

Synthèse de champ sonore par ambisonie adaptative d'ordres élevés (ANFC-HOA)

Stage de Master 2 ATIAM : Acoustique, Traitement du signal, Informatique Appliqués à la
Musique réalisé par

Pierre LECOMTE

Master 2 ATIAM, IRCAM-UPMC-TELECOM

du 01 Mars 2012 au 31 Juillet 2013 sous la tutelle de messieurs

Christophe LANGRENNE
Alexandre GARCIA

Ingénieur recherche, CNAM
Professeur, CNAM

TABLE DES MATIÈRES

1	Introduction	9
1.1	Resumé	9
1.2	Abstract	9
1.3	Remerciements	10
1.4	Contexte académique	10
1.5	Notes pour la lecture du rapport	10
2	Synthèse de champs sonores par méthodes ambisoniques	11
2.1	Introduction : principes généraux du formalisme ambisonique	11
2.1.1	Aperçu et enjeux de la méthode	11
2.1.2	Enregistrement ambisonique	12
2.1.3	Encodage	12
2.1.4	Décodage	13
2.1.5	Restitution	13
2.2	Formalisme ambisonique par harmoniques cylindriques	13
2.2.1	Décomposition d'une onde plane en harmoniques cylindriques	14
2.2.2	Formulation matricielle	15
2.2.3	Restitution par superposition d'ondes planes	15
2.2.4	Calcul des gains des haut-parleurs de restitution	16
2.2.5	Inversion de la matrice \mathbf{U}	16
2.2.6	Exemple en simulation	16
2.2.7	Inconvénients du formalisme en harmonique cylindrique dans le cadre du stage	17
2.3	Formalisme ambisoniques par harmoniques sphériques	17
2.3.1	Décomposition d'une onde plane sur une base d'harmoniques sphériques	17
2.3.2	Formulation matricielle	18
2.3.3	Décodage	19
2.4	Ambisonie de champ proche : prise en compte de la distance finie de la source et des haut-parleurs de restitution	19
2.4.1	Décomposition d'un onde sphérique sur une base d'harmoniques sphériques	19
2.4.2	Filtres de prise en compte du champ proche, filtres "Bass-Boost"	20
2.4.3	Filtres de compensation de champ proche, filtre "NFC : Near Field Compensated"	21
2.5	Restriction en 2 dimensions des harmoniques sphériques	21
2.6	Reconstruction par superposition d'ondes sphériques	23
2.7	Exemple en simulation	23
2.8	Caractérisation du champ acoustique reconstruit	23
2.8.1	Critère d'erreur	24
2.8.2	Taille de la zone d'écoute	24
2.8.3	Décodage "psychoacoustique"	27
2.8.4	Vecteur vitesse	27
2.8.5	Vecteur énergie	27
2.9	Conclusion	27
3	Correction des canaux de restitution au centre du dispositif	29
3.1	Introduction générale	29
3.2	Mesure du gain complexe et de la réponse impulsionnelle d'un canal à une position	30
3.2.1	Méthode d'estimation choisie	30
3.2.2	Exemple de mesure	30
3.3	Inversion de la réponse impulsionnelle d'un canal, calcul d'un filtre inverse.	31

Table des matières

3.3.1	Introduction	31
3.3.2	État de l'art	33
3.3.3	Techniques utilisées dans le cadre du stage	34
3.4	Quelques résultats	39
3.4.1	Filtres utilisés	39
3.4.2	Déconvolution sous MATLAB ©	41
3.4.3	Déconvolution en temps réel sur le système de restitution	41
3.5	Conclusions	43
4	Correction multicanale du système de restitution sur un ensemble de point de contrôle	45
4.1	Introduction	45
4.1.1	Rapide État de l'art	46
4.1.2	Contexte du stage	46
4.1.3	Algorithmes utilisés dans le cadre du stage	47
4.1.4	Inversion fréquentielle utilisant la régularisation	49
4.1.5	Inversion fréquentielle utilisant le formalisme issu de l'AWFS adapté à l'ambisonie : ANFC-HOA	49
4.1.6	Choix du paramètre de régularisation β	50
4.1.7	Introduction d'un délai	50
4.2	Un Exemple en simulation	50
4.2.1	Correction de l'effet de sol	50
4.2.2	Source synthétisée quelconque	52
4.3	Quelques résultats sur le dispositif de restitution	53
4.3.1	Conception des filtres	53
4.3.2	Résultats issus de la mesure	53
4.3.3	Notes sur le critère d'erreur	54
4.4	Conclusions	54
5	Réalisation pratique d'un système de synthèse de champ sonore 2D temps-réel par ANFC-HOA	61
5.1	Introduction	61
5.2	Choix de l'ordre de restitution	61
5.3	Disposition des enceintes de restitution	61
5.4	Transmission des signaux aux enceintes	62
5.5	Interface de commande du système de restitution	62
5.5.1	Programme Pure Data	62
5.6	Déconvolution de l'effet de salle en temps-réel à l'aide du logiciel BRUTEFIR	63
5.6.1	Présentation de BRUTEFIR	63
5.6.2	Fichier de configuration	64
5.7	Mesure de champs sonores	64
5.7.1	Configuration de l'antenne	64
5.7.2	Configuration de l'antenne globale	66
5.7.3	Acquisition des mesures	67
6	Conclusions	71
6.1	Conclusions des travaux menés	71
6.1.1	Remise en contexte	71
6.1.2	Inversion de la réponse de chaque canal de restitution	71
6.1.3	Inversion multicanale	72
6.1.4	Simulations et mesures	72
6.2	Futurs travaux	72
6.2.1	Formulation des filtres inverses	72
6.2.2	Ajout de sources pour aider la correction	72
6.2.3	Passage en salle réelle	73
6.2.4	Simulation d'effet de salles	73
6.2.5	Critère d'erreur	73

6.2.6	Caractérisation par décomposition en harmoniques sphériques	73
6.2.7	Caractérisation fines des sources de restitution	73
6.2.8	Mise en œuvre de décodages psychoacoustique	73
6.2.9	Caractérisation subjective des champs synthétisés	74
6.2.10	Comparaison avec d'autres techniques de synthèse de champs sonores	74
A	Estimation de la réponse impulsionnelle par filtrage de l'interspectre	75
B	Estimation de la réponse impulsionnelle par sinus glissant	77
C	Décomposition d'un champ acoustique en harmoniques cylindriques	79
C.1	Équation des ondes	79
C.2	Équation des ondes en coordonnées cylindriques	79
C.3	Harmoniques cylindriques	79
C.4	Solution de l'équation des ondes en harmoniques cylindriques	80
D	Décomposition d'un champ acoustique en harmoniques sphériques	81
D.1	Équation des ondes en coordonnées sphériques	81
D.2	Harmoniques sphériques	81
D.3	Solution de l'équation des ondes en séries de Fourier-Bessel	82

NOMENCLATURE

- A-NFCHOA Adaptive Near Field Compensated Higher Order Ambisonics : Ambisonie adaptative d'ordres élevés avec compensation de champ proche
- AWFS Adaptive Wave Field Synthesis : Synthèse de champ sonore adaptative par holophonie
- FFT Fast Fourier Transform : Transformée de Fourier rapide. Il s'agit d'un algorithme permettant de calculer rapidement la TFD.
- HOA Higher Order Ambisonics : Ambisoniques d'ordres élevés
- LMSSC Laboratoire de Mécanique des Structures, Matériaux et Système Couplés. Lieu dans lequel se déroule le stage
- LS Least Squares : Moindres Carrées, utilisé pour mentionner les filtres conçus à l'aide d'une minimisation aux sens des moindres carrés
- NFC Near Field Compensated : Compensation du champ proche
- NFC-HOA Near Field Compensated Higher Order Ambisonics : Ambisoniques d'ordres élevés avec compensation du champ proche
- RI Réponse Impulsionnelle
- RIF Réponse Impulsionnelle Finie
- RII Réponse Impulsionnelle Infinie
- TFD Transformée de Fourier Discrète
- WFS Wave Field Synthesis : Synthèse de champ sonore par holophonie

INTRODUCTION

1.1 RESUMÉ

Ce stage traite de la synthèse de champs sonores par une méthode d'ambisonie adaptative d'ordre élevé. Les méthodes ambisoniques permettent d'encoder des champs sonores complexes en trois dimensions et de les reproduire le plus fidèlement possible sur une zone d'écoute la plus large possible. À l'instar des techniques binaurales ou transaural, la zone d'écoute peut inclure plusieurs auditeurs.

Le formalisme mathématique repose sur la décomposition du champ sonore sur une base tronquée d'harmoniques sphériques. La restitution du champ sonore se fait à l'aide de plusieurs haut-parleurs disposés autour des auditeurs. L'approche présentée ici se focalise sur un système de synthèse adaptative en temps réel de champs sonores en deux dimensions. Le système de restitution est un cercle de haut-parleurs disposés autour de l'auditeur.

Le formalisme ambisonique classique requiert une condition de champ libre, ce qui n'est jamais le cas en situation réelle. On présente dans ce rapport une méthode qui se propose de corriger l'effet de salle de restitution : Les signaux ambisoniques sont préfiltrés par une matrice de filtre pour diminuer l'"effet de salle" sur un grand nombre de points de contrôle au niveau de la zone de synthèse. En extrapolant cette technique, on peut envisager de simuler n'importe quel effet de salle. Outre une application audio, ce système peut alors être envisagé pour les moyens d'essais acoustiques (chambre réverbérante, cabine alpha, etc.) et pourrait constituer une alternative séduisante pour de tels équipements dès lors que la reproduction du champ est suffisamment précise.

Le rapport s'articule de la manière suivante : après ce premier chapitre d'introduction, les principales méthodes ambisoniques sont présentées au second chapitre, du formalisme le plus classique (les haut-parleurs de restitution sont considérés comme générateurs d'ondes planes) à la NFC-HOA : Near Field Compensated Higher Order Ambisonics. Cette dernière technique permet de prendre en compte la distance finie des sources à synthétiser et du système de restitution. Les sources de restitution sont alors modélisées comme des monopôles. Les troisième et quatrième chapitres présentent les différentes techniques de correction d'effet de salle mises en oeuvre pendant le stage. Enfin, le dernier chapitre décrit l'implémentation temps réel d'un système de restitution 2D ANFC-HOA à l'ordre 7, les différents choix pour la réalisation et les moyens utilisés.

Les expériences sont réalisées au Laboratoire de Mécanique des Structures et Systèmes Couplés (LMSSC) du Conservatoire National des Arts et Métiers (CNAM) à Paris. Des cartographies du champ synthétisé sont réalisées à l'aide d'une antenne plane (une antenne sphérique est envisagée pour la suite).

Mots clés : acoustique, son 3D, ambisoniques d'ordres élevés, synthèse de champ sonore adaptative, déconvolution, conception de filtres, filtrage, traitement du signal.

1.2 ABSTRACT

This master thesis deals with adaptive sound fields synthesis using Higher Order Ambisonics (HOA) methods. These methods allow to encode complex sound fields in three dimensions and to reproduce them on a large area called the "sweet spot". In the contrary of others technologies such as binaural or transaural, the sweet spot can include several listener.

The mathematical theory behind ambisonics uses a decomposition of the sound field in a truncated spherical harmonics basis. At the decoding stage, the sound field is reconstructed using several speakers around the audience.

In this report, we focus on a reproduction system in two dimensions which is doing adaptive sound field synthesis in real time. The speakers are on a circle around the listeners.

HOA formalism requires free field condition, which never occurs in real situation. Thus in this report we present a method which try to remove the "room effet" in order to approche free field condition in the sweet spot. The ambisonics signals are pre-filtered by a matrix of filters to reduce the room effect on several control positions in the sweet spot. By extrapolation, we can imagine apply another room effet. At the end, this system could be used for audio application as well as a for acoustic testings (reverberant room, alpha cabin, etc.) if the sound field is physically well reconstructed.

The report is organized as follow : after this first chapter of introduction, the main ambisonics methods are investigated. From the more classical formulation, (the speakers are modeled as plane-waves generators) to the NFC-HOA : Near Field Compensated - Higher Order Ambisonics. In this last situation, the sources are modeled as monopoles, as well as the speakers that are in the near field.

The third and the fourth chapters deal with the room effect cancellation techniques we used for the system. Finally the last chapter presents the implementation of a 2D Adaptive NFC-HOA system at the 7th order, the choices made for the realization and means used.

Experiences are conducted in the laboratory "Laboratoire de Mécanique des Structures et Systèmes Couplés (LMSSC)" from the "Conservatoire National des Arts et Métiers (CNAM)" in Paris. Some sound field maps are made with a planar antenna (a spherical one is planned for later).

Keywords : acoustics, 3D sound, Higher Order Ambisonics, adaptive sound field synthesis, dereverberation, filter design, filtering, signal processing.

1.3 REMERCIEMENTS

Je tiens à remercier tout particulièrement mon maître de stage, Monsieur Christophe LANGRENNE, pour sa disponibilité et ses idées nombreuses, un vrai travail d'équipe, qui se poursuivra en thèse !

Merci également à toute l'équipe acoustique du LMSSC pour leur bonne humeur, leur précieux conseils et la bonne ambiance du bureau. Ainsi, merci à Manuel, Éric, Alexandre, Sarah, Marie, Stéphanie et Isabelle.

1.4 CONTEXTE ACADÉMIQUE

Ce stage Master 2 est le stage de fin d'étude de la formation Acoustique, Traitement du Signal, Informatique appliqués à la Musique (ATIAM). Ce master est issu d'un partenariat entre l'Université Pierre et Marie Curie - Paris VI (UPMC), l'Institut de Recherche et de Coordination Acoustique/Musique (IRCAM) et de Telecom-ParisTech. L'organisme d'accueil du stage est le LMSSC du CNAM. Le stage à lieu dans l'équipe Acoustique du laboratoire, sous la tutelle de Monsieur Christophe LANGRENNE, ingénieur de recherche.

1.5 NOTES POUR LA LECTURE DU RAPPORT

Le rapport présente de nombreuses figures annotées en anglais. Ceci est volontaire pour une éventuelle réutilisation ultérieure.

Les codes qui ont permis de faire les simulations, le dépouillement des mesures, les calculs sont mis à disposition du lecteur dans un répertoire disponible en ligne à cette adresse : <http://www.sekisushai.net/ANFCHOA/> Il s'agit principalement de fichiers *.wav, de données au format *.mat exploitable sous MATLAB®, de code MATLAB®*.m, de feuilles de calcul MATHEMATICA®*.nb et de patch PURE DATA *.pd

SYNTHÈSE DE CHAMPS SONORES PAR MÉTHODES AMBISONIQUES

2.1 INTRODUCTION : PRINCIPES GÉNÉRAUX DU FORMALISME AMBISONIQUE

Dans cette section, on présente de manière très générale les différentes étapes du formalisme ambisonique : ses enjeux, l'étape d'encodage, de décodage et de restitution. De manière volontaire, aucune équation n'apparaît, ceci étant fait dans les sections suivantes.

2.1.1 Aperçu et enjeux de la méthode

L'ambisonie a été introduite par Michael Gerzon dans les années 70. L'objectif premier est de pouvoir offrir une méthode d'encodage et de décodage fidèle et réaliste d'un champ sonore, en conservant au mieux les informations directionnelles de l'espace sonore (figure 2.1). La méthode tente de reconstruire physiquement le champ acoustique (du moins en basses fréquences) à l'aide de sources de restitution réparties autour de/des auditeur(s). Aux fréquences plus élevées, la reconstruction se base sur des considérations énergétiques et psychoacoustiques. En effet, la reconstruction physique devient coûteuse et peu utile en regard de la perception auditive humaine (voir 2.8). On cherche alors à donner une information perceptive de localisation. L'encodage ambisonique repose sur la décomposition du champ acoustique en harmoniques sphériques ou cylindriques. Ce formalisme mathématique solide et générique permet d'encoder le champ acoustique indépendamment du dispositif de restitution. Le décodage sera alors adapté suivant les situations de restitutions mais toutes partiront du même encodage. Cet avantage permet alors d'enregistrer une scène sonore à un endroit donné et de la restituer ailleurs suivant différentes configurations.

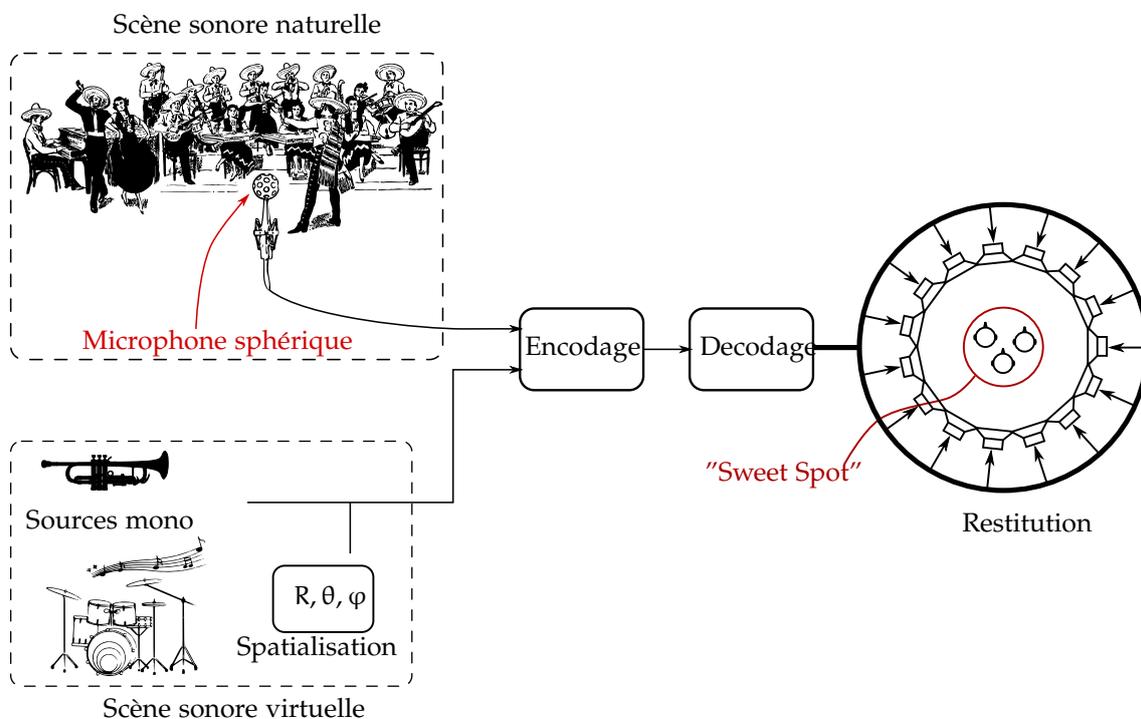


FIGURE 2.1.: Aperçu de la méthode ambisonique

L'ambisonie permet également de spatialiser des sources sonores à partir d'enregistrements plus classiques : mono, stéréo. En effet, il est possible de spécifier explicitement les informations

spatiales de localisation de la source et reconstruire le champ acoustique résultant au niveau de la zone d'écoute. Pour illustrer cela, on peut imaginer donner l'impression à l'auditeur qu'une trompette joue à un endroit précis de l'espace, en partant d'un enregistrement mono de la trompette (figure 2.1). Dans le cadre du stage c'est l'approche de spatialisation qui est étudiée.

2.1.2 Enregistrement ambisonique

Dans le cas où l'objectif est de restituer une scène sonore naturelle, une prise de son ambisonique est nécessaire. Une prise de son ambisonique consiste à obtenir en un point le maximum d'informations directionnelles du champ acoustique. Historiquement, le système *Soundfield*® conçu par Michael Gerzon et Peter Craven [Gerzon, 1975] permet de réaliser un enregistrement ambisonique à l'ordre 1. Il s'agit de quatre capsules de microphone formant un tétraèdre (figure 2.2(a)). En combinant les signaux des quatre capsules, on accède à l'information de pression p et aux trois composantes de son gradient v_x, v_y, v_z [Nicol, 1999, p. 73-79]. Ceci correspond à la troncature à l'ordre 1 de la décomposition du champ acoustique en harmoniques sphériques (voir 2.3).

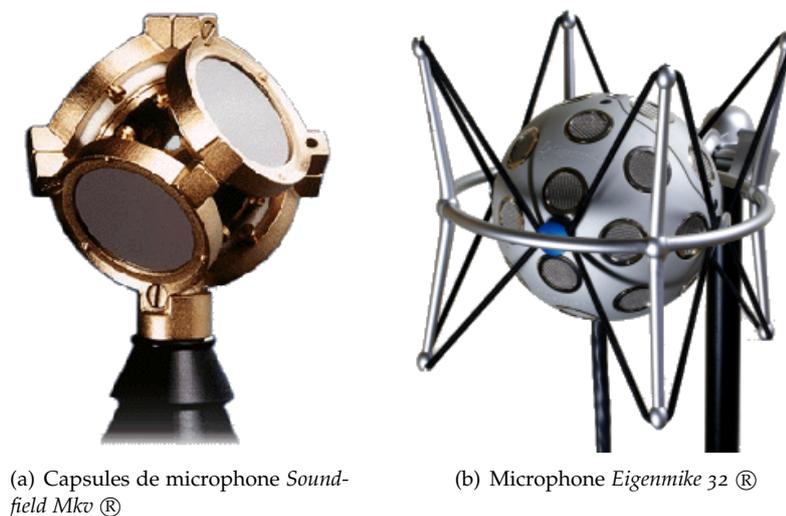


FIGURE 2.2.: Microphones pour prise de son ambisonique

Pour une prise de son permettant un encodage à des ordres supérieurs, il faut pouvoir obtenir le maximum d'informations de directions du champ acoustique au point d'enregistrement. Les signaux obtenus aux ordres supérieurs de la décomposition correspondent alors aux tenseurs d'ordre respectifs du champ [Daniel, 2000, p. 149]. Des microphones de forme sphériques sont utilisés pour l'enregistrement aux ordres élevés [Moreau, 2006, Chapitre 5]. Un exemple de ce type de microphone est visible sur la figure 2.2(b). Il s'agit du microphone *Eigenmike 32*® qui permet un enregistrement ambisonique jusqu'à l'ordre 4.

2.1.3 Encodage

Le champ acoustique original, après être enregistré, est décomposé sur une base d'harmoniques sphériques (voir 2.3). Plus sa connaissance au point d'enregistrement est précise, plus la décomposition peut être faite à un ordre élevé (voir 2.1.2), et plus la reconstruction sera valable sur une grande zone (voir 2.8.2).

On se demande comment la connaissance du champ en un seul point permet de reconstruire sur toute une zone le champ original. En fait, la connaissance des tenseurs d'ordre respectifs du champ en un point contraint les points aux alentours. Pour imager, le mouvement de la particule d'air situé au point d'enregistrement est corrélé à celui des particules environnantes.

Ainsi plus on caractérise finement ce mouvement, plus on contraint le mouvement des particules aux alentours.

Le formalisme de l'encodage permet également de spatialiser des sources : il est possible de définir volontairement les informations spatiales de la source à encoder et d'obtenir la décomposition correspondante à l'ordre souhaité.

La décomposition à l'ordre 1 a donné historiquement le format B [Gerzon, 1985]. Aux ordres supérieurs on parle d'ambisonie d'ordres élevés (*HOA : Higher Order Ambisonics*). Le format correspondant est le format B étendu [Daniel, 2000, Deuxième partie].

2.1.4 Décodage

Le décodage consiste à calculer les signaux qui alimentent les haut-parleurs servant à la restitution, à partir des données issues de l'encodage. La combinaison des différents signaux issus de ces haut-parleurs va reconstruire le champ sonore au point d'origine et aux alentours. À ce stade, la connaissance de la configuration géométrique du système de restitution est nécessaire.

Plusieurs stratégies de décodages sont envisageables suivant la gamme de fréquence. En effet, la présence des auditeurs peut perturber le champ reconstruit, surtout en hautes fréquences. Des considérations psychoacoustiques sont alors envisagées pour favoriser les directions perceptives [Daniel, 2000, p. 159].

Au moins autant de haut-parleurs que de canaux ambisoniques sont nécessaires pour le bon conditionnement du problème (voir 2.2.5), mais aucune restriction n'est faite sur leur répartition autour des utilisateurs. Cependant les configuration régulières ou semi-régulière facilite grandement le décodage [Daniel *et al.*, 1998].

2.1.5 Restitution

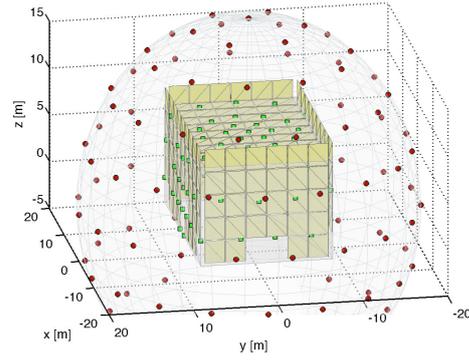
Les systèmes de restitutions peuvent être à deux ou à trois dimensions. Bien souvent, pour des besoins d'implémentation temps réel, la répartition des haut-parleurs est régulière : cercle ou sphère. (figure 2.3) Les systèmes à deux dimensions, comme celui élaboré pour ce stage (figure 2.3(c)), tentent à reconstruire le champ sonore dans le plan au niveau de l'oreille des auditeurs. On néglige alors toutes les contributions directionnelles en dehors de ce plan dans la matrice d'encodage.

A noter que la configuration du système de restitution peut également modifier la stratégie de décodage par des considérations énergétiques. Par exemple pour des configurations en demi-sphère, une correction énergétique peut être réalisée pour compenser la demi-sphère absente [Zotter *et al.*, 2010]. Les effets de "trous" entre les haut-parleurs peuvent également être compensés par filtrage spatial [Poletti, 2000]

Enfin, la salle de restitution et la proximité des haut-parleurs doit également être pris en compte. Effectivement, le formalisme du décodage ambisonique modélise les haut-parleurs de restitution comme des générateurs d'ondes planes en champ libre. Ce n'est évidemment pas le cas lorsque les haut-parleurs sont près des auditeurs et que la salle n'est pas anéchoïque. Ainsi des filtres de compensation de champs proches sont appliqués [Daniel, 2003] pour prendre en compte la proximité des haut-parleurs qui sont alors modélisés comme des monopôles. On parle alors d'ambisoniques avec compensation de champ proche (*NFC-HOA : Near Field Compensated Higher Order Ambisonics*). Une déconvolution de l'effet de salle peut être envisagée pour retrouver, au mieux, une condition de champ libre (par exemple, [Kirkeby *et al.*, 1998]). La mise en oeuvre de ces compensations est le principal objectif du stage.

2.2 FORMALISME AMBISONIQUE PAR HARMONIQUES CYLINDRIQUES

Dans cette section on présente le formalisme ambisonique de manière plus détaillée dans le cas de sources cylindriques. Le champ acoustique est supposé invariant dans la direction verticale.



(a) l'IEM-CUBE : Système de restitution HOA (b) Système de restitution HOA de l'IRCAM (Image 24 voie demi-sphérique (Image issue de l'article issue de l'article [Noisternig *et al.*, 2013]) [Zotter *et al.*, 2010])



(c) Système de restitution HOA-2D jusqu'à l'ordre 7 (Expérience montée lors du stage)

FIGURE 2.3.: Différents systèmes de restitution HOA

2.2.1 Décomposition d'une onde plane en harmoniques cylindriques

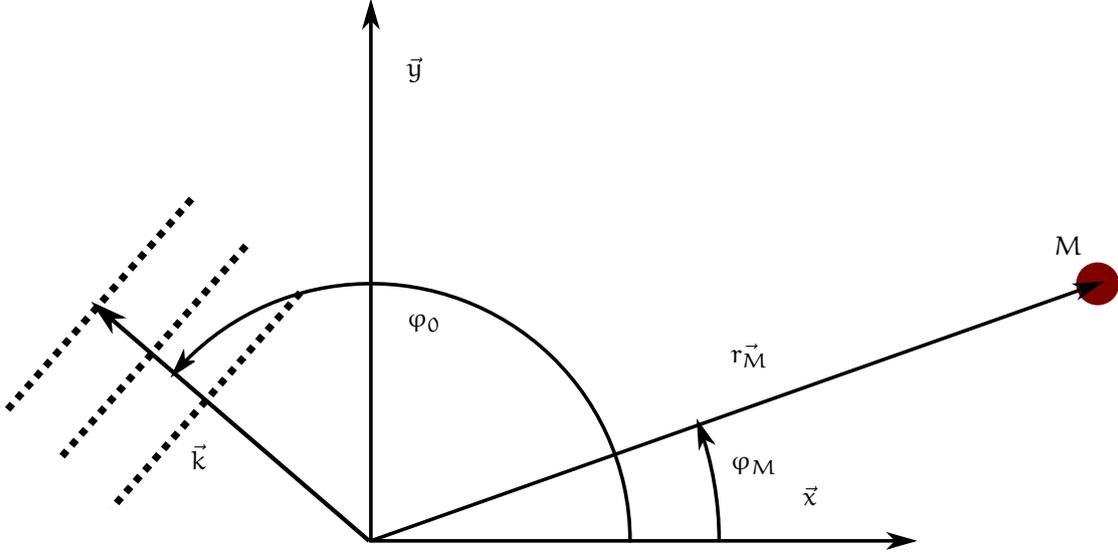
La première étape est de considérer le champ acoustique comme la superposition d'ondes planes et de décomposer chacune de ces ondes planes sur une base d'harmoniques cylindriques [Nicol, 1999, p. 86].

Soit une onde plane d'amplitude a se propageant horizontalement suivant la direction \vec{k} en coordonnées cylindrique (figure 2.4). Son vecteur d'onde en coordonnées cartésienne est :

$$\vec{k} = k \begin{vmatrix} \cos(\varphi_0) \\ \sin(\varphi_0) \\ 0 \end{vmatrix} \quad (2.1)$$

La pression générée par l'onde plane au point M de coordonnées cylindriques $r_M = (r_M, \varphi_M, z)$ s'exprime sur une base d'harmoniques cylindriques de la manière suivante [Morse et Ingard, 1968, p. 357-364], [Daniel, 2000, p. 302] :

$$\begin{aligned} p(r_M, \varphi_M) &= a e^{j\vec{k} \cdot r_M} \\ &= a J_0(kr_M) + 2a \sum_{m=1}^{+\infty} j^m J_m(kr_M) [\cos(m\varphi) \cos(m\varphi_0) + \sin(m\varphi) \sin(m\varphi_0)] \\ &= a \sum_{m \geq 0, \sigma = \pm 1}^{+\infty} \tilde{Y}_m^\sigma(\varphi_0) j^m \tilde{Y}_m^\sigma(\varphi_M) J_m(kr_M) \end{aligned} \quad (2.2)$$

FIGURE 2.4.: Une onde plane en coordonnées cylindriques. L'altitude z est quelconque

J_m représentent les fonctions de Bessel de première espèce. Cette décomposition est aussi connue sous le nom de développement en série de Fourier-Bessel. \tilde{Y}_m^σ sont les *harmoniques cylindriques* (ici normées, voir l'annexe C).

2.2.2 Formulation matricielle

En repartant de l'équation (2.2) on exprime la pression $p(r_M, \varphi_M)$ par un produit scalaire entre les deux vecteurs \vec{u} (ou \vec{B}_{cyl}) et \vec{v} définis par [Nicol, 1999, p. 87] :

$$\vec{u}^T(\varphi_0) = \vec{B}_{\text{cyl}}^T = [\tilde{Y}_0^1(\varphi_0) \quad \tilde{Y}_1^1(\varphi_0) \quad \tilde{Y}_1^{-1}(\varphi_0) \quad \cdots \quad \tilde{Y}_m^1(\varphi_0) \quad \tilde{Y}_m^{-1}(\varphi_0) \quad \cdots] \quad (2.3)$$

$$\vec{v}^T = \begin{bmatrix} J_0(kr_M) & j\tilde{Y}_1^1(\varphi_M)J_1(kr_M) & j\tilde{Y}_1^{-1}(\varphi_M)J_1(kr_M) & \cdots \\ j^m\tilde{Y}_m^1(\varphi_M)J_m(kr_M) & j^m\tilde{Y}_m^{-1}(\varphi_M)J_m(kr_M) & \cdots \end{bmatrix} \quad (2.4)$$

De cette manière, l'équation (2.2) s'exprime alors sous la forme :

$$p(r_M, \varphi_M) = \alpha \vec{u}^T \vec{v} \quad (2.5)$$

Le vecteur \vec{B}_{cyl} contient les informations directionnelles de l'onde plane et contient les coefficients de la décomposition en harmoniques cylindriques. Si l'on tronque ce vecteur à l'ordre 1, on retrouve le format B original sans la composante Z [Gerzon, 1985] [Nicol, 1999, p. 88]

2.2.3 Restitution par superposition d'ondes planes

Considérons à présent N haut-parleurs de restitution répartis autour de l'origine. On suppose que les haut-parleurs sont des générateurs d'ondes planes. Cette hypothèse peut s'approximer lorsque les haut-parleurs sont suffisamment loin de la zone de restitution. L'approximation de champ lointain est valable lorsque $kr_{\text{HP}} \gg 1$, où k est le nombre d'onde et r_{HP} la distance entre le point M et le haut-parleur [Chaigne *et al.*, 2008, p. 333].

Ces "générateurs d'ondes planes" sont définis spatialement par leur angles φ_i ($i = 1, \dots, N$). L'objectif est de reconstruire l'onde plane originale par superposition d'ondes planes provenant de chacun des haut-parleurs de restitution. La pression \tilde{p} "reconstruite" au point M résulte de la superposition des N ondes planes :

$$\tilde{p}(r_M, \varphi_M) = \sum_{i=1}^N g_i e^{jk r_M \cos(\varphi_M - \varphi_i)} \quad (2.6)$$

g_i représente le gain appliqué au haut-parleur i . Cette équation peut s'exprimer sous forme matriciel de la même manière qu'en 2.2.2 [Nicol, 1999, p. 88]

$$\begin{aligned}\tilde{p}(r_M, \varphi_M) &= \sum_{i=1}^N g_i \tilde{u}_i^T(\varphi_i) \vec{v} \\ &= \vec{g}^T \mathbf{U} \vec{v}\end{aligned}\quad (2.7)$$

avec

$$\vec{g}^T = [g_1 \quad g_2 \quad \cdots \quad g_N] \quad (2.8)$$

$$\mathbf{U}^T = [\tilde{u}_1(\varphi_1) \quad \tilde{u}_2(\varphi_2) \quad \cdots \quad \tilde{u}_N(\varphi_N)] \quad (2.9)$$

2.2.4 Calcul des gains des haut-parleurs de restitution

Pour reconstruire la pression acoustique $p(r_M, \varphi_M)$ à partir des N haut-parleurs, on cherche donc les gains \vec{g} qui donnent $\tilde{p}(r_M, \varphi_M) = p(r_M, \varphi_M)$

On a alors

$$a \tilde{u}^T \vec{v} = \vec{g}^T \mathbf{U} \vec{v} \quad (2.10)$$

Les gains \vec{g} sont alors données par :

$$\vec{g}^T = a \mathbf{U}^{-1} \tilde{u}^T \quad (2.11)$$

2.2.5 Inversion de la matrice \mathbf{U}

La difficulté réside dans l'inversion de la matrice \mathbf{U} : lorsque l'on tronque à l'ordre M la décomposition en harmoniques cylindriques, on obtient $2M + 1$ composantes dans la base tronquée. La matrice \mathbf{U} a alors pour dimensions $N \times 2M + 1$. Ainsi lorsque $N \neq 2M + 1$ l'inversion stricte n'existe pas et on a recours à une pseudo-inversion [Daniel *et al.*, 1998]. En pratique il est recommandé de choisir au moins autant de haut-parleurs que de canaux ambisoniques pour le bon conditionnement du problème (ici $N = 2M + 1$) [Nicol, 1999, p. 88]

Lorsque les N haut-parleurs sont équirépartis autour du point M , formant un polygone régulier, l'inversion prend une forme simple [Daniel *et al.*, 1998] :

$$\mathbf{U}^{-1} = \frac{1}{N} \mathbf{U}^T \quad (2.12)$$

et les gains des haut-parleurs de restitutions deviennent :

$$\vec{g} = \frac{a}{N} \mathbf{U} \tilde{u} \quad (2.13)$$

2.2.6 Exemple en simulation

Un résultat de simulation est présenté ci-dessous (figure 2.5). La source à reconstruire est une onde plane monochromatique de fréquence $f = 2000\text{Hz}$ et d'inclinaison 20° par rapport à l'axe \vec{x} . Les sources de restitution sont modélisées comme des générateurs d'ondes plane équirépartis, autour de l'origine. L'ordre de décomposition vaut ici $M = 13$. On choisit $N = 2M + 1 = 27$ sources de restitution. Les gains des sources sont calculés suivant la formule (2.13). Le trait gras noir représente la "zone d'écoute" dans laquelle l'erreur de reconstruction est inférieure à 20% (voir 2.8.1).

