instead the upsampling and downsampling of the wavelet filters of the previous level, we obtain a DWT with the desired time-invariance property. The situation for a generic i^{th} level of decomposition is reported in Figure 2.15.

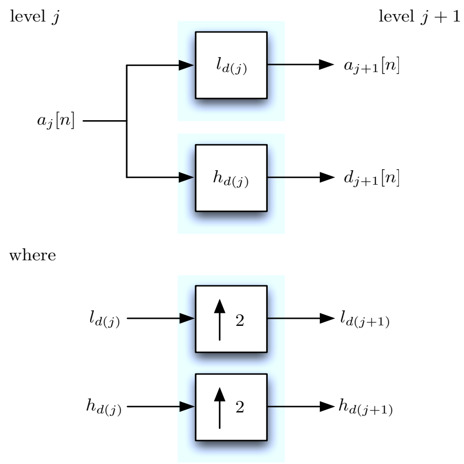


Figure 2.15: Generic SWT level decomposition and filters setup

Another important difference between the DWT and SWT is the length of the detail and approximation coefficients: in the first case it is the half on the input signal, whereas in the second case it is the same. The DWT and SWT are mainly used for signal denoising and sub-band processing, as in our case.

2.6 Concluding remarks

This chapter has introduced the theory required to explain the functioning of the work presented in this thesis. In the first part we have studied the physical model that describes the real world phenomena involved in the measurements. Starting from the wave equation, we have gradually introduced all the elements needed to arrive to the definition of the sound absorption coefficients. After that, we have taken into account some of the most interesting techniques for the sound absorption coefficients estimation, already present in the literature. Among all, we have described the Ducorneau's method, which is the one from which the work of this thesis takes the cue from. This technique presents a solution that uses a weighted *sum-and-delay* microphone array, in order to measure the energy of the direct and the reflected signals. Instead, our work performs the measures by means of a superdirectional differential microphone array.

The goal of this work is the implementation of a real measurement system, capable to get the sound absorption coefficients of a surface. The real devices involved in the measurements, such as the loudspeaker but also the microphones and the ADC / DAC board, introduce some undesired effects. The most significative is the linear-distortion. In order to compensate this problem, we require a method to exactly estimate the overall frequency response of the system. This purpose can be achieved by the spectral estimation method, presented in Section 4. Once this estimation is completed, we apply the corresponding whitening filter to the MLS source signal.

The need to process the input signals by sub-bands and, at the same time, the need to maintain temporal information about the signals after they are decomposed, has driven us to introduce the DWT. Moreover, the constraint about the time-invariance property of the transform has moved our attention to the SWT. A more detailed description of the system and the actual implementation are discussed in the next chapter.

Chapter 3

Absorption coefficients estimation

This chapter describes the structure and the implementation of the system proposed as the solution of the absorption coefficients estimation problem. In order to calculate the values of these coefficients we need to gather information about the direct and the reflected wavefronts. In particular, we are interested in the energy carried by each of these signals. The processes required to get the sound absorption coefficients are quite a few. Consequently, we adopt a top-down paradigm to describe the overall system. Beginning from a general description of the system, this chapter enters into the details of each sub-processing block reporting the inputs, the outputs and the operating principle. Moreover, all the theoretical elements presented in Chapter 2 are recalled to support the implementation choices.
Section 1 describes the system on the whole, highlighting the basic idea behind this work. Section 2 describes which kind of signal is used for the system measurements and its characteristics. Section 3 analyzes the multiband processing blocks, pointing out why we need them. Instead, the techniques used in the separation of the direct and the reflected wavefronts and the superdirectional microphone array implementation are discussed in Section 4. Once we get the direct and the reflected signals properly split, we need to elaborate them in order to get the absorption coefficient as function of the frequency. This procedure is reported in the last section.

3.1 System overview

The work presented in this thesis is involved in the estimation of the sound absorption coefficients of a surface. There are many methods to achieve such a result, as described in Section 2.2, but some of them are unreliable and the others are laboratory methods. Instead, we want to realize a system to estimate the absorption coefficients *in-situ* with a portable device, in a reliable and repeatable way. With this method we are able to infer the sound absorption characteristics of a surface under real conditions. For example, we could estimate the influence of a surface on the reverberation time of a certain room. Moreover, as the method does not require small portions of material, unlike the normal incidence methods, it is non destructive. Lastly, we want the method to be robust against the reverberation phenomena generated by objects or other surfaces adjacent to the one under consideration.

Under real conditions, the *power* sound absorption coefficient is defined as the energy portion of the incident wavefront that a surface does not reflect. As this quantity is represented by the energy absorbed by the wall (W^{abs}) and the one that is transmitted (W^{tr}), we should be able to estimate these quantities. Unfortunately, this task is quite difficult both empirically and theoretically, as we have described in Section 2.1. In fact, the physical models that describe the absorption and the transmission phenomena under real conditions have to keep into account a high number of variables. A more convenient way to address the problem is to apply the energy conservation law. Knowing the incident and the reflected energies, we get

$$\alpha = \left(\frac{W^{abs}}{W^{+}} + \frac{W^{tr}}{W^{+}} \right) = 1 - r, \tag{3.1}$$

where

$$r = \frac{W^{-}}{W^{+}}, \tag{3.2}$$

is the *power* reflection coefficient, whereas W^{+} and W^{-} are respectively the incident and the reflected energy. An additional information that could be retrieved from such a system is the dependency of the *power* sound absorption coefficient as function of the frequency. Formalizing this aspect, equation (3.1) reduces to

$$\alpha(f) = 1 - r(f) = 1 - \frac{W^{-}(f)}{W^{+}(f)}, \tag{3.3}$$

where f is the frequency, measured in Hz. It is clear that the fastest way to obtain a frequency dependent result is to stimulate the surface under test using a wideband signal, rather than repeating the same measurements for each frequency we are interested in. For this purpose we choose a Maximum Length Sequence (MLS) signal. It is a particular type of pseudorandom binary sequence that is generated using the Linear Feedback Shift Registers. The MLS signals benefit of interesting properties such as the ability to generate short impulse responses, provided that we use a pre-equalization technique. Further details about the MLS source signal and about the pre-equalization technique are given in the next section.
As we have seen above, the initial problem is converted into a new one, where the unknowns are represented by $W^{-}(f)$ and $W^{+}(f)$. Some methods, such as the Mommertz [3] one, measure these quantities in two different time instances. The reflected energy is measured during the estimation of the surface coefficient, whereas the incident one in a free-field condition. This situation could represent a problem since the environmental conditions may vary significantly in time. Small variations of the sound velocity, caused by temperature variations, or different conditions of sound propagation could deteriorate the results. Consequently, our choice is to measure the quantities we are interested in, in the same time instant. The idea is to place a measuring instrument between the sound source and the reflecting wall. This device has to be able to split the incident and the reflected signals, since they arrive to the measuring instrument superimposed. The acquiring device should also be robust against the reverberations. In fact, it is better to remove the undesired components at the source rather than trying to compensate them later. A solution that satisfies all these requirements is represented by a superdirectional microphone array, as the one theoretically introduced in Section 2.3.
If we place the source normal respect to the surface under test, the reflected signal is ideally mirrored in the opposite direction respect to the incident one. This is true only if we are dealing with plane waves (i.e. the loudspeaker is in the *far-field*). The angle of incidence is the same of the reflection one. In this way we reduce the reverberation phenomena, provided that there are not obstacles, except the measuring device, between the loudspeaker and the surface. Ideally no other wavefronts are generated during the measurement, since refraction and

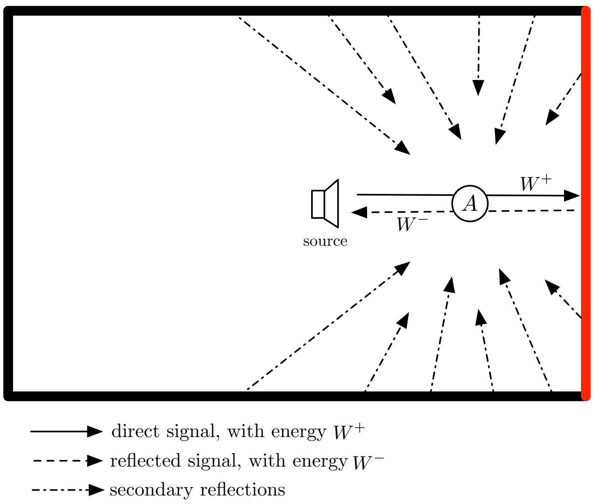


Figure 3.1: System placement and wavefronts involved in the propagation. As the measuring device A is omnidirectional, it receives also the secondary reflections, besides the direct wavefront and the one reflected by the surface.

diffraction phenomena are avoided using the normal incidence positioning. This method introduces robustness to reverberation but also the limitation of not being able to determine the absorption coefficient as function of the incidence angle. In our implementation we choose to be robust against reverberations, rather than allowing the estimation of the sound absorption coefficient as function of the incidence angle, also because we exploit the normal source-surface positioning in the implementation of the superdirectional microphone array. As we work mainly under the *near-field* condition, the loudspeaker has to be considered as a source of spherical waves. This assumption modifies the previously described situation, since part of the incident waves are not at normal incidence respect to the surface. So, these wavefronts propagate in other direction respect to the normal one, generating secondary reflections (i.e. the cause of the reverberation phenomenon). The *near-field* situation is reported in Figure 3.1. The red line represents the wall under consideration. The speaker is placed normally respect to the surface, whereas the measuring instrument A is located between the speaker and the surface. The wavefronts involved in the propagation are reported too. In addition to the direct and the reflected wave paths, there are also the reverberation wavefronts coming from arbitrary directions.

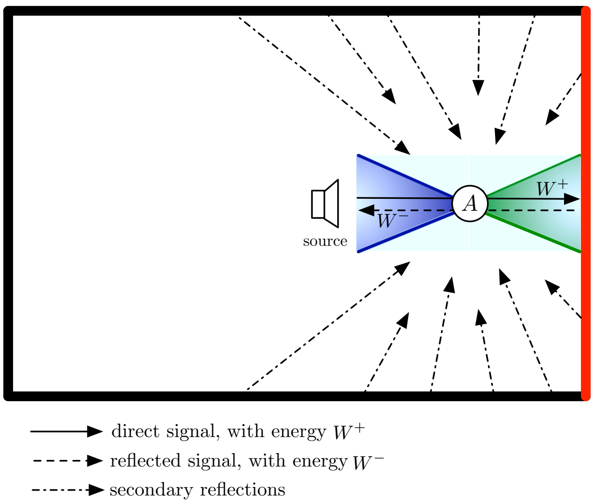


Figure 3.2: System placement and wavefronts involved in the propagation, using a device with a narrow beamwidth. The blue and the green angles represent respectively the direct and the reverse direction beamwidths. The wavefronts that arrive to the device with an incidence angle outside these beamwidths are to be considered null

The properties that the superdirectional microphone array should have are:

- narrow beamwidth, in order to neglect the reverberation wavefronts, introduced above. The situation that considers the narrow directivity of the array is represented in Figure 3.2. The beamwidths for the allowed *listening* directions are identified by the blue and the green angles. Ideally, the wavefronts arriving from paths that are outside the reported beamwidths are to be considered null.
- maximum front-to-back ratio, in order to make the system able to measure the direct and the reflected signals, at the same time.
- constant directivity on the desired frequency range.

The constant directivity is a weak requisite since our system does not really require it, even if it is anyway a desired property. The ratio to calculate the *power* absorption coefficient is a frequency-by-frequency operation, so a variation of the directivity as function of the frequency does non represent an issue. In fact, the energy carried by a certain frequency of the direct or the reflected wavefront, is received in the same way from the microphone array. Anyway, we want to satisfy the

constant directivity property for a possible reuse of the microphone array in other applications.
In order to build the superdirectional microphone array with the desired characteristics we have chosen to implement it as a fourth order differential array. A detailed discussion about the microphone array implementation is given in Section 3.4. As we know from the theory, a fourth order differential microphone array requires five real microphones. Consequently, the wavefronts are recorded in different places, at the same time. It is the so-called spatial sampling. As for the case of the temporal sampling we have to provide a condition that, if satisfied, guarantees the absence of aliasing problems. We can find this condition making a comparison with the Shannon's theorem, valid for the temporal case. The Shannon's theorem is formalized by the following inequality

$$T < \frac{\pi}{\omega_{max}}, \tag{3.4}$$

where T is the sampling period and ω_{max} is the maximum angular frequency. If we make the following associations

$$T \leftrightarrow d, \tag{3.5}$$

$$\omega \leftrightarrow \nu = 2\pi\frac{c}{\omega}, \tag{3.6}$$

where d is the spacing between the microphones, ν is the spatial frequency and c is the sound speed, knowing that the wavelength λ is defined as

$$\lambda = \tfrac{c}{f} = \tfrac{c}{2\pi\omega},$$

we can rearrange inequality (3.4) into

$$d \leq \frac{\lambda_{min}}{2}, \tag{3.7}$$

that is the spatial sampling theorem. According to this condition, the spacing between the microphones has to be not more than the half of the minimum wavelength λ_{min}. Moreover, we know that the beamwidth of a microphone array is proportional to the inter-element spacing d. Consequently, we encounter a trade-off situation in the choice of the spacing. On one side we require a small spacing in order to be able to process without aliasing high frequency signals. On the other side we require a larger spacing in order to get a narrower beamwidth. Considering that the sound absorption coefficients are usually estimated in the

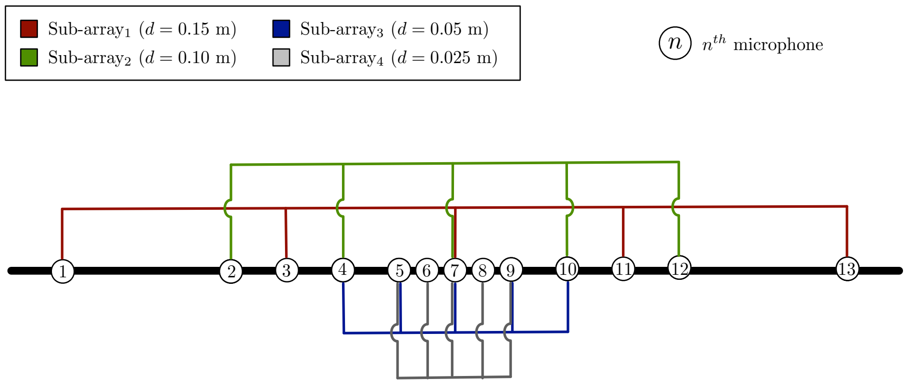


Figure 3.3: Sub-array overall system, with relative spacings. The microphones that belong to the same sub-array are identified with a line of the same color. The microphone no.7 is the array reference point, since it is in the middle of the microphone series and it is shared by all the sub-arrays.

$0 \div 4000$ Hz frequency range, it is clear that we can not achieve an optimal functioning, with appreciable results, using only one array. The idea is to use more microphone arrays with different spacings, making them working together. Each microphone array is able to process only a limited frequency range, according to its spacing d. The microphone input signals have to be split into the same sub-bands defined by the arrays. After that, each microphone sub-band has to be sent to the proper array (i.e. to the proper sub-array processing block).

The implementation of such functionality is realized by a sub-band processing technique that makes use of the stationary wavelet transform and of the wavelet filter banks. The detailed description of sub-band processing implementation is reported in Section 3.3.

