

École doctorale Sciences des Métiers de l'Ingénieur (Paris)

Laboratoire de Mécanique de Structures et des Systèmes Couplés

Université de Sherbrooke - Groupe d'Acoustique de l'Université de Sherbrooke

THÈSE DE DOCTORAT

présentée par : **Pierre LECOMTE**

soutenue le : **7 décembre 2016**

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers**

Discipline : **Mécanique**

Ambisonie d'ordre élevé en trois dimensions : captation, transformations et décodage adaptatif de champs sonores

THÈSE dirigée par

M. GARCIA Alexandre
M. BERRY Alain

Professeur des Universités, CNAM
Professeur titulaire, Université de Sherbrooke

RAPPORTEURS

M. HERZOG Philippe
Mme. NICOL Rozenn

Directeur de recherche, LMA-CNRS
Ingénieure de recherche, Orange Labs

PRÉSIDENT

M. PARIZET Étienne

Professeur des Universités, INSA Lyon

EXAMINATEURS

M. LEFEBVRE Roch
M. GAUTHIER Philippe-Aubert
M. LANGRENNE Christophe

Professeur titulaire, Université de Sherbrooke
Professionnel de recherche, Université de Sherbrooke
Ingénieur de recherche, CNAM

RÉSUMÉ

La synthèse de champs sonores est un domaine de recherche actif trouvant de nombreuses applications musicales, multimédias ou encore industrielles. Dans ce dernier cas, la reconstruction précise du champ sonore est souhaitée, ce qui implique de répondre à un certain nombre de questionnements scientifiques. À l'aide de réseaux de microphones et de haut-parleurs, la captation, la synthèse et la reconstruction précise de champs sonores sont théoriquement possibles. Seulement, pour des applications pratiques, la disposition des haut-parleurs et l'influence acoustique du lieu de restitution sont des facteurs cruciaux à prendre en compte pour s'assurer de la bonne reconstruction du champ sonore.

Dans ce contexte, cette thèse de doctorat propose des méthodes et des techniques pour la captation, la transformation et la reconstruction précise de champs sonores en trois dimensions en se basant sur la méthode ambisonique d'ordre élevé. Une configuration sphérique pour le réseau de microphones et de haut-parleurs est proposée. Elle suit un maillage de Lebedev à cinquante points qui permet la captation et la reconstruction du champ sonore jusqu'à l'ordre 5 avec le formalisme ambisonique. Les limitations de cette approche, tel le repliement spatial, sont étudiés en détails. De plus, une opération de transformation du champ sonore est présentée. Elle est établie dans le domaine des harmoniques sphériques et permet d'effectuer un filtrage directionnel avant le décodage pour privilégier certaines directions dans le champ sonore, suivant une fonction de directivité choisie. Pour la reconstruction, une approche originale, également établie dans le domaine des harmoniques sphériques, permet de prendre en compte l'influence acoustique du lieu de restitution, ainsi que les défauts du système de restitution. Ce traitement permet alors d'adapter la synthèse de champs sonores au lieu de restitution, en conservant le formalisme théorique établi en champ libre. Finalement, une validation expérimentale des méthodes et des techniques développées au cours de la thèse est faite. Dans ce contexte, une suite logicielle de synthèse et traitement en temps-réel des champs sonore est développée.

Mots-clés : acoustique, audio spatialisée, ambisonie d'ordre élevé, traitement du signal multicanaux

ABSTRACT

Sound field synthesis is an active research domain with various musical, multimedia or industrial applications. In the latter case, the accurate reconstruction of the sound field is targeted, which involves answering several scientific questions. Using arrays of microphones and loudspeakers, the capture, synthesis and accurate reconstruction of sound fields are theoretically possible. However, for practical applications, the arrangement of the loudspeakers and the acoustic influence of the restitution room are critical factors to consider in order to ensure the accurate reconstruction of the sound field.

In this context, this thesis proposes methods and techniques for the capture, transformations and accurate reconstruction of sound fields in three dimensions based on the Higher Order Ambisonics (HOA) method. A spherical configuration for the array of microphones and loudspeakers is proposed. It follows a fifty-node Lebedev grid that enables the capture and reconstruction of the sound field up to order 5 with HOA formalism. The limitations of this approach, such as the spatial aliasing, are studied in detail. A transformation operation of the sound field is also proposed. The formulation is established in the spherical harmonics domain and enables a directional filtering on the sound field prior to the decoding step. For the reconstruction of the sound field, an original approach, also established in the spherical harmonics domain, can take into account the acoustic influence of the restitution room and the defects of the playback system. This treatment then adapts the synthesis of sound fields to the restitution room, maintaining the theoretical formalism established in free field. Finally, an experimental validation of methods and techniques developed in the thesis is made. In this context, a digital signal processing toolkit is developed. It process in real-time the microphones, ambisonics, and loudspeaker signals for the sound field capture, transformations, and decoding.

Keywords: acoustic, spatial audio, higher order ambisonics, multichannel signal processing

REMERCIEMENTS

Sur un doctorat de trois années, en cotutelle, on rencontre du monde! C'est alors difficile de remercier sans oublier toutes celles et ceux qui ont contribué de près ou de loin à ce projet.

Voici ma tentative, évidemment non exhaustive :

Ça démarre naturellement par mon équipe encadrante : Alain, Alexandre, Philippe-Aubert et Christophe. Merci à vous quatre pour m'avoir suivi durant ces trois années et m'avoir apporté vos expériences, idées, corrections. J'ai appris énormément de vous sur le plus beau métier du monde : la recherche! J'ai apprécié tout particulièrement la grande liberté et confiance que vous m'avez accordé sur le projet.

Merci à Rozenn Nicol, Étienne Parizet, Roch Lefebvre, Philippe Herzog d'avoir accepté de faire partie de mon jury.

Merci à Alexandre et Alain de m'avoir fait confiance et avoir financé mes petits bricolages de Lebedev. Un coût modique je disais... Évidemment, de petites factures en petites factures ça a finalement donné un prix plus si modique que ça. D'ailleurs, à ce propos, je tiens également à remercier ceux qui m'ont donné des coups de mains indispensables à la réalisation des deux sphères! Pour le côté français, merci à Sarah Poirée, technicienne au LMSSC pour m'avoir aidé à emprunter ce bout de gouttière (abandonné je précise) pour commencer ce qui deviendra les anneaux de la SpherBedev 1.0, Merci également pour tous les coups de mains pour la construction de cette dernière, les aller-retour chez TARTAIX!! Merci également à Frédéric Guillermin, technicien au LMSSC, pour les usinages des pièces et les conseils techniques. À Eric Bavu pour ses très bon choix de matériel! Pour le côté québécois, un très grand merci à Patrick Lesveque, pour sa très grande expérience et sa réalisation des pièces de la SpherBedev 2.0! Pour le réseau de microphones MemBedev, merci énormément aux commanditaires, Alexandre et Christophe, et au constructeur Philippe Chenevez, magicien de l'électronique pour la réalisation de cette petite merveille en impression 3D! Merci à toute l'équipe du LMSSC et du GAUS où je me suis épanoui professionnellement mais aussi humainement durant ce doctorat : vive, entre autre, l'attirail et le 5@8.

Merci à tous mes colocos français, québécois, iranien, argentin, mexicain, de la maison bleue à Sherbrooke, pour ces belles rencontres de vie : Mohsen, Chloé, Geneviève, Jorge, Rocío, Lou, Ehsan, Roxanne. Merci également à mes colocos de Paris Adrien et Augustin pour votre folie!

Merci à mes potes d'ici et de là : Télina, Marc, Anthony, Gauthier, Stan, Alex, Léa, Antoine, Poli, et tout ceux et celles que je ne cite pas ici.

Merci à mes parents et ma soeur pour leur soutien, à mon père pour son expérience de professeur d'université, mais aussi de grand bricoleur!

La dernière année, on se rend compte qu'il faut finir, et évidemment on est en retard dans tout ce qu'on souhaiterai encore faire! Merci à Samuel Dupont pour ton coup de main providentiel lors de ton stage M2, notamment pour la campagne expérimentale présentée à la fin de la thèse.

Et enfin, merci Rocío pour ton soutien à la fin du doc, pour venir me chercher à point d'heure à l'université, pour ta patience...

TABLE DES MATIÈRES

1 INTRODUCTION	1
1.1 Contexte général	1
1.2 Problématique et objectifs	3
1.3 Contributions originales	5
1.4 Méthodologie et organisation du document	7
2 ÉTAT DE L'ART	11
2.1 Restitution du son spatialisé	11
2.2 Ambisonie d'ordre élevé	13
2.2.1 Aperçu de la méthode	13
2.2.2 Système de coordonnées	15
2.2.3 Produit scalaire sphérique	15
2.2.4 Harmoniques sphériques	16
2.2.5 Transformée de Fourier sphérique	19
2.2.6 Fonctions de Bessel et de Hankel sphériques	19
2.2.7 Décomposition d'un champ de pression acoustique sur les harmoniques sphériques	21
2.2.8 Encodage d'une onde plane et d'une onde sphérique	22
2.2.9 Captation et encodage d'un champ sonore	23
2.2.10 Troncature	27
2.2.11 Transformations	30
2.2.12 Décodage	34
2.2.13 Exemple en simulation	39
2.3 Synthèse de fronts d'ondes	41
2.3.1 Présentation générale	41
2.3.2 Principe de Huygens	41
2.3.3 Intégrale de Kirchhoff-Helmholtz	42
2.3.4 Élimination du terme dipolaire	43
2.3.5 Synthèse de fronts d'ondes en 2.5 dimensions	44
2.3.6 Sélection des sources secondaires	44
2.3.7 Sources focalisées	45
2.3.8 Artefacts de reproduction	45
2.3.9 Synthèse de fronts d'ondes adaptative	45
2.3.10 La synthèse de front d'ondes comme approximation hautes-fréquences de l'ambisonie	46
2.3.11 Exemple en simulation	47
2.3.12 Conclusions	48
3 UN MAILLAGE DE LEBEDEV A CINQUANTE NOEUDS ET SES APPLICATIONS À L'AMBISONIE	51
3.1 Avant-propos	51

3.2	Contribution au document	52
3.3	A Fifty-Node Lededev Grid and Its Applications to Ambisonics	53
3.4	Abstract	53
3.5	Introduction	53
3.6	Definitions and Notations	55
3.6.1	Spherical Coordinate System	55
3.6.2	Spherical Harmonics	55
3.6.3	Spherical Scalar Product	56
3.6.4	Spherical Fourier Transform	56
3.7	Ambisonics	56
3.7.1	Sound Field Representation	56
3.7.2	Truncated Decomposition and Truncation Error	58
3.7.3	Sound Field Capture	59
3.7.4	Sound Field Reproduction With Continuous Secondary Source Dis-	
	tributions	60
3.8	Discretization	61
3.8.1	Quadrature Rules	61
3.9	Rotationally-Invariant Quadratures for the Sphere	62
3.9.1	Rotationally-Invariant Quadrature	63
3.9.2	Sobolev's Theorems	63
3.9.3	Solving the Problem	65
3.9.4	Lebedev Quadrature Rule	66
3.9.5	Maximum Order of Integration for Ambisonics Purposes	68
3.10	Applications to Capture, Encoding, Decoding and Panning Law for Am-	
	bisonics	68
3.10.1	Capture, Encoding and Decoding	68
3.10.2	Encoding and Decoding of Plane or Spherical waves, Panning Law	69
3.11	Orthonormality Error	70
3.12	Aliasing Considerations	71
3.12.1	Aliasing at Reconstruction	71
3.12.2	Aliasing at Capture	74
3.13	Multi-Band Multi-Order Decoders	75
3.14	Conclusions	79
4	COMPENSATION DES RÉFLEXIONS DANS UNE SALLE SUR UNE	
	ZONE ÉTENDUE PAR UNE APPROCHE AMBISONIQUE	81
4.1	Avant-propos	81
4.2	Contribution au document	82
4.3	Cancellation Of Room Reflections Over An Extended Area Using Ambisonics	83
4.4	Abstract	83
4.5	Introduction	83
4.6	Definitions and notations	85
4.6.1	Spherical harmonics	85
4.6.2	Spherical Fourier Transform	86
4.6.3	Matrix and vectors notations	86

4.7	Ambisonics	86
4.7.1	Sound field representation for the interior problem	86
4.7.2	Truncation of the spherical harmonic expansion	87
4.7.3	Decoding	88
4.8	Cancellation of room reflections	90
4.8.1	Re-encoding in a room	90
4.9	Room simulation	93
4.10	Condition-number of the room reflections cancellation problem at low frequencies	96
4.11	Crossover, cross-order room reflections cancellation filters	97
4.12	Results of simulations	100
4.12.1	Frequency domain	100
4.12.2	Time domain	104
4.13	Practical considerations	108
4.14	Three-dimensional reverberation effect	111
4.15	Conclusions	112
5	FILTRAGE DIRECTIONNEL DANS UNE SCÈNE SONORE AMBI-	
	SONIQUE	113
5.1	Introduction	113
5.2	Formation de voies avec un réseau de microphones sphérique	114
5.3	Filtrage angulaire d'une scène sonore ambisonique	114
5.3.1	Fonction angulaire explicite	115
5.3.2	Dirac angulaire	116
5.3.3	Approche utilisant la décomposition de la fonction de directivité g	117
5.4	Simulations	119
5.4.1	Fonction hyper-cardioïde d'ordre $M' = 1$	119
5.4.2	Dirac angulaire	121
5.5	Mise en oeuvre temps-réel	121
5.6	Conclusions	122
6	DISPOSITIFS EXPÉRIMENTAUX ET MESURES DE CHAMPS SO-	
	NORES	123
6.1	Introduction	123
6.2	Sphère de restitution ambisonique « SpherBedev »	123
6.2.1	Configuration géométrique	124
6.2.2	Enceintes de reproduction	127
6.2.3	Structure	130
6.2.4	Chaînes audio-numériques	134
6.3	Antenne sphérique de microphone « MemsBedev »	135
6.3.1	Configuration géométrique	135
6.3.2	Support sphérique rigide	136
6.3.3	Capsules microphoniques	137
6.3.4	Bornier de connexions	137
6.3.5	Estimation des composantes ambisoniques	139

6.4	Campagne de mesure en chambre anéchoïque	142
6.4.1	Objectifs	142
6.4.2	Protocole expérimental	144
6.4.3	Antenne plane de microphones	144
6.4.4	Signaux d'excitation	147
6.4.5	Positionnement	149
6.4.6	Calibration	150
6.4.7	Post-traitement des signaux de mesures	151
6.4.8	Source réelles et virtuelles	154
6.4.9	Champ sonore avec filtrage directionnel	155
6.4.10	Conclusions	178
7	CONCLUSIONS	179
7.1	Principaux résultats	179
7.2	Perspectives	181
A	DIFFRACTION D'UNE SPHÈRE RIGIDE DANS UN CHAMP ACOUS-	
	TIQUE	183
B	FORMULATION SIMPLE SOURCE POUR UNE GÉOMÉTRIE SPHÉ-	
	RIQUE	185
B.1	Formulation simple source	185
B.2	Application à une géométrie sphérique	186
C	SYNTHÈSE TEMPS-RÉEL DE CHAMPS SONORES EN TROIS DI-	
	MENSIONS EN UTILISANT FAUST, PROCESSING, PURE DATA	
	ET OSC	191
C.1	Avant-propos	191
D	EXTRAIT DE LA DOCUMENTATION D'AMBITOOLS	201
D.1	Avant-propos	201
D.2	Introduction	201
D.2.1	Goals of <i>ambitools</i>	201
D.3	Installing <i>ambitools</i>	202
D.3.1	Retrieve <i>ambitools</i> repository	202
D.3.2	Dependencies	203
D.3.3	Compile the FAUST plug-ins	203
D.3.4	Local FAUST installation	203
D.3.5	FAUSTLIVE	203
D.3.6	Compile the PROCESSING VU-Meter	203
D.4	The different tools	204
D.4.1	FAUST	204
D.4.2	Processing	218
D.4.3	Jconvolver	220
D.4.4	Pure Data	223
D.5	Examples	223

TABLE DES MATIÈRES

xi

LISTE DES RÉFÉRENCES

225

LISTE DES FIGURES

2.1	Principe général de la méthode ambisonique.	14
2.2	Système de coordonnées sphériques.	15
2.3	Illustration des harmoniques sphériques jusqu'à l'ordre $M = 5$.	18
2.4	Fonctions de Bessel sphériques de la première espèce pour les ordres $m = 1$ à 5.	20
2.5	Gain des filtres de champ proches pour les ordres $m = 1$ à 5 pour l'encodage ambisonique d'une onde sphérique.	24
2.6	Les capsules d'un microphone SoundField™ MKV.	24
2.7	Gain des filtres d'égalisation E_m pour une sphère rigide de rayon $r_{\text{mic}} = 7$ cm, pour les ordres $m = 1$ à 5.	27
2.8	Exemples de réseaux de microphones sphériques répartis sur une sphère rigide.	28
2.9	Erreur de troncature moyennée $\bar{\epsilon}_M$ pour les ondes planes et différents ordres de troncature.	29
2.10	Partie réelle du champ de pression issu d'un monopôle à la position $\mathbf{r}_s = (2 \text{ m}, 0^\circ, 0^\circ)$ aux fréquences 500 Hz (colonne de gauche), 1000 Hz (colonne du milieu) et 1500 Hz (colonne de droite).	31
2.11	Partie réelle du champ de pression émis par un monopôle (en rouge).	40
2.12	Le principe de Huygens appliqué à un front d'onde.	41
2.13	Partie réelle du champ de pression émis par un monopôle (en rouge).	49
3.1	Spherical coordinate system in use.	55
3.2	G_8 transformations for the octahedron, $D = 24$.	65
3.3	The different types of nodes $a^{(1)}, a^{(2)}, a^{(3)}$ and $b^{(1)}$ for a Lebedev quadrature up to order $M' = 11$, shown in a single octant.	67
3.4	Orthonormality error matrices $10 \log_{10} (\mathbf{D}_{ij} _{1 \leq i \leq (M+1)^2, 1 \leq j \leq (M+1)^2})$ up to 10 th order and corresponding grids.	71
3.5	Average aliasing error $\bar{\epsilon}_{\text{alias}, M}$ for a unit amplitude point source at radius $r_s = 2$ m and frequency $f = 1000$ Hz.	73
3.6	Total average aliasing error at capture for plane waves for three grids.	75
3.7	Several orders decoders on the same Lebedev grid with $L = 50$ loudspeakers.	77
3.8	Real part of the target, truncated and reproduced sound field for a multi-frequency case.	78
4.1	Spherical coordinate system in use.	85
4.2	The Ambisonic playback system in a reflective parallelepipedic room.	94
4.3	(a): Magnitude of the frequency response function (B_{00}) (dB with reference 1) of the loudspeaker 2 (1.07 m, 0, 0) in the room. (b): Corresponding impulse response.	95
4.4	Condition number of matrix \mathbf{T} versus frequency for several truncation orders M .	96
4.5	Crossover cross-order filter design for a maximum truncation order $M = 5$.	99

4.6	(a): Impulse response of the filter $\mathbf{H}_{5,5}$. (b) Impulse response of the filter $\mathbf{H}_{25,18}$.	100
4.7	Real parts of sound pressure fields at 125 Hz in the horizontal plane, for a monopole at location $(r_s = 2 \text{ m}, \theta_s = 180^\circ, \delta_s = 0^\circ)$ with unit amplitude $S = 1$.	102
4.8	Real parts of sound pressure fields at 500 Hz in the horizontal plane, for a monopole at location $(r_s = 2 \text{ m}, \theta_s = -180^\circ, \delta_s = 0^\circ)$ with unit amplitude $S = 1$.	103
4.9	Sound pressure fields in the horizontal plane at $t \simeq 0.347 \text{ s}$.	105
4.10	Magnitude of sound pressure fields in the horizontal plane at $t \simeq 0.351 \text{ s}$.	107
4.11	Sound pressure at two points as function of time.	109
4.12	Sound pressure at two points as function of time (Fig. 4.11): frequency content.	110
5.1	Directivité obtenue pour le cas hyper-cardioïde d'ordre 1.	118
5.2	Partie réelle du champ de pression acoustique d'une onde plane.	120
5.3	Diagrammes de directivité dans le plan horizontal ($\delta = 0$).	121
5.4	Partie réelle du champ de pression acoustique filtré par \mathbf{M} , lorsque g est défini par un Dirac angulaire de direction $(20^\circ, 0^\circ)$.	122
6.1	Les 50 noeuds numérotés du maillage de Lebedev utilisé.	128
6.2	Détails d'une enceinte de la SpherBedev 1.0.	128
6.3	Amplitude normalisée de la FRF d'une enceinte de la SpherBedev 1.0 dans l'axe à 1 m, mesurée en chambre anéchoïque.	129
6.4	Directivité d'une enceinte de la SpherBedev 1.0 à différentes fréquences. Axe du haut-parleur à 0° .	129
6.5	Détails du montage de quelques enceintes de la SpherBedev 2.0.	131
6.6	FRF d'une enceinte de la SpherBedev 2.0 dans l'axe à 1 m, mesurée en chambre anéchoïque.	132
6.7	Détails de la structure pour la SpherBedev 2.0.	133
6.8	Sphère de restitution ambisonique jusqu'à l'ordre 5 SpherBedev.	134
6.9	Bâti de convertisseurs et d'amplificateurs de la SpherBedev 2.0.	136
6.10	Support en plastique de l'antenne MemsBedev réalisé en impression 3D.	138
6.11	Capsules microphoniques du MemsBedev.	139
6.12	Bornier de connexion du MemsBedev.	139
6.13	MemsBedev installé sur un pied de micro.	140
6.14	Gain des filtres d'égalisation de l'antenne MemsBedev de rayon $r_{\text{mic}} = 7 \text{ cm}$.	141
6.15	Dispositif expérimental installé dans la chambre anéchoïque du GAUS.	146
6.16	Configuration de l'antenne plane de microphones.	147
6.17	Positionnement des antennes de microphones dans la SpherBedev 2.0.	150
6.18	Calibration des réseaux de microphones.	151
6.19	Partie réelle des champs de pression à 500 Hz.	156
6.20	Partie réelle des champs de pression à 750 Hz.	157
6.21	Partie réelle des champs de pression à 1000 Hz.	158
6.22	Partie réelle des champs de pression à 1250 Hz.	159

6.23	Partie réelle des champs de pression à 1500 Hz.	160
6.24	Partie réelle des champs de pression à 1750 Hz.	161
6.25	Partie réelle des champs de pression à 2000 Hz.	162
6.26	Partie réelle des champs de pression à 750 Hz.	163
6.27	Partie réelle des champs de pression à 1000 Hz.	164
6.28	Partie réelle des champs de pression à 1250 Hz.	165
6.29	Partie réelle des champs de pression à 1500 Hz.	166
6.30	Partie réelle des champs de pression à 1750 Hz.	167
6.31	Partie réelle des champs de pression à 2000 Hz.	168
6.32	Partie réelle des champs de pression à 500 Hz.	171
6.33	Partie réelle des champs de pression à 750 Hz.	172
6.34	Partie réelle des champs de pression à 1000 Hz.	173
6.35	Partie réelle des champs de pression à 1250 Hz.	174
6.36	Partie réelle des champs de pression à 1500 Hz.	175
6.37	Partie réelle des champs de pression à 1750 Hz.	176
6.38	Partie réelle des champs de pression à 2000 Hz.	177
B.1	Configuration du problème pour l'approche par simple source.	185
D.1	FAUSTLIVE Export Manager.	204
D.2	hoa_encoder N sources plug-in.	205
D.3	hoa_mic_encoder_lebedev26 plug-in.	206
D.4	Connections in order to obtain the Ambisonics components from the signals of a rigid spherical microphone.	207
D.5	hoa_decoder_lebedev50 plug-in.	209
D.6	hoa_decoder_lebedev50_binaural plug-in.	210
D.7	hoa_panning N sources lebedev50 plug-in.	211
D.8	Connections between plug-ins hoa_encoder N sources, hoa_mirroring and hoa_decoder_lebedev50.	212
D.9	hoa_mirroring plug-in.	213
D.10	hoa_mirroring effects on a sound field.	213
D.11	hoa_azimuth_rotator plug-in.	214
D.12	hoa_azimuth_rotator effects on a sound field.	214
D.13	hoa_beamforming_hypercardioid_to_mono plug-in.	215
D.14	Regular hyper-cardioid beampatterns up to order 3.	215
D.15	hoa_beamforming_hypercardioid_to_hoa plug-in.	216
D.16	The effects of hoa_beamforming_hypercardioid_to_hoa on a sound field.	217
D.17	hoa_beamforming_dirac_to_hoa plug-in.	218
D.18	hoa_beamforming_dirac_to_hoa effects on a sound field.	219
D.19	Spherical_VU_Meter for a 50-loudspeaker Lebedev array and two virtual sources.	220
D.20	Connection between hoa_decoder_lebedev50 tool and <i>Jconvolver</i> for bin- aural rendering.	222
D.21	Scenshot of application andOSC on Android.	224
D.22	Graphical user interface of patch andOSC.pd.	224

LISTE DES TABLEAUX

4.1	Wall absorption coefficients for the room simulation and Sabine's estimated reverberation time in octave band.	93
5.1	Coefficients c_m pour un directivité de type hyper-cardioïde.	118
6.1	Types de points utilisés pour le maillage de la sphère.	125
6.2	Coordonnées cartésiennes et sphériques des différents noeuds de la Sphère-Bedev.	127
6.3	Détails des mesures du protocole expérimental.	145
D.1	Dependencies of <i>ambitools</i> .	203

LISTE DES ACRONYMES

Acronyme	Définition
ACN	<i>Ambisonic Channel Number</i> : Numéro de canal ambisonique
ADAT	<i>Alesis Digital Audio Tape</i>
CNAM	Conservatoire National des Arts et Métiers
DAW	<i>Digital Audio Workstation</i> : station de travail audio-numérique
DSFT	<i>Discrete Spherical Fourier Transform</i> : TFSD
FIR	<i>Finite Impulse Response</i> : RIF
FRF	Fonction de Réponse en Fréquence
GAUS	Groupe d'Acoustique de l'Université de Sherbrooke
HOA	<i>Higher Order Ambisonics</i> : Ambisonie d'ordres élevés
IIR	<i>Infinite Impulse Response</i> : RII
JACK	<i>Jack Audio Connection Kit</i> : serveur de son sous Linux
LMSSC	Laboratoire de Mécanique des Structures et des Systèmes Couplés
MADI	<i>Multichannel Audio Engineering Interface</i>
MEMS	<i>Micro Electro-Mechanical System</i> : MicroSystème Électro-Mécanique
NFC-HOA	<i>Near Field Compensated - Higher Order Ambisonics</i> : Ambisonie d'ordre élevé avec compensation du champ proche
RMS	<i>Root Mean Square</i> : valeur efficace
RIF	Réponse Impulsionnelle Finie
RII	Réponse Impulsionnelle Infinie
TFS	Transformée de Fourier Sphérique
TFSD	Transformée de Fourier Sphérique Discrète
SFT	<i>Spherical Fourier Transform</i> : TFS
UdeS	Université de Sherbrooke
UN	Unité numérique
VBAP	<i>Vector Base Amplitude Panning</i> : Loi panoramique d'amplitude basée sur une formulation vectorielle.
WFS	<i>Wave Field Synthesis</i> : synthèse de fronts d'ondes

CHAPITRE 1

INTRODUCTION

1.1 Contexte général

La reproduction du son est un domaine des plus actuels, où le contexte multimédia est de plus en plus omniprésent. Depuis que la technologie permet d'enregistrer et de reproduire des événements sonores, les systèmes n'ont cessés de s'améliorer pour donner aux auditeurs de plus en plus de réalisme et de fidélité à ce qu'ils écoutent.

Mais au-delà des applications musicales ou de divertissement, la reproduction du son est aussi devenue un enjeu industriel important. En effet, le son et le bruit sont présents au quotidien et leur contrôle est primordial, ne serait-ce que dans un contexte de santé publique. Mais est-il possible de contrôler la reproduction sonore finement, de pouvoir recréer avec précision un champ sonore qui aurait été enregistré ou bien synthétisé de toute pièces ? Si tel est le cas, les applications seraient alors nombreuses pour la reproduction sonore dans un contexte d'ingénierie ou d'acoustique industrielle : confort acoustique, études acoustiques de lieux ou de structures virtuelles, aide aux personnes à déficience visuelle, etc.

La problématique d'enregistrement, de synthèse et de reproduction de champs sonores est un domaine de recherche actif. Elle trouve des solutions analytiques avec, entre autres, l'ambisonie d'ordre élevés, (*Higher Order Ambisonics*, *HOA*) et la synthèse de fronts d'ondes, (*Wave Field Synthesis*, *WFS*). L'ambisonie est apparue avec les travaux de Gerzon dans les années 70 [Gerzon, 1973] puis a été étendue aux ordres élevés dans les années 2000, notamment par Daniel [Daniel, 2000]. La synthèse de fronts d'ondes, quant à elle, est apparue dans les années 1980 avec les travaux de Berkhout [Berkhout, 1988; Berkhout et al., 1993]. Ces deux approches permettent de reproduire en deux ou en trois dimensions des champs sonores sur une zone étendue de l'espace. Elles ont longtemps évolué séparément, de par leur formulations distinctes : l'ambisonie se base sur une décomposition du champ acoustique sur une base d'harmoniques cylindriques ou sphériques. La synthèse de fronts d'ondes, quant à elle, se base sur la formulation intégrale de Kirchhoff-Helmholtz. Cependant, des parallèles entre ces deux techniques ont été établis récemment [Ahrens et Spors, 2012; Ahrens et al., 2010; Daniel et al., 2003; Fazi et Nelson, 2013; Nicol et Emerit, 1999; Spors et Ahrens, 2008]. Dans les deux cas, les systèmes de restitution nécessaires

à la reconstruction du champ sonore requièrent un grand nombre de haut-parleurs, ainsi que d'importantes ressources matérielles et de calcul. Ainsi, il n'a été possible de voir apparaître que récemment des systèmes de restitution possédant plusieurs centaines de haut-parleurs, voir par exemple [Noisternig *et al.*, 2013].

Les applications multimédia de la synthèse de champs sonores ne requièrent pas que le champ de pression reconstruit soit physiquement identique au champ de pression cible. La reproduction des bons indices de localisation pour l'auditeur est déjà suffisante pour recréer une immersion sonore [Guastavino et Katz, 2004]. Ceci laisse une marge de manoeuvre quant à la précision physique de la reproduction pour le cas précis de ces applications. Cependant, en contexte d'ingénierie, une précision physique de reproduction est visée, ce qui est plus difficile à atteindre. Par exemple, il est facile de concevoir que le nombre de haut-parleurs et leur arrangement dans l'espace sont deux facteurs déterminants pour la bonne reconstruction du champ sonore (fondée sur la superposition des fronts d'onde des haut-parleurs) [Poletti, 2005; Zotter, 2009b; Zotter *et al.*, 2010].

À la lumière de ces contraintes d'applications spécifiques, certaines problématiques scientifiques furent déjà considérées mais mériteraient d'être plus approfondies pour ces applications en trois dimensions. Ainsi, premièrement, le repliement spatial dû à l'échantillonnage avec des haut-parleurs de la surface de source de reproduction, fut souvent étudié pour les configurations à deux dimensions, mais peu pour des configurations à trois dimensions [Zotter *et al.*, 2009]. Deuxièmement, les formalismes, autant pour l'ambisonie que pour la synthèse de fronts d'ondes, sont établis en champ libre, c'est-à-dire sans aucune prise en compte de l'effet acoustique de salle dans laquelle est installé le dispositif de restitution. Cependant, dans un contexte de reproduction physique de champs sonores, l'effet de salle n'est pas négligeable et doit être corrigé. Suivant cette problématique, la synthèse de fronts d'ondes adaptative l'*Adaptive Wave Field Synthesis* a été proposée par Gauthier et Berry [Gauthier et Berry, 2007] ou par Spors [Spors *et al.*, 2003]. Peu d'études sont encore proposées pour les approches ambisoniques [Betlehem et Abhayapala, 2005; Deprez *et al.*, 2010; Poletti, 2005], et elles sont souvent proposées pour le cas, peu réaliste, de deux dimensions. En ce qui concerne la captation des champs sonores, de nombreuses approches sont également proposées. Mais est-ce que les réseaux de microphones utilisés permettent de s'affranchir du dispositif de restitution? Ou autrement dit, est-il possible de représenter un champ sonore indépendamment du dispositif de captation et de restitution? De même, quelles sont les limitations pour la captation en terme de bande fréquentielle de fonctionnement, de repliement spatial [Gauthier *et al.*, 2011; Moreau *et al.*, 2006; Rafaely, 2005; Rafaely *et al.*, 2007]? Finalement, est-il possible d'opérer des transformations sur

les champs sonores captés ou synthétisés avant de les reconstruire, pour permettre de sélectionner une seule région de l'espace dans le champ sonore capté. Si tel est le cas, la reconstruction du champ sonore pourrait être sélective et permettre d'isoler tel ou tel événement sonore dans l'espace.

C'est dans ce contexte et dans cet état actuel des connaissances que s'inscrit ce projet de doctorat. Les principaux enjeux sont alors d'étudier et de proposer des solutions de captation ou de synthèse, de manipulation, et de reconstruction de champs sonores en trois dimensions. L'approche physique pour la synthèse de champs sonores est retenue ici et se justifie dans les applications en ingénierie.

Ce projet de doctorat est réalisé suite à une candidature de l'auteur pour un financement du Pôle de Recherche et d'Enseignement Supérieur « Hautes Études - Sorbonne - Arts et Métiers » (PRES heSam), dans le cadre du projet « Paris Nouveaux Mondes ». Une entente de cotutelle est alors mise en place entre le Laboratoire de Mécanique des Structures et Systèmes Couplés (LMSSC) du Conservatoire National des Arts et Métiers de Paris (CNAM) et le Groupe d'Acoustique de l'Université de Sherbrooke (GAUS) au Québec.

1.2 Problématique et objectifs

À la lumière de cette brève entrée en matière, la problématique de recherche principale de ce projet est la suivante :

Comment enregistrer, manipuler, puis reconstruire avec précision et sur une zone la plus large possible, tout ou une partie d'un champ acoustique en trois dimensions, quel que soit le lieu de restitution ?

Plusieurs questions de recherches sont alors dégagées de cette problématique :

- Quelle approche pour la synthèse de champs sonores en trois dimensions doit être retenue pour la mise en oeuvre du dispositif de restitution et de captation ? Quelle serait la meilleure approche pour la manipulation des champs sonores ?
- Pour la synthèse de champs sonores en trois dimensions, le choix de la géométrie sphérique semble le plus approprié, car cette géométrie ne privilégie aucune direction. Cependant, seule une réalisation discrète d'un réseau de haut-parleurs (ou de microphones) est possible suivant une surface sphérique. Alors, comment échantillonner de manière optimale une géométrie sphérique pour la captation et la restitution de champs sonores ?

- De même, l'échantillonnage spatial introduit inévitablement des artefacts dans la captation et la reproduction des champs sonores. Quels sont alors ces artefacts et comment détériorent-ils la reconstruction du champ sonore ?
- La compensation de salle a été proposée pour la synthèse de fronts d'ondes mais peu encore pour l'ambisonie. Ce sont souvent des approches établies en deux dimensions. Quelle stratégie de compensation de l'effet de salle doit être adoptée pour garantir la bonne reconstruction du champ sonore sur une zone étendue de l'espace ?
- Enfin en termes de manipulation du champ sonore, est-il possible de ne sélectionner qu'une partie du champ sonore capté ou synthétisé et ne reconstruire que cette dernière ?

Pour répondre à ces questions de recherche et à la problématique générale mentionnée ci-dessus, les objectifs du projet sont alors les suivants :

1. **Choisir la méthode pour la captation, la manipulation et la reconstruction des champs sonores en trois dimensions.** Plusieurs approches sont proposées dans la littérature pour la synthèse de champs sonores. Des approches analytiques tels que l'ambisonie ou la synthèse de fronts d'ondes, mais aussi des approches numériques, comme la synthèse par résolution d'un problème inverse acoustique. Il s'agit alors de déterminer quelle est l'approche la plus adaptée pour la synthèse de champs sonores en trois dimensions qui respecte un certain nombre de compromis : nombre de sources de restitution limité, flexibilité de mise en oeuvre.
2. **Proposer un échantillonnage de la sphère offrant les meilleurs compromis pour la synthèse de champs sonores en trois dimensions.** L'hypothèse d'une géométrie sphérique pour le dispositif de restitution est justifiée car cette configuration ne privilégie a priori aucune direction de l'espace. L'échantillonnage de la surface sphérique est une thématique de recherche en soi. Il s'agit alors de proposer un échantillonnage qui offre les meilleures performances pour la synthèse de champs sonores en offrant un nombre de noeuds minimum pour une reproduction efficace sur une zone étendue. L'étude comparative des effets de la discrétisation, tel que le repliement spatial, est également nécessaire pour choisir le maillage le plus adapté.
3. **Proposer une méthode de compensation des réflexions et des défauts du système.** La réalisation pratique d'un système de restitution peut amener des divergences avec les modèles théoriques développés. Le placement des haut-parleurs n'est pas parfait, leurs réponses en fréquence et leurs directivités ne sont pas celles de la théorie, et le lieu de restitution a une influence acoustique non négligeable sur la

reconstruction (réflexions, réfraction, diffraction). Il est alors nécessaire de mettre en oeuvre un traitement des signaux des haut-parleurs pour compenser les défauts du système et garantir la bonne reconstruction du champ sonore sur une zone étendue de l'espace.

4. **Proposer une approche de filtrage spatial dans un champ sonore pour ne sélectionner que certaines directions à reconstruire.** La manipulation du champ sonore avant reconstruction est souhaitable. Des opérations simples telles que des rotations ou symétries permettent par exemple de fixer les coordonnées pour la reconstruction. Une opération de filtrage spatial est également un atout car elle permet de sélectionner une partie seulement du champ sonore original et de ne reconstruire que celle-ci. Ceci offre alors la possibilité d'isoler des événements sonores d'une captation et de les reconstruire aux bonnes positions.
5. **Réaliser un dispositif de captation et de reconstruction pour la validation expérimentale des méthodes et techniques proposées.** Les méthodes et techniques développées pour répondre aux précédents objectifs doivent être validées expérimentalement. Ainsi, la réalisation d'une antenne de microphones pour la captation, et d'un réseau de haut-parleurs pour la restitution sont nécessaires. L'instrumentation de ces dispositifs est également à réaliser, pour permettre un contrôle de la synthèse. L'instrumentation en temps-réel serait un avantage certain, à travers une suite d'outils logiciels mettant en oeuvre les concepts développés. Une campagne de mesures utilisant la technologie réalisée est prévue pour confronter les résultats théoriques aux mesures pratiques.

1.3 Contributions originales

En réponse aux objectifs mentionnés à la section précédente, les contributions originales de ce projet de doctorat sont les suivantes :

- L'utilisation d'un maillage de Lebedev à cinquante noeuds pour la capture et la reconstruction de champs sonores en trois dimensions par approches ambisoniques. Ce maillage possède en effet plusieurs avantages dans le contexte de synthèse de champs sonores. Tout d'abord, le nombre de noeuds et leur répartition sur la sphère offre un bon compromis pour la synthèse de champs sonores par approches ambisoniques et pour la réalisation pratique. Par ailleurs, l'étude des artefacts de repliement spatial à la capture et à la reconstruction donne les meilleurs performances à la grille de Lebedev, en comparaison des résultats pour d'autres grilles avec sensiblement le même

nombre de noeuds. Finalement, la grille proposée contient deux sous grilles, offrant une grande flexibilité d'utilisation. Ces travaux ont donné lieu à une communication en conférence internationale avec comité de lecture et à un article de journal [Lecomte *et al.*, 2016a, 2015].

- La prise en compte des réflexions dans la salle de restitution ainsi que des défauts de réalisation du système de restitution, pour la reconstruction des champs sonores. L'approche proposée est entièrement formulée dans le domaine des harmoniques sphériques, ce qui permet de réduire le nombre de filtres nécessaires à la correction du système. L'utilisation de décodeurs ambisoniques conçus en condition de champ libre reste possible. La compensation des réflexions devient alors une manipulation du champ sonore avant son décodage sur le dispositif de restitution. Ces travaux ont donné lieu à un manuscrit soumis à *The Journal of Acoustical Society of America* [Lecomte *et al.*, 2016b].
- Le filtrage directionnel d'un champ sonore décrit dans un formalisme ambisonique. L'opération de filtrage directionnel permet alors de sélectionner certaines directions dans le champ sonore en fonction d'une fonction de directivité choisie. La formulation est réalisée dans le domaine des harmoniques sphériques, ce qui permet de s'affranchir du dispositif de captation et de restitution. À l'instar des techniques de formation de voies qui n'extraient qu'un signal monophonique à partir des composantes du champ sonore sur les harmoniques sphériques, l'approche proposée ici permet de conserver un champ sonore tridimensionnel et de le reconstruire par la suite sur un réseau de haut-parleurs. Ces travaux ont donné lieu à une communication au congrès français d'acoustique 2016 [Lecomte *et al.*, 2016c].
- Des résultats de mesures de champs sonores réalisés à l'aide de dispositifs expérimentaux réalisés pour l'occasion. Ainsi, deux sphères de restitutions ambisoniques jusqu'à l'ordre 5 « SpherBedev » et un prototype d'antenne sphérique rigide de microphone « MemsBedev » ont été réalisés au cours du doctorat. Une campagne de mesures fournit des cartes de champs sonores synthétisés et valide les concepts et techniques proposés dans cette thèse.
- La mise en oeuvre d'algorithmes de synthèse et de manipulation en temps-réel de champs sonores par méthodes ambisoniques, au travers d'une collection de greffons logiciels. Ces travaux ont donné lieu à une communication en conférence internatio-

nale avec comité de lecture [Lecomte et Gauthier, 2015] et à une suite de greffons disponibles en ligne [1].

1.4 Méthodologie et organisation du document

Afin d'établir une vision globale sur le sujet visé, il est nécessaire d'entrevoir une revue succincte des différentes techniques de la restitution sonore spatialisée. C'est ce que présente le chapitre [2], en détaillant notamment les deux approches analytiques que sont l'ambisonie et la synthèse de fronts d'ondes. Ce chapitre permet d'établir les principales équations utilisées au cours des chapitres suivants. A la fin du chapitre, l'approche ambisonique est retenue pour la synthèse de champs sonores en trois dimensions sur une géométrie sphérique. En effet, le chapitre montre que la formulation classique qu'utilise la synthèse de fronts d'ondes n'est pas la plus adaptée dans le cas d'une géométrie sphérique. Ce chapitre permet alors de répondre au premier objectif énoncé à la section précédente.

Pour l'échantillonnage spatial de sphère, la formulation est d'abord établie de manière continue et une attention particulière est faite aux artefacts qui apparaissent à la discrétisation. Parmi les différentes règles de quadrature de la sphère, le chapitre propose l'utilisation d'un maillage de Lebedev à cinquante noeuds pour la capture et la reconstruction de champs sonores par méthodes ambisoniques. La théorie des quadratures de la sphère par invariance rotationnelle est présentée pour comprendre la construction de la grille de Lebedev considérée. Le repliement spatial à la capture et à la reconstruction est étudiée comparativement avec deux autres grilles proposant quasiment le même nombre de noeuds. Finalement, une particularité due à la construction de la grille de Lebedev est présentée et exploitée pour le décodage d'un champ sonore. Il s'agit de la présence inhérente de plusieurs sous-grilles dans la grille principale. Ceci permet d'utiliser le maillage à différents ordres selon la fréquence. Ce chapitre permet de répondre au second objectif énoncé à la section [1.2]. Le chapitre [3] est un article en cours de parution dans le journal *Journal of Audio Engineering Society* [Lecomte et al., 2016a].

Afin de répondre à l'objectif de compensation des réflexions et des imperfections d'un système de restitution, un formalisme est d'abord établi dans le domaine des harmoniques sphériques. À cette fin, une écriture matricielle et un problème de minimisation de l'erreur de reproduction des composantes ambisoniques seront posés. Afin d'évaluer le conditionnement du problème, le potentiel de la méthode, et, surtout, la façon la plus adéquate de la mettre en oeuvre, des simulations théoriques dans le domaine de l'espace, du temps et

¹<http://www.sekisushai.net/ambitools>

des fréquences seront étudiés. Ces travaux sont présentés au Chapitre 4, ils permettent de répondre au troisième objectif énoncé à la section 1.2. Le Chapitre 4 est un manuscrit soumis pour publication à *The Journal of the Acoustical Society of America* le 17 Août 2016 [Lecomte et al., 2016b].

Le chapitre 5 propose une méthode de filtrage directionnel dans une scène sonore décrite dans un formalisme ambisonique. Le formalisme, établi dans le domaine des harmoniques sphériques, permet de s'affranchir du dispositif de captation et de restitution. Ceci permet plusieurs traitements intermédiaires dont l'application n'est pas limitée par la géométrie des réseaux utilisés. En utilisant une fonction de directivité choisie il est possible d'atténuer ou de renforcer certaines directions dans le champ sonore, avant sa reconstruction sur un réseau de haut-parleurs. Afin d'étudier cette possibilité, le formalisme doit être écrit et étudié dans le domaine des harmoniques sphériques. Pour évaluer l'efficacité de la méthode, des simulations théoriques et des validations expérimentales doivent être réalisées. Ainsi, le chapitre 5 propose une méthode de filtrage directionnel. Ce chapitre permet de répondre au quatrième objectif énoncé à la section 1.2.

L'ensemble des techniques et méthodes proposées doit être validé expérimentalement. À cette fin, des dispositifs expérimentaux doivent être réalisés. Un réseau sphérique de haut-parleurs et un réseau sphérique, rigide, de microphones seront réalisés. Le chapitre 6 présente ces dispositifs. La première partie du chapitre 6 décrit les dispositifs utilisés ou fabriqués. La seconde partie du même chapitre présente les validations expérimentales qui furent réalisées dans une salle anéchoïque. Ainsi, ce chapitre fait écho au cinquième objectif énoncé à la section 1.2.

Enfin, le chapitre 7 constitue la conclusion de ce mémoire de thèse. Les principaux résultats obtenus aux différents chapitres sont récapitulés pour conclure sur les contributions originales apportées par ce projet de doctorat. Les perspectives de recherches ouvertes à la suite de ces travaux sont présentées.

Quatres annexes sont fournies pour détailler quelques résultats utilisés dans la thèse et donner des informations supplémentaires sur les outils logiciels développés pendant le projet. Ainsi, l'annexe A reprend la démonstration de la diffraction d'une sphère rigide dans un champ acoustique. Ce résultat est utilisé pour capter un champ sonore en trois dimensions à l'aide d'une antenne sphérique rigide de microphones. L'annexe B détaille le calcul des signaux de commande pour une distribution continue de sources de types monopolaires sur une surface sphérique. Ce résultat permet d'établir l'équivalence, dans ce cas précis, de l'approche de synthèse de fronts d'ondes, basée sur l'intégrale de Kirchhoff-

Helmholtz et de l'approche ambisonique, basée sur la décomposition du champ sonore sur une base d'harmoniques sphériques. L'annexe [C](#) est un article de conférence [Lecomte et Gauthier, 2015](#) qui détaille la mise en oeuvre des équations pour une synthèse en temps-réel en utilisant les approches ambisoniques. L'annexe [D](#) est un extrait de la documentation de la suite de greffons *ambitools*. Avec l'article de l'annexe [C](#), ces deux dernières annexes offrent des informations supplémentaires sur la mise en oeuvre logicielle des concepts développés au cours du projet et présentent les outils utilisés au chapitre [6](#).

CHAPITRE 2

ÉTAT DE L'ART

2.1 Restitution du son spatialisé

Depuis l'invention du téléphone et du phonographe, les technologies ont beaucoup évolué pour la restitution du son. Les progrès conjoints sur la perception sonore, [Blauert, 1997; Hartmann, 1997] et les phénomènes acoustiques [Beranek, 1993; Bruneau, 1998; Kuttruff, 2009; Morse et Ingard, 1986] ont permis d'améliorer constamment les technologies de restitution du son, notamment avec un attribut spatial de plus en plus prononcé. Rapidement, on rappelle ci après quelques unes des principales approches pour la restitution sonore spatialisée. Pour une revue plus exhaustive, voir par exemple [Daniel, 2000; Gauthier, 2007; Nicol, 1999].

- Le format stéréophonique traditionnel, très largement utilisé, permet de restituer une prise de son ou de synthétiser une scène sonore avec deux enceintes formant avec l'auditeur un triangle équilatéral. L'illusion spatiale est basée sur une localisation additive de sources cohérentes. Par exemple, pour un signal monophonique, il est possible de jouer sur l'amplitude et le retard du signal avant de le transmettre aux deux enceintes [Blauert, 1997]. Une source fantôme, localisée de manière perceptive entre les deux enceintes, est créée. Cependant, lorsque la restitution stéréo est faite au casque d'écoute, la localisation des sources est souvent incertaine.
- L'illusion stéréophonique a ensuite été étendue sur un réseau plus important d'enceintes [Steinke, 1996; Theile, 1993]. Ainsi, le surround et les formats tels que le 5.1 ou 7.1¹ exploitent un système de restitution qui entoure l'auditoire sur un plan. Plus récemment, des canaux assignés à des enceintes en dehors du plan horizontal ont été introduits, tels que le format 22.2 ou le système Dolby Atmos [Sergi, 2013].
- L'approche de spatialisation VBAP (Vector Based Amplitude Panning) [Pulkki, 1997] permet de sélectionner dans un ensemble d'enceintes de géométrie quelconque, les trois enceintes les plus proches de la source virtuelle. Une loi panoramique en amplitude est utilisée en se basant sur une formulation vectorielle.

¹La nomenclature $x.y$ réfère à x canaux véhiculant des signaux large bande et y canaux pour les signaux basses fréquences, assignés à un ou plusieurs caissons de basses.

- Pour des écoutes aux casques comportant une illusion spatiale, la synthèse binaurale [Blauert, 2013; Møller, 1992] est une solution de plus en plus démocratisée. Elle se base sur les fonctions de transfert de la tête et du buste HRTF (Head-Related Transfert Functions). Ces fonctions de transfert décrivent les effets de diffraction de la tête et du buste d'un champ de pression acoustique aux tympanes de l'auditeur (ou d'un mannequin), en fonction de la direction. Elles sont mesurées pour un nombre fini de directions dans l'espace. L'illusion spatiale est alors réalisée en filtrant le signal à spatialiser par la paire des HRTF gauche et droite correspondant à la direction souhaitée. Le signal est ensuite délivré aux oreilles à l'aide d'un casque d'écoute. L'interpolation entre les différentes directions mesurées peut se faire par traitement sur les HRTFs [Blauert, 2013] ou par une approche Ambisonique [Noisternig *et al.*, 2003].
- La restitution binaurale peut se faire sur une paire d'enceintes [Bauck et Cooper, 1996]. On parle alors de technologie transaurale par réduction de la diaphonie. Cette approche requiert une matrice de filtres censée annuler l'effet des trajets croisés des enceintes sur les oreilles.

Pour la majorité des approches mentionnées ci-dessus, l'objectif est de recréer l'illusion sonore pour un ou plusieurs auditeurs, notamment en recréant les bons indices de localisation dans la zone d'écoute, ou perceptivement pour les techniques au casque d'écoute. Le champ sonore obtenu pour les approches avec une restitution sur enceinte n'est pas contrôlé directement au sens physique. En fait, ces techniques n'offrent pas directement la possibilité de synthétiser le champ de pression acoustique issu de sources avec un modèle de propagation choisi, comme des monopôles par exemple. De même, elles ne permettent pas de recréer à l'identique un champ sonore capté sur une zone étendue de l'espace.

Pour répondre à de tels objectifs, les approches ne doivent alors plus se focaliser sur l'illusion sonore pour un auditeur, mais plutôt sur les attributs physiques du champ sonore à reconstruire (pression, vitesse, intensité, etc.). Dans ce contexte, plusieurs approches sont proposées dans la littérature. Les deux grandes familles de description analytiques du problème de synthèse de champs sonore sont l'ambisonie d'ordre élevé ou HOA (Higher Order Ambisonics) et la synthèse de fronts d'ondes ou WFS (Wave Field Synthesis), détaillées dans les sections suivantes.

D'autres approches pour la synthèse de champ sonore sont également à mentionner, par exemple celles basées sur la résolution numérique d'un problème inverse [Camier *et al.*, 2010; Gauthier *et al.*, 2012, 2015; Kirkeby et Nelson, 1993; Nelson *et al.*, 1992], ou bien les

méthodes de génération de zone de « contraste acoustiques », proposée par Choi et Kim [Choi et Kim, 2002, 2012].

2.2 Ambisonie d'ordre élevé

2.2.1 Aperçu de la méthode

La méthode ambisonique d'ordre élevé (appelée simplement ambisonie pour la suite du document) est une technologie qui permet d'encoder et de restituer un champ sonore capté ou synthétisé en deux ou trois dimensions. Elle repose sur la description d'un champ de pression acoustique sur une base tronquée d'harmoniques cylindriques (deux dimensions) ou sphériques (trois dimensions). La méthode a été introduite par Gerzon dans les années 70 au premier ordre (harmoniques du premier ordre), [Gerzon, 1973, 1985]. Elle a ensuite été étendue aux ordres supérieurs par les travaux de Bamford [Bamford, 1995], Malham [Malham, 1992, 1999; Malham et Myatt, 1995] et Daniel [Daniel, 2000; Daniel *et al.*, 2003, 1998] notamment.

À noter que d'autres méthodes de décomposition de champs sonores utilisant les harmoniques sphériques / cylindriques et partageant de nombreuses analogies avec l'ambisonie ont été proposées, notamment par Poletti [Poletti, 1996, 2000] ou Ward et Abhayapala [Ward et Abhayapala, 2001].

La méthode ambisonique repose sur trois étapes qui sont : l'encodage, les transformations et le décodage sur un réseau de haut-parleurs ou au casque d'écoute. La figure 2.1 illustre ces différentes étapes.

L'encodage repose sur la description d'un champ de pression sur une base d'harmoniques sphérique. Cette décomposition permet d'obtenir les composantes ambisoniques, dénommées couramment format-B pour le premier ordre. Ces composantes peuvent être estimées lors de la captation d'un champ sonore, à l'aide d'un réseau sphérique de microphones par exemple, ou bien produites à l'aide d'un modèle de source, comme un monopôle par exemple. La description encodée du champ sonore sous forme de composantes ambisoniques présente de nombreux avantages pratiques. En effet, l'encodage est extensible (il peut être décrit à n'importe quel ordre) et flexible : il est relativement simple d'ajouter des composantes ambisoniques issus de différents champs sonores pour créer un champ sonore complexe. Enfin, le fait de travailler dans l'espace des harmoniques sphériques permet de s'affranchir du dispositif de restitution.