La figure 2.5 est réalisée à partir de la feuille de calcul MATHEMATICA@HOA_Cylindriques.nb disponible en ligne ¹. Ce code permet de synthétiser une onde plane monochromatique par harmoniques cylindriques à n'importe quel ordre de décomposition. Une version MATLAB@est également fournie ².

1. http://www.sekisushai.net/ANFCHOA/CHAP02/HOA_Cylindriques.nb

2. http://www.sekisushai.net/ANFCHOA/CHAP02/HOA_Cylindriques.m

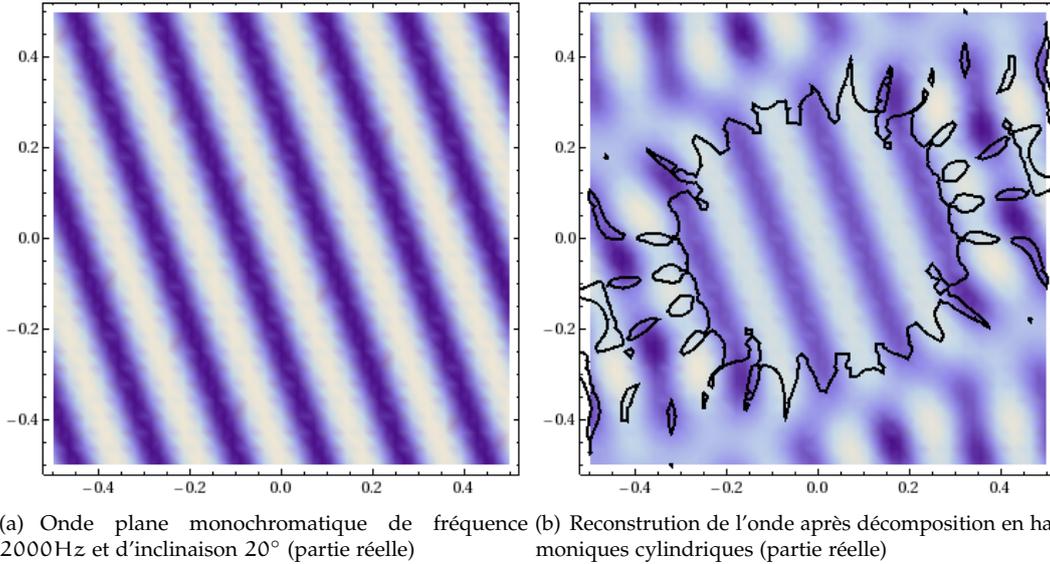


FIGURE 2.5.: Reconstruction d'une onde plane (a) après décomposition en harmoniques cylindriques à l'ordre $M = 13$ avec $N = 27$ générateurs d'ondes planes équirépartis autour de l'origine (b). La zone entourée du trait gras noir assure une reconstruction avec un critère d'erreur $\leq 20\%$ (voir 2.8.1). Les échelles sont en mètres.

2.2.7 Inconvénients du formalisme en harmonique cylindrique dans le cadre du stage

Le formalisme en harmoniques cylindriques présente quelques désavantages qui rendent son utilisation inappropriée à des fins de restitution ou de création réaliste d'environnement sonores. En effet, le formalisme est optimisé pour des sources sonores de type cylindrique, c'est à dire homogène suivant une direction. Ainsi il ne permet pas d'encoder des sources qui génèrent des ondes sphériques, ce qui représente la grande majorité des sources sonores que l'on souhaite encoder. La restitution se fait en deux dimensions ce qui ne permet que de générer un champ sonore sur un plan au niveau des oreilles des auditeurs.

Par ailleurs, les sources de restitution sont ici des générateurs d'ondes planes. Dans la réalité ce type de source ne s'approche qu'en champ lointain. On exclue alors les salles de restitution de taille petite et moyenne pour une utilisation audio.

Dans la section suivante est présenté le formalisme ambisonique par décomposition harmoniques sphériques. Cette approche permet de pallier en partie aux inconvénients mentionnés ci-dessus.

2.3 FORMALISME AMBISONIQUES PAR HARMONIQUES SPHÉRIQUES

Comme vu précédemment le formalisme en harmonique cylindrique n'est pas le plus approprié pour restituer un champ sonore complexe en 3 dimensions. A travers cette section, on décrit le formalisme ambisonique avec une approche par harmoniques sphériques. Pour comprendre l'origine du formalisme il est conseillé au lecteur de lire l'annexe D

2.3.1 Décomposition d'une onde plane sur une base d'harmoniques sphériques

De la même manière qu'en 2.2.1, on considère une onde plane d'amplitude a définie par le vecteur d'ondes $\vec{k} = (k\theta_0, k\varphi_0)$ en coordonnées sphériques (figure 2.6). La pression au point M de coordonnées sphériques $\vec{r}_M = (r_M, \theta_M, \varphi_M)$ est donnée par :

$$p(r_M, \theta_M, \varphi_M) = ae^{j\vec{k} \cdot \vec{r}_M} \quad (2.14)$$

ou encore [Morse et Ingard, 1968, p. 419] [Daniel, 2000, p.305] :

$$p(r_M, \theta_M, \varphi_M) = \sum_{m,n,\sigma} a \tilde{Y}_{mn}^\sigma(\theta_0, \varphi_0) \tilde{Y}_{mn}^\sigma(\theta_M, \varphi_M) j_m^m j_m(kr) \quad (2.15)$$

j_m représente les fonctions de Bessel sphériques de la première espèce d'ordre m . Cette décomposition est aussi connue sous le noms de développement en série de Fourier-Bessel. Les fonctions \tilde{Y}_{mn}^σ sont les *harmoniques sphériques* (ici normées, voir l'annexe D).

2.3.2 Formulation matricielle

De la même manière que pour les harmoniques cylindriques, on peut réécrire l'équation (2.15) comme le produit scalaire entre les vecteur \vec{B}_{sph} et \vec{v} définis par :

$$\vec{B}_{\text{sph}} = \begin{bmatrix} \tilde{Y}_{00}^1(\theta_0, \varphi_0) \\ \tilde{Y}_{11}^1(\theta_0, \varphi_0) \\ \tilde{Y}_{11}^{-1}(\theta_0, \varphi_0) \\ \dots \\ \tilde{Y}_{mm}^1(\theta_0, \varphi_0) \\ \tilde{Y}_{mm}^{-1}(\theta_0, \varphi_0) \\ \dots \\ \tilde{Y}_{mn}^1(\theta_0, \varphi_0) \\ \tilde{Y}_{m0}^{-1}(\theta_0, \varphi_0) \\ \dots \end{bmatrix} \quad \vec{v} = \begin{bmatrix} \tilde{Y}_{00}^1(\theta_M, \varphi_M) j_0(kr_M) \\ \tilde{Y}_{11}^1(\theta_M, \varphi_M) j_1^1(kr_M) \\ \tilde{Y}_{11}^{-1}(\theta_M, \varphi_M) j_1^{-1}(kr_M) \\ \dots \\ \tilde{Y}_{mm}^1(\theta_M, \varphi_M) j_m^m(kr_M) \\ \tilde{Y}_{mm}^{-1}(\theta_M, \varphi_M) j_m^{-m}(kr_M) \\ \dots \\ \tilde{Y}_{mn}^1(\theta_M, \varphi_M) j_m^m(kr_M) \\ \tilde{Y}_{m0}^{-1}(\theta_M, \varphi_M) j_m^{-m}(kr_M) \\ \dots \end{bmatrix} \quad (2.16)$$

(2.17)

De telle sorte que

$$p(r_M, \theta_M, \varphi_M) = a \vec{B}_{\text{sph}}^T \vec{v} \quad (2.18)$$

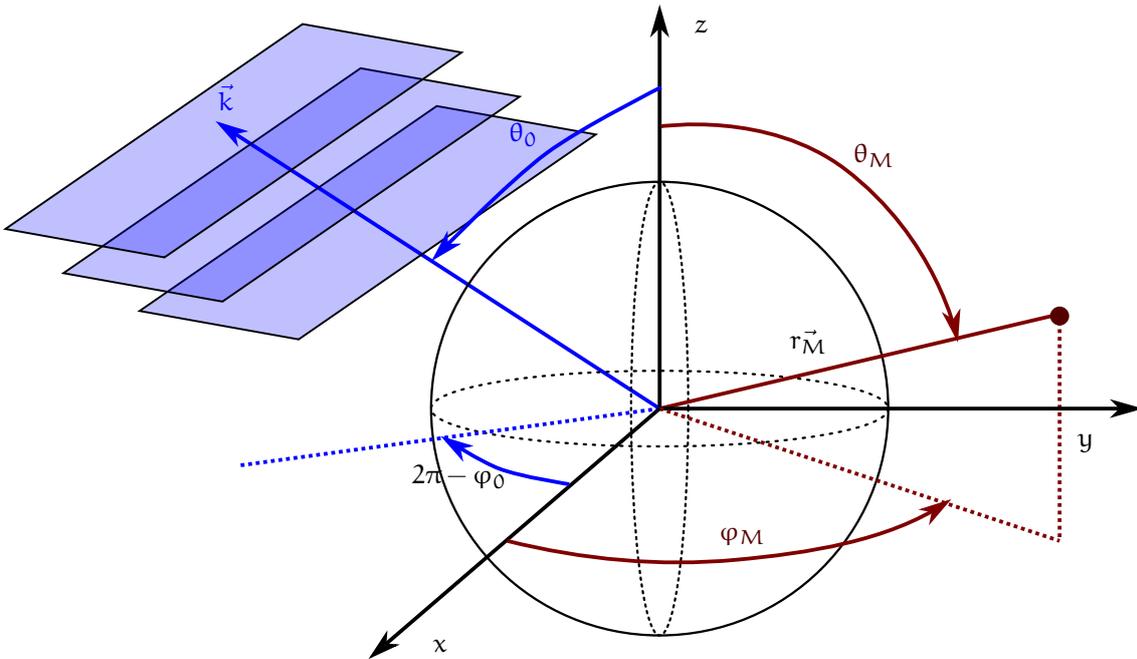


FIGURE 2.6.: Une onde plane en coordonnées sphériques

2.3.3 Décodage

De la même manière qu'en 2.2.3 et avec les mêmes notations, on peut considérer les sources de restitution comme générateurs d'ondes planes. Ainsi le problème de décodage se formule de la manière suivante.

$$\mathbf{a}\vec{\mathbf{B}}_{\text{sph}}^T \vec{\mathbf{v}} = \vec{\mathbf{g}}^T \mathbf{U} \vec{\mathbf{v}} \quad (2.19)$$

En tronquant la décomposition à l'ordre M , on a $2M + 1$ composantes par ordre (voir figure D.1, annexe D pour une représentation jusqu'à l'ordre $M = 3$). Ainsi à l'ordre M on obtient $K = (M + 1)^2$ canaux ambisoniques. En utilisant N sources de restitution. La matrice \mathbf{U} est alors de taille $N \times (M + 1)^2$ et définie par :

$$\mathbf{U}^T = [\vec{\mathbf{u}}_1(\theta_1, \varphi_1) \quad \vec{\mathbf{u}}_2(\theta_2, \varphi_2) \quad \cdots \quad \vec{\mathbf{u}}_N(\theta_N, \varphi_N)] \quad (2.20)$$

avec

$$\vec{\mathbf{u}}_N(\theta_N, \varphi_N) = \begin{bmatrix} \tilde{\mathbf{Y}}_{00}^1(\theta_N, \varphi_N) \\ \tilde{\mathbf{Y}}_{11}^1(\theta_N, \varphi_N) \\ \tilde{\mathbf{Y}}_{11}^{-1}(\theta_N, \varphi_N) \\ \cdots \\ \tilde{\mathbf{Y}}_{mm}^1(\theta_N, \varphi_N) \\ \tilde{\mathbf{Y}}_{mm}^{-1}(\theta_N, \varphi_N) \\ \cdots \\ \tilde{\mathbf{Y}}_{mn}^1(\theta_N, \varphi_N) \\ \tilde{\mathbf{Y}}_{m0}^{-1}(\theta_N, \varphi_N) \\ \cdots \end{bmatrix} \quad (2.21)$$

Les gains peuvent se trouver par inversion ou pseudo-inversion de la matrice \mathbf{U} , suivant le nombre N de sources de restitution et leur configuration géométrique :

$$\vec{\mathbf{g}} = \mathbf{a}\mathbf{U}^{-1}\vec{\mathbf{B}}_{\text{sph}} \quad (2.22)$$

Pour plus de détail, se référer à [Daniel, 2000, Deuxième partie] par exemple.

2.4 AMBISONIE DE CHAMP PROCHE : PRISE EN COMPTE DE LA DISTANCE FINIE DE LA SOURCE ET DES HAUT-PARLEURS DE RESTITUTION

Le formalisme ambisonique présenté aux sections précédentes suppose que les sources de restitution sont des générateurs d'ondes planes. Cette supposition peut s'avérer raisonnable dès lors que les haut-parleurs de restitution sont suffisamment éloignés de la zone d'écoute (approximation de champ lointain). Cependant c'est en pratique rarement le cas. Il faut alors prendre en compte l'effet de proximité des haut-parleurs, c'est à dire la courbure de leur fronts d'ondes. Les sources de restitution sont alors modélisés comme des monopôles à une distance proche du centre de la zone d'écoute. Ce formalisme ambisonique a été proposé par daniel [Daniel, 2003] et se trouve dans la littérature sous le nom *NFC-HOA : Near Field Compensated Higher Order Ambisonics*. Pour bien comprendre le formalisme, commençons par encoder une source de type monopôle sur une base d'harmoniques sphériques.

2.4.1 Décomposition d'une onde sphérique sur une base d'harmoniques sphériques

On considère une source monopolaire. On connaît la pression à une distance d proche du point O avec $\vec{\mathbf{r}}_0 = (r_0, \theta_0, \varphi_0)$ (figure 2.7). Avec la convention $e^{-j\omega t}$ La pression au point O est donnée à partir de la pression mesurée au point M , $\vec{\mathbf{r}}_M = (r_M, \theta_M, \varphi_M)$, ($r_M \leq r_0$) par [Morse et Ingard, 1968, p. 352]

$$p(r_M, \theta_M, \varphi_M) = p(r_0, \theta_0, \varphi_0) \frac{jkd}{4\pi} \sum_{m,n,\sigma} \tilde{\mathbf{Y}}_{mn}^\sigma(\theta_0, \varphi_0) \tilde{\mathbf{Y}}_{mn}^\sigma(\theta_M, \varphi_M) j_m(kr_M) h_m^-(kr_0) \quad (2.23)$$

En identifiant cette équation avec l'équation (D.7), annexe D, on obtient les coefficients de la décomposition [Daniel, 2003] :

$$\tilde{A}_{mn}^\sigma = p(r_0, \theta_0, \varphi_0) \tilde{Y}_{mn}^\sigma(\theta_0, \varphi_0) h_m^-(kr_0) \frac{kd}{4\pi} j^{-(m+1)} \quad (2.24)$$

où h_m^- est la fonction de Hankel de seconde espèce. On pose la fonction $\Gamma_m(kr) = \frac{kd}{4\pi} j^{-(m+1)} h_m^-(kr)$. La pression au point M est issue de celle la source après propagation d'une distance finie. Ceci est reflété par la fonction $\Gamma_0(k|r_0 - r_M|)$ de telle sorte que $p(r_M, \theta_M, \varphi_M) = p(r_0, \theta_0, \varphi_0) \Gamma(k|r_0 - r_M|)$. En plaçant le point d'encodage M à l'origine du repère, on obtient alors pour les coefficients ambisoniques :

$$\tilde{A}_{mn}^\sigma = p(r_M, \theta_M, \varphi_M) F_m(kr_0, \omega) \tilde{Y}_{mn}^\sigma(\theta_0, \varphi_0) \quad (2.25)$$

avec [Daniel, 2003]

$$F_m(kr_0, \omega) = \frac{\Gamma_m(kr_0, \omega)}{\Gamma_0(kr_0, \omega)} = \frac{j^{-m} h_m^-(kr_0)}{h_0^-} \quad (2.26)$$

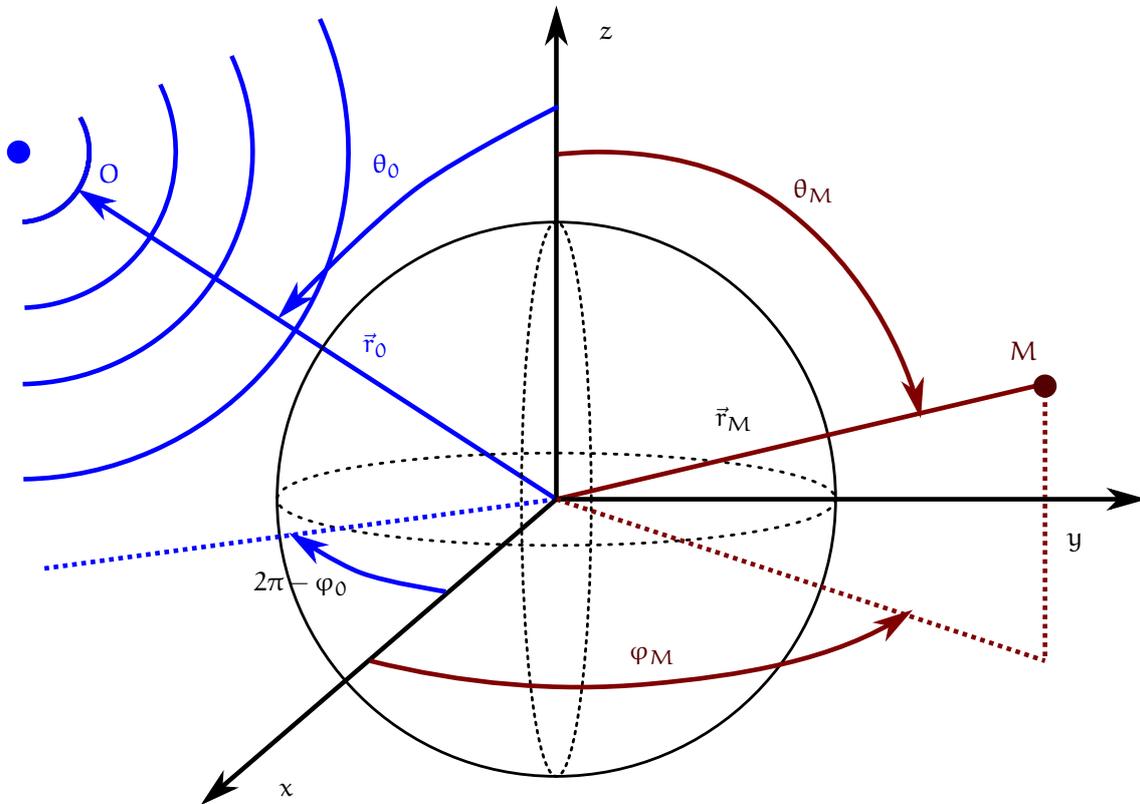


FIGURE 2.7.: Un monopôle en coordonnées sphériques

2.4.2 Filtres de prise en compte du champ proche, filtres "Bass-Boost"

La fonction de transfert $F_m(kr_0)$ renferme l'information de distance de la source à encoder. Elle affecte l'encodage ambisonique et agit comme un filtre qui amplifie les basses fréquences. On parle de filtre à effet "bass-boost" [Daniel, 2000, p. 153]. L'amplification est d'autant plus importante que la source à encoder est proche du point d'encodage et que l'ordre est grand. Sur la figure 2.8³ on peut observer l'allure fréquentiel de tels filtres. Ils sont de type intégrateur et instable par nature [Daniel, 2003]. On verra par la suite que l'utilisation de sources de restitution également monopôlaires permet de rendre le filtrage total stable.

³. Cette figure est réalisée à partir de la feuille de calcul MATHEMATICA@NFC_H0A.nb disponible en ligne : http://www.sekisushai.net/ANFCHOA/CHAP02/NFC_H0A.nb

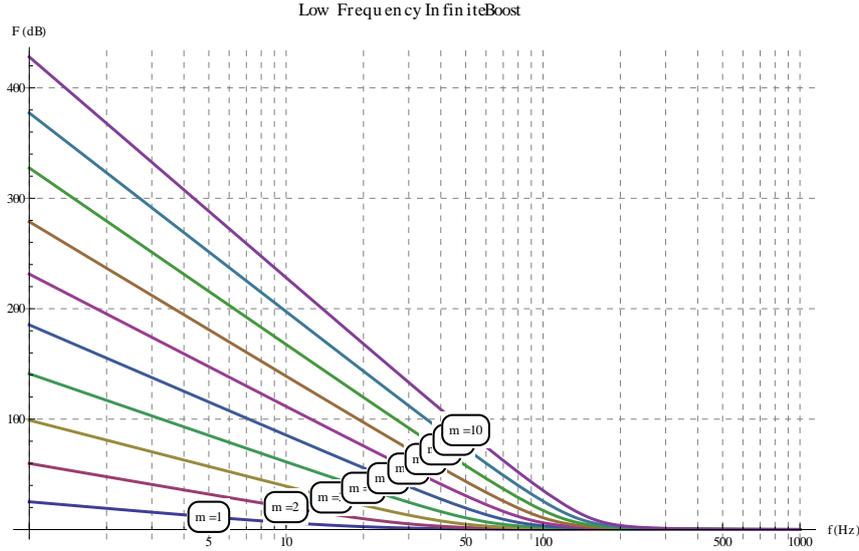


FIGURE 2.8.: Filtres "bass-boost" traduisant l'effet de champ proche. Les différentes courbes représentent les filtres à différents ordres pour une source à $r_0 = 3m$.

2.4.3 Filtres de compensation de champ proche, filtre "NFC : Near Field Compensated"

En utilisant des sources de restitution à proximité de la zone de restitution, il est nécessaire de prendre en compte l'effet de champ proche et de le corriger. Pour ce faire, on multiplie chacune des composantes du vecteur d'encodage \vec{B}_{sph} par l'inverse des filtres précédemment vus 2.4.2. Pour un système de restitution de plusieurs haut-parleurs à une distance r_{HP} de la source, les coefficients ambisoniques du vecteur \vec{B}_{sph} deviennent [Daniel, 2003] :

$$\tilde{A}_{mn}^{\sigma} = p(r_M, \theta_M, \varphi_M) H_m(kr_0, r_{\text{HP}}, \omega) \check{Y}_{mn}^{\sigma}(\theta_0, \varphi_0) \quad (2.27)$$

Avec

$$H_m(kr_0, r_{\text{HP}}, \omega) = \frac{F_m(kr_0, \omega)}{F_m(kr_{\text{HP}}, \omega)} \quad (2.28)$$

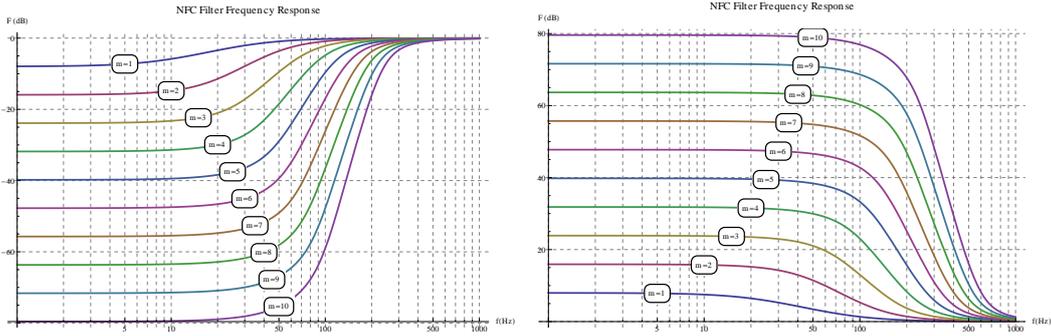
Ces filtres sont appelés filtres de compensation de champ proche ou "NFC : Near Field Compensated". Ils produisent une amplification positives pour des sources à l'intérieur du dispositif de restitution ($r_0 \leq r_{\text{HP}}$) et négatives pour des sources à l'extérieur ($r_0 \geq r_{\text{HP}}$). Sur la figure 2.9⁴ on peut observer les courbes de gains à différents ordres pour des sources de restitution à $r_{\text{HP}} = 1.25m$ avec une source à reproduire à $r_0 = 5m \geq r_{\text{HP}}$ 2.9(a) et $r_0 = 0.8m \leq r_{\text{HP}}$ 2.9(b)

Ainsi, la prise en compte de la distance finie des sources à spatialiser et des sources de restitution revient à filtrer les composantes ambisoniques classiques avec les filtres $H_m(kr_0, r_{\text{HP}}, \omega)$. Une réalisation stable de ces filtres est proposée par Daniel [Daniel, 2003] et Adriaensen [Adriaensen, 2006]. Dans le cadre du stage, seule l'implémentation de [Daniel, 2003] a été réalisée sous PURE DATA (voir 5.5.1).

2.5 RESTRICTION EN 2 DIMENSIONS DES HARMONIQUES SPHÉRIQUES

La réalisation d'un système de restitution en trois dimensions peut être parfois problématique. Outre les problèmes d'équirépartition des sources de restitution sur une sphère [Moreau, 2006, p. 110], il est parfois difficile de réaliser une sphère entière où l'auditeur prend place à proximité du centre (voir figures (2.3(b), 2.3(a))). Des adaptations au système de restitution à partir de considérations énergétiques et/ou psychoacoustiques sont alors réalisées au décodage pour compenser l'absence de certaines sources de restitution [Zotter *et al.*, 2010].

4. Cette figure est réalisée à partir de la feuille de calcul MATHEMATICA@NFC_HOA.nb disponible en ligne : http://www.sekisushai.net/ANFCHOA/CHAP02/NFC_HOA.nb



(a) Filtres "NFC" pour une sources à $r_0 = 5\text{m}$ et des sources de restitution à $r_{HP} = 1.25\text{m}$ (b) Filtres "NFC" pour une sources à $r_0 = 0.8\text{m}$ et des sources de restitution à $r_{HP} = 1.25\text{m}$

FIGURE 2.9.: Filtres "NFC" pour une source à l'extérieur du dispositif de restitution 2.9(a) et à l'extérieur 2.9(b).

Dans ce contexte, un système de restitution en 2 dimensions est envisagé pour le stage. En effet, sa réalisation pratique demande de répartir les sources de restitution sur un cercle autour de l'auditeur ce qui est beaucoup plus simple à réaliser (voir 5.3).

En se basant sur le formalisme en harmoniques sphériques, la restriction à deux dimensions revient à ne conserver que les harmoniques sphériques porteuses de l'information horizontale (voir figure D.1). Pour une décomposition à l'ordre M on a alors $K = 2M + 1$ composantes.

Pour une répartition en cercle de rayon r_{HP} autour de l'auditeur avec N haut-parleurs, le calcul des gains des sources de restitution est donné par [Daniel et Moreau, 2004] :

$$g_i = \frac{1}{N} (H_0(kr_0, r_{HP}) + 2 \sum_{m=0}^M \cos(\gamma_i) H_m(kr_0, r_{HP}, \omega)) \quad (2.29)$$

où γ_i est la différence d'angle entre le $i^{\text{ème}}$ haut-parleur et la source à encoder. On a $\gamma_i = \arccos(\vec{r}_0, \vec{r}_i)$ avec $i = [1 \dots N]$

Sur la figure 2.10 on peut observer quelques exemples de systèmes de restitution 2D. Le dispositif sur la figure 2.10(a) est celui élaboré dans le cadre du stage pour une restitution HOA. Le dispositif sur la figure 2.10(b) est utilisé dans le cadre de l'AWFS (*Adaptive Wave Field Synthesis*) dans [Gauthier et Berry, 2007].



(a) Dispositif de restitution 2D élaboré dans le cadre du stage



(b) Dispositif de restitution 2D utilisé pour l'AWFS dans [Gauthier et Berry, 2007]

FIGURE 2.10.: Exemples de dispositifs de restitution 2D

2.6 RECONSTRUCTION PAR SUPERPOSITION D'ONDES SPHÉRIQUES

On considère les N haut-parleurs de restitution, équirépartis autour de l'origine à un rayon R_{HP} . Le formalisme présenté ci-dessus modélise ces sources de restitution comme des monopôles. Une fois les gains de ces sources connus la pression, en un point $A(r_A, \theta_A, \varphi_A)$ est donnée par :

$$\tilde{p}_A = p(r_M, \theta_M, \varphi_M) \sum_{i=1}^N g_i \Gamma_0(k|r_{HP} - r_A|) \quad (2.30)$$

$p(r_M, \theta_M, \varphi_M)$ est la pression à l'origine (voir equation (2.25)) et $\Gamma_0(k|r_{HP} - r_A|)$ la fonction de propagation entre les haut-parleurs et le point A (fonction de Green en espace libre).

2.7 EXEMPLE EN SIMULATION

Un exemple en simulation de NFC-HOA en deux dimensions est montré sur la figure 2.11. Il s'agit de reconstruire le champ issu d'un monopôle 2.11(a) de débit unitaire situé à $r_0 = 3m$ de l'origine et $\theta_0 = 0^\circ$ et à une fréquence de $f = 1000Hz$ (figure 2.11(a)). Le champ reconstruit à l'ordre $M = 16$ par la méthode décrite ci-dessus est visible figure 2.11(b). Le trait noir représente la limite de la "zone d'écoute" à l'intérieur de laquelle la reconstruction physique respecte un critère d'erreur $e \leq 20\%$ (voir 2.8.1).

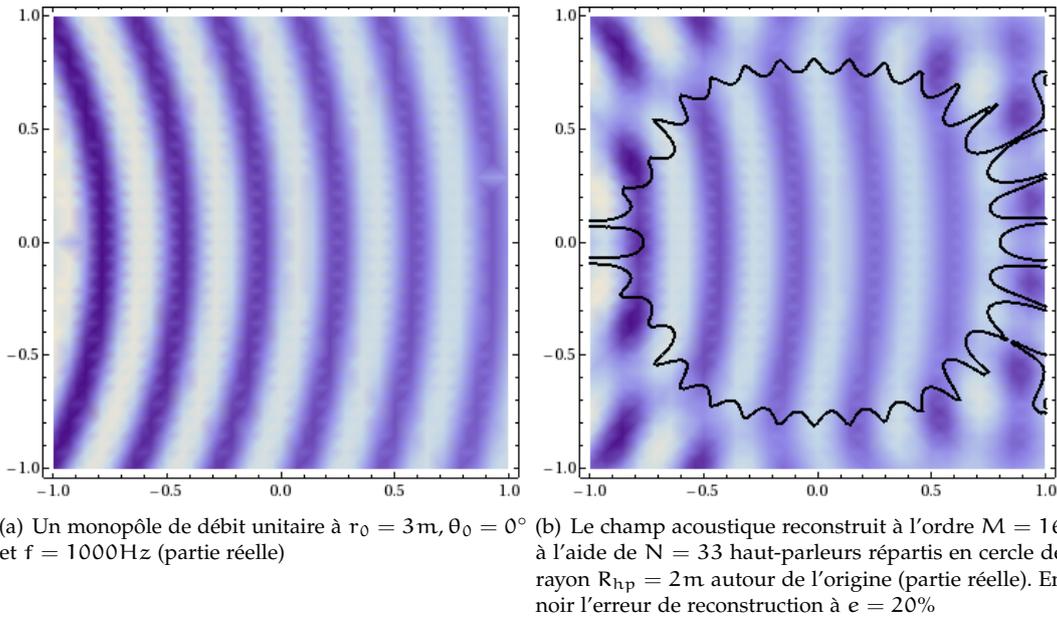


FIGURE 2.11.: Un exemple de simulation pour la NFC-HOA en deux dimensions

La figure 2.11 est réalisée à partir de la feuille de calcul MATHEMATICA©NFC_HOA.nb disponible en ligne⁵. Ce code permet de synthétiser un champ sonore en deux dimensions issu d'un monopôle par la méthode NFC-HOA. L'allure des filtres NFC est également disponible dans ce code.

2.8 CARACTÉRISATION DU CHAMP ACOUSTIQUE RECONSTRUIT

Le champ acoustique reconstruit doit être évalué objectivement et subjectivement pour estimer la qualité de la restitution ambisonique. Plusieurs critères d'évaluation existent dans cette optique, on présente seulement quelques uns d'entre eux dans cette section. Pour plus de détails, on pourra consulter [Daniel, 2000] par exemple.

5. http://www.sekisushai.net/ANFCHOA/CHAP02/NFC_HOA.nb

2.8.1 Critère d'erreur

Plusieurs critères d'erreur sont envisageables pour évaluer objectivement la qualité de la reconstruction physique du champ sonore [Daniel, 2000, p. 168]. Dans le cas présent on utilise, l'erreur relative entre le champ théorique et le champ obtenue (2.31). En fixant un seuil à $e = 20\%$ on obtient le contour d'une "zone d'écoute" où l'on estime que la reconstruction physique est satisfaisante :

$$e(x, y, z) = \frac{|\tilde{p}(x, y, z) - p(x, y, z)|}{|p(x, y, z)|} \times 100 \quad (2.31)$$

D'autres critères d'erreur existent pour notamment définir une zone d'écoute circulaire ([Daniel, 2000, p.160]).

2.8.2 Taille de la zone d'écoute

La taille de la zone d'écoute dépend de l'ordre de reconstruction et de la fréquence. Ainsi plus l'ordre est élevée, plus la zone d'écoute sera grande. De même, plus la fréquence est basse, plus la zone d'écoute est grande. Sur la figure 2.12 on peut observer la reconstruction d'un monopôle en $r_0 = 3\text{m}, \theta_0 = 0^\circ$ à un fréquence de $f = 1000\text{Hz}$ pour différents ordres de décomposition $M = [1 \dots 10]$. Le nombre de haut-parleurs est de $N = 2M + 1$. On observe que la taille de la zone d'écoute augmente avec l'ordre de décomposition.

De même, sur la figure 2.13, on fait varier la fréquence à un ordre de décomposition fixé $M = 7$ ($N = 15$). On observe que plus la fréquence augmente, plus la taille de la zone d'écoute diminue. En utilisant des critères d'erreur permettant de définir une zone d'écoute circulaire et d'avoir alors un rayon pour cette dernière, on peut définir (à un ordre M fixé) une fréquence limite au delà de laquelle la reconstruction physique n'est plus acceptable, la zone d'écoute devenant trop petite [Daniel, 2000, p.160]

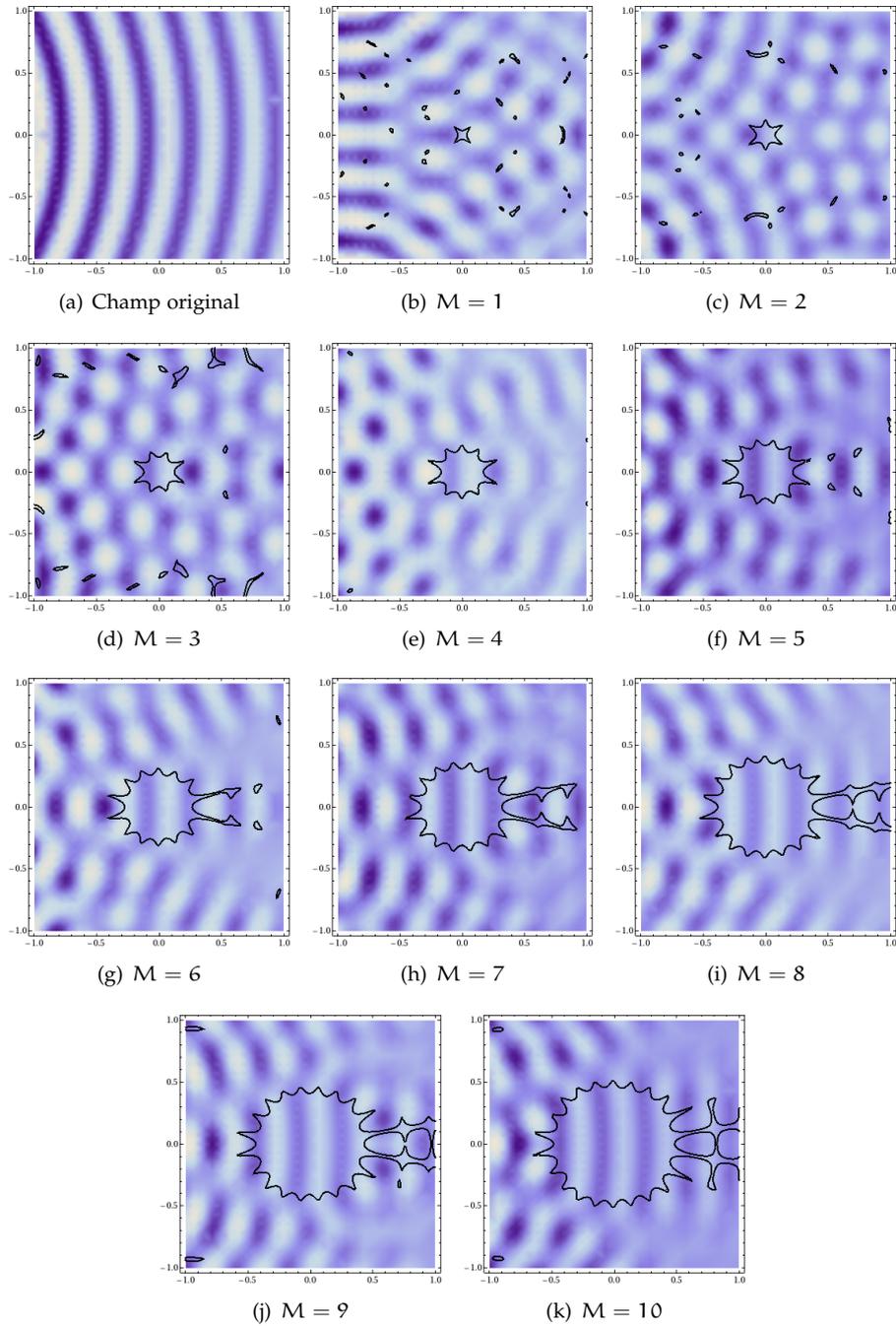


FIGURE 2.12.: Reconstruction à différents ordres d'un champ issu d'un monopôle en $r_0 = 3\text{m}$, $\theta_0 = 0^\circ$ à $f = 1000\text{Hz}$. Le champ original 2.12(a) (partie réelle), et les champs obtenus pour des ordres de décomposition $M = 1 \dots 10$ (partie réelle). Le trait noir délimite la zone où la reconstruction respecte un critère d'erreur $e \leq 20\%$. Les échelles sont en mètre.