The proposed system uses 4 microphone arrays, thus we have 4 sub-bands. As we want a system that is cheap and handy, we desire to reduce the number of microphones. The solution is given by the fact that different arrays share some microphones. This can be achieved setting the same reference point for all the arrays and choosing an appropriate spacing for each of them. In this way we are able to use 13 microphones globally, instead of the 20 required. This implementation is sketched in Figure 3.3. The colored lines represents the microphones belonging to the same sub-array[1]. If more than one line exits from the same microphone, it means that microphone is shared between two or

[1]From now, each array is referred as *sub-array*, whereas with the term array we refer to the whole system, made up of the 13 microphones.

more sub-arrays. Notice that the microphone no.7 is shared by all the sub-arrays since it is the central reference point.
Once the microphone sub-bands have been processed by the proper sub-array, at the output of each sub-array processing block we get the related sub-bands for the direct and the reflected signals. In order to proceed to the estimation of the sound absorption coefficient we have to reconstruct the direct and the reflected wavefronts, using the sub-band obtained at the previous step.
The detailed overall system block diagram is reported in Figure 3.4. There we can identify all the functionalities described up to now. The testing MLS signal is emitted by the loudspeaker, which is placed normally in front of the surface under consideration. All the wavefronts involved in the sound propagation are recorded by the array microphones (from $m(1)$ up to $m(13)$). As previously described, in order to correctly process all the frequencies in the range we are interested in guaranteeing a narrow beamwidth, we require more than one array. In fact, we have 4 sub-arrays, where each of them is delegated to process a particular frequency sub-band. The arrays with the smallest spacings are delegated to process high frequency signal components. On the contrary, the arrays with the largest spacings are designed to process low frequency signal components.
We have to send to a sub-array only the frequency components of the signals it is designed to deal with. To clarify this point we take into account the microphone in which is located the reference point. It is the central microphone ($m(7)$). Since this microphone is shared between all the arrays, we require that the signal acquired from it has to be split in 4 sub-bands, one for each sub-array. Generalizing, even if not all the microphones serves more than one array, we have to send the proper signal frequency component to its related sub-array. Consequently, all the signals coming from the microphones are split in 4 sub-bands each one, getting 13×4 input signal sub-bands, globally. This operation is realized by the *wavelet filter bank decomposition* block. The signal bands coming from the wavelet decomposition are passed to the proper sub-array processing block. We globally introduce 8 array sub-band processing blocks: 4 (*B1.D*, *B2.D*, *B3.D* and *B4.D*) for the direct signal array processing and 4 (*B1.R*, *B2.R*, *B3.R* and *B4.R*) for the reflected one. Each of these blocks has 5 inputs, since we have decided to implement our system with a fourth order differential microphone

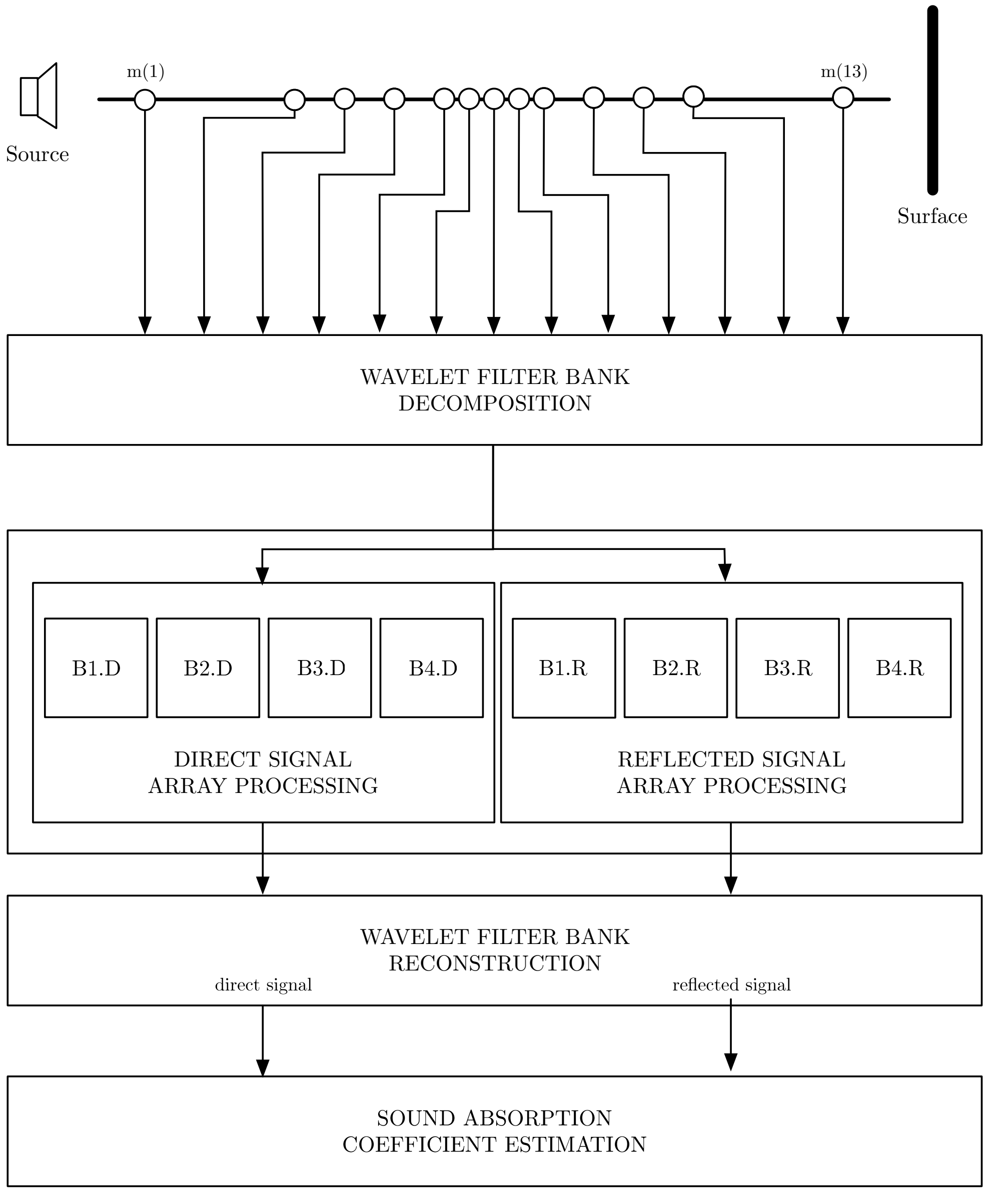


Figure 3.4: Overall system block diagram

array. The inputs are represented by the proper microphone sub-bands, described above. At the end of the array processing we get 4 sub-bands for the direct signal and 4 for the reflected one. A detailed description about the sub-array processing blocks, together with the actual microphone array implementation, is given in Section 3.4.
The next step is to reconstruct the desired wavefronts, the direct and the reflected one, using the *wavelet filter bank reconstruction* block that operates in the dual way respect to the decomposition one. These *wavelet filter bank* blocks are described in detail in the *Multiband processing* section.
Lastly, the sound absorption coefficient estimation is performed by the last block reported in Figure 3.4, where the implementation of equation (3.3) takes place. This block is discussed in detail in the last section of this chapter.

3.2 MLS source signal

In order to test a real environment over a broad frequency spectrum we desire a signal that carries the same energy content on all the frequencies. In this way we get a solution that does not depend on the frequency characteristics of the input signal, since its frequency response is ideally flat. The signals that satisfy this property are the so-called *white-noise like* signals. As the goal of the proposed work is to estimate the sound absorption coefficients of a surface as a function of the frequency in the $0 \div 4000$ Hz range, the use of a *white-noise* source signal is highly recommended. Moreover, with such a signal, we avoid the need to test the environment with multiple measurements, one for each frequency we are interested in. We introduce the MLS as a particular type of *white-noise*. The MLS signals own other interesting properties besides having a flat frequency response. This is the motivation for that we have chosen them as the system input source.
The MLS sequences are pseudorandom and periodic signals. With the *pseudo-random* term we mean that these signals are periodic with a finite length and so they only approximate a real random process. The longer the sequence is, the better is the obtained approximation.
The most common way to generate an MLS signal is to use the so-called Linear Feedback Shift Registers (LFSR). The LFSR are single bit shift registers, in which the outputs of some stages are used to calculate the

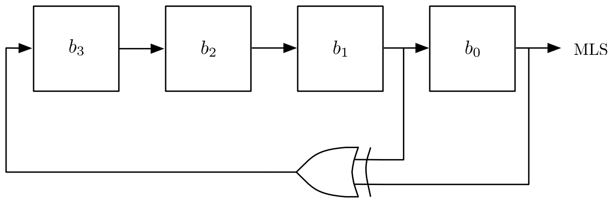


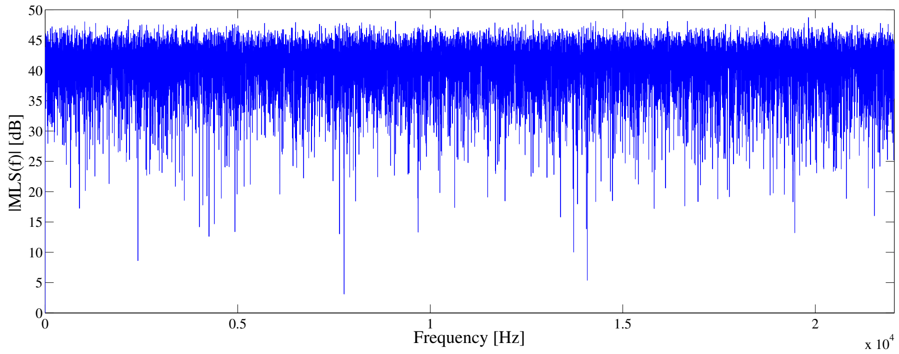
Figure 3.5: MLS generation using LFSR (N=4)

inputs of other stages, by means of xor ports. If the generating LFSR contains N bit blocks, the generated MLS is periodic with a period equal to $2^N - 1$. An example of the generation of an MLS signal by LFSR is represented in Figure 3.5.
The mathematical formalization for the $N = 4$ case is given by the following recursive definition

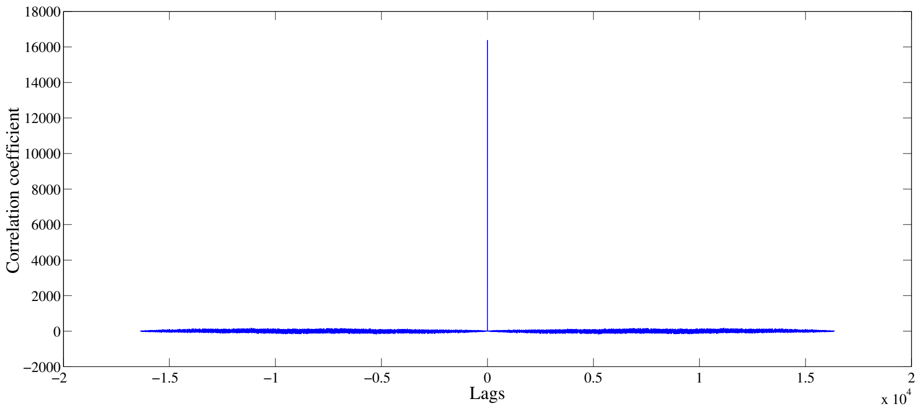
$$b_k[n+1] = \begin{cases} b_0[n] \oplus b_1[n], & \text{if } k = 3 \\ b_{k+1}[n], & \text{otherwise} \end{cases}, \tag{3.8}$$

where k is the shift register bit position index, n is the discrete time index and $\oplus$ is the xor operator. Since the result of the MLS generation is achieved using logical operations, we get a binary result. Consequently, the amplitude values of MLS sequences are defined in the $\{0, 1\}$ set. As we are dealing with signals used for acoustical purposes, we can convert the amplitudes of MLS sequences in the more appropriate $\{-1, 1\}$ set. In the implementation of our work we use an MLS sequence with $N = 14$, generating a signal with a period of length $2^{14} - 1$.
As anticipated the MLS benefits of some interesting properties. The first one is represented by its flat frequency response, as reported in Figure 3.6(a).
The second property is the one for which the MLS auto-correlation is with a good approximation a unit impulse, as reported in Figure 3.6(b). As described in the last section of this chapter, in order to get the sound absorption coefficients we have to know the impulse responses of the direct and the reflected wavefronts.
The starting point to get these values is the definition of the correlation between two generic signals x and y, defined as

$$c_{x,y} = \frac{1}{L} \sum_{k=0}^{L-1} x[n]\ y[n+k]. \tag{3.9}$$



(a) MLS flat spectrum



(b) MLS unitary auto-correlation

Figure 3.6: MLS properties, exploited during the processing calculations

Taking the Discrete Fourier Transform, we get

$$C_{x,y} = \frac{1}{L}\,\overline{X[k]} \cdot Y[k]. \tag{3.10}$$

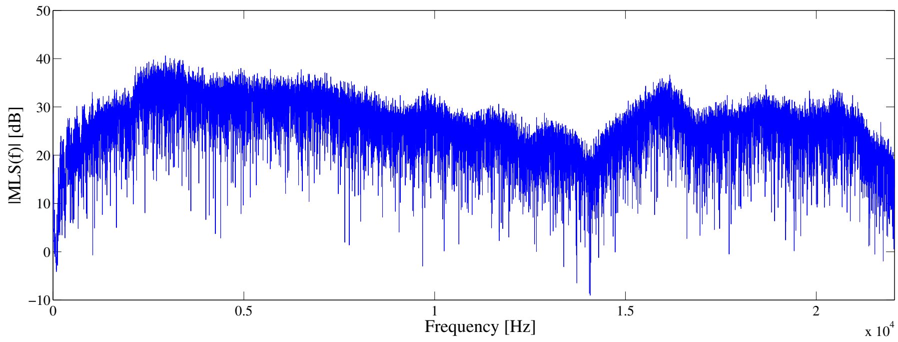
Assuming that the signal y is generated by x passing through a system with response h, we can write

$$C_{x,y} = \frac{1}{L}\,\overline{X[k]} \cdot Y[k] = \frac{1}{L}\,\overline{X[k]} \cdot X[k] \cdot H[k] = \frac{1}{L}\,|X[k]|^2 \cdot H[k]. \tag{3.11}$$

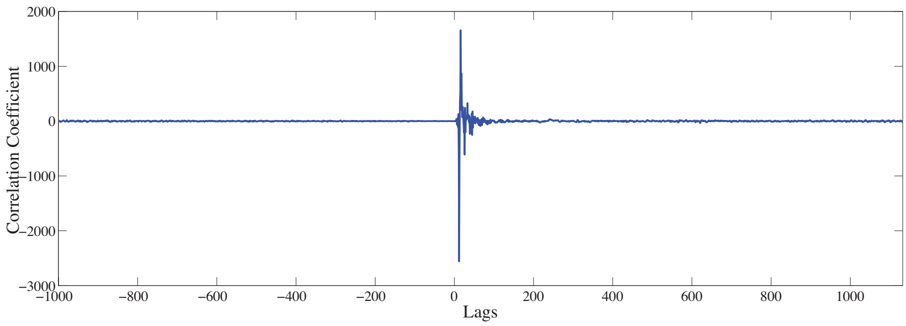
Rearranging the previous equation as function of $H[k]$, we get

$$H[k] = \frac{\overline{X[k] \cdot Y[k]}}{|X[k]|^2} = \frac{\overline{X[k]} \cdot Y[k]}{X[k] \cdot X[k]} = \frac{C_{x,y}}{C_{x,x}}, \tag{3.12}$$

where $C_{x,y}$ is the Fourier Transform of the correlation between the signal x and the signal y. Usually, it is also referred as the *cross-*



(a) Non-flat MLS spectrum



(b) Long MLS impulse response

Figure 3.7: Frequency and time domain artifacts for an MLS signal, due to the linear distortion introduced by the loudspeaker and the cascade of measuring instruments

correlation. Instead, $C_{x,x}$ is the Fourier Transform of the correlation between the signal x and itself and it usually referred as *auto-correlation*. In this way we have proved that in order to get the impulse response of a signal in a certain environment, it is sufficient to take the ratio between the Fourier transform of the cross-correlation between the input and the output signal, and the Fourier Transform of the auto-correlation of the input signal. Knowing that the MLS is the input signal of our system (i.e. the source signal) and that its auto-correlation is with a good approximation a unit response, equation (3.12) simply reduces to

$$H[k] = C_{x,y}. \tag{3.13}$$

We exploit this simplification during the computation of the sound absorption coefficients, since they are calculated taking the ratio between the energy carried by the impulse responses of the direct and the reflected wavefronts.

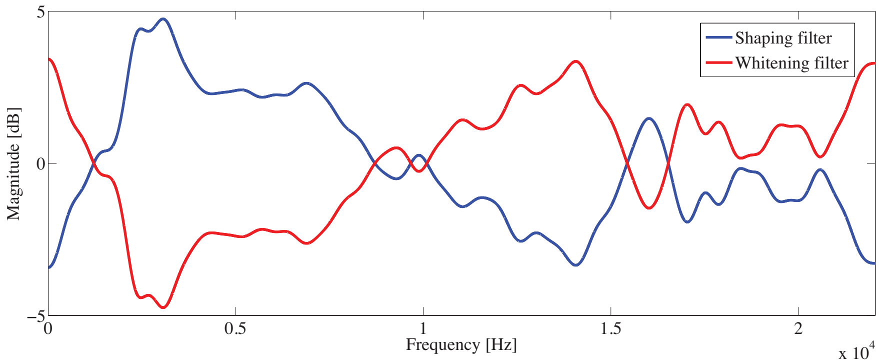


Figure 3.8: Shaping and whitening filters, used to achieve source signal pre-equalization.

Notice that the discussion about MLS signals presented up to now is purely theoretical since it does not take into account the linear distortion introduced by the loudspeaker and the cascade of the measuring instruments. This kind of distortion results in a non-flat MLS spectrum for what regards the frequency domain and in a longer impulse response respect to the desired one, for what regards the time domain. An example of time and frequency domain artifacts introduced by the linear distortion are represented in Figure 3.7.

As anticipated in Section 2.4, in order to compensate the linear distortion we can perform the spectral estimation of the overall system (i.e. the loudspeaker and the cascade of the measuring instruments), applying its inverse to the MLS source signal before it is emitted. Applying this pre-equalization technique to an MLS signal, it has been proved [12] [13] that the linear distortion phenomenon is compensated in a proper way. In fact, we obtain a nearly flat spectrum and, above all, short impulse responses.

The filter obtained by the spectral estimation, the so-called *shaping filter*, is represented in Figure 3.8, together with the inverse one. This one is the so-called *whitening filter* and it is the filter applied to the MLS source before it is emitted, achieving the desired source pre-equalization.

3.3 Multiband processing

In order to estimate the sound absorption coefficients over a frequency range between 0 and 4000 Hz, we have seen that one microphone array is not sufficient. This is due to the trade-off existing over d, between the spatial sampling constraint ($d \leq \lambda_{min}/2$) and the fact that the microphone spacing d has to be large enough to get a narrow beamwidth. Hence, we decide to introduce 4 sub-arrays where each of them is designed to process a certain sub-band, according to the spatial sampling constraint. The implementation of the differential microphone arrays is discussed in the next section. For the moment we simply recall the fact that each sub-array is a fourth order differential microphone array, made up of 5 real microphones. In order to obtain a cheap and a handy microphone array, we desire to minimize the number of microphones used in the implementation. We can achieve this optimization sharing some microphones between two or more sub-arrays.

Each microphone has to send a proper signal frequency component (i.e. a sub-band) to the sub-band processing units, related to it. Knowing that a microphone is shared by at most 4 sub-arrays, it is sufficient to split the signal coming from each microphone in 4 sub-bands. Notice that not all the sub-bands coming from the microphones bring useful information. For example, for a microphone belonging only to one sub-array, the sub-band related to that sub-array is kept, whereas the other 3 are discarded. The sub-band decomposition is implemented by the *wavelet*[2] *filter bank decomposition* block.