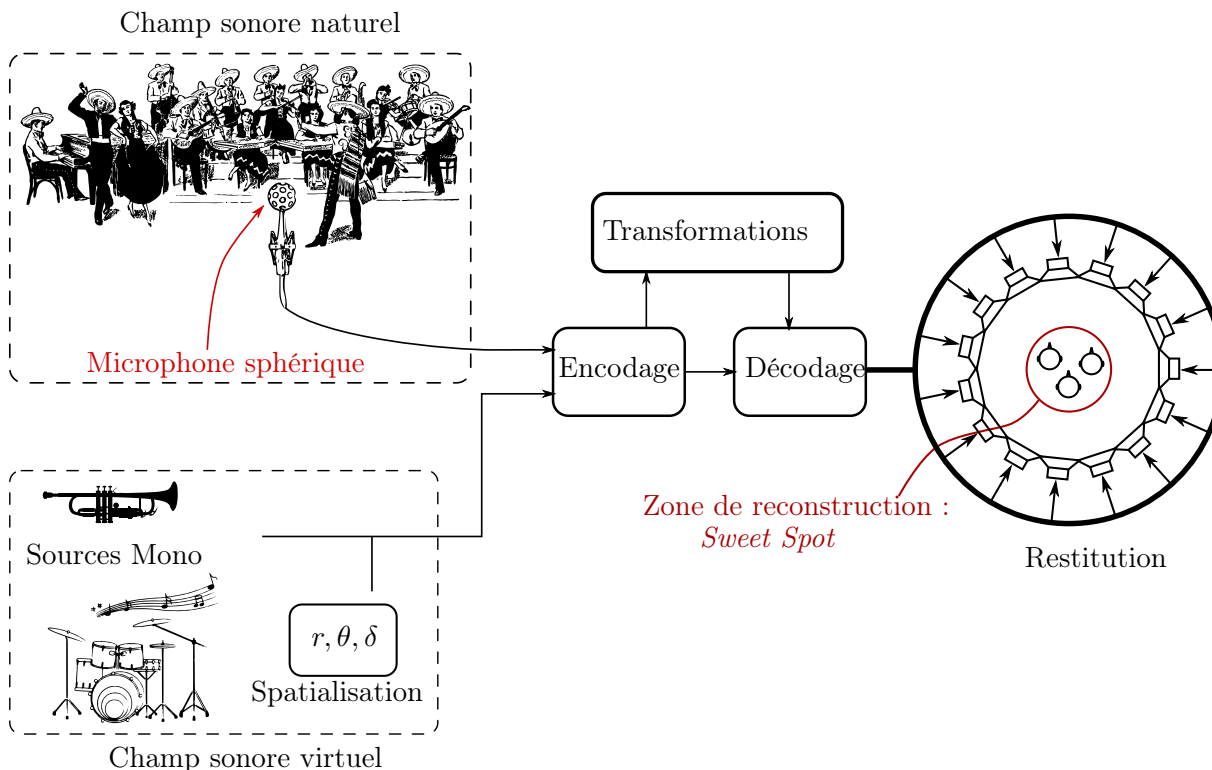


Figure 2.1 Principe général de la méthode ambisonique.

Dans le domaine ambisonique, des transformations peuvent être réalisées sur le champ sonore. Par exemple des transformations simples telles qu'une rotation [Daniel, 2000; Moreau, 2006] ou une symétrie [Kronlachner, 2014; Zotter, 2009a]. Des opérations plus complexes sont également possibles, comme une distorsion de la perspective [Daniel, 2000; Zotter et Pomberger, 2011], une égalisation comme présentée au chapitre 4 de ce document ou encore un filtrage directionnel tel que présenté au chapitre 5 de ce document.

Finalement, l'étape de décodage consiste à calculer les signaux des haut-parleurs du dispositif de restitution à partir des composantes ambisoniques encodées, pour reconstruire le champ sonore. Historiquement, l'ambisonie modélise les haut-parleurs du système de restitution comme générateur d'ondes planes, suffisamment éloignés de la zone de reproduction. Les travaux de Daniel [Daniel, 2003] ont inclus les sources de restitution en champ proche. Le formalisme permet alors d'encoder des ondes sphériques et de les restituer à l'aide de sources monopolaires. A noter que la configuration géométrique du réseau de haut-parleurs est importante pour une bonne reconstruction du champ sonore.

Les différentes étapes de la méthode ambisonique sont revues plus en détails dans les sections ci-après. Certaines conventions et notions mathématiques sont dans un premier temps introduites.

2.2.2 Système de coordonnées

Le système de coordonnées sphérique retenu pour l'ensemble de ce document suit la convention définie sur la figure 2.2 :

$$\begin{cases} x = r \cos(\theta) \cos(\delta) \\ y = r \sin(\theta) \cos(\delta) \\ z = r \sin(\delta) \end{cases} \quad (2.1)$$

Un point $P(x, y, z)$ en coordonnées cartésiennes est décrit en coordonnées sphériques par

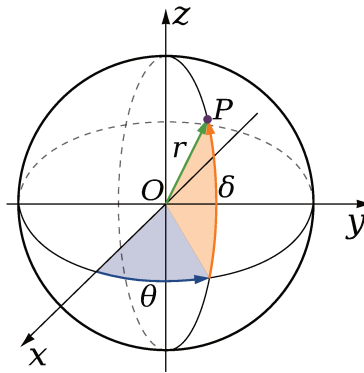


Figure 2.2 Système de coordonnées sphériques. Un point $P(x, y, z)$ est décrit par son rayon r son azimut θ et son élévation δ .

son rayon r , avec $r \in [0, \infty[$ son azimut θ , avec $\theta \in [0, 2\pi[$ et son élévation (ou site) δ , avec $\delta \in [-\pi/2, \pi/2]$. Le sens positif de rotation est choisi comme le sens trigonométrique (anti-horaire). Il existe une norme ISO pour le système de coordonnées sphériques [ISO8000-2, 2009] où l'élévation est décrite à parti du pôle nord. Cependant, dans la littérature de l'audio spatiale, il est plus commun de trouver le système de coordonnées présenté ci-dessus, qui décrit l'élévation à partir du plan de l'équateur [Blauert, 1997]. En effet, c'est généralement ce plan qui est choisi au niveau des oreilles des auditeurs dans le cadre d'une restitution pour un public. Les sources avec une élévation positive sont alors situées au dessus de l'équateur et celles avec une élévation négative, en dessous.

2.2.3 Produit scalaire sphérique

Dans ce document, l'espace Hilbertien \mathbb{L}^2 est défini comme l'espace des fonctions à valeurs réelles de carré intégrable sur la sphère unité $\Omega = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$.

Soient f et g deux fonctions définies dans \mathbb{L}^2 . f et g fonctions des deux variables, θ et δ . Le produit scalaire de ces deux fonctions est alors défini par :

$$\langle f, g \rangle = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\frac{\pi}{2}}^{+\frac{\pi}{2}} f(\theta, \delta)g(\theta, \delta) \cos(\delta)d\delta d\theta. \quad (2.2)$$

Ainsi, le produit scalaire sphérique $\langle f, g \rangle$ correspond à la multiplication des deux fonctions f et g et à leur intégration de surface sur la sphère unité.

Les deux fonctions f et g sont orthogonales si leur produit scalaire sphérique est nul.

La norme associée au produit scalaire est définie pour une fonction f par :

$$\| f \| = \sqrt{\langle f, f \rangle}. \quad (2.3)$$

2.2.4 Harmoniques sphériques

Définition

La famille d'harmoniques sphériques [Arfken et Weber, 2005; Morse et Ingard, 1986], dénotés Y_{mn} , forme une base orthonormée de \mathbb{L}^2 . Ils sont définis² par [Nachbar et al., 2011] :

$$Y_{mn}(\theta, \delta) = \sqrt{(2m+1)\epsilon_n \frac{(m-|n|)!}{(m+|n|)!}} P_{m,|n|}(\sin(\delta)) \times \begin{cases} \cos(|n|\theta) & \text{if } n \geq 0 \\ \sin(|n|\theta) & \text{if } n < 0 \end{cases}, \quad (2.4)$$

où $P_{m,|n|}$ sont les polynômes de Legendre associés d'ordre m et de degré $|n|$, $(m, n) \in (\mathbb{N}, \mathbb{Z})$, $m \leq |n|$, et $\epsilon_n = 1$ si $n = 0$, $\epsilon_n = 2$ si $|n| > 0$. Dans cette équation, m désigne l'ordre de l'harmonique sphérique et n son degré. Les polynômes de Legendre associés $P_{m,|n|}(x)$ sont définis pour $x \in [-1, 1]$ par :

$$P_{m,|n|}(x) = (1-x^2)^{\frac{|n|}{2}} \frac{d^{|n|}}{dx^{|n|}} P_m(x), \quad (2.5)$$

²Il existe plusieurs définitions possibles des harmoniques sphériques. Pour ce document, les harmoniques sphériques à valeurs réelles sont utilisés.

où $P_m(x)$ sont les polynômes de Legendre de première espèce d'ordre m . Finalement, ces polynômes peuvent se calculer par récurrence [Arfken et Weber, 2005] :

$$\begin{cases} P_0(x) & = 1 \\ P_1(x) & = x \\ (m+1)P_{m+1}(x) & = (2m+1)xP_m(x) - mP_{m-1}(x) \quad m > 1 \end{cases} . \quad (2.6)$$

À chaque ordre m , il y a $2m+1$ harmoniques sphériques. Une base tronquée à l'ordre M contient alors $(M+1)^2$ fonctions. Les harmoniques sphériques forment une base orthonormée de \mathbb{L}^2 . Ainsi, ils sont orthonormés deux à deux :

$$\langle Y_{mn}, Y_{m'n'} \rangle = \delta_{mm'} \delta_{nn'} . \quad (2.7)$$

Dans cette équation δ_{ij} représente le symbole de Kronecker avec $\delta_{ii} = 1 \forall i$ et $\delta_{ij} = 0, \forall i, j, i \neq j$.

Normalisation

Le facteur de normalisation utilisé dans l'équation (2.4) correspond à la normalisation N3D (Normalisation 3D), telle que définie par Daniel [Daniel, 2000]. Ainsi, ce facteur est tel que la norme d'un harmonique sphérique est unitaire et la base formée par la famille des harmoniques sphériques est orthonormée. Cependant, dans l'utilisation des harmoniques sphériques pour l'ambisonie, d'autres facteurs de normalisation sont utilisés, telles que la normalisation SN3D (Semi Normalisation 3D) ou FuMa (Furse-Malham, du nom des auteurs proposant cette normalisation [Malham, 2003]). Les bases correspondantes sont toujours orthogonales, mais plus orthonormées. Le choix d'autres facteurs de normalisation est motivé par des considérations notamment audio-numériques [Chapman et al., 2009]. La normalisation SN3D, par exemple, garantit que les signaux ambisoniques d'ordre $m > 0$ ont une amplitude maximale inférieure au signal d'ordre 0 [Kronlachner, 2014].

Ordonnement

Les harmoniques sphériques peuvent être rangés de plusieurs manières suivant les indices (m, n) . Dans ce document, ils sont ordonnés suivant leur ordre *ACN* (*Ambisonic Channel Number*³) croissant [Nachbar et al., 2011], avec :

$$ACN = m^2 + m + n . \quad (2.8)$$

³Voir notamment : <http://ambisonics.ch/standards/channels/>

La notation ACN permet de ne retenir qu'un seul indice pour identifier un harmonique sphérique. Les indices m et n correspondants peuvent être retrouvés à l'aide des formules suivantes :

$$\begin{aligned} m &= \lfloor \sqrt{ACN} \rfloor \\ n &= ACN - m^2 - m \end{aligned} \quad (2.9)$$

où $\lfloor \cdot \rfloor$ représente la partie entière. La figure 2.3 montre le diagramme de directivité des harmoniques sphériques jusqu'à l'ordre $M = 5$, suivant leur ordonnancement ACN croissant.

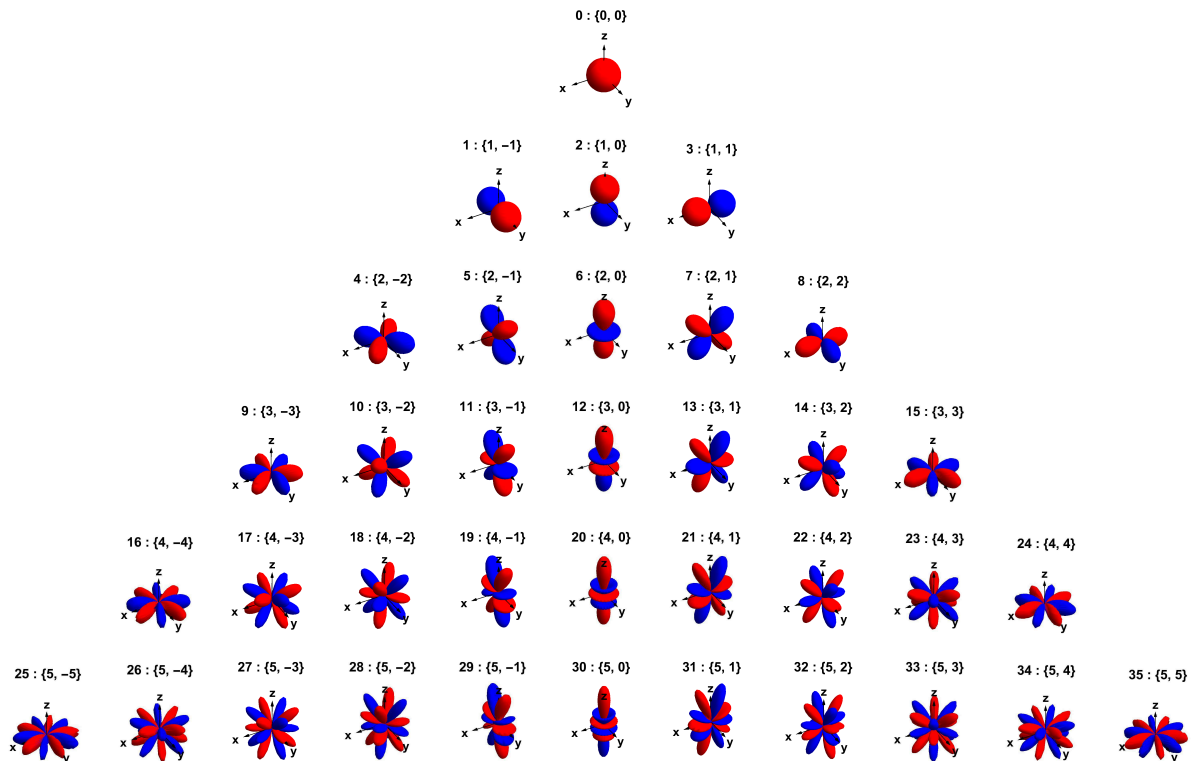


Figure 2.3 Illustration des harmoniques sphériques jusqu'à l'ordre $M = 5$. Les couleurs rouge et bleue indiquent respectivement des valeurs positives et négatives. Le numéro ACN et les indices (m, n) correspondants sont notés au dessus de chaque figure.

Théorème d'additivité

Un théorème utile pour la manipulation des harmoniques sphériques est le théorème d'additivité [Arfken et Weber, 2005] :

$$\sum_{n=-m}^m Y_{mn}(\theta_1, \delta_1) Y_{mn}(\theta_2, \delta_2) = P_m(\gamma), \quad (2.10)$$

où $\gamma_l = \cos(\delta_1) \cos(\delta_2) \cos(\theta_1 - \theta_2) + \sin(\delta_1) \sin(\delta_2)$.

2.2.5 Transformée de Fourier sphérique

La Transformée de Fourier Sphérique (TFS) permet de projeter une fonction angulaire f de \mathbb{L}^2 sur la base des harmoniques sphériques. Les coefficients f_{mn} issus de la projection de cette fonction sur la base orthonormée sont donnés par :

$$f_{mn} = \langle f, Y_{mn} \rangle = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} f(\theta, \delta) Y_{mn}(\theta, \delta) \cos(\delta) d\theta d\delta. \quad (2.11)$$

La transformée de Fourier sphérique inverse est donnée par :

$$f(\theta, \delta) = \sum_{m=0}^{\infty} \sum_{n=-m}^m f_{mn} Y_{mn}(\theta, \delta). \quad (2.12)$$

Ainsi, la fonction angulaire f peut être décrite par une combinaison linéaire d'harmoniques sphériques. Les coefficients de la combinaison linéaire sont donnés par la TFS.

Dans la pratique, l'intégrale dans l'équation (2.11) est approximée par une somme (éventuellement pondérée) sur L noeuds, répartis sur la sphère unité et la Transformée de Fourier Sphérique Discrète (TFSD) est obtenue. La discrétisation de la sphère est alors un sujet de recherche en soi, plus de détails à ce sujet sont donnés au chapitre 3.

2.2.6 Fonctions de Bessel et de Hankel sphériques

Comme il le sera montré à la section suivante, la dépendance radiale du champ de pression acoustique est exprimée avec les fonctions de Bessel et de Hankel sphériques, présentées à cette section.

Fonctions de Bessel sphériques

Les fonctions de Bessel sphériques de première et seconde espèces utilisées dans ce document sont définies par Williams, 1999 :

$$\begin{aligned} j_m(x) &= \sqrt{\frac{\pi}{2x}} J_{m+1/2}(x) \\ y_m(x) &= \sqrt{\frac{\pi}{2x}} Y_{m+1/2}(x) \end{aligned}, \quad (2.13)$$

où $J_m(x)$ et $Y_m(x)$ représentent la fonction de Bessel d'ordre m de première espèce et seconde espèce respectivement. La figure [2.4](#) montre le tracé des fonctions de Bessel $j_m(kr)$ pour différents ordres.

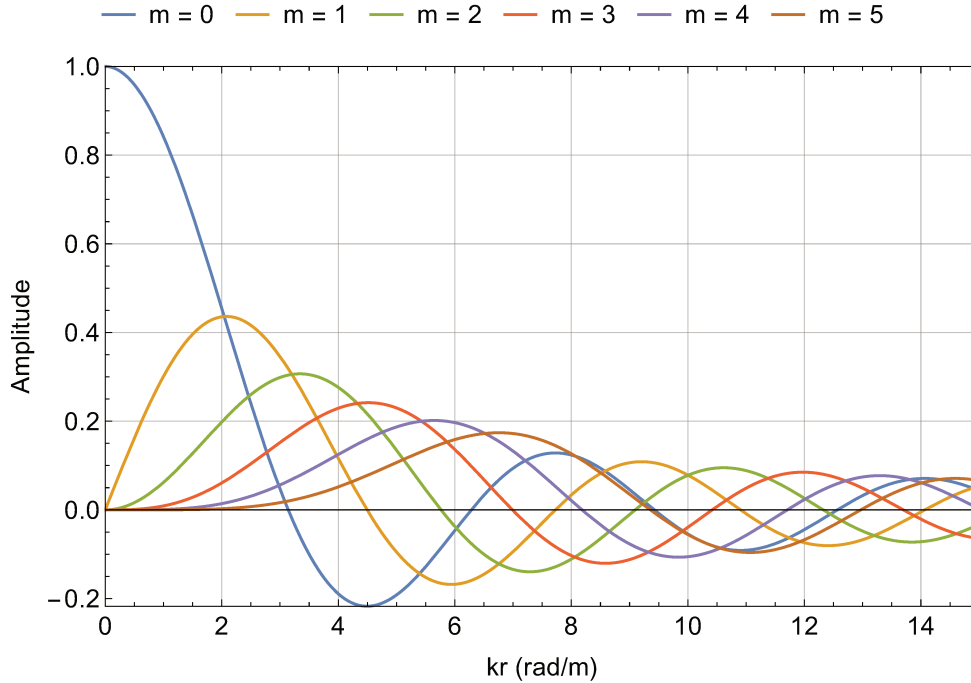


Figure 2.4 Fonctions de Bessel sphériques de la première espèce pour les ordres $m = 1$ à 5.

Fonctions de Hankel sphériques

Les fonctions de Hankel sphériques de première et seconde espèces utilisées dans ce document sont définies par :

$$\begin{aligned} h_m^{(1)}(x) &= j_m(x) + iy_m(x) \\ h_m^{(2)}(x) &= j_m(x) - iy_m(x) \end{aligned} \quad (2.14)$$

où i représente l'unité imaginaire $i = \sqrt{-1}$.

Wronskiens

Le Wronksien est utile pour relier les fonctions de Hankel et de Bessel sphériques à leur dérivée première [\[Abramowitz et Stegun, 1972\]](#) :

$$j_m(x)h_m^{(2)\prime}(x) - j_m^{\prime}(x)h_m^{(2)}(x) = \frac{-i}{x^2}. \quad (2.15)$$

$$j_m^{\prime}(x)h_m^{(2)}(x) - j_m(x)h_m^{(2)\prime}(x) = \frac{i}{x^2}. \quad (2.16)$$

Ces relations sont notamment utilisées pour trouver l'expression des filtres d'égalisation radiaux pour un réseau de microphones à la surface d'une sphère rigide (voir l'annexe [A](#)) et pour trouver l'expression des signaux des sources secondaires dans la formulation simple source (voir l'annexe [B](#) et la section [2.2.12](#)).

2.2.7 Décomposition d'un champ de pression acoustique sur les harmoniques sphériques

En utilisant le formalisme mathématique présenté ci-dessus, il est possible de décrire un champ de pression acoustique sur la base des harmoniques sphériques.

La convention temporelle utilisée dans ce document est choisie comme $e^{+i\omega t}$ où i est l'unité imaginaire, $\omega = 2\pi f$, la pulsation en radians par secondes avec f la fréquence en Hertz. Le temps est dénoté t et s'exprime en secondes.

Séries de Fourier-Bessel

Dans un domaine spatial sphérique, exempt de sources acoustiques, de rayon r_s , l'équation de Helmholtz homogène est donnée par [\[Williams, 1999, p. 218\]](#) :

$$\Delta p + k^2 p = 0. \quad (2.17)$$

Dans cette équation, Δp représente le laplacien de la pression acoustique. Le nombre d'onde $k = \omega/c$ s'exprime en radians par mètres où c est la vitesse du son dans l'air en mètres par secondes.

Les solutions de l'équation d'Helmholtz homogène dans le système de coordonnées sphériques se trouvent en utilisant la méthode de la séparation des variables r, θ, δ [\[Bruneau, 1998\]](#) et on aboutit à une description de la pression sous la forme d'une série de Fourier-Bessel :

$$p(kr, \theta, \delta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta), \quad (2.18)$$

Ainsi, la pression acoustique est entièrement décrite dans une sphère centrée à l'origine O de rayon r_s où la dépendance radiale et fréquentielle sont données par les fonctions de Bessel sphériques $j_m(kr)$, et la dépendance angulaire est donnée par la décomposition sur les harmoniques sphériques. Cette dépendance angulaire correspond à la TFS inverse du champ de pression acoustique, suivant l'équation [\(2.12\)](#). Les coefficients B_{mn} sont dénommés les composantes ambisoniques. Physiquement, ils correspondent aux dérivées spatiales successives de la pression acoustique prise à l'origine O [\[Daniel, 2003\]](#).

L'expression analytique des composantes ambisoniques B_{mn} se trouve pour des sources acoustiques simples comme une onde plane ou une onde sphérique. Ces deux cas sont traités ci-après.

2.2.8 Encodage d'une onde plane et d'une onde sphérique

Encodage d'une onde plane

Soit une onde plane monochromatique, de pulsation ω , d'amplitude S et de direction de propagation (θ_p, δ_p) . La pression acoustique générée par cette onde plane est donnée au point (r, θ, δ) par :

$$p(r, \theta, \delta) = S e^{-ikr \cos(\gamma)}, \quad (2.19)$$

où γ est l'angle entre la direction (θ, δ) et (θ_p, δ_p) . En exprimant cette équation sur une base d'harmoniques sphériques et en identifiant les termes avec la série de Fourier-Bessel de l'équation (2.18), on obtient les coefficients B_{mn} pour une onde plane [Rafaely, 2015] :

$$B_{mn} = S Y_{mn}(\theta_p, \delta_p). \quad (2.20)$$

Encodage d'une onde sphérique et filtres de champ proche

Soit, une onde sphérique émise par une monopôle placé à la position $\mathbf{r}_s = (r_s, \theta_s, \delta_s)$. La pression au point d'observation $\mathbf{r} = (r, \theta, \delta)$ est donnée par [Williams, 1999] :

$$p(r, \theta, \delta) = S \frac{e^{-ik|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|}. \quad (2.21)$$

La fonction de Green en espace libre $e^{-ik(|\mathbf{r}-\mathbf{r}_s|)}/(4\pi|\mathbf{r}-\mathbf{r}_s|)$ est exprimée en coordonnées sphériques [Morse et Ingard, 1986, p. 352], [Abramowitz et Stegun, 1972, eq. 10.1.17] :

$$S \frac{e^{-ik|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|} = S \frac{-ik}{4\pi} \sum_{m=0}^{\infty} j_m(kr) h_m^{(2)}(kr_s) \sum_{n=-m}^m Y_{mn}(\theta_s, \delta_s) Y_{mn}(\theta, \delta), \quad (2.22)$$

où $h_m^{(2)}$ représentent les fonctions de Hankel sphériques de seconde espèce. En identifiant cette équation avec la série de Fourier-Bessel de l'équation (2.18) on obtient les composantes ambisoniques B_{mn} pour une onde sphérique :

$$B_{mn} = S F_m(kr_s) Y_{mn}(\theta_s, \delta_s). \quad (2.23)$$

Dans l'équation précédente, $F_m(kr_s) = i^{-(m+1)}kh_m^{(2)}(kr_s)/(4\pi)$ sont les filtres de champs proches, introduits par Daniel [Daniel, 2003] dans le cadre de l'ambisonie. Ces filtres modélisent la distance finie de la source à l'origine O . Ils dépendent de la fréquence et de l'ordre m . À noter que dans la formulation présentée ici, l'amplitude S est reliée au débit volumique du monopôle Q [Kuttruff, 2009] par $S = i\rho_0ckQ$ (ρ_0 est la densité de l'air en kilogrammes par mètres cubes). Cependant, la formulation présentée par Daniel [Daniel, 2003] modélise S comme l'amplitude de la pression à l'origine O . Elle simplifie notamment la mise en oeuvre numérique des filtres de champs proches, car l'atténuation et le retard entre la source et l'origine sont supposés déjà modélisés [Lecomte et Gauthier, 2015]. En conséquence, les filtres de champs proches F_m exprimés dans ce document sont légèrement différents de ce que propose Daniel [Daniel, 2003].

La figure 2.5 montre le gain des filtres de champ proche pour différents ordres m . L'amplitude de ces filtres est infinie lorsque $kr_s = 0$ et $m > 0$. Ainsi l'encodage de sources sphériques est problématique en basses fréquences et à un faible rayon r_s . En effet, la mise en oeuvre de tel filtres est instable par nature et une trop grande amplification pour des faibles valeurs de kr_s peut détériorer les haut-parleurs du système de restitution. Pour stabiliser les filtres, Daniel propose de les coupler avec les filtres de prise en compte du champ proche des sources de restitution lors du décodage [Daniel, 2003] (voir la section 2.2.12).

2.2.9 Captation et encodage d'un champ sonore

La section précédente donne l'expression analytique des composantes ambisoniques pour des sources acoustiques simples. Pour des champs de pression acoustique complexes, comme un champ sonore enregistré, il est possible d'estimer les composantes ambisoniques au premier ordre, ou aux ordres plus élevés, en utilisant notamment un réseau sphérique de microphones.

Microphone Soundfield™ pour le premier ordre

Gerzon a développé, dans les années 70, un microphone basé sur l'association de quatre microphones à directivité cardioïde, disposés en tétraèdre [Gerzon, 1975]. Il s'agit du microphone Soundfield™ visible sur la figure 2.6. Ce microphone permet d'enregistrer la pression et ses gradients suivant les trois directions x, y, z , correspondant aux composantes ambisoniques B_{mn} jusqu'à l'ordre $M = 1$, en matriciant les signaux issus de chacune des capsules.

Microphone sphérique pour les ordres élevés

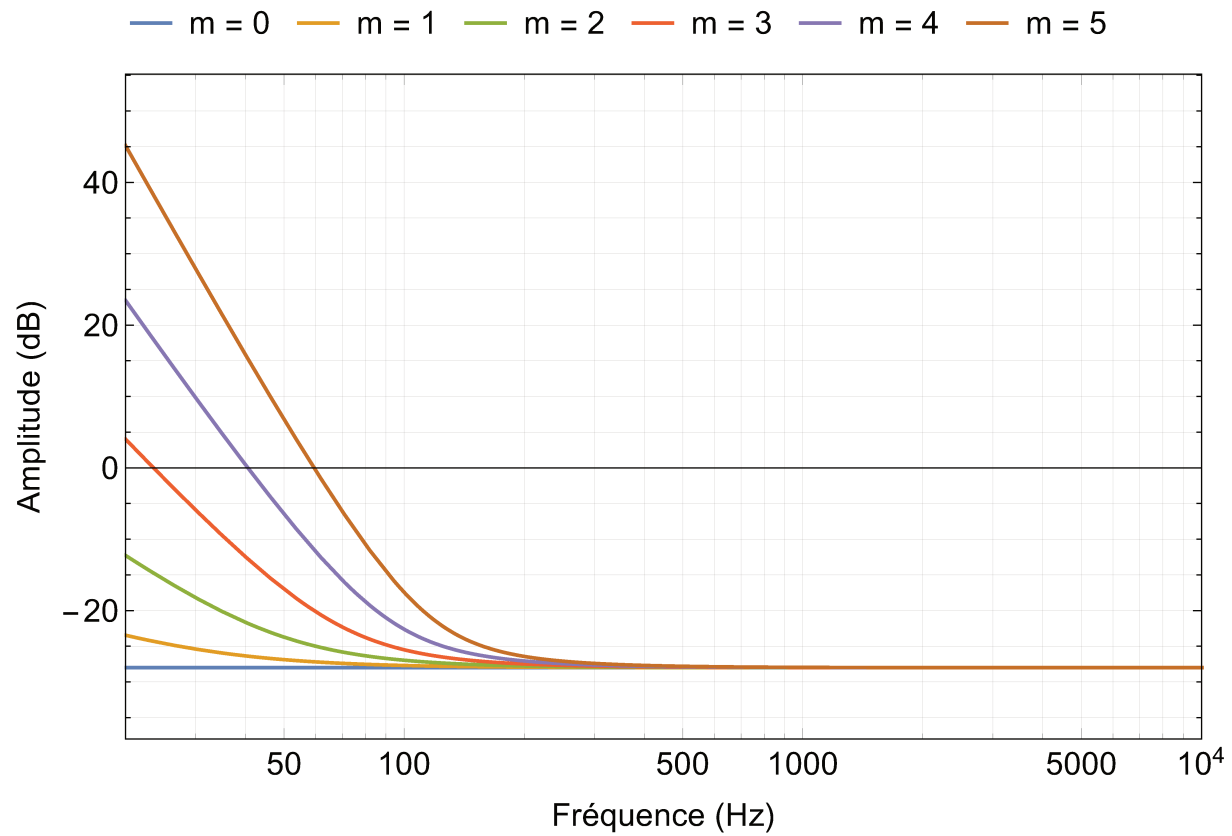


Figure 2.5 Gain des filtres de champ proches pour les ordres $m = 1$ à 5 pour l'encodage ambisonique d'une onde sphérique.



Figure 2.6 Les capsules d'un microphone SoundField™ MKV.

L'estimation des composantes ambisoniques aux ordres élevés peut se faire à partir de la connaissance de la pression sur une surface sphérique, ce qui permet de calculer la TFS. La formulation continue pour une surface sphérique ouverte et détaillée ci-après.

Surface sphérique simple couche. Le champ de pression acoustique $p(kr_{\text{mic}}, \theta, \delta)$ est supposé connu sur une surface sphérique de rayon r_{mic} . La TFS de la pression s'exprime de la manière suivante :

$$p_{mn} = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\frac{\pi}{2}}^{+\frac{\pi}{2}} p(kr_{\text{mic}}, \theta, \delta) Y_{mn}(\theta, \delta) \cos(\delta) d\delta d\theta. \quad (2.24)$$

En remplaçant $p(kr_{\text{mic}}, \theta, \delta)$ par son expression en série de Fourier-Bessel (voir l'équation (2.18)), on obtient :

$$p_{mn} = \sum_{m'=0}^{+\infty} i^{m'} j_{m'}(kr_{\text{mic}}) \sum_{n'=-m'}^{m'} B_{m'n'} \langle Y_{m'n'}(\theta, \delta) Y_{mn}(\theta, \delta) \rangle. \quad (2.25)$$

La propriété d'orthonormalité des harmoniques sphériques permet alors de simplifier la précédente équation :

$$p_{mn} = i^m j_m(kr_{\text{mic}}) B_{mn}. \quad (2.26)$$

Finalement,

$$B_{mn} = \frac{1}{i^m j_m(kr_{\text{mic}})} p_{mn}. \quad (2.27)$$

Ainsi, les composantes B_{mn} sont obtenues en multipliant la TFS de la pression p par les filtres d'égalisation $i^{-m}/j_m(kr_{\text{mic}})$. Cependant, les fonctions $j_m(kr)$ présentent des zéros, pour certaines valeurs de kr_{mic} (voir la figure 2.4). L'estimation des composantes B_{mn} est alors compromise par une potentielle division par 0 aux fréquences telles que $j_m(kr_{\text{mic}}) = 0$. Pour remédier à ce problème, d'autres formulations sont possibles en utilisant par exemple plusieurs surfaces sphériques concentriques [Braikia et al., 2013; Williams, 1999] où en exploitant les propriétés de diffraction d'une sphère rigide. Une revue complète des différentes approches est proposée par Plessas [Plessas, 2009; Rafaely, 2015].

Surface sphérique rigide. Dans le cas d'une sphère rigide, les composantes ambisoniques s'obtiennent par l'équation suivante :

$$B_{mn} = \underbrace{i^{-(m-1)} (kr_{\text{mic}})^2 h_m^{(2)}(kr_{\text{mic}})}_{E_m(kr_{\text{mic}})} p_{mn}, \quad (2.28)$$

où $h_m^{(2)}(kr_{\text{mic}})$ représente la dérivée des fonctions de Hankel de seconde espèce par rapport à la variable kr . La démonstration de ce résultat est disponible dans l'annexe [A](#).

Filtres d'égalisation. Pour la surface sphérique rigide, les filtres d'égalisation correspondant $E_m(kr_{\text{mic}})$ ne présentent pas de fréquences indéterminées à part à la fréquence nulle pour $m > 0$. La figure [2.7](#) présente le gain de ces filtres d'égalisation pour les ordres $m = 1$ à 5 pour une surface sphérique rigide de rayon $r_{\text{mic}} = 0.07$ m. Ces filtres présentent un gain excessif en basses fréquences et pour les faibles rayons de réseaux sphériques de microphones. Ceci est dû au fait qu'en basses fréquences, le champ sonore est dominé par l'ordre 0 et l'estimation des ordres élevés requiert une amplification excessive des composantes de la TFS d'autant plus importante que le rayon du réseau sphérique r_{mic} est faible. Pour des applications pratiques, les filtres d'égalisation sont stabilisés et le gain sur les composantes de la TFS est limité. Daniel et Moreau proposent de stabiliser ces filtres en les couplant aux filtres de compensation du champ proche des sources de restitution (voir la section [2.2.12](#)), ou bien en utilisant un filtre passe-haut dont la fréquence de coupure est judicieusement choisie [\[Daniel et Moreau, 2004\]](#). Moreau résout le problème inverse d'estimation des composantes à l'aide d'une régularisation de Thikhonov dont le paramètre de régularisation est relié à un rapport signal sur bruit désiré [\[Moreau, 2006; Moreau et al., 2006\]](#). Le chapitre [6](#) détaille cette approche dans le cadre du microphone MemsBedev qui est réalisé dans le cadre de la thèse. D'autres approches de mise en oeuvre stable de ces filtres d'égalisation sont proposées dans notamment [\[Baumgartner et al., 2011; Lösler et Zotter, 2015\]](#).

Discretisation et fréquence de repliement. En pratique, la connaissance du champ de pression sur une surface rigide n'est possible qu'en un nombre fini de positions. Ainsi, la surface sphérique rigide est discrétisée pour estimer la TFS par une TFSD. La règle de quadrature doit respecter l'orthonormalité des harmoniques sphériques pour garantir la validité du résultat. La discrétisation entraîne alors une fréquence de repliement au delà de laquelle l'estimation des composantes ambisoniques n'est plus possible. Cette fréquence est approximativement $f_{\text{alias}} = Mc/(2\pi r_{\text{mic}})$, [\[Rafaely, 2005; Rafaely et al., 2007\]](#), où M est l'ordre maximal pour lequel la TSFD est identique à la TSF. L'étude du repliement spatial pour les microphones sphériques est menée par Rafaely [\[Rafaely et al., 2007\]](#) pour quelques règles de quadratures. Le chapitre [3](#) détaille également le repliement spatial dans le contexte d'une captation ambisonique et le quantifie pour plusieurs règles de quadrature de la sphère.

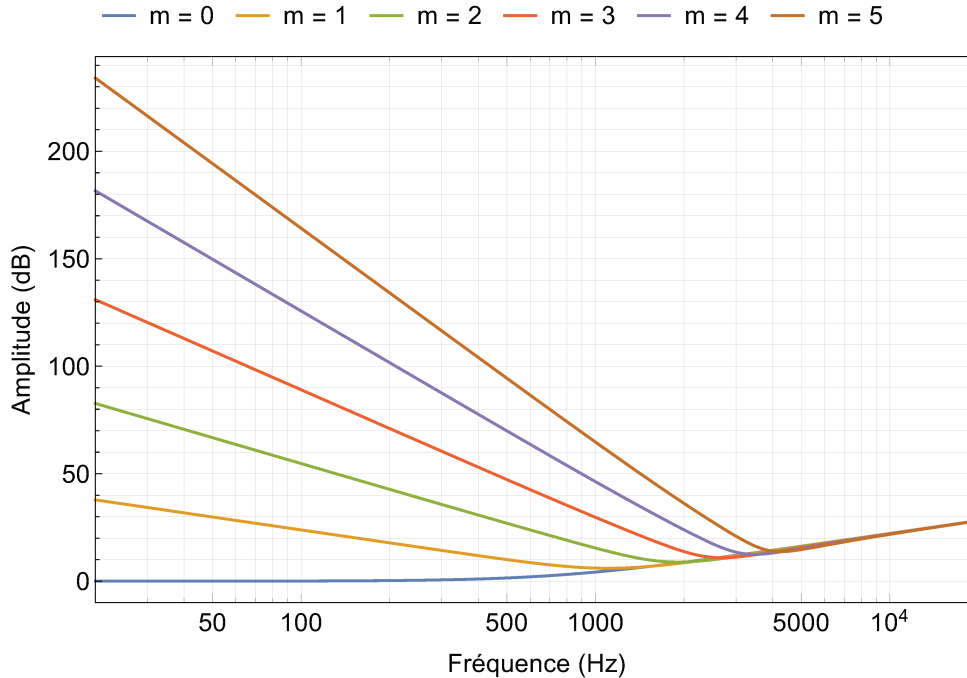


Figure 2.7 Gain des filtres d'égalisation E_m pour une sphère rigide de rayon $r_{\text{mic}} = 7$ cm, pour les ordres $m = 1$ à 5.

La figure 2.8, présente deux exemples de réseaux de microphones répartis sur une surface sphérique rigide. La figure 2.8(a) présente le microphone EigenMike 32TM commercialisé par MH AcousticsTM, avec 32 capsules microphoniques réparties suivant un pentakis-dodécaèdre [Moreau *et al.*, 2006]. Ce microphone peut réaliser l'estimation des composantes ambisoniques jusqu'à l'ordre $M = 4$. La figure 2.8(b) présente le prototype de MemsBedev, développé dans le cadre ce projet de doctorat, avec 50 capsules microphoniques réparties sur une sphère rigide en suivant une grille de Lebedev. Ce prototype est capable d'estimer les composantes ambisoniques jusqu'à l'ordre $M = 5$ et est présenté plus en détails au chapitre 6 de ce document. D'autres réalisations de réseaux sphériques rigides de microphones sont visibles par exemple dans [Moreau *et al.*, 2006; Plessas, 2009; Zotkin *et al.*, 2008].

2.2.10 Troncature

En pratique, la représentation d'un champ de pression acoustique à l'aide d'une série de Fourier-Bessel de l'Eq. (2.18) est tronquée à l'ordre M . Le champ est alors représenté à l'aide de $(M + 1)^2$ composantes ambisoniques. L'erreur de troncature normalisée entre le champ de pression théorique et le champ décrit par une série de Fourier-Bessel tronquée



Figure 2.8 Exemples de réseaux de microphones sphériques répartis sur une sphère rigide. (a) Microphone EigenMike 32 TM permettant une estimation des composantes ambisoniques jusqu'à l'ordre $M = 4$. (b) Microphone MemsBedeV permettant une estimation des composantes ambisoniques jusqu'à l'ordre $M = 5$.

est donné par [Poletti, 2005] :

$$\epsilon_M(k, r, \theta, \delta) = \frac{|p(k, r, \theta, \delta) - p_M(k, r, \theta, \delta)|^2}{|p(k, r, \theta, \delta)|^2}. \quad (2.29)$$

Cette erreur peut être moyennée pour toutes les directions (θ, δ) en intégrant l'équation (2.29) sur la sphère unité. La démonstration repose sur la décomposition de $p(k, r, \theta, \delta)$ et $p_M(k, r, \theta, \delta)$ sur les harmoniques sphériques. Elle est menée par Ward et al. [Ward et Abhayapala, 2001] pour le cas bidimensionnel avec des sources de type ondes planes et par Poletti [Poletti, 2005] pour le cas tridimensionnel avec des ondes sphériques. Une formulation élégante est donnée par Zotter pour les ondes sphériques [Zotter et al., 2009]. Les formules correspondantes sont rappelées ci-après :

$$\bar{\epsilon}_M = 1 - \sum_{m=0}^M (2m+1)(j_m(kr))^2, \quad (2.30)$$

pour les ondes planes et

$$\bar{\epsilon}_M = 1 - \frac{2k^2 r_s r}{\ln\left(\frac{r_s+r}{r_s-r}\right)} \sum_{m=0}^M (2m+1)(j_m(kr))^2 |h_m^{(2)}(kr_s)|^2, \quad (2.31)$$

pour les ondes sphériques. A titre d'exemple, la figure 2.9 présente l'erreur de troncature moyennée $\bar{\epsilon}_M$ pour les ondes planes et pour plusieurs ordres de troncature M . De ces réf-

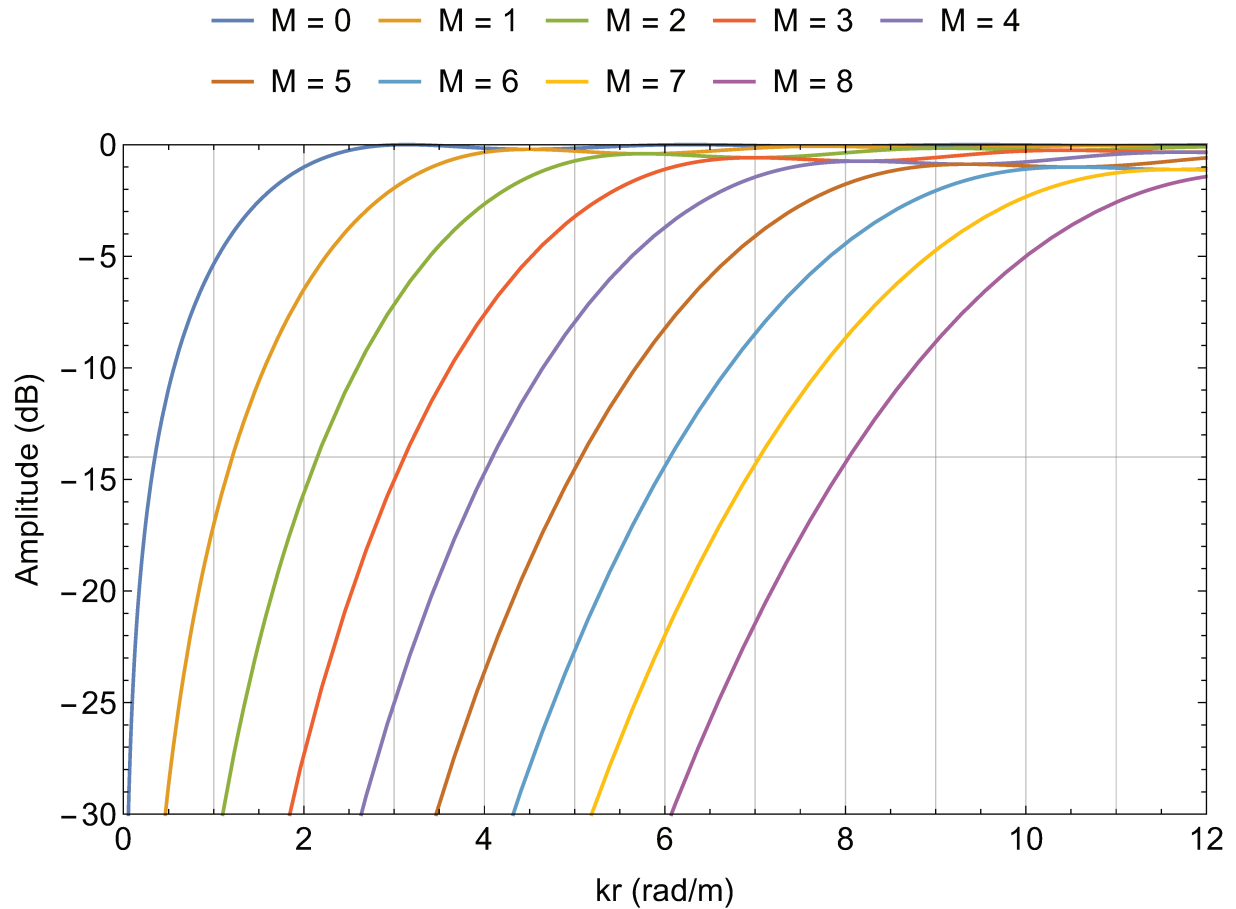


Figure 2.9 Erreur de troncature moyennée $\bar{\epsilon}_M$ pour les ondes ondes planes et différents ordres de troncature. L'erreur à -14 dB est indiquée par une ligne horizontale.

rences, une règle générale d'approximation est établie : pour une erreur $\bar{\epsilon}_M$ de 4 % (-14 dB) entre le champ de pression théorique et le champ de pression tronqué à l'ordre M , la taille de la zone de reconstruction est donnée par une sphère de rayon $r_{4\%}$ avec :

$$r_{4\%} = M/k. \quad (2.32)$$

Ainsi $r_{4\%}$ est proportionnel à l'ordre de troncature M et inversement proportionnel à la fréquence, via le nombre d'onde k . Ceci s'observe sur la figure 2.9 où l'erreur à -14 dB est atteinte pour $[kr] = M$ ($[\cdot]$ représente l'arrondi supérieur). À noter que dans le cas des ondes sphériques, cette règle approximative est valable dans la mesure où $r < r_s$. En effet, la série de Fourier-Bessel de l'équation (2.18) est définie pour le problème intérieur limité par r_s .

La figure 2.10 illustre l'effet de troncature pour un champ sonore issu d'un monopôle situé à la position $\mathbf{r}_s = (2 \text{ m}, 0^\circ, 0^\circ)$ pour différents ordres de troncature M . L'erreur $\epsilon_M = 4\%$ de l'équation (2.29) est montrée sous forme d'un contour noir. La cercle de rayon $r_{4\%}$ est en rouge pointillé. On observe que la zone de reconstruction à -14 dB est de la taille du cercle à $r_{4\%}$. Ainsi, plus la fréquence augmente, plus la taille de la zone de reconstruction diminue. De même, plus l'ordre de troncature augmente, plus la taille de la zone de reconstruction augmente. La troncature de la série de Fourier-Bessel est alors une limitation de la reproduction d'un champ de pression dans une zone étendue. En effet, l'ordre de troncature conditionne directement la taille de la zone de reconstruction.

Une fois la série de Fourier-Bessel tronquée, les composantes ambisoniques se représentent sous forme d'un vecteur $\mathbf{b}^{(M+1)^2 \times 1} = [b_0 \cdots b_i \cdots b_{(M+1)^2-1}]^T$, où b_i représente la $i^{\text{ème}}$ composante ambisonique, suivant son indice ACN, et T représente la transposée.

2.2.11 Transformations

D'après l'équation (2.18), le champ sonore est entièrement décrit par les composantes ambisoniques B_{mn} , décrites dans le domaine des harmoniques sphériques. Dans le cas d'une description tronquée à un ordre M fini, c'est le vecteur \mathbf{b} qui décrit le champ sonore. Il est possible d'effectuer des transformations sur le champ sonore, directement dans le domaine des harmoniques sphériques, en agissant sur les composantes du vecteur \mathbf{b} . Ces transformations s'appliquent par une opération matricielle sur le vecteur \mathbf{b} à l'aide d'une matrice de transformation \mathbf{M} . La matrice est généralement indépendante de la fréquence ou du temps, lorsqu'il s'agit d'une transformation géométrique dans le domaine angulaire. Quelques transformations simples sont décrites dans cette section.

Rotation

Une rotation sur un harmonique sphérique d'ordre m résulte en un autre harmonique sphérique, toujours d'ordre m . Ainsi, la matrice de rotation correspondante est diagonale par bloc [Moreau, 2006]. L'expression analytique de la matrice de rotation autour de l'axe Oz d'un angle θ est donnée ci-après [Daniel, 2000; Kronlachner, 2014; Moreau, 2006;

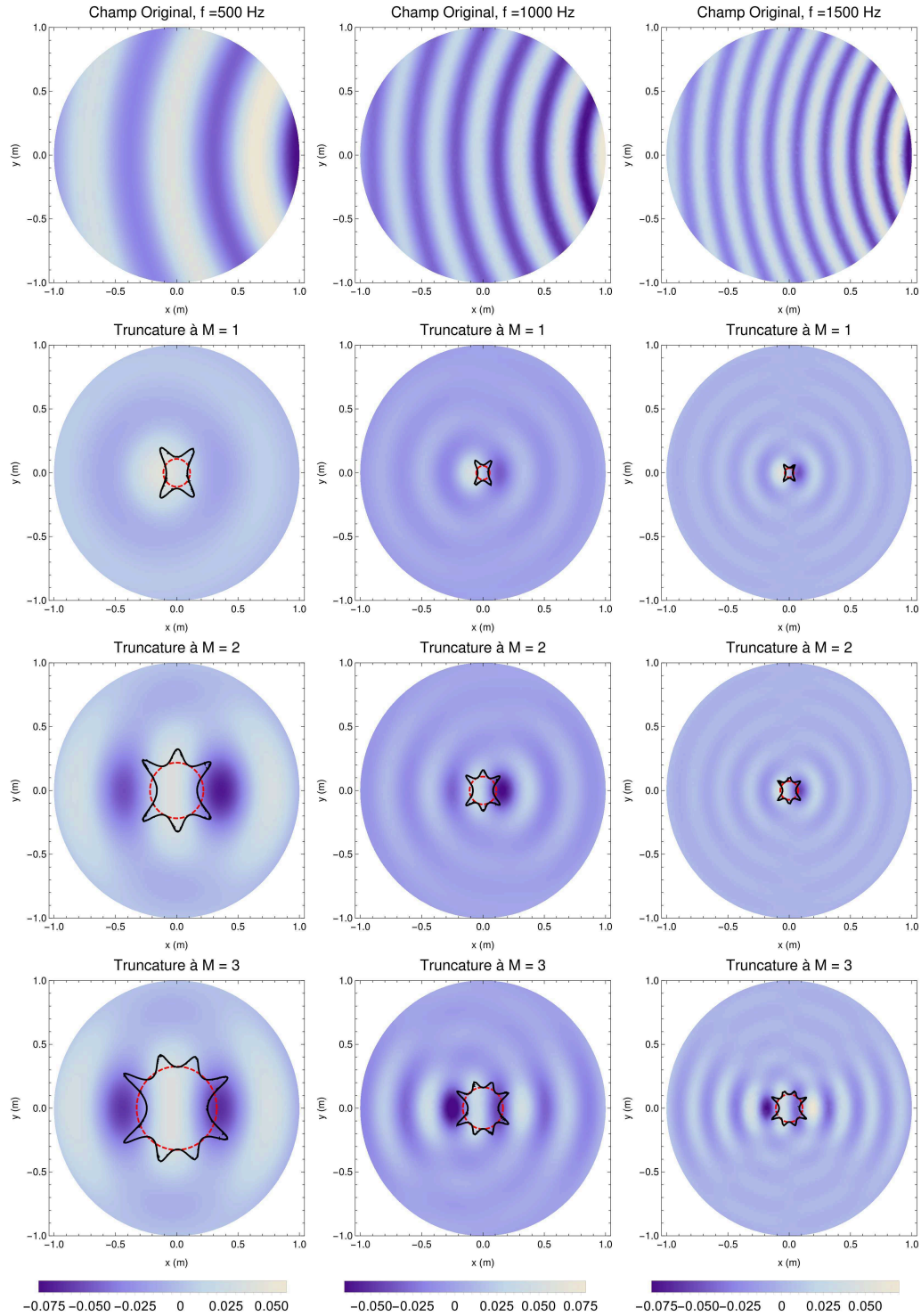


Figure 2.10 Partie réelle du champ de pression issu d'un monopôle à la position $\mathbf{r}_s = (2 \text{ m}, 0^\circ, 0^\circ)$ aux fréquences 500 Hz (colonne de gauche), 1000 Hz (colonne du milieu) et 1500 Hz (colonne de droite). Première ligne : champs théoriques. Lignes suivantes : champs tronqués aux ordres $M = 1$, $M = 2$ et $M = 3$. L'erreur ϵ_M entre le champ théorique et tronqué est représentée sous forme d'un contour noir. Le cercle de rayon $r_{4\%}$ est représenté en rouge pointillé.

Zotter, 2009a :

$$\mathbf{M}_{\text{rot}}^z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & \cos(\theta) & 0 & \sin(\theta) & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & -\sin(\theta) & 0 & \cos(\theta) & 0 & 0 & 0 & 0 & 0 & \\ \hline 0 & 0 & 0 & 0 & \cos(2\theta) & 0 & 0 & 0 & \sin(2\theta) & \\ 0 & 0 & 0 & 0 & 0 & \cos(\theta) & 0 & \sin(\theta) & 0 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & -\sin(\theta) & 0 & \cos(\theta) & 0 & \\ 0 & 0 & 0 & 0 & -\sin(2\theta) & 0 & 0 & 0 & \cos(2\theta) & \\ \hline \mathbf{0} & & \mathbf{0} & & & & \mathbf{0} & & & \cos(3\theta) \\ & & & & & & & & \ddots & \end{bmatrix}. \quad (2.33)$$

La matrice de rotation des autres axes x et y est plus difficile à obtenir. Pour obtenir une matrice de rotation $\mathbf{M}_{\text{rot}}^{xyz}$ quelconque, il est possible de calculer la matrice par récurrence [Ivanic et Ruedenberg, 1996], ou bien décomposer la rotation sous forme de trois rotations successives autour des axes Ox , Oy et Oz . Une approche par une combinaison de rotation de 90° autour de l'axe y jointe avec une rotation arbitraire autour de l'axe z est proposée par Zotter dans [Zotter, 2009a]. Une autre approche, proposée par Kronlachner et Zotter [Kronlachner, 2014] consiste à exprimer la scène sonore décrite par \mathbf{b} dans le domaine angulaire, en appliquant une TFSD inverse sur le champ sonore. Ceci est effectué par une opération matricielle sur le vecteur \mathbf{b} [Kronlachner, 2014]. Dans le domaine angulaire les matrices de rotations autour des axes Ox , Oy et Oz sont connues et appliquées successivement au champ sonore avant de retourner dans le domaine des harmoniques sphériques en effectuant une TFSD. Finalement, la matrice $\mathbf{M}_{\text{rot}}^{xyz}$ résultante est le produit de la matrice effectuant la TFSD inverse, des matrices de rotations dans le domaine angulaire et de la matrice effectuant la TFSD. Cependant, cette approche nécessite alors la mise en oeuvre de l'opération de TFSD. Suivant le maillage utilisé, la TFSD peut être inexacte au delà d'un certain ordre pour lequel la règle de quadrature de la sphère ne fonctionne plus. Plus de détails sont donnés à ce sujet au chapitre 3.