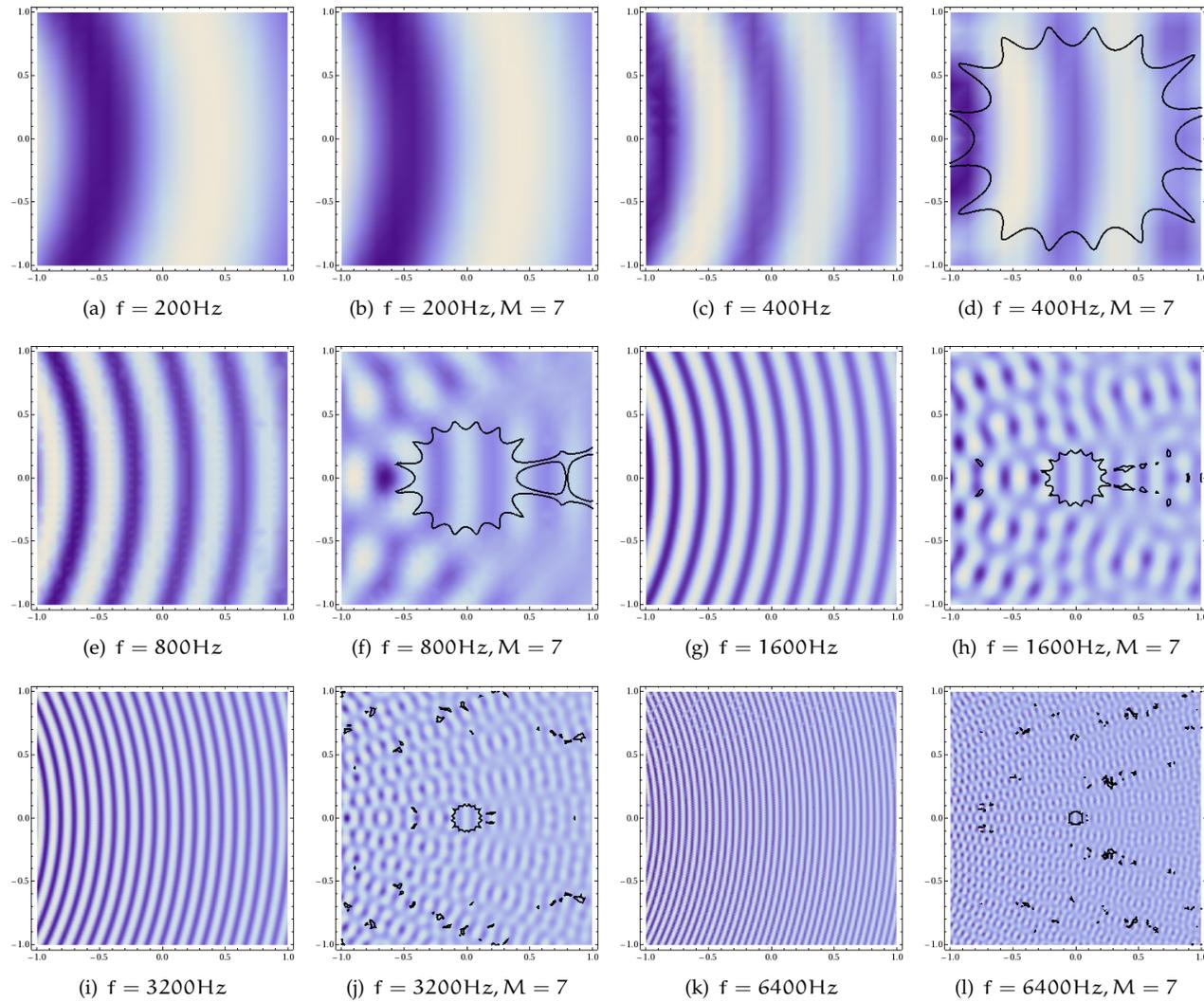


FIGURE 2.13.: Un monopôle à $r_0 = 3\text{m}$, $\theta_0 = 0^\circ$ (partie réelle) reconstruit à l'ordre $M = 7$ pour différentes fréquences $f = 200, 400, 800, 1600, 3200, 6400\text{Hz}$. La taille de la zone d'écoute diminue lorsque la fréquence augmente. Le trait noir délimite la zone où la reconstruction respecte un critère d'erreur $e \leq 20\%$. Les échelles sont en mètres.

2.8.3 Décodage "psychoacoustique"

Comme on vient de le voir la reconstruction physique du champ sonore en hautes fréquences est très coûteuse du fait de la nécessité d'un ordre élevé de décomposition. Qui plus est, la présence d'auditeurs dans la zone d'écoute perturbe le champ sonore synthétisé dans les hautes fréquences où les dimensions de la tête ne sont plus négligeables devant les longueurs d'ondes.

En reconstruisant physiquement le champ et en immergeant l'auditeur dans ce dernier on peut espérer naïvement que la perception obtenue serait la même que si le son provenait réellement des sources simulées. Ceci est vrai en basse fréquence où la tête de l'auditeur perturbe peu le champ sonore. En hautes fréquences la tête diffracte le champ incident et le perturbe. Là où le processus diffère entre le champ original et synthétisé, c'est que la perturbation n'est pas la même. En effet, le champ sonore est synthétisé par des sources qui ne sont pas aux mêmes endroits que les sources du champ sonore original. La combinaison des différents signaux acoustique de ces sources tend à reconstruire physiquement le champ sonore. Si un obstacle, qui se trouve être la tête de l'auditeur perturbe ce mélange de signaux, le champ obtenu n'est plus le même et la perception de ce dernier n'est a priori plus celle du champ original. Dans cette optique un décodage ambisonique prenant en compte les effets subjectifs de localisation est mis en place.

Ce décodage se base sur les indices de localisation de sources. Gerzon a mis en place deux indices permettant de traduire les informations spatiales contenues dans l'encodage ambisonique [Gerzon, 1992]. Il s'agit des vecteurs vitesse et énergie. Le décodage prend en compte ces indices spatiaux pour qu'ils coïncident au mieux avec ceux de la source originale.

2.8.4 Vecteur vitesse

Le vecteur vitesse est défini dans [Gerzon, 1992] par :

$$\vec{r}_v = \sum_{i=1}^N \Re \left[\frac{g_i}{\sum_{i=1}^N g_i} \right] \vec{r}_i \quad (2.32)$$

où g_i sont les gains complexes des différentes sources de restitution, \Re la partie réelle et \vec{r}_i les vecteurs unitaires donnant leur direction. Cette vecteur donne alors une information de direction apparente de la source synthétisée lorsque qu'on se place au centre du dispositif. Ce critère s'applique en basses fréquences pour $f \leq 700\text{Hz}$ [Nicol, 1999, p. 81].

2.8.5 Vecteur énergie

De même, le vecteur énergie est défini par :

$$\vec{r}_e = \sum_{i=1}^N \left[\frac{|g_i|^2}{\sum_{i=1}^N |g_i|^2} \right] \vec{r}_i \quad (2.33)$$

Cette quantité s'applique aux hautes fréquences et donne une information spatiale de direction moyenne d'où provient l'énergie du champ.

En utilisant ces deux indices, basses et hautes fréquence, on peut optimiser le décodage ambisonique classique pour faire coïncider au mieux ces grandeurs entre le champ original et synthétisé. Plusieurs décodeurs sont alors mis en parallèle suivant la zone fréquentielle et les signaux sont passés à travers des "shelf filters" avant d'être envoyé aux décodeurs. Pour plus de détails sur les décodeurs incluant la dimension psychoacoustique, se référer à Nicol [Nicol, 1999, p.79-83] où Daniel [Daniel, 2000].

2.9 CONCLUSION

L'ambisonie est une méthode de synthèse de champ sonore qui présente de nombreux avantages : le nombre de sources de restitution n'est pas nécessairement grand pour avoir de bons résultats (comparativement à la l'holophonie et la WFS (Wave Field Synthesis) par exemple) et la

configuration du système de restitution est très flexible. Enfin, une implémentation temps-réel est envisageable pour des configurations régulières.

Cependant, pour des applications audio, le décodage basique qui s'attache à reconstruire physiquement le champ sonore n'est souvent pas suffisant. En effet, la reconstruction en hautes fréquences est très coûteuse et les auditeurs perturbent le champ sonore. Ainsi des décodeurs mettant en avant le critère de perception au dessus de la reconstruction physique sont utilisés.

Dans le cadre du stage, on projette de mettre en place un système de restitution ambisonique qui utilise seulement le décodage basique, dans un premier temps. En effet, l'objectif serait de reconstruire des champs sonores de manière suffisamment précise pour être utilisés à des fins de mesures acoustiques. Le formalisme NFC-HOA est utilisé car il permet d'utiliser des sources de restitution en champ proche. Cependant, pour améliorer la précision de la reconstruction, il faut se rapprocher au mieux de la théorie et tenter de corriger la réponse des haut-parleurs pour se rapprocher du monopôle et diminuer l'effet de salle pour se rapprocher de la condition champ libre. Bien évidemment ces conditions sont difficiles à obtenir, aussi à l'aide de techniques de déconvolution d'effet de salle, on cherche par la suite à corriger la réponse du dispositif de restitution pour l'amener au plus proche de la théorie NFC-HOA. Ceci est l'objet des deux prochains chapitres.

A noter que quelques cartes de champs sonores mesurés sur le dispositif de restitution du stage son visibles à partir du chapitre 4.

CORRECTION DES CANAUX DE RESTITUTION AU CENTRE DU DISPOSITIF

3.1 INTRODUCTION GÉNÉRALE

Comme vu précédemment, le formalisme ambisonique requiert une condition de champ libre car à aucun moment l'effet de la salle de restitution n'est prise en considération. Une application en dehors d'une chambre anéchoïque peut détériorer les performances du système si ce dernier n'est pas "adapté" à la salle [Deprez *et al.*, 2010]. Effectivement, les réflexions des ondes sonores sur les parois de la salle perturbent le champ synthétisé par le système NFC-HOA. L'idée est alors de corriger cet effet de salle et plus généralement tout le système de restitution, incluant la réponse de amplificateurs et haut-parleurs, pour se rapprocher au mieux de la condition de champ libre. En effet, en plus de l'effet de salle, les haut-parleurs de restitution ne sont pas des monopôles et ne sont pas identiques entre eux. Leur réponse est alors à corriger également.

Plusieurs approches sont possibles pour diminuer l'effet de la salle sur la zone d'écoute. Une première idée est de tenter d'obtenir une réponse en fréquence "plate" sur un ensemble de points de contrôle situés dans la zone d'écoute. On se sert alors de tous les haut-parleurs à disposition pour corriger les positions. Une fois la réponse corrigée sur ces points, on applique les gains NFC-HOA classiques. Cependant la correction multicanale peut parfois dégrader la réponse en dehors des points de contrôle [Spors *et al.*, 2003].

Une autre approche est de venir corriger les gains NFC-HOA à partir de la mesure réelle du champ obtenu dans la salle. Ainsi, pour chaque position de source synthétisée on trouve une correction (par inversion dans L^2 par exemple) pour rapprocher le champ obtenu du champ théorique. Gautier et Berry proposent une démarche dans ce sens pour la WFS [Gauthier et Berry, 2007], Corteel également [Corteel, 2006] ou encore Stefanakis [Stefanakis *et al.*, 2010]. Cependant, ces techniques peuvent s'avérer très coûteuses si il faut calculer autant de correction que de positions de sources à synthétiser.

Les techniques de synthèse de champs sonores comme l'ambisonie d'ordre élevé ou la "Wave Field Synthesis" (WFS) permettent de contrôler le champ sur une zone étendue, la zone d'écoute. On peut alors tirer avantage de cette théorie pour superviser la correction multicanale. Notre objectif est d'obtenir une réponse de la salle la plus anéchoïque possible sur la zone d'écoute. Ainsi si l'on corrige la réponse de chaque canal de restitution en utilisant une égalisation multicanale combinée à une approche ambisonique, on obtient en théorie une correction valable non seulement sur les points de contrôles, mais également aux alentours [Stefanakis *et al.*, 2010].

Dans le cadre du stage, l'idée est alors de ne corriger que les canaux de restitution en amont de l'application de la NFC-HOA. Ainsi cette opération se ferait par étalonnage et non par contrôle actif. On corrige seulement les N positions de sources correspondants aux sources de restitution suivant l'approche proposée par Gauthier et Berry dans le cadre de l'"Adaptive Wave Field Synthesis" (AWFS) [Gauthier et Berry, 2007]. En théorie, une fois ces sources de restitution corrigées, on peut appliquer la NFC-HOA pour synthétiser n'importe quelle autre source, les gains étant alors filtrés par la matrice de filtres de correction. Par la suite on appellera cette technique l'ANFC-HOA : *Adaptive Near Field Compensated Higher Orders Ambisonics*.

Le système de restitution ANFC-HOA est installé dans la chambre semi-anéchoïque du LMSSC. En effet, on souhaite dans un premier temps corriger un effet de salle "minime" dû à la seule réflexion sur le sol. En cas de résultats convaincants, le travail sera poursuivi dans une salle réelle.

La stratégie adoptée est la suivante : En premier lieu, on souhaite corriger la réponse du système canal par canal, en un seul point de contrôle au centre du dispositif. L'objectif est de regarder l'éventuelle amélioration par rapport à la situation sans correction et de regarder la faisabilité temps-réel avec N filtres de correction sur une machine de type ordinateur de bureau. (3.3). Ceci fait l'objet du présent chapitre. Ensuite, le nombre de points de correction est augmenté autour

de du point d'origine pour étendre la zone de correction sur la zone d'écoute ambisonique, ce qui est le sujet du chapitre 4

La correction canal par canal consiste alors à corriger chaque canal par lui même (un canal est vu comme l'association amplificateur - haut-parleur - microphone). Pour la correction multicanale on se sert de tous les canaux à disposition pour corriger l'un d'entre eux sur un ensemble de points de contrôle. Par permutation on construit alors une matrice de filtres (voir 4).

Avant de créer les filtres de correction, il faut connaître la réponse de chaque canal de restitution sur les points de contrôle. Ceci est fait par la mesure de la réponse impulsionnelle des canaux sur les différents points de contrôle. La section suivante présente la manière dont la mesure de ces réponses est faite.

3.2 MESURE DU GAIN COMPLEXE ET DE LA RÉPONSE IMPULSIONNELLE D'UN CANAL À UNE POSITION

3.2.1 Méthode d'estimation choisie

Un canal est considéré comme l'association d'une voie d'amplificateur et d'un haut-parleur à une position dans la salle de restitution (voir figure 3.2).

Différentes techniques existent pour acquérir la réponse impulsionnelle d'un canal. Un aperçu des principales est présenté dans [Müller et Massarani, 2001]. La démarche est cependant commune pour les différentes techniques : un signal d'excitation contenant les fréquences d'intérêt est envoyé au système. Ce signal est considéré comme la référence. Après passage à travers l'amplificateur, le haut-parleur, la salle et le microphone, le signal est mesuré. Le signal mesuré est ensuite comparé à la référence pour mettre en évidence la fonction de transfert du système (voir A.1). Parmi les stimuli utilisés dans la littérature, on peut citer les séquences MLS ou encore le sinus glissant [Farina, 2000]. L'avantage du sinus glissant est de rejeter facilement les distorsions du système après leur mise en évidence [Farina, 2000]. Ceci permet alors d'augmenter le rapport signal sur bruit en faisant fonctionner l'amplificateur et le haut-parleur à volume élevé voire dans leur zone non linéaire.

Dans notre cas, la réponse impulsionnelle est mesurée en chambre semi-anéchoïque en premier lieu. Le rapport signal sur bruit semble déjà très suffisant. Les amplificateurs et haut-parleurs sont a priori excités dans leur zone linéaire ce qui limite les distorsions. Enfin, l'hypothèse de stationnarité est assez plausible dans les conditions de mesure et l'on utilise un bruit blanc comme signal d'excitation, selon les méthodes de mesure du laboratoire.

La démarche pour obtenir le gain complexe et la réponse impulsionnelle d'un canal est détaillée schématiquement sur la figure A.1, annexe A. La théorie associée est présentée dans l'annexe A. Un bruit blanc généré par la carte son est envoyé à l'amplificateur et re-numérisé en parallèle pour obtenir une référence synchrone. En effet, l'opération de conversion numérique-analogique n'est pas instantanée. Le fait de re-numériser le signal de référence en même temps que le signal de mesure permet alors d'enlever la latence de la carte. Après amplification, le signal est envoyé vers le haut-parleur et propagé acoustiquement dans la salle. Le microphone capte alors ce signal de mesure qui est pré-amplifié et numérisé. À partir des signaux de mesure et de référence, on estime le gain complexe et la réponse impulsionnelle suivant la démarche détaillée dans l'annexe A.

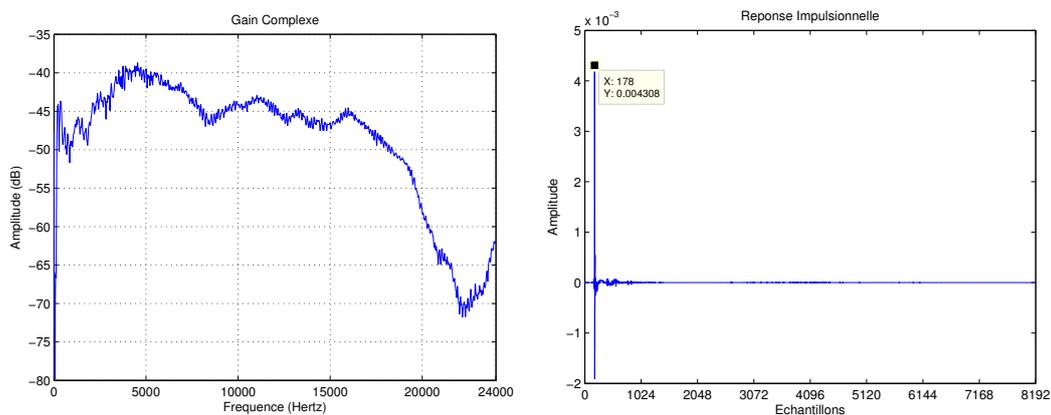
A noter qu'après calcul des filtres de correction, la mesure des réponses impulsionnelles par la technique présentée dans l'annexe A a fait défaut ! En effet, la cohérence estimée entre le signal électrique d'excitation et la mesure tombait à zéro, donnant un gain complexe inexploitable (voir 3.4.3). On s'est alors tourné vers une mesure par sinus glissant suivant la méthode de Farina [Farina, 2000] pour évaluer la qualité des filtrages. L'annexe B détaille cette autre méthode d'estimation des réponses impulsionnelles.

3.2.2 Exemple de mesure

Sur la figure 3.1(a) on peut observer le gain complexe du canal 1 et sa réponse impulsionnelle associée pour le système de restitution. On observe un premier pic de la réponse impulsionnelle à

$n = 178$ échantillons. Ceci correspond à l'arrivée de l'onde directe émise par le haut-parleur sur le microphone. En effet à une fréquence d'échantillonnage de $F_e = 48000\text{Hz}$, le signal arrive après un temps $t = n/F_e = 3.7\text{ms}$. En considérant la vitesse du son $c = 340\text{m/s}$ ce temps de propagation correspond à une distance de $d = ct = 1.26\text{m}$. On retrouve donc la distance du centre acoustique du haut-parleur par rapport au centre du dispositif de restitution. Cette observation peut être utile pour régler précisément la position radiale des haut-parleurs par méthode acoustique (sous réserve de connaître précisément la vitesse du son).

La salle de restitution est une chambre semi-anéchoïque. Le sol réfléchissant effectue alors un filtrage en peigne (voir 4.2). qui s'observe sur le gain complexe par une modulation de la réponse en fréquence et la présence d'un second pic à $n = 409$ correspondant à la réflexion sur le sol de l'onde acoustique. Par la méthode des sources images, on peut également retrouver la hauteur des haut parleur par rapport au sol ($h \approx 1.25\text{m}$).



(a) Gain du canal 1 au centre du système de restitution (b) Réponse Impulsionnelle du canal 2 au centre du système de restitution

FIGURE 3.1.: Module du gain complexe du canal 1 et réponse impulsionnelle associée

La figure 3.1 est obtenue à l'aide du code MATLAB©GainComplexe.m avec en entrée les deux fichiers WAV HP1Ref.wav et HP1Mes.wav. Ce code permet de calculer le gain complexe (et la réponse impulsionnelle) par filtrage de l'interspectre (voir annexe A). Il est disponible en ligne ainsi que les fichiers WAV¹.

Après mesure des N réponses impulsionnelles au centre du dispositif de restitution, on conçoit des filtres inverses pour corriger les canaux par eux même et obtenir une "réponse plate", ce qui fait l'objet de la section suivante.

3.3 INVERSION DE LA RÉPONSE IMPULSIONNELLE D'UN CANAL, CALCUL D'UN FILTRE INVERSE.

3.3.1 Introduction

L'inversion de la réponse impulsionnelle d'une salle, d'un haut-parleur ou plus généralement d'un canal de restitution est sujet à beaucoup d'articles dans la littérature. On parle également de déconvolution. L'objectif est de trouver un filtre qui annule, qui "déconvolue" l'effet d'un filtrage direct par propagation acoustique dans une salle, par le passage du signal dans un transducteur, etc. Différentes techniques sont utilisées pour concevoir un filtre inverse efficace. En effet, l'inversion "simple" conduit dans la majorité des cas à des filtres non causaux et instables. Ceci est du à la nature du filtre direct qui, dans l'absolu, n'est pas à minimum de phase (voir 3.3.2). Le problème de déconvolution peut être approché de manières fréquentielle ou temporelle.

1. <http://www.sekisushai.net/ANFCHOA/CHAP03/>

Formulation temporelle

Considérons le signal temporel discrétisé $y[n]$ provenant du microphone au point de mesure (le temps est indexé par n). Ce signal est obtenu à partir du signal $x[n]$ envoyé à l'amplificateur puis diffusé par le haut-parleur à travers la salle. Cette transformation du signal peut être alors vue comme un filtrage du signal d'entrée par la réponse impulsionnelle $c[n]$ du système amplificateur - haut-parleur - salle. Le filtrage correspond à l'opération de convolution linéaire dans le domaine temporel. On a de fait :

$$y[n] = c[n] \star x[n] \quad (3.1)$$

où \star est l'opérateur de convolution linéaire.

L'objectif est alors de trouver le filtre inverse $h[n]$ qui "déconvolue" le filtrage par $c[n]$ pour obtenir un signal $y[n]$, le plus proche de $x[n]$ (ou d'une autre consigne, $a[n]$). Le processus de filtrage est montré schématiquement figure 3.2). On a :

$$w[n] = h[n] \star c[n] \star x[n] \quad (3.2)$$

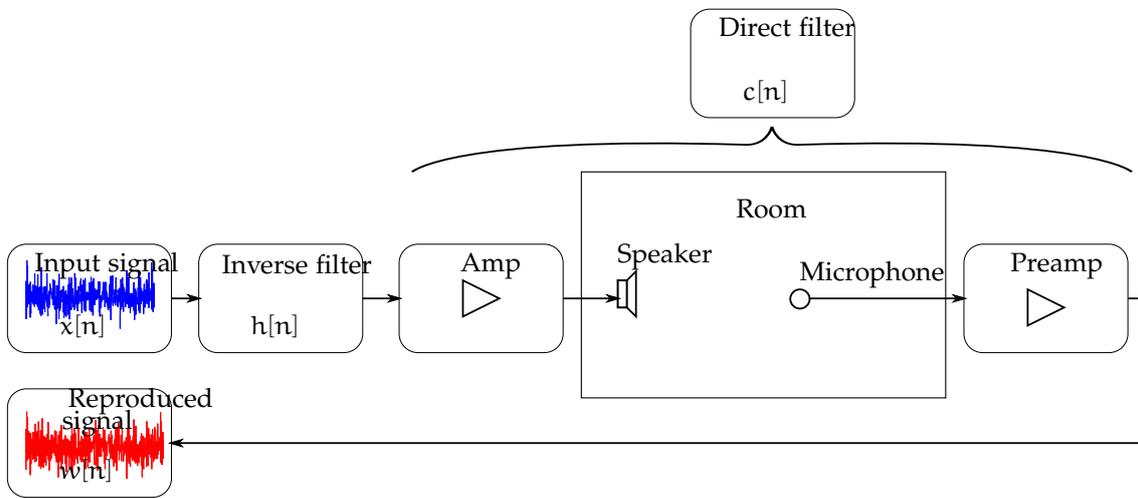


FIGURE 3.2.: Schéma de principe de la déconvolution sur un canal

Dans le cas où l'on veut retrouver le signal $w[n]$ égal à l'entrée $x[n]$ le problème revient à chercher $h[n]$ tel que :

$$\delta[n] = h[n] \star c[n] \quad (3.3)$$

Si l'on veut que le signal $w[n]$ s'approche d'un signal cible $a[n]$ on a

$$a[n] = h[n] \star c[n] \quad (3.4)$$

Formulation fréquentielle

Dans le domaine spectral, l'opération de filtrage revient à une simple multiplication. En notant k l'indice fréquentiel pour une transformée de Fourier discrète on a :

$$W[k] = C[k]H[k]X[k] \quad (3.5)$$

Dans le cas où l'on souhaite $W[k] = X[k]$ on a

$$H[k] = \frac{1}{C[k]} \quad (3.6)$$

et pour un signal cible $A[k]$:

$$H[k] = \frac{A[k]}{C[k]} \quad (3.7)$$

En partant de la formulation du problème dans le domaine temporel ou spectral, l'objectif est de trouver le filtre $h[n]$ (ou $H[k]$) qui déconvolue au mieux l'effet du filtrage par $c[n]$ (ou $C[k]$).

3.3.2 État de l'art

Filtre à minimum de phase

Le principal problème dans la conception d'un filtre inverse est qu'une inversion sans prise de précautions produit dans la majorité des cas des filtres non causaux et non stable. En effet, le filtre direct $c[n]$ n'est généralement pas à minimum de phase. Un filtre à minimum de phase possède tous les zéros de sa transformée en z à l'intérieur du cercle unité. [Blanchet et Charbit, 2001, p. 131]. Une inversion de la transformée en z remplace les pôles par des zéros. Dans le cas où des zéros sont à l'extérieur du cercle unité dans le filtre direct, le filtre inverse devient instable ! Or si le filtre est à minimum de phase, son inverse est assuré d'être stable et causal.

Neely et Allen [Neely et Allen, 1979] ont ainsi proposé de modéliser la réponse impulsionnelle d'une salle par la combinaison d'un filtre passe-tout et d'un filtre à minimum de phase, et d'inverser uniquement la partie à minimum de phase, ce qui assure au filtre inverse d'être stable et causal.

Inversion temporelle

L'inversion par méthode temporelle consiste à remplacer l'opération de convolution par un produit matriciel à l'aide d'une matrice de Toeplitz [Blanchet et Charbit, 2001, p. 629]. Le calcul du filtre inverse se fait alors par inversion d'une matrice, avec ou sans régularisation (voir 3.3.3). L'avantage de ce genre de méthode est qu'à aucun moment on ne passe dans le domaine spectral qui peut présenter des inconvénients lors de l'utilisation de TFD. En effet, l'opération de convolution est circulaire lorsque l'on utilise une TFD ou TFD inverse [Oppenheim et Schafer, 1975, p. 105]. Ceci introduit alors des effets de repliements lors du retour dans le domaine temporel par TFD inverse.

Le principal inconvénient des formulations temporelles est la taille importante des matrices mises en jeu. De telles méthodes ne sont alors pas envisageables lorsque la taille des filtres dépasse quelques milliers d'échantillons, ce qui est le cas dans le cadre du stage.

Kirkeby et Nelson [Kirkeby et Nelson, 1999] proposent une méthode temporelle de design de filtre inverse au sens des moindres carrés avec régularisation. Un retard est introduit pour compenser les composantes non minimum de phase du filtre direct (voir 3.3.3). Le paramètre de régularisation β est choisi dépendant de la fréquence, ceci pour limiter le gain du filtre inverse dans certaines régions fréquentielle. En effet, si le canal a une très faible réponse en basse fréquence, par exemple, le filtre inverse va tenter de corriger cette réponse en amplifiant énormément les basses fréquences, ce qui peut alors détériorer le canal, notamment le haut-parleur de restitution. Le paramètre β sert alors à contrôler l'inversion. Plus de détails sur cette approche sont donnés dans la section 3.3.3.

Inversion fréquentielle

L'inversion dans le domaine spectral est attrayante car peu coûteuse. En effet, l'opération de filtrage revient à une simple multiplication, mais le principal inconvénient est le caractère fini des signaux. L'opération de multiplication dans le domaine spectral correspond alors à une convolution circulaire et non pas linéaire dans le domaine temporel lorsque l'on utilise des TFDs [Oppenheim et Schafer, 1975, p. 105]. Lors du retour dans le domaine temporel à l'aide de TFDs inverse, des effets de repliements peuvent apparaître sur les réponses impulsionnelles obtenues.

La convolution linéaire peut être calculée à l'aide de TFD en insérant des zéros à la suite de la réponse impulsionnelle [Oppenheim et Schafer, 1975, p. 110], technique dite du "zéro-padding". Ce "zéro-padding" n'est cependant pas efficace dans le calcul des filtres inverses [Kirkeby *et al.*, 1998]. En voulant déconvoluer deux séquences, il faudrait insérer des zéros à l'infini pour annuler complètement les effets de convolution circulaire. Ainsi on utilise des formulations avec régularisation pour diminuer ces effets de repliements [Kirkeby *et al.*, 1998], ou bien on peut concevoir des filtres de grande longueur pour renvoyer les effets de repliements loin de la partie utile du filtre [Norcross *et al.*, 2002].

Inversion par modélisation

Une modélisation du filtre direct par un filtre à réponse impulsionnelle infinie (RII) et d'ordre peu élevé peut faciliter l'inversion [Mourjopoulos et Paraskevas, 1991]. Cependant, la simplification du filtre direct se fait au prix d'une baisse de performance de la correction [Fielder, 2001].

Une autre approche peut se faire par la méthode de la fenêtre [Blanchet et Charbit, 2001, p. 137], en ayant le gabarit fréquentiel du filtre inverse désiré et une modélisation analytique de ce dernier. Cette méthode est cependant très coûteuse car elle emploie une transformée de Fourier à temps continu. De plus le filtre obtenue ne corrige pas la réponse en phase.

Pour le stage, une simplification du gain complexe du canal est éventuellement opéré à l'aide d'un lissage cepstrale 3.3.3.

Évaluation subjective

Dans [Norcross *et al.*, 2002] on compare subjectivement plusieurs conceptions de filtres inverses, obtenus par la méthode des moindres carrés dans le domaine temporel et par méthodes fréquentielles. Il en ressort que les filtres inverses calculés au sens des moindres carrés sont plus performants subjectivement. Seuls les longs filtres calculés par méthode fréquentielle deviennent efficace. Cependant il est également important que la longueur des filtres inverses ne soit pas trop grande pour éviter des artefacts dans le domaine temporel [Fielder, 2001]

3.3.3 Techniques utilisées dans le cadre du stage

Dans le cadre du stage, les réponses impulsionnelles sont acquises sur $l = 8192$ échantillons. La réponse impulsionnelle dans la chambre semi-anéchoïque (figure 3.1(b)) est contenu majoritairement dans les 1024 premiers échantillons. Cependant l'objectif est d'établir des méthodes de déconvolution pour des salles réelles où les réponses impulsionnelles sont plus longues. Les méthodes de déconvolution temporelles sont alors vite abandonnées dans le cadre du stage, car elle mettent en jeu des matrices aux dimensions largement supérieures à $l \times l$. Pour la suite, le design des filtres inverses se fait dans le domaine fréquentiel. L'utilisation du logiciel BRUTEFIR pour l'implémentation temps-réel (voir 5.6.1) et les méthodes d'inversion utilisés nous ont poussés vers le choix du de filtres à réponses impulsionnelles finies (RIF).

Une fonction MATLAB®FIR.m, disponible en ligne ², permet de concevoir un filtre inverse RIF suivant plusieurs techniques présentées ci-après.

Ajout d'un passe bande

En observant le gain sur la figure 3.1(a) on remarque que la réponse est fortement atténuée en basses fréquences ($f_{\min} \leq 180\text{Hz}$) et hautes fréquences $f_{\max} \geq 20000\text{Hz}$. Ceci est dû à la réponse du haut-parleur qui est atténuée en dessous de f_{\min} (voir les haut-parleurs utilisés 5.3). En s'approchant de la fréquence de Nyquist ($F_e/2$) la réponse est également atténuée : Ceci est du au filtre anti-repliement présent lors de la numérisation [Blanchet et Charbit, 2001, p. 48].

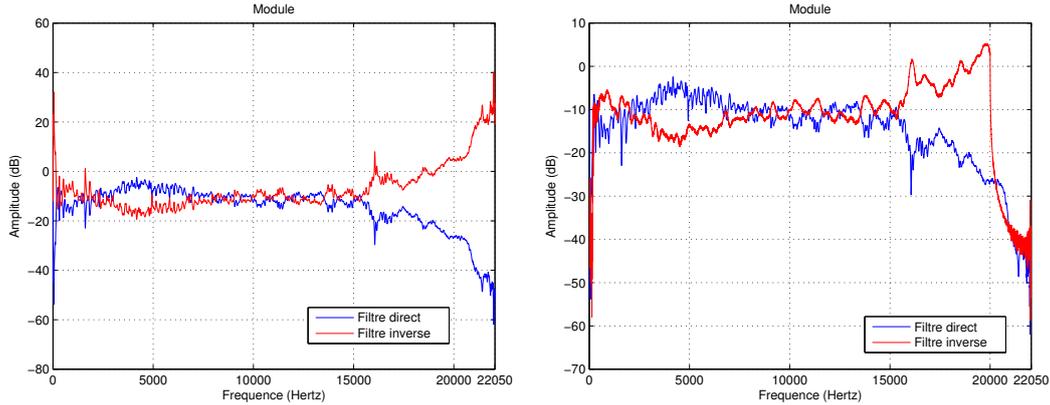
Lors de l'inversion, le filtre inverse aura un gain très important dans ces zones. Ceci risque de détériorer les haut-parleurs. Pour prévenir ce problème, on applique un filtre passe bande de bande passante $[f_{\min}, f_{\max}]$ par dessus le filtre inverse. Ceci est réalisé soit après l'inversion à proprement dite, soit pendant l'inversion dans le cas de l'inversion au sens des moindres carrés (voir 3.3.3). Sur la figure 3.3 on peut observer les gabarits fréquentiels du filtre inverse du canal 1 avec et sans filtre passe bande.

Le filtre passe bande est obtenu à l'aide de la fonction `firls` de MATLAB® et il est appliqué au filtre inverse à l'aide de la fonction `filter`.

Gain du filtre inverse

Pour éviter que le filtre inverse ait un gain moyen trop important, on fixe un gain moyen pour l'inversion. Par exemple, si la valeur moyenne du module du gain du filtre direct est à -20dB ,

2. <http://www.sekisushai.net/ANFCHOA/CHAP03/FIR.m>



(a) Gabarit fréquentiel du filtre direct (bleu) et inverse (rouge)
 (b) Gabarit fréquentiel du filtre direct (bleu) et inverse après l'ajout d'un passe-bande de bande passante 200Hz, 20000Hz(rouge)

FIGURE 3.3.: Gabarits fréquentiels du filtre inverse sans et avec passe-bande

le filtre inverse aura sa moyenne à +20dB si l'on utilise la formulation (3.6). En effet dans ces formulations, l'objectif est d'obtenir un gain de 0dB et une réponse plate. On fixe alors la valeur moyenne requise à la valeur moyenne maximale parmi les gains des filtres directs des 15 canaux. En appelant ce gain G , la formulation (3.6) devient :

$$H[k] = \frac{G}{C[k]} \quad (3.8)$$

Les N canaux de restitution auront alors un niveau moyen semblable après correction.