Before entering into the detail of the discussion about the system blocks involved in the multiband processing, we have to set the sub-array spacings, fulfilling the spatial sampling and the cost minimization constraints. Moreover the choice of d has to be compatible with the bands obtained by the wavelet filter bank decomposition. In fact, from Section 2.5 we know that at each decomposition level the band of the input signal is split in two sub-bands large the half of the input one. The situation for a 3-level wavelet decomposition is reported in Figure 2.14. In our implementation the frequency sampling rate has been set to 44100 Hz. By the Shannon's sampling theorem we know that using this rate we can properly describe signals over the $0 \div 22050$ Hz frequency range. At each decomposition step the input bandwidth is

[2]Henceforth, when we talk about the wavelet transform and the wavelet filter bank, we mean the SWT implementation.

divided recursively by 2, getting submultiples of 22050 Hz. In order to get 4 sub-bands from each microphone input, we could use a 3-level wavelet filter bank. Beginning the decomposition from 20050 Hz, we get the sub-bands reported in Table 3.1

Band no.	Frequency range [Hz]
1^{st}	$\approx (0 \div 2756)$
2^{nd}	$\approx (2757 \div 5512)$
3^{rd}	$\approx (5513 \div 11025)$
4^{th}	$\approx (11026 \div 22050)$

Table 3.1: Frequency ranges for the 3-level sub-band decomposition

This result does not satisfy our requirements since it provides frequency bands too large and out of the range we are interested in. To solve this issue there are two possible solutions. The first one is to downsample the original signal. In this way we proceed to the wavelet decomposition starting from a lower frequency and so getting narrower and lower frequency bands. Unfortunately this solution is lossy, in the sense that we lose frequency information that may result in having temporal aliasing artifacts. Instead, the second solution is represented by a higher level of the wavelet decomposition. Implementing a 6-level wavelet filter bank, we get the sub-bands reported in Table 3.2.

Band no.	Frequency range [Hz]
1^{st}	$\approx (0 \div 344)$
2^{nd}	$\approx (345 \div 689)$
3^{rd}	$\approx (690 \div 1378)$
4^{th}	$\approx (1379 \div 2576)$
5^{th}	$\approx (2757 \div 5512)$
6^{th}	$\approx (5513 \div 11025)$
7^{th}	$\approx (11026 \div 22050)$

Table 3.2: Frequency ranges for the 6-level sub-band decomposition

This solution is better than the previous one and does not introduce drawbacks. In fact we have 5 sub-bands that cover the frequency band we are interested in. Using the data obtained by the 6-level wavelet decomposition, together with the constraints previously introduced, we can easily set the microphone spacing. Notice that the sub-band widths are automatically defined by the wavelet sub-band decomposition. The

Band no.	d [m]	Frequency range [Hz]	Max freq. [Hz]
1^{st}	0.15	$\approx (0 \div 344)$	≈ 857
2^{nd}	0.10	$\approx (345 \div 689)$	≈ 1715
3^{rd}	0.05	$\approx (690 \div 1378)$	≈ 3430
4^{th}	0.025	$\approx (1379 \div 2756)$	≈ 6860
5^{th}	0.025	$\approx (2757 \div 5512)$	≈ 6860
6^{th}	n.a.	$\approx (5513 \div 11025)$	n.a.
7^{th}	n.a.	$\approx (11026 \div 22050)$	n.a.

Table 3.3: Microphone spacings and related sub-bands

situation is summarized in Table 3.3.
The *Max freq.* column represents the maximum frequency supported by the selected spacing d, obtained form the spatial sampling constraint. As we expect, small spacings are suitable to deal with high frequencies, whereas large spacings are suitable to deal with low frequencies. Another aspect to notice is that all the bands for which we have defined a spacing value d, except the 5^{th} one, are quite distant from the maximum allowed frequency, resulting in robustness respect to spatial aliasing problem.
The frequency bands we are interested in are the ones from the 1^{st} up to the 5^{th}. The 6^{th} and the 7^{th} frequency bands represent information that is not significative for our purposes. Moreover, since these bands cover high frequency ranges we should use very small spacings in order to fulfill the spatial sampling theorem, but arrays with such a small spacings are difficult to realize in practice due to the microphones dimensions. Thus, the 6^{th} and the 7^{th} bands have to be considered only in order to achieve the perfect reconstruction during the filter bank reconstruction. For this reason we neglect them in the following discussion.
The 5^{th} band requires more attention. It is in the middle of the selected and the discarded bands and a portion of its frequency range covers the frequencies we are interested in. Spatial aliasing problems could arise since the highest frequencies of this band are closer to its maximum allowed frequency. Consequently, we consider only the frequencies that are outside the range that could be affected by spatial aliasing issues. During the system test, we have found that this limit is around the 3000 Hz. Notice that the spacing chosen for the 5^{th} band is the same of the 4^{th} one. This means that the 5^{th} band is related to the same

sub-array of the 4^{th} one. The only difference is their frequency content, but for the rest one band is a clone of the other one, in the sense that they are processed in the same way, as well. Neglecting also the 5^{th} band, we obtain a convenient representation of the overall system that results in a perfect mapping between sub-bands and sub-arrays. In the actual system implementation we obviously take into account the fact that we have 7 sub-bands instead of the 4 presented in the discussion. With the spacings configuration presented in Table 3.3 we get a microphone array composed globally by 13 microphones, as the one reported in Figure 3.3. The microphones / sub-arrays mapping is reported in Table 3.4. From this table we can easily identify the microphone that is set as the system reference point. It is the microphone no. 7, since it is shared by all the sub-arrays and placed exactly in the middle of the microphone series. Further information about microphones positioning, sub-bands and sub-arrays mapping is provided in Appendix C.

Mic no.	1	2	3	4	5	6	7	8	9	10	11	12	13
Sub-array $_1$	$\times$		$\times$				$\times$				$\times$		$\times$
Sub-array $_2$		$\times$		$\times$			$\times$			$\times$		$\times$	
Sub-array $_3$				$\times$	$\times$		$\times$		$\times$	$\times$			
Sub-array $_4$					$\times$	$\times$	$\times$	$\times$	$\times$				

Table 3.4: Microphones / sub-arrays mapping

As we have defined the sub-bands required by the system implementation and the microphones / sub-arrays mapping, we are now able to discuss the blocks involved in the multiband processing. The first block we take into account is the *wavelet filter bank decomposition* one. The purpose of this block is to decompose each signal coming from the microphone array into 4 sub-bands. Consequently the inputs of this block are the signals coming from the microphones, whereas the outputs are represented the sub-bands obtained by the signal decomposition. Since we have 13 microphone signals as inputs and each of them generates 4 sub-bands, we globally obtain 13×4 sub-bands at the output of the block. The implementation of the wavelet filter bank decomposition is theoretically introduced in Section 2.6. In order to get the 4 desired sub-bands of the input signal, we need a 3-level wavelet decomposition. At each decomposition level the band of the input signal (or of the approximation coefficients of the previous level) is divided by 2 getting the detailed and the approximation coefficients for the new

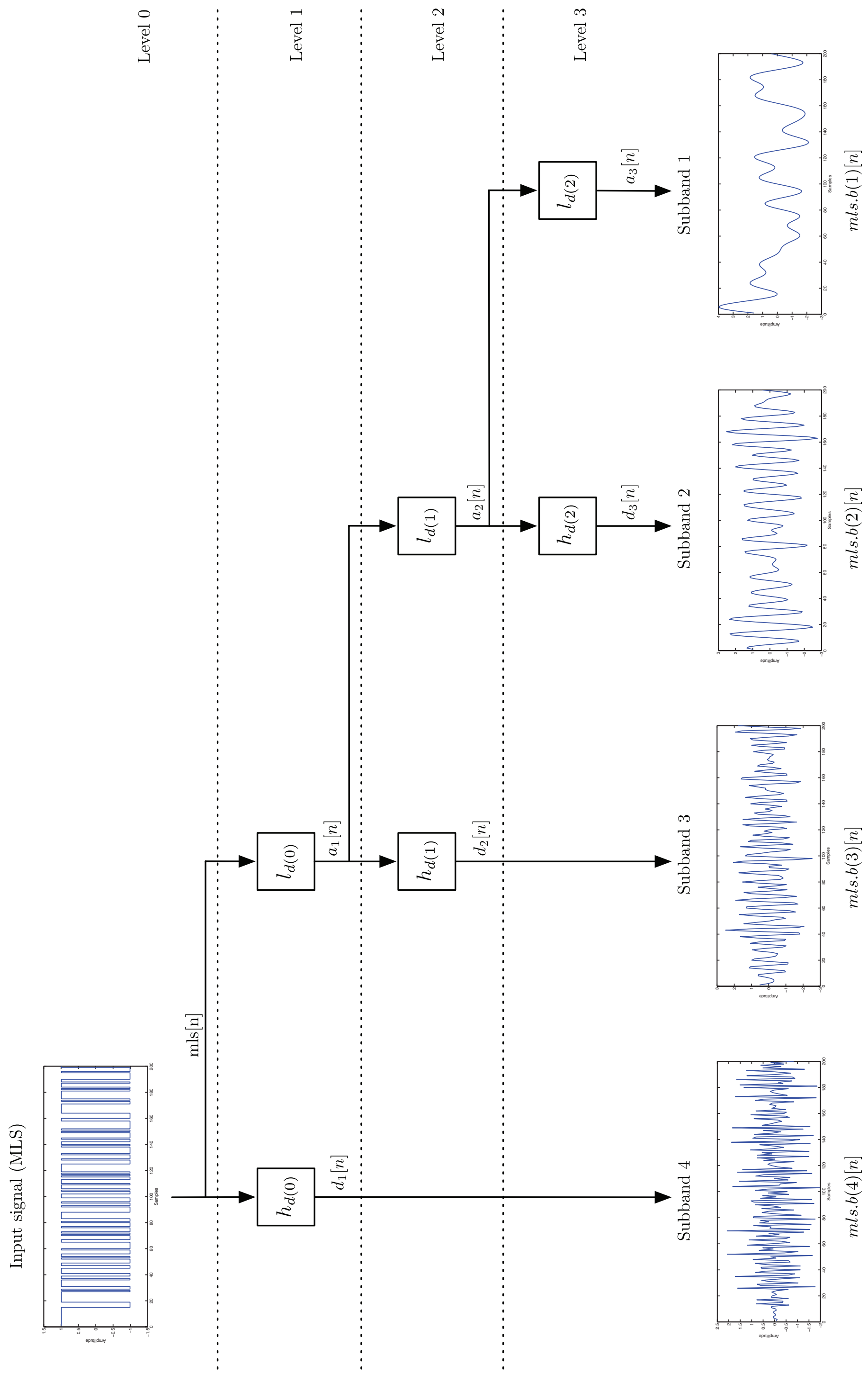


Figure 3.9: Actual wavelet filter bank decomposition. For each decomposition level, the required filters are reported. The time domain outputs for each sub-band are reported, too.

level. The bandwidths of these signals are the half of the input one. The situation for a 3-level wavelet decomposition with a generic MLS input signal is reported in Figure 3.9. Since we are using the SWT version of the wavelet decomposition, the obtained sub-bands have the same length of the input signal. The $a_i[n]$ and $d_i[n]$ terms represent the approximation and the detail signal coefficients for each decomposition level i, respectively. The low-pass ($l_{d(i+1)}$) and the high pass ($h_{d(i+1)}$) decomposition filters are built as described in Section 2.6: at each decomposition level they are obtained upsampling by a factor 2 the respective filters of the previous level. Moreover $l_{d(0)}$ and $h_{d(0)}$ are built using the wavelet base, which in our implementation is the discrete Mayer wavelet, also referred as *dmey*. It is a symmetric, orthogonal and biorthogonal wavelet. The coefficients for $l_{d(0)}$ and $h_{d(0)}$ decomposition filters obtained with such a wavelet base are represented in Figure 3.11, together with their frequency responses. In particular, from Figure 3.11(b) and 3.11(d) we can identify the reason for which at each decomposition level we get two output signals, whose frequency bandwidths are the half of the input one: in both the cases the normalized cut-off frequency is exactly at 0.5π $rad/samples$ that is the half of the input signal frequency bandwidth.
Once we have performed the sub-band decomposition for each signal coming from the microphones, we have to send the obtained sub-bands to the sub-array processing blocks, as reported in Figure 3.10.
The sub-array processing blocks are identified by $B1.D, B2.D, B3.D$, and $B4.D$ for the direct signal processing, and by $B1.R, B2.R$, $B3.R$, and $B4.R$ for the reflected one. As the two sets of blocks are equivalent, except for an implementation detail that allows to separate the direct and the reflected signals, in the following discussion we take into account just the direct one.
Each sub-array receives the signals coming from the microphones related to it. More precisely, each sub-array receives only a particular sub-band of the signals coming from the microphones related to it. The selected sub-band is the one that the sub-array is able to process correctly. For example, we know that the first sub-array ($d = 0.15m$) is associated with the first sub-band ($0 \div 344$ Hz). Thus, only the lowest frequency sub-band obtained by the decomposition of its related microphones is sent to the array. The complete mapping between the microphone sub-bands and the sub-arrays is reported in Table 3.5. No-

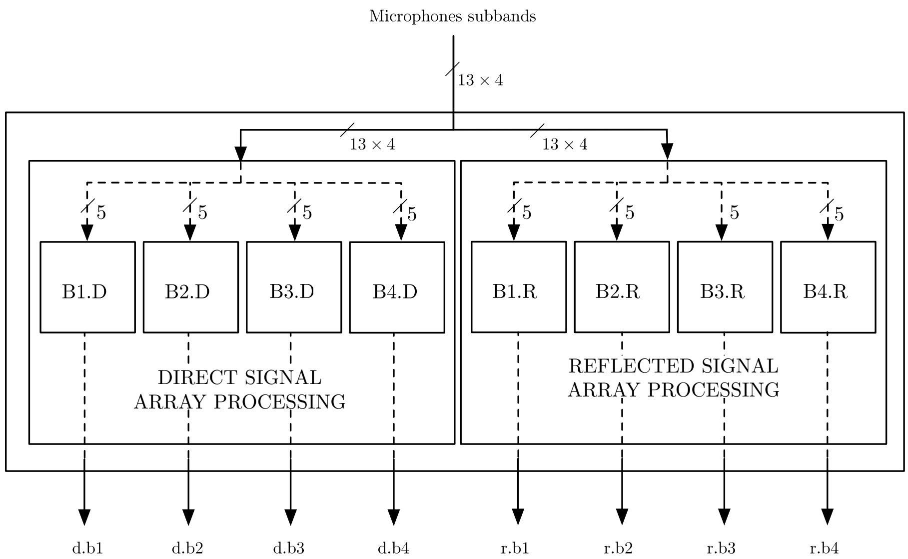


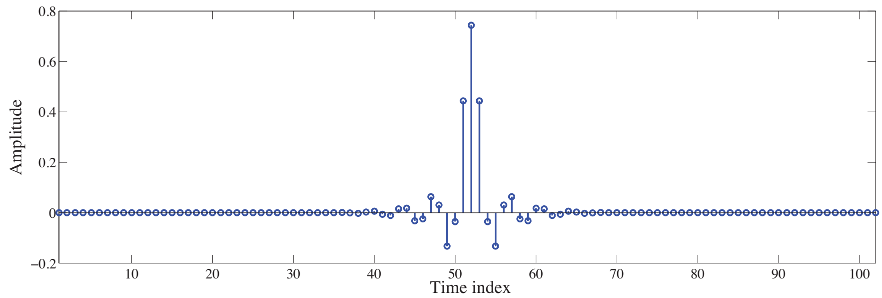
Figure 3.10: Microphone array sub-band processing blocks, for the direct and the reflected signals.

tice that in the case of the reflected signal array processing, the inputs of the sub-bands processing blocks are the same of the ones presented in Table 3.5.