La rotation du champ sonore peut être pilotée par un dispositif de suivi des mouvements de la tête de l'auditeur dans le cadre d'une restitution au casque. Les mouvements sont alors compensés en appliquant la rotation inverse au champ sonore [Lecomte et Gauthier, 2015; Noisternig *et al.*, 2003].

Symétrie

Une transformation simple sur le champ sonore consiste à exploiter les symétries des harmoniques sphériques, pour effectuer une symétrie sur le champ sonore [Chapman, 2009; Zotter, 2009a]. Ainsi, les directions frontale et arrière peuvent être interverties, ainsi que la gauche et la droite ou le haut et le bas. L'opération consiste à changer le signe des harmoniques sphériques concernées dans le vecteur \mathbf{b} . Les relations de symétrie des harmoniques sphériques sont données ci après (voir également la figure 2.3) :

Symétrie par rapport à l'axe (Ox). Les harmoniques sphériques d'indices :

$$n < 0 \text{ ET } n \text{ pair} \text{ OU } n \geq 0 \text{ ET } n \text{ impair}, \quad (2.34)$$

sont anti-symétriques (symétriques avec changement de signe) par rapport à l'axe (Ox).

Symétrie par rapport à l'axe (Oy). Les harmoniques sphériques d'indices :

$$n < 0, \quad (2.35)$$

sont anti-symétriques (symétriques avec changement de signe) par rapport à l'axe (Oy).

Symétrie par rapport à l'axe (Oz). Les harmoniques sphériques d'indices :

$$(m + n) \text{ impair}, \quad (2.36)$$

sont anti-symétriques (symétriques avec changement de signe) par rapport à l'axe (Oz).

Finalement la matrice de transformation par symétrie du champ sonore est diagonale unitaire avec des termes ± 1 sur la diagonale, suivant la symétrie (ou combinaison de symétries) appliquée.

Autres transformations

D'autres transformations sont également à mentionner tels que la distorsion de la perspective (*warping*) où une direction dans le domaine angulaire est amenée vers une autre. Cette notion a été proposée par Gerzon pour le premier ordre [Gerzon et Barton, 1992]. Malham la traite [Malham, 1990] et elle a été étendue aux ordres supérieurs par Zotter

et Pomberger [Zotter et Pomberger, 2011]. Kronlachner propose plusieurs transformations sur les champs sonores encodés au formalisme ambisonique [Kronlachner, 2014].

Finalement, le filtrage directionnel proposé au chapitre 5 de ce document est également une transformation effectuée dans le domaine des harmoniques sphériques au moyen d'une matrice. De même que l'égalisation du système de restitution ambisonique proposée au chapitre 4 de ce document, à ceci près que la transformation est dépendante de la fréquence dans ce cas.

2.2.12 Décodage

L'étape de décodage ambisonique vise à trouver les signaux des haut-parleurs de restitution, utilisés pour reconstruire le champ de pression cible, décrit par les composantes ambisoniques du vecteur \mathbf{b} . Dans la plupart des cas, la géométrie sphérique est la plus naturelle pour répartir les haut-parleurs autour de la zone de reconstruction. Deux approches sont proposées pour résoudre ce problème d'estimation des signaux. L'approche par *mode-matching*, telle que proposée par Daniel [Daniel, 2000; Daniel et al., 2003] ou l'approche par *simple source* [Williams, 1999], proposée plus récemment pour l'ambisonie [Ahrens et Spors, 2008a; Fazi et Nelson, 2013; Fazi et al., 2009; Poletti, 2005; Williams, 1999]. Les deux sections suivantes reprennent le principe de ces deux approches.

Mode-matching et principe de ré-encodage

Soient L haut-parleurs de restitution, modélisés comme des monopôles placés autour de l'origine et avec les amplitudes $\mathbf{s} = [s_1 \cdots s_L]^T$. La contribution de chacune des sources est exprimée à l'aide d'une série de Fourier-Bessel à l'origine (équations (2.18) et (2.22)). La somme de leur contribution doit être égale à la série de Fourier-Bessel du champ cible :

$$\sum_{l=1}^L s_l \sum_{m=0}^{\infty} i^m j_m(kr) F_m(kr_l) \sum_{n=-m}^m Y_{mn}(\theta_l, \delta_l) Y_{mn}(\theta, \delta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta). \quad (2.37)$$

La partie à gauche de l'égalité dans l'équation (2.37) représente la somme des champs sonores de chacun des haut-parleurs. La partie à droite de l'égalité représente le champ sonore cible. Les harmoniques sphériques étant orthonormés, chacun des termes (m, n) , (où modes ambisoniques) de cette somme peut être identifié séparément (d'où le nom *mode-matching*). Pour s'en convaincre, il suffit d'appliquer une TFS à chacun des membres de l'équation (2.37). Ainsi :

$$\sum_{l=1}^L s_l F_m(kr_l) Y_{mn}(\theta_l, \delta_l) = B_{mn}. \quad (2.38)$$

En travaillant à un ordre de troncature M , le problème peut s'écrire sous forme matricielle :

$$\mathbf{C}\mathbf{s} = \hat{\mathbf{b}}, \quad (2.39)$$

où $\mathbf{C}^{(M+1)^2 \times L}$ est appelée la matrice de ré-encodage [Daniel *et al.*, 2003]. Cette matrice, étant donné le vecteur de signaux \mathbf{s} , permet d'exprimer le champ sonore reconstruit sous forme d'un vecteur ambisonique $\hat{\mathbf{b}}$. Dans le cas d'une distribution sphérique des haut-parleurs, à un rayon r_{spk} de l'origine, la matrice de ré-encodage peut s'exprimer comme le produit d'une matrice diagonale $\mathbf{F}_{r_{\text{spk}}}^{(M+1)^2 \times (M+1)^2}$ et d'une matrice rectangulaire $\mathbf{Y}^{(M+1)^2 \times L}$:

$$\mathbf{C} = \mathbf{F}_{r_{\text{spk}}} \mathbf{Y}. \quad (2.40)$$

La matrice diagonale $\mathbf{F}_{r_{\text{spk}}}$ contient les filtres de champs proches $F_m(kr_{\text{spk}})$ (présentés à la section 2.2.8) sur sa diagonale. La matrice \mathbf{Y} contient les harmoniques sphériques jusqu'à l'ordre M , évalués dans la direction de chaque haut-parleur. Ainsi, la $l^{\text{ème}}$ colonne de \mathbf{Y} , notée \mathbf{y}_l , est $\mathbf{y}_l = [Y_{0,0}(\theta_l, \delta_l) \cdots Y_{M,M}(\theta_l, \delta_l)]^T$ et $\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_L]$.

Finalement, l'opération de décodage, étant donné les composantes ambisoniques cibles \mathbf{b} , fournit les signaux \mathbf{s} :

$$\mathbf{s} = \mathbf{D}\mathbf{b}, \quad (2.41)$$

où $\mathbf{D}^{L \times (M+1)^2}$ représente la matrice de décodage. Pour s'assurer que $\hat{\mathbf{b}} = \mathbf{b}$, la matrice \mathbf{C} doit être inversée :

$$\mathbf{s} = \mathbf{C}^{\#}\mathbf{b} = \mathbf{D}\mathbf{b}, \quad (2.42)$$

Dans l'équation (2.42), l'inverse est symboliquement noté $\#$ car sa définition peut varier suivant la disposition des haut-parleurs [Zotter, 2009b]. De plus, l'inversion de \mathbf{C} peut nécessiter une éventuelle régularisation si les haut-parleurs ne sont pas répartis de manière suffisamment *régulière* autour de la sphère [Poletti, 2005]. Enfin, le nombre de haut-parleurs doit excéder le nombre de modes harmoniques à évaluer pour être dans le cas d'un problème sur-déterminé avec au moins une solution exacte [Moreau, 2006; Poletti, 2005] : $L \geq (M+1)^2$. En pratique, lorsque la matrice de ré-encodage sépare la partie directionnelle de la partie radiale, selon l'équation (2.40), la matrice de décodage \mathbf{D} sépare également la partie radiale de la partie directionnelle. La partie directionnelle est indépendante de la fréquence et la partie radiale est appelée la matrice de compensation du champ proche ou *near field compensation matrix*. Ces filtres de compensation du champ proche, inverse des filtres de champs proches, sont implémentés à l'aide de filtres RII (Réponse Impulsionnelle Infinie) à l'aide de *biquads* avec peu de coefficients. [Daniel, 2003; Lecomte et Gauthier,

[2015]. Cette formulation du décodage à l'aide de monopôles et de compensation du champ proche est connue sous l'appellation *Near Field Compensated HOA (NFC-HOA)*, telle que proposée par Daniel [Daniel, 2003].

Approche par simple source

Les approches plus récentes pour le décodage ambisonique se basent sur la formulation par simple source [Ahrens et Spors, 2008a; Fazi et Nelson, 2013; Fazi *et al.*, 2009; Poletti, 2005; Williams, 1999]. Ainsi, au lieu de partir d'un nombre fini de haut-parleurs, la formulation est dérivée de manière continue et la discrétisation se fait par la suite. Considérons une distribution continue de sources monopolaires sur une surface sphérique de rayon r_{spk} avec la distribution de signal $s(\mathbf{r}_l)$, ($\mathbf{r}_l = (r_{\text{spk}}, \theta_l, \delta_l)$). La pression $\hat{p}(k, r, \theta, \delta)$ à l'intérieur du volume sphérique est donnée par :

$$\hat{p}(k, r, \theta, \delta) = \left\langle s(\theta_l, \delta_l), \frac{e^{-ik|\mathbf{r}-\mathbf{r}_l|}}{4\pi|\mathbf{r}-\mathbf{r}_l|} \right\rangle, \quad (2.43)$$

où $s(\theta_l, \delta_l)$ est l'amplitude du monopôle à la position $\mathbf{r}_l = (r_{\text{spk}}, \theta_l, \delta_l)$. En considérant un problème intérieur et extérieur conjointement $s(\theta_l, \delta_l)$ peut être calculé [Williams, 1999] (la démonstration est fournie à l'annexe [B]) :

$$s(\theta_l, \delta_l) = \sum_{m=0}^{+\infty} \sum_{n=-m}^m \frac{B_{mn} Y_{mn}(\theta_l, \delta_l)}{F_m(kr_{\text{spk}})}, \quad (2.44)$$

Finalement, en remplaçant les termes de l'équation [2.43] par leur expression dans les équations [2.22] et [2.44], et en appliquant le critère d'orthonormalité des harmoniques sphériques, on obtient la série de Fourier-Bessel de l'équation [2.18]. Ainsi, en fixant les composantes ambisoniques B_{mn} , le champ sonore cible est reproduit.

Lorsque la formulation précédente est discrétisée, il est important de conserver le critère d'orthonormalité des harmoniques sphériques pour s'assurer de la validité de la solution. Si la règle de quadrature de la sphère utilisée respecte ce critère (du moins jusqu'à un ordre M), la formulation simple source est équivalente à l'approche par mode-matching. La règle de quadrature présentée au chapitre [3] respecte le critère d'orthonormalité des harmoniques sphériques.

Loi panoramique équivalente

Dans le cadre d'une règle de quadrature de la sphère à L haut-parleurs, respectant le critère d'orthonormalité des harmoniques sphériques jusqu'à un ordre M , on a la relation

suivante :

$$\mathbf{Y}^T \mathbf{W} \mathbf{Y} = \mathbf{I}, \quad (2.45)$$

où $\mathbf{Y}^{L \times (M+1)^2}$ est la matrice des harmoniques sphériques évaluées aux directions de chaque haut-parleur, $\mathbf{I}^{(M+1)^2 \times (M+1)^2}$ est la matrice identité et $\mathbf{W}^{L \times L}$ est une matrice diagonale de poids pour chacun des haut-parleurs, dépendant de la règle de quadrature utilisée (voir chapitre 3). Ainsi, la matrice de décodage a la forme :

$$\mathbf{D} = \mathbf{W} \mathbf{Y} \mathbf{F}_{r_{\text{spk}}}^{-1}. \quad (2.46)$$

Finalement, dans le cadre de la spatialisation d'une source de type onde plane de direction de propagation (θ_s, δ_s) ou d'une onde sphérique provenant de la position $(r_s, \theta_s, \delta_s)$, les signaux des haut-parleurs sont donnés par :

$$\begin{aligned} \mathbf{s}_{\text{spk}} &= S \mathbf{W} \mathbf{Y} \mathbf{F}_{r_{\text{spk}}}^{-1} \mathbf{Y}_{(\theta_s, \delta_s)} && \text{Ondes planes,} \\ \mathbf{s}_{\text{spk}} &= S \mathbf{W} \mathbf{Y} \mathbf{F}_{r_s} \mathbf{F}_{r_{\text{spk}}}^{-1} \mathbf{Y}_{(\theta_s, \delta_s)} && \text{Ondes sphériques,} \end{aligned} \quad (2.47)$$

Il est possible d'utiliser le théorème d'addition des harmoniques sphériques (équation 2.10) pour simplifier cette équation et obtenir les lois panoramiques équivalentes :

$$\begin{aligned} s_l &= S w_l \sum_{m=0}^M \frac{(2m+1)}{F_m(k, r_{\text{spk}})} P_m(\gamma_l) && \text{Ondes planes,} \\ s_l &= S w_l \sum_{m=0}^M (2m+1) \frac{F_m(k, r_s)}{F_m(k, r_{\text{spk}})} P_m(\gamma_l) && \text{Ondes sphériques,} \end{aligned} \quad (2.48)$$

où $\gamma_l = \cos(\delta_s) \cos(\delta_l) \cos(\theta_s - \theta_l) + \sin(\delta_s) \sin(\delta_l)$ est l'angle entre la source virtuelle à la direction (θ_s, δ_s) et le haut-parleur l .

Décodages psycho-acoustiques

En utilisant les équations de décodage présentées ci-dessus, un décodage dit *basique* est utilisé. L'objectif est alors de reconstruire le champ sonore sur une zone étendue, centrée à l'origine. Cependant, plusieurs limitations pratiques limitent cet objectif. Ainsi, la série de décomposition du champ sonore est tronquée et la zone de reconstruction varie alors avec la fréquence. De même, le nombre fini de haut-parleurs utilisés pour la reconstruction introduit un repliement spatial comme il le sera montré plus en détail au chapitre 3. Enfin, le lieu de restitution perturbe la reconstruction (voir notamment le chapitre 4 à ce sujet). Pour des applications vouées à l'écoute par un public, des décodages motivés par des choix psycho-acoustiques sont proposés pour maximiser les effets perceptifs du

champ sonore reproduit. Il est possible, par exemple de donner un poids plus important aux haut-parleurs de restitution qui sont dans la même direction que la source à restituer, dans le cadre d'une spatialisation. Ainsi, les haut-parleurs dans la direction opposés sont progressivement éteints. Gerzon propose de s'appuyer sur des indices spatiaux de la source, comme le vecteur vitesse ou le vecteur énergie, pour optimiser le décodage [Gerzon, 1992]. Un décodeur en deux bandes de fréquences est alors conçu en maximisant progressivement l'un puis l'autre vecteur. Daniel propose un décodage qui maximise le vecteur énergie en pondérant les composantes harmoniques sphériques avant l'étape de décodage [Daniel *et al.*, 1998]. Plus récemment, Zotter et Frank proposent un décodeur *All-Round* [Zotter et Frank, 2012] qui allie un décodeur basique sur une grille virtuelle de haut-parleurs et un *mapping* des signaux virtuels sur les haut-parleurs réels par la technique VBAP.

Décodages sur des géométries non sphériques

Bien que la géométrie sphérique ne privilégie aucune direction, et est la plus directe pour concevoir un décodeur, elle n'est pas la géométrie la plus avantageuse pour une installation pratique. En effet, bien souvent on désire garder l'accès au sol et les salles accueillant de tels dispositifs sont parallélépipédiques. Le décodage sur des répartitions de haut-parleurs non sphériques est cependant possible. Une approche est de trouver une base de fonctions orthonormées adaptée à la géométrie du système de restitution. Ceci est proposé notamment par Zotter [Zotter *et al.*, 2010, 2012]. Un décodage basique sur une grille virtuelle de haut-parleurs suivi d'un *mapping* des signaux sur la géométrie réelle est également proposé par Zotter [Zotter et Frank, 2012].

Enfin, il est important de noter que les équations du décodage sont établies en champ libre. En réalité, la réverbération présente dans la salle où se situe le dispositif de restitution détériore la reconstruction. Betlehem et Abhayapala proposent alors un décodage qui prend en compte les réflexions de la salle de restitution en deux dimensions [Betlehem et Abhayapala, 2005]. Le chapitre 4 étend ce concept en trois dimensions avec une approche sensiblement différente.

Décodage binaural

La flexibilité de l'ambisonie permet de décoder le champ sonore décrit par le vecteur \mathbf{b} sur de nombreuses configurations de haut-parleurs, mais aussi au casque d'écoute pour une restitution binaurale [Daniel, 2000; Moreau, 2006; Noisternig *et al.*, 2003]. La scène sonore \mathbf{b} est dans un premier temps décodée sur un réseau virtuel de haut-parleurs, puis chacun des signaux des haut-parleurs virtuels est filtré par la paire d'HRTFs correspondantes. Finalement, les signaux filtrés de chacun des haut-parleurs virtuels sont sommés pour obtenir

le signal stéréo à transmettre au casque d'écoute. La rotation du champ sonore dans le domaine ambisonique permet de mettre facilement en place un dispositif de compensation des mouvements de la tête de l'auditeur, comme expliqué à la section [2.2.11](#).

2.2.13 Exemple en simulation

La figure [2.11](#) présente un résultat de simulation d'un champ sonore. Il s'agit d'un monopôle placé à la position $\mathbf{r}_s = (2 \text{ m}, 45^\circ, 35^\circ)$, avec une amplitude $S = 1$ et à une fréquence $f = 500 \text{ Hz}$. La partie réelle du champ sonore tronqué à l'ordre $M = 5$, suivant la série de Fourier-Bessel de l'équation [\(2.18\)](#), est montrée sur les deux figures du haut. Le champ sonore reconstruit à l'aide d'une grille de Lebedev de $L = 50$ haut-parleurs fonctionnant jusqu'à l'ordre $M = 5$ [\[Lebedev, 1975\]](#); [\[Lecomte et al., 2015\]](#) est montré sur les figures du bas. Cette grille est présentée en détails au chapitre [3](#) de ce document. Pour calculer les signaux des haut-parleurs, modélisés comme des monopôles, l'équation [\(2.48\)](#) est utilisée. On observe sur cette figure que le champ sonore reconstruit est équivalent au champ sonore tronqué de la série de Fourier-Bessel. Ainsi, si le dispositif de haut-parleurs est bien conçu, il est possible de reproduire un champ sonore équivalent à celui décrit par une série de Fourier-Bessel. La troncature du champ sonore décrite par la série de Fourier-Bessel est alors la principale limitation de la reproduction de champ sonore par la méthode ambisonique.

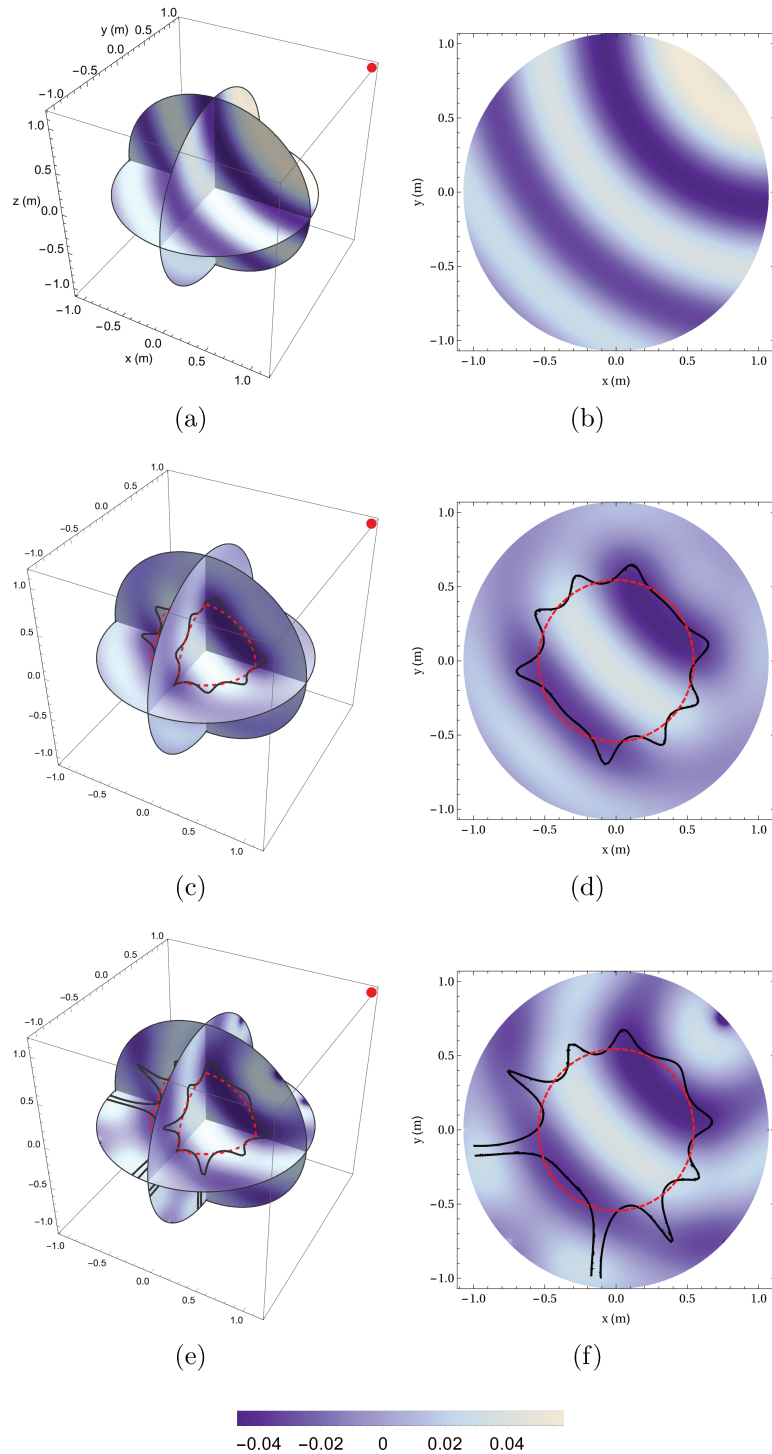


Figure 2.11 Partie réelle du champ de pression émis par un monopôle (en rouge). Amplitude $S = 1$, fréquence $f = 500$ Hz, position ($r_s = 2$ m, $\theta_s = 45^\circ$, $\delta_s = 35^\circ$). (a), (c) et (e) : vue en 3 dimensions dans les plans Oxy , Oxz et Oyz . (b), (d) et (f) : vue en 2 dimensions dans le plan horizontal Oxy . (a) et (b) : champ sonore cible. (c) et (d) : champ sonore tronqué à l'ordre $M = 5$. (e) et (f) : champ sonore reproduit à l'aide d'une grille de Lebedev de $L = 50$ haut-parleurs et de rayon de $r_{\text{spk}} = 1,07$ m. Le contour noir indique l'erreur quadratique de reconstruction $\epsilon = 0.04$. Le cercle en rouge pointillé est de rayon $r = M/k$.

2.3 Synthèse de fronts d'ondes

2.3.1 Présentation générale

La synthèse de fronts d'ondes ou Wave Field Synthesis (WFS) est une technique qui a été introduite par Berkhout *et al.* à la fin des années 80 [Berkhout, 1988; Berkhout *et al.*, 1993]. Cette approche permet de reconstruire ou de synthétiser des champs sonores sur une zone étendue de l'espace. La zone de reconstruction est délimitée par les haut-parleurs de restitution. La synthèse de fronts d'ondes repose sur deux principes fondamentaux : le principe de Huygens et les équations de Kirchhoff-Helmholtz. Une revue des équations de la synthèse de front d'ondes peut être trouvée par exemple dans les références suivantes : [Ahrens, 2012; Gauthier, 2007; Nicol, 1999; Spors *et al.*, 2008; Verheijen, 1998].

2.3.2 Principe de Huygens

Le principe de Huygens peut se formuler de la sorte : *Tout point atteint par l'ébranlement se comporte comme une source secondaire.* En appliquant ce principe à la propagation d'une onde acoustique, toute particule atteinte par l'onde devient une source secondaire. Autrement dit, on peut voir un front d'onde à l'instant t comme la superposition d'ondulettes rayonnées par des sources secondaires. Sur la figure 2.12 est illustré ce principe, et pour plus de détails, voir par exemple, [Nicol, 1999, p.99]. Ainsi, il est possible de re-

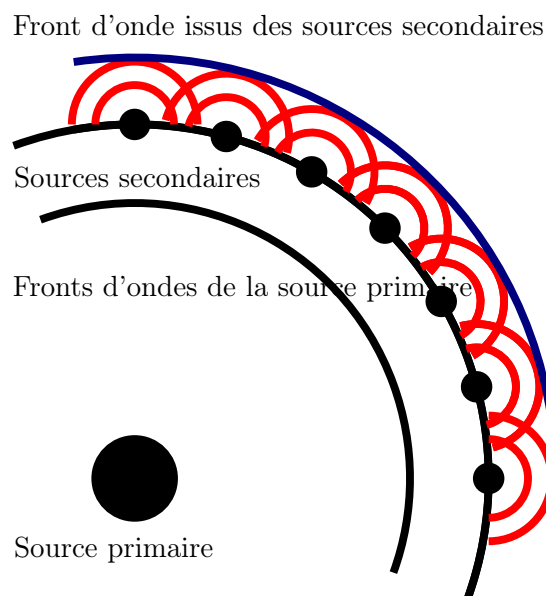


Figure 2.12 Le principe de Huygens appliqué à un front d'onde.

construire un front d'onde issu d'une source primaire à partir de la superposition d'ondes issues de sources secondaires.

2.3.3 Intégrale de Kirchhoff-Helmholtz

L'intégrale de Kirchhoff-Helmholtz est un outil important en acoustique. En effet, pour le problème intérieur ou extérieur [Williams, 1999], la formulation intégrale permet d'exprimer la pression en tout point $P(\mathbf{r})$ (avec $\mathbf{r} = (r, \theta, \delta)$) d'un volume V à partir de la connaissance de la pression p et de son gradient ∇p sur une surface ∂V délimitant le volume V . Un point sur la surface ∂V est de coordonnées \mathbf{r}_0 . Le problème de la détermination du champ acoustique dans un volume V (trois dimensions) peut donc se résoudre à partir d'une intégrale de surface (deux dimensions). Ce résultat est notamment à la base de la méthode des éléments finis de frontière en acoustique [Atalla et Sgard, 2015]. Pour la synthèse de champ sonores par synthèse de front d'ondes, on considère le problème intérieur : dans un domaine V exempt de sources acoustiques, la pression acoustique p au point $P(\mathbf{r})$ peut s'exprimer à l'aide de l'intégrale de Kirchhoff-Helmholtz :

$$\int_{\partial V} \left(p(\mathbf{r}_0) \frac{\partial G_0(\mathbf{r}|\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} - G_0(\mathbf{r}|\mathbf{r}_0) \frac{\partial p(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} \right) dS(\mathbf{r}_0) = \begin{cases} p(\mathbf{r}) & \mathbf{r} \in V \\ p(\mathbf{r})/2 & \mathbf{r} \in \partial V \\ 0 & \text{sinon} \end{cases} \quad (2.49)$$

Dans cette équation, $p(\mathbf{r})$ est la pression au point $P \in V$, $\mathbf{n}(\mathbf{r}_0)$ désigne la normale intérieure à la surface ∂V , au point \mathbf{r}_0 . La fonction de Green G_0 utilisée ici est la fonction de Green en espace libre définie par :

$$G_0(\mathbf{r}|\mathbf{r}_0) = \frac{1}{4\pi} \frac{e^{-ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|}. \quad (2.50)$$

L'équation (2.49) peut s'interpréter de la façon suivante : le champ acoustique à l'intérieur du volume V est entièrement décrit par le rayonnement d'une distribution de monopôles et de dipôles sur la surface ∂V . En reliant cela au principe de Huygens, on peut alors recréer un champ acoustique d'une source primaire virtuelle à partir de sources secondaires monopolaires et dipolaires réparties autour de la zone de reconstruction. Les signaux de commande des sources secondaires sont alors la pression issue de la source primaire pour le terme dipolaire et le gradient de pression pour le terme monopolaire [Nicol, 1999]. Le champ de pression à l'extérieur du volume V est identiquement nul.

2.3.4 Élimination du terme dipolaire

Pour des applications pratiques, il est d'usage d'éliminer le terme dipolaire $\partial G_0(\mathbf{r}|\mathbf{r}_0)/\partial \mathbf{n}(\mathbf{r}_0)$ dans l'équation (2.49). En effet, il est plus simple de réaliser des sources acoustiques à directivité monopolaire que dipolaire : une enceinte close par exemple. Pour ce faire, deux approches sont possibles [Spors *et al.*, 2008] : l'utilisation d'une fonction de Green de Neumann ou bien l'approche par simple source [Williams, 1999].

La fonction de Green de Neumann a une dérivée normale sur ∂V , nulle :

$$\frac{\partial G_N(\mathbf{r}|\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} = 0. \quad (2.51)$$

Elle doit également satisfaire à l'équation des ondes homogène à l'intérieur du volume V . Seules quelques géométries simples pour le volume V permettent de trouver une fonction de Green de Neumann analytique [Williams, 1999].

Intégrale de Rayleigh

Par exemple, lorsque la surface ∂V est un plan infini le terme dipolaire peut être éliminé par l'utilisation de la fonction de Green de Neumann $G_N(\mathbf{r}|\mathbf{r}_0) = 2G_0(\mathbf{r}|\mathbf{r}_0)$. L'équation de Kirchhoff-Helmholtz est alors substituée par l'intégrale de Rayleigh de type I :

$$p(\mathbf{r}) = - \int_{\partial V} 2G_0(\mathbf{r}|\mathbf{r}_0) \frac{\partial p(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} dS(\mathbf{r}_0). \quad (2.52)$$

Formulation simple source

Pour d'autres géométries, l'approche par simple source consiste à trouver la densité de simple couche $\mu(\mathbf{r}_0)$ telle que :

$$p(\mathbf{r}) = \int_{\partial V} \mu(\mathbf{r}_0) G_0(\mathbf{r}|\mathbf{r}_0) dS(\mathbf{r}_0). \quad (2.53)$$

Cela revient à poser conjointement un problème intérieur et extérieur, reliés par la même surface ∂V (la démonstration est faite en annexe B). La pression doit être continue sur la surface ∂V et le gradient de pression discontinu [Williams, 1999]. La fonction de Green utilisée est la fonction de Green en espace libre G_0 .

En utilisant ces formulations, il est possible de contrôler le champ acoustique dans le volume V à partir de sources de type monopolaire seulement.

2.3.5 Synthèse de fronts d'ondes en 2.5 dimensions

Dans la littérature, la synthèse de fronts d'ondes est souvent associée à un système de restitution en 2 dimensions. En effet, il est plus pratique et facilement réalisable d'entourer la zone de reproduction d'un contour fermé que d'une surface fermée.

En repartant de la formulation à 3 dimensions pour un plan infini (intégrale de Rayleigh de l'équation (2.52)), une approximation par phase stationnaire est utilisée pour réduire la dimension du problème et ne garder qu'une ligne horizontale des sources les plus utiles à la reconstruction [Verheijen, 1998]. La synthèse de fronts d'ondes en deux dimensions ne reproduit plus exactement les sources en dehors du plan dans lequel se situent les sources secondaires. Si l'on utilise des sources secondaires monopolaires, la reproduction en deux dimensions est faite avec des sources rayonnant dans l'espace à trois dimensions. Un terme correctif est alors introduit pour prendre en compte cette incohérence de dimensions : il s'agit d'une formulation 2.5 dimensions [Spors *et al.*, 2008].

2.3.6 Sélection des sources secondaires

Un critère de sélection des sources secondaires est introduit pour des géométries du système de restitution qui ne sont pas planes ou linéaires [Nicol et Emerit, 1999; Spors, 2007]. En effet, pour éliminer le terme dipolaire, on utilise souvent une fonction de Green de type Neumann pour un plan infini : $2G_0$. Si la géométrie du système de restitution est différente, cette fonction de Green n'est plus exacte. Il faut donc sélectionner au mieux les sources de restitution pour limiter les erreurs de reconstruction. Spors propose de sélectionner les sources suivant un critère d'intensité [Spors, 2007] : si le produit scalaire entre le vecteur d'intensité de la source primaire à la position de la source secondaire et la normale est positif, la source secondaire est active, sinon elle reste muette :

$$a(\mathbf{r}_0) = \begin{cases} 1, & \mathbf{I}(\mathbf{r}_0) \cdot \mathbf{n}(\mathbf{r}_0) > 0 \\ 0, & \text{sinon} \end{cases}, \quad (2.54)$$

où \cdot représente le produit scalaire Euclidien et \mathbf{I} le vecteur d'intensité acoustique. Ce dernier est défini par :

$$\mathbf{I}(\mathbf{r}) = \frac{1}{2} \Re(p(\mathbf{r})\mathbf{v}^*(\mathbf{r})), \quad (2.55)$$

où \Re représente la partie réelle et \mathbf{v}^* le conjugué de la vitesse acoustique \mathbf{v} . La formulation de synthèse de front d'ondes devient alors [Spors *et al.*, 2008] :

$$p(\mathbf{r}) = - \int_{\partial V} 2a(\mathbf{r}_0) \frac{\partial p(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} G_0(\mathbf{r}|\mathbf{r}_0) dS(\mathbf{r}_0). \quad (2.56)$$

2.3.7 Sources focalisées

Pour la synthèse de front d'ondes, (tout comme en ambisonie [Ahrens et Spors, 2008b; Daniel, 2003]), il est possible de reproduire le champ acoustique de sources situées à l'intérieur de la zone de restitution. Les signaux des haut-parleurs de restitution sont alors retournés temporellement [Yon *et al.*, 2003], par rapport à la situation où la source virtuelle n'est pas focalisée. Les front d'ondes convergent alors vers le point de focalisation avant de diverger à nouveau [Verheijen, 1998]. La zone de reconstruction est alors limitée par la position du point de focalisation. Lors de la reconstruction, des artefacts de types pré-échos peuvent survenir [Spors *et al.*, 2009; Wierstorf *et al.*, 2013].

2.3.8 Artefacts de reproduction

Les formulations présentées ci-dessus supposent une distribution continue des sources de restitution. Cependant, il n'est possible de n'avoir qu'un nombre fini de sources de restitution pour des applications pratiques. Les sources sont généralement des haut-parleurs montés en enceinte, pour imiter au mieux une source monopolaire. Elle sont réparties le long d'un segment linéaire pour la synthèse de fronts d'ondes 2.5 dimensions. Le caractère fini du réseau de haut-parleurs et l'espacement entre les sources, impliquent des artefacts dans le champ acoustique reconstruit. Ainsi, du repliement spatial est introduit du fait de la discrétisation spatiale. L'étude de ces effets de repliement spatial n'est faite de manière analytique que pour des géométries simples : des haut-parleurs régulièrement espacés, géométries linéaires ou circulaires [Spors et Rabenstein, 2006]. De même, le fait que le dispositif de restitution n'a pas de dimensions infinies, et n'est souvent pas un contour fermé, introduit une erreur de troncature et de diffraction par les coins du réseau [Boone *et al.*, 1995; Verheijen, 1998].

2.3.9 Synthèse de fronts d'ondes adaptative

Des méthodes de synthèse adaptative (adaptée au lieu de restitution) ont été appliquées à la synthèse de fronts d'ondes pour permettre de limiter les effets de la réverbération du lieu de restitution. Spors propose de passer par une décomposition du champ en ondes planes

[Spors *et al.*, 2003]. Corteel propose une approche basée sur une égalisation multi-points couplée à la synthèse de fronts d'ondes [Corteel, 2006]. Enfin, Gauthier et Berry proposent d'adapter le contrôle actif aux opérateurs de la synthèse de front d'ondes [Gauthier et Berry, 2007].

2.3.10 La synthèse de front d'ondes comme approximation hautes-fréquences de l'ambisonie

Les deux techniques d'ambisonie et de synthèse de fronts d'ondes ont historiquement évolué séparément. Cependant, plusieurs parallèles ont été établis récemment. Plusieurs études comparées sur différentes propriétés spatio-temporelles des champs reconstruits sont disponibles dans la littérature [Ahrens *et al.*, 2010; Daniel *et al.*, 2003; Spors et Ahrens, 2008].

Approche par simple source

Pour la synthèse de front d'ondes sur une géométrie sphérique, l'approche par simple source présentée section 2.3.4 est en réalité équivalente à l'approche ambisonique. En effet, la résolution de l'équation (2.53), en passant par les harmoniques sphérique, donne le même résultats que le décodage ambisonique. La démonstration est faite à l'annexe B.

Un problème de diffraction équivalent

Lorsqu'on utilise l'approche avec la fonction de Green de Neumann d'un plan infini, $2G_0$, avec un critère de sélection des sources actives tel que présenté à la section 2.3.6, les formulations ambisoniques et de synthèse de front d'ondes ne sont plus équivalentes. Fazi montre que la formulation de synthèse de front d'ondes est équivalente à un problème de diffraction d'un objet de volume V , dont la surface ∂V respecte une condition de Dirichlet ($p(\mathbf{r}) = 0$ sur ∂V) [Fazi et Nelson, 2013]. Si l'objet est convexe, comme une sphère par exemple, une approximation haute-fréquence permet de résoudre facilement l'équation de Kirchhoff-Helmholtz (2.49). La surface de l'objet diffractant peut être séparée en deux régions [Colton et Kress, 2013b] : une région illuminée par le champ incident ∂V_+ et une région obscure ∂V_- . Si le champ incident est une onde plane de direction de propagation \mathbf{r}_s , les régions illuminées et obscures sont définies par :

$$\begin{aligned} \partial V_+ &= \{\mathbf{r}_0 \in \partial V, \mathbf{n}(\mathbf{r}_0) \cdot \mathbf{r}_s > 0\} \\ \partial V_- &= \{\mathbf{r}_0 \in \partial V, \mathbf{n}(\mathbf{r}_0) \cdot \mathbf{r}_s \leq 0\} \end{aligned} \quad (2.57)$$

où \cdot représente le produit scalaire Euclidien. Pour les petites longueurs d'ondes, en première approximation, la pression $p(\mathbf{r}_0)$ et son gradient $\partial p(\mathbf{r}_0)/\partial \mathbf{n}(\mathbf{r}_0)$ sont nuls sur ∂V_- . De même, pour les petites longueurs d'ondes, la région ∂V_+ est alors vue comme plane pour le champ incident, et la formulation utilisant l'équation de Rayleigh de l'équation (2.52) peut être utilisée.

Ainsi, pour une géométrie convexe, l'utilisation d'une fonction de Green de type Neumann avec un critère de sélection des sources, est une approximation hautes-fréquences de l'équation de Kirchoff-Helmoltz, et pour la géométrie sphérique, de l'approche par simple source.

Pour la géométrie sphérique, l'approche par simple source est directement la formulation continue du décodage ambisonique (voir l'annexe B). Ainsi l'utilisation d'une formulation de synthèse de fronts d'ondes avec critère de sélection de sources, dans ce cas, permet d'économiser le nombre de sources actives, mais au détriment de la zone de reconstruction.

2.3.11 Exemple en simulation

Un exemple en simulation dans les mêmes conditions que pour l'ambisonie (voir la section 2.2.13) est traité ici. Le champ cible est celui issu par un monopôle, d'amplitude $S = 1$, de fréquence $f = 500$ Hz, à la position $\mathbf{r}_s = (2 \text{ m}, 45^\circ, 35^\circ)$. La surface sphérique est de rayon r_{spk} . La discrétisation de la sphère utilise une quadrature de Lebedev à $L = 50$ noeuds [Lebedev, 1975], de coordonnées \mathbf{r}_l , ($l \in [1 \dots 50]$), telle que présentée au chapitre 3. En utilisant une fonction de Green de Neumann $2G_0$, et un critère de sélection des sources actives, la pression à l'intérieur du volume sphérique délimité par les haut-parleurs est approximée par :

$$\tilde{p}(\mathbf{r}) = -r_{\text{spk}}^2 \sum_{l=1}^L 2a(l)w(\mathbf{r}_l)G_0(\mathbf{r}|\mathbf{r}_l) \frac{\partial p(\mathbf{r}_l)}{\partial \mathbf{n}(\mathbf{r}_l)}, \quad (2.58)$$

où $w(\mathbf{r}_l)$ représente un poids appliquée au $l^{\text{ième}}$ haut-parleur, selon la règle de quadrature de Lebedev (voir le chapitre 3) et $a(\mathbf{r}_l)$ est le critère de sélection du $l^{\text{ième}}$ haut-parleur. Le critère $a(\mathbf{r}_l)$ utilisé ici est donné par [Nicol et Emerit, 1999] :

$$a(\mathbf{r}_l) = \begin{cases} 1, & \langle \mathbf{r}_l, \mathbf{n}(\mathbf{r}_l) \rangle > 0 \\ 0, & \text{sinon} \end{cases} \quad (2.59)$$

Les signaux de commande des sources secondaires sont exprimés par le gradient de pression $\partial p(\mathbf{r}_l)/\partial \mathbf{n}(\mathbf{r}_l)$ de la source à reproduire, calculé à l'emplacement de la source secondaire \mathbf{r}_l .

L'erreur de reproduction est définie comme :

$$\epsilon(k, r, \theta, \delta) = \frac{|p(k, r, \theta, \delta) - \tilde{p}(k, r, \theta, \delta)|^2}{|p(k, r, \theta, \delta)|^2}. \quad (2.60)$$

La figure 2.13 présente les résultats de simulations. Les figures du haut montrent la partie réelle du champ de pression cible en deux et trois dimensions. Les figures du bas montrent la partie réelle du champ de pression, reproduit à l'aide de la formulation de l'équation 2.58). La figure la plus en bas présente la grille de haut-parleurs utilisée ainsi que les sources actives, suivant le critère de sélection de l'équation 2.59). Dans le cas de cette simulation, le champ de pression reconstruit n'est pas semblable au champ cible. La zone de reconstruction à $\epsilon = 4\%$, définie par l'équation 2.60) est petite et décentrée. Dans ce cas de simulation, très peu de sources sont actives et la formulation utilisée est une approximation hautes fréquences de la formulation simple source pour la sphère. Le champ sonore résultant n'est alors pas reconstruit de manière optimale. Cependant, il est probable que la localisation perceptive pour un auditeur qui se placerait dans la sphère soit améliorée par rapport à l'ambisonie, car ici seule une portion des sources secondaires est active, ce qui limite le nombre de pré-échos, et améliore alors la localisation [Blauert, 1997].

2.3.12 Conclusions

L'objectif de ce chapitre était de donner un aperçu des différentes techniques de reproduction de champs sonores en trois dimensions, sur une zone étendue de l'espace.

La première partie du chapitre a présenté une revue de la méthode ambisonique d'ordre élevé en trois dimensions. Les étapes d'encodage, de captation de champs sonores, de transformations dans le domaine des harmoniques sphériques et de décodage ont été présentées ainsi que les principales équations associées. Ainsi, l'ambisonie, basée sur un formalisme utilisant les harmoniques sphériques, permet de reconstruire un champ sonore synthétique ou capté sur une zone étendue de l'espace. L'encodage du champ sonore, découplé de l'étape de décodage, permet une grande flexibilité d'utilisation et autorise des transformations sur le champ sonore dans le domaine des harmoniques sphériques. Cependant, quelques compromis sont à faire sur la reconstruction et la réalisation pratique d'un système de synthèse physique de champs sonores par ambisonie d'ordre élevé. Tout d'abord, la troncature de la série de Fourier-Bessel décrivant le champ sonore dans un volume sphérique limite la taille de la zone de reconstruction : plus la décomposition se fait à ordre élevé, plus grande sera la zone de reconstruction. De même, la zone de reconstruction est

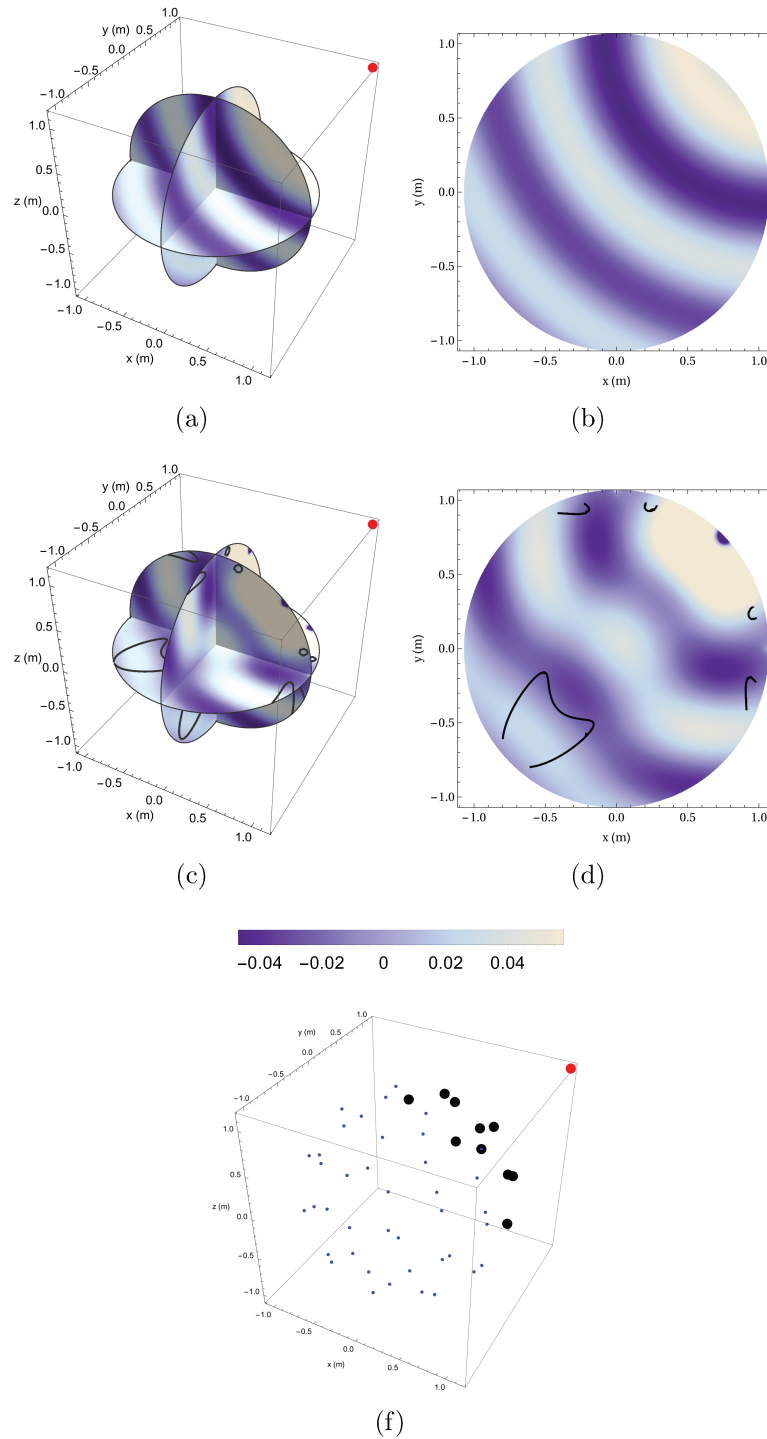


Figure 2.13 Partie réelle du champ de pression émis par un monopôle (en rouge). Amplitude $S = 1$, fréquence $f = 500$ Hz, position ($r_s = 2 \text{ tex}m, \theta_s = 45^\circ, \delta_s = 35^\circ$). (a) et (c) : vue en 3 dimensions dans les plans Oxy , Oxz et Oyz . (b) et (d) : vue en 2 dimensions dans le plan horizontal Oxy . (a) et (b) : champ sonore cible. (c) et (d) champ sonore reproduit par synthèse de front d'ondes sur une grille de Lebedev de $L = 50$ haut-parleurs et de rayon $r_{\text{spk}} = 1,07$ m, en utilisant la formulation de l'équation (2.58). Le contour noir indique l'erreur quadratique de reconstruction $\epsilon = 0.04$. (f) Les sources actives, selon l'équation (2.59) sont montrées en noir sur la vue en 3 dimensions de la grille de Lebedev utilisée.

dépendante de la fréquence : plus la fréquence est élevée, plus la taille de la zone diminue. Ainsi, l'ambisonie est rapidement limitée pour les reconstructions hautes-fréquences si l'ordre de troncature n'est pas suffisamment élevé. Les équations reposant sur un problème intérieur, la taille de la zone de reconstruction est alors délimitée par la position des sources de restitution, ou des sources encodées. Enfin, la captation, ou le décodage d'un champ sonore se fait en pratique à l'aide d'un nombre fini de transducteurs. La discrétisation des formulations continues, notamment la discrétisation des intégrales sur une surface sphérique introduit des artefacts de repliement spatial qui limitent également la reconstruction. Ces artefacts sont encore peu étudiés pour des géométries sphériques [Zotter *et al.*, 2009] et, à titre d'originalité, le chapitre 3 de ce document approfondit ce point notamment. Finalement, les systèmes de restitution 3D par ambisonie sont encore peu utilisés pour de la reconstruction physique de champ sonores, et plutôt orientés à des applications musicales ou multimédias. Ainsi, des résultats de mesure de champs sonores synthétisés ou reconstruits à partir d'enregistrements sont encore peu présents dans la littérature [Fazi, 2010]. En ce sens, le chapitre 6 comporte une originalité en apportant des résultats expérimentaux en cette voie.

La seconde partie de ce chapitre a présenté succinctement les équations de la synthèse de front d'ondes. Cette approche repose sur l'intégrale de Kirchhoff-Helmholtz et le principe de Huygens. La réalisation pratique de la formulation basée sur l'intégrale de Kirchhoff-Helmholtz est difficilement réalisable en trois dimensions car elle requiert des sources monopolaires et dipolaires coïncidentes, et en nombre très élevés. Ainsi, la littérature de synthèse de front d'ondes s'est alors plus orientée vers une reproduction en 2 (ou 2.5) dimensions. Pour la géométrie sphérique, l'approche par simple source cherchant à trouver une formulation n'utilisant que des sources monopolaires est équivalente à l'approche ambisonique.

La géométrie sphérique semble la plus naturelle pour la synthèse physique de champ sonore, car elle ne privilégie aucune direction de l'espace. Cette géométrie est retenue dans le présent document. Cependant, dans ce cas, l'utilisation d'une formulation de synthèse de front d'ondes reposant sur l'intégrale de Rayleigh et un critère de sélection de source est une approximation haute-fréquence de la formulation utilisée en ambisonie. Ainsi, la synthèse de champ sonore en trois dimensions à l'aide d'un réseau sphérique de haut-parleurs sera effectuée dans la suite du document avec les formulations de l'ambisonie.

CHAPITRE 3

UN MAILLAGE DE LEBEDEV A CINQUANTE NOEUDS ET SES APPLICATIONS À L'AMBISONIE

3.1 Avant-propos

Auteurs et affiliations :

- Pierre Lecomte, Candidat au doctorat, Université de Sherbrooke, Conservatoire National des Arts et Métiers.
- Philippe-Aubert Gauthier, Professionnel de recherche, Université de Sherbrooke.
- Christophe Langrenne, Ingénieur de recherche, Conservatoire National des Arts et Métiers.
- Alain Berry, Professeur, Université de Sherbrooke.
- Alexandre Garcia, Professeur, Conservatoire National des Arts et Métiers

Date d'acceptation : 22 Juillet 2016

État de l'acceptation Version finale acceptée, publiée.

Revue Journal of Audio Engineering Society

Référence [1] P. Lecomte, P.-A. Gauthier, C. Langrenne, A. Berry, and A. Garcia, "A Fifty-Node Lebedev Grid And Its Applications To Ambisonics," *J. Audio Eng. Soc.*, vol. 64, no. 11, 2016. [\[1\]](#)

¹Les travaux préliminaires à ce qui est présenté dans ce chapitre ont fait l'objet d'un article de conférence internationale : [Lecomte et al., 2015](#).

Résumé français Cet article étudie l'utilisation d'une grille de Lebedev à cinquante noeuds pour la capture et la reproduction de champs sonores par l'approche ambisonique. La grille en question est obtenue à l'aide d'une règle de quadrature pour la sphère invariante par rotation. La théorie associée à ce type de quadrature est rappelée. La grille de Lebedev est comparée avec une grille de Fliege et une grille *t-design* qui utilisent presque le même nombre de noeuds. La grille de Lebedev fournit alors les meilleures performances en terme de capture et reproduction de champs sonores. Finalement, un décodeur multi-bandes multi-ordres est présenté. Ce décodeur exploite avantageusement les sous-grilles incluses dans la grille de Lebedev à cinquantes noeuds.

3.2 Contribution au document

Comme mentionné à la fin du chapitre précédent, la géométrie sphérique est retenue pour la capture et la reproduction de champs sonores en trois dimensions, par approche ambisonique. La discrétisation des équations doit alors être faite avec précautions pour limiter les artefacts de repliement spatial et favoriser l'implémentation pratique. L'échantillonnage de la sphère fait l'objet de nombreuses recherches dans la littérature. En effet la répartition régulière de point sur la sphère ne peut se faire que dans les cas des polyèdres réguliers (solides platoniciens). Il est alors impossible d'équirépartir plus de vingt points à la surface d'une sphère, selon un dodécaèdre. Selon le champ d'application, plusieurs schémas de discrétisation sont alors proposés. Pour des applications ambisoniques, l'intégration des harmoniques sphériques est requise. Ce chapitre apporte alors des éléments de réponse à la problématique d'échantillonnage de la sphère dans le contexte de capture et reproduction de champs sonores par la méthode ambisonique. En reprenant les formulations continues, il est mis en évidence que la règle de quadrature utilisée doit respecter la propriété d'orthonormalité des harmoniques sphériques. Une règle de quadrature de Lebedev est choisie et comparée avec plusieurs autres fournissant quasiment le même nombre de noeuds. Le maillage de Lebedev se base sur les règles de quadrature de la sphère invariantes par rotation. La théorie de ces dernières, proposée par Sobolev, permet notamment d'intégrer exactement les harmoniques sphériques jusqu'à un ordre M avec un nombre fini de noeuds. Le chapitre présente également une étude détaillée du repliement spatial à la capture et à la reproduction. La dépendance fréquentielle du repliement est mise en évidence. Finalement, le maillage de Lebedev retenu semble alors être un bon choix pour l'ambisonie : il permet la capture et la synthèse de champ sonore jusqu'à l'ordre $M = 5$. Le repliement spatial est limité en comparaison à d'autres règles de quadrature utilisant quasiment le même nombre de noeuds. De plus, la répartition des noeuds sur la sphère

3.3. A FIFTY-NODE LEDEDEV GRID AND ITS APPLICATIONS TO AMBISONICS

est homogène, facilitant les contraintes de réalisation pratiques. Finalement, de par sa construction, la grille de Lebedev retenue abrite deux sous-grilles fonctionnant respectivement à l'ordre $M = 1$ et $M = 3$. Cette particularité autorise alors une grande flexibilité d'utilisation pour l'ambisonie. Pour une réalisation pratique, elle autorise notamment un décodage multi-bandes multi-ordres, qui permet d'utiliser des haut-parleurs adaptés à différentes bandes de fréquences, en nombre limité.