Liffrage cepstral

On intuite que le filtre inverse sera d'autant plus "simple" à réaliser que le filtre direct l'est aussi. Évidemment la simplification du filtre direct se fait au détriment des performances du filtre inverse mais permet d'obtenir des filtres inverse plus courts. L'effet est alors comme une régularisation lors de l'inversion dans le cadre de l'inversion au sens des moindres carrés, voir 3.3.3. Pour simplifier le filtre direct, on utilise une technique de liffrage cepstral. Beaucoup utilisé dans le traitement de la voix, le passage par le cepstre permet notamment de séparer la source du filtre dans le cas de modèles sources filtres du signal vocal [Noll, 1967]. De même, ceci peut permettre d'extraire l'enveloppe de la réponse en fréquence d'un filtre [Oppenheim et Schafer, 2004]. Dans notre cas l'effet serait d'extraire l'enveloppe du filtre direct $c[n]$ et de fait de diminuer l'effet du liffrage en peigne (voir 4.2). Un exemple est donné figure 3.4.

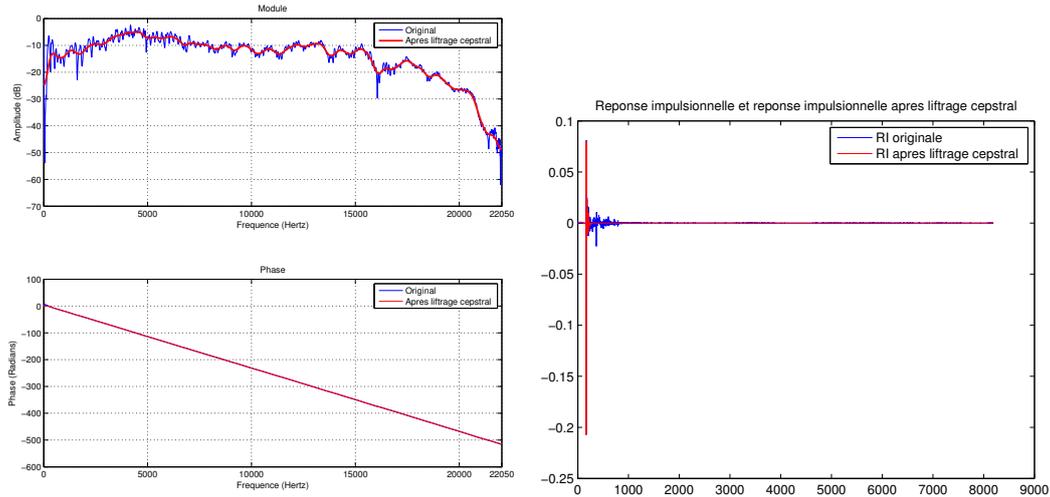
Le cepstre est défini de la manière suivante [Oppenheim et Schafer, 2004] :

$$\begin{cases} \tilde{x}_{\text{complex}}[\tau] = \text{TFD}^{-1}(\log(X[f])) & \text{pour le cepstre complexe} \\ \tilde{x}_{\text{real}}[\tau] = \text{TFD}^{-1}(\log(|X[f]|)) & \text{pour le cepstre réel} \end{cases} \quad (3.9)$$

où $X[f]$ représente la TFD du signal $x[n]$. L'indice τ est homogène à un temps. Dans notre cas, nous souhaitons également garder l'information de phase du filtre direct et l'on utilise le cepstre complexe $\tilde{x}_{\text{complex}}(f)$. Pour alléger la notation on notera le cepstre complexe $\tilde{x}(f)$ par la suite.

Dans le domaine cepstral, les premiers coefficients représentent les variations lentes du spectre (son enveloppe) alors que les coefficients plus élevés renferment ses variations fines (voir figure 3.5(a))

Le principe du liffrage utilisé ici consiste alors à atténuer les coefficients qui représentent les variations fines du spectre. On utilise pour ce faire un fenêtrage du cepstre à l'aide d'une demi fenêtrage de Blackman, qui tombe à 0 à ses extrémités.



(a) Gabarit fréquentiel du filtre direct sans modification (bleu) et après liftrage cepstral, amplitude et phase (bleu) et après liftrage cepstral, amplitude et phase impulsionnelle obtenue après liftrage cepstral (rouge)

FIGURE 3.4.: Module du gain du canal 1 et réponse impulsionnelle associée avec et sans liftrage cepstral. Coefficients cepstraux conservés : 128

La Fonction MATLAB©Liftrage.m, disponible en ligne³, réalise cette opération. Elle utilise la fonction cceps de MATLAB©pour calculer le cepstre complexe et icceps pour retrouver la réponse impulsionnelle après liftrage cepstral.

L'effet du liftrage sur le filtre inverse est montré à la section 3.4.

Inversion fréquentielle aux sens des moindres carrés

Une première approche de conception de filtre inverse est l'approche par minimisation d'une fonction au sens des moindres carrés. La formulation du problème dans le domaine spectral est schématisé figure 3.6. k représente alors l'indice fréquentiel avec l'utilisation de TFDs.

La longueur des vecteurs est de l . Le terme $D[k]$ correspond à la TFD d'un retard pur $d[n]$ introduit pour produire un filtre inverse causal [Blanchet et Charbit, 2001, p. 138] et compenser les composantes non minimum de phase du filtre direct [Kirkeby et Nelson, 1999] [Norcross *et al.*, 2002]. Il est fixé à $l/2$. $A[k]$ est la TFD du signal cible $a[n]$. Lorsque que l'on veut inverser le filtre direct pour retrouver le signal $x[n] = w[n]$ d'après (3.2), $A[k]$ est un filtre passe tout (TFD de $\delta[0]$). Cependant on peut vouloir viser un autre signal cible. Dans notre cas on souhaite corriger le canal pour obtenir la réponse d'un monopole en champ libre (au gain près, fixé par G). $A[k]$ est alors le TFD d'un retard pur correspondant au temps de propagation entre le haut-parleur et microphone.

Pour obtenir le filtre inverse $H[k]$, on cherche à minimiser une fonction coût définie par

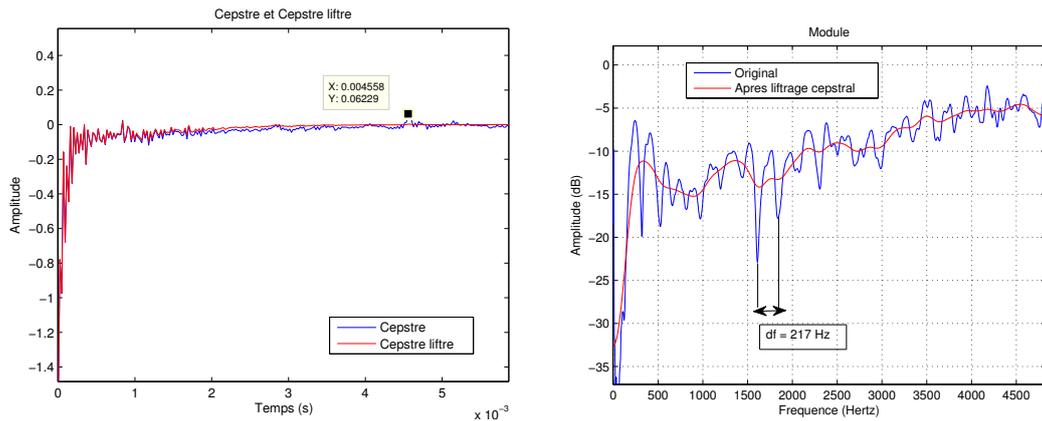
$$J = E[k] + \beta[k]V[k] \quad (3.10)$$

où $E[k]$ représente le signal d'erreur, $\beta[k]$ le paramètre de régularisation et $V[k]$ le signal en sortie du filtre inverse, terme d'effort. Ainsi le paramètre $\beta[k]$ permet de régler le compromis entre l'exactitude de la solution et son énergie [Kirkeby et Nelson, 1999].

L'algorithme d'inversion au sens des moindres carrées que l'on utilise est inspiré de l'article [Kirkeby et Nelson, 1999]. On utilise le paramètre de régularisation β de manière dépendant de la fréquence pour contrôler l'inversion et empêcher de trop amplifier pour les fréquences en dehors de f_{\min} et f_{\max} (voir 3.3.3), ce qui revient à utiliser un filtre de régularisation, dans notre cas un filtre coupe bande. La formulation utilisée est tirée de [Norcross *et al.*, 2002, formule 1.10] :

$$H[k] = \frac{D[k]C^*[k]}{C[k]C^*[k] + \beta B[k]B^*[k]} \quad (3.11)$$

3. <http://www.sekisushai.net/ANFCHOA/CHAP03/Liftrage.m>



(a) Cepstre de la réponse impulsionnelle du canal 1 en (b) Détail du module du gain complexe du canal 1. En bleu (en référence à la figure 3.4). En rouge le cepstre bleu avant lifrage cepstral, en rouge après lifrage. Le pic à $\tau \approx 4$ ms correspond à la modulation fréquentielle de $f = 217$ Hz sur la figure 3.5(b)

FIGURE 3.5.: Cepstre du canal 1 avant et après lifrage. Coefficients cepstraux conservés : 512. En atténuant le pic à $\tau \approx 4$ ms (a), le lifrage en peigne de $df = 217$ Hz est atténué (b)

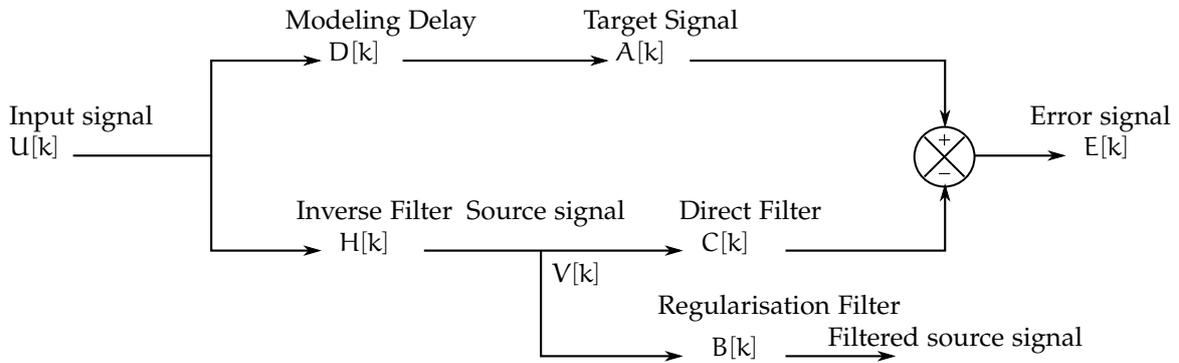


FIGURE 3.6.: Formulation du problème de déconvolution d'un canal en un point

$B[k]$ représente le filtre de régularisation. Il s'agit de la TFD d'un filtre coupe bande $b[n]$ de bande atténuée $[f_{\min}, f_{\max}]$. Finalement, à l'aide d'une TFD inverse on obtient la réponse impulsionnelle du filtre inverse $h[n]$. Le filtre est causal grâce au délai $D[k]$ introduit dans le domaine spectral et il est stable de par sa nature RIF [Blanchet et Charbit, 2001, p. 115].

Le paramètre β est choisi constant sur toute la bande de fréquence, le filtre de régularisation étant défini dans $B[k]$ et il contrôle alors l'énergie du filtre inverse [Kirkeby *et al.*, 1998].

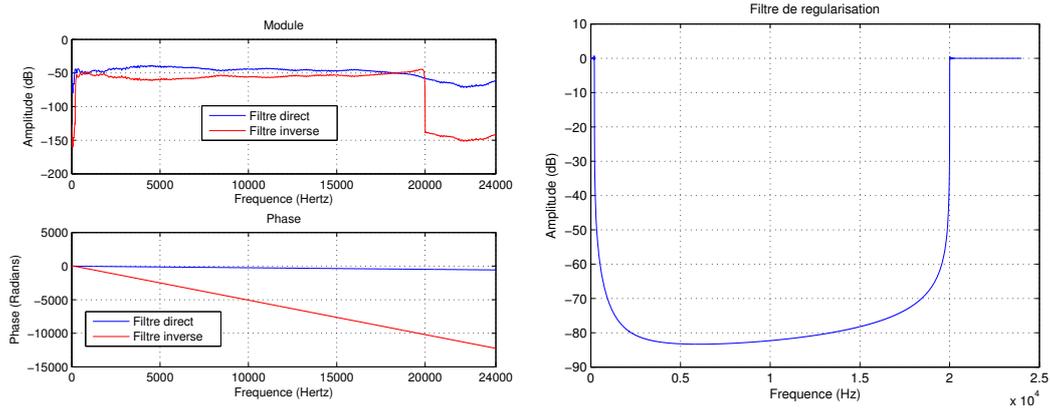
Sur la figure 3.7 on peut observer le gabarit fréquentiel d'un filtre obtenue par inversion au sens des moindres carrés avec $\beta = 0.1$ (figure 3.7(a) en rouge). Le filtre de régularisation utilisé est visible sur la figure 3.7(b). En fixant le signal cible $a[n] = \delta[0]$, on remarque que la réponse en phase du filtre inverse n'est pas l'opposé de celle du filtre direct. Ceci est normal dans la mesure où un retard est artificiellement introduit pour rendre la réponse impulsionnelle du filtre causal. En retirant ce retard, on obtient une réponse en phase opposée mais le filtre n'est plus implémentable car non causal.

La fonction MATLAB©LS_Filter.m, disponible en ligne⁴, réalise le calcul d'un filtre inverse au sens des moindres carrés à l'aide de la réponse impulsionnelle du filtre direct.

Inversion sans régularisation

Une inversion sans régularisation dans le domaine fréquentiel est opérée à l'aide de la formule (3.8). Pour produire un filtre causal et compenser les composantes non minimum de phase du

4. http://www.sekisushai.net/ANFCHOA/CHAP03/LS_Filter.m

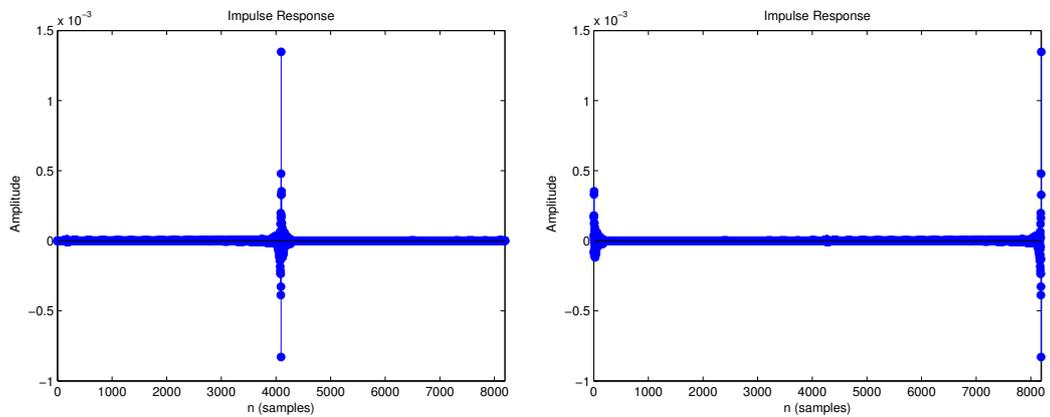


(a) Module et phase du filtre direct (bleu) et inverse (rouge) (b) Filtre de régularisation utilisé pendant l'inversion : coupe bande entre 200Hz et 20000Hz

FIGURE 3.7.: Conception d'un filtre inverse par méthode des moindres carrés avec régularisation et filtre de régularisation utilisé ($\beta = 0.1$)

filtre direct, on introduit un délai $d[n]$ après retour dans le domaine temporel. Ceci est réalisé en faisant une permutation circulaire de $l/2$ échantillons (voire figure 3.8). Cette opération est équivalente à appliquer $D[k]$ au numérateur dans la formule (3.8) [Kirkeby *et al.*, 1998].

Sur la figure 3.8 on peut observer l'effet de la permutation circulaire pour rendre le filtre causal. Dans le cas de la figure 3.8(b), la permutation n'a pas été faite, et on obtient un filtre non causal.



(a) Réponse impulsionnelle après permutation circulaire de $l/2$ (b) Réponse impulsionnelle sans permutation circulaire de $l/2$

FIGURE 3.8.: Réponse impulsionnelle du filtre inverse du canal 1 par inversion sans régularisation. Le délai initial de la réponse impulsionnelle du filtre direct a été enlevé dans le cas de cette figure.

Filtres à minimum de phase

Une autre approche pour trouver un filtre inverse stable et causal est de n'inverser que la partie à minimum de phase du filtre direct [Neely et Allen, 1979]. Pour trouver la partie à minimum de phase, on utilise le cepstre réel [Oppenheim et Schaffer, 1975, p. 503-504]. La fonction `rceps` de MATLAB® retourne le cepstre réel d'un signal mais également sa partie à minimum de phase. C'est ce qui nous intéresse ici, elle utilise une propriété du cepstre à cette

fin [Oppenheim et Schafer, 1975, p. 503-506] Une fois la partie minimum de phase obtenue, on inverse cette dernière dans le domaine fréquentiel suivant la formule suivante

$$H_{\min}[k] = \frac{G}{C_{\min}[k]} \quad (3.12)$$

On obtient alors un filtre inverse stable et causal, à minimum de phase. Étant donné que l'on applique par dessus un filtre passe bande (voir 3.3.3) le filtre inverse final n'est plus nécessairement à minimum de phase. On applique alors une seconde fois la fonction `rceps` pour obtenir un filtre inverse à minimum de phase.

Il est important de noter que le filtre à minimum de phase ne corrige que l'amplitude du filtre direct, mais pas sa phase [Norcross *et al.*, 2002].

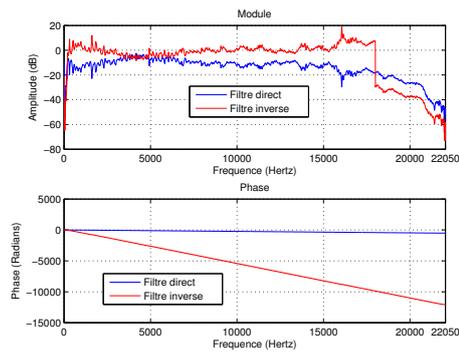
3.4 QUELQUES RÉSULTATS

3.4.1 *Filtres utilisés*

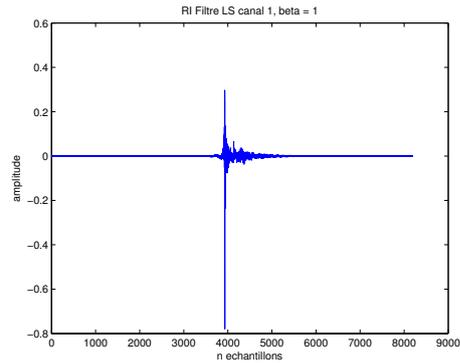
Pour notre installation, on teste 4 types de filtres :

- Inversion fréquentielle sans régularisation
- Inversion fréquentielle avec régularisation ($\beta = 1$). Par la suite, on note ce type de filtre "LS" (*Least Square*)
- Inversion de la partie minimum de phase
- Inversion fréquentielle sans régularisation après liftrage cepstral (fenêtrage de taille 512 échantillons)

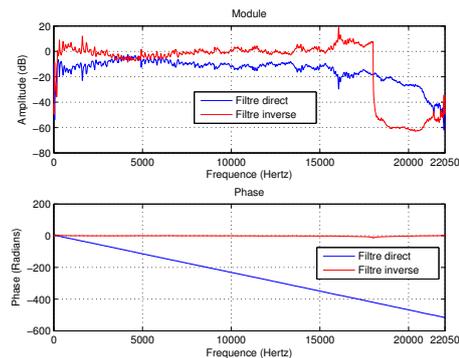
La longueur des filtres direct et inverse est de $l = 8192$ échantillons. Le nombre de points pour le calcul des TFD est également de $m = 8192$ échantillons. La fréquence d'échantillonnage est de $F_e = 44100\text{Hz}$.



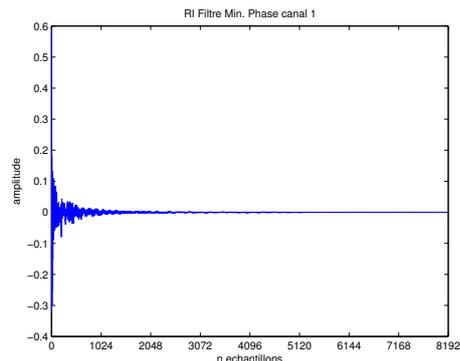
(a) Filtre LS, $\beta = 1$



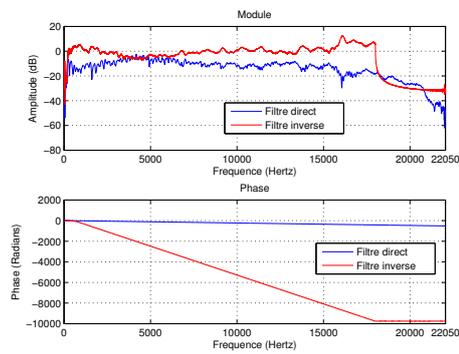
(b) Filtre LS, $\beta = 1$, RI



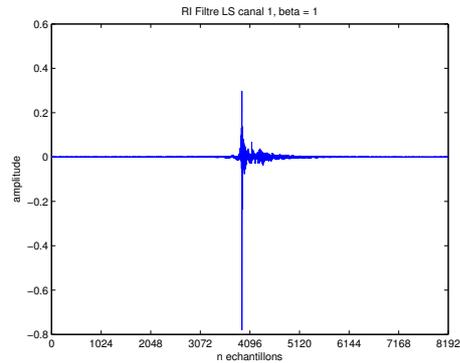
(c) Filtre Min. Phase



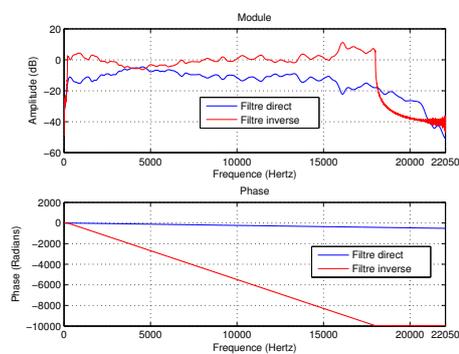
(d) Filtre Min. Phase, RI



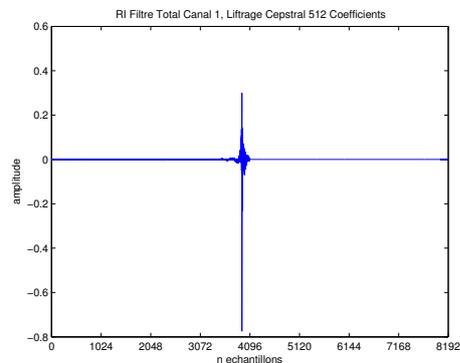
(e) Filtre sans régularisation



(f) Filtre sans régularisation, RI



(g) Filtre sans régularisation, liffrage à 512 coefficients



(h) Filtre sans régularisation, liffrage à 512 coefficients, RI

FIGURE 3.9.: Filtrés utilisés pour corriger le canal 1 au centre du dispositif

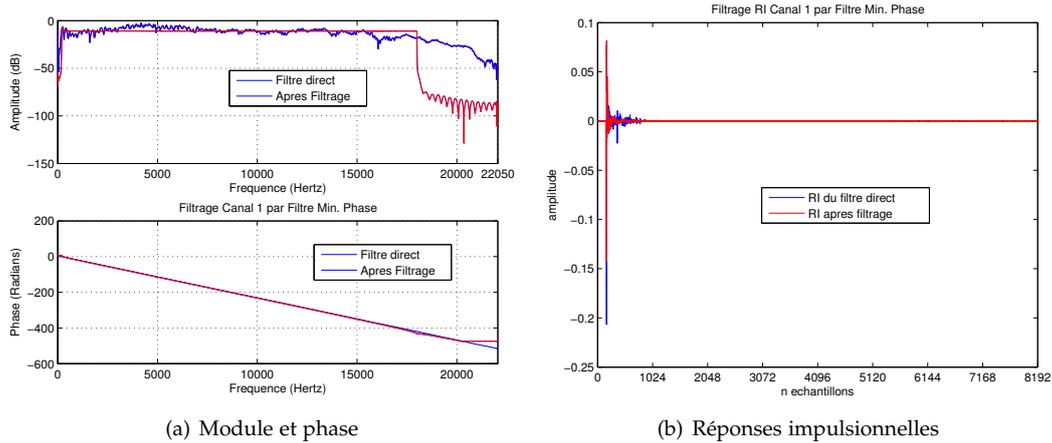


FIGURE 3.10.: Résultats de la déconvolution sous MATLAB © : en bleu la réponse avant filtrage et en rouge après déconvolution à l'aide d'un filtre inverse à minimum de phase.

Sur la figure 3.9 on peut voir la réponse en fréquence des 4 types de filtres (disponibles en ligne⁵) ainsi que leur réponse impulsionnelle pour le canal 1. Un filtre passe bande de bande passant [200Hz, 18000Hz] est appliqué dans tous les cas.

3.4.2 Déconvolution sous MATLAB ©

A titre d'exemple, on peut voir sur la figure 3.4.1 la simulation de la déconvolution à l'aide de la fonction `filter` de MATLAB ©. Le filtre direct $c[n]$ (disponible en ligne⁶) est filtré par le filtre inverse $h[n]$. Il s'agit ici du filtre inverse à minimum de phase qui corrige le canal 1. On remarque que la réponse en module et phase dans la bande passante et la réponse en phase est laissée inchangée (figure 3.10(a)). Au niveau des réponses impulsionnelles on remarque que le second pic correspondant à la réflexion sur la sol a disparu après déconvolution (figure 3.7(a)). Cependant, la réponse impulsionnelle après filtrage n'est pas un dirac retardé car un filtre passe bande est appliqué au filtre inverse. Ce filtre ne déconvolue alors pas le filtre direct sur toute la bande passante et en conséquence on n'obtient pas un dirac après déconvolution.

3.4.3 Déconvolution en temps réel sur le système de restitution

Mise en place en temps réel

A présent, on teste les 4 types de filtres présentés plus haut sur le système de restitution. Le filtrage est mis en place en temps réel à l'aide du logiciel BRUTEFIR (voir 5.6.1).

Pour chaque type, 15 filtres sont calculés pour corriger les 15 canaux au centre du dispositif. Lorsqu'un signal est envoyé à un seul canal, seul un filtre est actif. Par contre lorsqu'un signal est envoyé aux 15 canaux, ce qui est le cas lorsque l'on applique la NFC-HOA, les 15 filtres sont actifs en même temps.

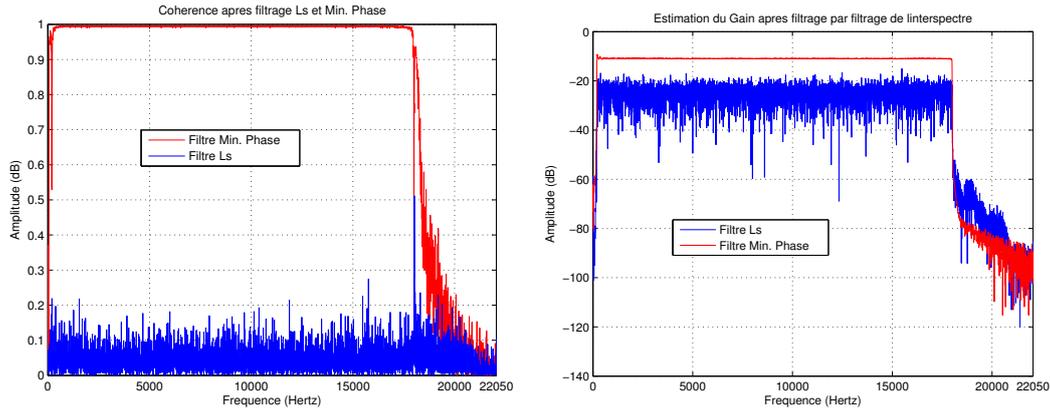
Mesure des fonctions de transfert

La mesure de la fonction de transfert à l'aide d'un bruit blanc et estimation spectrale de la fonction de transfert ne marche plus ici avec les paramètres utilisés ($m = 8192$, $df = \frac{44100}{8192}$), la cohérence (voir A) tombe en effet à 0 (!) sauf dans le cas minimum de phase. Pour s'en convaincre, observons la figure 3.11. Sous MATLAB ©, on filtre un bruit blanc d'une durée de 10 sec `BB10Sec.wav` (disponible en ligne⁷) avec respectivement le filtre inverse LS et le filtre à

5. <http://www.sekisushai.net/ANFCHOA/CHAP03/>, fichiers `x1_*.mat`

6. <http://www.sekisushai.net/ANFCHOA/CHAP03/RI1.mat>

7. <http://www.sekisushai.net/ANFCHOA/CHAP03/BB10Sec.wav>



(a) Cohérence obtenue avec un filtre inverse à minimum de phase et un autre LS

(b) Module de la fonction de transfert estimée

FIGURE 3.11.: Cohérence et module de la fonction de transfert estimée avec un filtre à minimum de phase (rouge) et un filtre LS (bleu)

minimum de phase pour le canal 1. On utilise la fonction `filter`. On obtient alors deux bruits filtrés `BB10SecFiltreLS.wav`⁸ et `BB10SecFiltreMin.wav`⁹. On filtre ensuite ces deux fichiers par le filtre direct du canal 1¹⁰. Finalement on obtient les deux fichiers `BB10SecMesLS.wav`¹¹ et `BB10SecMesMin.wav`¹². La fonction de transfert est alors estimée dans le domaine spectral suivant la méthode de l'annexe A entre les fichiers `BB10SecMesLS.wav` et `BB10SecMesMin.wav` avec `BB10Sec.wav`. On utilise le code `GainComplexe.m`¹³.

On remarque que la méthode d'estimation ne marche que pour le filtre inverse à minimum de phase (voir figure 3.11). Les autres types de filtres utilisés, (de même qu'un filtre passe bas très simple, réalisé à partir de la fonction `fir1` de MATLAB ©) font tomber la cohérence entre signal original et filtré à 0!

Après quelques tests rapides, il s'avère que le choix de la longueur des signaux pour l'estimation, avant moyennage dans le domaine spectral semble déterminant. De même que la longueur du filtre inverse (plus il est court, meilleur est la cohérence). Ceci reste à approfondir.

Pour obtenir des résultats rapidement au cours du stage, on fait alors le choix de changer de méthode d'estimation et d'utiliser la méthode proposée par Farina dans [Farina, 2000]. Cette méthode utilise un sinus glissant exponentiel, voir l'annexe B pour la méthode.

Résultats de mesures

Les résultats des mesures sont disponibles en ligne¹⁴ sous forme de fichier `*.mat`. Il s'agit des gains complexes entre les 143 microphones de l'antenne (voir 5.7.2) et le signal électrique envoyé. Les résultats sans filtrage sont également disponibles en ligne¹⁵. Le programme `MATHEMATICA @VisuFiltre1Seul.nb`¹⁶ permet de visualiser les résultats pour les quatre types de déconvolution. 2 sources synthétisées ont été mesurées : l'une à l'intérieur du cercle de haut-parleur $r_0 = 0.8\text{m}, \theta_0 = 228^\circ$; l'autre à l'extérieur $r_0 = 5\text{m}, \theta_0 = 324^\circ$. Dans le dossier `SansFiltres` une troisième source est mesurée $r_0 = 100\text{m}, \theta_0 = 0^\circ$ ainsi que la réponse de tous les haut-parleurs sur l'antenne : `H_1_143.mat` par exemple pour le premier haut-parleur. La mesure de la réponse des haut-parleur a été faite avec un bruit blanc, suivant la méthode de l'annexe A. Les mesures des gains complexes après filtrage sont réalisées à partir d'un sinus glissant (voir annexe B). Ce

8. <http://www.sekisushai.net/ANFCHOA/CHAP03/BB10SecFiltreLS.wav>

9. <http://www.sekisushai.net/ANFCHOA/CHAP03/BB10SecFiltreMin.wav>

10. <http://www.sekisushai.net/ANFCHOA/CHAP03/RI1.mat>

11. <http://www.sekisushai.net/ANFCHOA/CHAP03/BB10SecMesLS.wav>

12. <http://www.sekisushai.net/ANFCHOA/CHAP03/BB10SecMesMin.wav>

13. <http://www.sekisushai.net/ANFCHOA/CHAP03/GainComplexe.m>

14. <http://www.sekisushai.net/ANFCHOA/CHAP04/> dans les sous dossiers `LS`, `Min`, `Tot` et `Tot512`

15. <http://www.sekisushai.net/ANFCHOA/CHAP04/SansFiltrage/>

16. <http://www.sekisushai.net/ANFCHOA/CHAP03/VisuFiltre1Seul.nb>

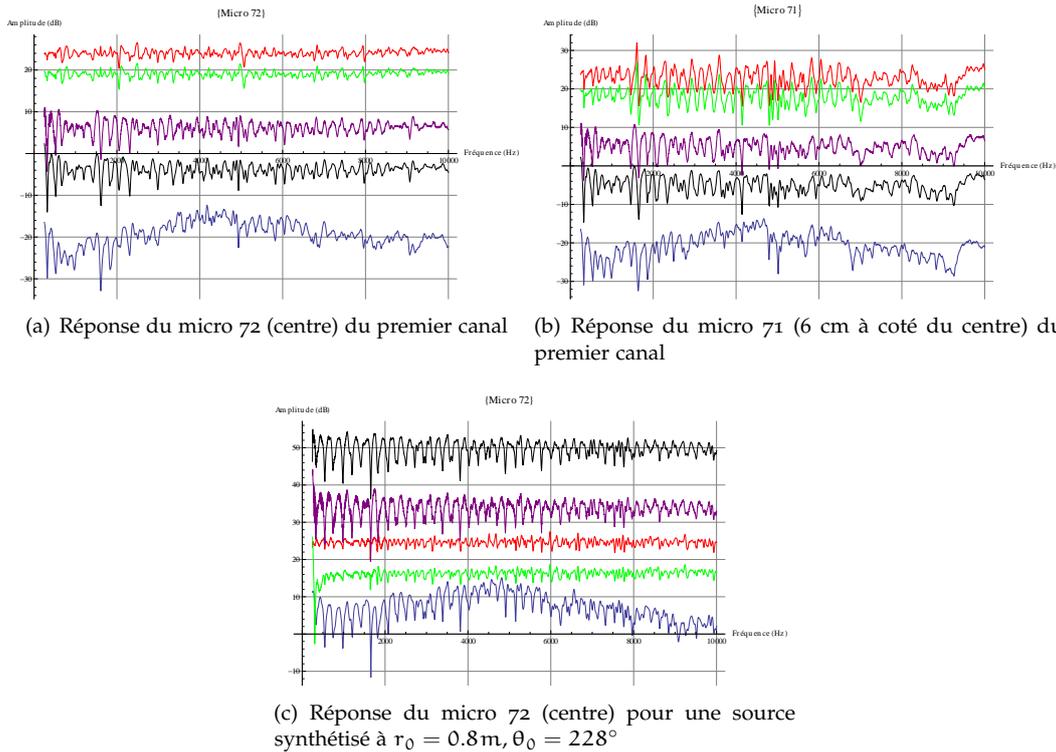


FIGURE 3.12.: Réponse des microphones 72 (centre) et 71 (6 cm à côté) pour les différents filtres. Rouge : Filtre LS, Vert : Filtre Min. Phase, Violet : Filtre sans régularisation, Noir : Filtre sans régularisation avec liftrage 512 coefficients, Bleu : Sans filtre. Les courbes sont décalés verticalement pour une meilleure observation.

changement soudain de méthode d'estimation se justifie à cause des problèmes relatés dans la section 3.4.3. Les gains complexes sont calculés sur une longueur de $m = 8192$ point à une fréquence d'échantillonnage de $F_e = 44100\text{Hz}$.

3.5 CONCLUSIONS

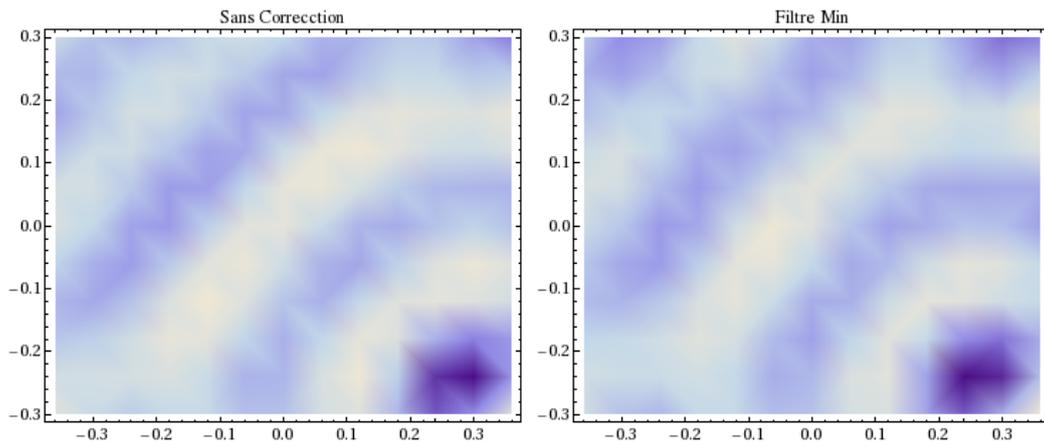
La première observation montre que le résultat de déconvolution sous MATLAB ©(figure 3.4.1) n'est qu'approché sur les mesures. En effet, seuls les filtres LS et à minimum de phase semblent s'approcher d'une réponse plate (figures 3.12(a) et 3.12(c)) à $\pm 2\text{dB}$. Le filtrage en peigne n'est cependant que partiellement atténué.