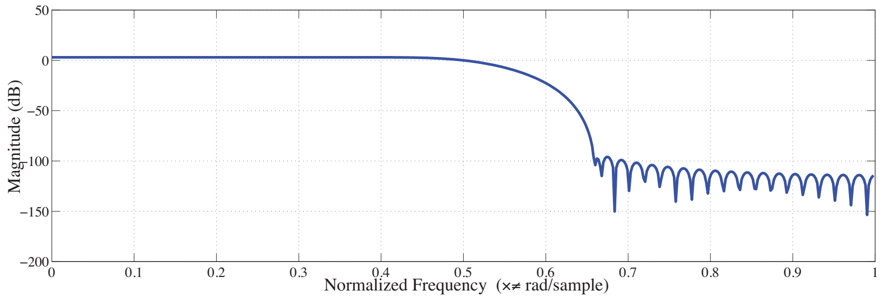
Block	sub-band	Inputs (sub-bands)
$B1$	1^{st}	$m(1).b(1), m(3).b(1), m(7).b(1), m(11).b(1), m(13).b(1)$
$B2$	2^{nd}	$m(2).b(2), m(4).b(2), m(7).b(2), m(10).b(2), m(12).b(2)$
$B3$	3^{rd}	$m(4).b(3), m(5).b(3), m(7).b(3), m(9).b(3), m(10).b(3)$
$B4$	4^{th}	$m(5).b(4), m(6).b(4), m(7).b(4), m(8).b(4), m(9).b(4)$

Table 3.5: Blocks / sub-band mapping

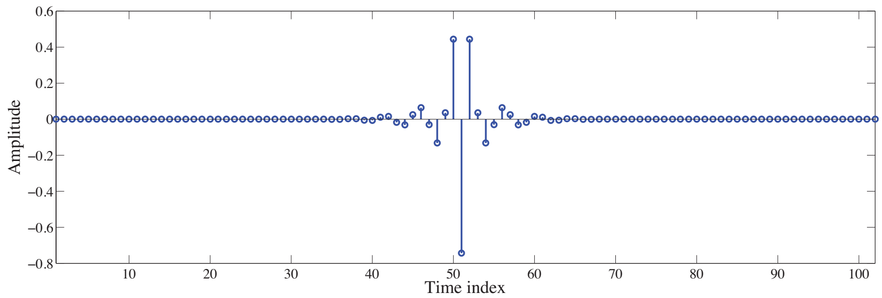
However, the *reflected signal array processing* blocks differ from the ones related to the direct signal processing due to their internal implementation, discussed in the next section. Their outputs are different, too. In one case these outputs represent the sub-bands related to the direct wavefront ($d.b1, d.b2, d.b3$ and $d.b4$), and in the other case they represent the sub-bands related to the reflected wavefront ($r.b1, r.b2, r.b3$ and $r.b4$). In order to estimate the sound absorption coefficients we have to know the direct and the reflected wavefronts. Consequently, we have to reconstruct these two signals, starting from their sub-bands. This task is realized by the *wavelet filter bank reconstruction* block, presented in Figure 3.12. It is the dual block respect to the *decomposition*



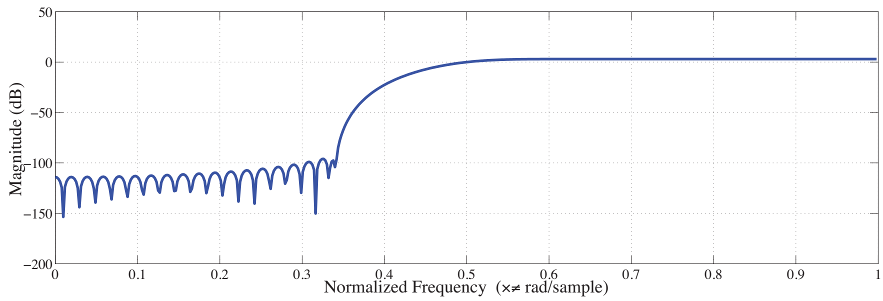
(a) Impulse response of the wavelet low-pass decomposition filter



(b) Frequency response of the wavelet low-pass decomposition filter



(c) Impulse response of the wavelet high-pass decomposition filter



(d) Frequency response of the wavelet high-pass decomposition filter

Figure 3.11: Time and frequency responses of the decomposition wavelet filters

one. The wavelet reconstruction block contains only two filter banks: one for the direct wavefront reconstruction and one for the reflected wavefront. As in the *decomposition* case, the internal filter banks are one the clone of the other. The implementation of the *reconstruction*

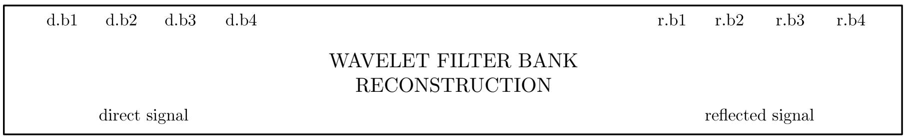


Figure 3.12: Wavelet filter bank reconstruction block

filter bank is dual respect what presented in Figure 3.9. The low-pass and high-pass reconstruction base filters are chosen in such a way that the perfect reconstruction is guaranteed. Once the direct and the reflected wavefronts have been reconstructed we are able to proceed to the sound absorption coefficients estimation, whose block is described in the last section of this chapter. Instead, the next section describes the implementation of the sub-array processing blocks, pointing out the difference between the direct and the reflected wavefront cases.

3.4 Superdirectional microphone array

In order to realize a superdirectional acquiring device, the sub-arrays have to be implemented using a fourth order differential microphone array. From Chapter 2 we know that such a device has a narrow beamwidth of about 57 degrees, in the ideal case. Moreover, in order to split in an efficient way the direct and the reflected wavefronts we desire that the microphone array maximizes the front-to-back ratio, also providing a technique to swap the direction in which we are *listening to*. The implementation for a generic direct signal sub-array processing block is reported in Figure 3.13. The inputs are represented by the proper sub-band of the microphones related to that sub-array. Instead, the output is the related sub-band of the direct signal. The boxes denoted by τ_i are simply time delay filters that, if properly chosen, allow to get the desired beamform.
Assuming small spacings between the microphones (i.e. $kd \ll \pi$ and $\omega\tau \ll \pi$), the directivity for a fourth order differential microphone

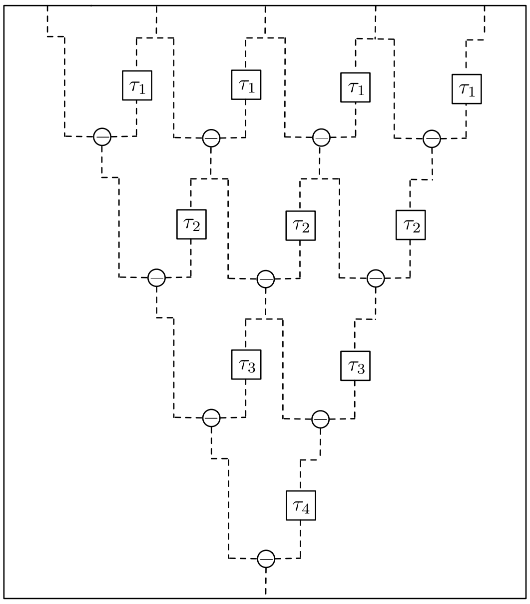


Figure 3.13: Fourth order differential microphone array implementation. The desired beamform pattern is achieved by choosing properly the tunable time delays τ_i

array, is given by

$$E(\omega, \theta) \approx P_0 \omega^4 \prod_{i=1}^{4} [\zeta_i + (1 - \zeta_i) \cos\theta], \tag{3.14}$$

where P_0 is the wave amplitude, ω is the angular frequency and ζ_i is the i^{th} tunable coefficient. As the value of τ_i depends on ζ_i, adjusting the values assumed by ζ_i we can achieve the desired beamform pattern. Expanding the product of equation (3.14), we get

$$E(\omega, \theta) = P_0 \omega^4 (a_0 + a_1 \cos\theta + a_2 \cos^2\theta + a_3 \cos^3\theta + a_4 \cos^4\theta). \tag{3.15}$$

Notice that ζ_i represents the roots of equation (3.15). In order to get the maximum front-to-back ratio for a fourth order differential array, it can be proved that the optimal solution is given by the following values of a_i:

$$\begin{aligned} a_0 &\approx 0.0036, \\ a_1 &\approx 0.0670, \\ a_2 &\approx 0.2870, \\ a_3 &\approx 0.4318, \\ a_4 &\approx 0.2107. \end{aligned} \tag{3.16}$$

Substituting these values in equation (3.15) we find the following roots:

$$\begin{aligned} \zeta_1 &\approx 0.0757, \\ \zeta_2 &\approx 0.4916, \\ \zeta_3 &\approx 0.2735, \\ \zeta_4 &\approx 0.4228. \end{aligned} \tag{3.17}$$

Notice that the ordering of the roots is not significative for our purposes. Knowing that the i^{th} time delay τ_i related to the $\imath^{th}$ tunable coefficient ζ_i is defined as

$$\tau_i = \frac{d\zeta_i}{c(1-\zeta_i)}, \tag{3.18}$$

where d is the microphone spacing and c is the sound speed, we are able to define all the elements that appear in the sub-array block implementation. The value of τ_i depends on the inter-element spacing d and on the sound propagation speed c, besides the value of ζ_i. Consequently, knowing the values of c, ζ_i and d, the calculation for the i^{th} tunable time delay τ_i is straightforward. Table 3.6 presents the values for the time delays related to each sub-array.

	Sub-array$_1$ (d=0.15m)	Sub-array$_2$ (d=0.10m)	Sub-array$_3$ (d=0.05m)	Sub-array$_4$ (d=0.025m)
τ_1 [s]	3.58×10^{-5}	4.23×10^{-4}	1.65×10^{-4}	3.20×10^{-4}
τ_2 [s]	2.39×10^{-5}	2.82×10^{-4}	1.10×10^{-4}	2.14×10^{-4}
τ_3 [s]	1.19×10^{-5}	1.41×10^{-4}	5.49×10^{-5}	1.07×10^{-4}
τ_4 [s]	5.97×10^{-6}	7.05×10^{-5}	2.74×10^{-5}	5.34×10^{-5}

Table 3.6: Time delays related to each sub-array

The calculation for the time delays presented above assumes $c = 343\ \frac{m}{s}$, which is the sound speed propagation in the air, measured at 20 °C. Since we are dealing with discrete signals sampled at 44100 Hz, we know that the sampling period is set to $T = 2.2676 \times 10^{-5}\ s$. Moreover, this value represents the minimum time delay that we are able to apply to a signal. As we can see from Table 3.6, some of the time delays are smaller than the sampling period value, or they are not exactly its multiples. In order to address this issue we have to apply fractional delays, avoiding errors due to the introduction of approximated values. Thus, the actual implementation provides a technique for applying fractional

delays to the signals, based on the cubic Lagrange interpolation [14].
What we have presented up to now is valid for the direct signal processing. Consequently, the previous discussion covers the sub-array processing blocks of Figure 3.10, identified by $B1.D$, $B2.D$, $B3.D$ and $B4.D$. We are now interested in finding a solution that, working on the same principles and using the same physical device, is able to *listen* to the opposite direction respect the one we have considered up to now. The simplest solution is the one in which we turn physically the microphone array by 180 degrees. This method is not valid since we have to observe the direct and the reflected wavefronts in the same time instant. Instead, what we can easily do is to turn the microphone array *by software*, inverting the order in which the sub-bands coming from the microphones are fed into the sub-array processing blocks. It can be proved that inverting the order of the incoming signals or applying negative time delays respect to the ones reported in Table 3.6, is perfectly equivalent. As the actual implementation of the fractional time delays algorithm requires positive values, we are forced to use the solution that inverts the order of the input signals. In order to realize such an implementation, the blocks for the direct signal array processing could be reused, but we decided to make this distinction clearer introducing independent reflected signal processing blocks. These blocks, identified by $B1.R$, $B2.R$, $B3.R$ and $B4.R$, are the clones of the direct ones, but they receive the input signals in the reverse order. In this way we are able to *listen* to a direction and to the opposite one in the same moment, getting the direct and the reflected wavefronts separately. The normalized directivity pattern for a direct signal array processing block and its inverted clone is reported in Figure 3.14. As we expect the pattern is perfectly symmetrical.
The output of each of these array processing blocks represents a particular band of the direct or reflected signal. Once the sub-bands are available at the output of each sub-array processing block, they are sent to the *wavelet filter bank reconstruction* block in order to reconstruct the whole direct and reflected signals, as described in the previous section.
Before proceeding to discuss the block involved in the estimation of the sound absorption coefficients, we want to better analyze the array directivity. Even if a constant directivity over the frequency range we are interested in is not strictly required by our application, it is how-

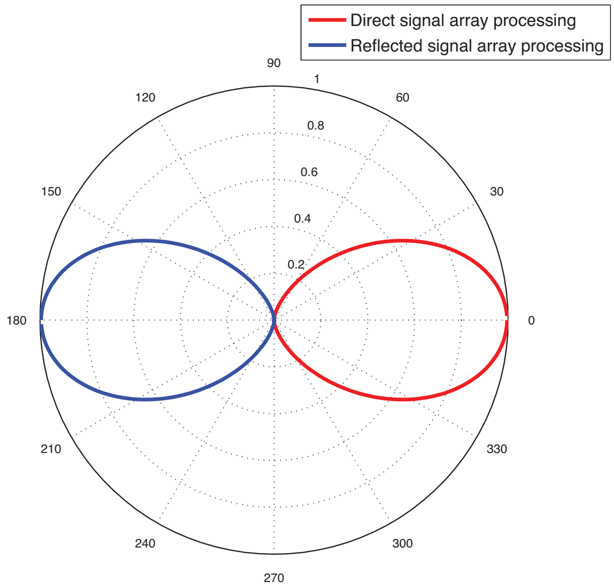


Figure 3.14: Theoretical normalized directivity pattern for a direct signal array processing block and its reverse version

ever a desirable property. In this way, the microphone array developed for this work could be easily reused for other applications. Neglect for a moment the ω^4 term that appears in equation (3.15). In this case, the constant directivity over the frequency range of each sub-array is achieved by choosing the maximum frequency of each range too far from the one allowed by the spatial sampling theorem. In fact, as we get closer to the upper bound of the frequency range to the spatial sampling limit frequency, as the beamwidth becomes larger. The previous constant directivity condition is fully satisfied since, as Table 3.3 reports, the maximum frequency of each sub-array range is always less than the half of the maximum frequency allowed by the sampling theorem. The only exception is represented by the 5^{th} sub-band, for which we consider only the frequencies up to 3000 Hz. The frequencies beyond this limit are safely discarded using a low-pass filter. Reintroduce now the ω^4 term of equation (3.15). Filtering the signals coming from each sub-array with a low-pass filter, whose frequency response is proportional to ω^{-4}, is sufficient to get the desired constant directivity.
Instead, some problems arise when considering the system as a whole (i.e. we are watching to the direct and reflected wavefronts at the output of the *wavelet filter bank reconstruction* block). In this case, the

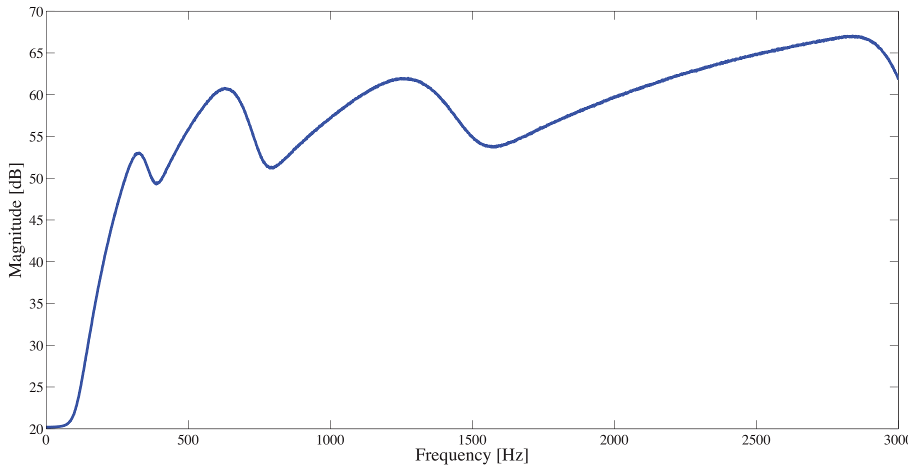


Figure 3.15: Microphone array directivity as function of the frequency, measured at 0 degrees for an MLS source signal (over multiple realizations).

dependency of the directivity to the ω^4 term is superimposed to the one introduced by the microphone spacing d term. The situation for an MLS wavefront (over a high number of realizations) arriving at zero degree respect to the array is reported in Figure 3.15.
The repetition of a scaled version of the same *pattern* for each sub-band is due to the d dependency, since the microphone spacing varies for each sub-array. Instead, the exponential-like behaviour for each *pattern* repetition is due to the ω^4 term. Notice that the ideal case is the one in which we obtain a constant line over the frequency axis, which represents a constant directivity along all the frequencies. In order to compensate the non-constant result presented in Figure 3.15, we proceed in the same way as we have previously done for the source signal pre-equalization: we use the parametric spectral estimation. Once we get the *shaping filter*, the calculations to get the *whitening* one are straightforward. The direct and the reflected wavefronts, coming from the wavelet reconstruction block, are filtered with the obtained *whitening filter*. After this compensation, the directivity for the fourth order differential microphone array is presented in Figure 3.16. The same result for all the angles of arrival in the $0 \div 180$ degree range is reported in Figure 3.17. Notice that for angles between 180 and 360 degrees the microphone array pattern is mirrored respect to the x-axis. After the compensation introduced above, the directivity along the frequency axis is more uniform, even if we have some leaks in the

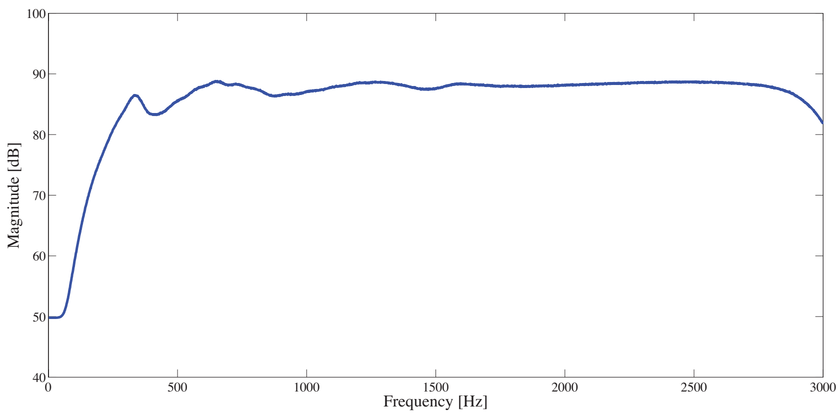


Figure 3.16: Microphone array directivity as function of the frequency, measured at 0 degrees for an MLS source signal, after the whitening filtering (over multiple realizations).

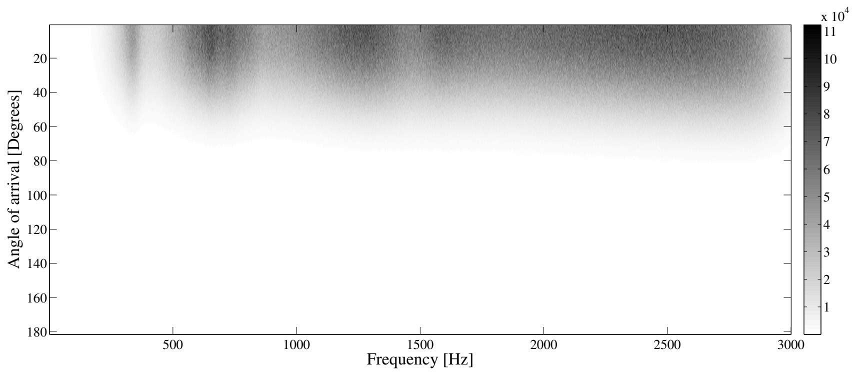


Figure 3.17: Microphone array directivity after the compensation, as function of the frequency and of the angle of arrival (over multiple realizations).

low frequency range. This phenomenon does not represent a problem for our purposes, since the sound absorption coefficients are calculated as the ratio of the direct and the reflected wavefronts, which are managed in the same way (for what regards the frequency response) by the differential microphone array. The beamwidth of the microphone array reaches its maximum at 0 degrees and it decreases gradually reaching its minimum at about 60 degrees, as we expect from the theoretical model.