3.3 A Fifty-Node Lededev Grid and Its Applications to Ambisonics

3.4 Abstract

This paper investigates the use of fifty-node Lebedev grid for sound field capture and reproduction using Ambisonic approach. This grid is derived from rotationally-invariant quadrature rules, for which the underlying theory is recalled. Special attention is paid to spatial aliasing artifacts at capture and reproduction steps. While comparing a fifty-node Lebedev grid with a Fliege and a t -design grid that both use almost the same number of nodes, it is shown that the Lebedev grid provides the best performance in terms of sound field capture and reproduction. Finally, a multi-band multi-order decoder is presented. These decoders take advantages of the inherent nested sub-grids when following the rotationally-invariant quadrature approach.

3.5 Introduction

Physical reconstruction or synthesis of three-dimensionnal sound fields is an active research topic. It can be implemented notably with Near Field Compensated Higher Order Ambisonics [Daniel, 2000; Poletti, 2005] (simply called Ambisonics for the remainder of this paper). The captured (or modeled) sound field is decomposed on an orthonormal basis of spherical harmonics using spherical Fourier transform. The reconstruction is achieved over a finite region of space using secondary sources arranged in a spherical array in most cases. When using spherical harmonics approach (i.e. Ambisonics), the physical recording, encoding and decoding processes imply integration over a spherical surface and rely on the orthonormality of the spherical harmonics to avoid aliasing artifacts [Poletti, 2005]. Although the most recent Ambisonic formulations are derived with continuous spherical microphone or secondary source distributions [Ahrens et Spors, 2008a; Fazi et Nelson,

[2013; Fazi *et al.*, 2009; Poletti, 2005; Williams, 1999], a discretization step is mandatory for practical implementation. During the spatial sampling process, the orthonormality criterion should be preserved in order to avoid aliasing between spherical harmonics orders [Rafaely *et al.*, 2007]. Indeed, since the sound pressure field is not band-limited in term of spherical harmonics decomposition, the inter-order aliasing results in frequency-dependent spatial aliasing artifacts. This occurs at capture [Rafaely, 2005; Rafaely *et al.*, 2007] and at reproduction [Ahrens *et al.*, 2010; Poletti, 2005; Zotter *et al.*, 2009] steps. Many quadrature rules for the sphere are proposed in the literature. However, if one excludes the least-squares solutions, only a few of them fulfill the orthonormality criterion [Zotter, 2009b] which is important when using discrete formulations derived from continuous ones. Furthermore, they are valid (in a sense of very small orthonormality error) up to a finite order. Among them one can cite uniform angular sampling [Driscoll et Healy, 1994], pseudo-regular sampling [Moreau *et al.*, 2006], irregular sampling [Fliege et Maier, 1996; Hardin et Sloane, 1996; Li *et al.*, 2004; Rafaely, 2005] and rotationally invariant sampling [Ahrens et Beylkin, 2009; Lebedev, 1975]. Furthermore, the choice of a particular sampling scheme has to be done with practical implementation considerations, for example the distribution of the nodes on the sphere and their number [Zotter, 2009b]. Indeed, some sampling schemes concentrate the nodes on the poles of the sphere [Driscoll et Healy, 1994] and thus are sensitive to noise level for long wavelength. A large number of nodes also inevitably leads to a large number of electro-acoustic channels in the sound field capture or synthesis. One of the aim of this paper is the comparison of sampling grids with a similar number of channels, but with different quadrature rules. How this can affect capture and reproduction using Ambisonics formalism is one of the motivations for this work.

In this context, this paper investigates the potentials of a Lebedev quadrature rule with fifty nodes when used for Ambisonics. More generally, the rotationally-invariant grids present several advantages in the context of sound field capture and reproduction, namely exact integration up to finite spherical harmonic order [Williams *et al.*, 2006], homogeneous distribution on the sphere, relatively low number of nodes for a given integration order [Lebedev, 1975], slow degradation due to aliasing in recording and reproduction situations, and nested sub-grids allowing the use of different transducers with frequency bands in reproduction situation.

After recalling the definitions and notations in Sec. 3.6, the main equations of Ambisonics are derived in Sec. 3.7 with a continuous formulation. Section 3.8 presents the discretization process with the corresponding equations. Section 3.9 focuses on the rotationally-invariant

quadrature rules theory and presents the Lebedev approach. In Sec. 3.11, the orthonormality errors which are the cause of spatial aliasing artifacts are analyzed and compared for three grids with approximately fifty nodes. The aliasing artifacts are further analyzed for reproduction and capture in Sec. 3.12. Finally, in Sec. 3.13 the potential of nested sub-grids for the Lebedev quadrature is exploited in a context of multi-band multi-order decoder at reproduction step.

3.6 Definitions and Notations

Throughout this paper, bold case \mathbf{v} denotes a vector and bold upper case \mathbf{M} a matrix.

3.6.1 Spherical Coordinate System

The following spherical coordinate system is used and is shown in Fig. 3.1:

$$x = r \cos(\theta) \cos(\delta), \quad y = r \sin(\theta) \cos(\delta), \quad z = r \sin(\delta). \quad (3.1)$$

This convention, albeit being different from ISO standard [ISO8000-2, 2009], is more

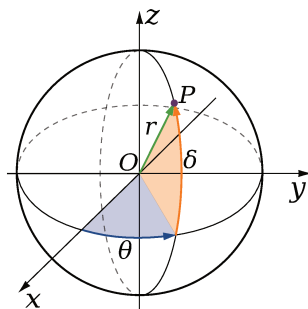


Figure 3.1 Spherical coordinate system in use. A point $P(x, y, z)$ is described by radius r , azimuth θ and elevation δ .

coherent with the literature on spatial hearing and spatial audio [Blauert, 1997].

3.6.2 Spherical Harmonics

The spherical harmonics used throughout this paper are the real-valued functions given by:

$$Y_{mn}(\theta, \delta) = \sqrt{(2m+1)\epsilon_n \frac{(m-|n|)!}{(m+|n|)!}} P_{m|n|}(\sin(\delta)) \times \begin{cases} \cos(|n|\theta) & \text{if } n \geq 0 \\ \sin(|n|\theta) & \text{if } n < 0 \end{cases}, \quad (3.2)$$

where $P_{m|n|}$ are the associated Legendre polynomials of order m and degree $|n|$, $(m, n) \in (\mathbb{N}, \mathbb{Z})$ with $|n| \leq m$, and $\epsilon_n = 1$ if $n = 0$, $\epsilon_n = 2$ if $|n| > 0$. In Eq. (3.2), m denotes the spherical harmonic order and n its degree. The family of spherical harmonics defines an orthonormal basis in \mathbb{L}^2 , the vectorial space of square-integrable functions on the unit sphere $\{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$. For each order m , there are $(2m + 1)$ spherical harmonics. Thus a basis truncated at order M contains $(M + 1)^2$ functions.

3.6.3 Spherical Scalar Product

Let f and g be two functions defined in \mathbb{L}^2 . The spherical scalar product of these two functions is defined as:

$$\langle f, g \rangle = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\frac{\pi}{2}}^{+\frac{\pi}{2}} f(\theta, \delta)g(\theta, \delta) \cos(\delta)d\delta d\theta. \quad (3.3)$$

The spherical harmonics are orthonormal in terms of the spherical scalar product:

$$\langle Y_{mn}, Y_{m'n'} \rangle = \delta_{mm'}\delta_{nn'}, \quad (3.4)$$

where δ_{ij} is the Kronecker symbol with $\delta_{ii} = 1 \forall i$ and $\delta_{ij} = 0 \forall i, j, i \neq j$.

3.6.4 Spherical Fourier Transform

The spherical Fourier transform (SFT) of a function f is defined as:

$$f_{mn} = \langle f, Y_{mn} \rangle. \quad (3.5)$$

Thus, f can be represented on the spherical harmonics basis as:

$$f(\theta, \delta) = \sum_{m=0}^{\infty} \sum_{n=-m}^m f_{mn}Y_{mn}(\theta, \delta). \quad (3.6)$$

3.7 Ambisonics

3.7.1 Sound Field Representation

The Ambisonic approach is based on the wave equation expressed in spherical coordinates. In sphere of radius r_s exempt of sources, with the implicit time dependence $e^{+i\omega t}$, the

Helmholtz equation is $\Delta p + k^2 p = 0$ ($\omega = 2\pi f$ with f the frequency, $k = 2\pi f/c$ the wave number, c the sound celerity, $i = \sqrt{-1}$). The spherical coordinate solution leads to the Fourier-Bessel description of the sound pressure field [Williams, 1999]:

$$p(k, r, \theta, \delta) = \sum_{m=0}^{+\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta). \quad (3.7)$$

Thus, the sound pressure field is fully described in a sphere centered at the origin O with a radial dependency given by the spherical Bessel functions $j_m(kr)$ and an angular dependency given by the spherical harmonics decomposition $B_{mn} Y_{mn}(\theta, \delta)$. This angular dependency, or spherical wave spectrum [Williams, 1999, p. 207], is the inverse spherical Fourier transform of the pressure sound field Eq. (3.6). The B_{mn} are called Ambisonic components and are frequency-dependent. They describe the angular spectrum of the sound field and correspond physically to the successive spatial derivatives of the pressure at the origin [Daniel, 2003].

One recalls the spherical harmonics decomposition of two types of simple waves:

Plane wave One assumes an harmonic plane wave of amplitude S , and propagating direction (θ_p, δ_p) . The B_{mn} components are given by [Rafaely, 2015]:

$$B_{mn} = S Y_{mn}(\theta_p, \delta_p). \quad (3.8)$$

Spherical wave For the interior problem, the expansion of a spherical wave of amplitude S and emitted by a point source at the position $(r_s, \theta_s, \delta_s)$ is given by [Williams, 1999]:

$$S \frac{e^{-ik|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|} = S \frac{-ik}{4\pi} \sum_{m=0}^{\infty} j_m(kr) h_m^{(2)}(kr_s) \sum_{n=-m}^m Y_{mn}(\theta_s, \delta_s) Y_{mn}(\theta, \delta), \quad (3.9)$$

where $\mathbf{r} = (r, \theta, \delta)$, is the observation and $\mathbf{r}_s = (r_s, \theta_s, \delta_s)$ is the position of the point source and $r < r_s$. The B_{mn} components are derived by identifying Eq. (3.9) with Eq. (3.7). They are given by [Rafaely, 2015]:

$$B_{mn} = S F_m(kr_s) Y_{mn}(\theta_s, \delta_s). \quad (3.10)$$

In the above equation, $F_m(kr_s) = i^{-(m+1)} k h_m^{(2)}(kr_s)/(4\pi)$ are the near-field filters introduced by Daniel [Daniel, 2003], where $h_m^{(2)}$ are the spherical Hankel functions of the second kind and of order m . The filters F_m take into account the finite distance of the source with

respect to the origin O . They are frequency-dependent and they depend on the order m . Note that S is the point source amplitude and not the amplitude of the wave at the origin O . Thus, the formulation of $F_m(kr_s)$ is slightly different to what Daniel proposes in [Daniel, 2003].

3.7.2 Truncated Decomposition and Truncation Error

In practical applications the sound field representation (Eq. (3.7)) is truncated up to order M . Thus, in three dimensions, the representation involves $L = (M + 1)^2$ spherical harmonics. The normalized error between the original sound field $p(k, r, \theta, \delta)$ and its truncated representation $p_M(k, r, \theta, \delta)$ is given by [Poletti, 2005]:

$$\epsilon_M(k, r, \theta, \delta) = \frac{|p(k, r, \theta, \delta) - p_M(k, r, \theta, \delta)|^2}{|p(k, r, \theta, \delta)|^2}. \quad (3.11)$$

The error presented above depends on the position (r, θ, δ) and frequency with k . The derivation of an average radial truncation error depending on k and r , is made by integrating Eq. (3.11) over all angles (θ, δ) . The demonstration involves the decomposition of $p(k, r, \theta, \delta)$ and $p_M(k, r, \theta, \delta)$ using a Fourier-Bessel representation (as in Eq. (3.7)). The derivation is made by Ward and Abhayapala in reference [Ward et Abhayapala, 2001] for a plane wave and by Zotter and *al.* in reference [Zotter *et al.*, 2009] for a spherical wave. The expressions are given by:

$$\bar{\epsilon}_M = 1 - \sum_{m=0}^M (2m + 1)(j_m(kr))^2 \quad (3.12)$$

for plane waves, and by:

$$\bar{\epsilon}_M = 1 - \frac{2k^2 r_s r}{\ln(\frac{r_s+r}{r_s-r})} \sum_{m=0}^M (2m + 1)(j_m(kr))^2 |h_m^{(2)}(kr_s)|^2 \quad (3.13)$$

for spherical waves. From the above equation, a “rule of thumb” is proposed by Ward and Abhayapala [Ward et Abhayapala, 2001] for a plane wave: $\bar{\epsilon}_M \approx 4\%$ (−14 dB) when:

$$[kr] = M, \quad (3.14)$$

where $[\cdot]$ represents the integer part of the argument. It means that if an order M is chosen for the truncation in Eq. (3.7), then the average radial error $\bar{\epsilon}_M$ is less than 4% when $kr \leq M$, and increases otherwise. Equation (3.14) was shown to work well for

spherical waves by Poletti in reference [Poletti, 2005]. It is worth noting that, in this case, the truncation error is valid as long as $r < r_s$, since the Fourier-Bessel representation of Eq. (3.7) is defined for the interior problem limited by r_s .

3.7.3 Sound Field Capture

An advantage of Ambisonics is the ability to record sound fields in three dimensions. Indeed, from Eq. (3.7) one notes that a sound pressure field is fully and uniquely described by the B_{mn} components. In order to estimate these components, one computes the SFT of the sound pressure field, known on a spherical surface of radius a :

$$p_{mn} = \langle p(k, a, \theta, \delta), Y_{mn}(\theta, \delta) \rangle. \quad (3.15)$$

If one knows the pressure on an open spherical surface, one could replace $p(k, a, \theta, \delta)$ by its expression Eq. (3.7) in Eq. (3.15) and obtain:

$$p_{mn, \text{open}} = \sum_{m'=0}^{\infty} i^{m'} j_{m'}(ka) \sum_{n'=-m'}^{m'} B_{m'n'} \langle Y_{m'n'}(\theta, \delta), Y_{mn}(\theta, \delta) \rangle. \quad (3.16)$$

Finally, the orthonormality property of spherical harmonics (see Eq. (3.4)) allows simplifying the previous equation:

$$B_{mn} = (i^m j_m(ka))^{-1} p_{mn, \text{open}}. \quad (3.17)$$

Thus, the successive Ambisonic components can be obtained by the projection of the pressure on the corresponding spherical harmonics and by multiplying by equalization functions $(i^m j_m(ka))^{-1}$. However, these equalization functions are not defined for some values of ka (the zeroes of $j_m(ka)$). In order to avoid undetermined values, one can use multiple-layer sensors [Williams, 1999], directive sensors, or a rigid sensing surface [Moreau *et al.*, 2006]. In the latter case, using the diffraction properties of a rigid sphere of radius a , and the orthonormality property of spherical harmonics, the Ambisonic components are obtained by projecting the pressure over a spherical rigid surface and applying equalization filters $E_m(ka)$ as follows:

$$B_{mn} = \underbrace{i^{-(m-1)} (ka)^2 h_m^{(2)}(ka)}_{E_m(ka)} p_{mn, \text{rigid}}, \quad (3.18)$$

where $h_m^{(2)}(kr)$ are the first derivatives relative to the variable kr of the function $h_m^{(2)}(kr)$. In the latter equation, the SFT of the pressure $p_{mn,\text{rigid}}$ is given by:

$$p_{mn,\text{rigid}} = \frac{1}{(ka)^2} \sum_{m'=0}^{\infty} \frac{i^{m'-1}}{h_{m'}^{(2)}(ka)} \sum_{n'=-m'}^{m'} B_{m'n'} \langle Y_{m'n'}(\theta, \delta), Y_{mn}(\theta, \delta) \rangle. \quad (3.19)$$

Note that in practice, for small ka and higher orders, the filters $E_m(ka)$ in Eq. (3.18) amplify by an unrealistic amount the SFT of the pressure $p_{mn,\text{rigid}}$ as the soundfield is dominated by first orders. Regularization methods would be used to circumvent this issue, as in [Moreau *et al.*, 2006] for instance.

3.7.4 Sound Field Reproduction With Continuous Secondary Source Distributions

The reproduction step aims at reconstructing the sound field knowing the B_{mn} components. Although the approach was first derived with a discrete number of secondary monopoles [Daniel, 2000], the formulation was established more recently using a continuous distribution of monopoles: It is known as the simple source approach [Ahrens et Spors, 2008a; Fazi et Nelson, 2013; Fazi *et al.*, 2009; Poletti, 2005; Williams, 1999]. By considering a continuous distribution of monopoles on a sphere of radius r_{spk} one gets:

$$\hat{p}(k, r, \theta, \delta) = \left\langle s(\theta_l, \delta_l), \frac{e^{-ik|\mathbf{r}-\mathbf{r}_l|}}{4\pi|\mathbf{r}-\mathbf{r}_l|} \right\rangle, \quad (3.20)$$

where $\hat{p}(k, r, \theta, \delta)$ is the reconstructed pressure, $s(\theta_l, \delta_l)$ is the driving signal of the point source and $\mathbf{r}_l = (r_{\text{spk}}, \theta_l, \delta_l)$. One could choose for the driving signal $s(\theta_l, \delta_l)$:

$$s(\theta_l, \delta_l) = \sum_{m=0}^{\infty} \sum_{n=-m}^m \frac{B_{mn} Y_{mn}(\theta_l, \delta_l)}{F_m(kr_{\text{spk}})}, \quad (3.21)$$

where B_{mn} are the Ambisonic components to be reproduced and F_m are the near field filters introduced in Eq. (3.10). By inserting Eq. (3.21) and Eq. (3.9) in Eq. (3.20) one

obtains:

$$\hat{p}(k, r, \theta, \delta) = \sum_{m=0}^{\infty} \sum_{m'=0}^{\infty} i^m j'_m(kr) \frac{h_{m'}^{(2)}(kr_{\text{spk}})}{h_m^{(2)}(kr_{\text{spk}})} \sum_{n=-m}^m \sum_{n'=-m'}^{m'} B_{mn} Y_{m'n'}(\theta, \delta) \times \langle Y_{mn}(\theta_l, \delta_l), Y_{m'n'}(\theta_l, \delta_l) \rangle. \quad (3.22)$$

The orthonormality of spherical harmonics allows for the simplification of Eq. (3.22) into a Fourier-Bessel series given by Eq. (3.7), so that $\hat{p}(k, r, \theta, \delta) = p(k, r, \theta, \delta)$.

3.8 Discretization

As shown in the previous sections, the orthonormality criterion of the spherical harmonics helps simplifying the continuous formulation in order to derive the Ambisonic components B_{mn} of a recorded sound field or driving signals for the secondary source distribution. In real-world applications, continuous realizations of spherical microphone or loudspeaker arrays are impossible. Thus the discretization scheme of continuous formulations, if not formulated as a least-squares problem, should preserve the orthonormality criterion of spherical harmonics up to an order M as high as possible. Indeed, orthonormality errors will lead to spatial artifacts in the encoded or reconstructed sound field according to Eqs. (3.16) and (3.22). Moreover, even if the sampling scheme satisfies orthonormality up to order M , the truncated decomposition will lead to an "operating frequency range" given by $kr \leq M$ [Rafaely *et al.*, 2007] (where r is either a fixed radius of the reconstruction zone or the radius of a spherical microphone array).

3.8.1 Quadrature Rules

The scalar product of Eq. (3.3) relies on an integration over a spherical surface. This integral should be computed as accurately as possible with only a finite number of samples over the spherical surface. Also, for Ambisonic reproduction, the discrete scalar product of two spherical harmonics should respect the orthonormality criterion.

Let φ be a function in \mathbb{L}^2 . On the unit sphere, a quadrature rule J is defined as the discrete approximation:

$$J(\varphi) \equiv \sum_{l=1}^L w_l \varphi(\theta_l, \delta_l) \cong \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\frac{\pi}{2}}^{+\frac{\pi}{2}} \varphi(\theta, \delta) \cos(\delta) d\delta d\theta. \quad (3.23)$$

When $\varphi = Y_{mn}Y_{m'n'}$ (in this case corresponding to $\langle Y_{mn}(\theta_l, \delta_l), Y_{m'n'}(\theta_l, \delta_l) \rangle$ as encountered in Eqs. (3.16) and (3.22)), one should have:

$$\sum_{l=1}^L w_l Y_{mn}(\theta_l, \delta_l) Y_{m'n'}(\theta_l, \delta_l) = \delta_{mm'} \delta_{nn'} + \epsilon_{mm'nn'}. \quad (3.24)$$

In Eqs. (3.23) and (3.24), L represents the number of nodes on the spherical surface, w_l are the weights “introduced to support the equality” [Rafaely, 2005] and (θ_l, δ_l) are the spherical coordinates of node l . The quadrature rule should be as accurate as possible up to an order M' as high as possible (the choice of the notation M' instead of M here will be clarified in Sec. 3.9.5). The function φ should be band-limited in its spherical harmonics decomposition to avoid aliasing effects [Rafaely *et al.*, 2007]. The term $\epsilon_{mm'nn'}$ represents the orthonormality error and should be as small as possible up to order $(m + m') \leq M'$.

Discrete Spherical Fourier Transform The discrete spherical Fourier transform (DSFT) of a function f is defined as:

$$\hat{f}_{mn} = \sum_{l=1}^L w_l f(\theta_l, \delta_l) Y_{mn}(\theta_l, \delta_l). \quad (3.25)$$

3.9 Rotationally-Invariant Quadratures for the Sphere

This section focuses on the rotationally-invariant quadratures for the sphere. These quadrature rules exactly integrate a linear combination of spherical harmonics up to order M' . Thus, in Eq. (3.23), the approximation symbol becomes an equality under order M' . The theory behind these quadratures was introduced by Sobolev [Sobolev, 2006, p. 461]. In this section, the main steps to solve a quadrature rule based on rotationally invariance properties are recalled. More details are provided in references [Ahrens et Beylkin, 2009; Sobolev et Vaskevich, 1997]. The applications of these types of quadrature rules for Ambisonics will be discussed in the next sections.

3.9.1 Rotationally-Invariant Quadrature

Finding a quadrature rule explicitly valid up to an order M' requires finding the number of nodes L , their positions (θ_l, δ_l) and their weights w_l to fulfill Eqs. (3.23) and (3.24). These conditions involve a large number of equations and unknowns. Several theorems derived by Sobolev help reduce the number of equations if one requires the quadrature to be rotationally invariant. Let G be a rotation group in \mathbb{L}^2 of dimension D , this group contains the rotation operators of the unit sphere g_1, g_2, \dots, g_D . (For example, g_1 could be a rotation of $+180^\circ$ around z -axis, g_2 a rotation of -90° around x -axis, and so on.) The quadrature rule J given by Eq. (3.23) is said rotationally-invariant on G or G -invariant if $\forall g \in G$:

$$\sum_{l=1}^L w_l \varphi(g(\theta_l, \delta_l)) = \sum_{l=1}^L w_l \varphi(\theta_l, \delta_l). \quad (3.26)$$

As a consequence, the L nodes of J all end up sitting at the same positions in space under the action of G . From now on, J is assumed to be G -invariant.

3.9.2 Sobolev's Theorems

Several Sobolev's theorems exploit the G -invariant property of J . They are recalled in this section with the formulation adapted in the context of the present paper.

First theorem [Sobolev et Vaskevich, 1997, p. 47, Theorem 2.3]

Let Ω be a $(2m + 1)$ -dimensional linear subspace in \mathbb{L}^2 spanned by linear combination of spherical harmonics of order m which define a basis of Ω .

J is exact for all functions in Ω if and only if it is exact for all invariant functions of this order m , denoted Y_m^* . i.e. the functions unchanged under rotation belonging to G :

$$Y_m^*(g(\theta, \delta)) = Y_m^*(\theta, \delta). \quad (3.27)$$

As stated in [Sobolev et Vaskevich, 1997], given a spherical harmonic Y_{mn} of degree m , with $|n| \leq m$, one can construct an invariant function Y_m^* by the formula:

$$Y_m^*(\theta, \delta) = \frac{1}{D} \sum_{j=1}^D Y_{mn}(g_j(\theta, \delta)) \quad (3.28)$$

Moreover, all invariant functions are generated using this equation. Thus, if one requires the quadrature rule to be exact up to an order M' , this theorem helps to drastically

reducing the number of equations to solve in order to explicitly find J [Ahrens et Beylkin, 2009]. The system to solve involves only the invariant functions and the result will apply for all functions φ in Eq. (3.23) as long as φ has no components in its SFT for order $> M'$. Before recalling the second Sobolev's theorem, one needs to calculate the dimension D of the group G .

Dimension D of the group G One considers a regular polyhedron among: Tetrahedron (4 faces), octahedron (8 faces), icosahedron (20 faces): the rotation group under which the polyhedron is kept invariant is denoted G_4, G_8 or G_{20} respectively (where the subscript stands for the number of faces of the polyhedron). If G is one of these groups, the rotation operators $g_1, g_2 \cdots g_D$ are explicitly given. G contains the identity operator and proper rotations around the following axes: from center through a vertex, from center through the center of a face and from center through a middle point of an edge. One calls t_1 the number of vertices of the polyhedron, t_2 the number of faces and t_3 the number of edges [Sobolev, 2006]. At a vertex, q_1 edges meet. The faces of the polyhedron are q_2 -gons. The edges are the axes of rotations of order $q_3 = 2$ [Sobolev, 2006, p. 462]. The respective rotation angles are given by $2\pi k_1/q_1, 2\pi k_2/q_2, 2\pi k_3/q_3$, where $k_j \in [0, \dots, q_j - 1]$. Finally the dimension D of G is given by [Sobolev, 2006, p. 462]:

$$D = t_1 q_1 = t_2 q_2 = t_3 q_3. \quad (3.29)$$

For example, let us consider the octahedron ($G = G_8$). One has $t_1 = 6, t_2 = 8, t_3 = 12, q_1 = 4, q_2 = 3, q_3 = 2$. Then, $D = 24$. The rotation axes (black arrows) are shown in Fig. 3.2 for the octahedral group G_8 . The octahedron drawn in this figure is invariant under the 23 rotations and identity operator.

Second theorem [Sobolev, 2006, p. 463, Theorem 2] [Sobolev et Vaskevich, 1997, p. 51, Theorem 2.6]

From the first theorem it was shown that it is sufficient to find J exact (that fulfills Eq. (3.23)) for G -invariant function. The following theorem provides the number S of G -invariant function for a given order m :

$$S(m) = \left\lfloor \frac{m}{q_1} \right\rfloor + \left\lfloor \frac{m}{q_2} \right\rfloor + \left\lfloor \frac{m}{q_3} \right\rfloor + 1 - m. \quad (3.30)$$

In the next section, one defines rotation groups with reflections. These groups exploit symmetry properties which reduce $S(m)$ depending on the parity of m .

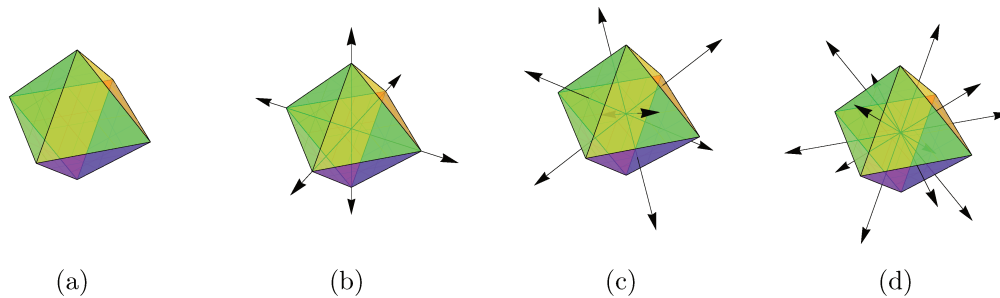


Figure 3.2 G_8 transformations for the octahedron, $D = 24$. The rotation axes are shown as black arrows. (a) Identity. (b) Six rotations of $+90^\circ$ and three rotations of $+180^\circ$ with rotation axes from origin to t_1 vertices. (c) Eight rotations of $+120^\circ$ with rotation axes from origin to center of t_2 faces. (d) Six rotations of $+180^\circ$ with rotation axes from origin to mid-point of t_3 edges.

Rotation group with reflections G^* The group G is of dimension D . By adding the point reflection operators with respect to the origin one constructs G^* which contains twice as many elements as G [Sobolev et Vaskevich, 1997, p. 61] (the reflection with respect to origin transforms a point of the unit sphere (θ, δ) into a point $(\pi - \theta, \pi + \delta)$). This group has a dimension $2D$. For example the octahedral group with reflections, G_8^* , has a dimension $D^* = 48$.

Third theorem [Sobolev et Vaskevich, 1997, p. 61]

The set of G^* -invariant functions of even order m coincides with the set of G -invariant functions. The group G^* has no invariant functions of odd order m . It means that:

$$\begin{aligned} S^*(m) &= S(m) && \text{for } m \text{ even,} \\ S^*(m) &= 0 && \text{for } m \text{ odd.} \end{aligned} \tag{3.31}$$

Thus, using a G^* group reduces again the number of equations to find J , by reducing the total number of invariant functions up to an order M' . Indeed, for odd orders, there are no invariant functions under the action of G^* .

3.9.3 Solving the Problem

One now looks for a quadrature rule J which exactly integrates spherical harmonics up to order M' . Thus, there are $(M' + 1)^2$ simultaneous equations to solve [Sobolev et Vaskevich, 1997, p. 66]. Hopefully, Sobolev's first theorem reduces the number of equations by restricting them to invariant functions. From the second theorem one can calculate $\sigma^*(M')$,

the total number of G^* -invariant functions up to order M' :

$$\sigma^*(M') = \sum_{m=0}^{M'} S^*(m). \quad (3.32)$$

The third theorem provides the relation between $S^*(m)$ and $S(m)$. To solve the system, one first chooses a G -invariant set of nodes, thus providing the number of unknowns to be found: Let U be this number. The matrix of the resulting system has dimensions $\sigma^*(M') \times U$. If one requires this matrix to be square, one finds the order M' by solving the following equation [Sobolev et Vaskevich, 1997, p. 67]:

$$\sigma^*(M') = U. \quad (3.33)$$

Some quadrature rules based on this theory are proposed for the group of rotations of the regular polyhedron among the tetrahedral group, the octahedral group and the icosahedral group [Ahrens et Beylkin, 2009; Sobolev et Vaskevich, 1997].

3.9.4 Lebedev Quadrature Rule

Lebedev proposed a method to construct quadrature formulas invariant under the octahedral rotation group with reflections G_8^* . Through elegant mathematical developments, he is able to reduce and solve the system of non-linear equations in order to find explicitly J [Lebedev, 1976]. In [Lebedev, 1975], he gives the results for several quadrature rules J up to $M' = 17$. In this paper, results of [Lebedev, 1975] are used for J .

Lebedev proposed to solve a Gauss-Markov type problem as defined in [Sobolev et Vaskevich, 1997, p. 44]:

"Given the nodes $x_l, l \in [1, \dots, L_1]$ with $L_1 \leq L$, find the weights w_1, \dots, w_{L_1} and the nodes x_{L_1+1}, \dots, x_L of quadrature formula so that it will be exact for all spherical harmonics up to the largest possible order M' ."

Considering the G_8^* group of rotations, a random point on the sphere generates 48 points under its action. However, if a point is the vertex of the inscribed octahedron, it generates only 6 distinct points. These points will be identified as $a^{(1)}$. In a similar way, a point which is the projection on the sphere of the mid-point of an edge generates 12 distinct points denoted $a^{(2)}$, a point which is the projection on the sphere of the center of a face generates 8 distinct points, denoted as $a^{(3)}$. Finally, particular points projected on the sphere from one bisectrix of a face generate 24 distinct points denoted $b^{(j)}$. The superscript (j) allows

several generating points in the same grid. Some of these particular points are shown in Fig. 3.3 for a single octant. Considering these particular points one of the quadrature rule

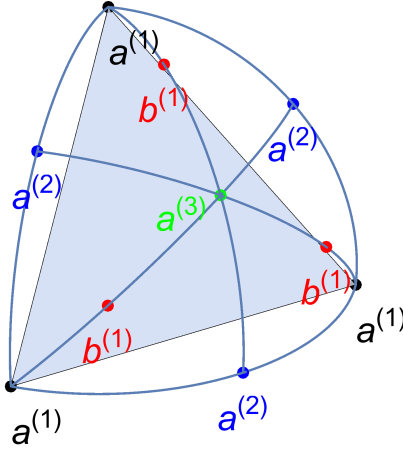


Figure 3.3 The different types of nodes $a^{(1)}, a^{(2)}, a^{(3)}$ and $b^{(1)}$ for a Lebedev quadrature up to order $M' = 11$, shown in a single octant. A face of the inscribed octahedron is displayed in light blue to help visualize the projections.

proposed in [Lebedev, 1975] is:

$$J_M(\varphi) = A_1 \sum_{i=1}^6 \varphi(a_i^{(1)}) + A_2 \sum_{i=1}^{12} \varphi(a_i^{(2)}) + A_3 \sum_{i=1}^8 \varphi(a_i^{(3)}) + \sum_{j=1}^{N_1} B_j \sum_{i=1}^{24} \varphi(b_i^{(j)}), \quad (3.34)$$

where

- $a_i^{(1)}$: $(0, 0, \pm 1)$, $(0, \pm 1, 0)$, $(\pm 1, 0, 0)$ are the six vertex points,
- $a_i^{(2)}$: $(\pm\sqrt{2}/2, \pm\sqrt{2}/2, 0)$, $(\pm\sqrt{2}/2, 0, \pm\sqrt{2}/2)$, $(0, \pm\sqrt{2}/2, \pm\sqrt{2}/2)$ are the twelve projections on the sphere of the middle of edges,
- $a_i^{(3)}$: $(\pm\sqrt{3}/3, \pm\sqrt{3}/3, 0)$, $(\pm\sqrt{3}/3, 0, \pm\sqrt{3}/3)$, $(0, \pm\sqrt{3}/3, \pm\sqrt{3}/3)$ are the eight projections on the sphere of the center of faces,
- $b_i^{(j)}$: $(\pm l_j, \pm l_j, \pm m_j)$, $(\pm l_j, \pm m_j, \pm l_j)$,
 $(\pm m_j, \pm l_j, \pm l_j)$, with $l_j = \sqrt{(1 - m_j^2)/2}$. m_j (or l_j) is given by the resolution of the problem [Lebedev, 1975].

Using the quadrature rule of Eq. (3.34) there are $L = 26 + 24N_1$ nodes but only $3 + 2N_1$ unknowns ($A_1, A_2, A_3, B_j, b_i^{(j)}$ (via l_j)). As an example, if one sets $N_1 = 1$. One has five unknowns: $A_1, A_2, A_3, B_1, b_1^1$. To have as many equations as unknowns one solves $\sigma^*(M') = 5$. For this example, from Eq. (3.32), $M' = 11$. From [Lebedev, 1975] the

following solution is provided:

$$\begin{aligned}
 N_1 = 1, \quad L = 50, \quad m_1 = 3/\sqrt{11}, \quad l_1 = 1/\sqrt{11}, \\
 A_1 = 4/315, \quad A_2 = 64/2835, \quad A_3 = 27/1280, \\
 B_1 = 11^4/725760. \quad (3.35)
 \end{aligned}$$

The grid obtained is shown in Fig. 3.4(d). It uses $L = 50$ points and it is able to integrate spherical harmonics up to order $M' = 11$. This grid has been used for vector intensity application [Williams *et al.*, 2006] or sound field reproduction with an integral approach [Choi et Kim, 2012].

3.9.5 Maximum Order of Integration for Ambisonics Purposes

When projecting the sound pressure $p(k, r, \theta, \delta)$ on the spherical harmonics basis using a quadrature rule, one tries to estimate the spherical Fourier transform components p_{mn} while computing the DSFT \tilde{p}_{mn} of Eq. (3.25). The grid constructed with rotationally-invariant quadrature allows exact integration of a linear combination of spherical harmonics up to order M' . From Eqs. (3.16) or (3.19), the function to integrate should be of order $m + m' \leq M'$ (product of two spherical harmonics functions of order m and m') to obtain an exact result [Rafaely, 2005; Williams *et al.*, 2006]. Thus, it is guaranteed that $\tilde{p}_{mn} = p_{mn}$ as long as m and $m' \leq \lfloor M'/2 \rfloor$. In the same way, for reproduction with finite numbers of loudspeakers the spherical harmonic decomposition of the reconstructed sound field $\hat{p}(k, r, \theta, \delta)$ in Eq. (3.22) matches the Fourier-Bessel decomposition of Eq. (3.7) as long as m and $m' \leq \lfloor M'/2 \rfloor$. In the case of the Lebedev quadrature Eq. (3.35), $\lfloor M'/2 \rfloor = M = 5$. For the remainder of this paper M is thus defined as the maximum order of integration in an Ambisonic context.

3.10 Applications to Capture, Encoding, Decoding and Panning Law for Ambisonics

3.10.1 Capture, Encoding and Decoding

In this section, one assumes a microphone and loudspeaker array using a grid with L nodes calculated from a quadrature rule exact up to order M . The recording, encoding

and decoding operations are summarized as follows in the frequency domain:

$$\mathbf{s}_{\text{spk}} = \underbrace{\mathbf{W}\mathbf{Y}\mathbf{F}_{\text{spk}}^{-1}}_{\text{decoding}} \underbrace{\mathbf{E}_{\text{mic}}\mathbf{Y}^T\mathbf{W}\mathbf{s}_{\text{mic}}}_{\text{encoding: } B_{mn}}, \quad (3.36)$$

where $\mathbf{s}_{\text{spk}}^{L \times 1} = [s_1, s_2, \dots, s_l, \dots, s_L]^T$ is the vector of input signals for the loudspeakers, $\mathbf{W}^{L \times L}$ is the diagonal matrix of quadrature weights, $\mathbf{Y}^{L \times (M+1)^2}$ is the matrix of spherical harmonics evaluated at the different nodes. $\mathbf{F}_{\text{spk}}^{-1 (M+1)^2 \times (M+1)^2}$ is the diagonal matrix of near field compensation filters with diagonal term $1/F_m(kr_{\text{spk}})$ (see Eq. (3.10)) where r_{spk} is the distance of the loudspeakers. $\mathbf{E}_{\text{mic}}^{(M+1)^2 \times (M+1)^2}$ is the diagonal matrix of equalization filters for the microphone array (see Eq. (3.18) for a rigid spherical microphone) and $\mathbf{s}_{\text{mic}}^{L \times 1}$ are the signals captured by the microphones. Transposition is indicated by the superscript T .

3.10.2 Encoding and Decoding of Plane or Spherical waves, Panning Law

For a virtual spatialization, i.e. without the use of a microphone array recording, the encoding and decoding of a plane or spherical wave is given by:

$$\begin{aligned} \mathbf{s}_{\text{spk}} &= S\mathbf{W}\mathbf{Y}\mathbf{F}_{\text{spk}}^{-1}\mathbf{Y}_{(\theta_s, \delta_s)} && \text{for plane waves,} \\ \mathbf{s}_{\text{spk}} &= S\mathbf{W}\mathbf{Y}\mathbf{H}\mathbf{Y}_{(\theta_s, \delta_s)} && \text{for spherical waves,} \end{aligned} \quad (3.37)$$

where S is the amplitude of the plane wave of propagating direction (θ_s, δ_s) or the amplitude of the monopole at position $(r_s, \theta_s, \delta_s)$ driving a spherical wave, $\mathbf{H}^{(M+1)^2 \times (M+1)^2}$ is the diagonal matrix of near field filters [Daniel, 2003] with diagonal terms $F_m(kr_s)/F_m(kr_{\text{spk}})$ and $\mathbf{Y}_{(\theta_s, \delta_s)}^{(M+1)^2 \times 1}$ is the vector of spherical harmonics evaluated in the direction (θ_s, δ_s) .

Using the additivity theorem of spherical harmonics [Arfken, 1985, pp. 693-698], the previous equation can be simplified as follows:

$$\begin{aligned} s_l &= Sw_l \sum_{m=0}^M \frac{(2m+1)}{F_m(k, r_{\text{spk}})} P_m(\gamma_l) && \text{for plane waves,} \\ s_l &= Sw_l \sum_{m=0}^M (2m+1) \frac{F_m(k, r_s)}{F_m(k, r_{\text{spk}})} P_m(\gamma_l) && \text{for spherical waves,} \end{aligned} \quad (3.38)$$

where $\gamma_l = \cos(\delta_s) \cos(\delta_l) \cos(\theta_s - \theta_l) + \sin(\delta_s) \sin(\delta_l)$ is related to the angle between the virtual source and the l -th loudspeaker, P_m is the Legendre polynomial of order m and w_l is the quadrature weight for the l -th loudspeaker.

3.11 Orthonormality Error

In order to display the orthonormality error (Eq. (3.24)) and evaluate aliasing errors, one computes the orthonormality error matrix \mathbf{D} :

$$\mathbf{D} = \mathbf{Y}^T \mathbf{W} \mathbf{Y} - \mathbf{I}, \quad (3.39)$$

where $\mathbf{I}^{(M+1)^2 \times (M+1)^2}$ is the identity matrix, $\mathbf{Y}^{L \times (M+1)^2}$, the matrix of spherical harmonics evaluated at the different nodes of the grid, and $\mathbf{W}^{L \times L}$, the matrix of quadrature weights. In this matrix, the spherical harmonics are ordered with increasing ACN (Ambisonic Channel Number [Nachbar *et al.*, 2011]) such that $\text{ACN}(m, n) = m^2 + m + n$. The error matrices $10 \log_{10} (|\mathbf{D}_{ij}|_{1 \leq i \leq (M+1)^2, 1 \leq j \leq (M+1)^2})$ are shown in Fig. 3.4 up to $M = 10$ for three grids with roughly the same number of nodes (around $L = 50$ nodes) but with different quadratures rules: Lebedev grid with $L = 50$ nodes Fig. 3.4(a), spherical t -design grid with $L = 50$ [Hardin et Sloane, 1996] Fig. 3.4(b) and Fliege grid with $L = 49$ [Fliege et Maier, 1996] Fig. 3.4(c). The corresponding grids are shown in Figs. 3.4(d), 3.4(e) and 3.4(f), respectively. The spherical t -design approach as an equally-weighted quadrature [Hardin et Sloane, 1996] and the Fliege grid is obtained by minimizing the potential energy of electrons on a conducting sphere [Fliege et Maier, 1996]. For the Lebedev grid (Fig. 3.4(a)), the orthonormality error vanishes up to order 5. More precisely, for all orders (m, m') such as $m + m' \leq 11$ or $m + m'$ is odd, the orthonormality error vanishes. This is related to the symmetry property of G_8^* presented in Sec. 3.9.2 and should be further analyzed. For the same number of nodes but a different quadrature rule, the t -design grid on Fig. 3.4(b) has its orthonormality error below -70 dB for order $m + m' \leq 8$. In the same way, the Fliege quadrature rule with $L = 49$ seems to work for order $m + m' \leq 6$. The choice of a maximum working order is a matter of how much aliasing due to orthonormality error is acceptable in the context of sound field recording or reconstruction. The following sections further investigates this point.

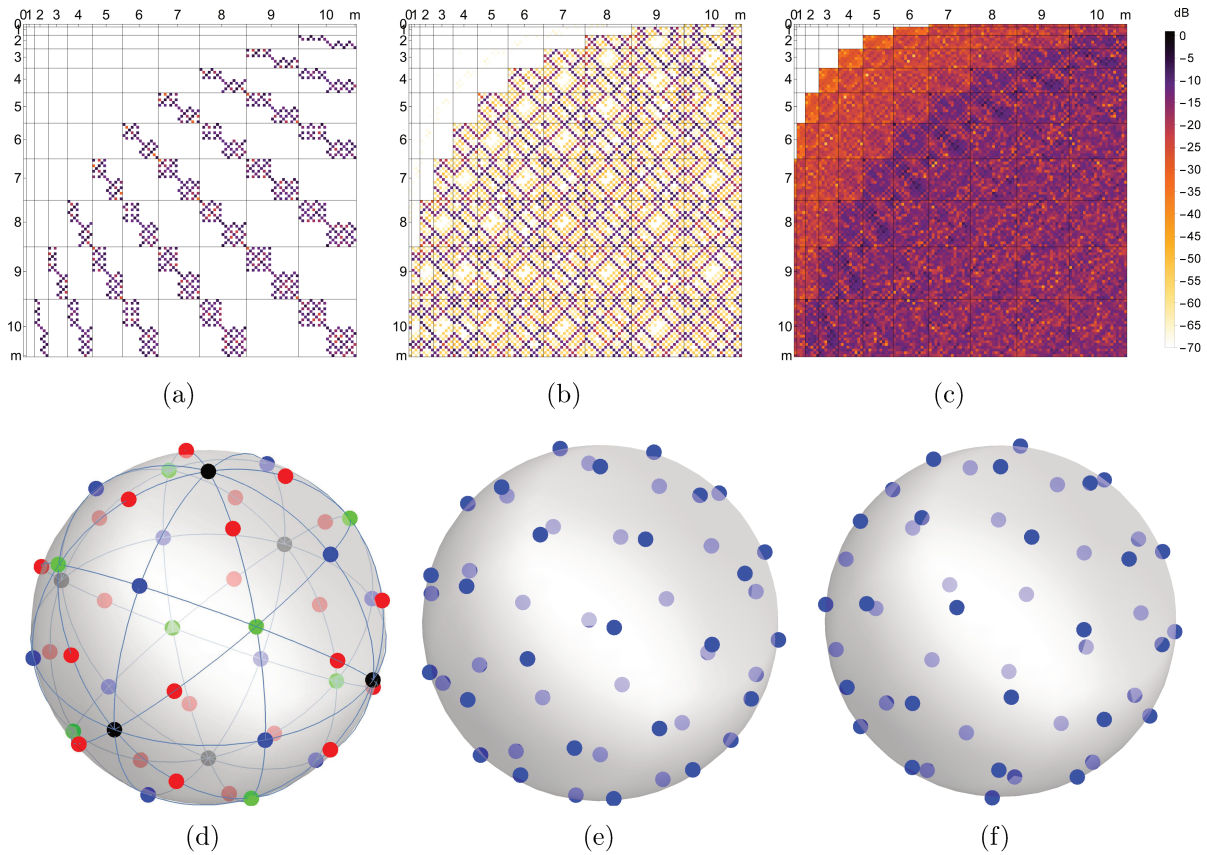


Figure 3.4 Orthonormality error matrices $10 \log_{10} (|\mathbf{D}_{ij}|_{1 \leq i \leq (M+1)^2, 1 \leq j \leq (M+1)^2})$ up to 10^{th} order and corresponding grids: Lebedev grid with 50 nodes ((a) and (d)), t -design grid with 50 nodes ((b) and (e)), Fliege grid with 49 nodes ((c) and (f)).

3.12 Aliasing Considerations

In this section, the aliasing is further illustrated. First, aliasing at reconstruction stage is discussed. Second, aliasing at capturing stage is presented.

3.12.1 Aliasing at Reconstruction

To evaluate the effect of discretization and orthonormality error on aliasing at reproduction, one recalls Eq. (3.22) with the quadrature rule Eq. (3.24) for the term $\langle Y_{mn}(\theta_l, \delta_l), Y_{m'n'}(\theta_l, \delta_l) \rangle$. The reconstructed pressure after truncation and discretization is

denoted \tilde{p} such that:

$$\tilde{p}(k, r, \theta, \delta) = p_M(k, r, \theta, \delta) + \sum_{m=0}^M \sum_{m'=0}^{\infty} i^m j_{m'}(kr) \frac{h_{m'}^{(2)}(kr_{\text{spk}})}{h_m^{(2)}(kr_{\text{spk}})} \sum_{n=-m}^m \sum_{n'=-m'}^{m'} B_{mn} Y_{m'n'}(\theta, \delta) \epsilon_{mm'n'n'}. \quad (3.40)$$

In this equation, p_M represents pressure of Eq. (3.7) truncated at the order M . Eq. (3.40) contains the effects of truncation and aliasing via discretization. More precisely, the second term on the right-hand side contains the aliasing effects due to discretization for a truncation order M . Note that the index m' in Eq. (3.40) starts at $m' = 0$ and not $m' = M + 1$. Indeed, even if the quadrature rule is said to work up to order M , the orthonormality errors for $0 \leq M$ is not necessary exactly null. In order to compare the three grids presented Fig. 3.4, one uses the following normalized quadratic error:

$$\epsilon_{\text{alias},M}(k, r, \theta, \delta) = \frac{|\tilde{p}(k, r, \theta, \delta) - p_M(k, r, \theta, \delta)|^2}{|p_M(k, r, \theta, \delta)|^2}. \quad (3.41)$$

This equation expresses the aliasing error due to discretization when working at a truncation order M . The average error is given by integrating Eq. (3.41) for all directions (θ, δ) . The orthonormality criterion of spherical harmonics applies and one obtains:

$$\bar{\epsilon}_{\text{alias},M}(k, r) = \frac{\sum_{m=0}^M \sum_{m'=0}^{\infty} (j_{m'}(kr))^2 \left| \frac{h_{m'}^{(2)}(kr_{\text{spk}})}{h_m^{(2)}(kr_{\text{spk}})} \right|^2 \sum_{n=-m}^m \sum_{n'=-m'}^{m'} |B_{mn}|^2 \epsilon_{mm'n'n'}^2}{\sum_{m=0}^M j_m(kr)^2 \sum_{n=-m}^m |B_{mn}|^2}. \quad (3.42)$$

Equation (3.42) can be further averaged if the sound field is a single plane wave of propagating direction (θ_s, δ_s) , amplitude S , and wave number k , or a point source at position $(r_s, \theta_s, \delta_s)$, amplitude S and wave number k . Integrating Eq. (3.42) for all (θ_s, δ_s) , one obtains $\bar{\bar{\epsilon}}_{\text{alias},M}$. Thus, the term $|B_{mn}|^2$ in Eq. (3.42) is replaced as follows:

$$\begin{aligned} \langle B_{mn}, B_{mn}^* \rangle &= |S|^2 && \text{for plane waves,} \\ \langle B_{mn}, B_{mn}^* \rangle &= |S|^2 \left(\frac{k}{4\pi} \right)^2 |h_m(kr_s)|^2 && \text{for spherical waves,} \end{aligned} \quad (3.43)$$

where $(.)^*$ indicates complex conjugate of the argument. Figure 3.5 shows the average error $\bar{\bar{\epsilon}}_{\text{alias},M}$ for a unit amplitude point source at radius $r_s = 2$ m and frequency $f = 1000$ Hz for the three grids presented above with $r_{\text{spk}} = 1$ m. Two orders of truncation are displayed $M = 5$ (thick curves) and $M = 10$ (thin curves). Note that Eq. (3.43) involves an infinite

summation over m' which represents the infinite spherical harmonics decomposition of the monopoles used as reconstruction sources. From Fig. 3.5, one notes that the Lebedev

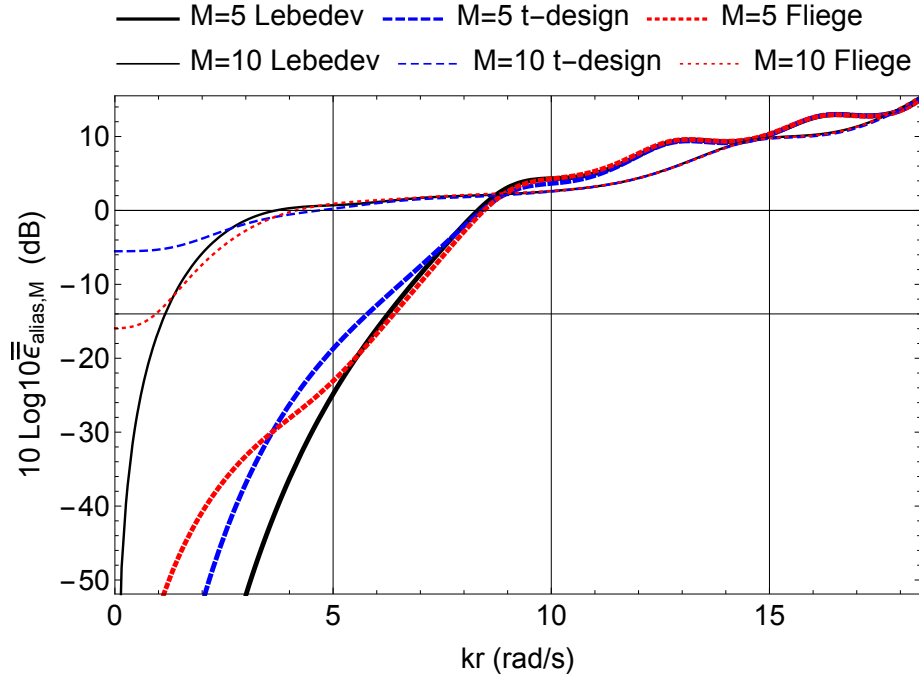


Figure 3.5 Average aliasing error $\bar{\epsilon}_{\text{alias},M}$ for a unit amplitude point source at radius $r_s = 2$ m and frequency $f = 1000$ Hz. Three grids are displayed: Lebedev grid with 50 nodes (black-continuous), Fliege grid with 49 nodes (red-dotted) and t -design grid with 50 nodes (blue-dashed) with $r_{\text{spk}} = 1$ m. Two truncation orders are displayed: $M = 5$ for the thick curves and $M = 10$ for the thin curves. Horizontal lines correspond to 0 and -14 dB.

suffers less from aliasing for both truncation orders $M = 5$ and $M = 10$. In the sweet-spot, as defined by Eq. (3.14), i.e. $kr \leq 5$ for $M = 5$, the aliasing error is below -18 dB for the t -design, -23 dB for Fliege grid and -25 dB for Lebedev grid and decreases with decreasing kr . This means that the aliasing error due to discretization is small in the sweet-spot compared to the truncated sound pressure field. However, when truncation order exceeds the maximum working order of the grids, the aliasing errors increase as well. Indeed, as shown in Fig. 3.4, the three grids are designed to work up maximum order $M = 3$ for Fliege grid, $M = 4$ for t -design and $M = 5$ for Lebedev. Thus for $M = 10$ the aliasing error in the sweet-spot is no longer negligible and around $+4$ dB for all grids when $kr = 10$. The aliasing errors become smaller with low kr , and the Lebedev grid seems to outperforms the other grids in this case. Beyond the sweet-spot, i.e. when $kr \gg M$, the truncated decomposition p_M tends to vanish [Rafaely, 2015] and the reconstruction is mainly due to aliasing errors.