Pour les filtres par inversion sans régularisation, avec ou sans liftrage cepstral, le résultat semble à peu près équivalent. Le filtrage en peigne n'est pas atténué mais la réponse direct est "redressée". Pour obtenir une différence plus notable, il faudrait peut être liftrer la réponse direct de manière plus importante.

Si l'on s'éloigne du point de correction de 6cm en regardant la réponse du micro 71 (voir sur la disposition de l'antenne 5.7(b)) on remarque que la correction n'est plus valable pour aucun des filtres (figure 3.12(b)). On ne retrouve pas pour autant la réponse direct, sans filtrage. Ainsi on ne contrôle pas ce qui se passe en dehors du point de correction !

La mesure des réponses sans filtrage et avec filtrage n'a pas été réalisée le même jour et par la même technique. Ce changement de conditions contribue à la baisse de performance. En effet, la température, la disposition de l'antenne, etc. font varier sensiblement les réponses observées. Un algorithme de filtrage adaptatif pourrait être envisagé.

En utilisant la NFC-HOA par dessus ces filtres (figure 3.12(c)), les 15 filtres des 15 haut-parleurs sont actifs en même temps et leur réponses sont alors additionnées. Les réponses de chacun d'eux étant "plates" au centre, on obtient toujours une réponse "plate". Encore une fois, seuls les filtres LS et à minimum de phase semblent convenir.



(a) Champ sonore d'une source synthétisée à $R_O = 0.8\text{m}$, $\theta_0 = 228^\circ$, $f = 1340\text{Hz}$ (partie réelle). Sans aucun filtrage
 (b) Champ sonore d'une source synthétisée à $R_O = 0.8\text{m}$, $\theta_0 = 228^\circ$, $f = 1340\text{Hz}$ (partie réelle). Filtrage au micro 72 (centre) avec le filtre à minimum de phase

FIGURE 3.13.: Champs sonore d'une source synthétisée à $R_O = 0.8\text{m}$, $\theta_0 = 228^\circ$ sans filtrage et avec un filtrage à minimum de phase sur le micro 72 (centre)

Les cartes des champs sonores restent cohérentes par rapport à la situation sans filtrage (figure 3.13). Cependant le principal inconvénient est que l'on contrôle la réponse des canaux en un seul point seulement. En dehors de ce point, la correction n'est plus valable. Ainsi, cette technique de déconvolution est abandonnée au profit d'une déconvolution multicanale, ce qui fait l'objet du chapitre suivant.

CORRECTION MULTICANALE DU SYSTÈME DE RESTITUTION SUR UN ENSEMBLE DE POINT DE CONTRÔLE

4.1 INTRODUCTION

À l'instar de la correction canal par canal, on cherche ici à corriger le système sur un ensemble de points de contrôle choisis dans la zone d'écoute. L'objectif est alors de contrôler l'acoustique de la salle de restitution sur une zone à l'aide de tous les haut-parleurs disponibles.

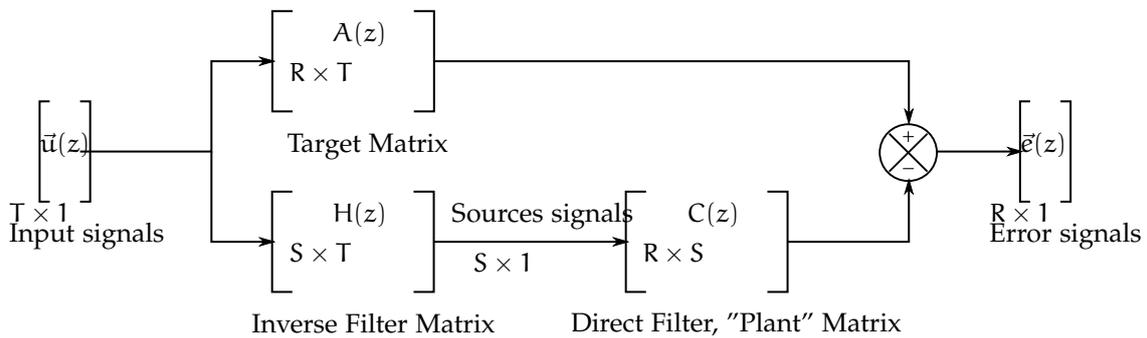


FIGURE 4.1.: Problème d'égalisation multicanale

Le principe est illustré sur la figure 4.1. Le système fonctionnant en temps discret, on utilise la notation en transformée en z [Blanchet et Charbit, 2001, p. 106]. On s'aide de tous les canaux de restitution disponibles pour corriger l'un d'entre eux. La correction vise à obtenir un champ sonore "cible" sur les points de contrôle. La cible est définie par la matrice cible $A(z)$. Avec S canaux de restitution, on conçoit S filtres qui corrigent la réponse d'un canal sur un ensemble de points de contrôle. Si l'on veut corriger T canaux de restitution, on cherche au final une matrice $H(z)$ de $T \times S$ filtres de correction.

Pour mieux comprendre, considérons dans un premier temps un seul canal de restitution à corriger à l'aide des S canaux ($T = 1$). Le signal $u(z)$ est envoyé à ce canal. Après propagation dans la salle de restitution on obtient $\vec{R}_{\text{mes}}(z)$ signaux mesurés sur l'ensemble des R micros. Parallèlement, le signal est envoyé au vecteur de filtres $A(z)$ de dimensions $R \times 1$. Ce vecteur de filtres peut par exemple contenir le modèle de propagation d'un monopôle en champ libre, le monopôle étant placé au même endroit que le haut-parleur du canal à corriger. On obtient alors un vecteur de signaux cibles $\vec{R}_{\text{cible}}(z)$: la réponse d'un monopôle en champ libre sur les R micros d'après la configuration géométrique du dispositif. Finalement, on calcule l'erreur $\vec{e}(z) = \vec{R}_{\text{cible}}(z) - \vec{R}_{\text{mes}}(z)$.

On suppose connu la réponse des S sources utilisés pour la correction sur les R microphones de contrôle, contenue dans la matrice $C(z)$. L'objectif est alors de construire un vecteur de filtre inverses $H(z)$ de dimensions $S \times 1$ pour minimiser $\vec{e}(z)$. Autrement dit, on cherche $H(z)$ tel que :

$$C(z)H(z) = A(z) \quad (4.1)$$

ou encore

$$H(z) = C^{-1}(z)A(z) \quad (4.2)$$

Ainsi, le signal $u(z)$ est filtré par le vecteur de filtres $H(z)$ et on obtient alors S signaux de commandes, envoyés aux S canaux utilisés pour la correction. Ces signaux sont ensuite propagés dans la salle de restitution suivant la réponse de chacun des S canaux comme définit dans la matrice C . On obtient alors les $\vec{R}_{\text{mes}}(z)$ signaux mesurés, correspondant à la superposition de chacun des signaux issus des S canaux.

En extrapolant ce résultat pour d'autres sources à corriger ($T \geq 1$), on construit la matrice cible \mathbf{A} où chacune des colonnes contient le vecteur cible pour le canal i ($i = 1 \dots T$). De même la matrice de filtre inverse \mathbf{H} est de dimensions $S \times T$ et chacune des colonnes correspond au vecteur de filtres qui corrigent le canal i .

Dans la suite du rapport on fixe $S = T$ c'est à dire qu'on utilise les S sources disponibles pour corriger $T = S$ sources. Une augmentation du nombre de source $S \geq T$ utiles à la correction serait envisageable par la suite.

4.1.1 Rapide État de l'art

Pour résoudre le problème de déconvolution multicanale présenté précédemment, on remarque d'après l'équation (4.2) que la difficulté réside dans l'inversion de la matrice \mathbf{C} . Cette inversion doit se faire en respectant plusieurs critères sur les filtres \mathbf{H} obtenus. En effet, pour implémenter de manière pratique ces filtres, ils se doivent d'être stable et causaux. A ces fins, il existe de nombreux algorithmes utilisant un approche temporelle ou fréquentielle.

Théorème "MINT"

Miyoshi et al. introduisent les premiers dans [Miyoshi et Kaneda, 1988] un théorème "*Multi Input/Output Inverse Theorem, MINT*". Ce théorème stipule qu'il est possible de trouver la matrice de filtres RIF qui déconvolue l'effet de salle sur un ensemble de point.

Quelques algorithmes

Dans [Kirkeby et Nelson, 1999] ou [Miyoshi et Kaneda, 1988] on propose une approche temporelle pour trouver la matrice de filtre inverse. Cependant la taille des matrices mise en jeu devient considérable lorsque les filtres sont d'ordres élevés. Ces approches deviennent très coûteuse dans le cadre de la conception de long filtres.

Pour pallier ce problème, Guillaume [Guillaume *et al.*, 2005] propose un algorithme itératif pour calculer la matrice de filtres inverse.

Kikerby et al. [Kirkeby *et al.*, 1998] proposent une méthode de déconvolution par transformée de Fourier rapide (FFT). C'est cette approche que l'on retient dans le cadre du stage.

Approche par holophonie

Comme il a été mentionné dans l'introduction générale du problème de déconvolution (voir 3.1), plusieurs approches sont proposées pour diminuer l'effet de salle dans le cadre de la WFS. On peut citer par exemple [Corteel, 2006], [Spors *et al.*, 2003], [Stefanakis *et al.*, 2010], [Gauthier et Berry, 2007]. Ces approches tentent de minimiser un fonction coût qui contient l'information de localisation de la source à synthétiser. Ainsi on oblige la solution obtenue à ne pas trop s'éloigner de la théorie WFS. Plus de détails seront donnés par la suite (voir 4.1.5).

4.1.2 Contexte du stage

Dans notre cas on cherche à ce que le système de restitution s'approche au mieux de la théorie de la NFC-HOA. Ainsi l'objectif est d'avoir des sources de restitution qui s'approchent au mieux de monopôles en champ libre. On corrige alors à la fois l'effet de salle et la réponse des haut-parleurs.

Une fois la matrice \mathbf{H} obtenue, les signaux de NFC-HOA sont simplement filtrés par cette matrice \mathbf{H} avant d'être envoyés au système de restitution. On gagne alors de ne pas devoir connaître au préalable la source à synthétiser, contrairement à la plupart des approches actuelles de l'holophonie adaptative. En effet, ces dernières doivent obtenir un filtre pour chaque position de la source à synthétiser.

Dans notre cas on effectue un unique étalonnage. Par exemple pour l'ordre 7 et $N = 15$ haut-parleurs, on conçoit la matrice de correction \mathbf{H} de taille 15×15 filtres. Cette phase d'étalonnage

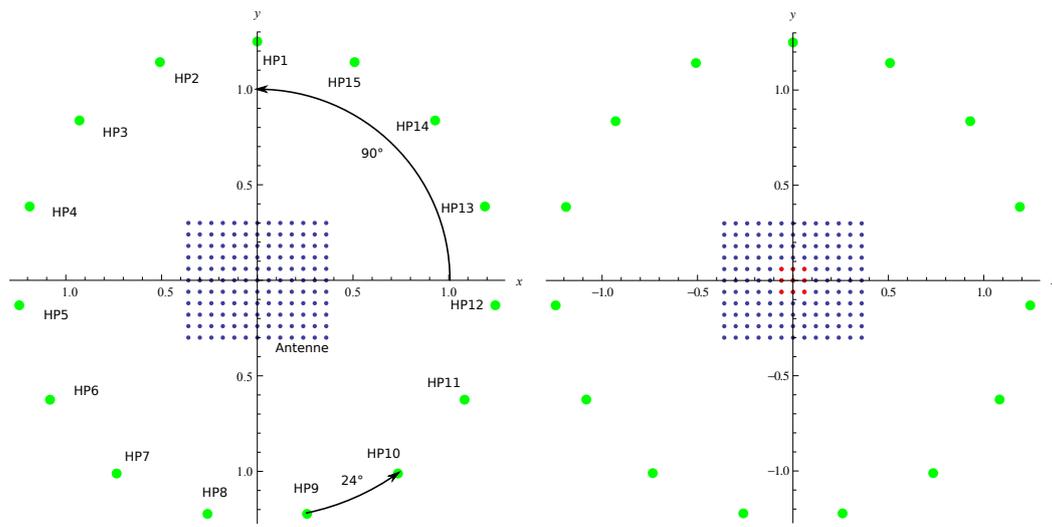
terminée, on peut alors synthétiser une source à n'importe quelle endroit selon la NFC-HOA classique.

4.1.3 Algorithmes utilisés dans le cadre du stage

Dans le cadre du stage, deux types d'inversion ont été implémentées. La première est une inversion dans le domaine fréquentiel en utilisant la régularisation comme décrit dans [Tokuno *et al.*, 1997] ou [Kirkeby *et al.*, 1998]. La seconde technique utilise une fonction coût inspirée de l'AWFS proposé par [Gauthier et Berry, 2007] et adaptée à l'ambisonie.

Configuration géométrique du système à corriger

La configuration géométrique du système de restitution est présentée figure 4.2(a). Pour plus de détails voir la section 5.7. La mesure des réponses impulsionnelle des $N = 15$ haut-parleurs



(a) Vue de dessus du dispositif de restitution. En vert les haut-parleurs. En bleu l'antenne de mesure. Les dis- tances sont en mètres. L'écart angulaire est de 24° entre deux haut-parleurs. Le premier haut-parleur est aux coordonées cartésiennes $HP_1(0, 1.25)$.

FIGURE 4.2.: Configuration géométrique du système et de l'antenne

est réalisée sur toute l'antenne à l'aide de la technique présentée 3.2. On choisit ensuite le nombre de points à corriger sur cette antenne. Dans le cas de la figure 4.2(b), on cherche à corriger $R = 9$ positions sur l'antenne.

Matrice cible

On cherche à trouver la matrice de filtre \mathbf{H} qui permet d'approcher au mieux la matrice cible suivant la formule (4.1). Dans notre cas la matrice cible contient la réponse de monopôles placés aux différents position des $T = 15$ haut-parleurs sur la grille de correction (voir figure 4.2(b)).

Pour rappel, la réponse au point M d'un monopôle de débit U en champ libre, placé au point O , tel que $|\vec{OM}| = r$ est donnée par [Chaigne *et al.*, 2008, p.555] :

$$p(kr) = j\omega\rho U G_\infty(kr) \quad (4.3)$$

où $G_\infty(r)$ est la fonction de Green en champ libre

$$G_\infty(kr) = \frac{e^{-jkr}}{4\pi r} \quad (4.4)$$

Dans notre cas on fixe le débit du monopôle unitaire en fonction de la fréquence, ainsi

$$j\omega\rho U = 1 \quad \forall\omega \quad (4.5)$$

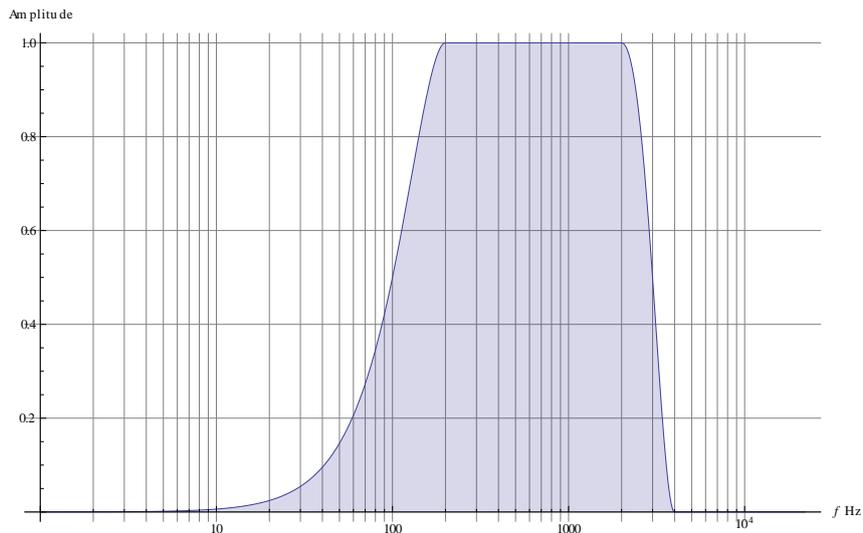
La matrice \mathbf{A} est de dimensions $R \times T \times \frac{N_{fft}}{2} + 1$. En effet, comme on le verra dans les sections suivantes, l'inversion se fait fréquence par fréquence, le nombre de fréquence étant de $\frac{N_{fft}}{2} + 1$.

Filtre passe bande par fenêtrage dans le domaine spectral

Pour les mêmes raisons qu'évoquées dans la section 3.3.3, on souhaite appliquer un filtre passe bande. Cependant, on choisit une fréquence de coupure supérieure beaucoup plus basse que dans le cas d'un seul point de correction. En effet, la taille de la zone d'écoute est très petite en haute fréquence. Qui plus est l'antenne de mesure à une résolution spatiale qui ne permet pas d'observer finement le champ en haute fréquence. Enfin, sa diffraction ne devient plus négligeable (voir 5.7). Pour ces raisons on a choisi dans un premier temps de filtrer les signaux en dehors de la bande passante $f_{min} = 200\text{Hz}$, $f_{max} = 2000\text{Hz}$. À l'instar des techniques présentées dans le cas de la correction d'un seul point, on choisit ici d'appliquer ce filtre passe bande directement dans la matrice cible. Ainsi la réponse du monopôle en champ libre est fenêtrée dans le domaine spectral pour être atténuée en dehors de la bande passante. La fenêtre utilisée est inspirée d'une fenêtre de Tukey [Harris, 1978, p. 66-67], mais elle n'est pas symétrique. Son équation est donnée ci-après :

$$w(f) = \begin{cases} \frac{1}{2} \left(1 + \cos\left(\frac{\pi}{f_{min}}(f - f_{min})\right)\right) & f \leq f_{min} \\ 1 & f_{min} \leq f \leq f_{max} \\ \frac{1}{2} \left(1 + \cos\left(\frac{\pi}{1 - f_{max}}(f - f_{max})\right)\right) & f_{max} \leq f \leq 2f_{max} \\ 0 & f \geq 2f_{max} \end{cases} \quad (4.6)$$

Un exemple d'allure de cette fenêtre est visible figure 4.1.3. Pour cet exemple $f_{min} = 200\text{Hz}$, $f_{max} = 2000\text{Hz}$.



Finalement, le terme générique de \mathbf{A} vaut donc :

$$\mathbf{A}_{i,j,f} = p(kr_{i,j}) \times w(f) \quad (4.7)$$

où $k = \frac{2\pi f}{c}$ est le nombre d'onde et c la vitesse du son ($c \approx 340\text{m/s}$).

4.1.4 Inversion fréquentielle utilisant la régularisation

Pour trouver la matrice de filtres inverse \mathbf{H} on effectue l'inversion fréquence par fréquence. En faisant varier la fréquence f on crée alors le gabarit fréquentiel des filtres inverses.

De la même manière qu'en 3.3.3 La fonction coût à minimiser est définie par :

$$J(f) = \bar{\mathbf{e}}^*(f)\bar{\mathbf{e}}(f) + \beta\bar{\mathbf{v}}^*(f)\bar{\mathbf{v}}(f) \quad (4.8)$$

où $\bar{\mathbf{e}}(f)$ représente le vecteur d'erreur et $\bar{\mathbf{v}}$ le terme d'effort proportionnel à l'énergie des sources [Kirkeby *et al.*, 1998]. * désigne l'opérateur transposé-conjugué. Le paramètre β fixe alors le compromis entre exactitude de la solution au sens minimisation de l'erreur quadratique et l'énergie de la solution par le terme d'effort.

La solution de cette minimisation fonctionnelle est donnée par [Kirkeby *et al.*, 1998] :

$$\mathbf{H}(f) = [\mathbf{C}^*(f)\mathbf{C}(f) + \beta\mathbf{I}]^{-1}\mathbf{C}^*(f)\mathbf{A}(f) \quad (4.9)$$

où la matrice \mathbf{I} est la matrice identité.

Cette formulation n'ajoute pas de contrainte sur l'énergie des filtres de chacune des sources. Ceci peut être un problème dans le cadre de notre utilisation : Par exemple si les $S = 15$ haut-parleurs corrigent le premier haut-parleur, rien ici n'oblige le premier haut-parleur à avoir le terme d'effort le plus important. Ainsi si l'on envoie un signal sur le canal 1, les $S = 15$ canaux vont filtrer ce signal. Au niveau de la zone d'écoute, la localisation perçue ne sera pas forcément en direction du haut-parleur numéro 1.

La section suivante présente une formulation qui tente de pallier ce problème. Dans le cas de notre exemple, l'inversion va contraindre à ce que ce soit le canal 1 qui émette le plus fort lorsque c'est lui que l'on corrige.

4.1.5 Inversion fréquentielle utilisant le formalisme issu de l'AWFS adapté à l'ambisonie : ANFC-HOA

La WFS ou l'ambisonie permettent de contrôler le champ sonore sur une zone d'écoute. L'idée est alors d'utiliser cette théorie pour superviser la correction multicanale en rajoutant une contrainte issue de la WFS (ou de l'ambisonie). La fonction coût à minimiser devient alors [Gauthier et Berry, 2007] :

$$J_{AWFS} = \bar{\mathbf{e}}^*\bar{\mathbf{e}} + \beta[\bar{\mathbf{v}} - \bar{\mathbf{v}}_{WFS}]^*[\bar{\mathbf{v}} - \bar{\mathbf{v}}_{WFS}] \quad (4.10)$$

et dans le cas de la NFC-HOA la fonction devient simplement :

$$J_{ANFC-HOA} = \bar{\mathbf{e}}^*\bar{\mathbf{e}} + \beta[\bar{\mathbf{v}} - \bar{\mathbf{g}}_{NFC-HOA}]^*[\bar{\mathbf{v}} - \bar{\mathbf{g}}_{NFC-HOA}] \quad (4.11)$$

Ainsi, on contraint l'énergie de la solution à ne pas trop s'écarter de la solution donnée par la NFC-HOA (le vecteur $\bar{\mathbf{g}}_{NFC-HOA}$ représentent les gains ambisoniques issus de la NFC-HOA).

En s'inspirant de [Kirkeby *et al.*, 1998] et [Gauthier et Berry, 2007], on obtient à une fréquence f fixée, la solution de la minimisation :

$$\mathbf{H}_{ANFC-HOA} = [\mathbf{C}^*\mathbf{C} + \beta\mathbf{I}]^{-1}[\mathbf{C}^*\mathbf{A} + \beta\mathbf{G}_{NFC-HOA}] \quad (4.12)$$

Dans notre cas, on ne corrige que les $T = 15$ sources de restitution, ainsi les gains $\bar{\mathbf{g}}_{NFC-HOA}$ doivent être calculés pour les 15 positions de ces 15 sources, suivant la figure 4.2(a). Comme on le verra dans le chapitre suivant (voir 5.2), on a choisi $N = 2M + 1$ sources de restitution. La dimension d'un vecteur $\bar{\mathbf{g}}_{NFC-HOA}$ est alors de $2M + 1 \times 1$ et dans ce cas particulier, les gains NFC-HOA sont données par :

$$\begin{cases} \mathbf{g}_{HP1}^T = [1 & 0 & \dots & 0] \\ \mathbf{g}_{HP2}^T = [0 & 1 & \dots & 0] \\ \dots \\ \mathbf{g}_{HP15}^T = [0 & 0 & \dots & 1] \end{cases} \quad (4.13)$$

Finalement la matrice $\mathbf{G}_{NFC-HOA}$ est une matrice identité de dimensions $S \times T$.

4.1.6 Choix du paramètre de régularisation β

A présent, il convient de choisir une valeur de β . Suivant les formulations (4.12) et (4.9), on peut trouver un β optimal pour chaque fréquence. Des méthodes comme la validation croisée généralisée (GCV) ou la courbe en L permettent de trouver la solution qui assure le meilleur compromis entre terme d'effort et performance [?].

Dans notre cas, il faut également considérer les effets temporels dus au choix de β . En effet, ce paramètre de régularisation contrôle l'énergie des filtres inverses, mais aussi la durée de leur réponse impulsionnelle. Dans [Kirkeby et Nelson, 1999] on montre pour un filtre que cela revient à régler la distance entre le cercle unité et les pôles dans sa transformée en z . On se sert alors de cette propriété pour trouver une valeur optimale de β . Cette méthode n'est hélas valable que dans le cas d'une inversion d'un seul canal en un seul point de contrôle.

Dans notre cas, on utilise alors un β unique sur toute la gamme de fréquence, et l'on observe les réponses impulsionnelles obtenues pour que leur énergie soit approximativement contenue dans la moitié de la longueur du filtre $l/2$, comme il est conseillé dans [Kirkeby et Nelson, 1999].

On s'assure alors que les filtres inverses décroissent assez rapidement pour considérer les effets de repliements temporels négligeables. Pour rappel, ces effets sont dûs à l'utilisation d'une méthode basée sur un gabarit fréquentiel discrétisé, et l'utilisation de TFD.

Cette méthode d'essai-erreur reste cependant à améliorer.

4.1.7 Introduction d'un délai

Après retour dans le domaine temporel, on introduit un délai pour rendre les filtres causaux. Ceci est fait à l'aide d'une permutation circulaire de la même manière qu'en 3.3.3

4.2 UN EXEMPLE EN SIMULATION

Pour illustrer l'approche de correction présentée ci-dessus, on traite un exemple en simulation. Ces simulations sont réalisées à partir de la feuille de calcul MATHEMATICA ©CorrectionSol.nb disponible en ligne¹.

On se place dans une salle semi-anéchoïque (sol réfléchissant), comme c'est le cas pour le système de restitution du stage. Un "filtrage en peigne" est alors réalisé, provenant des interférences avec les ondes réfléchies sur le sol.

L'ordre de décomposition NFC-HOA est fixé à $M = 7$ et $N = 2M + 1 = 15$ sources de restitution monopôlaires sont utilisées. La configuration est la même que pour le dispositif du stage. Ainsi le rayon du cercle de haut-parleurs est $r_{HP} = 1.25m$, la hauteur des haut-parleurs par rapport au $h_{HP} = 1.25m$, et les haut-parleurs sont équirépartis tous les 24° suivant la figure 4.2(a)

4.2.1 Correction de l'effet de sol

On cherche à corriger l'effet du sol sur une grille de correction en suivant les deux approches présentées en 4.1.5 et 4.1.4. Il est utile de se référer à la figure 4.1 pour bien comprendre la démarche.

Sur la figure 4.3(b) on peut observer l'effet du filtrage en peigne sur le champ sonore pour le haut-parleur 3. La zone où l'erreur relative, suivant la formule (2.31), est inférieure à 20% est délimitée en trait noir. Le champ original en champ libre est montré figure 4.3(a). La réponse au point d'origine (0,0) est montrée figure 4.3(c)

La simulation est effectuée à **une** seule fréquence, "défavorable", de $f = 1218Hz$. A cette fréquence, le filtrage en peigne atténue fortement la réponse au point d'origine du repère (figure 4.3(c))

L'objectif est de corriger $R = 9$ points de correction disposés en carré autour de l'origine, comme montré sur la figure 4.2(b).

1. <http://www.sekisushai.net/ANFCHOA/CHAP04/CorrectionSol.nb>

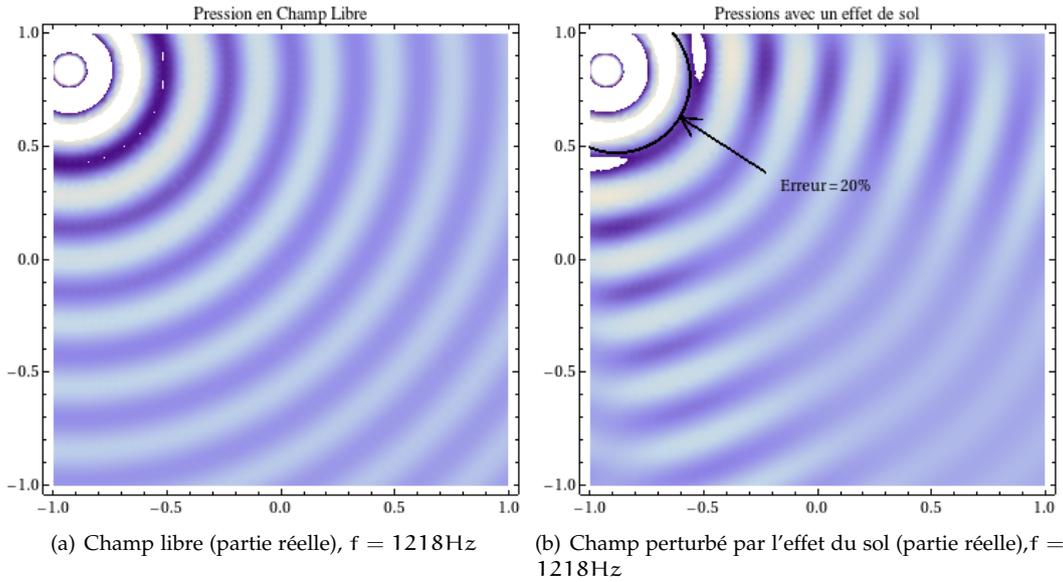


FIGURE 4.3.: Effet du filtrage en peigne par la réflexion sur le sol. Les dimensions des cartes sont en mètres.

La matrice \mathbf{A} est de taille 9×15 et contient alors la réponse **en champ libre** des $T = 15$ hauts-parleurs sur les R microphones. La matrice \mathbf{C} est de taille 9×15 et contient la réponse **perturbée par l'effet du sol** des $S = 15$ hauts-parleurs.

On calcule les matrices de filtre inverse \mathbf{H} et $\mathbf{H}_{\text{ANFC-HOA}}$ suivant les formules (4.9) et (4.12) respectivement. Ici $\beta = 10^{-8}$ est pour les deux inversions. Les matrices obtenues sont de tailles 15×15 .

Les gains issus de la NFC-HOA sont ensuite filtrés par les matrices \mathbf{H} et $\mathbf{H}_{\text{ANFC-HOA}}$:

$$\begin{cases} \vec{Z} = \mathbf{H}\vec{S} \\ \vec{Z}_{\text{ANFC-HOA}} = \mathbf{H}_{\text{ANFC-HOA}}\vec{S} \end{cases} \quad (4.14)$$

Par exemple, les gains obtenus pour corriger le haut-parleur 3 sont visibles sur la figure 4.4. On remarque que la solution ANFC-HOA (courbe en bleu) favorise plus le haut-parleur 3 pour la correction. En effet, la formulation impose une solution du problème proche de la solution NFC-HOA pour une source placée sur le haut-parleur 3.

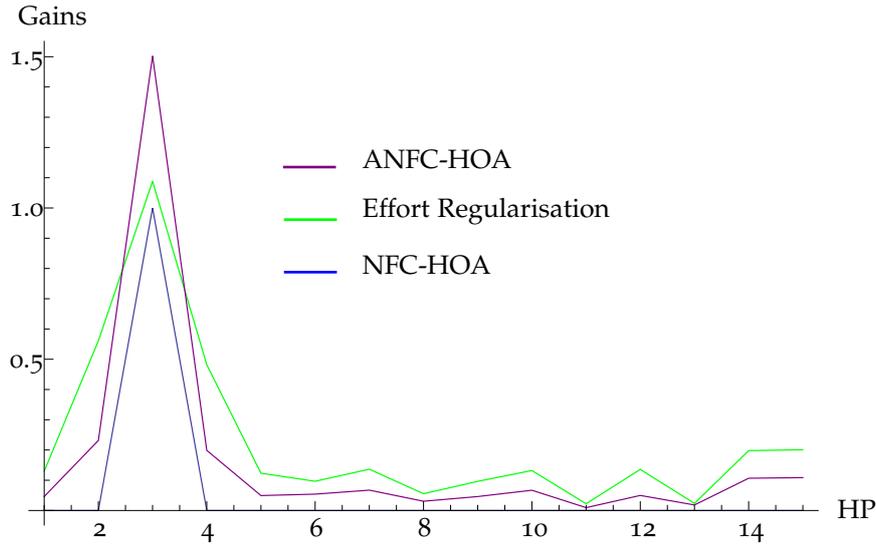
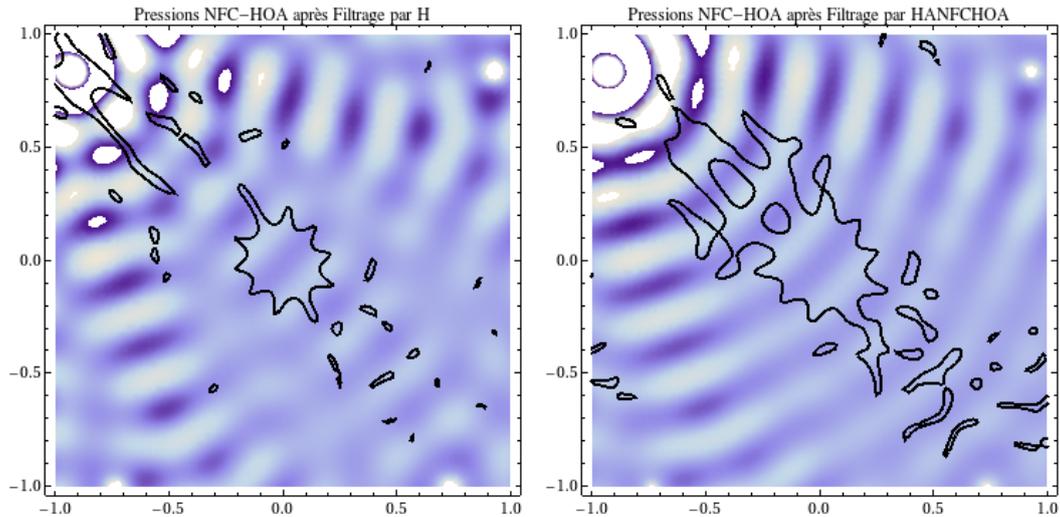


FIGURE 4.4.: Gains appliqués aux différents haut-parleurs pour corriger le haut-parleur 3

Finalement les champs obtenus après corrections du haut-parleur 3 sont visible sur la figure 4.5. On remarque que la solution ANFCHOA étend légèrement la zone par rapport à la solution qui utilise la régularisation classique.



(a) Champ obtenu après filtrage par inversion fréquentielle avec régularisation (partie réelle), $f = 1218\text{Hz}$
 (b) Champ obtenu après filtrage par ANFC-HOA (partie réelle), $f = 1218\text{Hz}$

FIGURE 4.5.: Champ obtenu après filtrage par inversion fréquentielle avec régularisation et par ANFC-HOA (partie réelle). Les dimensions sont en mètres. L'erreur $e = 20\%$ par rapport à la source en champ libre est en trait noir.

4.2.2 Source synthétisée quelconque

Les matrices \mathbf{H} et $\mathbf{H}_{\text{ANFC-HOA}}$ contiennent les coefficients pour la déconvolution des $T = 15$ haut-parleurs sur la grille de $R = 9$ points. En gardant ces matrices, on filtre les coefficients NFC-HOA pour une source synthétisée quelconque. On peut observer sur les figures 4.6 les résultats obtenus pour une source à synthétiser avec $R_s = 0.8\text{m}$, $\theta_s = 228^\circ$. Le paramètre de régularisation est de $\beta = 10^{-8}$ pour les deux types d'inversion. On observe dans cet exemple que l'ANFC-HOA étend légèrement la zone de correction par rapport à l'inversion standard.

Cependant il est important de noter que tout se joue au niveau du paramètre de régularisation β qui est à choisir avec une grande précaution. Sur ces figures est tracé en rouge pointillé la taille de la zone d'écoute dans le cas NFC-HOA en champ libre. L'ANFC-HOA tend alors vers cette zone lorsque la correction est efficace.

4.3 QUELQUES RÉSULTATS SUR LE DISPOSITIF DE RESTITUTION

L'ANFC-HOA ainsi que la déconvolution par inversion et régularisation affichent des résultats prometteurs en simulation. Dans cette section, on présente quelques mesures qui ont été effectuées sur le dispositif de restitution monté pour le stage (voir chapitre 5).