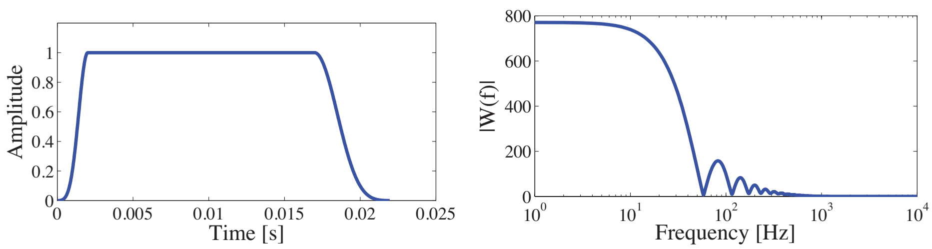
Once the direct and the reflected wavefronts have been calculated and compensated in order to get a frequency response as flat as possible, we are able to proceed to the discussion of the sound absorption coefficients estimation.

3.5 Absorption coefficients

In this section we are going to discuss the last block presented in Figure 3.4, whose goal is to take the estimation of the sound absorption coefficients, as function of the frequency. The inputs of this block are represented by the direct and the reflected wavefronts, properly split. From equation (3.3) we know that sound absorption coefficient, as function of the frequency, is defined as

$$\alpha(f) = 1 - r(f) = 1 - \frac{W^-(f)}{W^+(f)}, \tag{3.19}$$

where $W^-(f)$ and $W^+(f)$ are respectively the energy related to the reflected and the incident wavefront. In the actual implementation $W^-(f)$ and $W^+(f)$ are the energies related to the impulse responses obtained from the respective signals. From Section 4.2 we know that the impulse response of a signal obtained through an MLS source signal is simply defined as the cross-correlation between the output and the input signal. In this case, the output signal is represented by the direct wavefront in one case and by the reflected wavefront in the other one, whereas the input signal is represented in both the cases by the MLS source signal. In order to get the estimation result as function of the frequency, we have to transform the impulse responses from the time domain representation to the frequency one. This task is performed using the Fourier Transform (FT). Moreover, as we are interested only in a small portion of the impulse responses, before taking the FT we have to select them using a proper window. For this kind of application, it has been proved [15] that the ideal window is represented by the so-called *Adrienne* window. It is constructed using a Blackmann-Harris window for the first 2 ms, followed by a 15 ms rectangular window



(a) Time domain response (b) Frequency domain response

Figure 3.18: Adrienne window

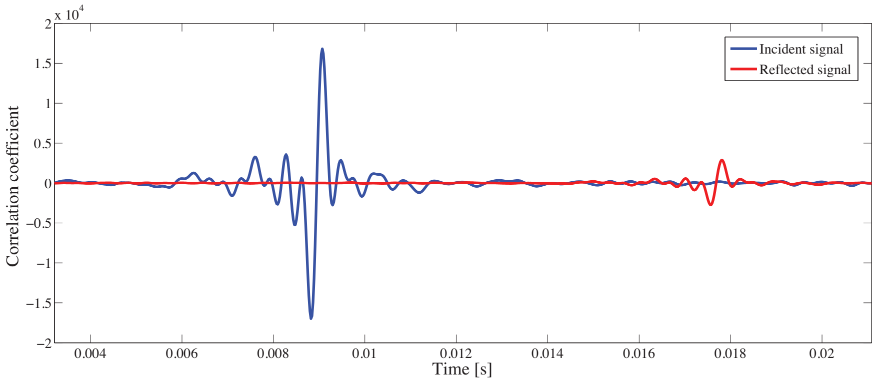


Figure 3.19: Cross-correlations of the incident and the reflected wavefronts, as received at the microphone array reference point.

and a final 5 ms Blackmann-Harris window. The result is reported in Figure 3.18.

Such a window has to be applied 2 ms before the occurrence of the first peak in the incident or reflected impulse response. Introducing the new features discussed above, equation (3.19) reduces to

$$\alpha(f) = 1 - r(f) = 1 - \left[\frac{|FT[c_{dir,in}\ W(t)]|}{|FT[c_{ref,in}\ W(t)]|} \right]^2, \tag{3.20}$$

where $c_{dir,in}$ and $c_{ref,in}$ are respectively the cross-correlation between the incident and the source MLS signal in one case and the cross-correlation between the reflected and the source MLS signal in the other one, and $W(t)$ is the *Adrienne* widow in the time domain.
In the simulation case in which the absorption coefficient has been set to 0.5 over all the frequency range, the cross-correlations of the incident and the reflected wavefronts are those reported in Figure 3.19.
Notice that the time delay between the peak of the direct and the reflected wavefronts, is due to the different path lengths of each of the two signals. Knowing the sound speed c, the calculation of the path difference (in meters) is given by

$$\Delta d = c\ \Delta t, \tag{3.21}$$

where Δt is inter-peak time difference. If the situation is the one reported in Figure 3.20, where A represents the reference point of the microphone array (i.e. microphone no.7), the distance between the surface under analysis and the microphone array reference point is given

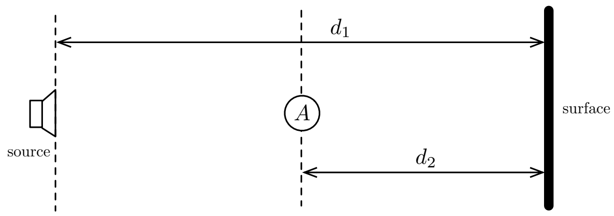


Figure 3.20: Definition of the distances between the loudspeaker, the microphone array and the surface under analysis

by

$$d_2 = \frac{\Delta d}{2}, \tag{3.22}$$

where Δd is the one defined in equation (3.21).
This is because the reflected wavefront has to cover the d_2 path both in forward and reversed direction before it could be received by the microphone array. Consequently, the distance Δt, calculated using the time delay between the direct and the reflected responses, is exactly twice the value of d_2. Notice that in this calculation we can neglect the delay introduced by the processing system, since both the signals are processed in the same way and so they are affected by the same time delay.
In the discussion presented up to now we have not considered any aspect related to the sound decay, related to the wavefronts propagation. From Chapter 2 we know that the amplitude of a spherical wave decreases as the inverse of its distance from the source. From Figure 3.20 we can see that the direct wavefront covers a distance equal to $(d_1 - d_2)$, whereas the reflected one covers a distance equal to $(d_1 + d_2)$. In order to find a solution for the sound absorption coefficient estimation problem that takes into account the sound decay due to the wavefront propagation, we have to modify equation (3.20) in the proper way. Assuming that the wavefronts involved in the measurements are spherical and that the situation is the one reported in Figure 3.20, equation (3.20) reduces to

$$\alpha(f) = 1 - r(f) = 1 - \left[\frac{|FT[c_{dir,in}\, W(t)]|}{|FT[c_{ref,in}\, W(t)]|}\, \frac{(d_1 + d_2)}{(d_1 - d_2)}\right]^2, \tag{3.23}$$

where the

$$\frac{(d_1+d_2)}{(d_1-d_2)}$$

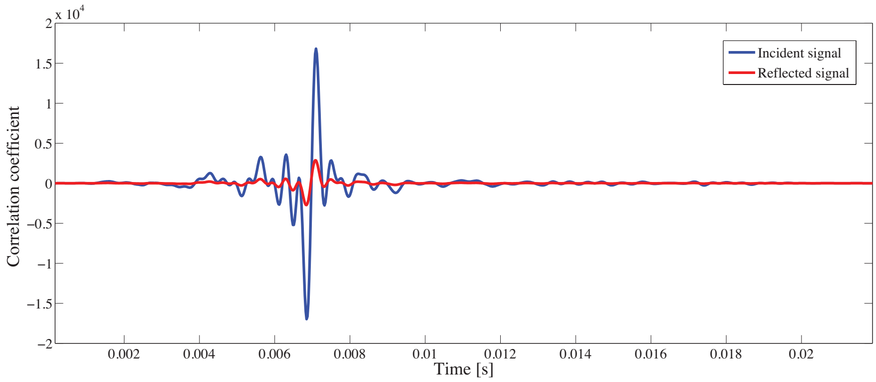


Figure 3.21: Impulse responses of the direct and the reflected wavefronts, after the windowing operation.

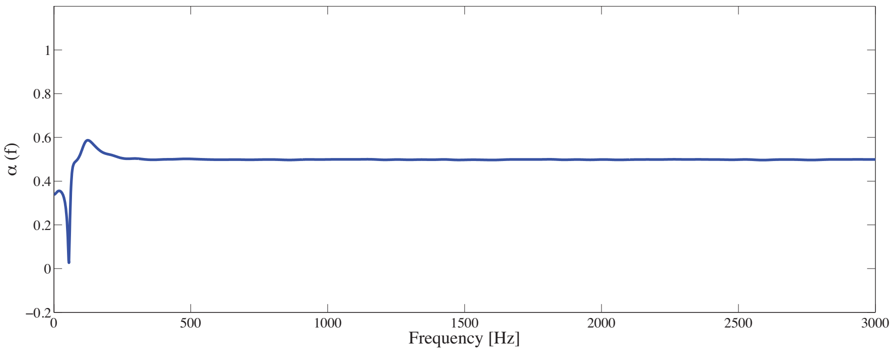


Figure 3.22: Sound absorption coefficient, as function of the frequency, for a simulation test case (α = 0.5).

ratio is the amplitude decay compensation term. In order to get a more appreciable solution, in the actual implementation the distances d_1 and d_2 are not the nominal ones, but those that are estimated using the time delays between the direct and the reflected impulse responses. Moreover, in order to prove the correctness of the proposed compensation factor, the real propagation and decay model has been measured for the environment in which we have taken the measurements, as described in theChapter 4.

Going back to the example previously introduced, the situation after the windowing operation realized by the *Adrienne* window is reported in Figure 3.21. As we expect, the impulse responses are perfectly superimposed. Lastly, applying the calculation reported in equation

(3.23), we get the desired result as reported in Figure 3.22. This figure represents the sound absorption coefficient as function of the frequency. Notice that in the low frequencies we encounter some artifacts that the proposed system is not able to avoid. In addition to that, also under real conditions these low frequencies could represent a problem since the small diameter loudspeaker could not radiate them in a proper way. In order to avoid a wrong interpretation of the data related to these frequencies, we choose to limit the frequency range of the sound absorption coefficient estimation in the $300 \div 3000$ Hz interval. Consequently, the signals involved in the calculations presented above are processed with a high-pass filter, whose cut-off frequency is set to 300 Hz.

Chapter 4

Validation

This chapter presents the results for the sound absorption coefficient estimation technique, proposed by this thesis. Both the simulation and the experimental results are presented. In particular, in Section 1 we take into account the simulations and the detailed description of the involved model. Details about the variables and real world phenomena included in the model are provided, as well.
Instead, Section 2 presents the results coming from the real world experiments. Moreover, the results of some *auxiliary* measurements, required in order to make the system working properly, are presented too.

4.1 Simulations

In order to perform the simulations required to characterize the proposed solution, we require a model that approximates the real world situation. The desired model has to describe just the variables and the sound phenomena significative for our purposes. The goal is to obtain the signals received by each microphone belonging to the microphone array. As under real conditions the signals measured by the microphones are the superimposition between the incident and reflected wavefront, we have to virtually generate both. Notice that each microphone receives the same version of the direct and reflected wavefronts, simply delayed in time as function of its position. The incident and reflected signals have to be superimposed (i.e. summed up over time) in the simulation case, as well. Moreover, we have to decide how the reflected signal is generated respect to the incident one. The energy

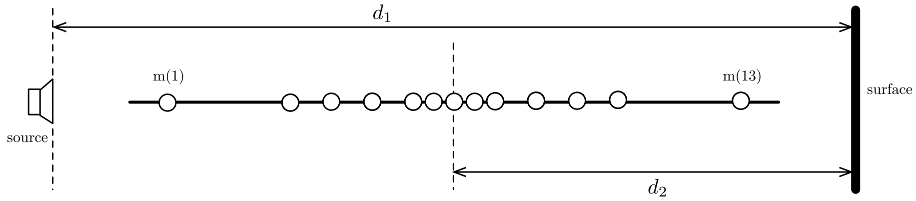


Figure 4.1: Simulation configuration

portion of the incident signal, which is reflected by the wall, is the one that influences the sound absorption coefficient estimation. In the following discussion we take into account the cases for which the reflection coefficient is constant and the one for which it varies over the frequency range.

Consider the situation reported in Figure 4.1. The distances d_1 and d_2 refer to the distance between the virtual source and the surface, and to the distance between the microphone array reference point and the surface under consideration, respectively. If we define $\mathrm{mls}(t)$ as the MLS signal generated at time t by the virtual source, the direct wavefront received by the i^{th} microphone $\mathrm{dir}_i(t)$ and the reflected one $\mathrm{rev}_i(t)$, are simply a delayed version of the source signal. The introduced time delay is proportional to the path length covered by each signal, in order to travel from the source to the microphone. The distances covered by the source signal for the direct and the reflected paths, in order to arrive to the i^{th} microphone, are respectively defined as

$$\Delta d_{dir,i} = d_1 - d_2 + rp_i, \tag{4.1}$$

and

$$\Delta d_{rev,i} = d_1 + d_2 - rp_i, \tag{4.2}$$

where the distances d_1 and d_2 are the ones defined in Figure 4.1, whereas rp_i is the relative positioning for the i^{th} microphone, respect to the microphone array reference point. The values for each rp_i, valid for our configuration, are reported in Appendix C. Once we have introduced the path lengths for the direct and the reflected wavefronts, we are able to define the direct signal received by the i^{th} microphone as

$$\mathrm{dir}_i(t) = \mathrm{mls}(t - \Delta d_{dir,i}/c), \tag{4.3}$$

and the reflected one as

$$\mathrm{rev}_i(t) = \sqrt{\hat{r}}\ \mathrm{mls}(t - \Delta d_{rev,i}/c), \tag{4.4}$$

where c is the sound propagation speed, while $\hat{r}$ is the *power* sound reflection coefficient associated to the virtual surface to estimate. Thus, $\sqrt{\hat{r}}$ represents the *pressure* sound reflection coefficient, suitable to be directly applied to the time-domain pressure signals. Notice that, for the moment, $\hat{r}$ is defined as a constant over all the frequency range. Moreover, from equation (2.43) we know that corresponding *power* absorption coefficient is defined as

$$\hat{\alpha} = 1 - \hat{r}. \tag{4.5}$$

The goal of the simulations is to estimate the value of $\hat{\alpha}$ but, as we are modifying directly the reflected signals, what appear in the model equations is the *power* reflection coefficient $\hat{r}$. However, knowing one coefficient, the calculation of the other one is straightforward.
As the global signal received by the i^{th} microphone is the superimposition of the direct and the reflected wavefronts, it can be defined as

$$\mathrm{mic}_i(t) = \mathrm{dir}_i(t) + \mathrm{rev}_i(t). \tag{4.6}$$

Equation (4.6) is the base equation for the creation of the signals involved in our simulation and it can be gradually refined. This equation does not take into account phenomena such as the sound decay due to the wave propagation, or the variation of sound absorption coefficients as a function of frequency.
Before proceeding to the discussion of a more detailed simulation model, we have to consider an issue that occurs introducing time discrete signals, instead of the continuous ones. The time delay terms $\Delta d_{rev,i}/c$ and $\Delta d_{dir,i}/c$, which appear respectively in equation (4.4) and (4.3), could be smaller than the sampling period T. As we are dealing with discrete signals, we have to find a method in order to apply fractional delays. As well as the situation presented in Section 3.4, where we had to apply fractional delays to the microphone signals in order to get the desired beamform, the solution is represented by the Lagrange cubic interpolation.
If we desire to refine our simulation model in order to consider the decay due to the sound propagation, knowing that for spherical waves the

sound pressure decays as the inverse of the distance from the source, we can rearrange equations (4.3) and (4.4) respectively as

$$\mathrm{dir}_i(t) = \frac{1}{\Delta d_{dir,i}} \ \mathrm{mls}(t - \Delta d_{dir,i}/c), \tag{4.7}$$

and

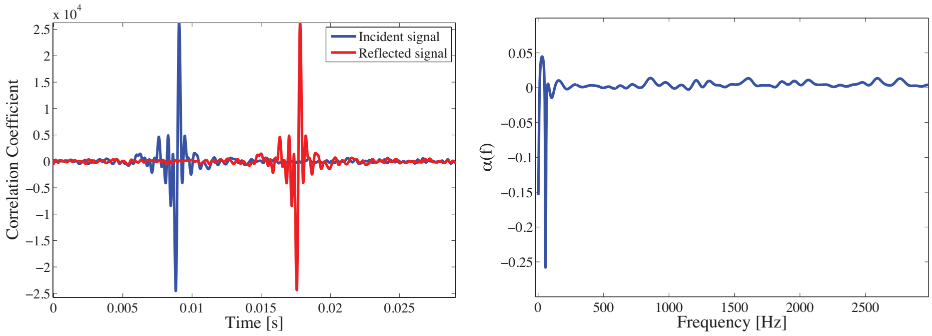
$$\mathrm{rev}_i(t) = \frac{1}{\Delta d_{rev,i}} \sqrt{\hat{r}} \ \mathrm{mls}(t - \Delta d_{rev,i}/c), \tag{4.8}$$

where $1/\Delta d_{dir,i}$ and $1/\Delta d_{rev,i}$ are the terms that introduce the desired decay. Notice how, even if we apply the attenuation decay factor microphone-by-microphone, the decay compensation factor is applied once at the array reference point, as described by equation (3.23). A comparison between the two situations shows that this approximation introduces a negligible error.