3.12.2 Aliasing at Capture

The analysis at capture follows the same logic as in the previous section. For simplicity, the analysis is solely provided for the rigid spherical microphone where the radial filters involved have no singularities (see Eq. (3.18)). The estimated Ambisonic component \tilde{B}_{mn} with m and n fixed, is given by Eq. (3.18). By inserting Eq. (3.19) in (3.18) with quadrature rule Eq. (3.24) for the term $\langle Y_{mn}(\theta_l, \delta_l), Y_{m'n'}(\theta_l, \delta_l) \rangle$, \tilde{B}_{mn} is expressed as:

$$\tilde{B}_{mn} = B_{mn} + i^{-(m-1)} h_m^{(2)}(ka) \sum_{m'=0}^{\infty} \frac{i^{m'-1}}{h_{m'}^{(2)}(ka)} \sum_{n'=-m'}^{m'} B_{m'n'} \epsilon_{mm'n'n'}. \quad (3.44)$$

This equation emphasizes the fact that higher orders component $B_{m'n'}$ alias on lower orders component B_{mn} due to orthonormality errors $\epsilon_{mm'n'n'}$. The normalized quadratic error at capture is then defined as:

$$\epsilon_{\text{capt},mn} = \frac{|\tilde{B}_{mn} - B_{mn}|^2}{|B_{mn}|^2}. \quad (3.45)$$

It is possible to average this error over source directions (θ_s, δ_s) as in the previous section. Finally, the total average capture error for an estimation up to order M is given by

$$\bar{\epsilon}_{\text{capt},M} = \frac{\|\tilde{\mathbf{b}} - \mathbf{b}\|^2}{\|\mathbf{b}\|^2}, \quad (3.46)$$

where $\mathbf{b} = [B_{00} \cdots B_{MM}]^T$ and $\tilde{\mathbf{b}} = [\tilde{B}_{00} \cdots \tilde{B}_{MM}]^T$. Figure 3.6 presents the average aliasing error $\bar{\epsilon}_{\text{capt},M}$ for plane waves for the three aforementioned grids when using capture orders $M = 3$ (thick curves) and $M = 5$ (thin curves). For an estimation of Ambisonic components $\tilde{\mathbf{b}}$ up to order $M = 3$, one observes on Fig. 3.6 that the total average error $\bar{\epsilon}_{\text{capt},3}$ is lower for the Lebedev grid. Indeed, although sharing almost the same number of nodes, this grid is conceived to work up to higher orders compared to Fliege or t -design grids. The two vertical lines in Fig. 3.6 correspond to the aliasing wave number for $M = 3$ and $M = 5$ using Eq. (3.14). The three grids present an error $\bar{\epsilon}_{\text{capt},3} \leq -14$ dB for $ka \leq 3$. Moreover, the error curves cross this line at $ka \simeq 6.5$ for the t -design grid and at $ka \simeq 7$ for Lebedev and Fliege grids, therefore improving performance at high frequencies. However, the estimated Ambisonic components $\tilde{\mathbf{b}}$ at an order higher than the working order of the grid in consideration can present a total average aliasing error which drastically increases, especially for low ka . Indeed, according to the thin curves in Fig. 3.6 for an estimation up to order $M = 5$, only the Lebedev grid shows low values of error. One recalls that this grid is the only one which works up to this order. Compared to the case of $M = 3$ (thick curves), the Fliege and t -design grids present higher errors, due to the higher components

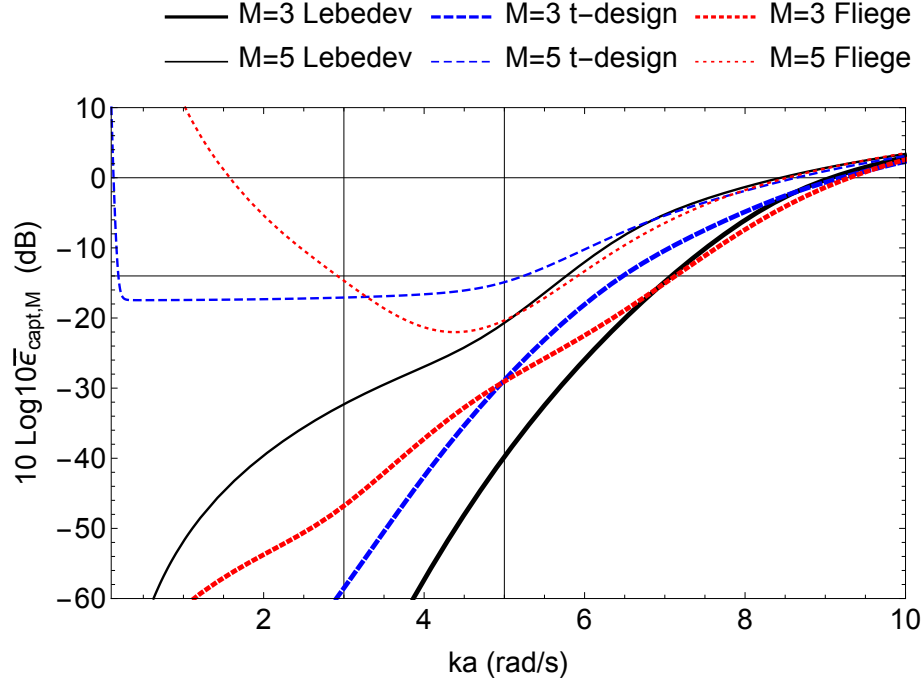


Figure 3.6 Total average aliasing error at capture for plane waves for three grids: Lebedev grid with 50 nodes (black-continuous), Fliege grid with 49 nodes (red-dotted) and t -design grid with 50 nodes (blue-dashed). Two maximum capture orders are displayed: $M = 3$ for the thick curves and $M = 6$ for the thin curves. Horizontal lines correspond to 0 and -14 dB.

aliasing error. Having a closer look at Eq. (3.44) from which Eq. (3.46) is constructed, one observes that the aliasing term is proportional to $(h_m^{(2)}/h_{m'}^{(2)})(ka)\epsilon_{mm'nn'}$. A simple function analysis shows that by decreasing ka and m' , and increasing m , the ratio of Hankel functions derivatives increase. Thus, if the orthonormality error $\epsilon_{mm'nn'}$ which weights this term is not small enough to cancel it, the resulting aliasing term will be high. In other words, the estimation of higher order components at low ka requires a grid which presents sufficiently small orthonormality errors at those higher orders to produce acceptable aliasing errors. This confirms the observations made from Fig. 3.6 for the Fliege and t -design grids at maximum estimation order $M = 5$.

3.13 Multi-Band Multi-Order Decoders

Rotationally-invariant quadratures have interesting advantages for Ambisonics. Indeed, the quadratures are classically constructed by stacking several sets of nodes which are invariant under the rotation group, as explained in Sec. 3.9.4. This results in several sub-grids nested in the whole resulting grid. Recalling the Lebedev quadrature rule Eq. (3.34) as an example,

one notes that the nodes $a_i^{(1)}, a_i^{(2)}, a_i^{(3)}$ are fixed, and shared between some configurations depending on the order of integration [Lebedev, 1975]. This grid contains two nested sub-grids: The first works up to order $M = 1$ with $L = 6$ nodes of type $a_i^{(1)}$ (only the black nodes in Fig. 3.4(d)) and parameters $A_1 = 1/6, A_2 = A_3 = B_k = 0 \forall k$. The second works up to order $M = 3$ with $L = 26$ nodes of types $a_i^{(1)}, a_i^{(2)}$ and $a_i^{(3)}$ (black, blue and green nodes in Fig. 3.4(d)) and parameters $A_1 = 1/21, A_2 = 4/105, A_3 = 9/280, B_k = 0 \forall k$. Thus, the position of the nodes are fixed and only the weight values change between each configurations for the nodes in common. This nesting property provides a practical set-up to decode a sound pressure field up to an order M with different numbers of nodes. For example, using this setup, it is possible to encode or decode a sound pressure field up to order 3 with either 26 or 50 microphones and loudspeakers.

Another advantage of several nested grids at rendering stage is the use of different types of loudspeakers as function of order and frequency. According to the rule of thumb Eq. (3.14), if one targets a reconstruction radius r , the lower orders contribute to the reconstruction at low frequencies and as the frequency increases the order should increase as well. Using small loudspeakers acting as monopole sources is realistic for mid frequencies. However, their response is often poor at low frequencies because of their small size. This limitation can be circumvented by having larger loudspeakers with an increased bandwidth. Only a few are required to physically reconstruct the sound pressure field at low frequencies, according to Eq. (3.14). For instance, on the Lebedev grid shown in Fig. 3.4(d), it is possible to design a mixed-band decoder. For the case of a two-band decoder one could use a layout with only six loudspeakers (at black nodes position in Fig 3.4(d)) for the low frequencies up to order $M = 1$ and then rely on those six loudspeakers combined with the 44 remaining (at blue, green and red nodes position in Fig. 3.4(d)) for higher frequencies up to order $M = 5$. The signals of the six larger loudspeakers are the summation of the decoding signals up to 1st order on a six loudspeakers layout (low frequencies) and those for order up to 5 on a 50 loudspeakers layout (higher frequencies), as shown in Fig. 3.7. Figure 3.8 shows some simulation results exploiting these nested grids. The target sound pressure field is composed of two point-sources at respective positions ($r_{s_1} = 1.5 \text{ m}, \theta_{s_1} = 90^\circ, \delta_{s_1} = 0^\circ$) and ($r_{s_2} = 1.5 \text{ m}, \theta_{s_2} = 0^\circ, \delta_{s_2} = 0^\circ$) and amplitudes $S_1 = 1, S_2 = -1$. Furthermore, these sources emit simultaneously multiple frequencies: from 250 Hz to 2000 Hz by 250 Hz steps. The target sound pressure field is visible on Fig. 3.8(a) (real part). As an indication, the corresponding truncated sound pressure field at order $M = 5$ is visible on Fig. 3.8(b) (real part). The extent of the region for which the truncation error ϵ_M as defined by Eq. (3.11) is smaller to -14 dB , is shown as a black contour line. The reproduced sound pressure field at order $M = 5$ using a Lebedev grid with $L = 50$ perfect point sources is shown in

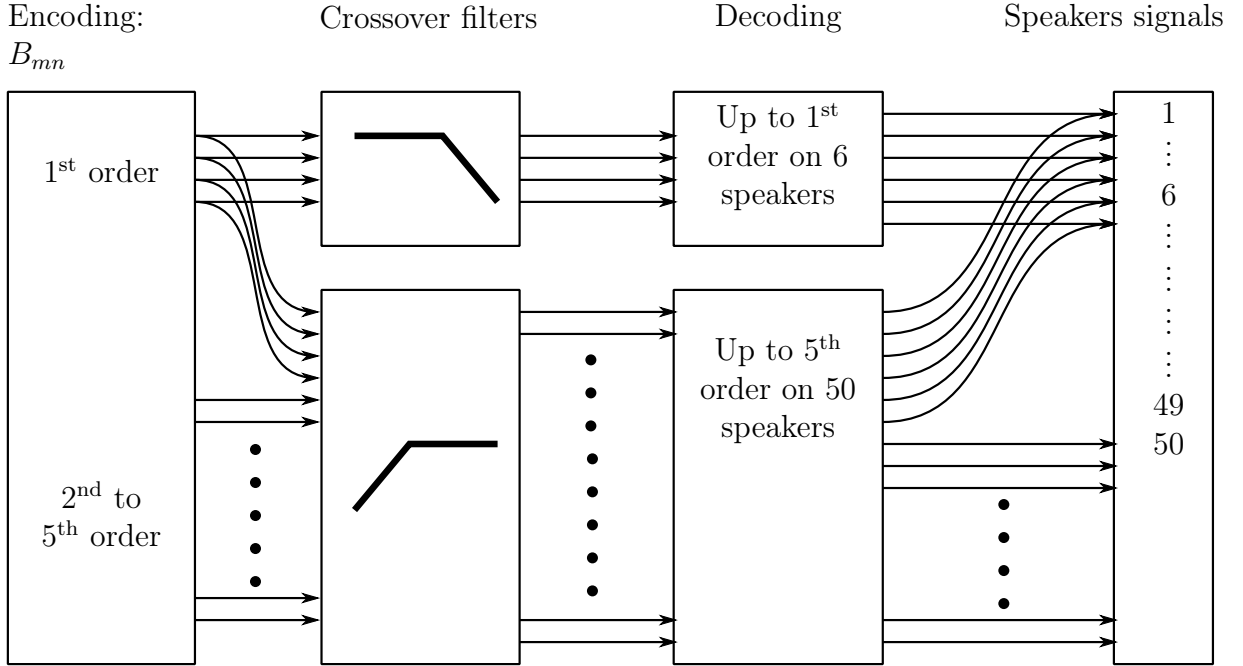


Figure 3.7 Several orders decoders on the same Lebedev grid with $L = 50$ loudspeakers. Two-band case: A decoder up to 1st order reconstructs the sound field in low frequency with 6 loudspeakers. A decoder up to 5th order reconstructs the sound field in higher frequencies with 50 loudspeakers.

Fig.3.8(c). The reproduction error is defined as follows :

$$\epsilon_{\text{recons}}(k, r, \theta, \delta) = \frac{|p(k, r, \theta, \delta) - \tilde{p}(k, r, \theta, \delta)|^2}{|p(k, r, \theta, \delta)|^2}. \quad (3.47)$$

The extent of the region for which the reproduction error ϵ_{recons} as defined by Eq. (3.47) is smaller to -14 dB, is shown as a black contour line in Fig.3.8(c). As mentioned in Sec. 3.12.1, the aliasing error for $kr \leq 5$, is negligible, thus the reconstruction zone fits well with the reproduction zone in the truncated-only case Fig.3.8(b). Furthermore, the reproduced sound field outside this zone is mainly due to aliasing errors in regards of the vanishing truncated sound field in Fig. 3.8(b) and Eq. (3.40).

To take advantages of the three nested sub-grids of the Lebedev quadrature aforementioned, one designs here a three-band decoder. The target radius for the reconstructed sound field is chosen as the smallest radius according to Eq. (3.14) and the frequency content of the sound field. Thus, this radius is given by

$$r_{\text{target}} = \frac{Mc}{2\pi f_{\text{max}}}, \quad (3.48)$$

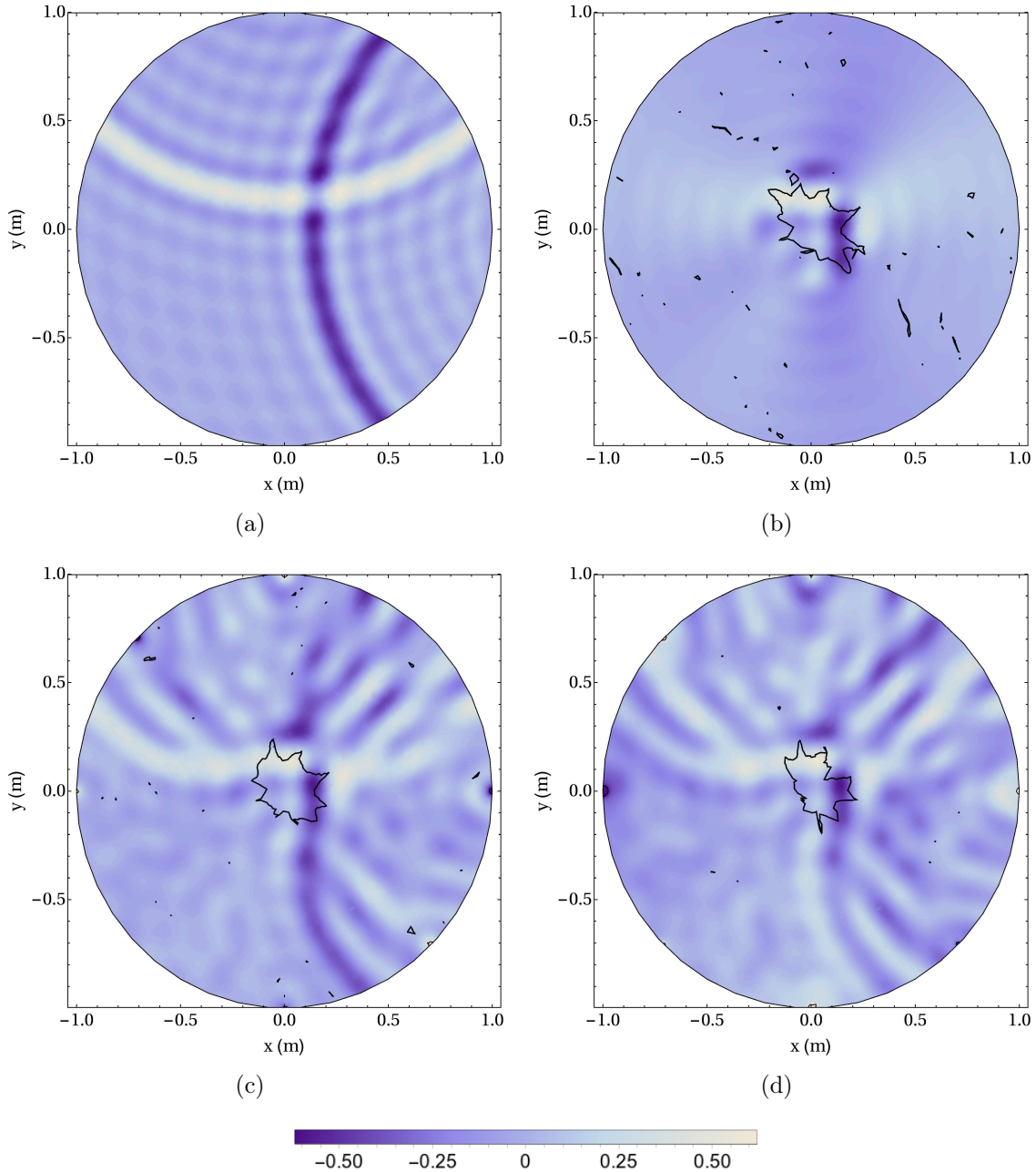


Figure 3.8 Real part of the target, truncated and reproduced sound field for a multi-frequency case: (a) target sound field. (b) truncated sound field at order $M = 5$. (c) reproduced sound field at order $M = 5$ on a fifty-nodes full-band Lebedev grid. (d) multi-band multi-order reconstructed sound field : band 0-400 Hz reproduced at order $M = 1$ on a six-nodes Lebedev grid, band 400-1200 Hz reproduced at order $M = 3$ on a twenty-six-nodes Lebedev grid and band 1200-2500 Hz reproduced at order $M = 5$ on a fifty-nodes Lebedev grid.

for $M = 5$ and $f_{\max} = 2000$ Hz, $r_{\text{target}} \simeq 0.13$ m. Once r_{target} is fixed, one then chooses the cutoff frequencies for each band such that Eq. (3.14) is respected for $r = r_{\text{target}}$, $M = 1$

and $M = 3$ which correspond to the maximum order of integration for the two nested sub-grids with $L = 6$ and $L = 26$ nodes respectively. The corresponding frequencies are $f_{M=1} = 400$ Hz and $f_{M=3} = 1200$ Hz. Finally, the signal in the band $0 - 400$ Hz is decoded with a 1st-order decoder with $L = 6$ loudspeakers, the signal in the band $400 - 1200$ Hz with a 3rd-order decoder and the signal in the band $1200 - 2000$ with a 5th-order decoder as explained in Fig. 3.7 for a two-band case. In the case of the proposed sound pressure field, the multi-band decoder implies that the 6 first loudspeakers in the Lebedev quadrature (3.35) should be full-band $0 - 2000$ Hz, the next 20 loudspeakers should have a response in the band $400 - 2000$ Hz and the remaining 24 loudspeakers a response in the band $1200 - 2000$ Hz. The resulting reconstructed sound field is visible Fig. 3.8(d) (real-part). The extent of the region for which the reproduction error ϵ_{recons} as defined by Eq. (3.47) is smaller to -14 dB, is shown as a black contour line in Fig. 3.8(d). One observes that the targeted radius for the reproduction zone is comparable to Fig. 3.8(b) and (c). However, in this case, a practical implementation would require only six full-band loudspeakers (which are more expensive). The remaining loudspeakers could have a more limited bandwidth and the results would be comparable in the target reconstruction zone to a situation with fifty full-band loudspeakers.

3.14 Conclusions

In this paper, the theory of rotationally-invariant quadratures rules was presented, and an example of a Lebedev grid using fifty nodes was investigated for Ambisonics. The importance of orthonormality of the spherical harmonics was highlighted in a context of physical encoding or reconstruction of a sound field with Ambisonics approach. During the discretization process, the rotationally-invariant quadrature rules guarantee exact orthonormality, up to an order M . The Lebedev grid with fifty nodes was then compared with two other grids in terms of spherical harmonics orthonormality error and aliasing for sound field reconstruction and capture. A further analysis of how aliasing occurs was also provided. The two other grids were a t -design grid with fifty nodes and a Fliege grid with forty-nine nodes. It appears that the Lebedev grid has the best performance at capture and reproduction. Finally, the Lebedev grid working up to fifth order with fifty nodes contains two nested sub-grids effective for first and third orders with six and twenty-six nodes respectively. This property enables several decoding configurations for physical reconstruction depending on the frequency and order, and thus the use of different transducers in different frequency bands. Simulation results were provided for the case of a three-band decoder using the three grids contained in the Lebedev grid. It was found that

a multi-frequency sound field can be reproduced accurately in the sweet-spot by using a combination of low-order decoder for low frequency and higher-order decoder for higher frequency. The cutoff frequency between each band, derived here using the classical rule of thumb $kr = M$, could be further refined. Moreover, the artifacts outside the reconstruction zone using a multi-band multi-order decoder should be subject of further investigations. Further perspectives of this work include comparison with more rotationally-invariant quadratures, in particular when working with other group of Platonic geometries such as tetrahedral or icosahedral rotation groups.

CHAPITRE 4

COMPENSATION DES RÉFLEXIONS DANS UNE SALLE SUR UNE ZONE ÉTENDUE PAR UNE APPROCHE AMBISONIQUE

4.1 Avant-propos

Auteurs et affiliations :

- Pierre Lecomte, Candidat au doctorat, Université de Sherbrooke, Conservatoire National des Arts et Métiers.
- Philippe-Aubert Gauthier, Professionnel de recherche, Université de Sherbrooke.
- Christophe Langrenne, Ingénieur de recherche, Conservatoire National des Arts et Métiers.
- Alain Berry, Professeur, Université de Sherbrooke.
- Alexandre Garcia, Professeur, Conservatoire National des Arts et Métiers

Date d'acceptation : Inconnue

État de l'acceptation Version soumise le 17 Août 2016.

Revue The Journal of Acoustical Society Of America

Référence Inconnue

Résumé français Cet article propose une méthode originale de compensation des réflexions dans une salle, basée sur une approche ambisonique. Une égalisation multi-canaux du système de restitution est proposée en gain et en phase. Les filtres de compensation des réflexions sont conçus pour agir dans le domaine des harmoniques sphériques, en amont de

l'étape de décodage. La conception de tels filtres requiert l'inversion d'une matrice qui peut être mal conditionnée aux basses fréquences et pour les ordres élevés. Une approche de transition entre chaque ordre en fonction de la fréquence est proposée pour contourner le problème et limiter la quantité de régularisation. Des résultats de simulations sont présentés dans le domaine spatial, temporel et fréquentiel, pour une large bande de fréquences. Il est alors montré que la méthode proposée peut réduire les erreurs de reproduction à -14 dB au maximum dans la zone de reconstruction définie en champ libre. Un aperçu de l'extension de la méthode proposée pour créer des effets de réverbération tridimensionnels basée sur des captures de réponses de salles ambisoniques est discuté.

4.2 Contribution au document

Les équations de décodage de l'ambisonie sont établies en champ libre, avec des sources de restitution monopolaires (voir le chapitre 2). Pour faciliter le calcul des signaux des sources de restitution, la disposition géométrique des haut-parleurs est importante, comme il est montré au chapitre 3. Cependant, pour des applications pratiques, des facteurs supplémentaires sont à prendre en compte pour la bonne reconstruction du champ sonore.

En effet, la disposition des enceintes de restitution est faite en pratique avec une marge d'erreur qui n'est pas nulle. De même, leur rayonnement n'est pas celui d'un monopole en champ libre : leur réponse fréquentielle en amplitude n'est pas plate, leur réponse en phase n'est pas linéaire. Aussi, la structure qui supporte les enceintes, ainsi que le lieu où se trouve le dispositif de restitution induisent des phénomènes acoustiques de réflexions, réfraction, diffusion, etc. Tout ces facteurs détériorent alors la reconstruction précise du champ sonore si les signaux théoriques, calculés pour un modèle en champ libre, sont utilisés.

Dans ce contexte, ce chapitre présente une compensation des différents phénomènes qui écartent le système de restitution réel du modèle théorique. À l'instar des approches d'égalisation multi-points, qui utilisent un réseau discret de microphones pour compenser les défauts du système de restitution dans le domaine physique, l'approche proposée ici, à titre d'originalité, repose sur une formulation dans le domaine des harmoniques sphériques. Ce formalisme permet de compenser le système de restitution sur une zone continue et étendue de l'espace. Comme il le sera montré au cours du chapitre, cette approche devient alors un traitement du champ sonore décrit par un vecteur de composantes ambisoniques avant le décodage. Le décodeur utilisé est un décodeur basique conçu en champ libre (voir le chapitre 2) qui est facilement mis en oeuvre [Lecomte et Gauthier, 2015]. Alors, le système

de restitution réel peut être compensé pour permettre l'utilisation des formulations théoriques et permettre une reconstruction précise des champs sonores. Finalement, l'extension de l'approche présentée dans ce chapitre permet de concevoir des effets de réverbération tridimensionnelle, dans le domaine des harmoniques sphériques.

4.3 Cancellation Of Room Reflections Over An Extended Area Using Ambisonics

4.4 Abstract

This paper investigates the compensation of room reflections based on Ambisonics. A multichannel room equalization method for Ambisonic playback systems is proposed. The equalization filters are designed to operate in the spherical harmonics domain, prior to the decoding step. The compensation filters design requires the inversion of a matrix which can be ill-conditioned at low frequencies and for higher orders. A crossover and cross-order method is proposed to circumvent this problem and to reduce the amount of necessary regularization. Simulations results are presented in frequency, space and time domains, over a wide range of frequencies. It is shown that the proposed method is efficient and can reduce the reproduction down to -14 dB in the reconstruction area defined in free field. A glimpse into the extension of the proposed method to create three-dimensional measurement-based Ambisonic reverberation is discussed.

4.5 Introduction

Sound Field Reproduction (SFR) over an extended area is nowadays a realistic technological possibility typically implemented using dense loudspeaker arrays. The leading approaches using analytical sound field analysis are Wave Field Synthesis (WFS) [Berkhout, 1988] and Higher Order Ambisonics (HOA) [Daniel, 2000] (simply called Ambisonics throughout this paper). However, the classic formulations are derived for a free-field scenario with idealistic monopole sources as loudspeakers. When physical accuracy of SFR is of concern, loudspeaker responses and room responses have to be taken into account.

The problem of room equalization over an extended area has been addressed in the literature. One strategy, given the set of loudspeakers, is to spatially sample the reproduction zone and equalize the playback system transfer function accordingly [Corteel, 2006; Kirkeby et Nelson, 1999; Kirkeby *et al.*, 1998; Tokuno *et al.*, 1997]. However, these multi-

point approaches may result in poor equalization between the control points [Fielder, 2001]. Using spectral band-limitation of the plenacoustic function [Ajdler *et al.*, 2006], the sampling can be chosen to have a solution valid between the control points [Kolundzija *et al.*, 2011]. Other recent approaches envision the problem from a different perspective by using, for example, directive loudspeakers to limit reflections [Poletti *et al.*, 2010], or by exploiting the room reflections to create the desired sound field [Betlehem et Poletti, 2014].

WFS and HOA are interesting approaches to address the room reflection compensation problem. Indeed, both methods can control the sound field over an extended continuous area and wide frequency range. For WFS, methods for room reflection compensation were proposed using plane wave decomposition [Spors *et al.*, 2003], multi-point equalization mixed with wave field synthesis [Corteel, 2006] or inverse problems [Gauthier et Berry, 2006; Stefanakis *et al.*, 2010]. For the Ambisonics approach with spherical/cylindrical harmonics, room equalization was proposed by modeling the playback system and its reflections in a room over a basis of cylindrical harmonics [Betlehem et Abhayapala, 2005]. Then, a decoder adapted to the room is derived by inverting the plant matrix. However, for practical applications and in comparison with free-field decoders [Daniel, 2000; Zotter et Frank, 2012], this approach may require a considerable amount of equalization filters.

In this context, the present paper extends the approach presented for two-dimensional situations in [Betlehem et Abhayapala, 2005] to the three-dimensional case and proposes the correction of the room response in the Ambisonic domain, prior to decoding. This is a marked difference with respect to the work of Bethlehem [Betlehem et Abhayapala, 2005] in which the room equalization filters also perform the Ambisonic decoding. A great advantage of this method is that a conventional free-field decoder can be used. Furthermore, it will be shown that, globally, less room correction filters are required.

The paper is organized as follows. Definitions and notations are introduced in Sec. 4.6. The main equations of Ambisonics are derived in Sec. 4.7. In Sec. 4.8, the reflections cancellation approach in the Ambisonic domain is presented. Image source simulations in Sec. 4.9 emphasize ill-conditioning of the matrix to be inverted at low frequencies and for high Ambisonic orders, as discussed in Sec 4.10. To solve this problem, a method with crossover cross-order combined with Truncated Singular Value Decomposition (TSVD) regularization is presented in Sec. 4.11. Frequency- and time-domain results of simulation for the proposed room correction approach are presented in Sec. 4.12. Finally, practical considerations in a measurement context are discussed in Sec. 4.13 and a possible extension

of the proposed approach to the design of three-dimensional Ambisonic reverberation effect is presented in Sec. [4.14](#).

4.6 Definitions and notations

Throughout this paper, the following spherical coordinate system is used (see Fig. [4.1](#)):

$$x = r \cos(\theta) \cos(\delta), \quad y = r \sin(\theta) \cos(\delta), \quad z = r \sin(\delta). \quad (4.1)$$

Albeit different from ISO standard [ISO8000-2, 2009](#), this convention is found to be more

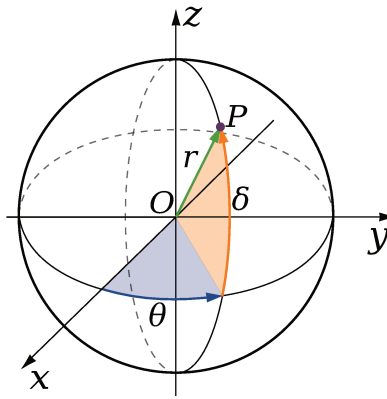


Figure 4.1 Spherical coordinate system in use. A point $P(x, y, z)$ is described by radius r , azimuth θ and elevation δ .

coherent with the literature on spatial hearing and spatial audio [\[Blauert, 1997\]](#).

4.6.1 Spherical harmonics

The real-valued spherical harmonics definition is retained in this paper:

$$Y_{mn}(\theta, \delta) = \sqrt{(2m+1)\epsilon_n \frac{(m-|n|)!}{(m+|n|)!}} P_{m|n}(\sin(\delta)) \times \begin{cases} \cos(|n|\theta) & \text{if } n \geq 0 \\ \sin(|n|\theta) & \text{if } n < 0 \end{cases}. \quad (4.2)$$

In Eq. [\(4.2\)](#), $(m, n) \in (\mathbb{N}, \mathbb{Z})$ with $|n| \leq m$, $P_{m|n}$ are the associated Legendre polynomials of order m and degree $|n|$ and $\epsilon_n = 1$ if $n = 0$, $\epsilon_n = 2$ if $|n| \geq 1$. The spherical harmonic of order m and degree n is denoted Y_{mn} . The family of spherical harmonics constitutes an orthonormal basis in \mathbb{L}^2 , the vectorial space of square-integrable functions on the unit sphere $\{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$. For each order m , there are $(2m+1)$ spherical harmonics. Thus, a basis truncated at order M contains $(M+1)^2$ functions.

4.6.2 Spherical Fourier Transform

The Spherical Fourier Transform (SFT) is the operation which computes the expansion over spherical harmonics of an angular function $f(\theta, \delta)$. It is defined as:

$$f_{mn} = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\frac{\pi}{2}}^{+\frac{\pi}{2}} f(\theta, \delta) Y_{mn}(\theta, \delta) \cos(\delta) d\delta d\theta. \quad (4.3)$$

The Inverse Spherical Fourier Transform (ISFT) describes the function $f(\theta, \delta)$ as a linear combination of spherical harmonics:

$$f(\theta, \delta) = \sum_{m=0}^{\infty} \sum_{n=-m}^m f_{mn} Y_{mn}(\theta, \delta). \quad (4.4)$$

4.6.3 Matrix and vectors notations

Throughout the paper, a matrix \mathbf{M} is denoted with capital bold letter, a vector \mathbf{v} with lower case bold letter. The transpose of matrix \mathbf{M} is denoted as \mathbf{M}^T and the Hermitian conjugate transpose \mathbf{M}^* .

4.7 Ambisonics

4.7.1 Sound field representation for the interior problem

Ambisonics starts with the expansion of a sound pressure field over spherical harmonics for the interior problem. In a sphere of radius r_s exempt of sources, the homogeneous Helmholtz equation is $\Delta p + k^2 p = 0$, with implicit time dependance $e^{i\omega t}$. In the latter equation, $\omega = 2\pi f$ with frequency f , wave number $k = \omega/c$, sound celerity c , and imaginary number $i = \sqrt{-1}$. The spherical coordinate solution of the Helmholtz equation leads to the Fourier-Bessel description of the sound pressure field:

$$p(k, r, \theta, \delta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta). \quad (4.5)$$

Therefore, a sound pressure field is fully described in a sphere centered at the origin O by Eq. (4.5). The radial and angular dependencies are separated in Eq. (4.5): The radial dependency is given by spherical Bessel functions $j_m(kr)$ and the angular dependency, or spherical wave spectrum [Williams, 1999], is described by the ISFT of the expanded

sound pressure field. The B_{mn} terms are called Ambisonic components and are frequency dependent [Daniel, 2000]. Therefore, the sound pressure field p is fully described by the Ambisonic components B_{mn} . Physically, these components correspond to the successive spatial derivatives of the sound pressure field at the origin [Daniel, 2000].

As an example, one recalls the Ambisonic components of a spherical wave of amplitude S , emitted by a point source at position $\mathbf{r}_s = (r_s, \theta_s, \delta_s)$ [Williams, 1999]:

$$S \frac{e^{-ik|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|} = -S \frac{ik}{4\pi} \sum_{m=0}^{\infty} j_m(kr) h_m^{(2)}(kr_s) \sum_{n=-m}^m Y_{mn}(\theta_s, \delta_s) Y_{mn}(\theta, \delta), \quad (4.6)$$

where $\mathbf{r} = (r, \theta, \delta)$, is the observation point and $r < r_s$. The B_{mn} components are derived by identifying Eq. (4.6) with Eq. (4.5). They are given by [Rafaely, 2015]:

$$B_{mn} = SF_m(kr_s) Y_{mn}(\theta_s, \delta_s). \quad (4.7)$$

In the above equation, $F_m(kr_s) = i^{-(m+1)} k h_m^{(2)}(kr_s) / (4\pi)$ are the near-field filters introduced by Daniel [Daniel, 2003], where $h_m^{(2)}$ are the spherical Hankel functions of the second kind and of order m . The filters F_m take into account the finite distance of the source with respect to the origin O .

4.7.2 Truncation of the spherical harmonic expansion

For practical applications, the Fourier-Bessel series of Eq. (4.5) is truncated up to order $m = M$. This truncation spatially limits the region over which the sound field is accurately described by Eq. (4.5). The average truncation error ϵ_M is defined as the deviation between $p(k, r, \theta, \delta)$ and its truncated decomposition to order M integrated over a spherical surface of radius r . For spherical waves, it is given by [Zotter *et al.*, 2009]:

$$\bar{\epsilon}_M = 1 - \frac{2k^2 r_s r}{\ln\left(\frac{r_s+r}{r_s-r}\right)} \sum_{m=0}^M (2m+1) (j_m(kr))^2 |h_m^{(2)}(kr_s)|^2. \quad (4.8)$$

Ward and Abhayapala established a rule of thumb in term of truncation error ϵ_M for plane waves in [Ward et Abhayapala, 2001]. A value of $\epsilon_M \simeq 4\%$ (-14 dB) is reached when:

$$[kr] = M, \quad (4.9)$$

where $[\cdot]$ represents the integer part of the argument. This was confirmed by Poletti [Poletti, 2005] for the three-dimensional case of spherical waves. Therefore, Eq. (4.9) presents

the accurate reproduction zone of Ambisonics when working with a finite spherical harmonic order M . Indeed, for a fixed order M , the representation zone decreases when the frequency increases. This is the lower bound of sound field reproduction with Ambisonics [Poletti, 2005].

4.7.3 Decoding

The decoding step in Ambisonics aims at deriving the driving signals of the loudspeaker array which is used to reproduce the target sound pressure field represented by its Ambisonic components. In many cases, the most straightforward array geometry is spherical. Modern formulations of the decoding step rely on a discretization of a spherical continuous distribution of secondary sources, modeled as monopoles. This approach is called the simple-source approach [Poletti, 2005; Williams, 1999; Zotter *et al.*, 2009]. Another way of designing a decoder is by a mode-matching approach, as previously reported by [Daniel, 2000]. This method matches the Ambisonic components of the reproduced sound pressure field with the discrete set of loudspeaker and the Ambisonic components of the target sound pressure field. This approach requires the inversion of a re-encoding matrix [Daniel *et al.*, 2003; Poletti, 2005; Zotter *et al.*, 2009] where the re-encoding process describes the encoded responses at the origin for each loudspeaker. For some discretization schemes of the sphere the two approaches are equivalent [Lecomte *et al.*, 2015].

As the re-encoding principle will be used for the room reflections treatment, it is further detailed in this section. The principle goes as follows: A set of L loudspeakers with driving signals $\mathbf{s} = [s_1 \cdots s_L]^T$ should recreate at the center of the array the same Ambisonic components as the target sound pressure field. The loudspeakers are modeled as monopole sources. The contribution of each loudspeaker is expanded at the center of the array using Eqs. (4.5) and (4.7). The summation of their contributions must equal the expansion of the target sound pressure field, therefore:

$$\sum_{l=1}^L s_l \sum_{m=0}^{\infty} i^m j_m(kr) F_m(kr_l) \sum_{n=-m}^m Y_{mn}(\theta_l, \delta_l) Y_{mn}(\theta, \delta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta). \quad (4.10)$$

The left-hand side of Eq. (4.10) represents the sum of the loudspeaker sound fields and the right-hand side represents the target sound field. Also, s_l is the complex gain of the l^{th} loudspeaker. As the spherical harmonics are orthonormal, each term (m, n) of this sum

can be identified separately (hence the mode-matching expression), leading to:

$$\sum_{l=1}^L s_l F_m(kr_l) Y_{mn}(\theta_l, \delta_l) = B_{mn}. \quad (4.11)$$

Working up to finite order M , the problem can be written in matrix form:

$$\mathbf{C}\mathbf{s} = \mathbf{b}, \quad (4.12)$$

where $\mathbf{C}^{(M+1)^2 \times L}$ is called the re-encoding matrix [Daniel *et al.*, 2003] which, given the complex loudspeaker inputs \mathbf{s} , provides the resulting Ambisonic components vector $\mathbf{b} = [B_{00} \cdots B_{MM}]^T$ at the center of the array. For an array where all the loudspeakers are at the same distance r_{spk} from the center, the re-encoding matrix \mathbf{C} can be expressed as the product of a diagonal matrix $\mathbf{F}^{(M+1)^2 \times (M+1)^2}$ and a rectangular matrix $\mathbf{Y}^{(M+1)^2 \times L}$:

$$\mathbf{C} = \mathbf{F}\mathbf{Y}. \quad (4.13)$$

The diagonal matrix \mathbf{F} contains the near field filters with $F_m(kr_{\text{spk}})$ on the diagonal. The matrix \mathbf{Y} contains the spherical harmonics up to order M , evaluated at each loudspeaker direction. Therefore, the l^{th} column of \mathbf{Y} , denoted \mathbf{y}_l , is $\mathbf{y}_l = [Y_{0,0}(\theta_l, \delta_l) \cdots Y_{M,M}(\theta_l, \delta_l)]^T$ and $\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_L]$. Finally, the decoding operation, given the target Ambisonic components \mathbf{b} , computes the driving signals \mathbf{s} by the inversion of \mathbf{C} :

$$\mathbf{s} = \mathbf{C}^{\#}\mathbf{b} = \mathbf{D}\mathbf{b}, \quad (4.14)$$

In Eq. (4.14), $\mathbf{D}^{L \times (M+1)^2}$ is called the decoder matrix. The inverse is symbolically denoted $\#$ because its definition may vary depending on the loudspeaker grid used [Zotter, 2009b]. Moreover, the inversion of \mathbf{C} may require regularization [Poletti, 2005]. Finally, the number of loudspeakers should exceed the number of modes to match [Poletti, 2005]: $L \geq (M + 1)^2$. In practice, since the re-encoding matrix separates the directional dependance from the radial dependance in Eq. (4.13), the decoder also separates the directional and the radial dependencies. The directional dependance is frequency independent and the radial dependance is called the near field compensation matrix and is implemented with recursive filters with a few coefficients [Daniel, 2003; Lecomte et Gauthier, 2015].

4.8 Cancellation of room reflections

4.8.1 Re-encoding in a room

So far, the spherical array of loudspeakers is modeled as L perfect monopoles in free field. However, in practical situations, the restitution system operates in a room with inherent reverberation. Moreover, the loudspeakers differ from ideal monopoles both in terms of frequency response and directivity. Also, their respective positions can suffer of misalignment compared to the theoretical loudspeaker grid. Consequently, the reverberation, the responses of the loudspeakers, and their misplacement have to be taken into account for accurate SFR.

Following the re-encoding principle presented in Sec. [4.7.3](#), one first defines the Ambisonic components of the sound pressure field of the real loudspeakers in a given room. Assuming that the l -th loudspeaker has a spherical harmonic expansion at the origin given by $\mathbf{b}_l = [B_{00,l} \cdots B_{MM,l}]^T$, one constructs the re-encoding matrix in the room, denoted $\mathbf{R}^{(M+1)^2 \times L} = [\mathbf{b}_1 \cdots \mathbf{b}_L]$. This re-encoding matrix includes the room reverberation, loudspeaker response and misalignment of the L loudspeakers. In the listening room with reflections and non-ideal loudspeaker frequency responses, the reproduced sound pressure field differs from the target sound pressure field when driving signals are provided by a free-field decoder $\mathbf{s} = \mathbf{D}\mathbf{b}$. The reproduced Ambisonic vector in the room is denoted $\hat{\mathbf{b}}^{(M+1)^2 \times 1}$, such that:

$$\mathbf{R}\mathbf{s} = \hat{\mathbf{b}} \quad (4.15)$$

The complex Ambisonic reproduction error $\mathbf{e}^{(M+1)^2 \times 1}$ is defined as the difference between the target Ambisonic component vector \mathbf{b} and the reproduced Ambisonic vector in the room $\hat{\mathbf{b}}$:

$$\mathbf{e} = \mathbf{b} - \hat{\mathbf{b}} \quad (4.16)$$

One then defines the cost function J :

$$J = \|\mathbf{b} - \hat{\mathbf{b}}\|_2^2 = \mathbf{e}^* \mathbf{e}, \quad (4.17)$$

where $\|\cdot\|_2$ indicates the vector argument 2-norm (Euclidean norm) and $*$ is the Hermitian transpose. When replacing $\hat{\mathbf{b}}$ by its expression in Eq. [\(4.15\)](#), one then looks for the signals $\mathbf{s}^{(LS)}$ which minimize J (superscript ^(LS) stands for "Least-Square"). This procedure is described by Bethlehem and Abhayapala in [\[Bethlehem et Abhayapala, 2005\]](#). The solution is then an integrated decoder specifically adapted to the set of L loudspeakers in the given room. However, as the reverberation is frequency-dependent, so will be the

decoder. For practical implementations, this results in a matrix of $L \times (M + 1)^2$ FIR (Finite Impulse Response) filters for the decoder coefficients. As mentioned in Sec. 4.7.3, for the majority of Ambisonic decoders designed in free field, the decoding operations are frequency-independent and consist of a constant matrix multiplication, which is easily implemented in real-time [Lecomte et Gauthier, 2015]. If the loudspeakers are near-field compensated, the Ambisonic signals are filtered before decoding (see Eq. (4.13)). These near-field compensation filters can be easily implemented with efficient parametric Infinite Impulse Response (IIR) filters using biquads with few coefficients [Daniel, 2003; Lecomte et Gauthier, 2015]. This highlights one of the practical limitations of the room-specific and integrated decoder as proposed by [Betlehem et Abhayapala, 2005]. In fact, it would be interesting to achieve non-integrated room compensation and decoding in the Ambisonic format so that the user can create or arrange with its favorite Ambisonic decoder. In this context, to do so, the approach presented in this paper is based on a matrix of filters in the spherical harmonics domain, in order to conserve the decoder designed in free field. The resulting correction matrix provides a filtering operation on the Ambisonic components prior to their decoding, as an extra operation which can be easily inserted in an existing Ambisonic signals flow without disrupting the chain and the possibility of using conventional Ambisonics-based tools or effect processing. The first step is the introduction of a filtering matrix $\mathbf{H}^{(M+1)^2 \times (M+1)^2}$ on \mathbf{b} , to give new Ambisonics components $\mathbf{b}_{\mathbf{H}}^{(M+1)^2 \times 1}$ which compensate for the room reflections:

$$\mathbf{b}_{\mathbf{H}} = \mathbf{H}\mathbf{b}. \quad (4.18)$$

Then, the filtered Ambisonic components $\mathbf{b}_{\mathbf{H}}$ are decoded using the free-field decoder \mathbf{D} , to give the loudspeakers input gains $\mathbf{s}_{\mathbf{H}}$:

$$\mathbf{s}_{\mathbf{H}} = \mathbf{D}\mathbf{b}_{\mathbf{H}} = \mathbf{D}\mathbf{H}\mathbf{b}. \quad (4.19)$$

Finally, the signals $\mathbf{s}_{\mathbf{H}}$ are filtered by the re-encoding matrix \mathbf{R} to give the reproduced Ambisonic vector in the room $\hat{\mathbf{b}}$:

$$\mathbf{R}\mathbf{s}_{\mathbf{H}} = \mathbf{R}\mathbf{D}\mathbf{b}_{\mathbf{H}} = \hat{\mathbf{b}}. \quad (4.20)$$

At this step, the matrix $\mathbf{T}^{(M+1)^2 \times (M+1)^2}$ is defined by:

$$\mathbf{R} = \mathbf{T}\mathbf{C}, \quad (4.21)$$

where \mathbf{C} is the "inverse" of the free-field decoder \mathbf{D} . The matrix \mathbf{T} "distorts" the free field re-encoding matrix \mathbf{C} into \mathbf{R} . That is to say, in free field with ideal and perfectly-placed monopoles as reproduction sources, \mathbf{T} is the identity matrix. When dealing with real reflective environment and non-ideal loudspeakers, \mathbf{T} describes all the transformations that the ideal free-field playback will go through in the actual room. Specifically, the term $\mathbf{T}_{i,j}$ represents the coupling between the Ambisonic component j as observed in free field to the Ambisonic component i as observed in the room, or how the j -th Ambisonic component will affect the i -th Ambisonic component in the reproduced Ambisonic vector $\tilde{\mathbf{b}}$. The decoder \mathbf{D} is designed such that $\mathbf{CD} = \mathbf{I}$ where \mathbf{I} is the identity matrix. Thus, multiplying both sides of Eq. 4.21 by \mathbf{D} one obtains

$$\mathbf{RD} = \mathbf{T}, \quad (4.22)$$

and replacing this result in Eq. 4.20, the latter can be expressed as:

$$\mathbf{T}\mathbf{b}_H = \hat{\mathbf{b}}. \quad (4.23)$$

Finally, $\hat{\mathbf{b}}$ as expressed in Eq. 4.23 is replaced in Eq. 4.17, $J = \|\mathbf{b} - \mathbf{T}\mathbf{b}_H\|_2^2$. The cost function J is then expressed as a quadratic function of compensated Ambisonic components \mathbf{b}_H [Elliott, 2001, p. 478]:

$$J = \mathbf{b}_H^* \mathbf{T}^* \mathbf{T} \mathbf{b}_H - \mathbf{b}_H^* \mathbf{T}^* \mathbf{b} - \mathbf{b}^* \mathbf{T} \mathbf{b}_H + \mathbf{b}^* \mathbf{b} \quad (4.24)$$

If the matrix $\mathbf{T}^* \mathbf{T}$ is non-singular, the cost function J possesses a global minimum given by [Elliott, 2001, p. 479]:

$$\mathbf{b}_H^{(LS)} = (\mathbf{T}^* \mathbf{T})^{-1} \mathbf{T}^* \mathbf{b} = \mathbf{H}^{(LS)} \mathbf{b}, \quad (4.25)$$

where the matrix $\mathbf{H}^{(LS)}$ corresponds to the pseudo inverse \mathbf{T}^\dagger of the matrix \mathbf{T} , such that:

$$\mathbf{H}^{(LS)} = \mathbf{T}^\dagger = (\mathbf{T}^* \mathbf{T})^{-1} \mathbf{T}^* \quad (4.26)$$

Moreover, as the matrix \mathbf{T} is square, if \mathbf{T} is non-singular, the pseudo-inverse corresponds to the inverse of \mathbf{T} [Golub et Van Loan, 2012], $\mathbf{H}^{(LS)} = \mathbf{T}^{-1}$. From now on, \mathbf{H} will always refer to $\mathbf{H}^{(LS)}$, and \mathbf{b}_H to $\mathbf{b}_H^{(LS)}$. Note that the filtering matrix \mathbf{H} operates in the spherical harmonics domain and its size is $(M+1)^2 \times (M+1)^2$, which diminishes the total number of required FIR filters in comparison with the approach presented in [Betlehem

et Abhayapala, 2005] in which a decoder adapted to the room will have $L \times (M + 1)^2$ FIR filters with $L > (M + 1)^2$.

4.9 Room simulation

In order to provide guidelines for the compensation filters and to investigate the effectiveness of the proposed method, a simulation of an Ambisonic playback system in a room is now introduced. Note that only the room reflections are simulated, i.e. the loudspeakers are modeled as perfect monopoles with exact positions in space. and flat magnitude, linear phase responses.

Loudspeaker and room configuration The Ambisonic playback system includes $L = 50$ loudspeakers arranged on a spherical Lebedev grid that corresponds to an actual 3D implementation described in [Lecomte *et al.*, 2015]. This playback system works up to order $M = 5$ with spherical Ambisonics. The spherical array has a radius $r_{\text{spk}} = 1.07$ m and is placed in a parallelepipedic room of dimensions $5 \text{ m} \times 3.60 \text{ m} \times 3 \text{ m}$, as shown on Fig. 4.2. The origin of the coordinate system is at the center of the sphere. The top-front-left corner (from the listener position placed at the origin and looking along the positive x -axis), is represented as a green dot in Fig. 4.2 and has coordinates (3.43 m, 1.4 m, 1.71 m). Note that the spherical array is off-centered with respect to x, y, z . The absorption coefficients of the walls are constant per octave band and are taken from [Vorländer, 2007]. The simulated wall materials and their random incidence absorption coefficients α are summarized in Tab. 4.1. The estimated reverberation time per octave band from Sabine’s theory [Sabine et Egan, 1994] is also provided in Tab. 4.1. For frequencies below

Wall number	Material	Octave band central frequency (Hz)					
		125	250	500	1000	2000	4000
1-4	Studio Curtains, 22 cm from wall	0.36	0.26	0.51	0.45	0.62	0.76
5	Fissured ceiling tile	0.49	0.53	0.53	0.75	0.91	0.99
6	Hairy carpet on 3 mm felt	0.11	0.14	0.37	0.43	0.27	0.25
Sabine’s reverberation time (s)		0.296	0.341	0.204	0.196	0.163	0.141

Table 4.1 Wall absorption coefficients for the room simulation and Sabine’s estimated reverberation time in octave band.

$125/\sqrt{2} \text{ Hz} \simeq 88 \text{ Hz}$, which corresponds to the region before the first octave band, α is set to the first octave band value. For frequencies above $4000\sqrt{2} \text{ Hz} \simeq 5657 \text{ Hz}$, i.e. the last octave band upper frequency, a perfect absorption $\alpha = 1$ is assumed, implying that the playback system is in free field. In those conditions the Schroeder frequency is

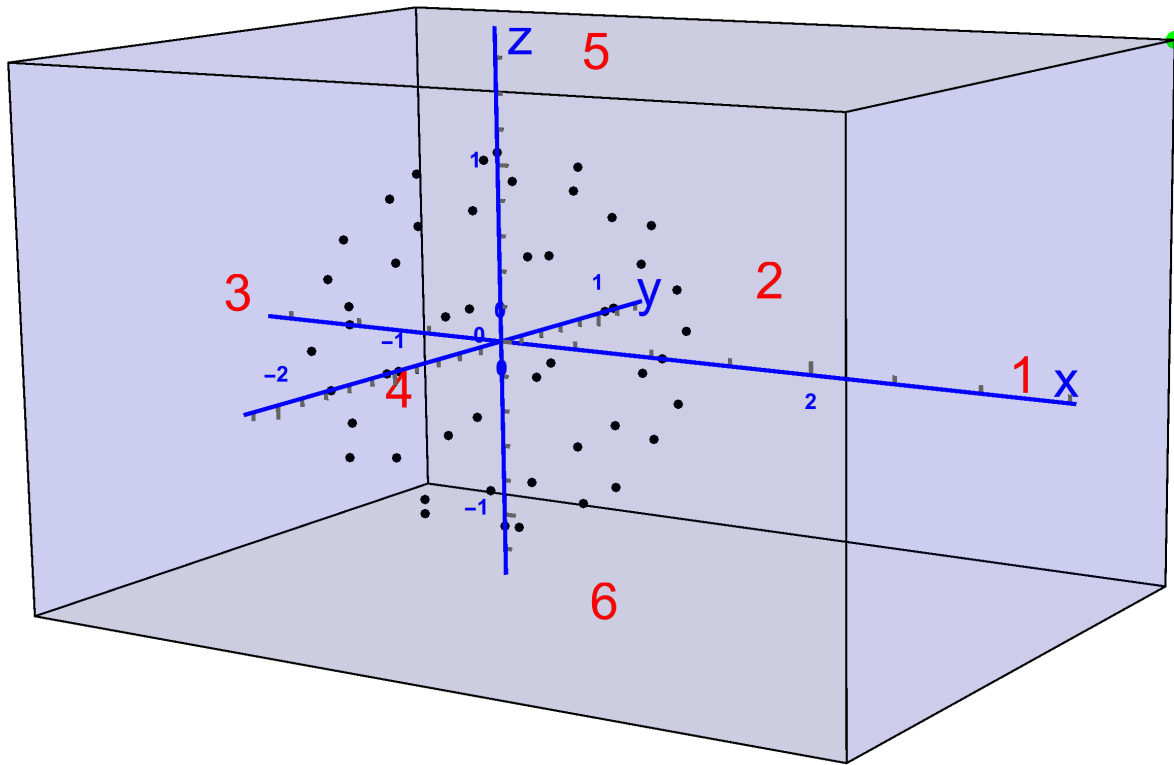


Figure 4.2 The Ambisonic playback system in a reflective parallelepipedic room: The loudspeakers positions are marked as black dots. The six walls of the room are identified by red numbers. Cartesian coordinates axes are shown in blue.