8 matrices de filtres ont été calculées pour faire les mesures. Les deux premières corrigent le dispositif sur le micro du centre avec l'ANFC-HOA et par inversion-régularisation. Les deux suivantes sur 9 micros, puis 25 micros et enfin 143 micros correspondant à toute l'antenne de mesure. Le paramètre de régularisation β a été choisi arbitrairement pour que la partie utile des filtres inverses soit approximativement contenue dans la moitié de leur longueur (voir 4.1.6). Ainsi les valeurs respectives de β sont $\beta_1 = 0.0001$, $\beta_9 = 0.001$, $\beta_{25} = 0.1$ et $\beta_{143} = 2$.

4.3.1 Conception des filtres

Le gabarit fréquentiel des filtres est conçu sous MATHEMATICA © avec la feuille de calcul `FiltresMultiChannel.nb` disponible en ligne². Ce programme permet de visualiser la déconvolution fréquence par fréquence et la réponse des microphones obtenues. La matrice qui contient les réponses des haut-parleurs sur l'antenne de mesure est disponible en ligne³ sous forme d'un fichier `*.mat`. Sur les figures 4.7 et 4.8 un exemple est montré pour la conception des filtres qui corrigent sur 9 points. On observe que la réponse est globalement bien corrigée lorsqu'on est dans la zone de correction et se dégrade dès que l'on en sort (figure 4.8). Il est assez difficile d'obtenir de meilleures performances avec l'ANFC-HOA plutôt qu'avec l'inversion et régularisation. Le choix d'un bon paramètre β est primordial. Cependant, des considérations de traitement du signal sont à prendre en compte 4.1.6. Un compromis doit être fait à ce niveau et la méthode d'un choix par essai-erreur reste à améliorer !

4.3.2 Résultats issus de la mesure

Les résultats des mesures sont disponibles en ligne⁴ sur forme de fichier `*.mat`. Le programme MATHEMATICA © `VisuFiltres.nb`⁵ permet de visualiser les résultats pour les deux types de déconvolution. 2 sources synthétisées ont été mesurées : l'une à l'intérieur du cercle de haut-parleur $r_0 = 0.8\text{m}, \theta_0 = 228^\circ$; l'autre à l'extérieur $r_0 = 5\text{m}, \theta_0 = 324^\circ$. Dans le dossier `SansFiltres` une troisième source est mesurée $r_0 = 100\text{m}, \theta_0 = 0^\circ$ ainsi que la réponse de tous les haut-parleurs sur l'antenne : `H_1_143.mat` par exemple pour le premier haut-parleur.

La mesure de la réponse des haut-parleur a été faite avec un bruit blanc, suivant la méthode de l'annexe A. Les mesures des gains complexes après filtrage sont réalisées à partir d'un sinus glissant (voir annexe B). Ce changement soudain de méthode d'estimation se justifie à cause des problèmes relatés dans la section 3.4.3. Les gains complexes sont calculés sur une longueur de $m = 8192$ point à une fréquence d'échantillonnage de $F_e = 44100\text{Hz}$.

Un exemple de résultat est montré sur les figures 4.9 pour une source à $r_0 = 0.8\text{m}, \theta_0 = 228^\circ$. La partie réelle des champs obtenus est tracée pour $f \approx 1340\text{Hz}$. Suivant les situations, la déconvolution semble mieux marcher pour le filtrage par inversion et régularisation que l'ANFC-HOA. Encore une fois, le paramètre de régularisation β est primordial sur le résultat obtenu. En augmentant le nombre de points de correction, β augmente mais la performance globale baisse.

2. <http://www.sekisushai.net/ANFCHOA/CHAP04/FiltresMultiChannel.nb>

3. <http://www.sekisushai.net/ANFCHOA/CHAP04/PlantTot.mat>

4. <http://www.sekisushai.net/ANFCHOA/CHAP04/> dans les sous dossiers `Filtres1 Filtres9 Filtres25 et Filtres143 et SansFiltres`

5. <http://www.sekisushai.net/ANFCHOA/CHAP04/VisuFiltres.nb>

Un autre exemple de mesure pour une source à $r_0 = 5\text{m}$, $\theta_0 = 324^\circ$ est visible figure 4.10. La partie réelle des champs obtenus est tracé pour $f \approx 1071\text{Hz}$. On montre ici les résultats des filtrages pour une correction au centre du dispositif (micro 72 sur l'antenne). La réponse à ce microphone est visible sur la figure 4.10(f). Les courbes sont décalées pour une meilleure observation. On remarque que la déconvolution par inversion et régularisation marche bien ici pour la réponse au micro 72. En effet, ici les 15 haut-parleurs servent à corriger seulement un point. On se rapproche alors de la situation évoquées dans [Miyoshi et Kaneda, 1988] : le filtrage inverse est quasi-exact. Cependant en regardant l'allure du champ obtenu sur la figure 4.10(c), on remarque que la correction au centre se fait au détriment de la NFC-HOA. La correction par ANFC-HOA devient alors beaucoup plus performante (figure 4.10(d)) sur l'allure du champ. Cependant, elle l'est moins sur la réponse au micro 72 4.10(f).

4.3.3 Notes sur le critère d'erreur

On ne trace pas ici un critère d'erreur relative par rapport à la source théorique en champ libre. Effectivement, les données de mesures sont bruitées et les champ théoriques et mesurés ne sont pas en phase dans nos mesures. Ceci s'explique notamment par la façon dont les réponses impulsionnelles corrigées sont estimées (voir annexe B).

Par ailleurs, l'acquisition des réponses impulsionnelles et des réponses impulsionnelles corrigées n'ont pas été faite le même jour. Alors, la température et les conditions de mesures varient sensiblement, ce qui a pour effet de baisser également la performance.

Enfin, la position des haut-parleurs dans la salle et les 4 positions de l'antenne de mesure (voir 5.7) varient à une précision de l'ordre de quelques cm. Ainsi le champ théorique sur une grille de points mathématiquement défini semble un choix de comparaison un peu trop dur.

Pour le moment, seul un aperçu de l'allure du champ et de la réponse aux microphones permet d'estimer la qualité de la déconvolution. Bien évidemment ceci reste à améliorer, par exemple en comparant par rapport à la salle simulée ou bien par rapport à une vraie source placée physiquement au lieu virtuel de la source synthétisée.

4.4 CONCLUSIONS

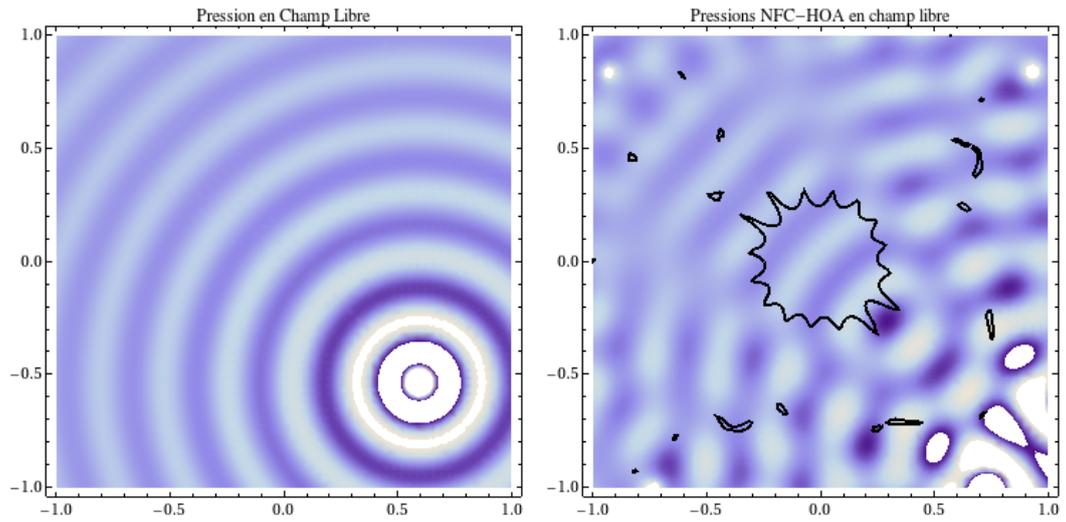
Les méthodes de déconvolution multicanales présentées ci-dessus semblent prometteuses pour réaliser une synthèse de champ sonore en dehors d'une chambre anéchoïque (champ libre).

L'ANFC-HOA est basée sur l'adaptation de l'AWFS pour l'ambisonie. Cette approche donne une information supplémentaire de provenance de la source pour la conception des filtres de correction. À l'audition on entend effectivement une différence. Les filtres de déconvolution par inversion et régularisation peuvent parfois dégrader fortement l'impression de localisation de la source. En effet, aucune direction n'est privilégiée dans la formulation.

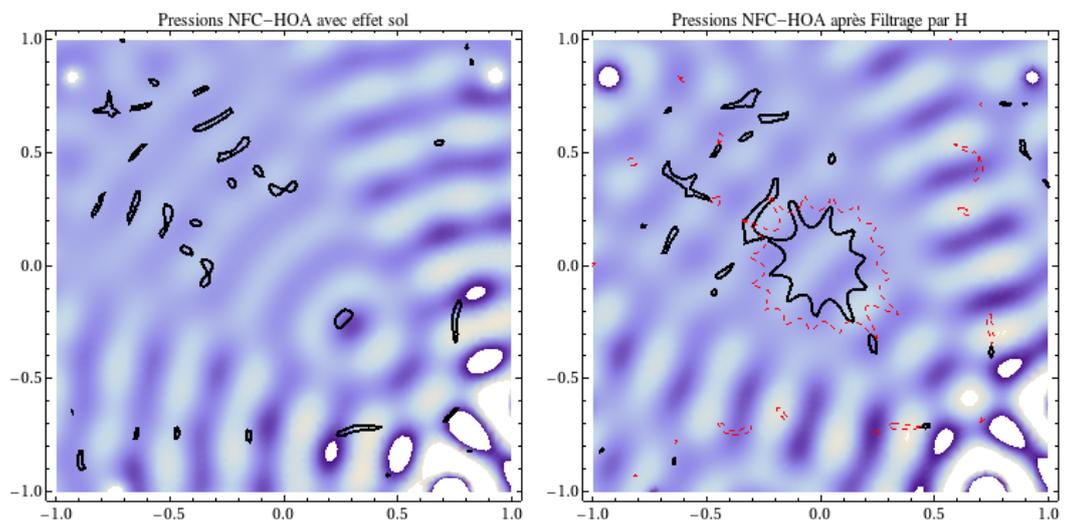
Le paramètre de régularisation est à choisir avec beaucoup de précaution. En effet, il faut à la fois qu'il soit optimal en terme de reconstruction physique (minimisation de l'erreur), d'énergie (terme d'effort), mais aussi d'un point de vue traitement du signal pour l'allure des filtres inverse. Des améliorations sont à faire par rapport à la méthode actuelle d'essai-erreur.

La prochaine étape de cette démarche est de passer le système de restitution dans une vraie salle. On s'attend à ce que les améliorations dues à la correction soient plus "visibles" car les champ mesurés sans correction devraient être perturbés de manière significative par rapport à la chambre semi-anéchoïque.

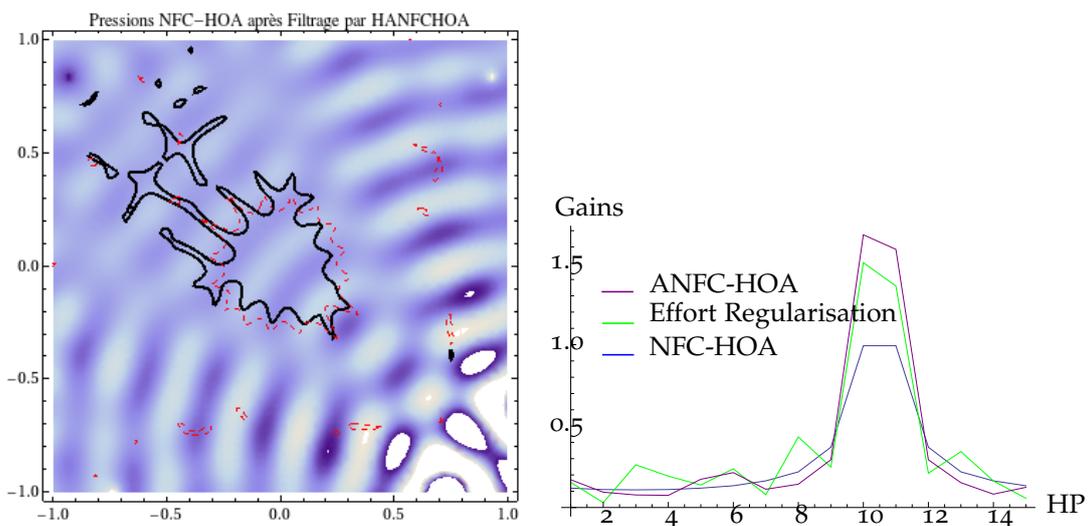
Une éventuelle amélioration seraient de rajouter des sources qui aideraient alors la correction : on ne corrigerait plus 15 haut-parleurs par 15 haut-parleurs, mais par 20 ou plus. En effet, pour le moment on réalise une "auto-corrrection" du système. Le fait de rajouter des haut-parleurs supplémentaires permettrait une meilleure correction, des premières simulations montrent en effet une amélioration dans ce sens. Par contre, le nombre et la position de ces haut-parleurs supplémentaires restent difficiles à préciser à l'heure actuelle. Ce travail est à poursuivre.



(a) Source en champ libre (partie réelle), $f = 1218\text{Hz}$ (b) Champ obtenu par NFC-HOA en champ libre (partie réelle), $f = 1218\text{Hz}$



(c) Champ obtenu par NFC-HOA perturbé par l'effet de sol (partie réelle), $f = 1218\text{Hz}$ (d) Champ obtenu après filtrage par inversion fréquentielle avec régularisation (partie réelle), $f = 1218\text{Hz}$. En rouge pointillé l'erreur à 20% pour la NFC-HOA en champ libre



(e) Champ obtenu après filtrage par ANFC-HOA (partie réelle), $f = 1218\text{Hz}$. En rouge pointillé l'erreur à 20% synthèse de la source, $f = 1218\text{Hz}$ pour la NFC-HOA en champ libre (f) Gains appliqués aux différents haut-parleurs pour la

FIGURE 4.6.: Synthèse adaptative d'une source à $R_s = 0.8\text{m}$, $\theta_s = 228^\circ$

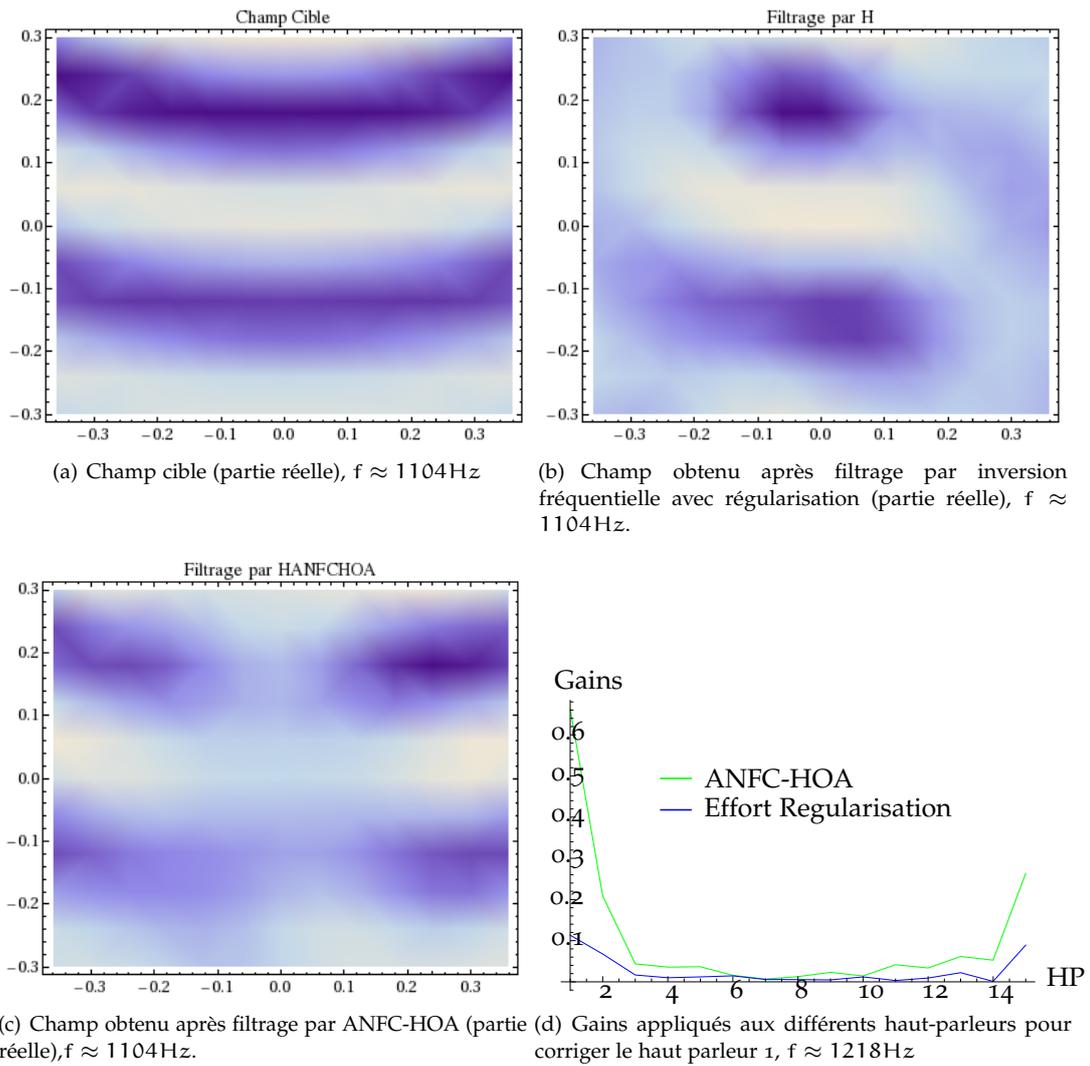


FIGURE 4.7.: Exemple pour la conception des filtres, ici le haut-parleur 1 d'après la mesure

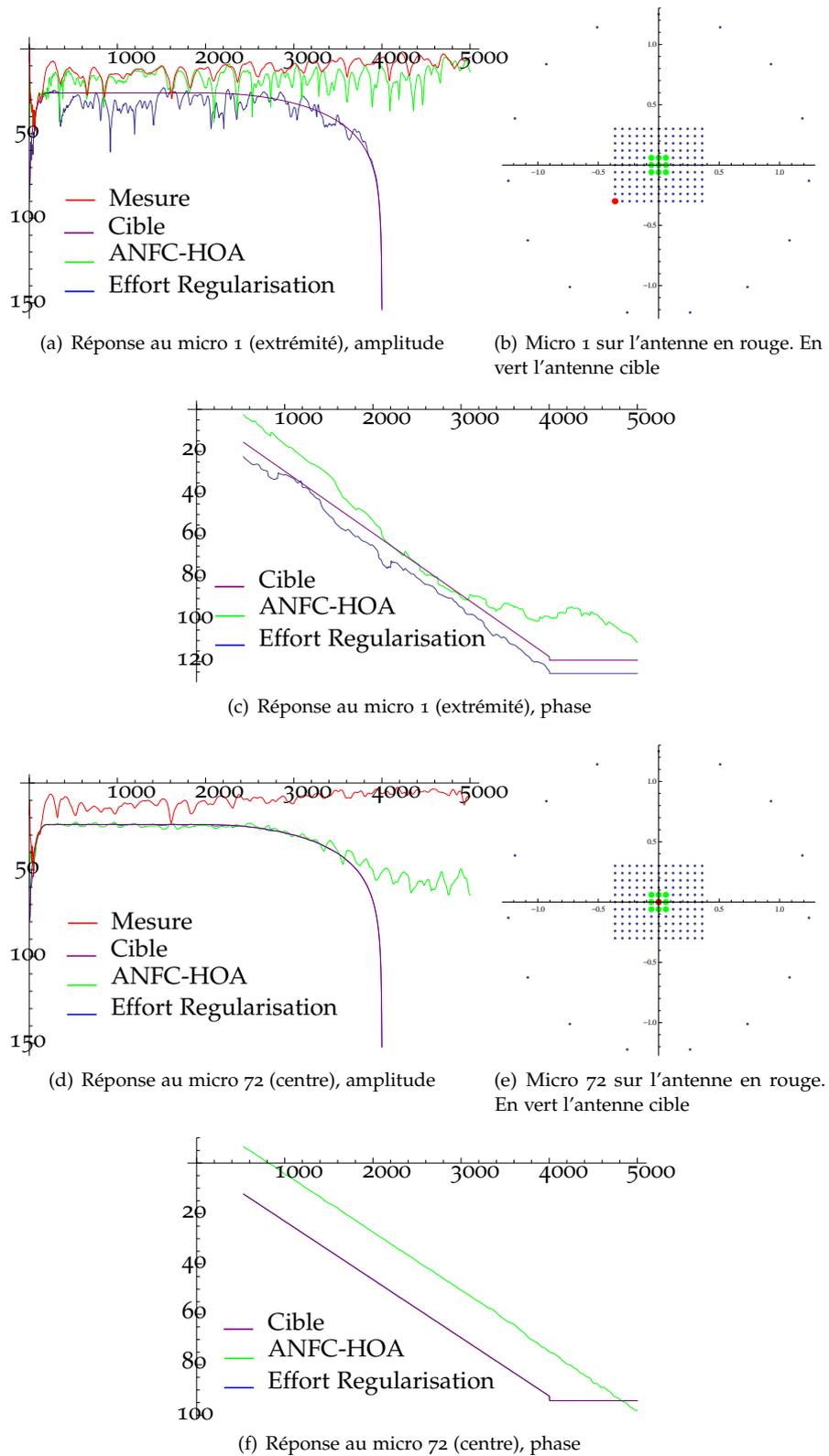
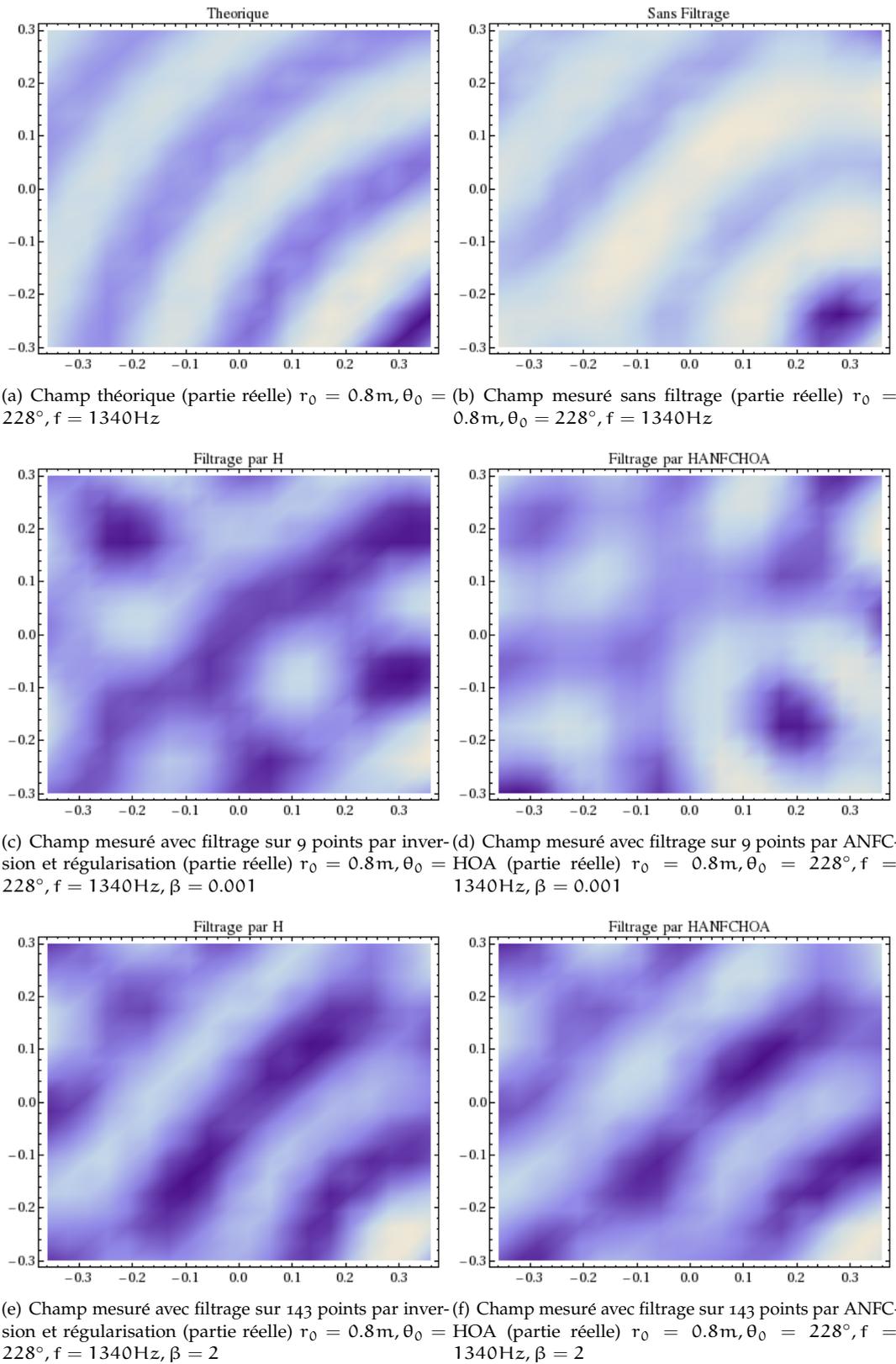
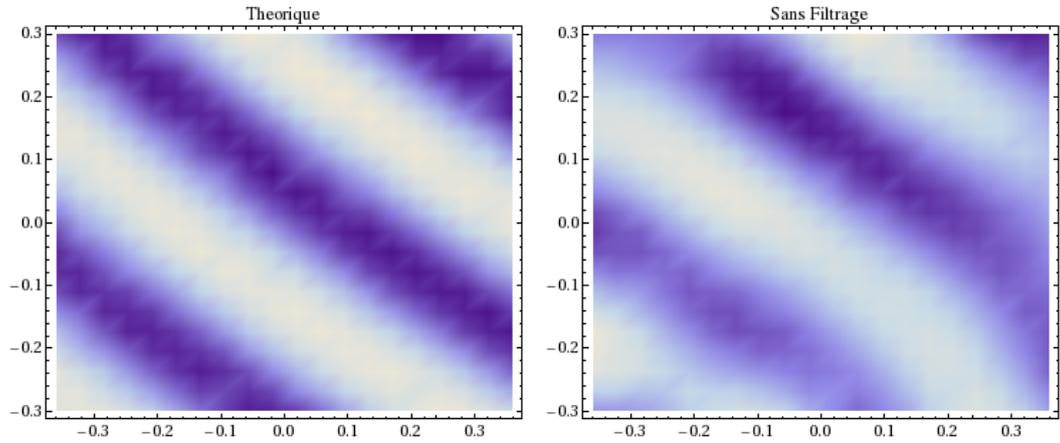
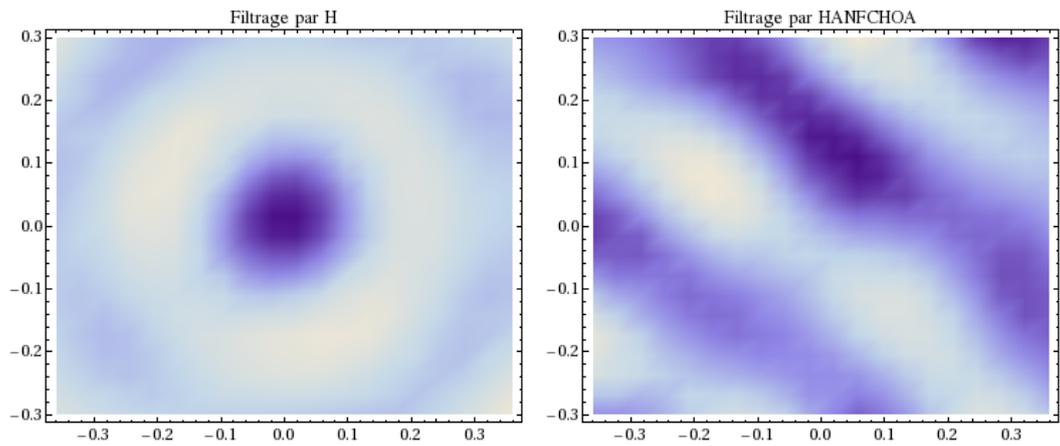


FIGURE 4.8.: Réponse à différents microphones micro 1 hors de la zone de correction (a) et micro 72 dans la zone de correction (d)

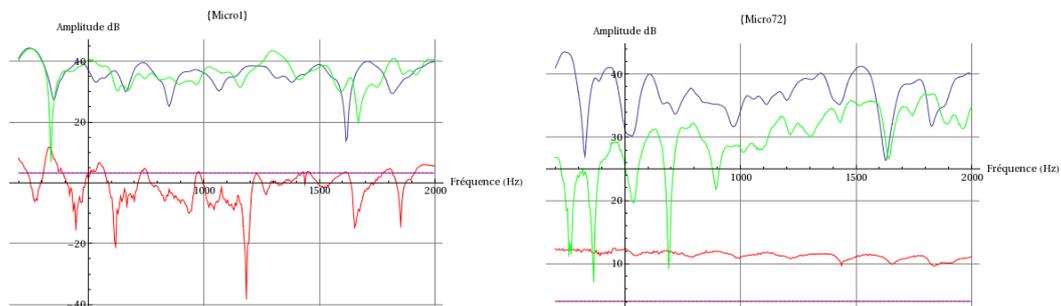
FIGURE 4.9.: Mesure d'une source synthétisée à $r_0 = 0.8\text{m}, \theta_0 = 228^\circ, f = 1340\text{Hz}$



(a) Champ théorique (partie réelle) $r_0 = 5\text{ m}$, $\theta_0 = 324^\circ$, $f = 1071\text{ Hz}$ (b) Champ mesuré sans filtrage (partie réelle) $r_0 = 5\text{ m}$, $\theta_0 = 324^\circ$, $f = 1071\text{ Hz}$



(c) Champ mesuré avec filtrage sur 1 point par inversion et régularisation (partie réelle) $r_0 = 5\text{ m}$, $\theta_0 = 324^\circ$, $f = 1071\text{ Hz}$, $\beta = 0.0001$ (d) Champ mesuré avec filtrage sur 1 point par ANFCHOA (partie réelle) $r_0 = 5\text{ m}$, $\theta_0 = 324^\circ$, $f = 1071\text{ Hz}$, $\beta = 0.0001$



(e) Réponse au micro 1 (extrémité) avec filtrage en un point (centre). Les courbes sont décalées selon l'axe vertical (centre). Les courbes sont décalées selon l'axe vertical cal pour une meilleure observation. Violet : cible, Rouge : Filtrage par inversion et régularisation, Vert : Filtrage ANFC-HOA, Bleu : Sans Filtrage (f) Réponse au micro 72 (centre) avec filtrage en un point (centre). Les courbes sont décalées selon l'axe vertical cal pour une meilleure observation. Violet : cible, Rouge : Filtrage par inversion et régularisation, Vert : Filtrage ANFC-HOA, Bleu : Sans Filtrage

FIGURE 4.10.: Mesure d'une source synthétisée à $r_0 = 5\text{ m}$, $\theta_0 = 324^\circ$, $f = 1071\text{ Hz}$

RÉALISATION PRATIQUE D'UN SYSTÈME DE SYNTHÈSE DE CHAMP SONORE 2D TEMPS-RÉEL PAR ANFC-HOA

5.1 INTRODUCTION

L'objectif premier du stage est de mener une implémentation pratique d'un système de restitution 2D de ANFC-HOA. Ce système est dans un premier temps utilisé pour recréer physiquement des champs acoustiques à des fins d'utilisation pour la mesure. Le décodage en hautes fréquences, utilisant des critères psychoacoustiques de type vecteur vitesse et énergie (voir 2.8), n'est pas encore implémenté à ce jour. Au commencement du stage, le système de restitution n'existait pas ! Plusieurs haut-parleurs, cartes son, des moyens de mesures et de calcul étaient à disposition. Ce chapitre présente l'implémentation pas à pas du système en présentant tous les éléments et les choix qui ont été faits.

5.2 CHOIX DE L'ORDRE DE RESTITUTION

Compte tenu des moyens matériels existants actuellement au laboratoire, une implémentation 2D a été retenue pour obtenir rapidement des résultats dans la durée du stage. De même, pour une implémentation en temps réel on choisit une équirépartition des haut-parleurs en cercle. En effet, cette équirépartition évite d'inverser la matrice "U" (voir (2.22)), les gains étant donnés par la formule (2.29). La salle de restitution est la chambre semi-anéchoïque (sol réfléchissant) du LMSSC. Ses dimensions permettent de répartir les haut-parleurs de restitution sur un cercle de rayon $r_{HP} = 1.25\text{m}$. Le choix de l'ordre de restitution est un compromis entre le nombre de haut-parleur disponibles, leur équirépartition et la puissance de calcul de l'ordinateur de commande qui calcule les signaux. Ainsi le choix s'est fait pour une restitution à l'ordre $M = 7$, mettant en jeu $N = 2M + 1 = 15$ haut-parleurs équirépartis tous les 24° . Un rapide calcul donne en effet que $360^\circ/15 = 24^\circ$

5.3 DISPOSITION DES ENCEINTES DE RESTITUTION

Les haut-parleurs utilisés sont des Aura Sound NSW2-326-8A©¹. Pour éviter le court-circuit acoustique, Les haut-parleurs sont encastrés dans un tube PVC fermé et rempli de mousse formant ainsi une mini-enceinte de restitution (voir figure 5.1(a)). La mousse sert à amortir les modes du tube PVC. Ces minis enceintes sont ensuite fixées sur un pied de micro réglage en hauteur (voir figure 5.1(b)). La hauteur est réglée à $h = 1.25\text{m}$, ce qui correspond environ à la hauteur du plan des oreilles d'un auditeur assis sur une chaise.

Le positionnement des enceintes se fait à l'aide d'un plateau tournant gradué sur lequel est fixé une perche de longueur 1.25m (figure 5.2(a)). Un laser fixe dans la salle vient affleurer le coin du plateau tournant (figure 5.2(d)). On tourne ce dernier jusqu'à ramener la graduation sur le point du laser (figure 5.2(c)). Ceci permet alors d'équirépartir les enceintes tous les 24° sur un cercle de rayon $r_{HP} = 1.25\text{m}$, en faisant affleurer le haut-parleur à l'extrémité de la perche. Le centre du plateau tournant est à l'aplomb de l'origine du repère marquée au sol. On utilise un fil à plomb et un niveau pour régler le dispositif (figure 5.2(b)).

1. Fiche technique disponible en ligne : <http://www.sekisushai.net/ANFCHOA/CHAP05/NSW2-326-8A.pdf>

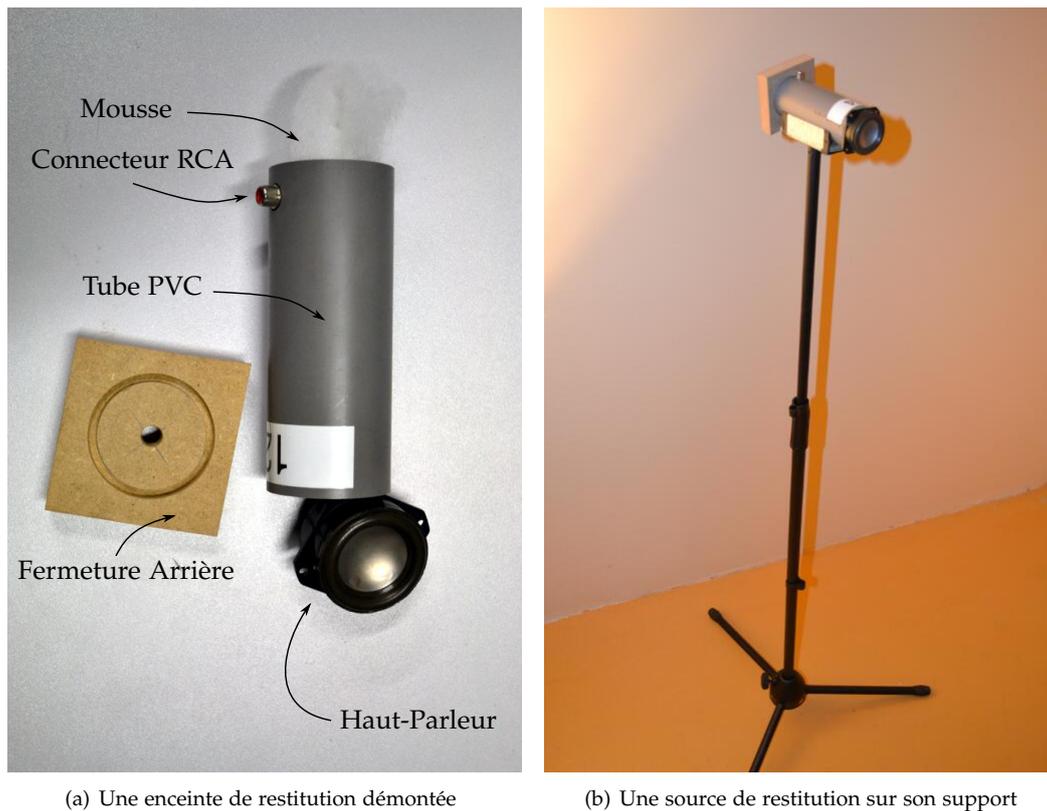


FIGURE 5.1.: Sources de restitution

5.4 TRANSMISSION DES SIGNAUX AUX ENCEINTES

Les haut-parleurs reçoivent leurs signaux depuis 2 cartes son firewire AudioFire 12 ©² chaînées, dont chacune des sorties est soudée à un amplificateur. Le gain de l'amplificateur est réglé à son maximum. Les connecteurs utilisés sont de type RCA (figure 5.3)

5.5 INTERFACE DE COMMANDE DU SYSTÈME DE RESTITUTION

La commande du système de restitution se fait en temps réel sous un ordinateur opérant sous Linux KUBUNTU 12.10³. Un noyau basse latence est utilisé : `linux-lowlatency-3.8*`. Les cartes sons et tous les logiciels utilisés sont pilotés par le serveur de son JACK⁴. La suite des logiciels de la distribution KXStudio⁵ est utilisé pour commander JACK et interconnecter les différents logiciels entre eux. Le logiciel CARLA permet de démarrer et contrôler les paramètres de JACK tel la latence, fréquence d'échantillonnage de la carte, etc. Le logiciel CLAUDIA permet de gérer les sessions et d'effectuer les connections entre les logiciels. Sur la figure 5.4 un exemple des connections entre les différents éléments sous CLAUDIA est montré.