For the most of real materials, the sound absorption coefficient is not constant over the frequency range. Thus, it is desirable to introduce this characteristic in the simulation model. We recall the fact that the sound absorption coefficient is defined as the frequency-by-frequency ratio between the impulse responses of the reflected and the direct wavefronts, as described in equation (3.23). Consequently, in order to obtain an absorption coefficient that varies over the frequency, it is sufficient to modify the frequency content of the reflected wavefront. This operation is simply performed filtering the reflected signal with a filter, whose frequency response is the one we want to map to absorption coefficient $\hat{\alpha}(f)$. Notice that, before the filtering operation both the reflected and the incident signals are characterized by a flat spectrum. In fact, their wavefronts are a delayed version of the source MLS signal. Defining a generic modelling filter $h(t)$, and modifying equation (4.8), we get:

$$\mathrm{rev}_i(t) = \frac{1}{\Delta d_{rev,i}} \sqrt{\hat{r}} \left[\mathrm{mls}(t - \Delta d_{rev,i}/c) \ * \ h(t)\right], \tag{4.9}$$

where $*$ is the convolution operator. The equation required to describe the direct wavefront is the same of the (4.7) one. In order to simplify the simulations, $h(t)$ was chosen to be a low-pass filter. From equation (4.9) we can notice that the *pressure* reflection coefficient is applied uniformly over all the frequency range. However, as we have introduced the low-pass filter $h(t)$(or equivalently $H(f)$), only the frequencies behind the filter cut-off frequency are influenced by the constant reflection

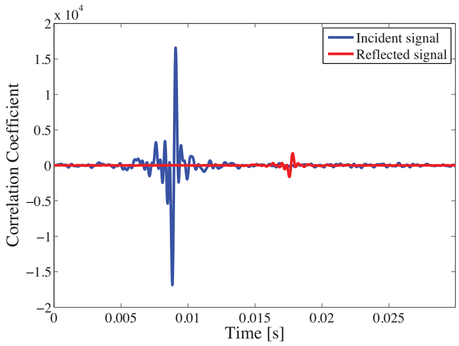


(a) Incident and reflected signal impulse responses (b) Sound absorption coefficient, as function of the frequency

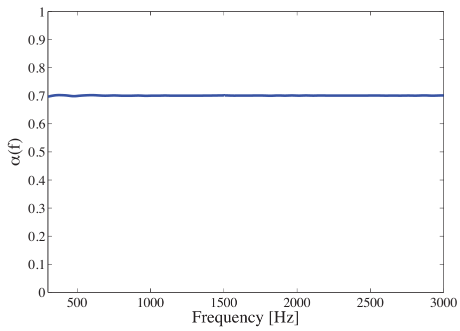
Figure 4.2: Simulation results for a totally reflecting material ($\hat{\alpha} = 0$), with $d_1 = 3m$, $d_2 = 1.5m$ and incident wavefront normal respect to the wall, related to the simplest model described by equations (4.3) and (4.4) (no sound amplitude decay as function of the distance from the source, absorption coefficient constant over the frequency range)

coefficient $\sqrt{\hat{r}}$. The energy carried by all the other frequencies is to be considered null, due to the filtering operation. Thus, we expect to obtain an absorption coefficient close to $1 - \hat{r}$ for what regards the frequencies behind the $H(z)$ cut-off frequency, and close 1 for the other ones.

Once we have defined all the possible models involved in the simulations, we proceed to the discussion about the obtained results. Notice that the sampling frequency is set to 44100 Hz, both in the simulation and in the experiment cases. Figure 4.2 presents the simulation results obtained by choosing as model the one described by equations (4.3) and (4.4). In this case the sound amplitude decay is not taken in to account and the sound absorption coefficient $\hat{\alpha}$ is set to 0, over all the $0 \div 3000$ Hz frequency range (i.e. the wall is totally reflective). The distances d_1 and d_2 are respectively set to 3 m and 1.5 m, while the incidence angle of the direct wavefront is normal respect to the wall. Figure 4.2(a) presents the impulse responses for the incident and the reflected signals, while Figure 4.2(b) presents the sound absorption coefficient estimation. As we can notice, the estimated values are very close to the nominal ones, except for some spurious components in the lowest frequency range. Some values are even negative and this fact represents an error, since the values of the absorption coefficient have to be limited in the $[0, 1]$ range. As anticipated in the previous chapter, in order to avoid a wrong interpretation of the data related to these



(a) Incident and reflected signal impulse responses



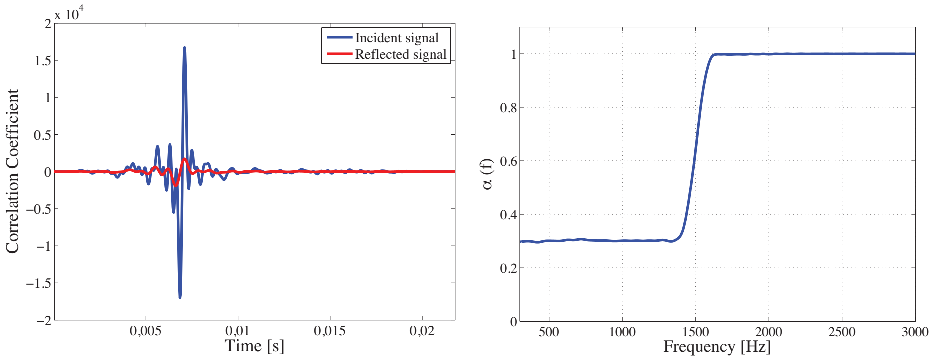
(b) Sound absorption coefficient, as function of the frequency

Figure 4.3: Simulation results for an absorbing material ($\hat{\alpha} = 0.7$), with $d_1 = 3m$, $d_2 = 1.5m$ and incident wavefront normal respect to the wall, related to the model described by equations (4.7) and (4.8) (sound amplitude decay as function of the distance from the source, absorption coefficient constant over the frequency range)

frequencies, we chose to limit the frequency range in the $300 \div 3000$ Hz interval. This operation is done by filtering the signals involved in the calculations with a high-pass filter, whose cut-off frequency is set to 300 Hz. For this reason, the presented results are limited in the $300 \div 3000$ Hz range.

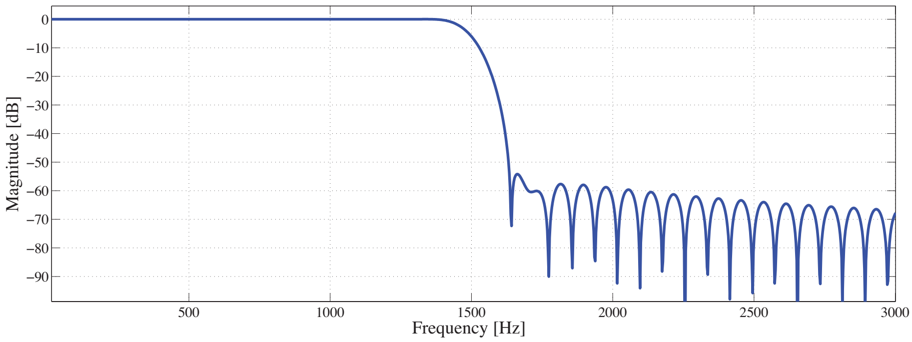
Figure 4.3 presents the simulation result for an absorbing material, where $\hat{\alpha}$ is set to 0.7 over all the frequency range. The distances d_1 and d_2 are the same of the previous simulation, but in this case the chosen model is the one that takes into account the sound amplitude decay, due to the wavefront propagation. The equations that realize such a model are the ones presented in (4.7) and (4.8). From Figure 4.3(a) we immediately deduce that the surface under test is an absorbing material, since the impulse response of the reflected wavefront is very small respect to the incident one. The absorption coefficient estimation is presented in Figure 4.3(b). Even in this case, the estimated value is almost identical to the expected one.

Lastly, the most interesting simulation case is presented in Figure 4.4, where the sound absorption coefficient varies over the frequency range. This feature is achieved using the model equation (4.9), in which the frequency content of the reflected wavefront is modelled by a generic filter $h(t)$ or, equivalently, by its Fourier Transform $H(f)$. The filter chosen for this simulation is a low-pass filter, whose frequency response is reported in Figure 4.4(c). As we can see, the cut-off frequency is set



(a) Incident and reflected signal impulse responses, superimposed

(b) Sound absorption coefficient, as function of the frequency



(c) Modelling filter H(f), used to modify the frequency content of the reflected wavefront. The cut-off frequency is set to 1500 Hz

Figure 4.4: Simulation results for an absorbing material ($\hat{\alpha}(f) = 0.3$ for $f \leq 1500$ Hz and $\hat{\alpha}(f) = 1$ for $f > 1500$ Hz), with $d_1 = 3m$, $d_2 = 1.5m$ and incident wavefront normal respect to the wall, related to the model described by equations (4.8) and (4.9) (sound amplitude decay as function of the distance from the source, absorption coefficient that varies over the frequency range)

to 1500 Hz, which is exactly in the middle of the frequency range we are interested in. Furthermore, $\hat{\alpha}$ is set to 0.3 for the frequencies able to pass through the filter ($f < 1500$ Hz). This situation results in having a complete sound absorption for the frequencies above the 1500 Hz, and in having a sound absorption equal to $\hat{\alpha}$ for the frequencies below that limit. As for the other simulation cases, d_1 and d_2 are respectively set to 3 m and 1.5 m. The simulation results are reported in Figure 4.4(a) and in Figure 4.4(b). Notice that in this case we prefer to report the impulse responses of the direct and the reflected wavefronts superimposed, rather than in the *original* time position. In fact, the filtering operation realized by the digital filter $h(t)$ introduces a time delay on the reflected signal, which makes the usual representation

meaningless. The sound absorption coefficient estimation is very close to the expected values. The gradual transition of $\hat{\alpha}(f)$ in the 1500 Hz region is simply due to the $H(f)$ filter characteristics.
The simulation tests presented above cover all the models introduced in the first part of this section. In the simulation case, the proposed system provides appreciable results, as proved by the presented data. The discussion about the results obtained in the case of real world measurements is presented in the next section.

4.2 Experiments

This section describes the results obtained from the real measurements, taken in a semi-anechoic room, for some common materials. Some *auxiliary* measurements, such as the ones gathered in order to study the actual sound decay, and the ones acquired to achieve the gain matching between different microphones, are presented, as well.
From Chapter 3, we know that the most complete equation in order to get the sound absorption coefficient estimation is given by

$$\alpha(f) = 1 - r(f) = 1 - \left[\frac{|FT[c_{dir,in}\ W(t)]|}{|FT[c_{ref,in}\ W(t)]|}\ \frac{(d_1 + d_2)}{(d_1 - d_2)} \right]^2, \qquad (4.10)$$

where the $(d_1 + d_2)/(d_1 - d_2)$ ratio is used to compensate the wavefront amplitude decay. The distances d_1 and d_2 are the same defined in Figure 4.1. The compensation is done assuming that the wavefront amplitude decays proportionally respect to the inverse of the distance from the source. Consequently, the sound power decays linearly respect to the inverse of the distance square. In order to be sure to apply the correct decay model for the actual environment, we have estimated it. The measurements was taken testing all the microphones belonging to the array, at some significative distances in the $50 \div 270$ cm range. For each microphone-distance configuration, 100 realizations of a 3-period MLS signal, have been recorded. The signal energy per time interval (i.e. the sound power) for the i^{th} microphone at distance d from the source, averaged over 100 observations, and defined as

$$P_{i,d} = \frac{1}{O} \sum_{o=1}^{O} \left[\frac{1}{T_{obs}} \int_{-T_{obs}/2}^{T_{obs}/2} |x_{(i,d,o)}(t)|^2\ dt \right], \qquad (4.11)$$

has been calculated. Notice that $x_{(i,d,o)}$ is the o^{th} observation taken by the i^{th} microphone at distance d from the source, T_{obs} is the time inter-

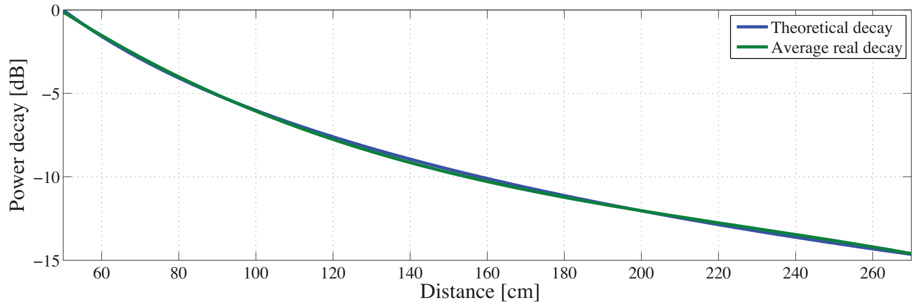


Figure 4.5: Comparison between the theoretical sound power decay and the real one, interpolated over the acquired data.

val in which the signal is observed, and O is the number of realizations. As we are dealing with time discrete signals, equation (4.11) has to be rearranged as

$$P_{i,d} = \frac{1}{O} \sum_{o=1}^{O} \left[\frac{1}{N} \sum_{n=0}^{N-1} |x_{(i,o)}[n]|^2 \right], \tag{4.12}$$

where N is the number of samples for which the recorded signal is observed and n is the discrete time index. The sound power decay PD for each distance d, averaged over all the microphones, can be expressed as

$$PD_d = 10 \log_{10} \left(\frac{\frac{1}{M} \sum_{i=1}^{M} P_{i,d}}{P_{ref}} \right), \tag{4.13}$$

where M is the number of microphones required by the array and P_{ref} is the reference sound power averaged over all the microphones, defined as

$$P_{ref} = \left[\frac{1}{M} \sum_{i=1}^{M} P_{i,d} \right]_{d=50cm}. \tag{4.14}$$

The results obtained from equation (4.13), expressed in dB, are reported in Figure 4.5, together with the theoretical expected values. As we can notice, the real and the theoretical decay models are very close. Consequently, we can apply the compensation factor included in equation (4.10), introducing a neglegible error.

As we have seen that the decay model chosen by assumption for the amplitude decay compensation is coherent with the real one, we can proceed to the discussion of the experimental setup, reported in Figure 4.6. The MLS sequence is sent from the PC to the ADC / DAC board and, once the DAC conversion is completed, the signal is ampli-

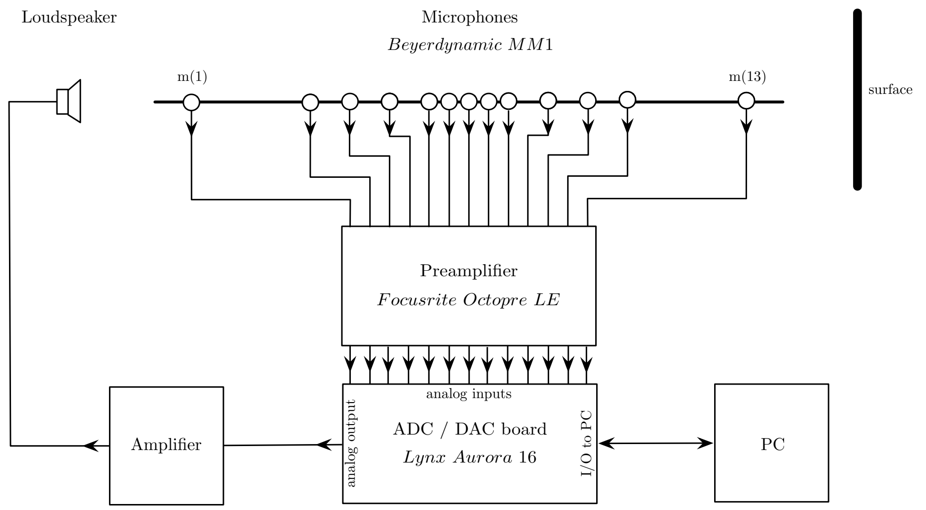


Figure 4.6: Experimental setup. The digital MLS signal, generated by the PC, is converted into an analog one by the ADC/DAC board. Once the conversion is completed, the signal is preamplified and emitted by the loudspeaker, placed normally respect to the surface under test. The analog signals received by the array microphones are preamplified and converted into digital ones, by the ADC/DAC board. Finally, the digital signals are sent to the PC for the required calculations.

fied and sent to the loudspeaker. Notice that, as the theory requires, the loudspeaker is placed in such a way that its emitted wavefronts are normally incident respect to the surface under consideration. Instead, the superdirectional microphone array is placed between the loudspeaker and the surface. For simplicity, the distances between the various elements are the same of the ones defined in Figure 4.1. Once the signal is emitted, the array microphones receive the generated wavefronts. Each signal is preamplified and sent to the ADC / DAC board. Lastly, when the conversion is completed, the signals are sent to the PC in order to achieve the required calculations. The specifications for the devices involved in the experimental setup are reported in Appendix D.
Before proceeding to the discussion of the experimental results, a further step is required. In order to emit an MLS signal with the desired properties (i.e. flat spectrum and short impulse responses) also in the real world case, a pre-equalization technique is required. With such a technique, we are able to compensate the linear distortion introduced by the loudspeaker and the cascade of measuring devices, as discussed in Section 2.4 and 3.2. In the actual setup, the MLS signal is pre-equalized at the PC-level, using a digital *whitening* filter. Another

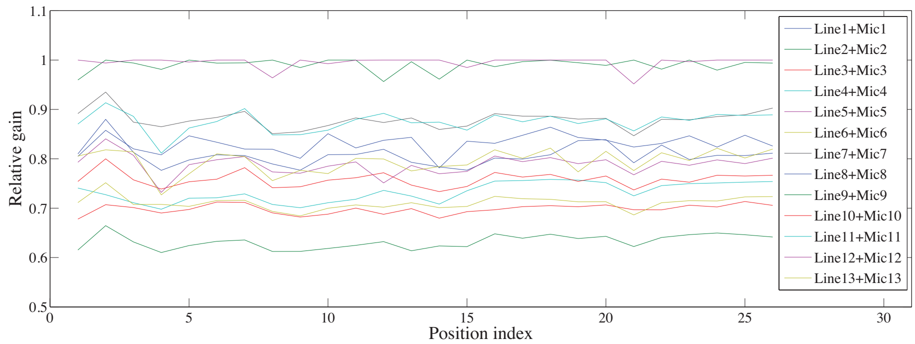


Figure 4.7: Relative line-microphone gains, at some significative source-microphone distances. For each position, moving on the y-axis, we can seen how each microphone gain is scaled respect to the others. Instead, moving on the x-axis we can see how the relative gain for each microphone varies as function of the position. Lower position indexes refer to microphones closer to the source.