$f_{sch} = 129$ Hz and the critical distance r_c where the direct and reverberant energy densities are equal is $r_c = 0.90$ m [Kuttruff, 2009].

Image sources model An image sources model [Allen et Berkley, 1979] is used to simulate the room reflections. The implementation is herein inspired from [Lehmann et Johansson, 2008]: The simulation is achieved in the frequency domain using free-field Green functions. Note that the propagation time associated to the path from an image source to the receiver can be non-integer when working in discrete time. This results in a truncated sinc-like fractional delay FIR filter for the room impulse responses [Laakso et Välimäki, 1996; Lehmann et Johansson, 2008]. Lastly, each wall is associated with a real, negative reflection coefficient $-\sqrt{1 - \alpha}$. Such a negative reflection coefficient produces impulse responses more realistic than when using a positive reflection coefficient $\sqrt{1 - \alpha}$ as compared to actual measurements [António *et al.*, 2002; Lehmann et Johansson, 2008].

Signal processing parameters For the reported cases, simulations are achieved in the 0 to 6 kHz frequency range with a sampling frequency $F_s = 12$ kHz. In order to prevent time-domain aliasing, the filters in matrix \mathbf{H} should be long enough. In the following, the filter length is set to twice the longest predicted reverberation time found in Tab. 4.1 (0.341 s). Accordingly, 8192 points are used to sample the frequency response up to F_s , which gives a filter length of $8192/F_s \simeq 0.682$ s. For the reported cases, the first 0.17 s of the room reverberation is modeled with image sources, which represents roughly half of the total reverberation time (see Tab. 4.1). Therefore, room reflections within the maximum propagation delay of 0.17 s will be modeled and compensated. To do so, reflections are modeled up to order 11 which results in 50×2046 image sources. Note that each of the 50 loudspeakers produces its own image source network of 2046 image sources. An example of the resulting frequency response function and impulse response function from loudspeaker $l = 2$ at position (1.07 m, 0 m, 0 m) to the sound pressure at origin B_{00} is shown in Fig. 4.3.

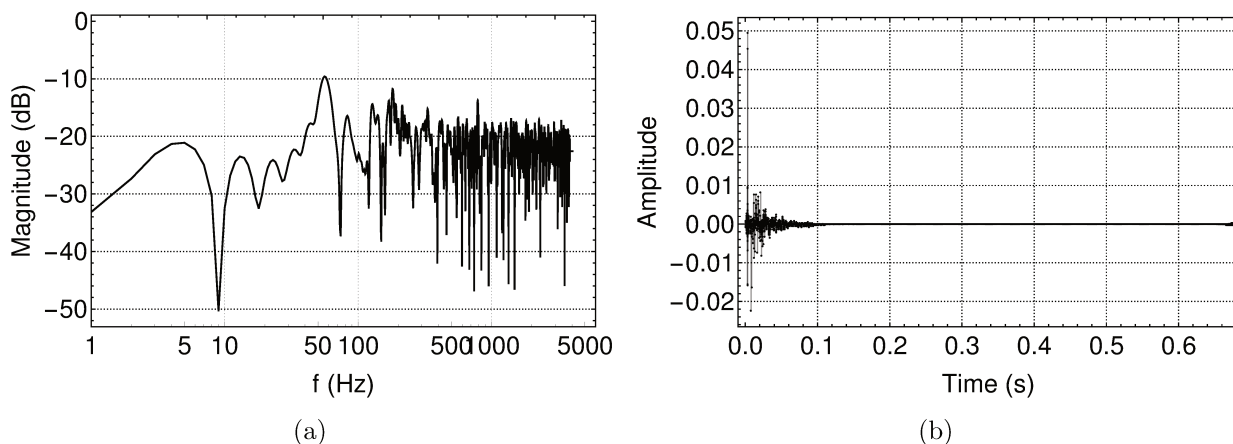


Figure 4.3 (a): Magnitude of the frequency response function (B_{00}) (dB with reference 1) of the loudspeaker 2 (1.07 m, 0, 0) in the room. (b) Corresponding impulse response.

Re-encoding matrix \mathbf{R} The impulse response function discussed previously for the l -th source only includes the sound pressure at origin O with its reflections. This corresponds to the zero-th order Ambisonic component in the room. In order to obtain all Ambisonics components up to order M , the l -th source is encoded as a monopole using Eq. 4.7 and summed with the encoding of all its associated image sources. The re-encoding matrix \mathbf{R} in the room is finally obtained by repeating this process for the L loudspeakers.

4.10 Condition-number of the room reflections cancellation problem at low frequencies

Once the re-encoding matrix \mathbf{R} is modeled, one obtains the transition matrix \mathbf{T} by multiplying \mathbf{R} with the free-field decoder \mathbf{D} as pointed by Eq. (4.22). From Eq. (4.26), the matrix \mathbf{T} is to be inverted to find the room reflections cancellation matrix \mathbf{H} . However, the inversion of matrix \mathbf{T} is problematic at low frequencies and can result in large gains for the compensation filters, beyond the limitations of practical implementations. Indeed, at low frequencies, the matrix \mathbf{T} is ill-conditioned. The condition number increases as the Ambisonic order M increases. To illustrate this, Fig. 4.4 shows the condition number of the matrix \mathbf{T} for the room simulation presented in Sec. 4.9, as function of frequency for various truncation orders from $M = 1$ to $M = 5$. One observes that the condition number increases drastically as the order increases and as the frequency decreases. The coefficients

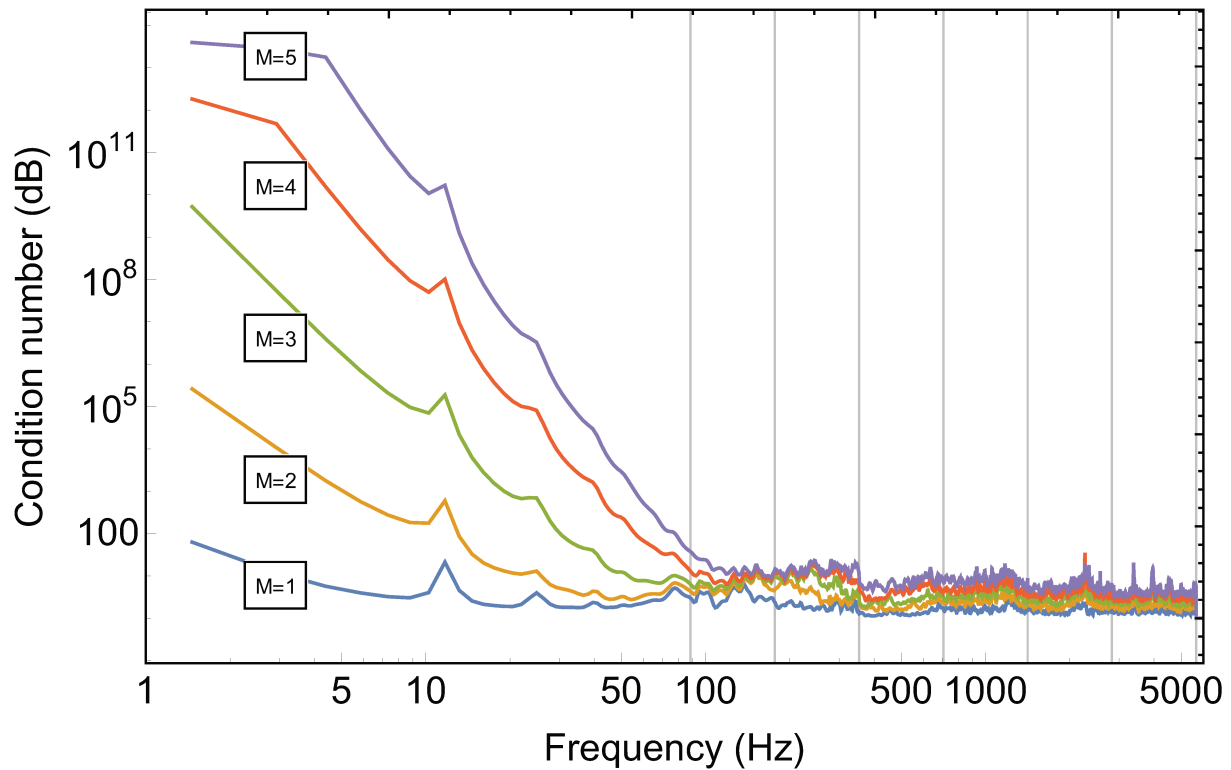


Figure 4.4 Condition number of matrix \mathbf{T} versus frequency for several truncation orders M . The vertical grid lines correspond to octave band limits used for room absorption coefficients.

of the re-encoding matrix \mathbf{R} at low frequencies take high values at high order, due to the amplification of Hankel functions in Eq. (4.7) for the image sources. In addition, the coefficients of the near-field compensated decoder matrix \mathbf{D} take low values at high order,

due to values of the inverse of Hankel functions used in near-field compensation [Daniel, 2003]. Consequently, the matrix \mathbf{T} , obtained with Eq. (4.22) has a high dynamic range at low frequency which explains the high condition number.

To assert this ill-conditioning problem, a pseudo-inversion method based on Truncated Singular Value Decomposition (TSVD) [Hansen, 1998] is used, as in [Kolundzija *et al.*, 2011] for instance. The Singular Value Decomposition (SVD) of \mathbf{T} is given by:

$$\mathbf{T} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*, \quad (4.27)$$

where $\mathbf{U}^{(M+1)^2 \times (M+1)^2}$ is the unitary matrix whose columns are the left singular vectors \mathbf{u}_i of \mathbf{T} , $\mathbf{V}^{(M+1)^2 \times (M+1)^2}$ is the unitary matrix whose columns are the right singular vectors \mathbf{v}_i of \mathbf{T} and $\mathbf{\Sigma}^{(M+1)^2 \times (M+1)^2}$ is the diagonal matrix with real, positive singular values σ_i of \mathbf{T} , ranked in decreasing order along the diagonal. The TSVD pseudo-inverse of \mathbf{T} is then given by [Hansen, 1998]:

$$\mathbf{T}^\dagger = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^*. \quad (4.28)$$

In Eq. (4.28), the diagonal matrix $\mathbf{\Sigma}^\dagger$ is constructed by setting to 0 the singular values in $\mathbf{\Sigma}$ which are below a given threshold ϵ . The remaining terms of $\mathbf{\Sigma}^\dagger$ are the inverse of the singular values in $\mathbf{\Sigma}$. In this paper, the threshold is adapted at each frequency as a fraction of the largest singular value σ_1 of \mathbf{T} :

$$\epsilon = \frac{\sigma_1}{\kappa}, \quad (4.29)$$

with $\kappa \geq 1$. Thus, at low frequencies when the span of singular value exceeds κ , the TSVD excludes the lowest singular values and the inversion is regularized. At higher frequencies when the ratio between the first and last singular value is larger than the threshold ϵ , the TSVD retains all the singular values and the inversion of \mathbf{T} is as in Eq. (4.26). In the reported cases, the threshold is set with $\kappa = 100$.

4.11 Crossover, cross-order room reflections cancellation filters

In the previous section, the inversion and regularization of \mathbf{T} with TSVD allows inversion with limited penalization on the gain of the resulting filters. However, this penalization can increase the reproduction quadratic error of Eq. (4.16) [Hansen, 1998]. In fact, as the regularization increases, the filter gain is attenuated (which is of interest for practical

implementations) but the quadratic error Eq. (4.16) increases, thus the reproduced sound field can be far from the target sound field.

Hopefully, at low frequencies, where this problem arises, only the first spherical harmonics orders are useful to describe the sound field over an extended area, as pointed in Sec. 4.7.2 and Eq. (4.9). In this frequency range, from Fig. 4.4, one notes that the condition number of \mathbf{T} is lower for low truncation orders. Therefore, if one defines a prescribed radius as a delimitation of the reproduction zone, the maximum frequency at which the reproduction error does not exceed -14 dB can be computed using equation Eq. (4.9). As Ambisonics for the interior problem is spatially bounded by the loudspeaker array, one can set the reconstruction sphere radius to the loudspeaker array radius r_{spk} and compute the maximum frequency according to Eq. (4.9), such that:

$$f_m = \frac{mc}{2\pi r_{\text{spk}}}. \quad (4.30)$$

In Eq. (4.30), the frequency f_m , for a given order m is the highest frequency at which the truncation order $M = m$ is sufficient to describe the sound field in the reproduction zone of radius r_{spk} . It is also the lowest frequency for which the truncation order $M = m + 1$ is necessary to describe the sound field in the reproduction zone. Following this concept, at low frequencies, only the inversion of a matrix \mathbf{T} defined at low truncation order is necessary to correct the reflections in the prescribed reproduction zone. The filters of matrix \mathbf{H} switch from the m order correction matrix to $m + 1$ order at cutoff frequency f_m . Consequently, according to Fig. 4.4, the inversion of the matrix \mathbf{T} at order $m < M$ requires less regularization than the inversion at truncation order M and the solution is still valid in the prescribed reproduction zone.

Following these guidelines, the matrix \mathbf{H} of compensation filters is constructed according to Fig. 4.5. The matrix \mathbf{T} is first inverted and regularized using TSVD for all truncation orders $1 \leq m \leq M$ for the entire frequency range. One obtains the filtering matrices $\mathbf{H}_1^{4 \times 4}, \dots, \mathbf{H}_M^{(M+1)^2 \times (M+1)^2}$. Then, for each cutoff frequency f_m with $1 \leq m \leq M$ in Eq. (4.30), a pair of crossover filters is designed by creating a Low-Pass (LP) and a High-Pass (HP) filters, both at cutoff frequency f_m . Then, the various filtering matrices $\mathbf{H}_1 \dots \mathbf{H}_M$ are combined using the crossover filters in cascade, hence the name crossover cross-order. The output of this process gives the filtering matrix $\mathbf{H}^{(M+1)^2 \times (M+1)^2}$. Note that the crossover filtering is implemented in the frequency domain with multiplication and addition operations. To avoid any phase distortion with the crossovers, only the square

magnitudes of the crossover filters are used. This operation is equivalent to a zero-phase forward-backward filtering in time domain [Oppenheim et Schafer, 2010, p. 320].

Finally the resulting matrix \mathbf{H} is given in the frequency domain by:

$$\mathbf{H}(f) = \mathbf{H}_1(f)|LP(f)|^2 + \mathbf{H}_2(f)|HP_{f_1}(f)|^2|LP_{f_2}(f)|^2 + \dots + \mathbf{H}_M|HP_{f_{M-1}}(f)|^2 \quad (4.31)$$

For the reported simulation cases, the target radius is $r_{spk} = 1.07$ m, the corresponding cutoff frequencies using Eq. (4.30) are $f_c(1) \simeq 51$ Hz, $f_c(2) \simeq 102$ Hz, $f_c(3) = 153$ Hz, $f_c(4) = 204$ Hz. The crossover filters are inspired by Linkwitz-Riley filters [Linkwitz, 1976], and are pairs of low-pass and high-pass Butterworth filters of order 8.

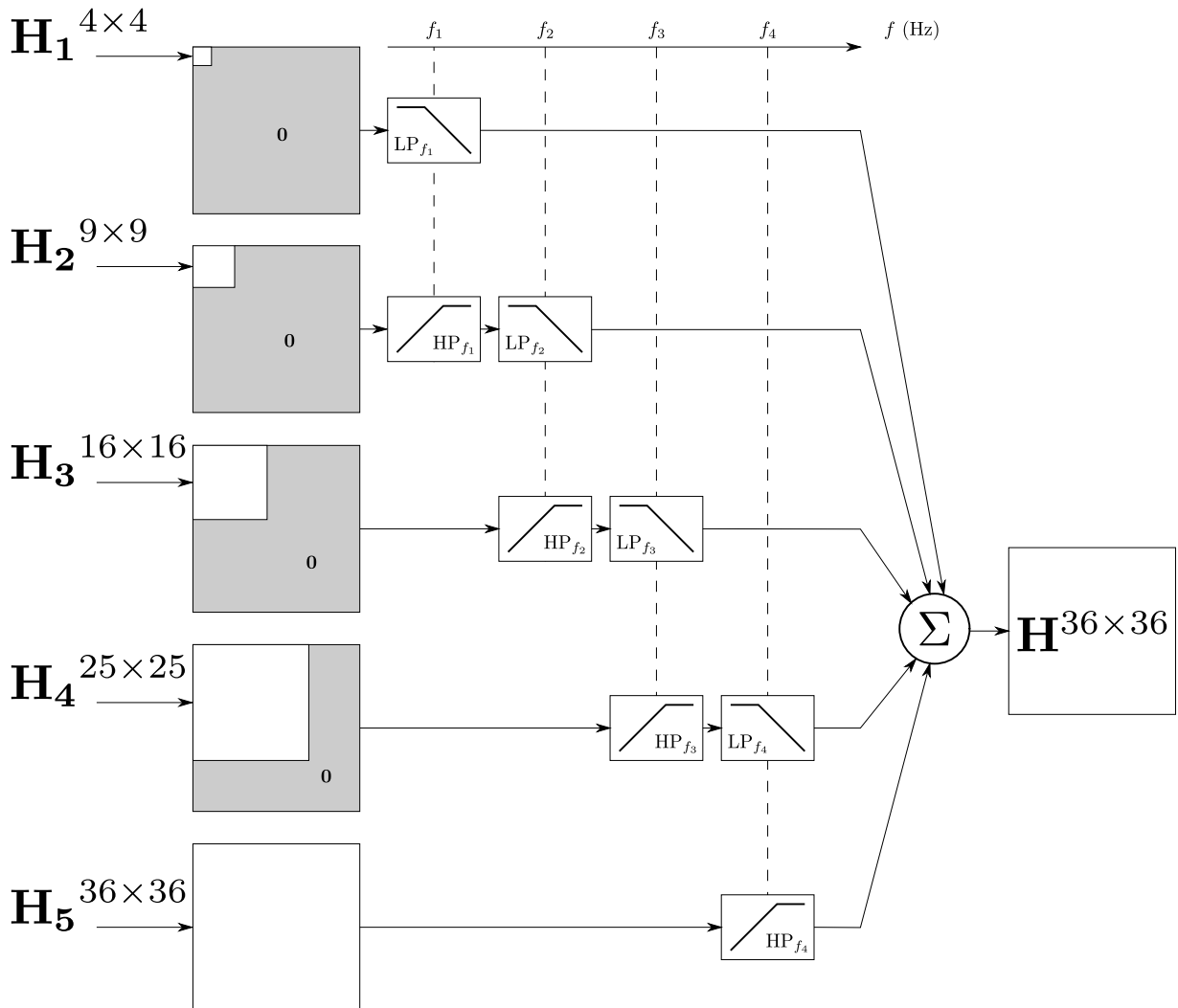


Figure 4.5 Crossover cross-order filter design for a maximum truncation order $M = 5$.

Two examples of resulting compensation filters, namely the diagonal filter $\mathbf{H}_{5,5}$ and the off-diagonal filter $\mathbf{H}_{25,18}$, are shown in Fig. 4.6. Note that a modeling delay of half the filter length is introduced to recenter the time origin of the impulse response. For the reported cases, this modeling delay is set to 4096 samples. The diagonal filter $\mathbf{H}_{5,5}$ presents a main

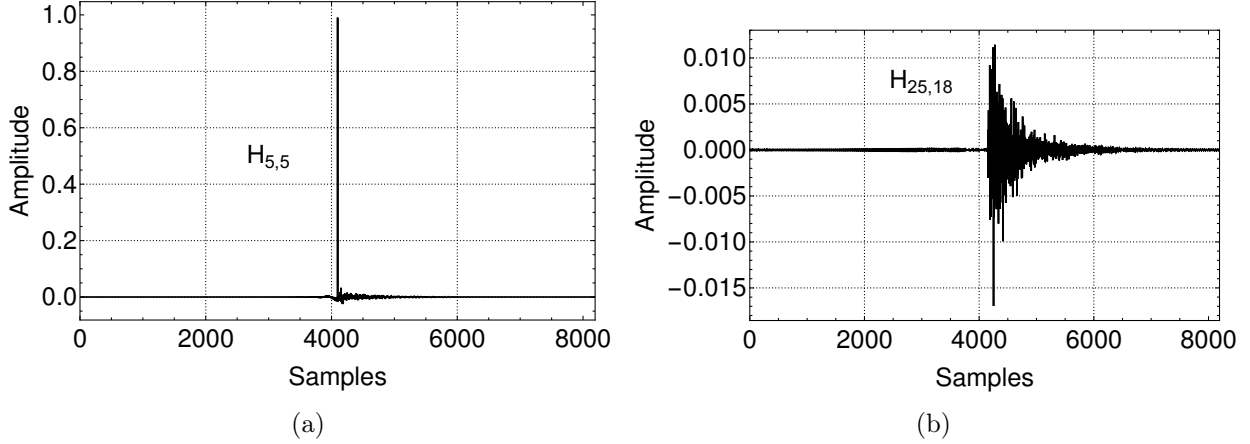


Figure 4.6 (a): Impulse response of the filter $\mathbf{H}_{5,5}$. (b) Impulse response of the filter $\mathbf{H}_{25,18}$.

peak prominent in comparison with the tail of the impulse response. This illustrates the fact that the matrix \mathbf{H} has its diagonal terms close to unity and larger than the off-diagonal terms at the majority of the frequencies considered in the reported case. For the two filters, the filter response is not perfectly zero before the time origin (4096-th sample), which is expected for the compensation of room reverberation. However, the majority of the energy in the impulse response is after the main peak. This suggests that the compensation filters of the matrix \mathbf{H} are "almost" causal. This should be a topic of further investigations.

4.12 Results of simulations

This section presents results of simulations in both frequency and time domains for the reproduction of a monopole source at location $\mathbf{r}_s = (r_s = 2 \text{ m}, \theta_s = 180^\circ, \delta_s = 0^\circ)$ with unit amplitude, $S = 1$.

4.12.1 Frequency domain

In the frequency domain, the target sound pressure field $p(k, r, \theta, \delta)$ is defined according to Eq. (4.6). We define the relative reproduction error as:

$$\epsilon(k, r, \theta, \delta) = \frac{|p(k, r, \theta, \delta) - \tilde{p}(k, r, \theta, \delta)|^2}{|p(k, r, \theta, \delta)|^2}, \quad (4.32)$$

where \tilde{p} is the reproduced sound pressure. For the reported cases, three reproduction situations are simulated. The first situation is a reproduction at order $M = 5$ with the 50 loudspeaker Lebedev grid in free field. In this situation, the Ambisonic components vector \mathbf{b} is obtained with Eq. (4.7) and the driving signals of the loudspeakers \mathbf{s} are given after the decoding step by Eq. (4.14). In the second situation, the same free field driving signals \mathbf{s} are used but the loudspeaker array is in the room, as displayed on Fig. 4.2. In this situation, reverberation occurs and degrades the reconstruction of the sound pressure field \tilde{p} . To model this, all the image sources of a given loudspeaker l (2046 image sources in the reported cases) are driven by the same signal as loudspeaker l . The resulting sound pressure field is obtained by summing the sound field of each loudspeaker and its image sources, using free field Green functions of Eq. (4.6). Finally, the third situation involves the cancellation of room reflections. In this situation the driving signals of the loudspeakers \mathbf{s}_H are derived using Eq. (4.19). Figures 4.7 and 4.8 show the real part of sound pressure field in the horizontal plane at two frequencies (125 Hz and 500 Hz) for the target sound pressure field and the three aforementioned reproduction situations. The 125 Hz case corresponds to a situation where the effective area in free field is larger than the array at order $M = 5$. Therefore, according to the crossover cross-order design of Sec. 4.11, the matrix \mathbf{H} contains only compensation filters up to the third order. Indeed, at this frequency, the correction up to third order is sufficient to correct the whole volume bounded by the loudspeaker array. For the 500 Hz case where the effective area is smaller than the array, the filters up to fifth order are used in \mathbf{H} .

The relative error $\epsilon(k, r, \theta, \delta) = 0.04$ is displayed as a black contour line. For the 500 Hz case, the truncation error at $M = 5$, $\epsilon_5(k, r) = 0.04$ is displayed as a red dashed circle of radius $r = 5/k$, according to the rule of thumb of Eq. (4.9). On Fig. 4.8(b), one observes that the effective reproduction zone defined by $\epsilon(k, r, \theta, \delta) = 0.04$ is roughly the same as the reproduction zone defined by the rule of thumb $r = 5/k$ and illustrated by the red dashed circle. In fact, as the 50 loudspeakers Lebedev grid is designed to work up to order $M = 5$ [Lecomte *et al.*, 2015], the reproduced sound field is identical to the sound field truncated to $M = 5$ in the reproduction zone. For the 125 Hz case, the theoretical reproduction zone in free field exceeds the loudspeaker array radius. Therefore, the reproduction is effective in all the volume bounded the loudspeaker array. As the theory is derived for the interior problem, some reproduction artifacts arise close the loudspeakers which bound the interior volume and where the pressure is theoretically infinite, according to Eq. (4.6). The black contour lines in Fig. 4.7(b) illustrate this.

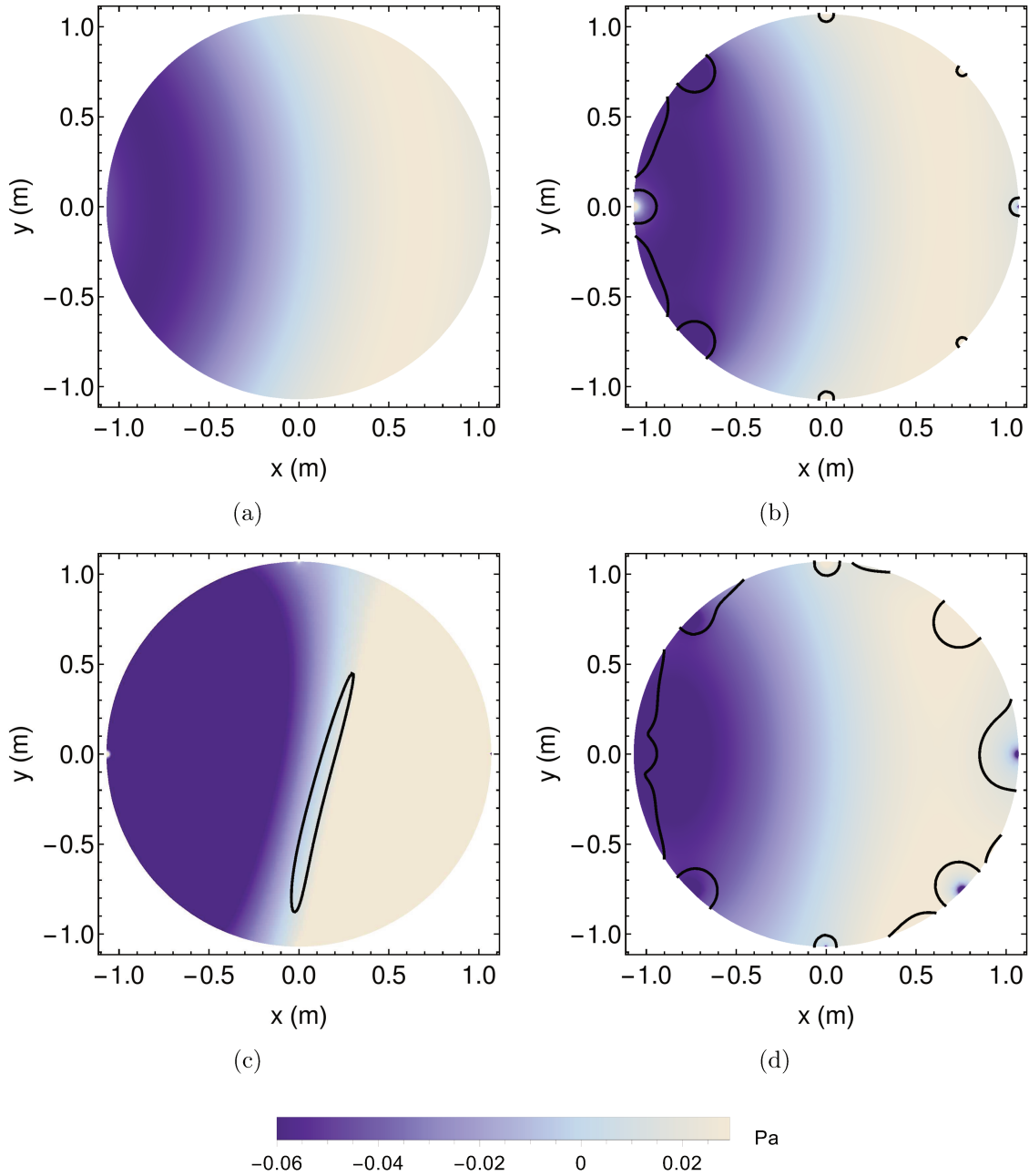


Figure 4.7 Real parts of sound pressure fields at 125 Hz in the horizontal plane, for a monopole at location $(r_s = 2 \text{ m}, \theta_s = 180^\circ, \delta_s = 0^\circ)$ with unit amplitude $S = 1$. (a): target sound field in free field, (b): reproduced sound field in free field at order $M = 5$ with 50 node Lebedev grid, (c): reproduced in the room, (d): reproduced sound field in the room with room reflections cancellation. The black contour line indicates the relative error at 4%.

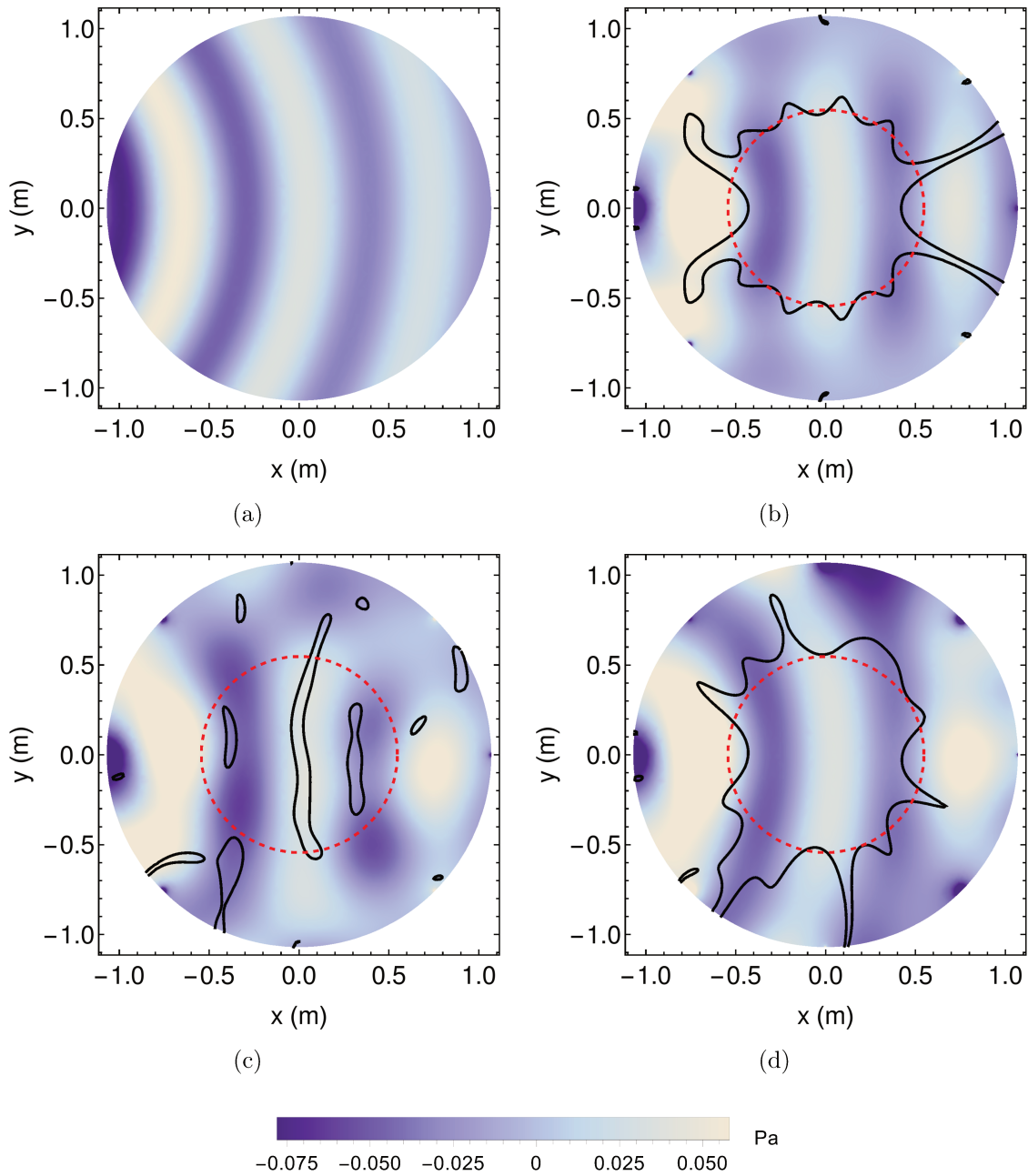


Figure 4.8 Real parts of sound pressure fields at 500 Hz in the horizontal plane, for a monopole at location $(r_s = 2 \text{ m}, \theta_s = -180^\circ, \delta_s = 0^\circ)$ with unit amplitude $S = 1$. (a) target sound field in free field, (b) reproduced sound field in free field at order $M = 5$ with 50 loudspeakers Lebedev grid, (c) reproduced sound field in the room, (d) reproduced sound field in the room with room reflections cancellation. The black contour line indicates the relative error at 4%.

The results of SFR in the room with driving signals derived from the free-field situation present a degradation of the reproduction zone at both frequencies, as shown in Figs. 4.7(c) and 4.8(c). Indeed, the reverberation is not taken into account in the playback system and the successive reflections interfere in the reconstruction of the sound field. Finally, Figs. 4.7(d) and 4.8(d) show the sound pressure field with cancellation of room reflections (in which case the driving signals are obtained according to Eq. (4.19)). In this case the reproduction zone extends to the same radius as in the free field (Figs. 4.7(b) and 4.8(b)).

4.12.2 Time domain

For the time domain simulations, the target sound pressure field is chosen as emitted by a monopole at position \mathbf{r}_s with a half-band driving signal in the frequency domain. That is to say, the monopole discrete time signal is derived from the Inverse Discrete Fourier Transform (IDFT) of a monopole at position \mathbf{r}_s with the following complex gain in the frequency domain:

$$\begin{cases} S = \frac{e^{-ik|\mathbf{r}_s - \mathbf{r}|}}{4\pi|\mathbf{r}_s - \mathbf{r}|} & 0 \text{ Hz} \leq f \leq 3 \text{ kHz} \\ S = 0 & 3 \text{ kHz} < f \leq 6 \text{ kHz} \end{cases} . \quad (4.33)$$

To reproduce the sound field in the time domain using Ambisonics, the driving signals of the loudspeakers are computed in the frequency domain at each frequency bin. Equation (4.14) is used for situations in free field and with room reflections. For room reflections cancellation situation, Eq. (4.19) is used. Finally, an IDFT gives the discrete time domain loudspeakers signals $\mathbf{s}[n]$. The reconstruction of a continuous time signal is done using the ideal reconstruction filter given by [Oppenheim et Schafer, 2010, p. 163]. For example, the continuous-time driving signal of the l -th loudspeaker is given by:

$$s_l(t) = \sum_{n=1}^{8192} s_l[n] \frac{\sin(\pi F_s(t - n/F_s))}{\pi F_s(t - n/F_s)}; \quad (4.34)$$

where F_s is the sampling frequency ($F_s = 12$ kHz in this paper). Finally, the sound pressure at position $\mathbf{r} = (r, \theta, \delta)$ and time t is given by the summation of all active sources using time domain free field Green functions:

$$p(r, \theta, \delta, t) = \sum_{i=1}^{L_{\text{tot}}} \frac{s_i(t - |\mathbf{r}_i - \mathbf{r}|/c)}{4\pi|\mathbf{r}_i - \mathbf{r}|}. \quad (4.35)$$

In Eq. (4.35) the total number of active source L_{tot} is either $L_{\text{tot}} = 1$ for the target sound pressure field, $L_{\text{tot}} = L = 50$ for the free field simulation or $L_{\text{tot}} = 50 \times 2046$ for the computation of the simulated SFR in the room.

Two sound pressure fields at time $t \simeq 0.347$ s are shown in Fig. 4.9: the target sound pressure field in Fig. 4.9(a) and the reconstructed sound pressure field with the 50 loudspeakers Lebedev grid in free field in Fig. 4.9(b). At this time, the wave front of the target sound pressure field arrives at the origin. Note that a modeling delay of 4096 samples was added during the computation of the signals. Thus, $t \simeq 0.347$ s corresponds to the propagation time from position \mathbf{r}_s to the origin, plus the modeling delay. In Fig. 4.9(b), the loudspeaker driving signals were computed in the frequency domain with Eq. (4.14). In

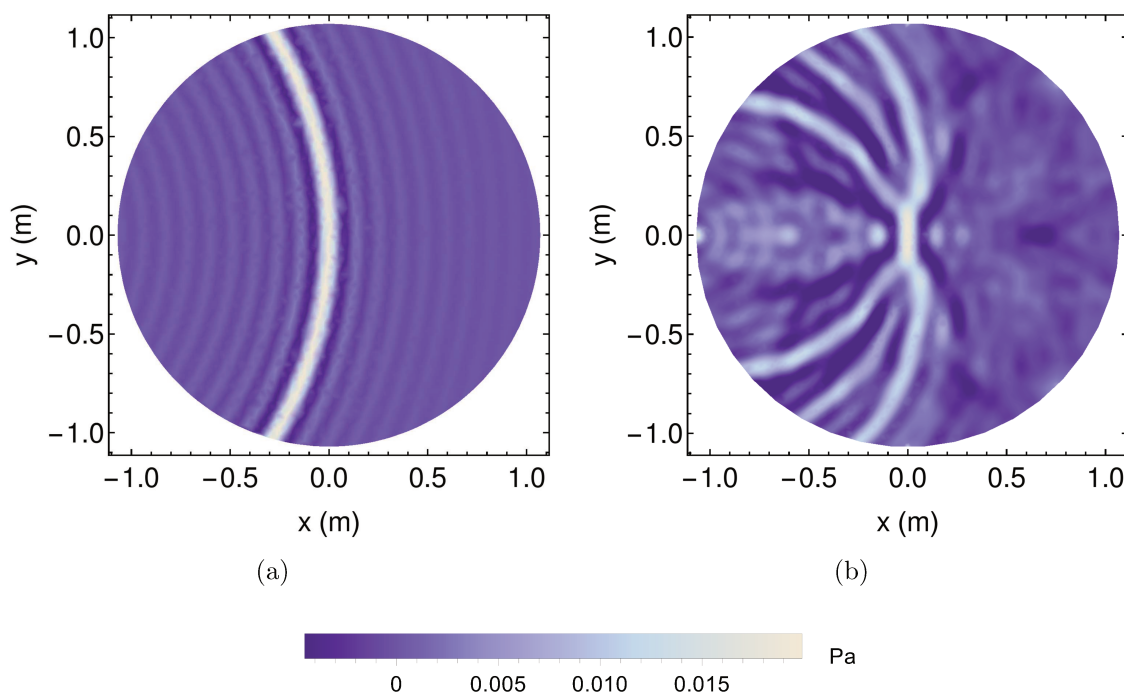


Figure 4.9 Sound pressure fields in the horizontal plane at $t \simeq 0.347$ s. (a) Target sound pressure field. (b) Fifth-order Ambisonics reconstructed sound pressure field with 50 loudspeakers Lebedev grid in free field.

Fig. 4.9(b), the wave fronts of the 50 loudspeakers of the array combine to reconstruct the target wave front in an extended area centered at the origin. In fact, as the reconstruction zone is frequency-dependent according to Eq. (4.9), the highest frequency present in the target sound field limits the size of the zone over which the reconstruction error is smaller than -14 dB.

At $t \simeq 0.347$ s, none of the room reflections have reached the observation horizontal plane of Fig. 4.9. Thus, in order to observe the effectiveness of the room reflections cancellation in the time domain, a later time is reported for which early reflections reach the origin. Figure 4.10 shows the magnitude of several sound pressure fields at $t \simeq 0.351$ s. The target sound pressure field is shown in Fig. 4.10(a). The reconstructed sound pressure field in free field with the Lebedev array is shown in Fig. 4.10(b). At this time, the first (target) wave front has left the observation horizontal plane. Therefore, in free field, no reflection is visible in the observation area of Figs. 4.10(a) and 4.10(b). However, in Figs. 4.10(c) and 4.10(d), the loudspeaker array is simulated in the room, and if the room reflections are not taken into account in the playback system, the reproduced sound field is distorted by the room as shown in Fig. 4.10(c). When the room reflections cancellation approach as proposed in this paper is used to compute loudspeaker signals with Eq. (4.19), the reproduced sound pressure field is shown in Fig. 4.10(d). For this result, one notes that room reflections are canceled in an area centered at the origin. As the room reflections are frequency-dependent, as shown in Sec. 4.12.1, the zone over which the reproduction error is smaller than -14 dB is limited by the highest frequency content of the target signal (as defined Eq. (4.9)). This is confirmed by observation of Fig. 4.10(d) where the red dashed circle is of radius $r = 5/k$ at frequency of 3000 Hz.

In fact, as the distance from the origin increases, the effect of the room reflections cancellation decreases at higher frequencies. To illustrate this observation, Fig. 4.11 shows the time-domain signals for the target and reproduced sound fields at two observation points, $\mathbf{r} = (0, 0, 0)$ and $\mathbf{r} = (0.2 \text{ m}, 45^\circ, 45^\circ)$. Figure 4.12 shows the corresponding frequency content. For the observation point at the origin, one observes that Ambisonics in free field accurately reproduces the sound pressure in time and frequency as shown in Figs. 4.11(c) and 4.12(c). However, as soon as the Ambisonic playback system is located in the reflective room (Figs. 4.11(e) and 4.12(e)), the successive reflections after the direct wave front disturb the accurate SFR. When room reflections cancellation is taken into account (Fig. 4.11(g) and 4.12(g)), the sound pressure is reproduced as in the free field situation. For an off-centered observation point, the situation is different. Since free-field Ambisonics have a frequency-dependent accurate reproduction zone, the reproduced sound pressure is not accurate for the frequency content above the criterion of Eq. (4.9) (Fig. 4.12(d)). This translates in the time domain by several wave fronts generated by the different loudspeakers which merge not synchronously to reconstruct the target sound pressure field. This is hardly visible in Fig. 4.11(d). The situation is worse when the room reflections occur, as shown in Figs. 4.11(f) and 4.12(f). Finally, when the room reflections cancellation is active (see Fig. 4.11(h) and Fig. 4.12(h)) the reflections are attenuated but not entirely

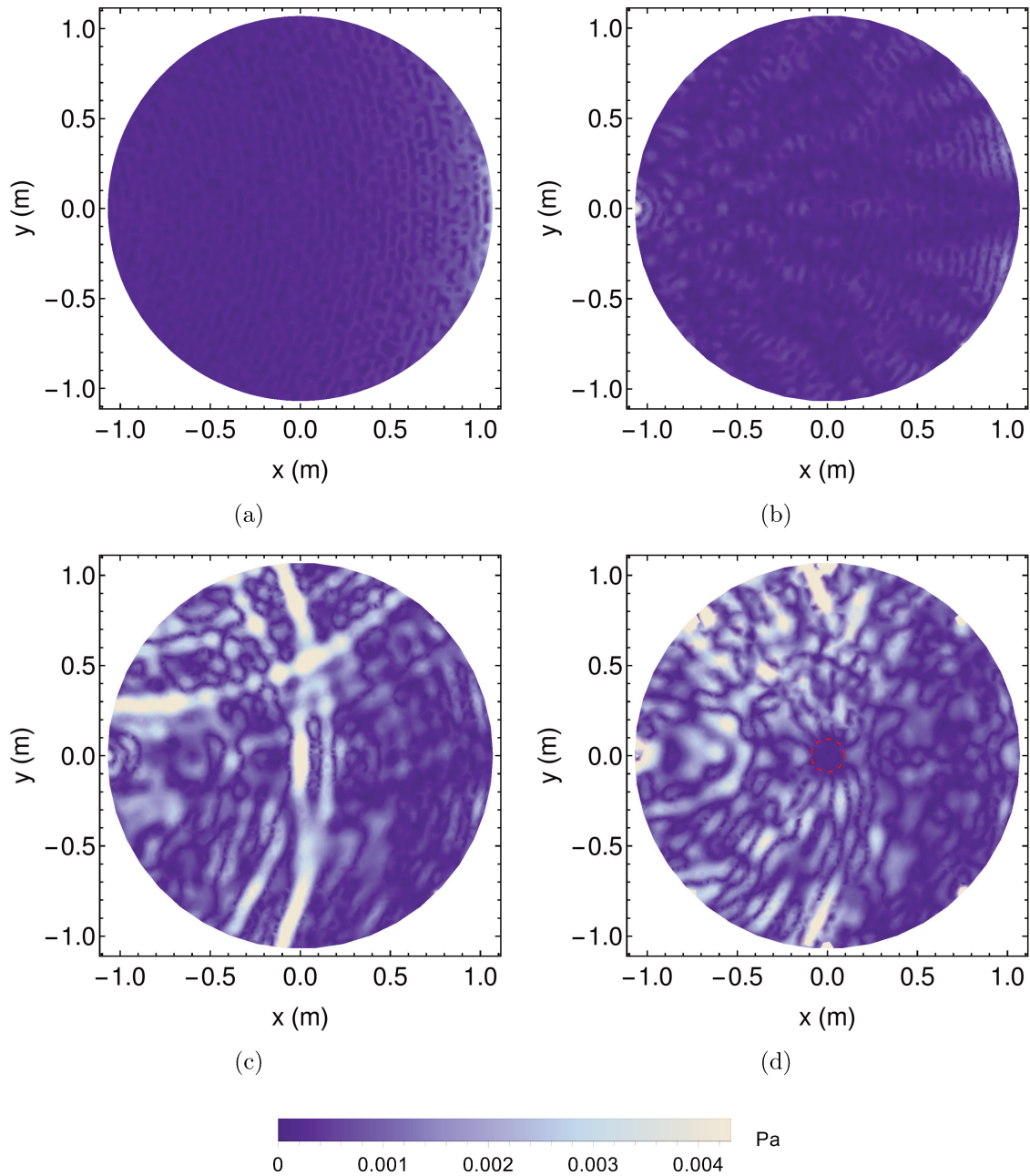


Figure 4.10 Magnitude of sound pressure fields in the horizontal plane at $t \simeq 0.351$ s. (a) target sound pressure field. (b) 5-th order Ambisonic reconstructed sound pressure field with 50 loudspeakers Lebedev grid in free field. (c) same as (b) but the loudspeaker array is in the room. (d) same as (c) with room reflections cancellation. In red dashed is shown the circle of radius $r = 5/k$ at frequency 3000 Hz.

canceled. In fact, the high-frequency content of reflections that does not respect criterion of (4.9) is not entirely accounted for in the cancellation process. For the considered point $\mathbf{r} = (0.2 \text{ m}, 45^\circ, 45^\circ)$, according to Eq. (4.9) the frequency content of the reflections above 1362 Hz is not removed from the reflections as shown in Fig. 4.12(h).

4.13 Practical considerations

This section discusses and extends the proposed method to practical implementations. The derivation of matrix \mathbf{T} implies estimating the re-encoding matrix in the room \mathbf{R} . This matrix contains Ambisonic responses at the center of the loudspeaker array, when each of the reproduction loudspeakers is active in the room. In a measurement context, the retrieval of these Ambisonic responses can be achieved with a spherical microphone array at the origin [Moreau *et al.*, 2006; Rafaely, 2015]. Indeed, if the acoustic pressure is known on a spherical surface, it is possible to compute the discrete SFT of the sound pressure field, p_{mn} , according to Eq. (4.3). Then, by applying radial equalization filters which depend of the microphone array radius a and configuration (single layer open sphere, rigid sphere, double layer open sphere), it is possible to retrieve the Ambisonic components as a function of frequency. For instance, if one uses a rigid spherical microphone of radius a , the Ambisonic components estimation are given by:

$$B_{mn} = i^{-m+1}(ka)^2 h_m^{(2)'}(ka) p_{mn}, \quad (4.36)$$

where $h_m^{(2)'}(ka)$ is the first derivative with respect to variable ka of the Hankel functions $h_m^{(2)}$, and p_{mn} is the SFT of the pressure sound field measured on a spherical rigid surface. For practical implementations, the SFT operation is estimated with discrete positions on the spherical surface, which introduces an aliasing frequency approximated by $f_{\text{alias}} \simeq ac/(2\pi M)$ [Rafaely *et al.*, 2007], depending on the discretization scheme used [Lecomte *et al.*, 2016a]. The radial equalization filters, which are the terms multiplying p_{mn} in Eq. (4.36), have an excessive gain at low values of ka and for higher orders [Moreau *et al.*, 2006]. Physically, at low frequencies the sound field is dominated by the zeroth-order Ambisonic component, thus, retrieving the higher Ambisonic orders implies an excessive amplification of the corresponding SFT coefficients p_{mn} . Accordingly the re-encoding matrix \mathbf{R} at low frequencies takes high values at high orders, as in the simulation context. Therefore the matrix \mathbf{T} for the low frequency range is ill-conditioned (see Sec. 4.10). A crossover cross-order approach is therefore recommended to limit the amount of regularization when inverting \mathbf{T} . Moreover, in measurement situations, the spherical microphone

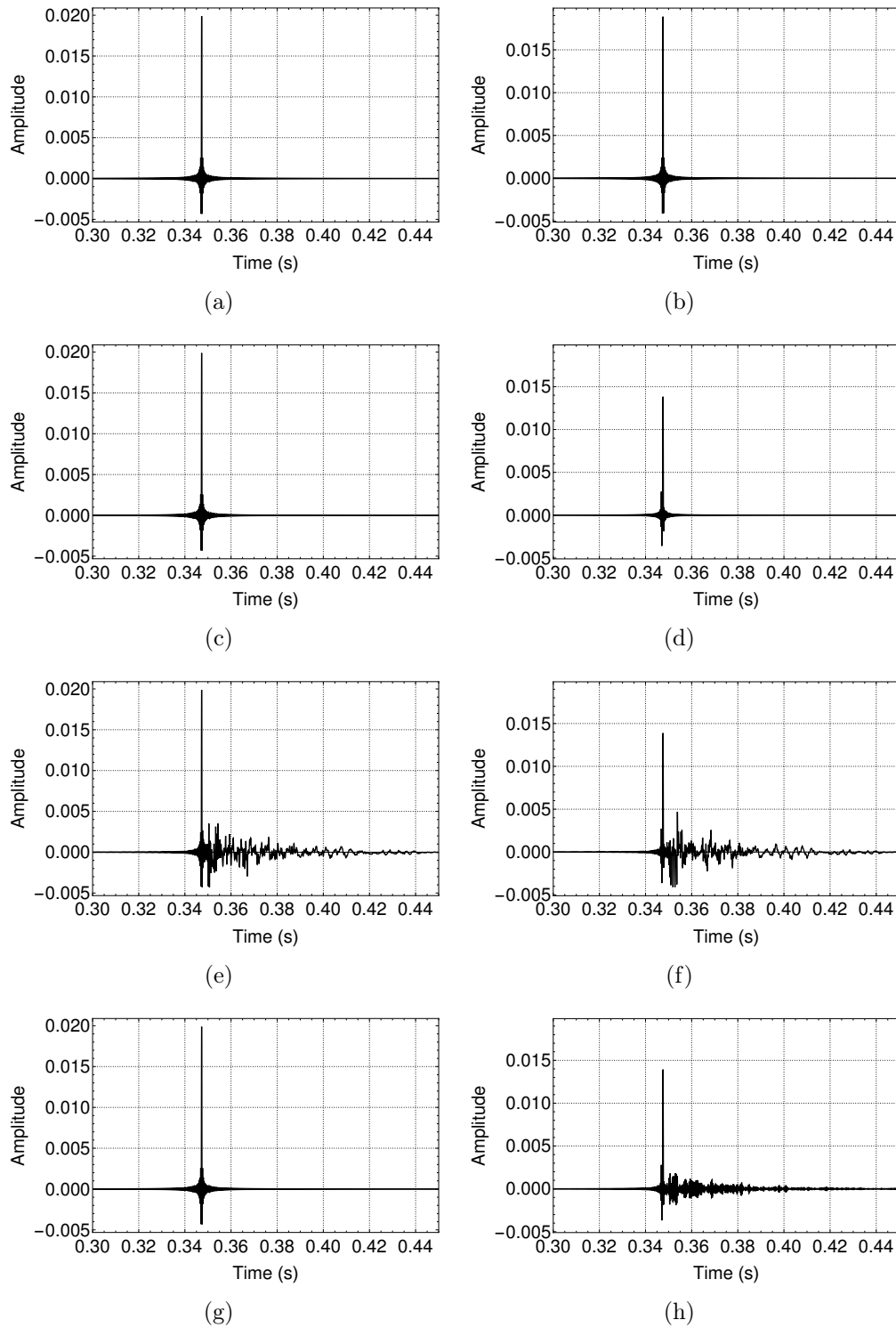


Figure 4.11 Sound pressure at two points as function of time. (a), (c), (e), (g) Sound pressure at $\mathbf{r} = (0 \text{ m}, 0^\circ, 0^\circ)$. (b), (d), (f), (h) Sound pressure at $\mathbf{r} = (0.2 \text{ m}, 45^\circ, 45^\circ)$. (a) (b) Target. (c) (d) Ambisonic reconstruction in free field. (e) (f) Ambisonic reconstruction in the room. (g) (h) Ambisonic reconstruction in the room with room reflections cancellation.