5.5.1 Programme Pure Data

Le calcul des gains des différents haut-parleurs est fait en temps réel à partir d'un patch PURE DATA développé au laboratoire⁶. Ce patch permet de calculer les signaux NFC-HOA 2D jusqu'à l'ordre 7 pour une source définie par un rayon et un angle. Sur la figure 5.5 est présenté l'interface

2. Documentation technique disponible en ligne : <http://www.sekisushai.net/ANFCHOA/CHAP05/Audiofire.pdf>

3. <http://www.kubuntu.org>

4. <http://jackaudio.org/>

5. <http://kxstudio.sourceforge.net/>

6. Patch disponible en ligne : <http://www.sekisushai.net/ANFCHOA/CHAP05/NFCHOA.pd>

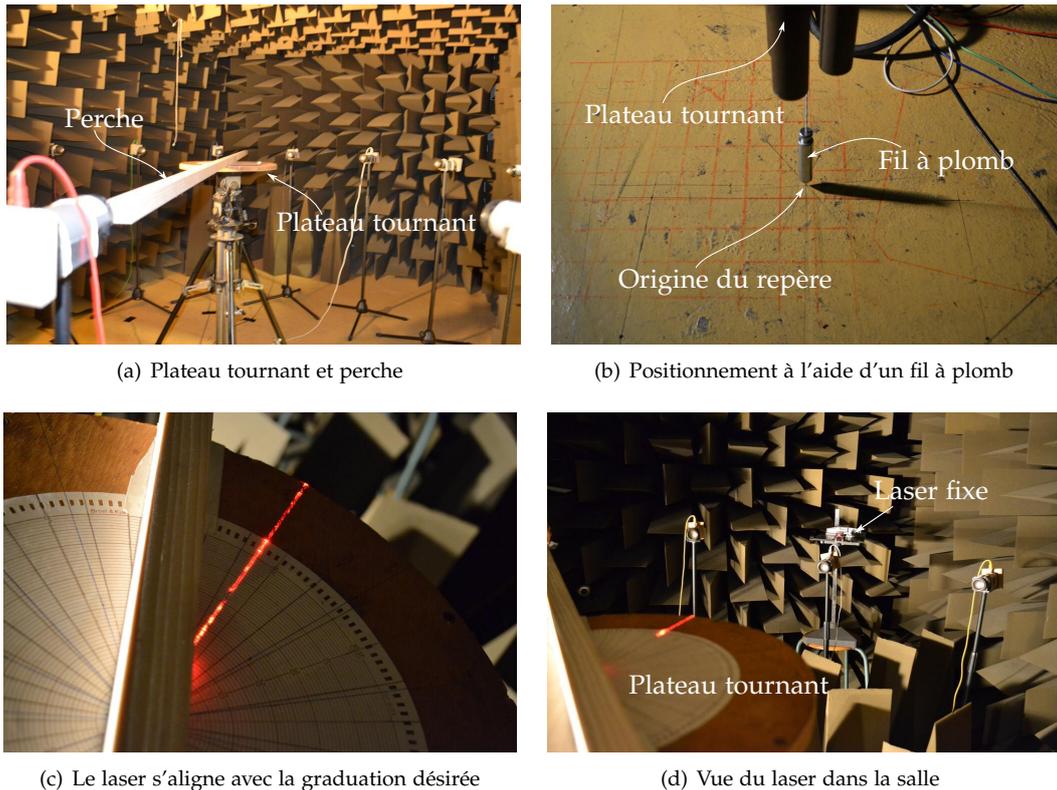


FIGURE 5.2.: Système de placement des enceintes dans la salle

principal du programme (on explique les différents éléments à proximité des chiffres en bleus sur la figure). La configuration du système de restitution est entrée dans la partie 1 : le rayon du cercle des haut-parleurs et éventuellement un gain différent pour chacun d'entre eux. La sélection du signal à spatialiser est faite en 2. Pour le son pur, sa fréquence peut être réglée. Un lecteur avec les boutons "play", "stop" et "mute" et "master-level" est disponible en 4. La spatialisation de la source se fait en 5. Le système étant en deux dimensions, on règle la distance et l'angle par rapport à l'origine (voir figure 4.2(a)). L'ordre de restitution est fixé en 8. Le système disposant de $N = 15$ haut-parleurs, on sort de la configuration $N = 2M + 1$ pour $M \neq 7$. En 6 un bouton "panoramique" permet de faire tourner la source autour de l'auditeur en faisant varier l'angle. La fréquence d'échantillonnage est choisie en 7. Enfin, les différents patches utilisés sont listé en 3.

L'implémentation des filtres NFC est faite suivant l'article [Daniel, 2003] à l'aide de cellules bi-quad dans les patches "pd ordre 1...7".

5.6 DÉCONVOLUTION DE L'EFFET DE SALLE EN TEMPS-RÉEL À L'AIDE DU LOGICIEL BRUTEFIR

5.6.1 Présentation de BRUTEFIR

BRUTEFIR⁷ est un logiciel de convolution en temps réel conçu par Anders Torger's. Il permet d'effectuer le filtrage en temps-réel de plusieurs flux audios avec des filtres RIF. L'opération de convolution est réalisée extrêmement rapidement à l'aide d'un algorithme de convolution partitionnée [Torger,]. Le passage dans le domaine spectral par FFT permet d'accélérer le processus de filtrage. Les filtres ayant une longueur de l échantillons, les signaux "zéros-paddés" pour appliquer la TFD sur au moins $2l - 1$ échantillons de manière à obtenir une convolution linéaire par TFD [Oppenheim et Schafer, 1975, p. 110]

7. <http://www.ludd.luth.se/~torger/brutefir.html>

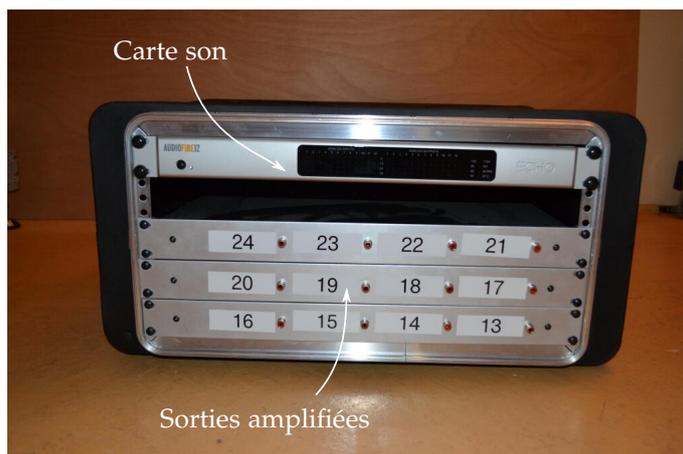


FIGURE 5.3.: Carte son et amplificateur utilisé pour le système ANFC-HOA

5.6.2 Fichier de configuration

BRUTEFIR se pilote par un fichier de configuration qui s'articule en plusieurs sections. Dans la première on précise toutes les informations générales nécessaires au filtrage : fréquence d'échantillonnage, longueur des filtres, nombre d'entrées et de sorties. La seconde section permet de définir les coefficients des réponses impulsionnelles des filtres. Dans notre cas ces coefficients sont stockés dans des fichiers *.txt. Enfin, la dernière section permet de définir quel filtre s'applique à quelle entrée vers quelle sortie.

Dans le cadre du stage deux types de filtrages sont opérés. Le premier corrige les canaux par eux même au centre du dispositif (voir 3). Ainsi on chargera 15 filtres qui filtreront respectivement les signaux des 15 entrées vers les 15 sorties. Le deuxième type de filtrage met en oeuvre une matrice de filtre H de dimensions 15×15 . Ainsi 225 filtres sont utilisés! Le filtre d'indices i, j dans la matrice H filtre le signal de l'entrée i et renvoie le signal filtré vers la sortie j , où $i, j \in [1 \dots 15] \cup [1 \dots 15]$.

Deux exemples de fichiers de configuration sont disponibles en ligne : un pour la correction canal par canal au centre du dispositif⁸ et un autre pour la correction multicanale.⁹

5.7 MESURE DE CHAMPS SONORES

5.7.1 Configuration de l'antenne

Le laboratoire dispose de 42 microphones "doubles couches". Ces microphones et leurs connectiques ont été fabriqués au Centre de Transfert des Technologies du Mans (CTTM)¹⁰ (figure 5.6). Les capsules de ces sondes doubles couches sont des microphones à électret Senheiser Ke4¹¹. L'avantage de ce genre de microphone réside dans la double couche. En effet, deux capsules sont superposées et espacées d'une distance de 3cm ceci permet de calculer le gradient de pression par différence finie jusqu'à quelques milliers de Hertz. Cette information peut être utile pour calculer des lignes de flux acoustique par exemple. Cependant, dans le cadre de l'utilisation actuelle en antenne horizontale, il n'est pas judicieux de calculer le gradient de pression dans l'axe vertical. Une antenne sphérique serait alors utile pour procéder à des mesures de flux ou à des décompositions en harmoniques sphériques pour caractériser le champ acoustique synthétisé. Ceci fait parties des ouvertures du stage, à explorer.

Une antenne plane de 7×6 microphones est utilisée pour réaliser des mesures de champ acoustique (figure 5.7). Une seule des deux couches des microphones est utilisée pour mesurer

8. <http://www.sekisushai.net/ANFCHOA/CHAP05/BruteFir.txt>

9. <http://www.sekisushai.net/ANFCHOA/CHAP05/BruteFirMulti.txt>

10. <http://www.cttm-lemans.com>

11. Fiche technique disponible en ligne : <http://www.sekisushai.net/ANFCHOA/CHAP05/Ke4.pdf>

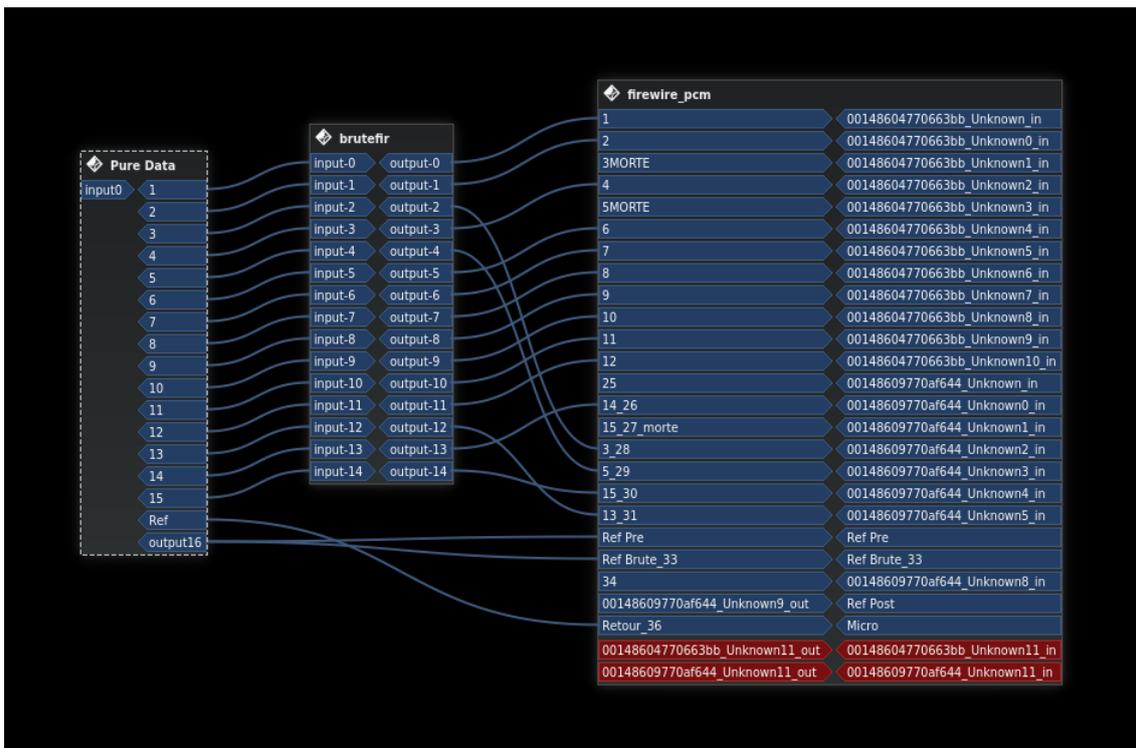


FIGURE 5.4.: Exemple des connexions pour une session ANFC-HOA sous JACK/CLAUDIA. A gauche les sorties de PURE DATA. Les 15 signaux NFC-HOA sont envoyés vers BRUTEFIR (au milieu) et sont filtrés en temps-réel. En sortie de BRUTEFIR les signaux filtrés sont envoyés à la carte son (droite) reliée au 15 haut-parleurs dans la salle de restitution. Les signaux de référence sont envoyés sur des voies annexes de la carte son pour une numérisation synchrone (voir 3.2)

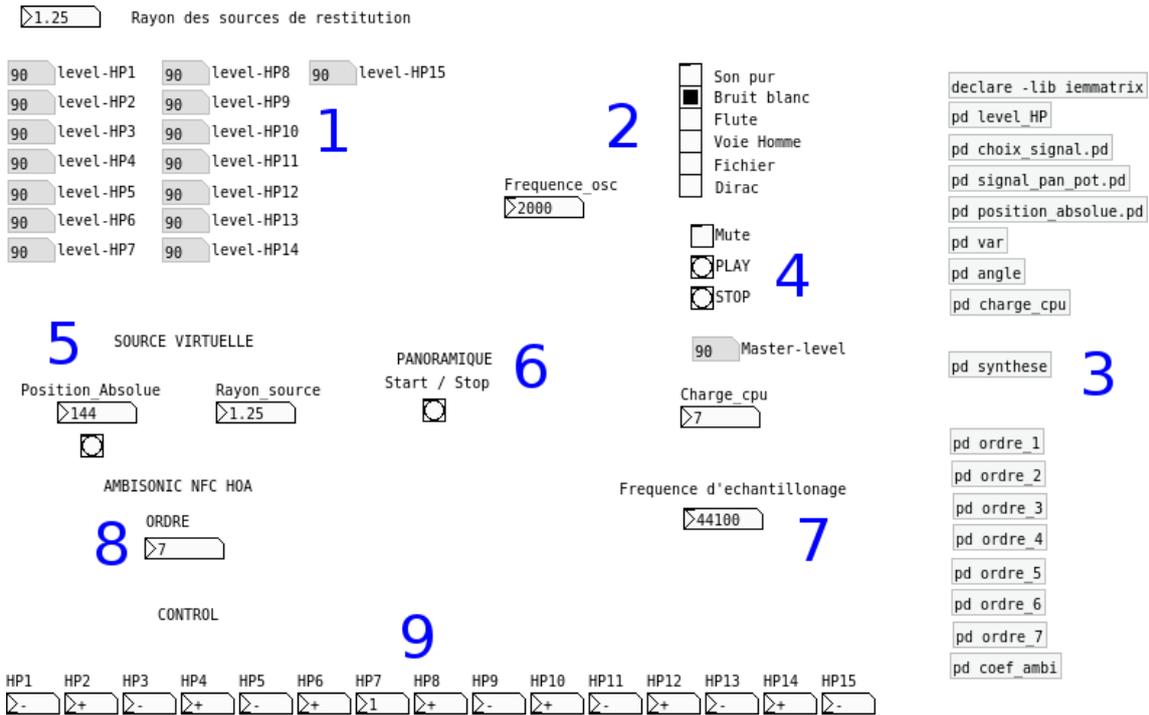


FIGURE 5.5.: Interface de contrôle principal sous Pure Data. Voir le texte pour l’explication des différents contrôles

l’information de pression. L’altitude de l’antenne est alors égale à la hauteur du plan de restitution $h = 1.25\text{m}$ (voir 5.3). Les paramètres géométriques de l’antenne sont les suivants :

$$\left\{ \begin{array}{l} N_x = 7 \\ N_y = 6 \\ \Delta_x = 6\text{cm} \\ \Delta_y = 6\text{cm} \\ L_x = 36\text{cm} \\ L_y = 30\text{cm} \end{array} \right. \quad (5.1)$$

où N_x est le nombre de micros suivant l’axe x , N_y suivant l’axe y . Δ_x est le pas entre deux microphones suivant l’axe x , Δ_y suivant l’axe y . L_x est la longueur de l’antenne suivant l’axe x , L_y suivant l’axe y

5.7.2 Configuration de l’antenne globale

Les dimensions de l’antenne seule ne permettent pas d’observer plusieurs longueurs d’ondes dans les basses fréquences. La dimension la plus défavorable est $L_y = 30\text{cm}$. Un rapide calcul montre qu’une demi-longueur d’onde est observable à partir de $f = \frac{c}{2L_y} \approx 570\text{Hz}$. En dessous le champ est trop homogène pour observer des variations. Dans ce contexte, on décide d’augmenter artificiellement la taille de l’antenne en réalisant 4 mesures avec différents placements de l’antenne pour créer une antenne globale en “recollant” les 4 mesures. Cette technique est possible car les mesures sont réalisées de manière synchrone avec une référence de phase électrique (voir A et B). On peut alors réaliser les mesures des 4 antennes à différents moments.

La configuration de l’antenne globale est présentée figure 5.8. Pour la réalisation pratique, l’antenne est positionnée à l’aide de deux fils à plomb qui sont à l’aplomb d’un marquage au sol (figure 5.7(b)). Le passage de la position 1 à la position 2 se fait par translation de L_x . Le passage de la position 2 à la position 3 se fait par une rotation de l’antenne de 180° et une translation de $-L_y$. Enfin le passage de la position 3 à la position 4 se fait par translation de $-L_x$ (figure 5.8(a))

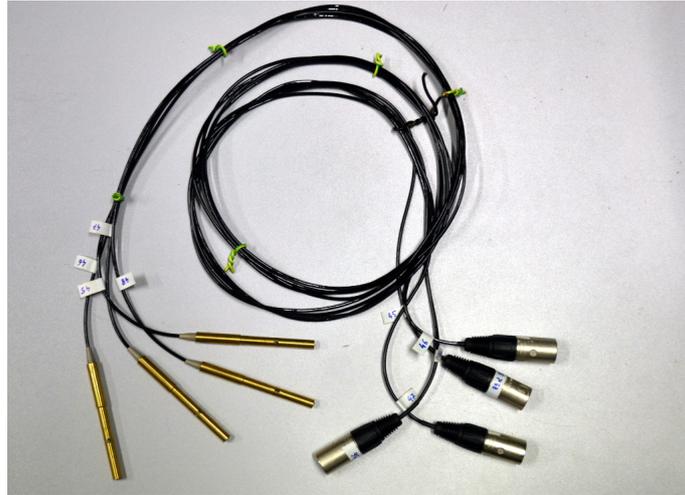


FIGURE 5.6.: Micros Ke4 double couches

On remarque que les différentes antennes se recoupent suivant un motif en croix pour obtenir une antenne globale dont la configuration est la suivante :

$$\left\{ \begin{array}{l} N_x = 13 \\ N_y = 11 \\ \Delta_x = 6\text{cm} \\ \Delta_y = 6\text{cm} \\ L_x = 72\text{cm} \\ L_y = 60\text{cm} \end{array} \right. \quad (5.2)$$

A présent, on peut observer au moins une longueur d'onde à partir de $f \approx 285\text{Hz}$. Au dépouillement des mesures, un feuille de calcul MATHEMATICA ©recolle les mesures sur l'antenne globale et moyenne les mesures suivant le motif en croix. Un exemple de cette opération est montré figure 5.9. Pour effectuer les mesures, le signal d'excitation est un bruit blanc, ou un sinus glissant suivant les situations (voir 3.4.3). Ceci permet d'obtenir toutes les fréquences en une seule mesure. La configuration géométrique du système de restitution avec l'antenne de mesure est montrée figure 4.2(a).

5.7.3 Acquisition des mesures

L'acquisition des mesures se fait sur un ordinateur dédié à l'aide de l'antenne de mesure et un frontal d'acquisition NATIONAL INSTRUMENTS PXI ©(figure 5.10(a)). Le pilotage logiciel se fait sous LABVIEW ©(figure 5.10(b)).

Le signal d'excitation est envoyé via le programme PURE DATA (voir 5.5.1) sur un autre ordinateur. Suivant le type de signal deux stratégies pour le déclenchement des mesures sont adoptés.

Si le signal d'excitation est un bruit blanc, la mesure peut être déclenchée à n'importe quel moment, étant en régime stationnaire. En effet avec la technique d'estimation des réponses impulsionnelle utilisée (voir 3.2) un moyennage est effectué sur les signaux.

Par contre dans le cas d'un sinus glissant (voir annexe B), la mesure doit être déclenché précisément au même instant que le signal est envoyé au système par PURE DATA. Ainsi, un signal supplémentaire de type front montant est émis simultanément avec le sinus glissant. Ce signal sert à déclencher la mesure sous LABVIEW ©par détection du front montant. Sur la figure 5.11 on peut observer un sinus glissant et son "Trigger" associé.

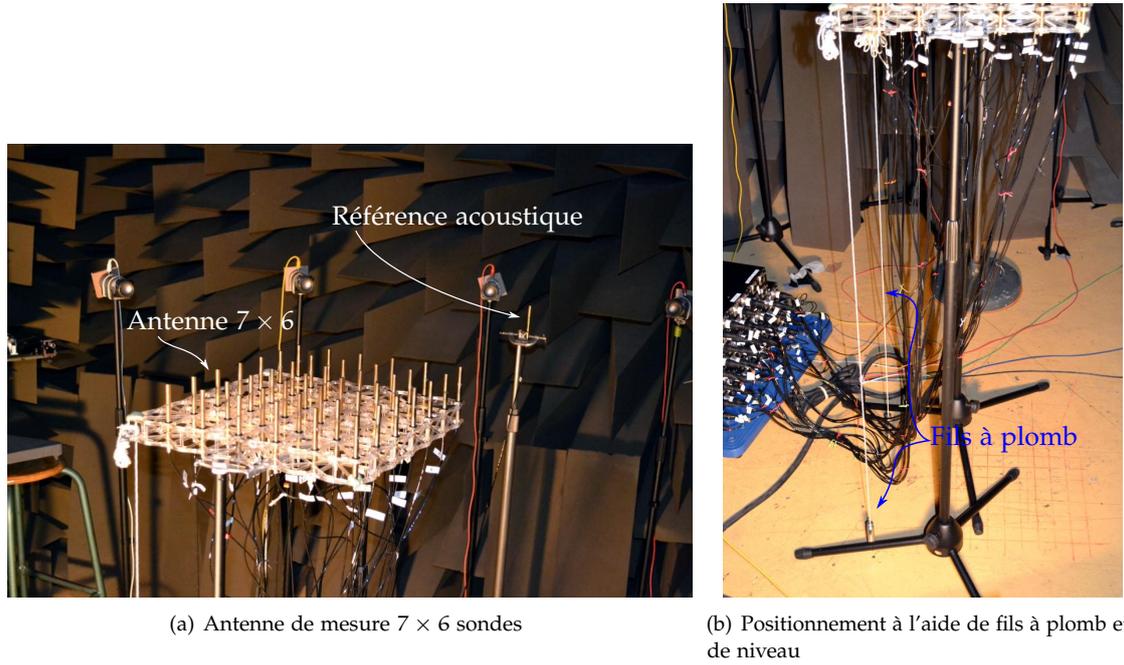


FIGURE 5.7.: Configuration géométrique de l'antenne globale

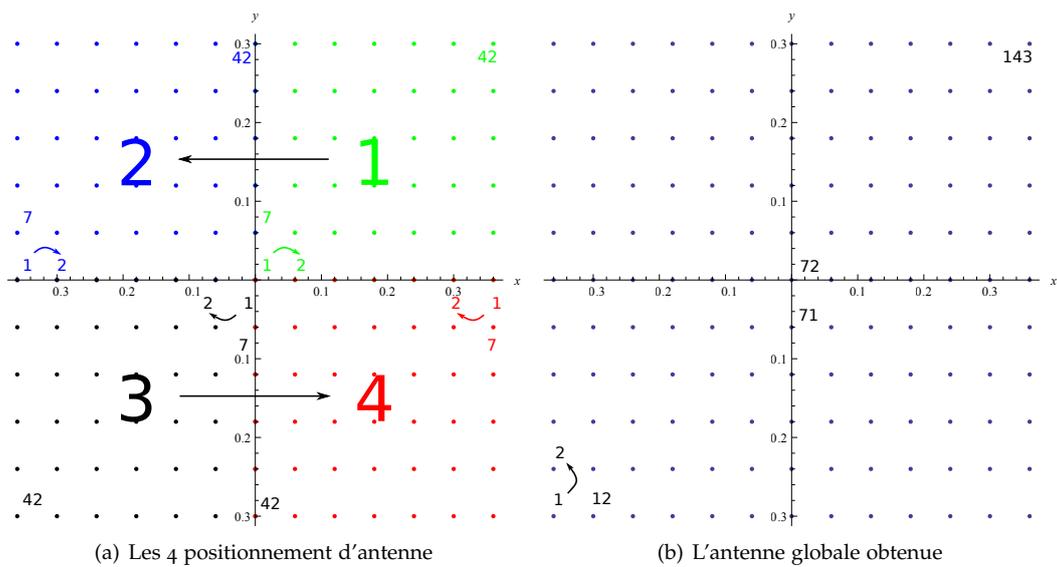
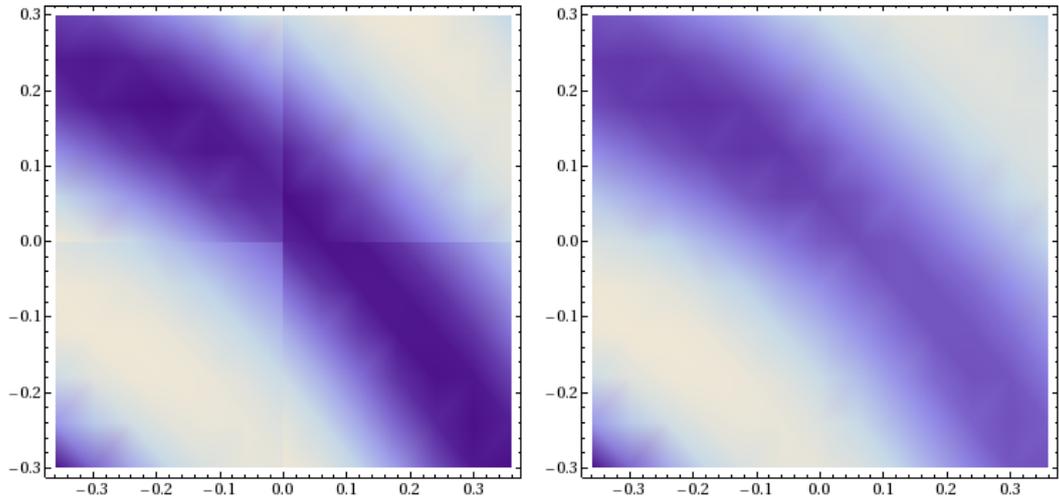


FIGURE 5.8.: Configuration géométrique de l'antenne globale

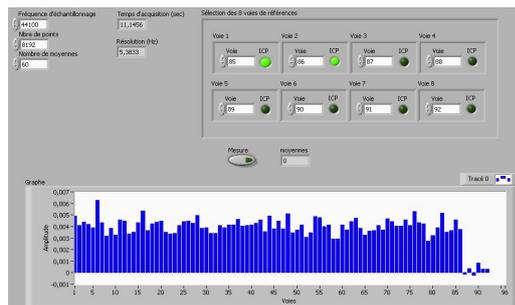


(a) Les 4 mesures obtenues pour les 4 positions d'antenne. Ici le haut-parleur 7
 (b) La mesure obtenue après recollage, Ici le haut-parleur 7

FIGURE 5.9.: Un exemple de mesure : Champ issu du haut-parleur 7 à $f = 533\text{Hz}$. La partie réelle du champ est affichée. Il s'agit en réalité de la fonction de transfert entre les microphones et la référence électrique



(a) Frontal d'acquisition NATIONAL INSTRUMENTS ©PXI



(b) Acquisition sous LABVIEW

FIGURE 5.10.: Acquisition des mesures à l'aide du frontal PXI et de LABVIEW

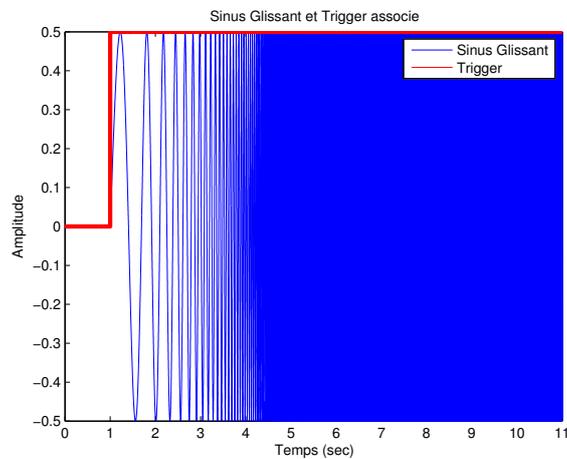


FIGURE 5.11.: Sinus glissant et son front montant associé pour le déclenchement synchrone de la mesure

CONCLUSIONS

6.1 CONCLUSIONS DES TRAVAUX MENÉS

6.1.1 *Remise en contexte*

Les méthodes ambisoniques d'ordres élevés présentées dans ce rapport permettent la synthèse réaliste, en basses fréquences, de champs sonores. À l'instar de méthodes comme la WFS, elles permettent d'avoir des résultats significatifs avec relativement peu de haut-parleurs. Dans ce contexte, un dispositif de restitution à l'ordre 7 utilisant 15 haut-parleurs équirépartis sur un cercle de rayon $r_{HP} = 1.25\text{m}$ a été installé dans la chambre semi-anéchoïque du LMSSC. La synthèse de champs sonores est alors contrôlée en temps réel à l'aide d'une station opérant sous Linux/JACK avec le logiciel PURE DATA.

Le formalisme ambisonique requiert une condition de champ libre pour reconstruire efficacement un champ sonore sur une zone d'écoute relativement large. Le critère d'erreur relative à 20% entre le champ reconstruit et le champ original a été choisi pour caractériser la reconstruction. De même, les sources de restitution sont modélisées comme des monopôles pour l'approche NFC-HOA qui a été retenue dans le cadre du stage.

En situation réelle, la condition de champ libre n'est pas réaliste, et les haut-parleurs de restitution ne sont pas des monopôles. Alors, le champ obtenu est perturbé par l'acoustique du lieu de restitution. La taille de la zone d'écoute s'en trouve diminuée, elle devient même inexistante en utilisant le critère d'erreur relative à 20%. Pour améliorer les performances du système, on propose des techniques de correction du système de restitution. Plus exactement on essaie de ramener la réponse du système la plus proche possible de monopôles en champ libre.

Le cas d'étude du stage est une chambre semi anéchoïque, où seuls les réponses des haut-parleurs et l'effet du sol doivent être corrigées.

6.1.2 *Inversion de la réponse de chaque canal de restitution*

Dans un premier temps, on a cherché à corriger la réponse de chacun des 15 canaux de restitution du système au centre du dispositif. Plusieurs techniques de conception de filtres inverses ont alors été testées en simulation à partir de la mesure des réponses impulsionnelles des 15 canaux au centre du dispositif. Après conception de 4 types de filtres, ces derniers ont été implémentés sous le logiciel BRUTEFIR pour leur application en temps réel. Au niveau du système de restitution, ces 15 filtres sont actifs entre la sortie de PURE DATA et la carte son.

Les 4 types de filtres conçus sont :

- Filtres à minimum de phase
- Filtres conçus à l'aide d'une inversion au sens des moindres carrées avec régularisation (LS).
- Filtres conçu par inversion fréquentielle sans régularisation
- Filtres conçu par inversion fréquentielle sans régularisation après lissage de la réponse des filtres direct par méthode cepstrale.

Des précautions ont dû être prises lors de la conception de ces filtres. En effet, l'inversion directe des réponses des canaux ne marche généralement pas. Ceci est dû à la nature des réponses des canaux qui ne sont pas à minimum de phase.

Après mesure sur la salle de restitution, seuls les filtres à minimum de phase et "LS" ont donné de bons résultats, bien qu'en simulation les 4 types de filtres semblaient fonctionner.

L'observation des champs sonores obtenus, ainsi que de la réponse des filtres en dehors du point de correction, nous montrent que l'on contrôle l'acoustique qu'au seul point de correction. En dehors, on ne prédit pas ce qui se passe. L'effet de salle n'est alors diminué qu'en un seul point. Ceci n'étant pas suffisant pour notre critère d'erreur on décide alors de se tourner vers des techniques d'inversion multicanales.

6.1.3 *Inversion multicanale*

L'inversion multicanale consiste à corriger le système de restitution sur un ensemble de points choisi l'intérieur de la zone d'écoute théorique. Ainsi les 15 hauts-parleurs de restitution servent à corriger l'un d'entre eux sur l'ensemble des points de contrôle. Par permutation, on constitue une matrice de filtres de correction.

Deux approches ont été retenues. La première utilise une inversion fréquentielle avec régularisation. La seconde utilise l'approche proposée dans le cadre de l'AWFS mais adaptée à l'ambisonie. L'objectif de ces deux approches est le même : corriger la réponse de chacun des canaux sur un ensemble de points de contrôles à l'aide de tous les canaux disponibles. Une fois cet objectif atteint, les gains NFC-HOA sont appliqués. Ceci revient alors à filtrer les signaux émis par chacun des haut-parleurs de restitution à l'aide de la matrice de filtres conçue. Sur le système de restitution, le logiciel BRUTEFIR vient s'intercaler entre la sortie de PURE DATA et la carte son pour effectuer ce filtrage.

En simulation, l'ANFC-HOA donne des résultats prometteurs pour une situation avec un effet de sol. Effectivement la zone d'écoute une fois corrigée tend à s'approcher de celle en champ libre. Seulement la performance est très grandement dépendante du choix du paramètre de régularisation β . Ce dernier doit être choisi au meilleur compromis entre les contraintes de minimisation d'erreur, de l'énergie de la solution, mais aussi de l'allure des filtres inverses !

En mesure, l'ANFC-HOA comme l'inversion plus classique donnent des résultats assez prometteurs. D'un point de vue perceptif, les premières impressions donnent l'avantage à l'ANFC-HOA pour ce qui est de la localisation de la source synthétisée.

6.1.4 *Simulations et mesures*

Il est important de noter qu'il existe un grand écart entre les résultats de simulations et le monde de la mesure. En effet, bien que les filtres inverses aient été conçus à partir de vraies mesures, les résultats obtenus sont toujours moins bons que lors de la simulation. De nombreux facteurs expliquent cela. Les mesures de réponses impulsionnelles et de réponses impulsionnelles n'ont pas été faites le même jour ni avec les mêmes techniques. À cela s'ajoutent le bruit de mesure, la précision du placement de l'antenne de mesure, etc.

6.2 FUTURS TRAVAUX

Ce stage a permis de lever de nombreuses problématiques pour d'éventuels travaux futurs sur le sujet. Ci-après nous mentionnons quelques améliorations envisageables pour la suite.

6.2.1 *Formulation des filtres inverses*

Tout d'abord la formulation des filtres inverses pour la déconvolution ANFC-HOA doit être améliorée. En effet, actuellement le choix du paramètre de régularisation se fait par méthode essai-erreur, pour trouver "à la main" le meilleur compromis entre des contraintes de problèmes inverses et de traitement du signal. Un algorithme prenant en compte toutes ces contraintes et proposant alors une solution optimale doit être développé.