issue is represented by the fact that the system works under the assumption that all the microphones are characterized by the same gain: the assumption is valid for the theoretical case, but not for the experimental one. In fact, due to factory differences, the microphones have (minimal) different characteristics, even if they are the same model, produced by the same manufacturer. The situation is further worsened by the use of the preamplifier, where the lines related to each microphone input have their own gain knob. As the preamplifier knobs are analog, it is really difficult to set the same gain for all the microphone lines. This situation, in addition to the one presented above, results in having a different gain for each microphone. Thus we require a way to normalize the line-microphone gains. Starting from the data gathered for the study of the sound decay and considering equation (4.12), the relative gain G for each i^{th} line-microphone at distance d from the source is defined as

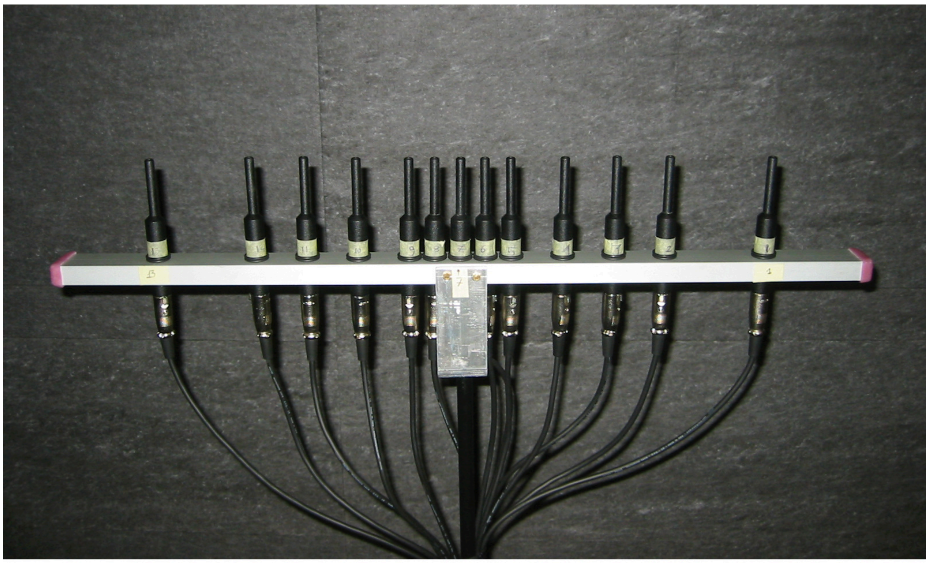
$$G_{i,d} = \sqrt{\frac{P_{i,d}}{\max(P_{i,d})}}, \tag{4.15}$$

where $\max(P_{i,d})$ is the maximum value of $P_{i,d}$ over all the microphones, measured at the same distance d from the source. The results are reported in Figure 4.7 where, in order to get a more compact representation, the distances are substituted by a position index. The distances at which the measurements are taken, are the ones at which the incident or the reflected signals are received by each microphone. The direct and the reflected path lengths are calculated considering the actual d_1 and d_2 distances and the positioning of each microphone on the array. Considering Figure 4.7, moving on the y-axis we can see how the gain of each microphone is scaled respect to the others while, moving on the x-axis, we can see how the relative gain of each microphone varies as function of the position. Lower position indexes refer to microphones closer to the source, and vice versa. As we expect for each position we encounter a different line-microphone gain. This issue can be solved introducing the amplitude normalization factor $N_{i,d}$, defined as

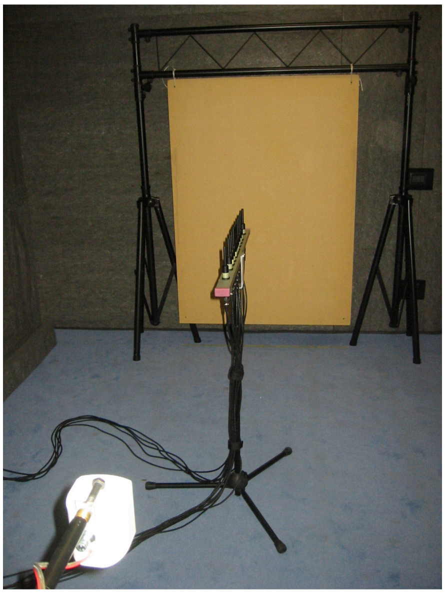
$$N_{i,d} = \frac{1}{G_{i,d}}, \tag{4.16}$$

which has to be applied to each amplitude sample of the signal acquired by the t^{th} microphone, at position d from the source. Another important thing to notice is that the microphone relative gains vary as function of the position. This means that the sensitivity of each microphone is not constant as a function of distance from the source. The artifacts introduced by this issue can degrade the solution, since an efficient compensation technique is not available. In fact, as during the experimental measurements each microphone is placed in a different position respect to the other ones, the only thing that can be done is to apply a gain normalization factor, calculated as the mean over different distances.

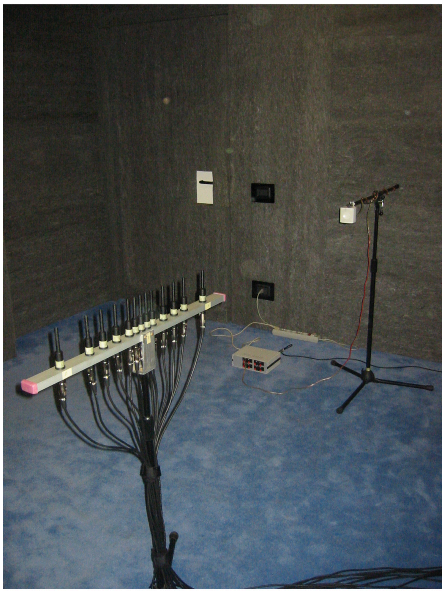
By means of the techniques presented above, we are able to compensate, or at least to limit, the artifacts that arise in the experimental case. Thus, we can proceed to the discussion of the results obtained by the real world measurements. All these measurements have been taken in a semi-anechoic room (4 m $\times$ 4 m $\times$ 3 m). A semi-anechoic room is a chamber in which the walls are almost totally sound absorbing. In this way the reverberation phenomenon is minimized and the propagation conditions are closer to the free field ones. The configuration is the same discussed up to now, both in theory and in simulation cases. The loudspeaker is placed in front of the surface under test, in order to make the emitted wavefronts normally incident respect to the surface, as reported in Figure 4.8(b) and in Figure 4.8(c). The superdirectional microphone array (Figure 4.8(a)) is placed between the loudspeaker and the wall, inline respect to the travelling direction of the incident wavefront . The definition of the d_1 and d_2 distances is the same provided for the simulation cases, and it is reported in Figure 4.1. For all



(a) Actual superdirectional microphone array, used during the experiments. It is made up of 13 microphones (Beyerdynamic MM1), and each of them is linked to the preamplifier unit using standard XLR cables

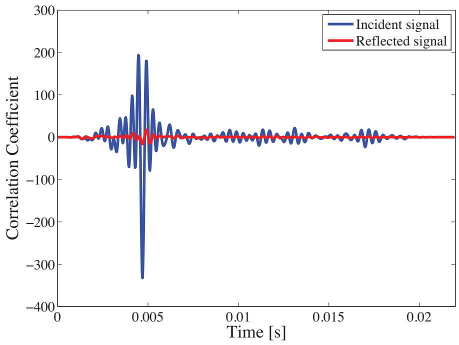


(b) Experimental configuration. The surface under test is placed normally respect to the loudspeaker, while the microphone array is placed between them

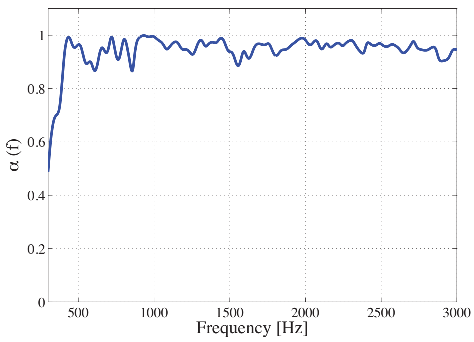


(c) Experimental configuration. The microphone array is placed inline respect to the loudspeaker emitting direction

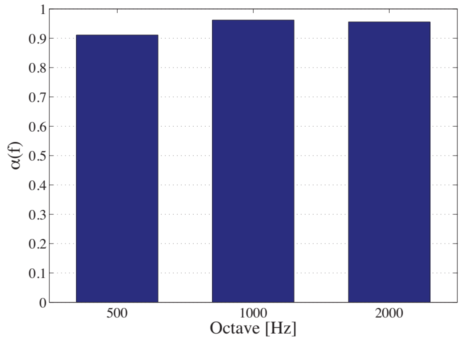
Figure 4.8: Superdirectional microphone array and experimental configuration



(a) Incident and reflected signal impulse responses, superimposed



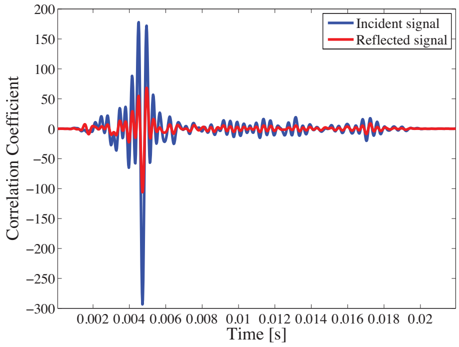
(b) Estimated sound absorption coefficient, as function of the frequency



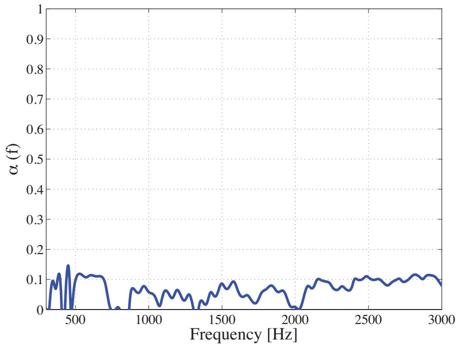
(c) Estimated sound absorption coefficient, averaged over octave bands

Figure 4.9: Absorption coefficient estimation results, for a PVC absorbing wall, tested at $d_1 = 1.6m$, $d_2 = 0.8m$ and with an incident wavefront normal respect to the surface

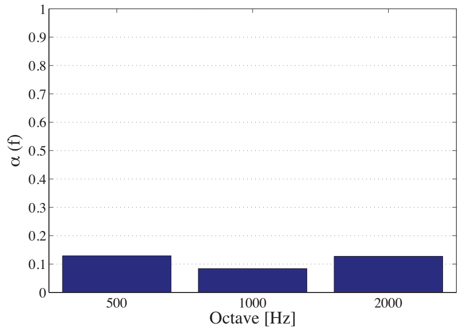
the experiments, d_1 and d_2 are respectively set to 1.6 m and 0.8 m. The first case taken into account is the one for which we want to estimate the sound absorption coefficient for the PVC material that covers the walls of the semi-anechoic room. It is clear that we expect this material should be classified as a higly absorbing one. The results are presented in Figure 4.9. In particular, Figure 4.9(a) presents the impulse responses of the direct and the reflected wavefronts, superimposed, while Figure 4.9(b) presents the sound absorption coefficient estimation, as a function of the frequency. The obtained results are in-line respect to our initial guess. In fact, the material is highly absorbing over all the frequency range, except for the lowest frequencies. As we can notice from the data presented in Appendix A, in the practice, the sound absorption coefficients are reported as function of octave bands, rather than simple frequencies. We adopt the octave band representa-



(a) Incident and reflected signal impulse responses, superimposed



(b) Estimated sound absorption coefficient, as function of the frequency



(c) Estimated sound absorption coefficient, averaged over octave bands

Figure 4.10: Absorption coefficient estimation results, for a plywood panel (2 m $\times$1 m, 10 mm thick), tested at $d_1 = 1.6m$, $d_2 = 0.8m$, and with an incident wavefront normal respect to the surface

tion, as reported in Figure 4.9(c). Unfortunately, we cannot provide the results for the 250 Hz and 4000 Hz octave bands, since the frequency range we are dealing with is limited in the $300 \div 3000$ Hz interval.

The second set of measurements regards the estimation of the sound absorption coefficient for a plywood panel (2 m $\times$1 m, 10 mm thick). From Table A1, we know that this material is classified as a reflecting one. The results obtained from our measurements are presented in Figure 4.10. Although the octave band results are coherent respect to the nominal values, we can notice a problem in the detailed coefficient-frequency representation. In fact, at some frequencies the sound absorption coefficients are negative. These values are physically meaningless, since we know that the absorption coefficients must be limited in the [0,1] range. Notice that during the calculation of the mean sound absorption coefficient, for each octave band, the negative values are neglected. The undesired phenomenon described above is caused by multiple co-existing effects:

- the size of the panel is smaller than the required one. In fact, if the incident wavefront covers an area that is larger than the panel one, part of the wavefront hits the edges of the panel, generating diffraction phenomena. These phenomena deteriorate the reflected signal impulse response. A possible solution is represented by choosing a larger panel, or by putting the loudspeaker closer to the surface. The latter solution is the one we have chosen to limit the described artifact in both the experiments. However, in the first experiment we do not encounter this issue at all, since the surface under test was large as an entire room wall.

- The panel is not fixed in a proper way to the wall. Under this condition, the surface is subject to small vibrations, due to external agents, even if it is not stimulated by an incident wavefront. The undesired vibration phenomenon deteriorates the reflected signal impulse response, as well. In the first experiment measurements this artifact is not present, since the surface under test was a fixed wall, whereas in the second one the artifact is avoided applying proper supports to the surface.

- The maximum front-to-back ratio is not perfectly achieved by the actual implementation of the superdirectional microphone array. The gain mismatch between the microphones and the lack of an efficient technique to compensate them, results in having a microphone array whose directivity pattern is not the same of the nominal one. An example of this situation is reported in Figure 4.11, where a small portion of the direct wavefront is superimposed to the reflected one. This kind of artifact affects both the experimental cases.

- For highly reflecting materials, the sound absorption estimation is strongly affected by measurement noise. In fact, the frequency response of the reflected signal, after applying the decay compensation factor, is similar to the incident one. This fact results in having the ratio between the direct and the reflected frequency impulse responses very sensitive to the measurement noise. On the contrary, for absorbing materials, the difference between the direct and the reflected wavefronts is quite large and this fact guarantees more robustness to measurement noise.

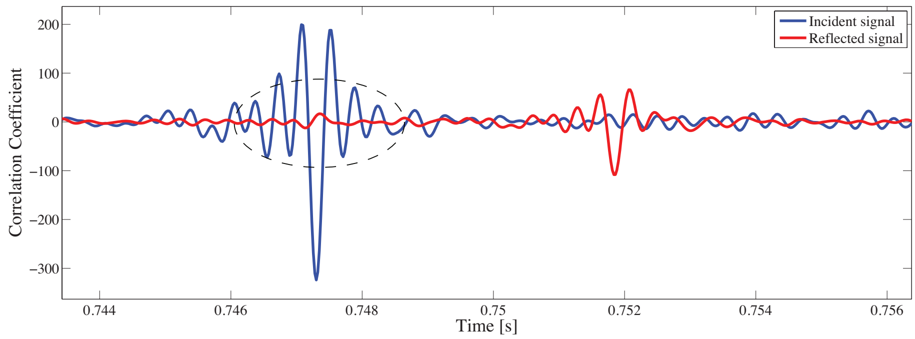


Figure 4.11: Non-optimal separation between the direct and the reflected wavefront. As reported in the highlighted area, a small portion of the direct wavefront is superimposed to the reflected one, due to the microphone gain mismatch.

The quality difference between the results obtained from the two experiments is mainly due to the situation described in the latter point. A further prove of the system behaviour, respect to reflecting materials, is given by the set of measurements taken for an MDF panel. As in the case of the plywood panel, we get some erroneous negative values. Moreover, the same issue affects other estimation techniques and, unfortunately, it is still unsolved even in our solution. Summarizing what we have presented in this section, we can say that the proposed system works in a proper way, achieving appreciable results, in the case of absorbing materials. Instead, in the case of reflecting materials, we may encounter some problems, represented by α negative values. If this issue arises, only the octave band representation can be presented as valid output.

Chapter 5

Conclusions and future directions

5.1 Conclusions

In this thesis we have presented a method for the sound absorption coefficients estimation. The knowledge of these acoustical properties assumes great relevance in the environmental acoustic field and could also be used for sound emitting systems, capable of self-calibration according to the characteristics of the surrounding environment. The proposed method is based on *in-situ* measurements, in order to test the surfaces directly in the environments in which they are located. Notice that the surfaces have to be regular and made up of the same material. The measurements are taken by means of a handy device, easy to carry around and to set up. This is the main characteristic that makes this technique different from the laboratory ones, for which small portions of material have to be removed from their original context and tested under highly controlled conditions. The surface under test is stimulated by means of a white-noise like signal. The loudspeaker that emits the test signal is placed in such a way that its outcoming wavefronts are at normal incidence respect to the surface. The sound absorption coefficients are estimated by measuring the energy carried by the incident and the reflected signals. The measurements are taken by means of a superdirectional microphone array, which is able to separate the direct and the reflected signals in an efficient way. Moreover, as this device has a high directivity, it is able to *listen to* the signals coming from a limited set of incidence angles, corresponding to the ones coming from a small spot on the surface under test. In this way, the method is robust against the reverberation phenomena, generated by undesired secondary reflections. Notice, however, that the real microphone array directivity has some differences respect to the theoretical one, due to the gain mismatch between different microphones. Thus, a compensation technique is provided in order to make this difference negligible. The results for the sound absorption coefficient estimation are provided as function of the frequency, in the $300 \div 3000$ Hz range. Outside this frequency range we encounter some artifacts that make the results not valid, or even meaningless. Moreover, as the method configuration requires that the loudspeaker is placed at normal incidence respect to the surface, the solution does not provide the estimation of the sound absorption coefficients as a function of the incidence angle. From Chapter 4, we have seen that the proposed solution is able to estimate with appreciable results the sound absorption coefficients for absorbing materials, whereas some issues may arise in the case of highly reflecting ones. This is due to the fact that in this case the reflected signal is very similar to the incident one and, as the sound absorption coefficient estimation is calculated as the ratio between them, it results in having a higher sensitivity to measurement noise. The first improvements of the proposed solution are the ones that remove the actual system limitations. These ones and few other refinements are proposed in the next section.