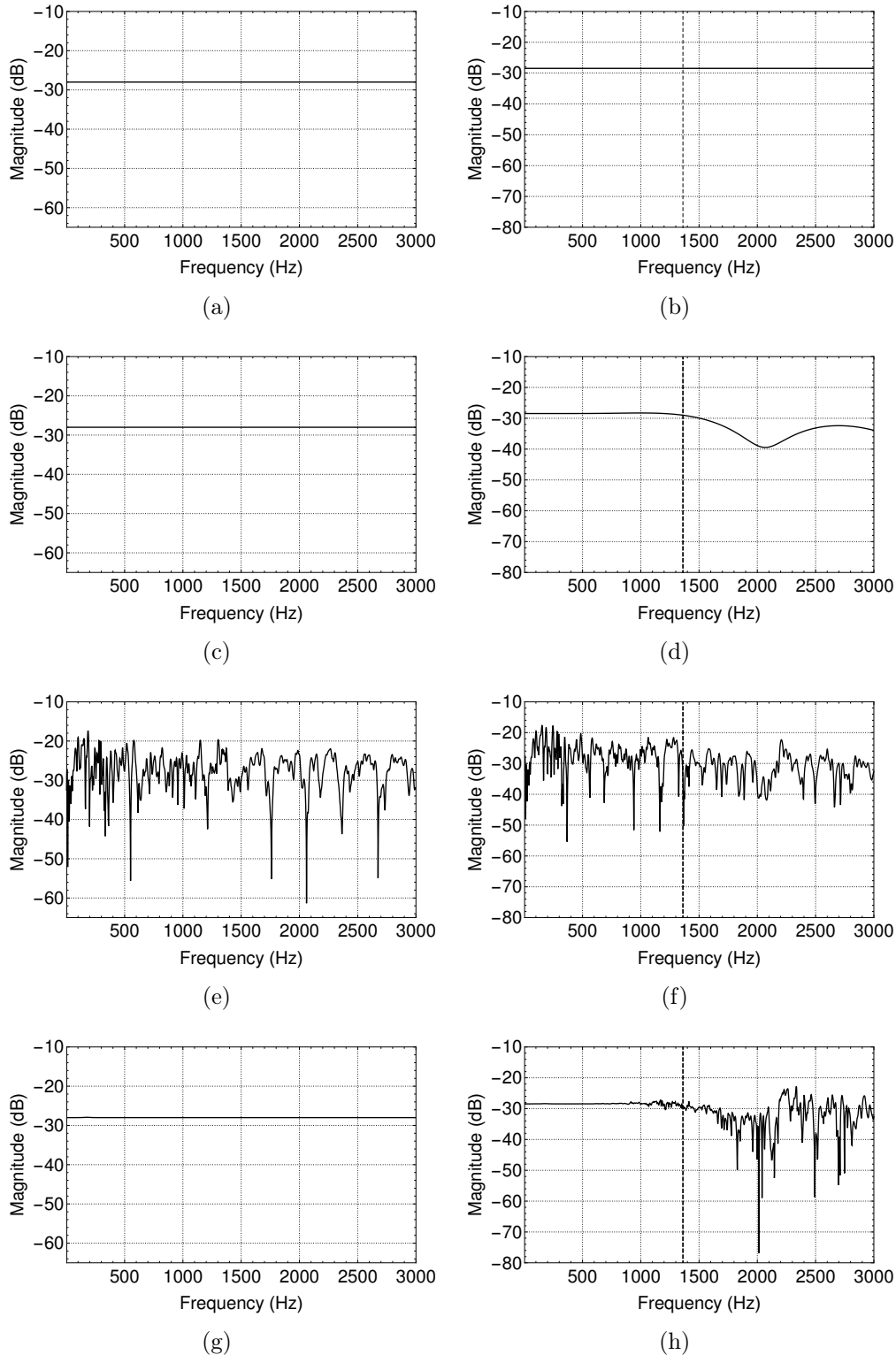


Figure 4.12 Sound pressure at two points as function of time (Fig. 4.11): frequency content. (a), (c), (e), (g) Magnitude of the pressure versus frequency at point $\mathbf{r} = (0 \text{ m}, 0^\circ, 0^\circ)$. (b), (d), (f), (h) Magnitude of the pressure versus frequency $\mathbf{r} = (0.2 \text{ m}, 45^\circ, 45^\circ)$. The frequency according to criterion of Eq. (4.9) is marked as a straight dashed line. (a) (b) Target. (c) (d) Ambisonic reconstruction in free field. (e) (f) Ambisonic reconstruction in the room. (g) (h) Ambisonic reconstruction in the room with room reflections cancellation.

presents an aliasing frequency f_{alias} above which the estimation of Ambisonic components is not longer accurate. Consequently, the room reflections cancellation is not reliable above this aliasing frequency. In this case, an additional upper-band crossover filter pair at f_{alias} should be included in Fig. 4.5 in order to progressively ensure a transition from the room compensation matrix $\mathbf{H}_{\mathbf{M}}^{(M+1)^2 \times (M+1)^2}$ to the identity matrix $\mathbf{I}^{(M+1)^2 \times (M+1)^2}$. When $\mathbf{H} = \mathbf{I}$ the room reflections cancellation is therefore bypassed. Some precautions should be taken to adequately merge the two filtering matrices in order to avoid abrupt changes in the phase and gain of final filters of \mathbf{H} by, for instance, using an approach proposed by Bernschütz, 2013. Finally, the room reflections cancellation within an applied experimental context implies that the spherical microphone array used to retrieve the Ambisonic components is perfect in its working frequency range. This assumption is not realistic and in fact the estimated re-encoding matrix \mathbf{R} is biased. The study of robustness of the compensation filter is a topic of future research.

4.14 Three-dimensional reverberation effect

In this section, other applications of the proposed method are introduced. A possible extension of the room reflections cancellation presented in this paper is the design of three-dimensional reverberation effects. Indeed, matrix \mathbf{T} includes the information of the room reflections which alter the idealistic free-field playback system into the system in the actual room. Thus, it is possible to insert this matrix \mathbf{T} in an Ambisonic signal flow, as a three-dimensional reverberation effect that would correspond to the negative footprint of the room in which \mathbf{T} was measured. As an example, if the playback system is measured in a room A the corresponding transition matrix is $\mathbf{T}_{\mathbf{A}}$ and the room reflection cancellation matrix is $\mathbf{H}_{\mathbf{A}}$. If the playback system is moved to a room B with the corresponding matrices $\mathbf{T}_{\mathbf{B}}$ and $\mathbf{H}_{\mathbf{B}}$, it is possible to inherit the reverberation of room A in room B by inserting the matrix $\mathbf{T}_{\mathbf{A}}\mathbf{H}_{\mathbf{B}}$ before the free-field decoder in the Ambisonic signal flow. In this case, the loudspeakers signals are given by:

$$\mathbf{s} = \mathbf{D}\mathbf{T}_{\mathbf{A}}\mathbf{H}_{\mathbf{B}}\mathbf{b} \quad (4.37)$$

Note that the matrix $\mathbf{T}_{\mathbf{A}}$ has an excessive gain at low frequencies when using the theoretical equalization filters in Eq. (4.36), as outlined in Sec. 4.13. Thus, as the matrix $\mathbf{T}_{\mathbf{A}}$ is not inverted in the presented example it should be regularized to allow for a practical implementation with FIR filters. This can be performed using a regularized version of

the re-encoding matrix \mathbf{R}_A following the approach presented in [Moreau *et al.*, 2006](#) for instance.

4.15 Conclusions

In this paper, an approach for room reflections cancellation within an Ambisonics paradigm was presented. The formulation relies on a model of the playback system responses in a reverberant environment in the spherical harmonic domain. The matrix of room compensation filters is also designed in the spherical harmonic domain and is inserted before the decoding step. This approach reduces the number of FIR filters needed for the room reflections cancellation as compared to compensation filters implemented at the decoding step. Moreover, the room compensation filters design requires a matrix inversion which is ill-conditioned at low frequencies. Thus, relying on the fact that only the first Ambisonic orders are necessary to accurately describe the sound pressure field at low frequencies, a crossover cross-order approach was proposed in order to reduce the amount of regularization. Based on results of simulation, it was shown that the compensation is effective over an area centered at the origin of the playback system. The size of the control area is frequency dependent, like the reconstruction zone in free-field Ambisonics. Simulation results in the frequency and time domains show the effectiveness of the proposed approach : the reconstruction zone over which the reproduction error is smaller than -14 dB is identical to the one in free field when the proposed method is applied. Future works includes the investigation of the compensation filters (causality and robustness) and experimental validation of the proposed method using spherical and planar microphone arrays in an actual listening room.

CHAPITRE 5

FILTRAGE DIRECTIONNEL DANS UNE SCÈNE SONORE AMBISONIQUE

5.1 Introduction

Ce chapitre [\[1\]](#) présente une technique de filtrage directionnel d'un champ sonore encodée dans un formalisme ambisonique. L'approche est entièrement faite dans le domaine des harmoniques sphériques, ce qui permet de s'affranchir du dispositif de capture et de restitution du champ sonore, comme discuté au chapitre [\[2\]](#). Un filtrage directionnel permet alors de favoriser certaines directions du champ sonore original, suivant une fonction de directivité choisie. Il est alors possible pour le décodage de ne reconstruire que certaines directions du champ sonore original. L'intérêt est alors de pouvoir effectuer des captations ambisoniques et de ne les restituer que partiellement pour favoriser certaines sources dans certaines directions. L'approche proposée peut présenter un certains nombres d'analogies avec la formation de voies (*beamforming*) dans le domaine des harmoniques sphériques, qui permettent d'extraire un signal monophonique dans une scène sonore tridimensionnelle [\[Meyer et Elko, 2002; Rafaely, 2015\]](#). Cependant, l'opération de filtrage directionnel présenté ici permet d'obtenir un nouveau champ sonore encodé dans le formalisme ambisonique et non pas un signal monophonique. Ceci est la contribution originale présentée à ce chapitre. Le champ sonore est alors filtré dans le domaine angulaire avant d'être décomposé dans le domaine des harmoniques sphériques à l'aide d'une TFS. Cette approche se résume à une opération matricielle lorsque le champ sonore original et la fonction de directivité sont tronqués. Ainsi cette opération de filtrage est une manipulation du champ sonore dans le domaine des harmoniques sphériques, au même titre que les transformations simples présentée au chapitre [\[2\]](#).

Le chapitre s'articule de la manière suivante : une présentation de la formation de voies sur réseau de microphones sphérique est faite à la section [\[5.2\]](#). Puis, l'opération de filtrage de la scène sonore est décrite à la section [\[5.3\]](#). Des résultats de simulations, présentés à l'appui section [\[5.4\]](#) viennent conforter la théorie mise en avant. Finalement, une mise en oeuvre en

¹Le contenu de ce chapitre a fait l'objet d'un article de conférence au congrès français d'acoustique 2016 [\[Lecomte et al., 2016c\]](#).

temps-réel du filtrage spatial est présentée à la section 5.5. La validation expérimentale de l'opération de filtrage directionnel est faite au chapitre 6.

5.2 Formation de voies avec un réseau de microphones sphérique

Les réseaux sphériques de microphones sont typiquement utilisés pour extraire un signal monophonique résultant d'une écoute dans une ou plusieurs directions privilégiées. En combinant les signaux des différents microphones il est possible de simuler une prise de son directionnelle suivant une fonction de directivité déterminée (*beam pattern*). Cependant, la géométrie sphérique permet de travailler directement dans le domaine des harmoniques sphériques après encodage du champ, suivant l'équation (2.28). Cette approche est notamment proposée dans [Meyer et Elko, 2002]. La combinaison des signaux se fait alors sur les composantes ambisoniques. De manière générale, le signal monophonique y , extrait du réseau de microphones sphérique est donné par [Rafaely, 2005] :

$$y = \sum_{m=0}^{\infty} \sum_{n=-m}^m w_{mn} B_{mn} \quad (5.1)$$

où w_{mn} représente les poids associés à la directivité désirée et B_{mn} , les composantes ambisoniques du champ de pression enregistré. Plusieurs approches pour la conception des fonctions de directivité (c'est à dire la détermination des poids w_{mn}) sont proposées dans la littérature parmi lesquels le *delay-and-sum*, par approche Dolph-Chebyshev [Koretz et Rafaely, 2009], ou par approches régulières [Li et Duraiswami, 2007; Meyer et Elko, 2002]. Une bonne revue des différentes techniques est faite dans [Cohen *et al.*, 2009, Chapitre 11] ou [Rafaely, 2015]. À titre d'exemple, la section 5.3.3 présente le cas d'une directivité prescrite simple, soit la fonction hyper-cardioïde.

5.3 Filtrage angulaire d'une scène sonore ambisonique

Dans cette section, la scène sonore captée et décrite par les signaux ambisoniques B_{mn} est filtrée à l'aide d'une fonction angulaire $g(\theta, \delta)$ (une directivité choisie) pour obtenir, non pas un signal monophonique, mais une scène sonore tridimensionnelle complète pour laquelle certaines directions sont atténuées. En d'autres termes, à la différence d'une approche par formation de voies où le signal de sortie est monophonique, (l'information spatiale autre que dans la direction de pointage (θ_0, δ_0) est perdue), l'approche décrite ci-après permet

de rester dans le domaine ambisonique après combinaison, et d'obtenir une scène sonore spatialement filtrée qui pourrait être par la suite reproduite. Cet aspect représente la contribution originale de ce chapitre. Afin d'alléger les notations, la notation ACN , telle que présentée à la section 2.2.4 est utilisée dans cette section et les harmoniques sphériques sont identifiés à l'aide d'un seul indice.

5.3.1 Fonction angulaire explicite

L'objectif est de filtrer la scène sonore $\mathbf{b} = [b_0 \cdots b_j, \cdots b_{(M+1)^2-1}]^T$ par une fonction angulaire $g(\theta, \delta)$. L'opération de filtrage s'effectue dans le domaine angulaire :

$$\tilde{b}(\theta, \delta) = g(\theta, \delta)b(\theta, \delta) \quad (5.2)$$

où \tilde{b} est la scène sonore obtenue dans le domaine angulaire après filtrage. b correspond à la scène sonore initiale dans le domaine angulaire, obtenue à l'aide d'une TFS inverse :

$$b(\theta, \delta) = \sum_{j=0}^{\infty} b_j Y_j(\theta, \delta) \quad (5.3)$$

Chaque composante \tilde{b}_i de la scène sonore filtrée est obtenue en appliquant une TFS sur l'équation (5.3) :

$$\tilde{b}_i = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} g(\theta, \delta) \left(\sum_{j=0}^{\infty} b_j Y_j(\theta, \delta) \right) Y_i(\theta, \delta) \cos(\delta) d\theta d\delta \quad (5.4)$$

soit :

$$\tilde{b}_i = \frac{1}{4\pi} \sum_{j=0}^{\infty} b_j \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} g(\theta, \delta) Y_j(\theta, \delta) Y_i(\theta, \delta) \cos(\delta) d\theta d\delta \quad (5.5)$$

Lorsque la scène sonore \mathbf{b} est tronquée jusqu'à l'ordre M , et qu'un vecteur ambisonique $\tilde{\mathbf{b}}$ filtré jusqu'à l'ordre \tilde{M} est désiré, la dernière équation peut s'écrire sous la forme d'un produit d'une matrice et d'un vecteur :

$$\tilde{\mathbf{b}} = \mathbf{M}\mathbf{b} \quad (5.6)$$

où $\mathbf{M}^{(\tilde{M}+1)^2 \times (M+1)^2}$ est la matrice de filtrage de terme courant

$$\mathbf{M}_{ij} = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} g(\theta, \delta) Y_j(\theta, \delta) Y_i(\theta, \delta) \cos(\delta) d\theta d\delta \quad (5.7)$$

Il est important de noter le fait que l'équation (5.6) représente un filtrage mais aussi un changement d'ordre de par la taille potentiellement rectangulaire de \mathbf{M} . Le calcul de \mathbf{M}_{ij} peut poser des difficultés. Il suppose notamment de connaître de manière analytique la fonction angulaire g . Néanmoins, il est possible de calculer l'intégrale en utilisant une règle de quadrature pour la sphère, qui respecte l'orthonormalité des harmoniques sphériques [Lecomte *et al.*, 2015]. Cette approche est proposée par Kronlachner et Zotter [Kronlachner, 2014] à l'aide de maillages de type t -design [Hardin et Sloane, 1996]. Il devient alors possible de définir une fonction angulaire g arbitraire. L'intégration utilisant une quadrature donnera un résultat exact jusqu'à un ordre M' pour laquelle elle fonctionne, mais si g possède dans sa TFS des coefficients non-nuls pour les ordres plus grands que M' , ces derniers peuvent se replier sur les ordres inférieurs selon le comportement de la règle de quadrature. La matrice \mathbf{M} étant de dimension finie, le champ sonore filtré $\tilde{\mathbf{b}}$ est tronqué à l'ordre \tilde{M} . Le contenu du spectre de Fourier sphérique de g peut nécessiter un ordre de redécomposition \tilde{M} assez important pour obtenir le résultat escompté.

5.3.2 Dirac angulaire

Un cas particulier de l'équation (5.7), $g(\theta, \delta) = \delta(\theta - \theta_0, \delta - \delta_0)$, est un Dirac angulaire dans la direction (θ_0, δ_0) , et le terme \mathbf{M}_{ij} de l'équation (5.7) se simplifie alors comme suit :

$$\mathbf{M}_{ij} = \frac{1}{4\pi} Y_i(\theta_0, \delta_0) Y_j(\theta_0, \delta_0) \quad (5.8)$$

Dans ce cas précis, le filtrage angulaire est le plus sélectif possible. La scène sonore obtenue après filtrage $\tilde{\mathbf{b}}$ correspond à la décomposition en harmoniques sphériques tronquée à l'ordre \tilde{M} d'une onde plane d'amplitude unitaire pondérée par la valeur en (θ_0, δ_0) de la fonction angulaire b . En effet, les composantes \tilde{b}_i de $\tilde{\mathbf{b}}$ sont données par :

$$\tilde{b}_i = \frac{1}{4\pi} Y_i(\theta_0, \delta_0) \sum_{j=0}^{(M+1)^2-1} b_j Y_j(\theta_0, \delta_0) \quad (5.9)$$

soit

$$\tilde{b}_i = \frac{1}{4\pi} Y_i(\theta_0, \delta_0) b(\theta_0, \delta_0) \quad (5.10)$$

5.3.3 Approche utilisant la décomposition de la fonction de directivité g

En repartant de l'équation (5.4), la fonction g peut également être décrite par sa transformée de Fourier sphérique inverse :

$$g(\theta, \delta) = \sum_{k=0}^{\infty} g_k Y_k(\theta, \delta) \quad (5.11)$$

Si la fonction g est décrite jusqu'à l'ordre M' , le terme \mathbf{M}_{ij} de l'équation (5.6) devient :

$$\mathbf{M}_{ij} = \sum_{k=0}^{(M'+1)^2-1} g_k C_{ijk} \quad (5.12)$$

où C_{ijk} représente l'intégrale du produit de trois harmoniques sphériques :

$$C_{ijk} = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} Y_i(\theta, \delta) Y_j(\theta, \delta) Y_k(\theta, \delta) \cos(\delta) d\theta d\delta \quad (5.13)$$

Cette intégrale est indépendante des fonctions b et g et est reliée aux coefficients de Clebsch-Gordan qui peuvent être pré-calculés pour une mise en oeuvre efficace [Arfken et Weber, 2005, pp. 803-805].

L'ordre de redécomposition \tilde{M} nécessaire pour ne pas omettre de composantes ambisoniques dans le résultat filtré est tel que $\tilde{M} = M + M'$. En effet, en remplaçant g dans l'équation (5.4) par son expression à équation (5.11), il est plus évident que la fonction à projeter sur les harmoniques sphériques est d'un ordre maximal $M + M'$:

$$\tilde{b}_i = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} \left(\sum_{k=0}^{(M'+1)^2-1} g_k Y_k(\theta, \delta) \right) \left(\sum_{j=0}^{(M+1)^2-1} b_j Y_j(\theta, \delta) \right) Y_i(\theta, \delta) \cos(\delta) d\theta d\delta \quad (5.14)$$

Ou autrement dit, la scène sonore filtrée \tilde{b} peut s'exprimer à l'aide de la TFS inverse comme un polynôme d'ordre maximal $M + M'$ sur les harmoniques sphériques.

Avec cette approche, la littérature du *beamforming* pour réseaux sphériques de microphones devient un atout pour calculer la matrice \mathbf{M} de l'équation (5.6). En effet, les coefficients g_k de l'équation (5.11) sont fournis dans la littérature pour des fonctions di-

rectionnelles optimisées, par exemple par Meyer et Elko [Meyer et Elko, 2002] pour une directivité hyper-cardioïde, présentée dans la section suivante.

Fonction de directivité régulière de type hyper-cardioïde

La directivité régulière de type hyper-cardioïde avec un angle de pointage donné par (θ_0, δ_0) est décrite suivant l'équation (5.11) avec les poids g_{mn} donnés par :

$$g_{mn} = c_m \sqrt{\frac{1}{2m+1}} Y_{mn}(\theta_0, \delta_0) \quad (5.15)$$

Les coefficients c_m sont récapitulés par le Tableau (5.1), selon [Meyer et Elko, 2002], à ceci près qu'une normalisation par la valeur maximale du lobe principal est effectuée ici. L'obtention de ces coefficients est décrite par Teutsch [Teutsch, 2007, Chapitre 5].

Tableau 5.1 Coefficients c_m pour un directivité de type hyper-cardioïde.

Ordre	c_0	c_1	c_2	c_3
1	0,249993	0,433017	0	0
2	0,11112	0,19245	0,248448	0
3	0,0625128	0,108241	0,139751	0,165365

La figure 5.1 montre une fonction de directivité résultant des coefficients choisis pour le cas considéré.

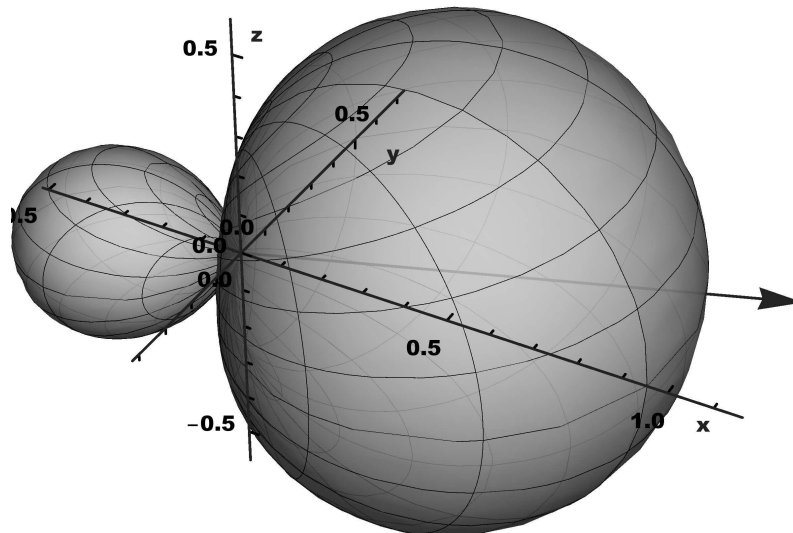


Figure 5.1 Directivité obtenue pour le cas hyper-cardioïde d'ordre 1, pointant vers la direction $(20^\circ, 0^\circ)$ indiquée par la flèche.

5.4 Simulations

Dans cette section, des résultats de simulation sont présentés pour un filtrage angulaire par deux fonctions de directivité : d'une part la fonction de directivité hyper-cardioïde d'ordre 1, telle que présentée à la Section [5.3.3](#), et, d'autre part, un Dirac angulaire tel que présenté à la section [5.3.2](#).

La scène sonore à filtrer, représentée par le vecteur \mathbf{b} est composée de 2 ondes planes monochromatiques à 500 Hz, d'amplitudes $S_1 = S_2 = 1$ et de directions de propagation respectives de $(\theta_1 = 20^\circ, \delta_1 = 0^\circ)$ et $(\theta_2 = 110^\circ, \delta_2 = 0^\circ)$. La scène sonore est décrite jusqu'à l'ordre $M = 4$. Le champ de pression acoustique tel que représenté par la série tronquée à l'ordre $M = 4$ est montré par la figure [5.2\(c\)](#). Le diagramme de directivité de la scène sonore est montré par la figure [5.3](#) en rouge pointillée d'après l'équation [\(5.3\)](#).

5.4.1 Fonction hyper-cardioïde d'ordre $M' = 1$

La matrice \mathbf{M} est construite pour une fonction de directivité hyper-cardioïde d'ordre 1, telle qu'indiquée par l'équation [\(5.12\)](#). L'hyper-cardioïde en question est montré par la figure [5.1](#) pour angle de pointage de $(20^\circ, 0^\circ)$. Elle est également visible en noir pointillé à la figure [5.3](#). L'ordre de redécomposition choisi pour le filtrage est de $\tilde{M} = 5$. Ceci dans le but d'éviter d'omettre de l'énergie dans les composantes ambisoniques après filtrage, tel que mentionné à la section [5.3.1](#). Ainsi, la matrice \mathbf{M} de l'équation [\(5.7\)](#) est de dimensions $(5 + 1)^2 \times (4 + 1)^2$. Le résultat après filtrage est montré par la figure [5.2\(d\)](#). Le diagramme de directivité correspondant est montré par la figure [5.3](#) en bleu.

Il est possible d'observer à la figure [5.2\(d\)](#) que le résultat retrouvé ressemble à celui d'une seule onde plane provenant de la direction $(20^\circ, 0^\circ)$, montré par la figure [5.2\(b\)](#). Cependant, d'après les diagrammes de directivité figure [5.3](#), l'onde plane provenant de la direction $(110^\circ, 0^\circ)$ est encore légèrement présente dans la scène sonore filtrée. Ceci est dû au fait que la fonction de directivité utilisée ici n'est pas très directive. Même si la fonction de directivité g choisie était plus directive, elle ne permettrait pas d'isoler totalement l'onde plane provenant de $(20^\circ, 0^\circ)$, car des lobes secondaires de l'onde plane $(110^\circ, 0^\circ)$ sont également présent vers $(20^\circ, 0^\circ)$. Ceci est une conséquence du fait que la scène sonore est tronquée à un ordre trop faible pour limiter les lobes secondaires associés à une onde plane d'une direction de propagation donnée (ici, $M = 4$).

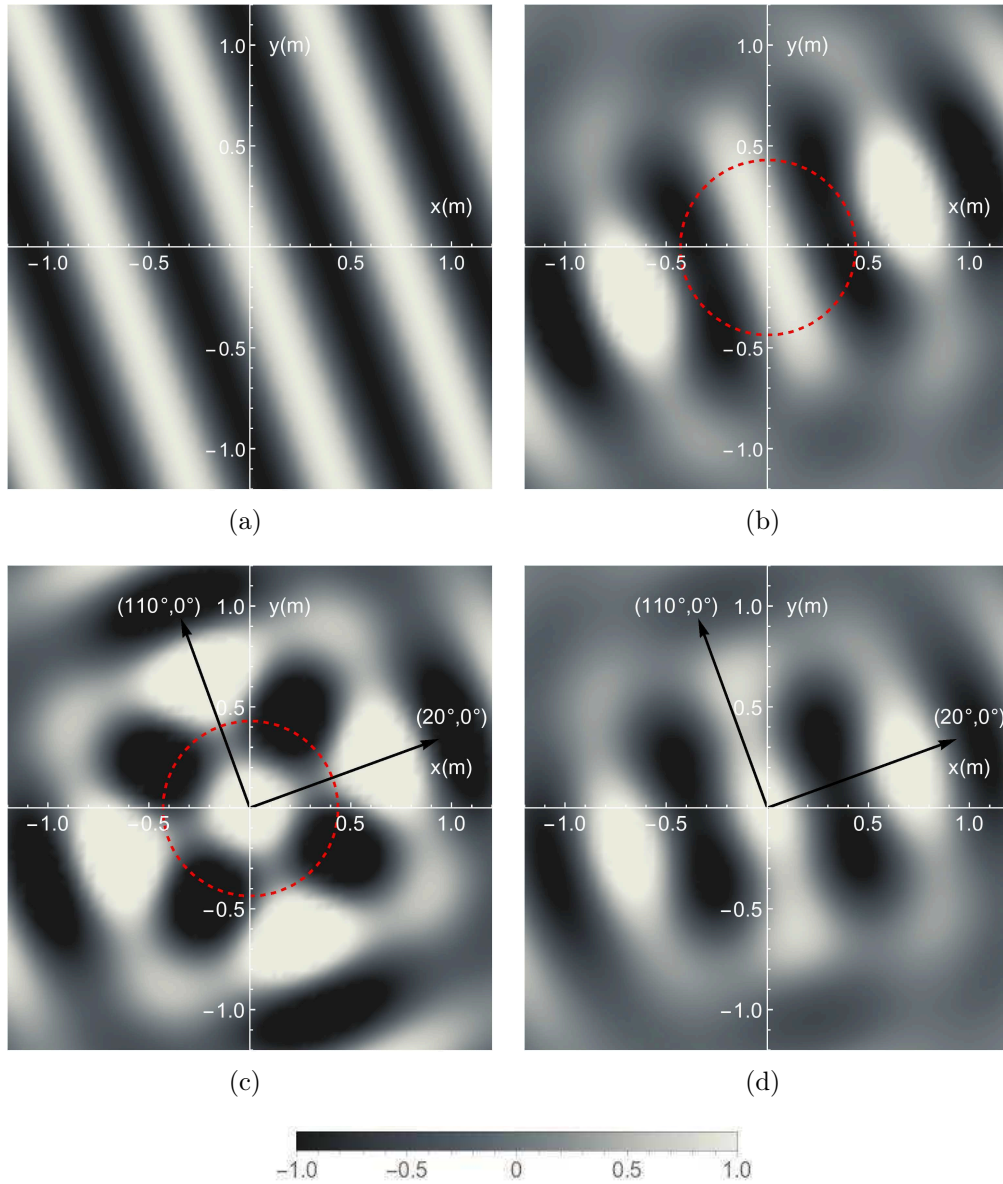


Figure 5.2 (a) Partie réelle du champ de pression acoustique d'une onde plane. (b) Représentation tronquée de la même onde plane à l'ordre $M = 4$. Le cercle en rouge donne la taille de la zone de reconstruction suivant l'équation (2.32). (c) Partie réelle du champ de pression acoustique, tronqué à l'ordre $M = 4$ pour le cas de deux ondes planes de direction de provenance telles qu'indiquées par les flèches noires. Le cercle rouge montre la taille de la zone efficace de reconstruction suivant l'équation (2.32). (d) Partie réelle du champ de pression acoustique montré en (c) filtré par \mathbf{M} , lorsque défini pour une hyper-cardioïde d'ordre $M' = 1$ visant à $(20^\circ, 0^\circ)$. L'ordre de redécomposition après filtrage est de $\tilde{M} = M + M' = 5$.

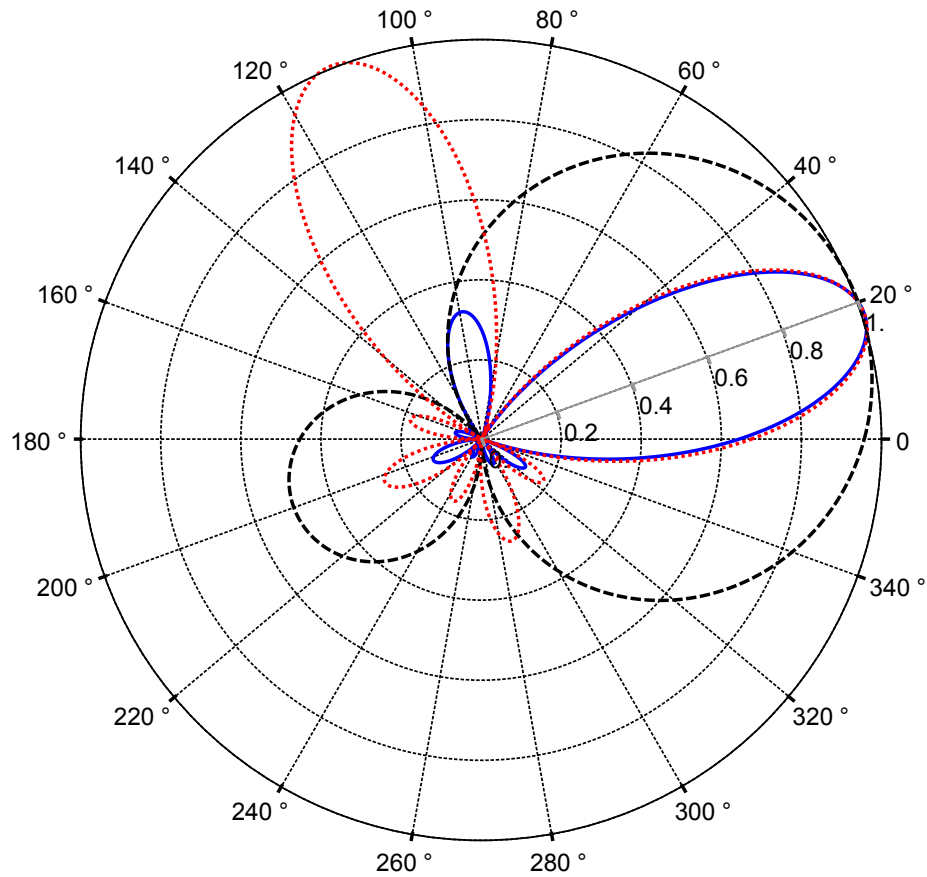


Figure 5.3 Diagrammes de directivité dans le plan horizontal ($\delta = 0$) de la scène sonore à filtrer $b(\theta, 0)$ (en rouge pointillé), de la fonction de directivité utilisée pour le filtrage $g(\theta, 0)$ (en noir pointillé) et le résultat après filtrage $\tilde{b}(\theta, 0)$ (en bleu). Les diagrammes sont normalisés par leur valeur absolue maximale.

5.4.2 Dirac angulaire

Un deuxième cas de simulation utilisant un Dirac angulaire pour g , comme décrit à la section 5.3.2, est montré par la figure 5.4 pour une représentation tronquée à l'ordre $\tilde{M} = 5$. Le Dirac vise dans la direction $(20^\circ, 0^\circ)$.

Tel que pressenti par l'équation (5.10), ce résultat de filtrage montre la représentation d'une onde plane tronquée à l'ordre $\tilde{M} = 5$, et d'amplitude $\frac{1}{4\pi}b(20^\circ, 0^\circ)$.

5.5 Mise en oeuvre temps-réel

La matrice \mathbf{M} est implémentée dans la suite logicielle *ambitools* [Lecomte et Gauthier, 2015] (voir les annexes C et D), pour un Dirac angulaire (outil `hoa_beamformin_dirac_to_hoa`, équation (5.8)), et pour des fonctions de directivités hyper-cardioïdes jusqu'à l'ordre

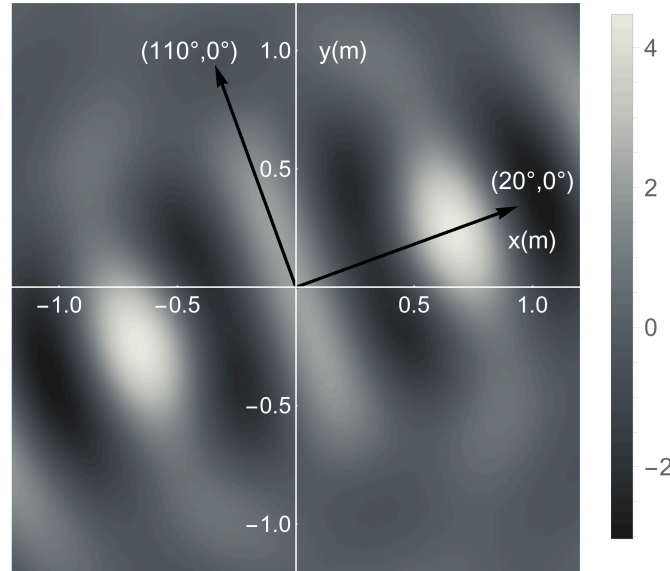


Figure 5.4 Partie réelle du champ de pression acoustique filtré par \mathbf{M} , lorsque g est défini par un Dirac angulaire de direction $(20^\circ, 0^\circ)$. Représentation tronquée à l'ordre $\tilde{M} = 5$. Les flèches indiquent la direction de provenance des deux ondes planes qui composent la scène sonore \mathbf{b} .

$M' = 3$, (outil `hoa_beamforming_hypercardioid_to_hoa`, équation (5.12)). Les dimensions de \mathbf{M} sont telles que $M = \tilde{M} = 5$. Cette matrice de filtrage s'insère dans le domaine ambisonique avant le décodage (c'est à dire avant la restitution sur réseau de haut-parleurs du champ sonore filtré spatialement).

5.6 Conclusions

Ce chapitre a présenté une approche pour le filtrage spatial directif d'une scène sonore tridimensionnelle encodée dans un formalisme ambisonique d'ordre élevé. En utilisant des fonctions de directivité issues de la littérature de la formation de voies, il est possible de filtrer spatialement la scène sonore pour obtenir une nouvelle scène sonore tridimensionnelle prête à être décodée sur un réseau de haut-parleurs. Les résultats de simulations montrent qu'il est possible de sélectionner certaines directions dans la champ sonore original. Les compromis quant à l'ordre de troncature et de recomposition nécessaire ont également été discutés. La validation expérimentale de cette transformation sur les champs sonores dans le domaine des harmoniques sphériques est réalisée au chapitre 6 avec l'outil `hoa_beamforming_hypercardioid_to_hoa`.

CHAPITRE 6

DISPOSITIFS EXPÉRIMENTAUX ET MESURES DE CHAMPS SONORES

6.1 Introduction

Le but principal de ce chapitre est d'effectuer la validation expérimentale et scientifique des méthodes et techniques proposées dans les chapitres précédents, à l'exception du chapitre 4 pour le moment. À cette fin, la géométrie sphérique a été retenue pour la captation et reconstruction tridimensionnelle de champs sonores car elle ne privilégie aucune direction et elle est la plus directe pour les approches ambisoniques.

Le chapitre est scindé en deux parties. Dans un premier temps la réalisation de deux sphères de restitution ambisonique « SpherBedev » est détaillée, ainsi que la réalisation d'une antenne sphérique de microphones « MemsBedev ». À l'aide de ces dispositifs il devient alors possible de capter un champ sonore en trois dimensions et de le restituer sur une zone étendue de l'espace.

Ainsi, dans un second temps, le chapitre relate une campagne de mesure effectuée au GAUS en chambre anéchoïque. L'objectif est alors d'utiliser les outils présentés à la première partie du chapitre pour encoder, synthétiser, reproduire et mesurer expérimentalement des champs sonores, en utilisant la théorie développée au cours de chapitres précédents. Le traitement des signaux de commande des haut-parleurs est fait en temps-réel à l'aide de la suite logicielle *ambitools*, développée dans le cadre du doctorat (voir les annexes C et D).

6.2 Sphère de restitution ambisonique « SpherBedev »

Dans le contexte de la cotutelle du doctorat, deux sphères de restitution ambisonique sont réalisées : l'une au LMSSC du CNAM en France et l'autre au GAUS à l'UdeS au Québec. Ces deux dispositifs ont sensiblement les mêmes dimensions, mais diffèrent sur quelques points détaillés ci-après. Pour la suite, la sphère du LMSSC sera dénommée « SpherBedev 1.0 » et la sphère du GAUS « SpherBedev 2.0 ».

La SpherBedev 1.0 a été réalisée au LMSSC entre novembre 2013 et mars 2014. Ce premier prototype, qui a été conçu de façon modulaire, a été envoyé au GAUS lors du premier séjour de l’auteur à l’UdeS, entre juillet et décembre 2014. Au retour de la SpherBedev 1.0 en France, le GAUS a investi pour la construction d’un second prototype comprenant quelques améliorations. Les pièces ont été usinées au cours de l’été 2015 et la sphère a été assemblée au début du second séjour de l’auteur à l’UdeS, en janvier 2016.

6.2.1 Configuration géométrique

Les deux prototypes SpherBedev se basent sur un maillage de Lebedev à cinquante noeuds, tel que présenté au chapitre 3. En effet, comme il a été montré dans ce chapitre, cette grille présente plusieurs avantages pour la restitution sonore spatialisée par approches ambisoniques, notamment :

- L’intégration exacte de produits d’harmoniques sphérique jusqu’à l’ordre $M = 5$ avec 50 noeuds.
- Un bon rapport du nombre de noeuds avec le nombre de fonctions à intégrer (50 noeuds pouvant intégrer 36 modes ambisoniques).
- Une flexibilité d’utilisation : trois grilles en une (6 noeuds jusqu’à $M = 1$, 26 noeuds jusqu’à $M = 3$ et 50 noeuds jusqu’à $M = 5$).
- Une répartition homogène sur la sphère, sans concentration excessive de noeuds dans une région spatiale.

Types de noeuds

Selon la règle de quadrature de Lebedev Eq. (3.35), présentée au Chapitre 3, 4 types de noeuds sont utilisés pour construire la grille :

- Les 6 sommets d’un octaèdre. Pour la suite, ce type de noeud est désigné par *Vertex* et la couleur **noire** (SpherBedev 1.0) ou **jaune** (SpherBedev 2.0) lui est associé.
- Les 12 projections sur la sphère du milieu des arêtes d’un octaèdre. Pour la suite, ce type de noeud est désigné par *Edge* et la couleur **bleue** lui est associé.
- Les 8 projections sur la sphère du milieu des faces d’un octaèdre. Pour la suite, ce type de noeud est désigné par *Face* et la couleur **verte** lui est associé.

- Les 24 projections de points situés sur la bissectrice du triangle équilatéral formé par la face d'un octaèdre (voir la figure [3.3](#)). Pour la suite, ce type de noeud est désigné par *LM* et la couleur **rouge** lui est associée.

Le tableau [6.1](#) récapitule les spécificités des 4 types de noeuds.

Désignation	Nombre	Couleur
<i>Vertex</i>	6	noire (SpherBedev 1.0), jaune (SpherBedev 2.0)
<i>Edge</i>	12	Bleu
<i>Face</i>	8	Vert
<i>LM</i>	24	Rouge
Total		50

Tableau 6.1 Types de points utilisés pour le maillage de la sphère.

Numérotation et coordonnées

Pour numéroter les noeuds, la convention suivante est choisie : pour chaque type de noeud, la numérotation est faite par décroissance de l'élévation δ et croissance de l'azimut θ . Suivant cette convention, les noeuds 1 à 6 sont de types *Vertex*, les noeuds de 7 à 18 sont de type *Edge*, les noeuds 19 à 26 sont de type *Face* et les noeuds de type 27 à 50 de type *LM*. La grille contenant les 50 noeuds et leur numérotation est visible à la figure [6.1](#). Cette numérotation permet d'exploiter facilement les sous-grilles incluses dans la grille de 50 noeuds. En effet, en utilisant des noeuds 1 à 6, la grille fonctionnant à l'ordre $M = 1$ est décrite. De même, en utilisant les noeuds de 1 à 26, la grille fonctionnant à l'ordre $M = 3$ est décrite.

Les deux prototypes SpherBedev ont un rayon $r_{\text{spk}} = 1,07$ m. Ce rayon est choisi comme un compromis entre un faible encombrement et une taille suffisante pour permettre à quelques utilisateurs de prendre place au centre de la sphère. En effet, il est voulu de pouvoir installer la SpherBedev dans une salle aux dimensions standards d'un bureau ou espace de laboratoire. Les coordonnées cartésiennes et sphériques des noeuds sont données au tableau [6.2](#).

Noeuds	x (m)	y (m)	z (m)	r (m)	θ (rad)	δ (rad)
1	0	0	1,07	1,07	0	$\frac{\pi}{2}$
2	1,07	0	0	1,07	0	0
3	0	1,07	0	1,07	$\frac{\pi}{2}$	0
4	-1,07	0	0	1,07	π	0
5	0	-1,07	0	1,07	$\frac{3\pi}{2}$	0
6	0	0	-1,07	1,07	0	$-\frac{\pi}{2}$

7	0,756604	0	0,756604	1,07	0	$\frac{\pi}{4}$
8	0	0,756604	0,756604	1,07	$\frac{\pi}{2}$	$\frac{\pi}{4}$
9	-0,756604	0	0,756604	1,07	π	$\frac{\pi}{4}$
10	0	-0,756604	0,756604	1,07	$\frac{3\pi}{2}$	$\frac{\pi}{4}$
11	0,756604	0,756604	0	1,07	$\frac{\pi}{4}$	0
12	-0,756604	0,756604	0	1,07	$\frac{3\pi}{4}$	0
13	-0,756604	-0,756604	0	1,07	$\frac{5\pi}{4}$	0
14	0,756604	-0,756604	0	1,07	$\frac{7\pi}{4}$	0
15	0,756604	0	-0,756604	1,07	0	$-\frac{\pi}{4}$
16	0	0,756604	-0,756604	1,07	$\frac{\pi}{2}$	$-\frac{\pi}{4}$
17	-0,756604	0	-0,756604	1,07	π	$-\frac{\pi}{4}$
18	0	-0,756604	-0,756604	1,07	$\frac{3\pi}{2}$	$-\frac{\pi}{4}$
19	0,617765	0,617765	0,617765	1,07	$\frac{\pi}{4}$	0,61548
20	-0,617765	0,617765	0,617765	1,07	$\frac{3\pi}{4}$	0,61548
21	-0,617765	-0,617765	0,617765	1,07	$\frac{5\pi}{4}$	0,61548
22	0,617765	-0,617765	0,617765	1,07	$\frac{7\pi}{4}$	0,61548
23	0,617765	0,617765	-0,617765	1,07	$\frac{\pi}{4}$	-0,61548
24	-0,617765	0,617765	-0,617765	1,07	$\frac{3\pi}{4}$	-0,61548
25	-0,617765	-0,617765	-0,617765	1,07	$\frac{5\pi}{4}$	-0,61548
26	0,617765	-0,617765	-0,617765	1,07	$\frac{7\pi}{4}$	-0,61548
27	0,322617	0,322617	0,967851	1,07	0,785398	1,13029
28	-0,322617	0,322617	0,967851	1,07	2,35619	1,13029
29	-0,322617	-0,322617	0,967851	1,07	3,92699	1,13029
30	0,322617	-0,322617	0,967851	1,07	5,49779	1,13029
31	0,967851	0,322617	0,322617	1,07	0,321751	0,306277
32	0,322617	0,967851	0,322617	1,07	1,24905	0,306277
33	-0,322617	0,967851	0,322617	1,07	1,89255	0,306277
34	-0,967851	0,322617	0,322617	1,07	2,81984	0,306277
35	-0,967851	-0,322617	0,322617	1,07	3,46334	0,306277
36	-0,322617	-0,967851	0,322617	1,07	4,39064	0,306277
37	0,322617	-0,967851	0,322617	1,07	5,03414	0,306277
38	0,967851	-0,322617	0,322617	1,07	5,96143	0,306277
39	0,967851	0,322617	-0,322617	1,07	0,321751	-0,306277
40	0,322617	0,967851	-0,322617	1,07	1,24905	-0,306277
41	-0,322617	0,967851	-0,322617	1,07	1,89255	-0,306277

42	-0,967851	0,322617	-0,322617	1,07	2,81984	-0,306277
43	-0,967851	-0,322617	-0,322617	1,07	3,46334	-0,306277
44	-0,322617	-0,967851	-0,322617	1,07	4,39064	-0,306277
45	0,322617	-0,967851	-0,322617	1,07	5,03414	-0,306277
46	0,967851	-0,322617	-0,322617	1,07	5,96143	-0,306277
47	0,322617	0,322617	-0,967851	1,07	0,785398	-1,13029
48	-0,322617	0,322617	-0,967851	1,07	2,35619	-1,13029
49	-0,322617	-0,322617	-0,967851	1,07	3,92699	-1,13029
50	0,322617	-0,322617	-0,967851	1,07	5,49779	-1,13029

Tableau 6.2 Coordonnées cartésiennes et sphériques des différents noeuds de la SpherBedev.

6.2.2 Enceintes de reproduction

SpherBedev 1.0

Pour le premier prototype SpherBedev 1.0, les haut-parleurs sont des AURA SOUND NSW2-380-8A¹. Il s'agit de haut-parleurs 2 po à large bande. Les haut-parleurs sont montés en enceinte close à l'aide d'un manchon en PVC de 63 mm de long (voir la figure 6.2(a)). L'enceinte est fermée à l'arrière par un capuchon en PVC. L'étanchéité est réalisée à l'aide de mousse néoprène autour du haut-parleur. La cavité de l'enceinte est remplie avec de la mousse synthétique (voir la figure 6.2(b)). Finalement, une connectique RCA sur le capuchon arrière sert à connecter l'enceinte sur une sortie d'amplificateur.

L'amplitude normalisée de la FRF à 1 m dans l'axe du haut-parleur est mesurée en chambre anéchoïque. Elle est visible à la figure 6.3. D'après cette figure, l'enceinte possède une bande passante allant de 200 Hz à 15 kHz environ. Entre 200 Hz et 1,5 kHz environ, la FRF possède une très faible dynamique. Cependant à partir de 1,5 kHz et jusqu'à 15 kHz, la FRF présente une plus grande dynamique avec notamment un accident à 6 kHz.

Par ailleurs, les équations du décodage de l'Ambisonie modélisent les sources de restitution comme des monopoles. Il est alors important d'avoir une figure de directivité la plus omnidirectionnelle possible pour les enceintes de restitution. Une mesure de directivité a été effectuée dans le cadre d'un stage de Master M1 IMDEA au Mans [Chouquet et Hamoui, 2014]. La figure 6.4 montre la directivité d'une enceinte de la SpherBedev 1.0 à différentes fréquences. D'après cette figure, l'enceinte possède une directivité omnidirectionnelle jusqu'à 2000 Hz et devient légèrement plus directive au delà. Les enceintes de

¹Données du constructeur : <http://sekisushai.net/THESE/SpherBedev1.0/NSW2-326-8A.pdf>

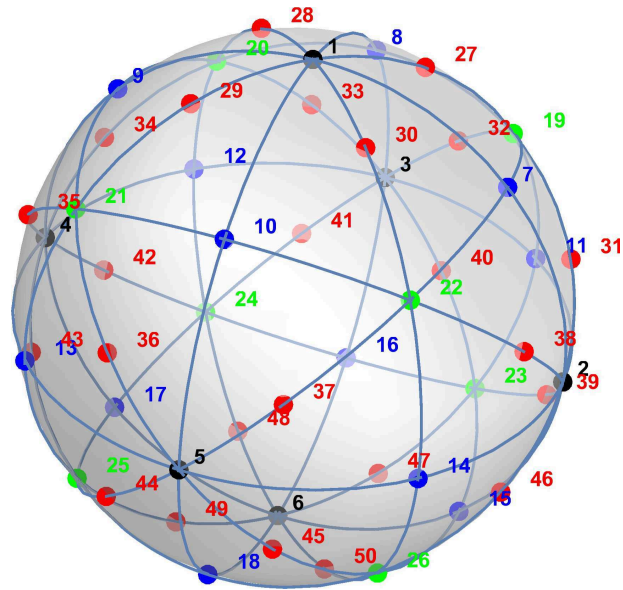


Figure 6.1 Les 50 noeuds numérotés du maillage de Lebedev utilisé : 1 à 6 de type *Vertex*, 7 à 18 de type *Edge*, 19 à 26 de type *Face* et 27 à 50 de type *LM*.

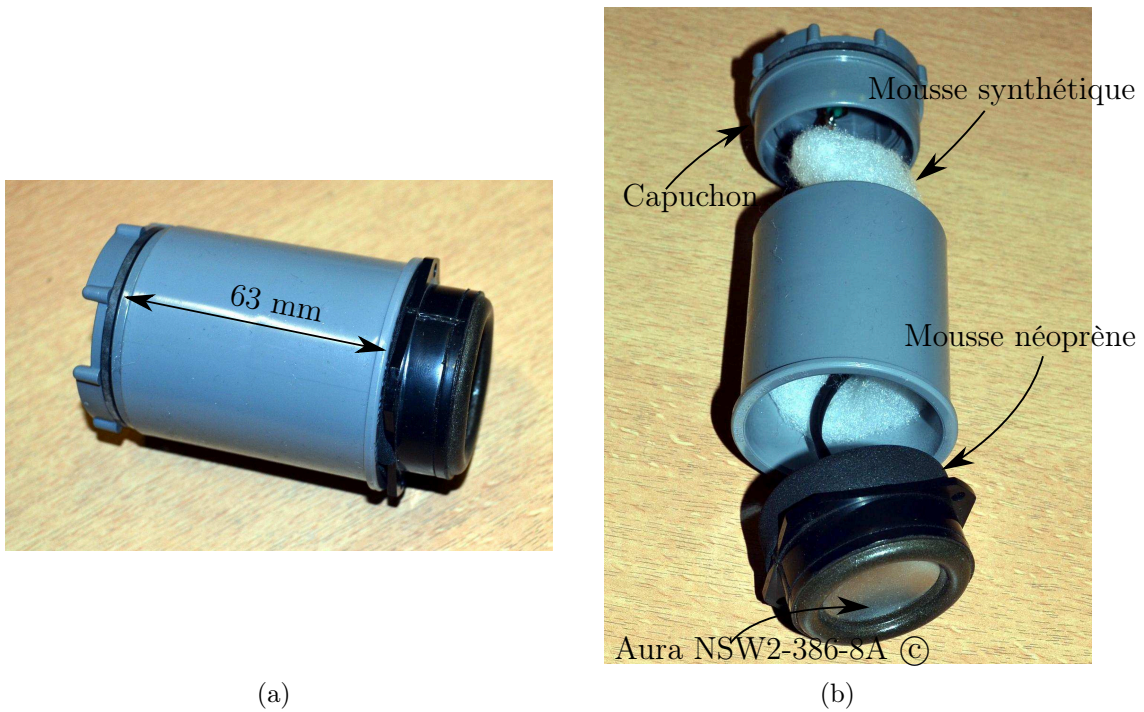


Figure 6.2 Détails d'une enceinte de la SpherBedev 1.0. (a) Enceinte montée. (b) Vue éclatée.

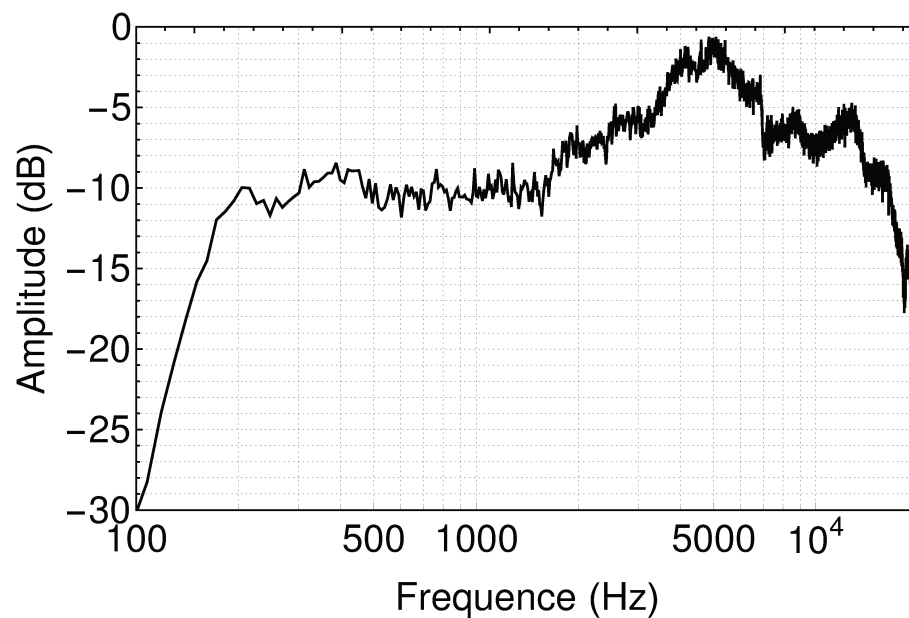


Figure 6.3 Amplitude normalisée de la FRF d'une enceinte de la SpherBedev 1.0 dans l'axe à 1 m, mesurée en chambre anéchoïque.