6.2.2 *Ajout de sources pour aider la correction*

Après de premiers essais convaincants en simulation, l'ajout de sources supplémentaires pour aider la correction est envisageable. Effectivement, le problème de déconvolution se base sur une formulation matricielle et le conditionnement de cette matrice joue un rôle important pour l'inversion. Ce conditionnement dépend de la qualité des mesures mais surtout du nombre de points à corriger par rapport au nombre de sources disponibles. L'ajout de source permettrait alors de mieux conditionner le problème.

6.2.3 *Passage en salle réelle*

La prochaine étape est d'installer le système de restitution dans une salle réelle. Ainsi l'acoustique ne serait plus aussi contrôlée que dans une chambre semi-anechoïque. La dégradation des champs synthétisés sans filtrage s'en ferait alors ressentir. Les réponses impulsionnelles seront alors de plus grande longueur, demandant un filtrage plus conséquent pour diminuer l'effet de salle. L'intérêt majeur est de se rapprocher de la situation réelle d'utilisation de la NFC-HOA. En effet, le système de restitution n'est généralement pas dans une salle où l'acoustique est finement contrôlée.

6.2.4 *Simulation d'effet de salles*

Lors de la formulation des filtres inverses, la cible est une réponse du système en champ libre. Cependant rien n'empêche de définir une autre cible qui serait la réponse du système dans une autre salle ! On pourrait alors simuler des effets de salle de restitution au niveau de la zone d'écoute. Bien entendu l'application serait ici orientée audio car la reconstruction physique serait plus dégradée que lorsque la cible est une réponse en champ libre.

6.2.5 *Critère d'erreur*

Lors du stage, il n'a pas été possible de tracer un critère d'erreur relative à 20% entre le champ théorique et la mesure. En effet ceci est trop dur à atteindre directement. Il serait alors judicieux de fixer un critère d'erreur par rapport à une réponse dans la salle simulée, ou alors par rapport à la réponse réelle d'une source placée à l'endroit physique ou l'on souhaite synthétiser une source.

6.2.6 *Caractérisation par décomposition en harmoniques sphériques*

Pour le moment, seule une antenne plane a été utilisée pour observer l'allure des champs synthétisés. L'ambisonie utilisant un formalisme en harmoniques sphériques, il serait alors judicieux d'utiliser une antenne sphérique permettant de refaire une décomposition en harmoniques sphériques du champ synthétisé, et alors de caractériser ce dernier en comparant la décomposition originale et obtenue après synthèse. Dans cette optique une installation d'un système de synthèse de champ sonore en trois dimensions serait envisagée.

6.2.7 *Caractérisation fines des sources de restitution*

Le LMSSC met en oeuvre des techniques de caractérisation fine des sources. On pourrait alors se servir de ces techniques pour caractériser finement les haut-parleurs de restitution et utiliser cette caractérisation dès l'encodage ambisonique. Ainsi là où la NFC-HOA modélise les haut-parleurs de restitution comme des monopôles, on pourrait utiliser directement le modèle réaliste des haut-parleurs. L'encodage ambisonique contiendrait alors l'information de la réponse réelle des sources de restitution. Ceci relâcherait donc une contrainte sur la formulation des filtres inverses qui ne serviraient plus qu'à corriger l'effet de salle !

6.2.8 *Mise en œuvre de décodages psychoacoustique*

Actuellement, le système de restitution utilise le décodage classique ambisonique qui se base sur la reconstruction physique du champ original. Cependant comme mentionné dans l'introduction les performances de ce décodage s'effondrent en hautes fréquences. Il serait alors envisageable, dans le cadre d'une utilisation audio, de mettre en œuvre les décodeurs psychoacoustiques existants.

6.2.9 *Caractérisation subjective des champs synthétisés*

Enfin, dans le cadre d'une utilisation audio, toute une étude de caractérisation psychoacoustique des champs synthétisés doit être menée, en vue d'améliorer le réalisme perceptif des environnements sonores simulés.

6.2.10 *Comparaison avec d'autres techniques de synthèse de champs sonores*

Une étude comparative de la qualité objective et subjective entre les différentes technologies de synthèse de champ sonore, notamment la WFS peut être également intéressante à mener.

Comme mentionné ci dessus, il reste encore beaucoup d'axes de recherche dans le domaine de la synthèse de champ sonore. Dans ce contexte, ce travail de stage va se poursuivre dès la rentrée 2013 en thèse de doctorat, dans le cadre d'une cotutelle entre le LMSSC et le Groupe d'Acoustique de l'Université de Sherbrooke, Québec, Canada.

ESTIMATION DE LA RÉPONSE IMPULSIONNELLE PAR FILTRAGE DE L'INTERSPECTRE

Le signal d'excitation $x[n]$ est un bruit blanc centré, représentant la réalisation d'un processus $X[n]$ centré stationnaire au sens large (SSL). On suppose que le filtrage $c[n]$ opéré par le système amplificateur - haut-parleur - salle est stable . Alors le signal mesuré $y[n]$ est également centré et représente une réalisation d'un processus centré $Y[n]$ SSL. [Blanchet et Charbit, 2001, p. 297]

Les signaux $x[n]$ et $y[n]$ sont de covariance R_{yx} , stationnaire et centrée. Les formules d'autocovariance et de covariance sont données ci-après

$$\mathbb{E}\{y[n+k]x[n]^*\} = R_{yx}[k] \quad (\text{A.1})$$

$$\mathbb{E}\{x[n+k]x[n]^*\} = R_{xx}[k] \quad (\text{A.2})$$

Le filtrage de la séquence $x[n]$ par $c[n]$ se traduit au niveau de l'autocovariance par [Blanchet et Charbit, 2001, p. 298]

$$R_{yx}[k] = \sum_{n=-\infty}^{+\infty} c[n]R_{xx}[k-n] \quad (\text{A.3})$$

En appliquant une transformée de Fourier à temps discret (TFTD) à la formule A.3 on obtient alors :

$$S_{yx}(f) = H(f) \cdot S_{xx}(f) \quad (\text{A.4})$$

De fait le gain complexe $H(f)$ peut être estimé par :

$$H(f) = \frac{S_{yx}(f)}{S_{xx}(f)} \quad (\text{A.5})$$

Cependant, dans notre cas les signaux sont de longueur finie et c'est une transformée de Fourier discrète (TFD) via l'algorithme `fft` de MATLAB© qui est appliquée pour calculer les autospectres et interspectres. On calcule $X[f] = \text{TFD}(x[n])$ et $Y[f] = \text{TFD}(y[n])$. Le produit $X[f]X[f]^*$ (respectivement $Y[f]X[f]^*$) correspond alors à la convolution circulaire des séquences $x[n]$ et $x[-n]$ (ou $y[-n]$) [Oppenheim et Schafer, 1975, p. 557]. Pour que l'estimation corresponde à ce qu'elle aurait été dans le cas d'une convolution linéaire les signaux sont complétés par des zéros avant l'application de la TFD [Oppenheim et Schafer, 1975, p. 111].

Il est important de noter que la longueur des signaux $x[n]$ et $y[n]$ doivent être supérieurs à la longueur de la réponse impulsionnelle de la salle pour pouvoir estimer cette dernière dans son intégralité.

Les signaux $x[n]$ et $y[n]$ représentent en réalité une seule réalisation des processus $X[n]$ et $Y[n]$. On moyenne alors plusieurs de ces réalisations pour améliorer l'estimation des autospectre et interspectre. Effectivement, bien que l'estimation est a priori non biaisée de manière asymptotique, la variance ne tend pas vers 0 en augmentant la longueur de la TFD. On a alors recours au fenêtrage et moyennage de plusieurs réalisations pour diminuer la variance (diminution du bruit de fond décorrélé avec le bruit d'excitation) et améliorer les estimations [Oppenheim et Schafer, 1975, p.555].

Une fois les gains complexes estimés par la formule (A.5), une TFD inverse donne la réponse impulsionnelle.

Le schéma A.1 résume la procédure.

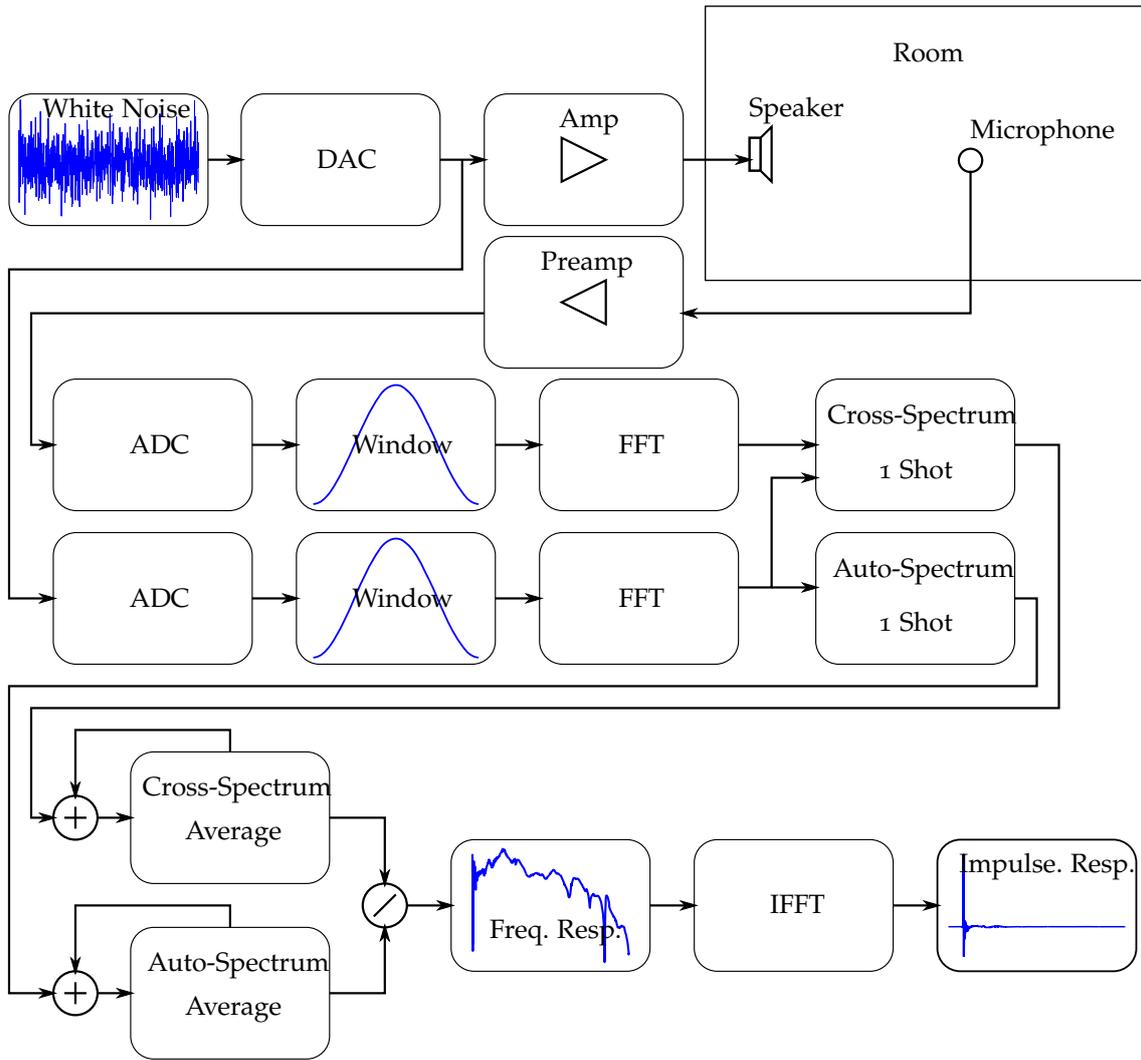


FIGURE A.1.: Mesure de la réponse impulsionnelle d'un canal. *White Noise* : Bruit Blanc ; *DAC* : *Digital Analog Conversion* : Conversion Analogique Numérique ; *ADC* : *Analog Digital Conversion* : Conversion Analogique Numérique ; *Auto-Spectrum* : Autospectre ; *Cross-Spectrum* : Interspectre

COHÉRENCE

En appliquant cette méthode, on peut également estimer au passage la cohérence¹ entre les deux signaux $X[f]$ et $Y[f]$. La cohérence est donnée par :

$$C_{xy}(f) = \frac{|S_{yx}(f)|^2}{S_{xx}(f)S_{yy}(f)} \tag{A.6}$$

On a $0 \leq C_{xy} \leq 1$.

1. [http://en.wikipedia.org/wiki/Coherence_\(signal_processing\)](http://en.wikipedia.org/wiki/Coherence_(signal_processing))

ESTIMATION DE LA RÉPONSE IMPULSIONNELLE PAR SINUS GLISSANT

La mesure de réponse impulsionnelle par sinus glissant présentée ici est directement tirée de [Farina, 2000]. Le schéma B.1 résume la procédure.

Un sinus glissant exponentiel excite le canal. L'équation de ce sinus est donnée ci-après :

$$x(t) = \sin \left[\frac{\omega_1 T}{\ln \left(\frac{\omega_2}{\omega_1} \right)} \left(e^{\frac{t}{T} \ln \left(\frac{\omega_2}{\omega_1} \right)} - 1 \right) \right] \quad (\text{B.1})$$

Cette équation représente un sinus glissant exponentiel d'une durée de T secondes, commençant à la pulsation ω_1 et finissant à la pulsation ω_2 .

On obtient la mesure du signal $y[n]$ obtenu après filtrage par $c[n]$ (amplificateur-salle-microphone).

En ayant

$$y[n] = x[n] \star h[n] \quad (\text{B.2})$$

et

$$x[n] \star f[n] = \delta[n] \quad (\text{B.3})$$

La réponse impulsionnelle $c[n]$ est donnée par [Farina, 2000] :

$$c[n] = y[n] \star f[n] \quad (\text{B.4})$$

où $f[n]$ est simplement le signal $x[n]$ retourné temporellement et filtré avec un filtre d'accentuation de +3dB par octave. En effet le signal $x[n]$ ne possède pas un spectre plat, mais un gain qui décroît de -3dB par octave.

Si les signaux d'entrée $x[n]$ et de sortie $y[n]$ ont une longueur de l , la réponse impulsionnelle obtenue est de longueur $2l - 1$ (convolution linéaire). Ainsi, on la tronque pour ne retenir que la partie utile. Cette troncature modifie la réponse en phase.

Dans le cadre du stage, on tronque à partir du même échantillon toutes les réponses impulsionnelles sur l'antenne. Ceci permet de garder la phase relative entre les microphones. De même les convolutions linéaires sont calculées par TFD en insérant des zéros sur les signaux [Oppenheim et Schaffer, 1975, p. 110].

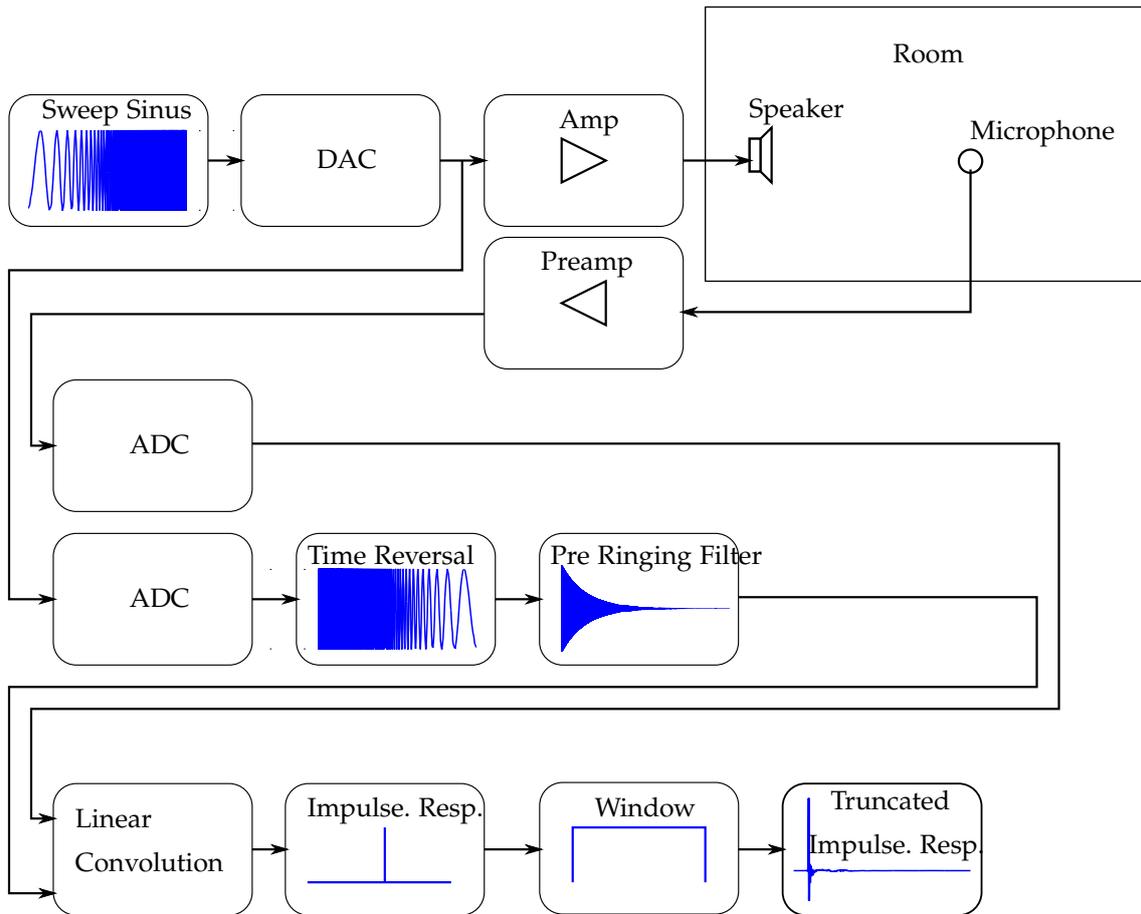


FIGURE B.1.: Mesure de la réponse impulsionnelle d'un canal. *Sweep sinus* : Sinus glissant; *DAC* : *Digital Analog Conversion* : Conversion Analogique Numérique; *ADC* : *Analog Digital Conversion* : Conversion Analogique Numérique; *Time Reversal* :Retournement temporel; *Pre Ringin Filter* : Filtre d'accentuation; *Linear Convolution* : Convolution linéaire

DÉCOMPOSITION D'UN CHAMP ACOUSTIQUE EN HARMONIQUES CYLINDRIQUES

C.1 ÉQUATION DES ONDES

Soit un point O de l'espace définit dans un système de coordonnées. On considère un domaine exempt de sources sonores (les sources sont à l'extérieur de ce domaine). L'équation des ondes en dehors des sources s'écrit [Bruneau, 1998, p. 171] en temporel

$$\Delta p(t) = \frac{1}{c^2} \frac{\partial^2 p(t)}{\partial t^2} \quad (\text{C.1})$$

C.2 ÉQUATION DES ONDES EN COORDONNÉES CYLINDRIQUES

Soit un point M de l'espace définit par ses coordonnées cylindrique r, θ, z (voir figure 2.4) où la pression acoustique vérifie l'équation (C.1). Cette equation s'exprime alors dans le système de coordonnée cylindriques de la manière suivante [Bruneau, 1998, p.224]

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial p}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} (p) + \frac{\partial^2 p}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \quad (\text{C.2})$$

Cette équation admet des solutions à variables séparées de la forme [Bruneau, 1998] :

$$p = R(r)\Phi(\varphi)Z(z)T(t) \quad (\text{C.3})$$

La résolution donne

$$\begin{cases} T(t) = e^{j\omega t} \\ R(r) = A_{1m}J_m(k_w r) + A_{2m}N_m(k_w r) \quad \text{ou} \quad B_{1m}H_m^+(k_w r) + B_{2m}H_m^-(k_w r) \\ \Phi(\varphi) = C_1(\cos(m\varphi)) + C_2(\sin(m\varphi)) \\ Z(z) = D_1(\cos(k_z z)) + D_2(\sin(k_z z)) \end{cases} \quad (\text{C.4})$$

Les fonctions J_m sont les fonctions de Bessel cylindriques de première espèce d'ordre m et les fonctions N_m sont celles de deuxième espèce (ou fonctions de Neumann cylindriques d'ordre m). Les fonctions H_m^+ et H_m^- sont les fonctions de Hankel convergentes et divergentes. La relation de dispersion est donnée par :

$$k_w^2 = k^2 - k_z^2 \quad (\text{C.5})$$

C.3 HARMONIQUES CYLINDRIQUES

On définit les harmoniques cylindriques de la manière suivante [Daniel, 2000, p. 302]

$$Y_m^\sigma(\varphi) = \begin{cases} \cos(m\varphi) & m \geq 0 \quad \sigma = 1 \\ \sin(m\varphi) & m \geq 1 \quad \sigma = -1 \end{cases} \quad (\text{C.6})$$

Elles forment une base orthogonale qui peut servir à décomposer une onde. Une version orthonormée est donnée par :

$$\tilde{Y}_m^\sigma = \sqrt{\epsilon_m} Y_m^\sigma, \quad \sigma = \pm 1, \quad \begin{cases} \epsilon_0 = 1 \\ \epsilon_m = 2 \quad m \geq 1 \end{cases} \quad (\text{C.7})$$

C.4 SOLUTION DE L'ÉQUATION DES ONDES EN HARMONIQUES CYLINDRIQUES

La solution de l'équation des ondes (C.2) devient alors :

$$p(r, \varphi) = \sum_{m, \sigma} \tilde{Y}_m^\sigma(\varphi) j^m (A_m^\sigma J_m(kr) + j B_m^\sigma N_m(kr)) \quad (\text{C.8})$$

Dans le cas d'un champ exempt de sources sur un cylindre infini de rayon r_2 , $B_m = 0$ et l'équation est valable sur $0 \leq r \leq r_2$. Dans le cas contraire apparaissent les fonctions de Hankel convergentes ou divergentes (C.4) suivant le sens de propagation des ondes. On considère le cas sans sources dans le cadre du stage. Ainsi l'équation devient :

$$p(r, \varphi) = \sum_{m, \sigma} \tilde{Y}_m^\sigma(\varphi) j^m A_m^\sigma J_m(kr) \quad (\text{C.9})$$

DÉCOMPOSITION D'UN CHAMP ACOUSTIQUE EN HARMONIQUES SPHÉRIQUES

D.1 ÉQUATION DES ONDES EN COORDONNÉES SPHÉRIQUES

La première étape de la décomposition en harmoniques sphériques consiste à exprimer l'équation des ondes C.1 en coordonnées sphériques.

Soit un point M de l'espace défini par ses coordonnées sphériques r, θ, φ (voir figure 2.6 où la pression acoustique vérifie l'équation (C.1)). Cette équation s'exprime alors dans le système de coordonnées sphériques de la manière suivante [Morse et Ingard, 1968, p. 334] :

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial p}{\partial r} \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial p}{\partial \theta} \right) + \frac{1}{r^2 \sin^2(\theta)} \frac{\partial^2 p}{\partial \varphi^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \quad (\text{D.1})$$

Cette équation admet des solutions à variables séparées [Morse et Ingard, 1968, p. 333] [Bruneau, 1998, p. 241] sous la forme :

$$p(r, \theta, \varphi) = R(r)\Theta(\theta)\Phi(\varphi)T(t) \quad (\text{D.2})$$

La résolution donne [Bruneau, 1998, p. 242] :

$$\begin{cases} T(t) = e^{j\omega t} \\ R(r) = A_{1n} j_n(kr) + A_{2n} n_n(kr) & \text{ou } B_{1n} h_n^-(kr) + B_{2n} h_n^+(kr) \\ \Theta(\theta) = P_{nm}(\cos(\theta)) \\ \Phi(\varphi) = \alpha_{1m} \cos(m\varphi) + \alpha_{2m} \sin(m\varphi) \end{cases} \quad (\text{D.3})$$

La solution de l'équation (D.1) est alors une combinaison linéaire des fonctions ci-dessus en faisant varier m et n . j_n représentent les fonctions de Bessel sphériques de première espèce d'ordre n ; n_n les fonctions de Neumann sphériques; h_n^- les fonctions de Hankel sphériques d'ordre n divergentes; h_n^+ les fonctions de Hankel sphériques d'ordre n convergentes; P_{nm} les fonctions de Legendre qui sont exprimables avec les polynômes de Legendre associé [Daniel, 2000, p. 307].

D.2 HARMONIQUES SPHÉRIQUES

Le produit des fonctions polaires et azimutales définit les harmoniques sphériques [Bruneau, 1998, p. 242] :

$$Y_{mn}^\sigma(\theta, \varphi) = P_{mn}(\cos(\theta)) \times \begin{cases} \cos(n\varphi) & \text{pour } \sigma = 1 \\ \sin(n\varphi) & \text{pour } \sigma = -1 \text{ et } n \geq 1 \end{cases} \quad (\text{D.4})$$

Avec $m \geq 0$ et $0 \leq n \leq m$. Ces fonctions forment une base de décomposition pour les solutions de l'équation des ondes sphériques (D.1). Dans l'usage, on définit une base orthonormée de d'harmoniques sphériques de la manière suivante [Daniel, 2000, p.304] :

$$\tilde{Y}_{mn}^\sigma(\theta, \varphi) = \sqrt{2m+1} \sqrt{\epsilon_n \frac{(m-n)!}{(m+n)!}} Y_{mn}^\sigma(\theta, \varphi) \quad (\text{D.5})$$

avec $\epsilon_0 = 1$ et $\epsilon_n = 2$ pour $n \geq 1$.

Sur la figure D.1 on peut observer une représentation des harmoniques sphériques jusqu'à l'ordre $M = 3$.

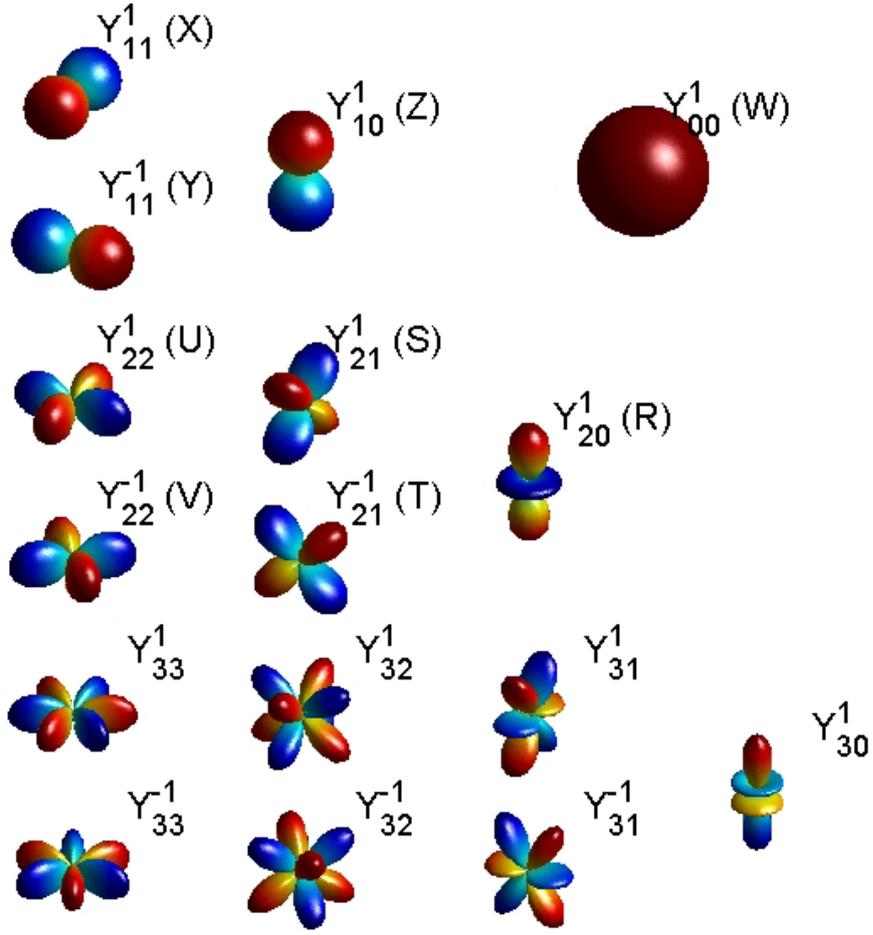


FIGURE D.1.: Une représentation des harmoniques sphériques jusqu'à l'ordre $M = 3$ les parties rouges et bleus correspondent respectivement aux valeurs positives et négatives des fonctions harmoniques sphériques. Figure tirée de l'article [Daniel, 2003].

D.3 SOLUTION DE L'ÉQUATION DES ONDES EN SÉRIES DE FOURIER-BESSEL

Finalement, les solutions de l'équation (D.1) peuvent s'exprimer sous la forme du développement en séries de Fourier-Bessel suivant :

$$p(r, \theta, \varphi) = \sum_{m,n,\sigma} \tilde{Y}_{mn}^{\sigma}(\theta, \varphi) j^m (\tilde{\Lambda}_{1mn}^{\sigma} j_m(kr) + j \tilde{\Lambda}_{2mn}^{\sigma} n_m(kr)) \quad (D.6)$$

Dans le cas où l'on considère un domaine exempt de sources on a $\tilde{\Lambda}_{2mn}^{\sigma} = 0$. [Daniel, 2000, p. 302]. Le cas contraire fait apparaître les fonctions de Hankel convergentes ou divergentes suivant le sens de propagation des ondes. Dans notre cas, on considère le domaine exempt de sources et l'équation devient donc :

$$p(r, \theta, \varphi) = \sum_{m,n,\sigma} \tilde{Y}_{mn}^{\sigma}(\theta, \varphi) j^m \tilde{\Lambda}_{mn}^{\sigma} j_m(kr) \quad (D.7)$$

BIBLIOGRAPHIE

- [Adriaensen, 2006] ADRIAENSEN, F. (2006). Near field filters for higher order ambisonics.
- [Blanchet et Charbit, 2001] BLANCHET, G. et CHARBIT, M. (2001). Signaux et images sous matlab. *HERMES Science Europe Ltd.*
- [Bruneau, 1998] BRUNEAU, M. (1998). *Manuel d'acoustique fondamentale*. Hermes.
- [Chaigne et al., 2008] CHAIGNE, A., KERGOMARD, J. et al. (2008). *Acoustique des instruments de musique*. Belin.
- [Corteel, 2006] CORTEEL, E. (2006). Equalization in an extended area using multichannel inversion and wave field synthesis. *Journal of the audio Engineering Society*, 54(12):1140–1161.
- [Daniel, 2000] DANIEL, J. (2000). *Représentation de champs acoustiques, application à la transmission et à la restitution de scènes sonores complexes dans un contexte multimédia*. Thèse de doctorat.
- [Daniel, 2003] DANIEL, J. (2003). Spatial sound encoding including near field effect : Introducing distance coding filters and a viable, new ambisonic format. In *Audio Engineering Society Conference : 23rd International Conference : Signal Processing in Audio Recording and Reproduction*.
- [Daniel et Moreau, 2004] DANIEL, J. et MOREAU, S. (2004). Further study of sound field coding with higher order ambisonics. In *Audio Engineering Society Convention 116*.
- [Daniel et al., 1998] DANIEL, J., RAULT, J.-B. et POLACK, J.-D. (1998). Ambisonics encoding of other audio formats for multiple listening conditions. In *Audio Engineering Society Convention 105*.
- [Deprez et al., 2010] DEPREZ, R., NICOL, R. et FRIOT, E. (2010). Validation théorique de la correction des réflexions sur la base d'une représentation en harmoniques sphériques. *10ème Congrès Français d'Acoustique*.
- [Farina, 2000] FARINA, A. (2000). Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *Audio Engineering Society Convention 108*.
- [Fielder, 2001] FIELDER, L. D. (2001). Practical limits for room equalization. In *Audio Engineering Society Convention 111*.
- [Gauthier et Berry, 2007] GAUTHIER, P.-A. et BERRY, A. (2007). Adaptive wave field synthesis for sound field reproduction : Theory, experiments, and future perspectives. In *Audio Engineering Society Convention 123*.
- [Gerzon, 1975] GERZON, M. A. (1975). The design of precisely coincident microphone arrays for stereo and surround sound. In *Audio Engineering Society Convention 50*.
- [Gerzon, 1985] GERZON, M. A. (1985). Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society*, 33(11):859–871.
- [Gerzon, 1992] GERZON, M. A. (1992). General metatheory of auditory localisation. In *Audio Engineering Society Convention 92*.
- [Guillaume et al., 2005] GUILLAUME, M., GRENIER, Y. et RICHARD, G. (2005). Iterative algorithms for multichannel equalization in sound reproduction systems. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, volume 3, pages iii–269. IEEE.
- [Harris, 1978] HARRIS, F. J. (1978). On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–83.
- [Kirkeby et Nelson, 1999] KIRKEBY, O. et NELSON, P. A. (1999). Digital filter design for inversion problems in sound reproduction. *J. Audio Eng. Soc.*, 47(7/8):583–595.
- [Kirkeby et al., 1998] KIRKEBY, O., NELSON, P. A., HAMADA, H. et ORDUNA-BUSTAMANTE, F. (1998). Fast deconvolution of multichannel systems using regularization. *Speech and Audio Processing, IEEE Transactions on*, 6(2):189–194.
- [Miyoshi et Kaneda, 1988] MIYOSHI, M. et KANEDA, Y. (1988). Inverse filtering of room acoustics. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(2):145–152.

- [Moreau, 2006] MOREAU, S. (2006). *Étude et réalisation d'outils avancés d'encodage spatial pour la technique de spatialisation sonore Higher Order Ambisonics : microphone 3D et contrôle de distance*. Thèse de doctorat, Université du Maine.
- [Morse et Ingard, 1968] MORSE, P. M. et INGARD, K. U. (1968). *Theoretical acoustics*. Princeton University Press.
- [Mourjopoulos et Paraskevas, 1991] MOURJOPOULOS, J. et PARASKEVAS, M. (1991). Pole and zero modeling of room transfer functions. *Journal of Sound and Vibration*, 146(2):281–302.
- [Müller et Massarani, 2001] MÜLLER, S. et MASSARANI, P. (2001). Transfer-function measurement with sweeps. *Journal of the Audio Engineering Society*, 49(6):443–471.
- [Neely et Allen, 1979] NEELY, S. T. et ALLEN, J. B. (1979). Invertibility of a room impulse response. *The Journal of the Acoustical Society of America*, 66:165.
- [Nicol, 1999] NICOL, R. (1999). *Restitution sonore spatialisée sur une zone étendue : application à la téléprésence*. Thèse de doctorat.
- [Noisternig et al., 2013] NOISTERNIG, M., CARPENTIER, T. et WARUSFEL, O. (2013). Dispositif de spatialisation sonore 3d à l'espace de projection de l'ircam - un réseau de 345 haut-parleurs pour une restitution par wfs et hoa. *Acoustique & Techniques*, 71:30–39.
- [Noll, 1967] NOLL, A. M. (1967). Cepstrum pitch determination. *The journal of the acoustical society of America*, 41:293.
- [Norcross et al., 2002] NORCROSS, S. G., SOULODRE, G. A. et LAVOIE, M. C. (2002). Evaluation of inverse filtering techniques for room/speaker equalization. *In Audio Engineering Society Convention 113*.
- [Oppenheim et Schaffer, 1975] OPPENHEIM, A. V. et SCHAFER, R. W. (1975). *Digital signal processing*. Prentice-Hall, Englewood Cliffs, NJ.
- [Oppenheim et Schaffer, 2004] OPPENHEIM, A. V. et SCHAFER, R. W. (2004). From frequency to quefrequency : A history of the cepstrum. *Signal Processing Magazine, IEEE*, 21(5):95–106.
- [Poletti, 2000] POLETTI, M. A. (2000). A unified theory of horizontal holographic sound systems. *Journal of the Audio Engineering Society*, 48(12):1155–1182.
- [Spors et al., 2003] SPORS, S., KUNTZ, A. et RABENSTEIN, R. (2003). An approach to listening room compensation with wave field synthesis. *In Audio Engineering Society Conference : 24th International Conference : Multichannel Audio, The New Reality*. Audio Engineering Society.
- [Stefanakis et al., 2010] STEFANAKIS, N., JACOBSEN, F. et SARRIS, J. (2010). Effort variation regularization in sound field reproduction. *The Journal of the Acoustical Society of America*, 128:740.
- [Tokuno et al., 1997] TOKUNO, H., KIRKEBY, O., NELSON, P. A. et HAMADA, H. (1997). Inverse filter of sound reproduction systems using regularization. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 80(5):809–820.
- [Torger,] TORGER, A. Brutefir : a software for applying long fir filters to multi-channel digital audio. <http://www.ludd.luth.se/~torger/brutefir.html>.
- [Zotter et al., 2010] ZOTTER, F., POMBERGER, H. et NOISTERNIG, M. (2010). Ambisonic decoding with and without mode-matching : A case study using the hemisphere. *In 2nd Ambisonics Symposium, Paris*.