5.2 Future directions

The most important refinements to apply to the proposed solution are the ones that remove its actual limitations. In particular:

- extension of the actual frequency range ($300 \div 3000$ Hz), for which the absorption coefficients are estimated, to a wider one. As in the practice the absorption coefficients are reported for the $125, 250, 500, 1000, 2000$ and 4000 Hz octave bands, it is desirable to extend both lower and upper bound limits. For example, a suitable frequency range could be represented by the $0 \div 5000$ Hz one. However, it is difficult to achieve this frequency range for practical reasons. In particular, the loudspeaker suffers of severe distortions in the emission of the lower frequencies, whereas the upper bound is limited by the spatial sampling theorem (i.e. small array inter-element spacing) in conjunction with the microphone dimensions.

- Introduction of a more efficient technique for the compensation of the microphone gain mismatch. In this way the real microphone array directivity pattern is closer to the theoretical one. Thus, a more precise direct-reflected signal separation can be achieved.

- Refinement of the actual method in order to make the results less sensitive to measurement noise, also in the case of highly reflecting materials.

Other refinements can be introduced, instead, in order to extend the actual method features. In particular:

- ability of the system to estimate the sound absorption coefficients even for non-regular surfaces, or for surfaces made up of different materials. This feature requires a more complex model, which takes into account also diffraction and refraction phenomena.

- Ability of the system to give the results even as a function of the angle of incidence, besides as a function of the frequency.

- Ability of the system to classify the materials of the studied surfaces into categories, according to the obtained results.

- Ability of the system to work in combination with the one capable of getting the geometry estimation for a generic environment. In this way we could be able to get the complete acoustical characterization for a generic environment. The gathered data could be used for a wide set of applications and, above all, for the ones related to self-calibration sound emitting systems.

Bibliography

[1] T. D. Rossing, editor. *Handbook of acoustics*. Springer, 2007.

[2] http://pcfarina.eng.unipr.it/.

[3] E. Mommertz. Angle-dependent in-situ measurements of reflection coefficients using a subtraction technique. *Applied Acoustics*, (46):251–265, 1995.

[4] Gary W. Elko. Differential microphone arrays. In *Audio signal processing for next-generation multimedia communication*. Kluwer Academic Publishers, 2004.

[5] W.C. Sabine. *Collected papers on acoustics*. Dover Publications, Inc, New York, 1964.

[6] Adrienne Research Team. Test methods for the acoustic performance of road traffic noise reducing devices. Technical Report MAT1-CT94049, European Commission, 1997.

[7] J. Ducorneau, V. Planeau, J. Chatillon, and A. Nejade. Measurement of sound absorption coefficients of flat surfaces in a workshop. *Applied Acoustics*, 70(3):710–721, 2009.

[8] J. Ducourneau, V. Planeau, and A. Nejade. Design of a multipolar weighting for acoustic antennae. *Applied Acoustics*, 70(3):484–492, 2009.

[9] ISO. Determination of sound absorption coefficient and impedance in impedance tubes: method using standing wave ratio. Technical Report ISO 10534-1, International Organization for Standardization, 1996.

[10] ISO. Determination of sound absorption coefficient and impedance in impedance tubes: transfer-function method. Technical Report ISO 10534-2, International Organization for Standardization, 1998.

[11] ISO. Measurement of sound absorption properties of road surfaces in situ: extended surface method. Technical Report ISO 13472-1, International Organization for Standardization, 2002.

[12] E. Mommertz and S. Müller. Measuring impulse responses with digitally pre-emphasized pseudorandom noise derived from maximum-length sequences. *Applied Acoustics*, 44(3):195–214, 1995.

[13] P. Cobo, A. Fernànedz, and M. Cuesta. Measuring short impulse responses with inverse filtered maximum-length sequences. *Applied Acoustics*, 68(7), 2006.

[14] T.I. Laakso, V. Valimaki, M. Karjalainen, and U.K. Laine. Splitting the unit delay [fir/all pass filters design]. *Signal Processing Magazine, IEEE*, 13(1):30–60, 1996.

[15] M. Garai and P. Guidorzi. European methodology for testing the airborne sound insulation characteristics of noise barriers in situ: experimental verification and comparison with laboratory data. *Journal of the Acoustical Society of America*, 108(3):1054–1067, 2000.

[16] J.C. Pequet, H. Krim, and H. Carfantan. Time-invariant orthonormal wavelet representations. *IEEE Transactions on signal processing*, 44(8):1964–1970, 1996.

[17] S. Cavaliere. Caratterizzazione di ambienti di ascolto: metodi e misure. `http://www.federica.unina.it/smfn/elaborazione-dei-segnali-per-la-multimedialita/caratterizzazione-ambienti-ascolto/`.

[18] M. Tagliasacchi. Lecture notes on foundamentals of audio-video signal processing, 2007.

[19] S. Mallat. *A wavelet tour of signal processing*. Elsevier, 3rd edition, 2009.

[20] S. K. Mitra. *Digital signal processing. A computer based approach.* Mc Graw Hill, 3rd edition, 2006.

[21] W. C. Elmore and M. A. Heald. *Physics of waves.* Dover Publications, Inc, 1985.

[22] D. R. Raichel. *The science and applications of acoustic.* Springer, 2nd edition, 2006.

[23] D. T. Blackstock. *Fundamentals of physical acoustics.* Wiley-Interscience, 2000.

Appendix A

Absorption coefficients for some common materials

	Octave-Band Center Frequency (Hz)					
	125	250	500	1000	2000	4000
Brick, unglazed	0.03	0.03	0.03	0.04	0.05	0.07
Brick, unglazed, painted	0.01	0.01	0.02	0.02	0.02	0.03
Carpet on foam rubber	0.08	0.24	0.57	0.69	0.71	0.73
Carpet on concrete	0.02	0.06	0.14	0.37	0.60	0.65
Concrete block, coarse	0.36	0.44	0.31	0.29	0.39	0.25
Concrete block, painted	0.10	0.05	0.06	0.07	0.09	0.08
Floors, concrete or terrazzo	0.01	0.01	0.015	0.02	0.02	0.02
Floors, hardwood	0.15	0.11	0.10	0.07	0.06	0.07
Glass, heavy plate	0.18	0.06	0.04	0.03	0.02	0.02
Glass, standard window	0.35	0.25	0.18	0.12	0.07	0.04
Gypsum, board 0.5 in.	0.29	0.10	0.04	0.05	0.07	0.09
Panels, fiberglass, 1.5 in thick	0.86	0.91	0.80	0.89	0.62	0.47
Panels, plywood 3/8 in.	0.28	0.22	0.17	0.09	0.10	0.11
Panels, plywood 1/8 in.	0.15	0.25	0.12	0.08	0.08	0.08
Wood, solid, 2 in. thick	0.01	0.05	0.05	0.04	0.04	0.04
Tile, marble or glazed	0.01	0.01	0.01	0.01	0.02	0.02
Wood, solid, 2 in. thick	0.01	0.05	0.05	0.04	0.04	0.04
Water surface	nil	nil	nil	0.003	0.007	0.02
Air	nil	nil	nil	0.003	0.007	0.03
One person	0.18	0.4	0.46	0.46	0.51	0.46

Table A.1: Absorption coefficients

Appendix B

Maximum eigenvalues and eigenvectors for differential arrays with maximum front to back ratio

Order	Max eigenvalue	Eigenvector
1^{st}	$7 + 4\sqrt{3}$	$\left[\frac{1}{1+\sqrt{3}} \quad \frac{1}{1+\sqrt{3}}\right]$
2^{nd}	$127 + 48\sqrt{7}$	$\left[\frac{1}{2(3+\sqrt{7}} \quad \frac{\sqrt{7}}{3+\sqrt{7}} \quad \frac{5}{2(3+\sqrt{7}}\right]$
3^{rd}	≈ 5875	$\approx [0.0184 \;\; 0.2004 \;\; 0.2870 \;\; 0.4750 \;\; 0.3061]$
4^{th}	≈ 151695	$\approx [0.0036 \;\; 0.0670 \;\; 0.2870 \;\; 0.4318 \;\; 0.2107]$

Table B.1: Maximum eigenvalues and eigenvectors corresponding to differential arrays with maximum front to back ratio, for spherically isotropic fields

Order	Max eigenvalue	Eigenvector
1^{st}	$7 + 4\sqrt{3}$	$\left[(\sqrt{2} - 1) \;\; (2 - \sqrt{2})\right]$
2^{nd}	$\frac{9\pi^2+12\sqrt{22}\pi+88}{9\pi^2-88}$	$\approx [0.103 \;\; 0.484 \;\; 0.413]$
3^{rd}	≈ 11556	$\approx [0.002 \;\; 0.217 \;\; 0.475 \;\; 0.286]$
4^{th}	≈ 336035	$\approx [0.0043 \;\; 0.0743 \;\; 0.2991 \;\; 0.4252 \;\; 0.1971]$

Table B.2: Maximum eigenvalues and eigenvectors corresponding to differential arrays with maximum front to back ratio, for cilindrically isotropic fields

Appendix C

Detailed sub-arrays configuration

Mic. no.	Sub-array$_1$ d=0.15m	Sub-array$_2$ d=0.10m	Sub-array$_3$ d=0.05m	Sub-array$_4$ d=0.025m	Relative positioning (m)
1	×				-0.30
2		×			-0.20
3	×				-0.15
4		×	×		-0.10
5			×	×	-0.05
6				×	-0.025
7	×	×	×	×	0
8				×	+0.025
9			×	×	+0.05
10		×	×		+0.10
11	×				+0.15
12		×			+0.20
13	×				+0.30

Table C.1: Detailed sub-arrays configuration and relative microphones positioning

Appendix D

Devices specifications

D.1 Lynx Aurora 16

Available I/O	Sixteen inputs and sixteen outputs
Type	Electronically balanced or unbalanced
Level	+4 dBu nominal / +20 dBu max. or -10 dBV nominal / +6 dBV max
Input Impedance	Balanced mode: 24kΩ Unbalanced mode: 12kΩ
Output Impedance	Balanced mode: 100Ω Unbalanced mode: 50Ω
Output Drive	600Ω impedance, 0.2 μF capacitance
A/D and D/A Type	24-bit multi-level, delta-sigma

Table D.1: Analog I/O

Frequency Response	20 Hz - 20 kHz, +0/-0.1 dB
Dynamic Range	117 dB, A-weighted
Channel Crosstalk	-120 dB maximum, 1kHz signal, -1 dBFS
THD + N	-108 dB (0.0004%) @ -1 DBFS -104 dB (0.0006%) @ -6 DBFS 1 kHz signal, 22 Hz - 22 kHz BW

Table D.2: Analog In Performance

Frequency Response	20 Hz - 20 kHz, +0/-0.1 dB
Dynamic Range	117 dB, A-weighted
Channel Crosstalk	-120 dB maximum, 1kHz signal, -1 dBFS
THD + N	-107 dB (0.00045%) @ -1 DBFS -106 dB (0.00050%) @ -6 DBFS 1 kHz signal, 22 Hz - 22 kHz BW

Table D.3: Analog Out Performance

Number / Type	16 inputs and 16 outputs 24 bit AES/EBU format, transformer coupled
Channels	16 in/out in single-wire mode 8 in/out in dual-wire mode
Sample Rates	All standard rates and variable rates up to 192 kHz in both single-wire and dual-wire modes

Table D.4: Digital I/O

Digital I/O Ports	25-pin female D-sub connectors Port A: channels 1-8 I/O Port B: channels 9-16 I/O Yamaha pinout standard
Analog I/O Ports	25-pin female D-sub connectors Analog In 1-8 Analog In 9-16 Analog Out 1-8 Analog Out 9 - 16 Tascam pinout standard

Table D.5: Connections

D.2 Focusrite Octopre LE

Gain	+13 dB to + 60 dB
Input Impedance	2.5 kΩ / 150Ω on LoZ (Ch1 + Ch2)
EIN	124 dB @ 60 dB Gain with 150 Ω termination and 22 Hz - 22 kHz filter
THD + N @ Min Gain (+13 dB)	0.0006% with 0 dBu input and 22 Hz - 22 kHz filter
THD + N @ Max Gain (+60 dB)	0.003% with -36 dBu input and 22 Hz - 22 kHz filter
THD + N @ Max Input (+9 dBu)	0.0008% with 22 Hz - 22 kHz filter
Frequency Response @ Min Gain (+13 dB)	-0.4 dB @ 10 Hz and -3 dB @ 122 kHz
Frequency Response @ Max Gain (+60 dB)	-2.3 dB @ 10 Hz and -3 dB @ 67 kHz
CMRR @ Max Gain (+60 dB)	80 dB

Table D.6: Mic Input Response

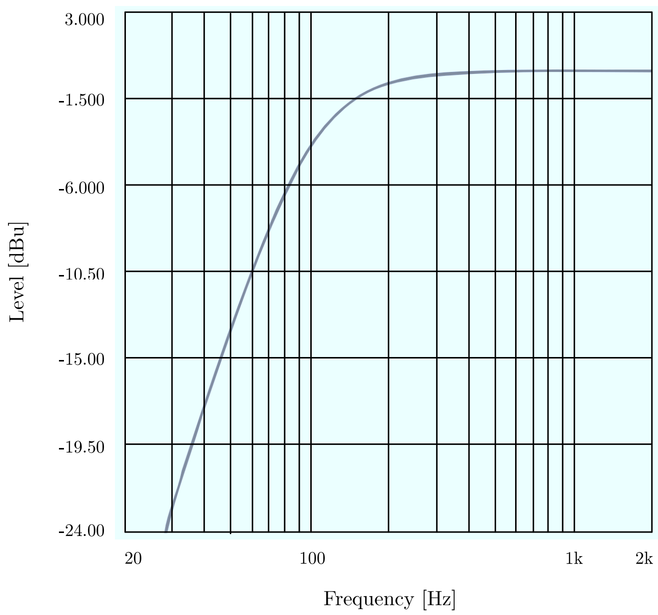


Figure D.1: Frequency response

D.3 Beyerdynamic MM1

Transducer Type	Condenser (back electret)
Operating Principle	Pressure
Frequency Response	20 Hz - 20 kHz (50 Hz - 16 kHz) ± 1.5 dB
Polar Pattern	Omnidirection, diffuse field calibrated
Open Circuit Voltage (1 kHz)	15 mV/Pa (-36.5 dBV) ± 1 dB
Nominal Impedance	330 Ω
Nominal Load Impedance	≥ 2.2 kΩ
Maximum SPL at f = 1 kHz, k = 1%, R_l = 2.2 kΩ	128 dB_{SPL}
S/N ratio, relative to 1 Pa	> 57 dB
A-weighted equivalent SPL	≈ 28 dB(A)
Power Supply	12 - 48 V phantom supply
Current Consumption	≈ 3.4 mA
Output	Transformer balanced
Connection	3-pin XLR male
Dimensions	Length: 133 mm Shaft diameter: 19/9 mm Head diameter: 9 mm
Weight	88 g

Table D.7: Analog Out Performance

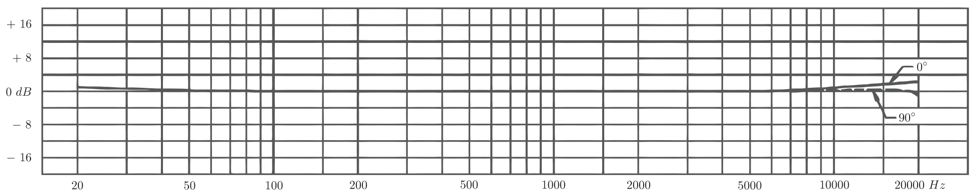


Figure D.2: Frequency response

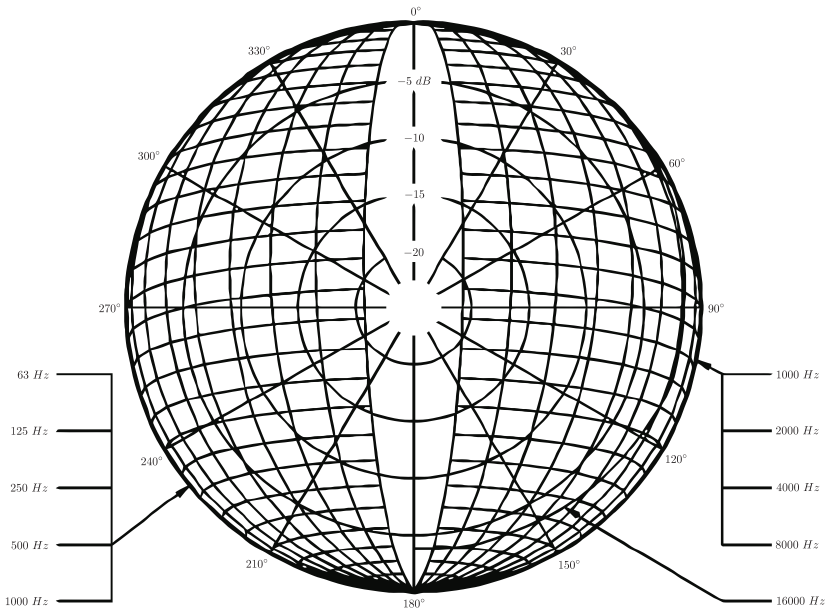


Figure D.3: Polar Diagram