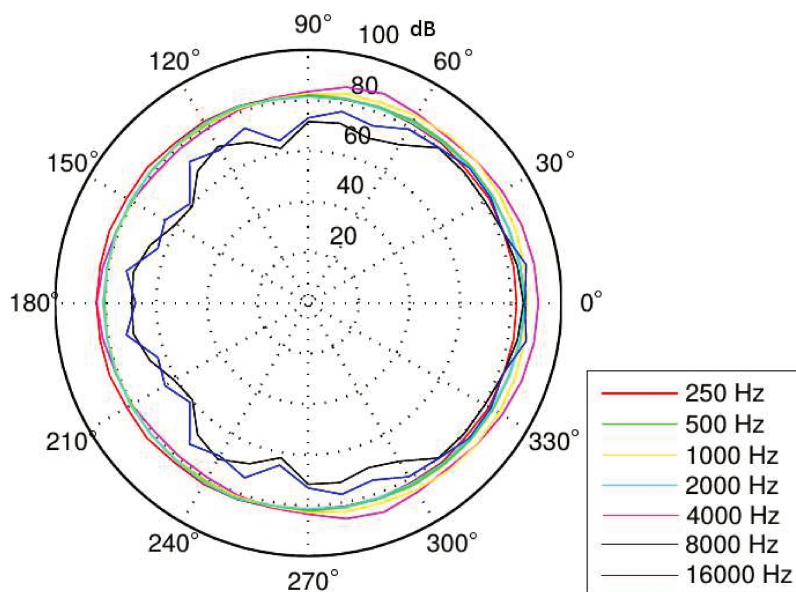


Figure 6.4 Directivité d'une enceinte de la SpherBedev 1.0 à différentes fréquences. Axe du haut-parleur à 0°.

la SpherBedev 1.0 sont ainsi exploitables entre 200 Hz et 1,5 kHz environ pour la reproduction ambisonique sans égalisation importante de leurs FRF. Cependant, la principale limite de ces enceintes est leurs réponses en basses fréquences, en dessous de 200 Hz. Dans les applications pratiques, un caisson de basse est ajouté à côté du dispositif avec comme signal d'entrée la composante ambisonique d'ordre 0 (B_{00} , pression à l'origine O), filtrée à l'aide d'un passe-bas avec une fréquence de coupure de 200 Hz.

SpherBedev 2.0

Les haut-parleurs de la SpherBedev 2.0 sont des AURA SOUND 3 po NS3-193-8A². Ces haut-parleurs présentent une bande-passante étendue dans les basses fréquences par rapport aux haut-parleurs de la SpherBedev 1.0 (voir section 6.2.2). Cependant ils nécessitent un plus grand volume pour une montage en enceinte (1.6 L). La figure 6.5 présente plusieurs vues du montage de quelques enceintes. L'enceinte close est réalisée à l'aide d'un tube en carton de longueur 31 cm. Ce tube est rempli de mousse synthétique. L'arrière est clos par un capuchon de plastique au milieu duquel un connecteur RCA est monté. Un support en plastique permet de solidariser le haut-parleur avec le tube en carton. L'étanchéité acoustique de l'enceinte est faite à l'aide de silicone. Finalement, une bande de mousse néoprène permet de limiter les transmissions solidiennes entre l'enceinte et la structure, une fois le dispositif monté.

La FRF d'une enceinte est mesurée à 1 m dans l'axe en chambre anéchoïque. Elle est visible sur la figure 6.6. La bande passante à -10 dB annoncée par la fiche du constructeur est de 80 Hz-15 kHz. Elle semble ici plutôt de 100 Hz-15 kHz. Un accident notable dans la FRF est visible à 7 kHz, qui peut être dû à une résonance d'un des composants du haut-parleur lui-même, étant donné que cet accident est également visible dans la FRF donnée par la fiche du constructeur. A noter également que la réponse en phase n'est pas linéaire.

6.2.3 Structure

La structure des deux SpherBedevs 1.0 et 2.0 repose sur le même principe : un cadre fait de barres d'aluminium permet de soutenir la structure sphérique par les 8 noeuds de son équateur. Les noeuds sont matérialisés par des anneaux métalliques dans lesquels les enceintes cylindriques sont montées de façon concentrique à l'aide de trois vis à 120° . Finalement, les anneaux sont reliés entre eux par des barres cylindriques en aluminium au bout desquelles des barres en acier plié permettent de solidariser les anneaux aux barres

²Données du constructeur : <http://sekisushai.net/THESE/SpherBedev2.0/NS3-193-8A.pdf>

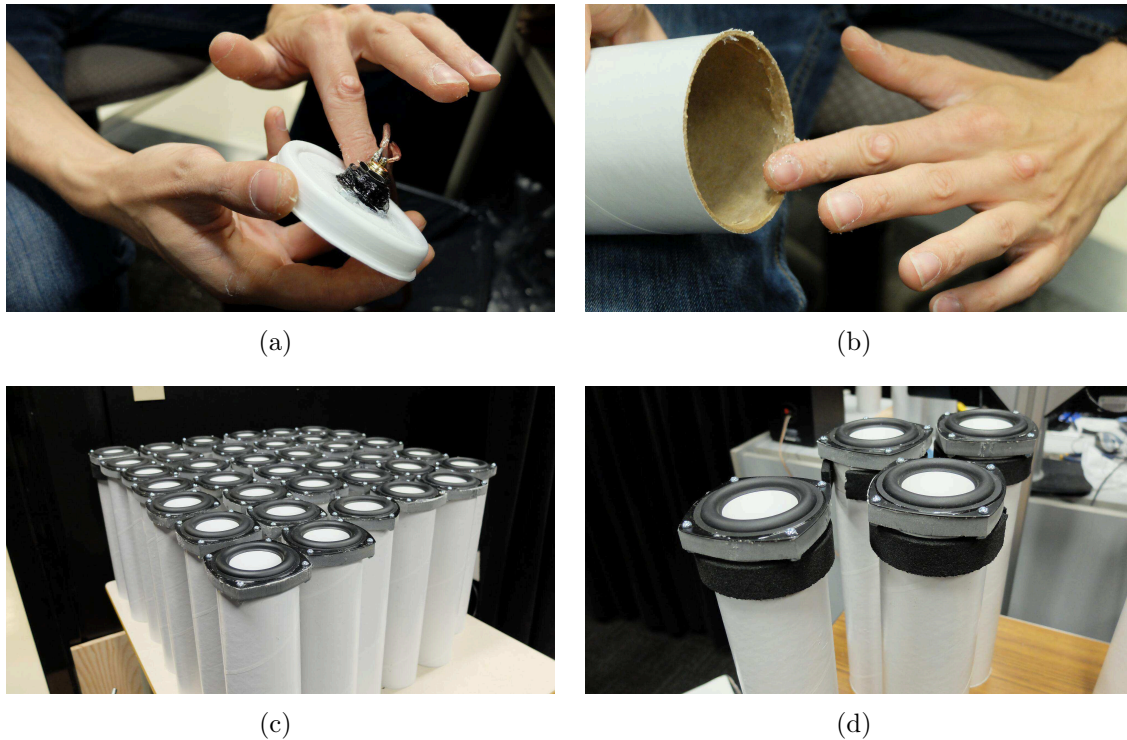


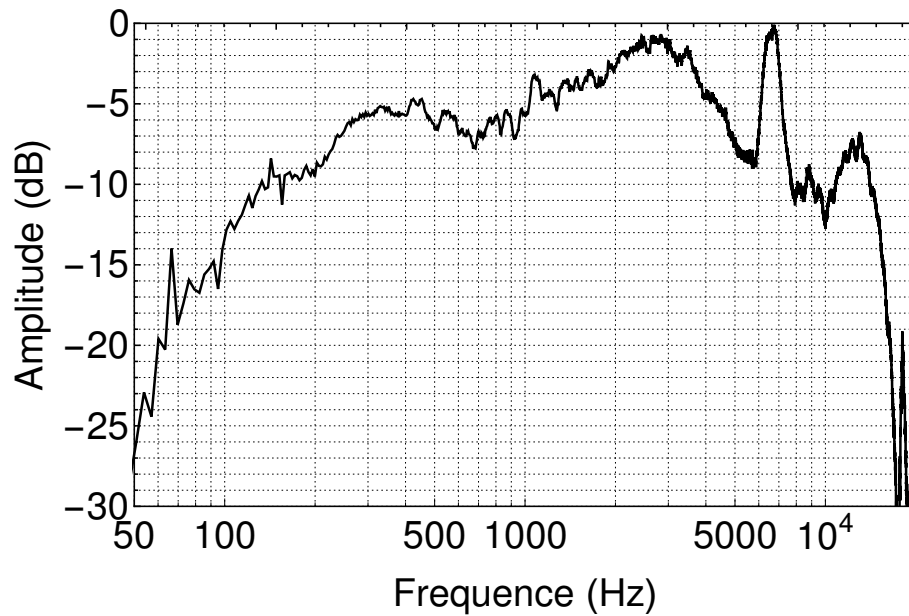
Figure 6.5 Détails du montage de quelques enceintes de la SpherBedev 2.0. (a) Étanchéité d'un capuchon arrière à l'aide de silicone. Le connecteur RCA est visible au milieu du capuchon. (b) Étanchéité d'un tube en carton à l'aide de silicone. (c) Quelques enceintes une fois assemblées. (d) Quelques enceintes assemblées avec une bande de néoprène pour limiter les transmissions solidiennes entre l'enceinte et la structure.

cylindriques. La figure 6.7 présente quelques détails de la structure pour la SpherBedev 2.0.

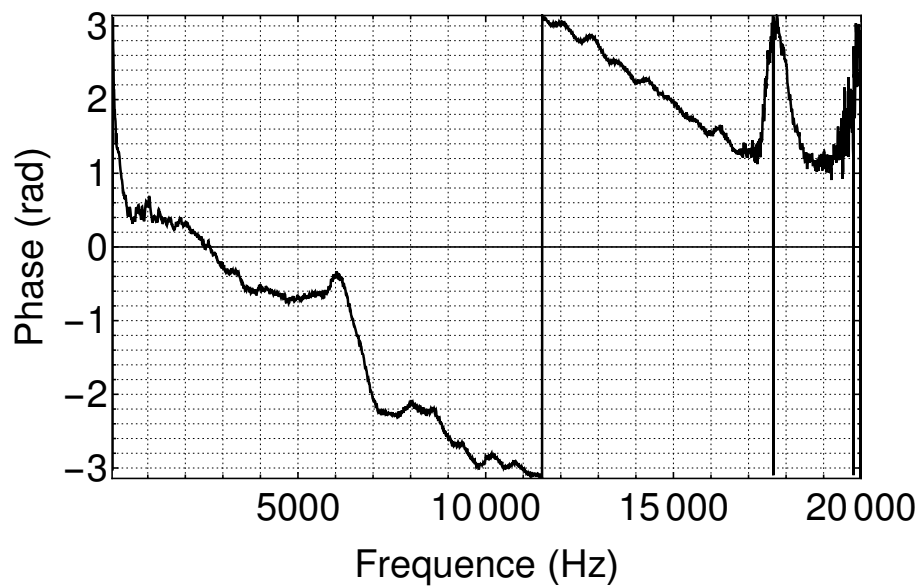
La hauteur du cadre d'aluminium est telle que l'équateur de la sphère est à 1.78 m du sol. Cette hauteur est choisie comme la hauteur moyenne d'un être humain de type européen [Grasgruber *et al.*, 2014]. Pour les personnes plus petites et souhaitant mettre leurs oreilles au plan de l'équateur, un piédestal peut être utilisé.

Pour entrer dans la SpherBedev, une ouverture est réalisée en enlevant une des barres en aluminium qui relie deux anneaux dans l'hémisphère inférieur. Ceci ne compromet pas la rigidité de la structure puisque les anneaux sont fixés au minimum à 3 barres d'aluminium.

La structure est conçue de manière entièrement modulaire. Par exemple, il est tout à fait possible de ne monter que l'hémisphère supérieur pour obtenir un dôme et dégager ainsi l'espace inférieur pour y placer plus d'auditeurs. Une vidéo « accéléré » (*timelaps*) du



(a)



(b)

Figure 6.6 FRF d'une enceinte de la SpherBedev 2.0 dans l'axe à 1 m, mesurée en chambre anéchoïque. (a) Amplitude normalisée. (b) Phase avec le retard de propagation compensé.

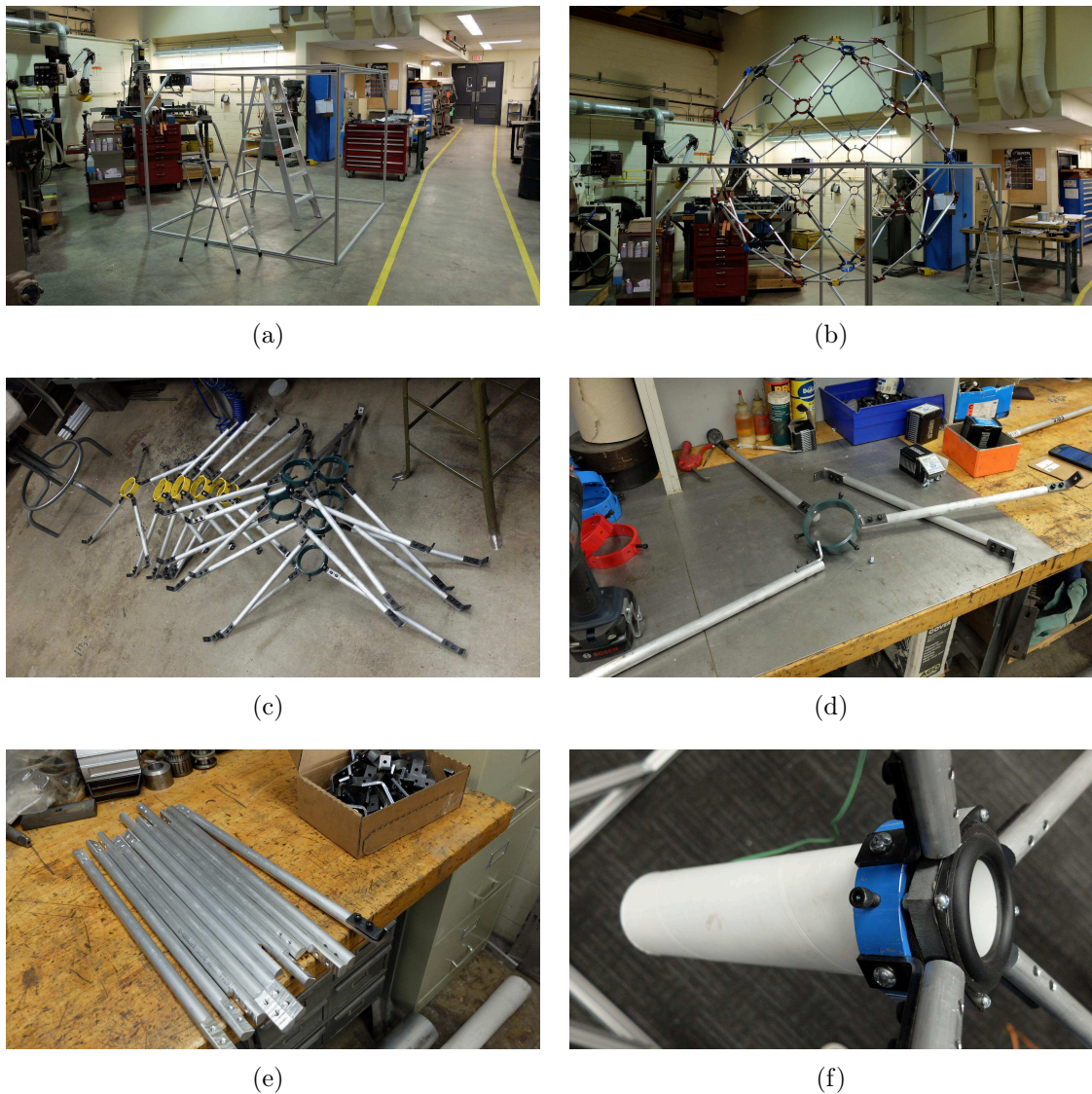


Figure 6.7 Détails de la structure pour la SpherBedev 2.0. (a) Cadre de support en aluminium dans les ateliers de l'UdeS. (b) Structure une fois montée, sans les enceintes. (c) Détails des anneaux auxquels sont montés les barres d'aluminium. (d) Trois vis à 120° permettent de maintenir et centrer les enceintes. (e) Détails des barres en aluminium qui relient les anneaux entre eux. (f) Une enceinte montée dans un anneau de la structure.

montage est disponible en ligne pour chacune des SpherBedev 1.0³ et 2.0⁴. Finalement, la figure 6.8 présente une photo de la SpherBedev 1.0 et 2.0 une fois montées et câblées.

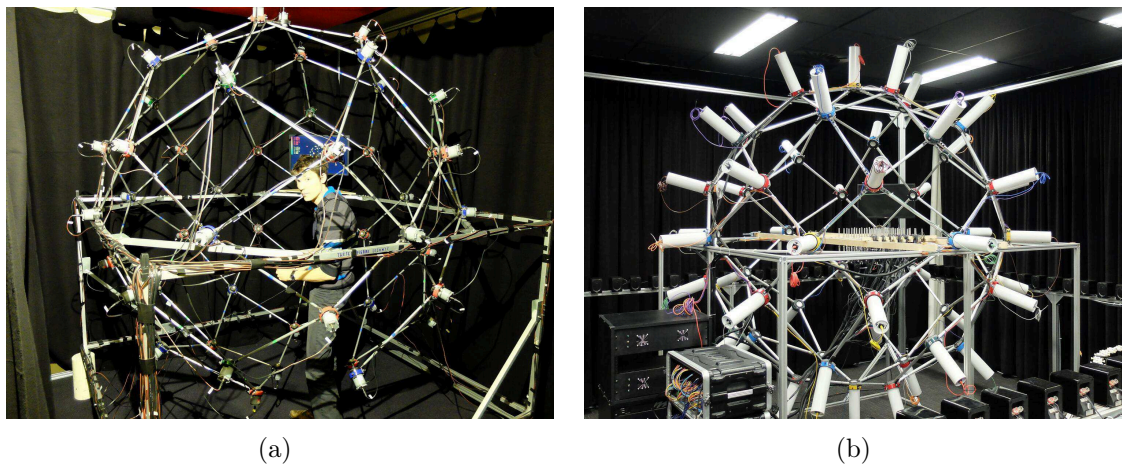


Figure 6.8 Sphère de restitution ambisonique jusqu'à l'ordre 5 SpherBedev. (a) SpherBedev 1.0 au CNAM. L'auteur se tient au centre du dispositif. (b) SpherBedev 1.0 à l'UdeS.

6.2.4 Chaînes audio-numériques

Le protocole MADI (Multichannel Audio Digital Interface) [AES, 1991] est retenu pour transmettre les signaux numériques des enceintes de la SpherBedev. En effet, ce protocole permet de manipuler un grand nombre de signaux audio-numériques et de les transmettre à l'aide d'un seul câble coaxial ou d'une fibre optique. Le protocole MADI supporte jusqu'à 64 canaux à 24 Bits - 48 kHz par câble, ce qui est suffisant dans le cas présent, où 50 voies sont nécessaires. Pour les deux sphères SpherBedev, l'instrumentation est faite en temps-réel sous Linux à l'aide de la suite logicielle *ambitools*, présentée aux annexes C et D.

SpherBedev 1.0

La carte son utilisée pour la SpherBedev 1.0 est une RME HDSPe MADIface⁵. La sortie numérique MADI est reliée à deux convertisseurs numériques analogiques RME M-32 DA⁶ à l'aide de fibres optiques. Les sorties analogiques sont reliées à un bâti de 5 amplificateurs DAYTON AUDIO MA1240a⁷ à l'aide de câbles jack TRS (Tip, Ring, Sleeve) 1/4 po vers RCA. Ces amplificateurs sont de classe AB avec une puissance de 40 Watts à 8 Ohms.

³<http://sekisushai.net/THESE/SpherBedev1.0/SpherBedev1.avi>

⁴<http://sekisushai.net/THESE/SpherBedev2.0/SpherBedev2.avi>

⁵Fiche technique : <http://www.sekisushai.net/THESE/SpherBedev1.0/HDSPE-MADIFace.pdf>

⁶Fiche technique : <http://www.sekisushai.net/THESE/SpherBedev1.0/M-32-DA.pdf>

⁷Fiche technique : <http://www.sekisushai.net/THESE/SpherBedev1.0/MA1240a.pdf>

Finalement, les sorties des amplificateurs sont reliées aux enceintes à l'aide de câbles connecteurs Phoenix vers RCA, réalisés au LMSSC.

SpherBedev 2.0

La carte son utilisée pour la SpherBedev 2.0 est une RME HDSPe MADI⁸. La sortie numérique MADI est reliée à un convertisseur numérique-numérique RME ADI-648⁹ qui convertit des signaux MADI vers ADAT (Alesis Digital Audio Tape) dans les deux sens. Le protocole ADAT permet de transmettre jusqu'à 8 signaux numériques à 24 Bits - 48 kHz à l'aide de fibre optiques. Ainsi, les sorties ADAT du convertisseur ADI-648 sont reliées à 7 convertisseurs numériques-analogiques BEHRINGER ADA8200¹⁰ à l'aide de fibre optiques. Les sorties analogiques des ADA8200 sont reliées à l'aide de câbles jacks TRS 1/4 po à un bâti de 7 amplificateurs, réalisées au GAUS. Les amplificateurs sont réalisés à partir de cartes SURE TK2050¹¹ 4 × 100 Watts de classe D. Un pont diviseur de facteur 10 est appliqué en entrée de chacune des voies des amplificateurs pour limiter la puissance de ces derniers et prévenir les risques de détérioration des enceintes en cas d'erreur de manipulation des signaux de commande des haut-parleurs. La figure 6.9 montre le bâti de convertisseurs et d'amplificateurs pour la SpherBedev 2.0.

6.3 Antenne sphérique de microphone « MemsBedev »

Un prototype d'antenne sphérique baptisé « MemsBedev » est réalisé dans le cadre du doctorat par une collaboration entre le CNAM et l'entreprise CINELA¹². Il s'agit d'un réseau de 50 capsules microphoniques de type MEMS (Micro Electro-Mechanical System) réparties selon un maillage de Lebedev (d'où le nom « MemsBedev ») sur une surface sphérique en plastique supposée acoustiquement rigide.

6.3.1 Configuration géométrique

Le maillage utilisé est une grille de Lebedev à 50 noeuds telle que présentée au chapitre 3. La numérotation et les coordonnées des noeuds sont les mêmes que pour la SpherBedev (voir section 6.2), à ceci près que le rayon du microphone est de $r_{\text{mic}} = 7$ cm. Ce rayon est choisi comme un compromis entre l'amplification maximale acceptable des signaux en basses fréquences lors de l'encodage ambisonique et la fréquence de repliement théorique

⁸Fiche technique : <http://www.sekisushai.net/THESE/SpherBedev2.0/HDSPe-MADI.pdf>

⁹Fiche technique : <http://www.sekisushai.net/THESE/SpherBedev2.0/ADI-648.pdf>

¹⁰Fiche technique : <http://www.sekisushai.net/THESE/SpherBedev2.0/ADA8200.pdf>

¹¹Guide utilisateur : <http://www.sekisushai.net/THESE/SpherBedev2.0/TK2050.pdf>

¹²<http://www.cinela.fr>

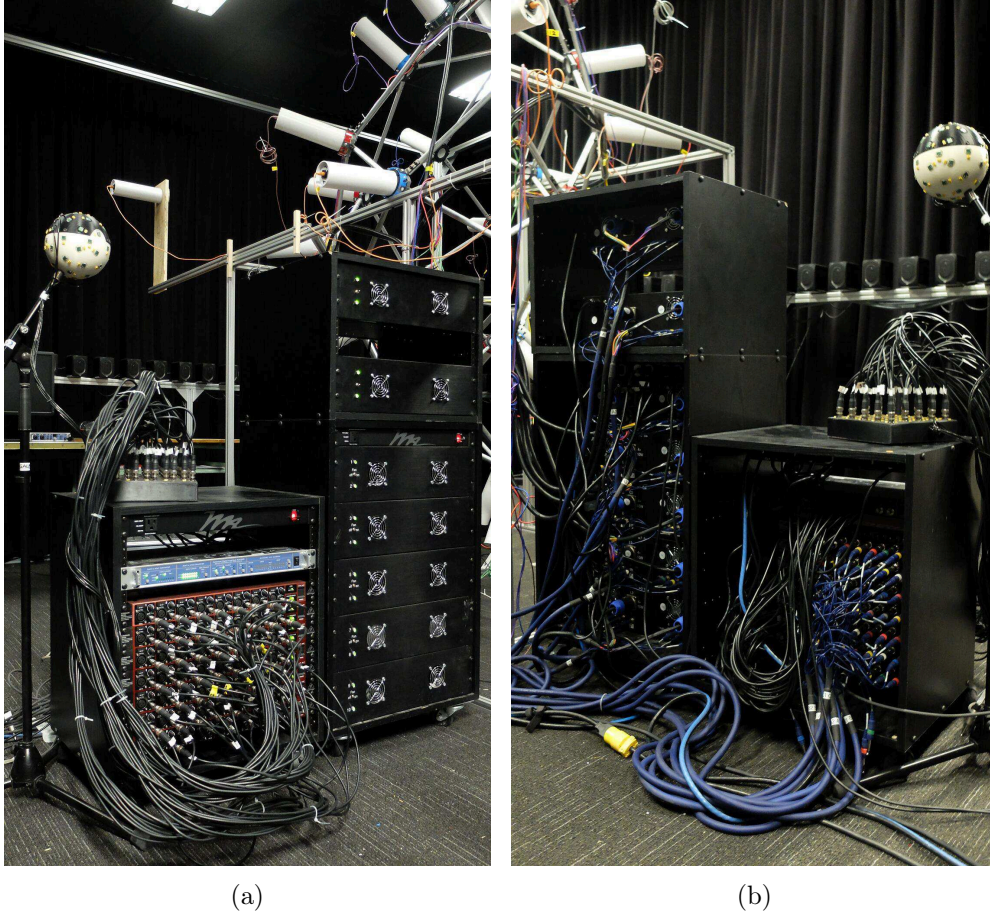


Figure 6.9 Bâti de convertisseurs et d’amplificateurs de la SpherBedev 2.0. (a) Vue de devant : le bâti des convertisseurs sur la gauche de l’image, le bâti d’amplificateurs sur la droite. Le microphone MemsBedev, dans le coin supérieur gauche de l’image, est câblé à l’entrée des convertisseurs. (b) Même situation vue de derrière.

du microphone (voir section. 6.3.5). De la même manière, le maillage de Lebedev retenu permet une utilisation du microphone jusqu’à l’ordre $M = 1$ avec les 6 premières capsules microphoniques, ou bien jusqu’à l’ordre $M = 3$ avec les 26 premières capsules ou finalement jusqu’à l’ordre $M = 5$ avec les 50 capsules.

6.3.2 Support sphérique rigide

Le support sphérique sur lequel sont disposés les capsules est réalisé en plastique à l’aide d’une imprimante 3D. Il s’agit d’une boule creuse d’une épaisseur de 1 cm environ réalisée en deux parties. Cette configuration donne assez d’épaisseur pour supposer la sphère acoustiquement rigide, et laisse suffisamment d’espace pour disposer les circuits électroniques

nécessaires au conditionnement des signaux issus des capsules. L'impression 3D permet une grande précision dans le positionnement des capsules microphoniques. Une rainure cylindrique est réalisée autour de chacun des emplacements des capsules microphoniques. Ceci permet d'insérer un calibre de type piston-phone et autorise ainsi la calibration des capsules lors de mesures expérimentales. Finalement, un insert métallique taraudé sur la demi-boule inférieure permet de visser un pied de microphone. À la figure [6.10](#), sont montrés quelques détails du support en plastique.

6.3.3 Capsules microphoniques

Les capsules microphoniques sont réalisées à l'aide de 4 microphones MEMS SPH0611LR5H-1¹³ électriquement montés en parallèle. D'après les données du constructeur, le rapport signal à bruit d'un microphone est de -65 dB à 94 dB SPL, 1 kHz. Le fait de mettre en parallèle 4 microphones permet d'augmenter ce rapport de -6 dB ($\simeq 10 \log(4)$), donnant ainsi un rapport signal à bruit à -71 dB pour une capsule microphonique. Finalement, les capsules microphoniques sont câblées par 10 sur une carte électronique insérée au centre de la boule. La figure [6.11](#) présente quelques détails des capsules microphoniques et de leurs connexions.

6.3.4 Bornier de connexions

Les 50 signaux analogiques issus des capsules microphoniques sont conditionnés dans 5 câbles véhiculant chacun 10 signaux. Ces câbles sont reliés sur un bornier de connexion réalisé en impression 3D. Les signaux peuvent être prélevés à l'aide de connecteurs BNC pour des câbles coaxiaux. Une alimentation de 4 mA par capsule est nécessaire. Cette dernière est incluse dans le bornier qui se branche sur le secteur. Une photo du bornier est montrée sur la figure [6.12](#). Dans une utilisation avec la SpherBedev 2.0, les signaux des microphones sont numérisés à l'aide du bâti de convertisseurs BEHRINGER ADA8200 (voir section [6.2.4](#)). À l'aide de la suite logicielle *ambitools*, il est alors possible d'utiliser l'antenne MemsBedev pour réaliser la captation d'un champ sonore, l'encodage et la restitution dans la SpherBedev 2.0 en temps-réel.

Finalement, une vue d'ensemble du microphone monté sur un pied est visible sur la figure [6.13](#).

¹³Fiche technique : <http://www.sekisushai.net/THESE/MemsBedev/SPH0611LR5H-1-RevA.pdf>



Figure 6.10 Support en plastique de l'antenne MemsBedev réalisé en impression 3D. (a) Vue en transparence sous SolidWorks (b) Vue de la demi-boule inférieure avec 4 circuits électroniques matérialisés par des plaques vertes (c) Demi-boule inférieure partiellement câblée. (d) Demi-boule inférieure câblée et montée sur un pied de microphone.

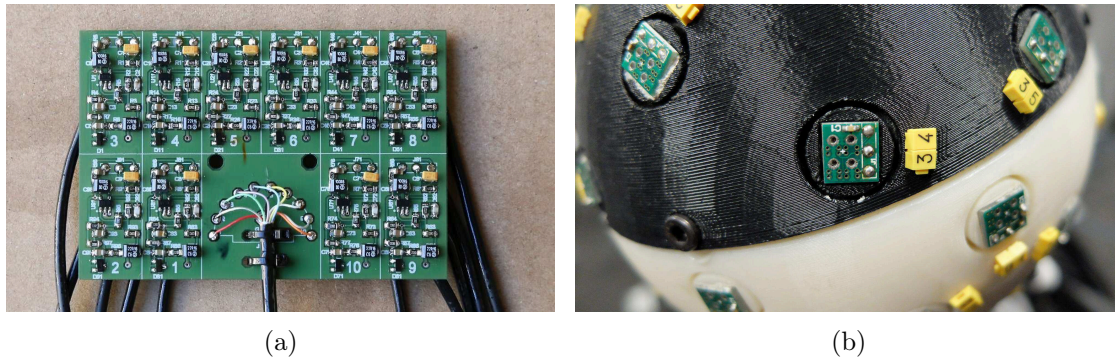


Figure 6.11 Capsules microphoniques du MemsBedev. (a) Circuit électronique pour 10 capsules microphoniques. (b) Détail d'une capsule microphonique montée sur le support sphérique rigide.

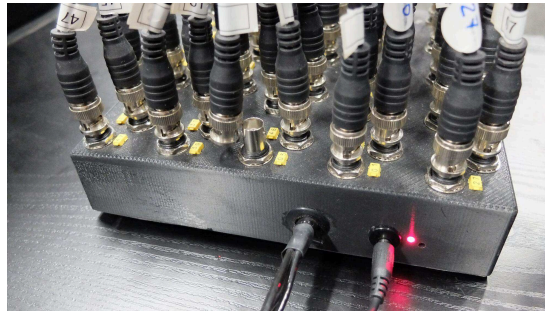


Figure 6.12 Bornier de connexion du MemsBedev. Les signaux sont prélevés sur le dessus du bornier à l'aide de câbles coaxiaux. Une LED rouge indique si l'ensemble des capsules microphoniques est alimenté et opérationnel.

6.3.5 Estimation des composantes ambisoniques

L'estimation des composantes ambisoniques à l'aide de l'antenne MemsBedev est possible sur une large bande fréquentielle. Cependant, deux contraintes limitent l'estimation en basses et hautes fréquences.

Limitation en basses fréquences - régularisation de Tikhonov

En basses fréquences, les filtres théoriques d'égalisation, présentés à l'équation (2.28), présentent un gain excessif aux ordres élevés pour être appliqués sur des signaux mesurés. En effet, le champ sonore est dominé par les premiers ordres en basses fréquences et l'estimation des ordres supérieurs nécessite alors d'amplifier excessivement les composantes issues de la TFSD du champ de pression capté par le MemsBedev. Cette amplification est d'autant plus importante que la fréquence est basse et le rayon du microphone petit. Le problème d'estimation à l'aide des filtres de l'équation (2.28) est alors mal conditionné en basses fréquences [Moreau *et al.*, 2006]. Dans un contexte de mesure, les signaux sont



Figure 6.13 MemsBedev installé sur un pied de micro.

inévitablement bruités, et une amplification excessive donne une estimation erronée des composantes d'ordres élevés en basses fréquences. Pour remédier à ce problème, un choix possible est l'utilisation d'une régularisation de Tikhonov, comme suggérée par Moreau [Moreau *et al.*, 2006], avec des filtres d'égalisation régularisés donnés par :

$$E_{m,\lambda} = \frac{(1/E_m(kr_{\text{mic}}))^*}{|1/E_m(kr_{\text{mic}})|^2 + \lambda^2} \quad (6.1)$$

Le paramètre de régularisation λ est relié au facteur d'amplification maximal des filtres $E_m(kr_{\text{mic}})$ dénoté g :

$$\lambda = \frac{1 - \sqrt{1 - 1/g^2}}{1 + \sqrt{1 - 1/g^2}} \quad (6.2)$$

où $g \geq 1$. Ce facteur peut être choisi identique pour tous les ordres m , en se basant sur le niveau d'amplification du bruit des microphones du MemsBedev et de leur nombre, par exemple [Moreau *et al.*, 2006]. En effet, en utilisant plusieurs capsules microphoniques le rapport signal à bruit peut être amélioré théoriquement de $10 \log(L)$, où L est le nombre de capsules utilisées. Ainsi, en dénotant g_s le niveau d'amplification du bruit de fond

acceptable, le facteur d'amplification maximal g est donné par :

$$g = \sqrt{L} 10^{\frac{g_s}{20}} \quad (6.3)$$

Dans le cas de l'antenne MemsBedev, $L = 50$. Le facteur d'amplification maximal du bruit est choisi par méthode essai-erreur, et vaut $g_s = +3$ dB. Le paramètre de régularisation correspondant est de $\lambda = 0,00251857$. La figure 6.14 présente le gain des filtres d'égalisation utilisés dans le cadre de la campagne de mesure présentée à la section 6.4. Ces filtres

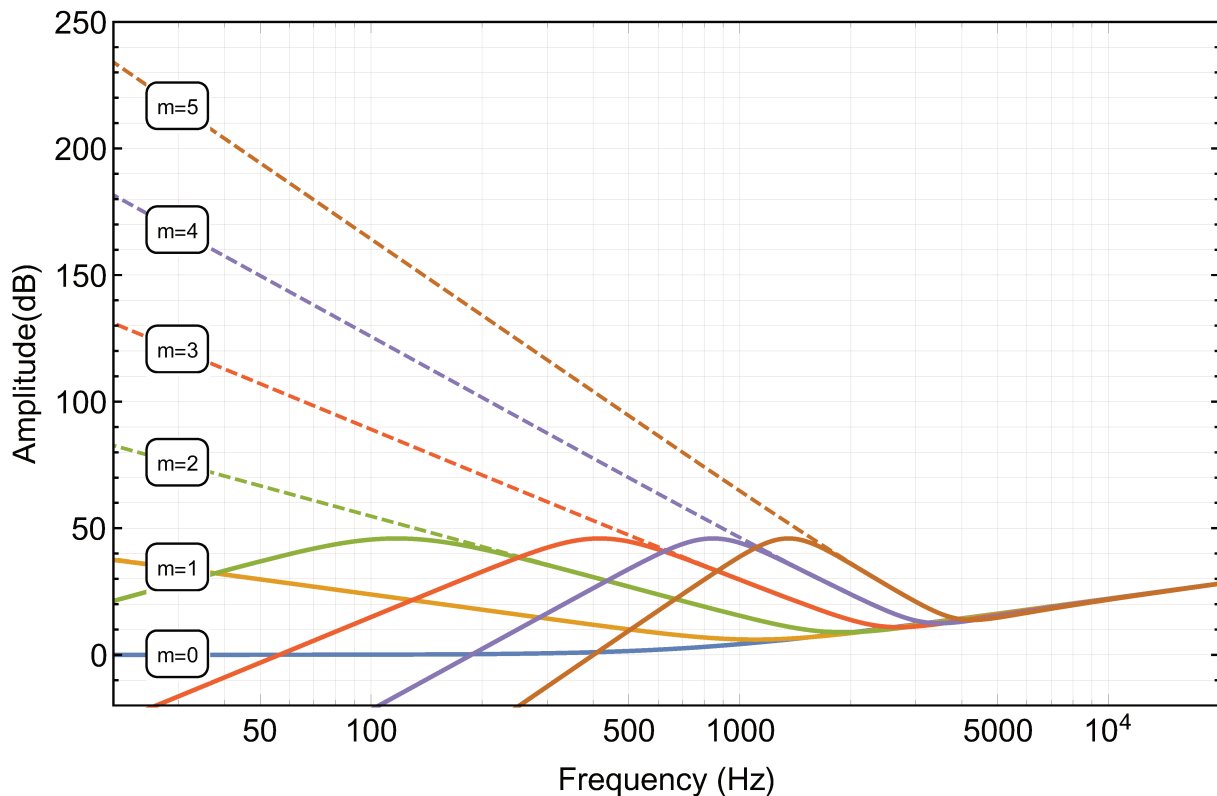


Figure 6.14 Gain des filtres d'égalisation de l'antenne MemsBedev de rayon $r_{\text{mic}} = 7$ cm. Filtres régularisés en traits pleins avec un paramètre de régularisation $\lambda = 0,00251857$. Filtres théoriques en pointillés.

sont implémentés dans la suite logicielle *ambitools* sous la forme de filtres RIF (Réponse Impulsionnelle Finie) de 4096 échantillons. Ils sont normalisés de telle sorte que le gain maximal soit de 0 dB. La convolution en temps-réel est faite à l'aide du logiciel *Jconvolver* sous Linux. Ainsi, il est possible d'estimer les composantes ambisoniques en temps-réel en filtrant les signaux en sortie de l'outil `hoa_mic_encoder_lebedev50`, qui réalise la TFSD des signaux de l'antenne (voir l'annexe D).

A noter que les filtres d'égalisation étant régularisés, l'estimation des composantes en basses fréquences ne se fait pas jusqu'à l'ordre $M = 5$ pour l'antenne MemsBedev. Ainsi

la zone de reconstruction théorique donnée par l'équation (2.32) est plus petite en basses fréquences. D'après le gabarit des filtres régularisés utilisés, l'estimation à l'ordre $M = 5$ est complète à partir de 1600 Hz environ.

Limitation hautes fréquences

L'estimation des composantes ambisoniques en hautes fréquences est limitée par la fréquence de repliement de l'antenne MemsBedev. Cette fréquence de repliement est induite par l'échantillonnage du champ de pression à la surface du micro avec un nombre fini de capsules. Ainsi, la TFS du champ de pression est approximée par la TFSD et une fréquence de repliement apparaît, au delà de laquelle les composantes ambisoniques aux ordres élevés du champ de pression se replient sur les composantes aux ordres plus faibles, suivant l'équation (3.44). Plus de détails à ce sujet sont donnés au chapitre 3. La fréquence de repliement théorique se proche de $f_{\text{alias}} \simeq \frac{Mc}{2\pi r_{\text{mic}}}$ [Rafaely, 2005], et pour le cas du microphone MemsBedev elle est estimée à $f_{\text{alias}} = 3890$ Hz.

6.4 Campagne de mesure en chambre anéchoïque

6.4.1 Objectifs

Une campagne de mesure en chambre anéchoïque est effectuée avec la SpherBedev 2.0 et l'antenne MemsBedev, présentées dans les sections précédentes. Cette campagne a notamment fait l'objet d'un stage de master [Dupont, 2016].

Le premier objectif de cette campagne est de produire des cartes de champs de pression synthétisés par approches ambisoniques et de les comparer qualitativement avec des champs de pression de sources réelles. En procédant ainsi, nous fournissons la première évaluation expérimentale et physique de la performance de la synthèse de champ sonores par ambisonie avec grille de Lebedev. C'est, en fait, une validation de la théorie présentée à titre de contribution originale au chapitre 3. Pour ce faire, la SpherBedev 2.0 est utilisée comme un réseau de haut-parleurs, et les signaux de commandes des haut-parleurs sont calculés à l'aide de la suite logicielle *ambitools* (voir annexes C et D). Deux situations sont envisagés :

- Simulation d'une source ponctuelle à la position de l'enceinte 3 de la SpherBedev 2.0 et comparaison avec le champ de pression généré par cette dernière.
- Simulation d'une source ponctuelle à la position de l'enceinte 51. Cette enceinte est positionnée à l'extérieur de la SpherBedev et ne fait pas partie de cette dernière.

Comparaison du champ synthétisé avec le champ de pression généré par l'enceinte 51.

Une antenne plane de microphones est aussi utilisée pour cartographier les champs de pression générés dans le plan horizontal autour du centre de la SpherBedev [6.4.3](#). Une bonne reproduction qualitative validera l'aptitude de synthèse de champ sonore expérimentale à l'aide du réseau de haut-parleurs SpherBedev et de la suite logicielle *ambitools*.

Un autre objectif est la validation de l'encodage des champs sonores à l'aide de l'antenne MemsBedev. Ainsi, en complément de l'antenne de géométrie plane, l'antenne MemsBedev est utilisée au centre de la SpherBedev pour la captation et l'encodage des champs sonores suivant le formalisme ambisonique. Les champs sonores encodés sont ensuite décodés sur la SpherBedev et comparés avec les champs sonores initiaux.

Finalement, un troisième objectif est la validation expérimentale du filtrage directionnel présenté au chapitre [5](#). Pour ce faire, un champ sonore est capté par l'antenne MemsBedev, puis filtré dans le domaine des harmoniques sphériques après encodage et avant décodage. Le champ sonore original capté par l'antenne MemsBedev est créé en faisant fonctionner simultanément deux enceintes de la SpherBedev qui sont dans des directions perpendiculaires : les enceintes 2 et 3. La capture ambisonique est alors filtrée après l'encodage par une fonction de directivité hyper-cardioïde d'ordre $M' = 1$ pointant dans la direction de l'enceinte 2. Le champ sonore après filtrage est finalement décodé sur la SpherBedev puis re-capté par les deux antennes plane et MemsBedev.

Il est important de noter que le système de reproduction ambisonique n'est pas égalisé lors de cette campagne. Ainsi, bien que les logiciels utilisés se basent sur un modèle parfait d'une grille de Lebedev à 50 monopoles en champ libre, le système réel s'écarte de ce modèle en plusieurs points. La position des haut-parleurs n'est pas parfaite et leurs réponses en fréquence et directivité ne sont pas monopolaires. De plus, la structure de la SpherBedev, le plancher, le bâti d'amplificateurs et les supports des microphones, peuvent avoir une influence acoustique qui interfère à la bonne reconstruction du champ sonore (réflexions, diffraction, réfraction). Pour une reproduction précise, une égalisation est envisagée pour le futur, par la méthode présentée au chapitre [4](#). Ainsi, cette campagne de mesure cherche avant tout à produire des résultats qualitatifs sur la reproduction de champ sonores et non quantitatif à ce stade.

Cette section s'organise de la manière suivante. Dans un premier temps, la section [6.4.2](#) détaille le protocole expérimental de la campagne de mesures. Puis, la section [6.4.3](#) présente l'antenne plane de microphones utilisée en complément de l'antenne MemsBedev. Les

signaux d'excitation utilisés pour les haut-parleurs de la SpherBedev sont détaillés pour chaque mesure à la section [6.4.4](#). La section [6.4.5](#) détaille le positionnement des antennes de microphones au centre de la SpherBedev et la section [6.4.6](#) la calibration de ces dernières. La section [6.4.7](#) présente les opérations de post-traitement des signaux de mesures et les sections [6.4.8](#) et [6.4.9](#) les résultats de mesures ainsi que leurs interprétations.

6.4.2 Protocole expérimental

Le protocole expérimental est détaillé ci-après. Il détaille les mesures faites à la fois sur l'antenne MemsBedev et sur l'antenne plane présentée à la section [6.4.3](#). Les mesures sont récapitulées dans le tableau [6.3](#). Avant de démarrer la campagne de mesure, la SpherBedev 2.0 est installée au centre de la chambre anéchoïque du GAUS (voir figure [6.15](#)). De la mousse absorbante est utilisée pour couvrir le bâti d'amplificateurs et de convertisseurs, ainsi que la passerelle sur laquelle est posée la SpherBedev afin de limiter l'influence acoustiques de ces éléments. Par ailleurs, une 51^{ème} enceinte, située à l'extérieur de la SpherBedev est installée à l'aide d'une barre en aluminium fixée au cadre de la SpherBedev 2.0. Les coordonnées de cette enceinte sont (2.23 m, 73°, 12.8°). Elle est visible à droite sur la figure [6.15](#). Finalement, les premières mesures sont effectuée avec l'antenne MemsBedev qui est calibrée et positionnée au centre de la SpherBedev 2.0 (voir les sections [6.4.6](#) et [6.4.5](#)). Juste avant de réaliser la mesure 9 (voir tableau [6.3](#)), l'antenne MemsBedev est remplacée par l'antenne plane qui est calibrée et positionnée au centre de la SpherBedev 2.0 (voir les sections [6.4.6](#) et [6.4.5](#)).

6.4.3 Antenne plane de microphones

Configuration géométrique

Une antenne plane de microphones est réalisée pour une utilisation complémentaire de l'antenne MemsBedev. L'objectif est de cartographier les champs de pression dans le plan horizontal autour du centre de la SpherBedev, et de comparer les résultats avec l'extrapolation dans le plan horizontal du champ sonore capté sur le MemsBedev. Cinquante-cinq microphones utilisant des capsules électret WN61A de 6 mm de diamètre dans un châssis cylindrique en aluminium sont montés dans un cadre en bois suivant une géométrie rectangulaire. Un pas de 2.5 po ($\simeq 6,35$ cm) est choisi dans les deux directions x et y . L'antenne possède 6 lignes de 8 microphones et une ligne de 7 microphones. (la 58^{ème} voie du bâti de convertisseurs est réservée à un signal de synchronisation pour estimer la latence de

Tableau 6.3 Détails des mesures du protocole expérimental.

Mesure	Description du champ sonore	Antenne	Objectif
1	Enceinte 3 (1,07 m, 90°, 0°) de la SpherBedev.	MemsBedev	Reproduction de champs sonores : champ sonore cible.
10		Plane	
2	Encodage de la mesure 1 à l'ordre $M = 5$ et décodage sur la SpherBedev.	MemsBedev	Reproduction de champs sonores : champ sonore reproduit après encodage et décodage d'une captation ambisonique.
11		Plane	
3	Source virtuelle à l'ordre $M = 5$ à la position de l'enceinte 3 (1,07 m, 90°, 0°).	MemsBedev	Reproduction de champs sonores : champ sonore reproduit pour une source virtuelle.
12		Plane	
4	Enceinte 51 (2.23 m, 73°, 12.8°) à l'extérieur de la SpherBedev	MemsBedev	Reproduction de champs sonores : champ sonore cible.
13		Plane	
5	Encodage de la mesure 4 à l'ordre $M = 5$ et décodage sur la SpherBedev.	MemsBedev	Reproduction de champs sonores : champ sonore reproduit après encodage et décodage d'une captation ambisonique.
14		Plane	
6	Source virtuelle à l'ordre $M = 5$ à la position de l'enceinte 51 (2.23 m, 73°, 12.8°).	MemsBedev	Reproduction de champs sonores : champ sonore reproduit pour une source virtuelle.
15		Plane	
7	Enceinte 2 (1,07 m, 0°, 0°) et enceinte 3 (1,07 m, 90°, 0°) alimentées simultanément.	MemsBedev	Filtrage directionnel : champ sonore complexe.
8	Encodage de la mesure 7 à l'ordre $M = 4$ et décodage sur la SpherBedev.	MemsBedev	Filtrage directionnel : champ sonore reproduit après encodage et décodage d'une captation ambisonique.
16		Plane	
9	Encodage de la mesure 7 à l'ordre $M = 4$. Filtrage directionnel dans le domaine des harmoniques sphérique : utilisation d'une fonction de directivité hyper-cardioïde d'ordre $M' = 1$ dont l'angle de pointage est fixé à (0°, 0°). Décodage du champ sonore filtré sur la SpherBedev.	MemsBedev	Filtrage directionnel : champ sonore reproduit après encodage d'une captation ambisonique, filtrage directionnel, et décodage.
17		Plane	

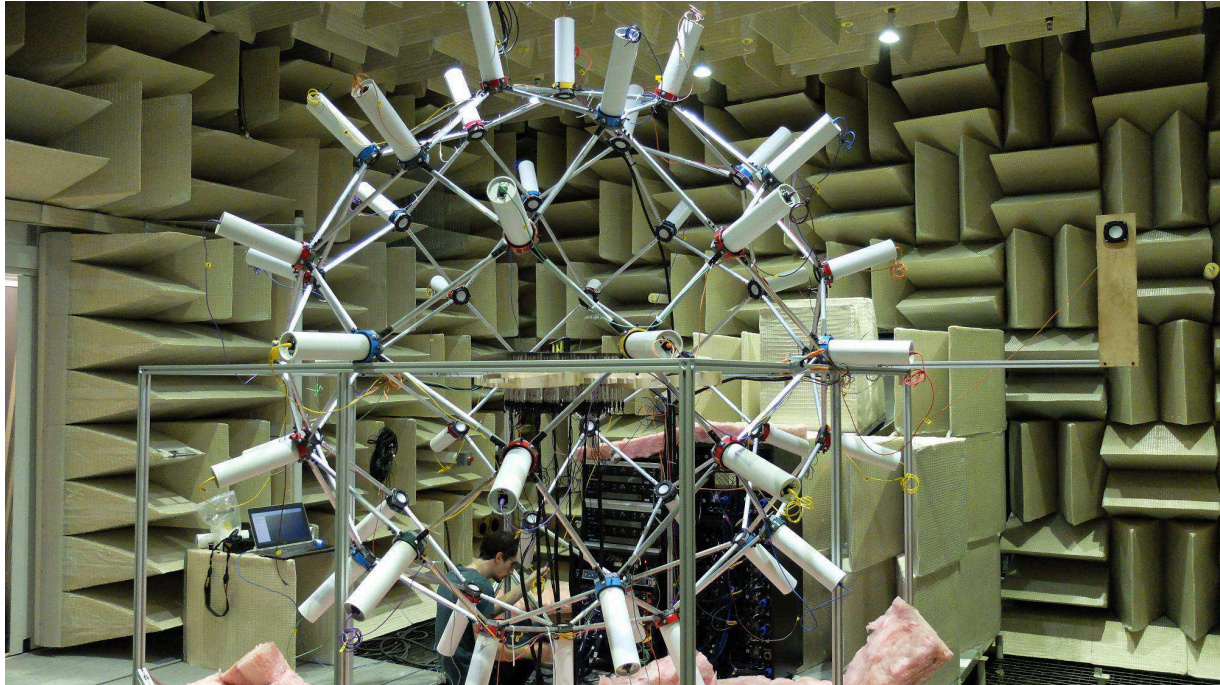


Figure 6.15 Dispositif expérimental installé dans la chambre anéchoïque du GAUS. L'antenne plane est positionnée au centre de la SpherBedev 2.0. Une 51^{ème} source est positionnée à l'extérieur de la sphère aux coordonnées (2.23 m, 73°, 12.8°) (à droite sur l'image).

la conversion numérique analogique). Le microphone 29 est à l'origine O du repère. La géométrie résultante est visible sur la figure [6.16\(b\)](#).

Connectiques

Les microphones de l'antenne possèdent une connectique RCA et sont regroupés par 8 dans un câble faisceau avec un connecteur DB-9 visible au premier plan de la figure [6.16\(a\)](#). Des préamplificateurs réalisés au GAUS, visibles à l'arrière plan de la figure [6.16\(a\)](#), permettent de conditionner les signaux avant de les numériser à l'aide des convertisseurs BEHRINGER ADA8200 (voir section [6.2.4](#)).

Fréquence de repliement

Le pas de l'antenne conditionne la fréquence de repliement de cette dernière. En effet, l'échantillonnage spatial introduit une périodicité dans la transformée en nombre d'ondes. D'après le théorème d'échantillonnage de Shannon, il est nécessaire d'avoir une période d'échantillonnage spatiale inférieure à la moitié de la longueur d'onde observée pour éviter le repliement. Ainsi la fréquence de repliement pour l'antenne de pas 6,35 cm est de 2695 Hz environ.

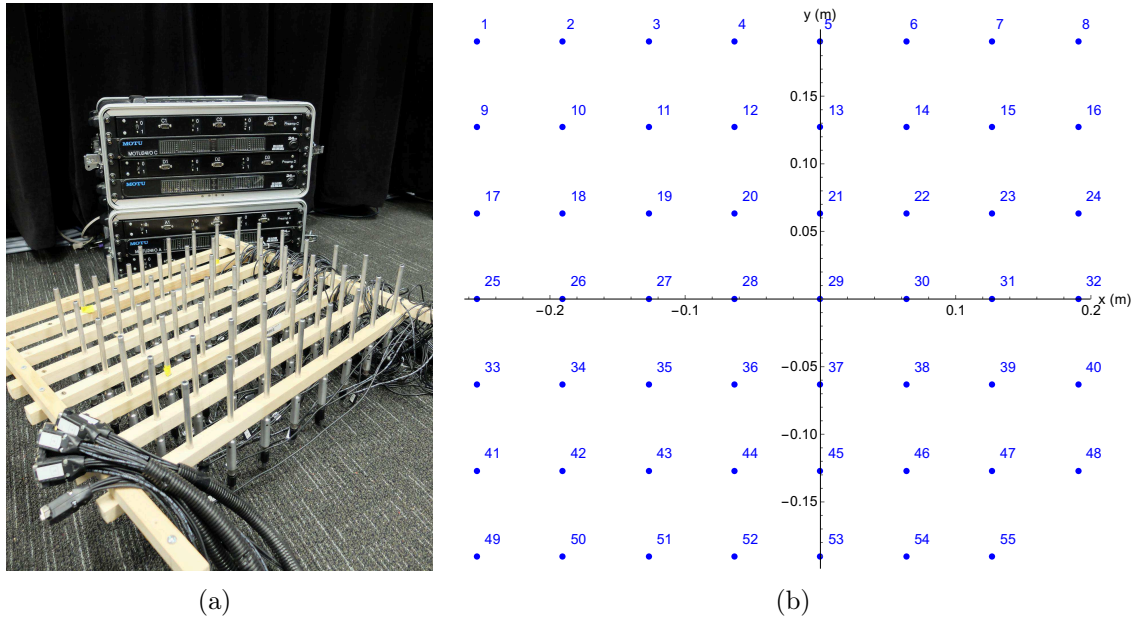


Figure 6.16 Configuration de l’antenne plane de microphones. (a) Antenne plane de microphones avec les connectiques au premier plan et les pré-amplificateurs en arrière plan. (b) Coordonnées et numérotation des microphones de l’antenne.

6.4.4 Signaux d’excitation

Pour toutes les mesures effectuées lors de cette campagne, la fréquence d’échantillonnage est fixée à $F_e = 48$ kHz, et le même signal d’excitation est utilisé pour estimer les FRFs. Il s’agit d’un sinus glissant logarithmique [Farina, 2000] allant de 0.01 Hz à 24 kHz d’une durée de 0.45 s. Pour maximiser le rapport signal à bruit, le signal d’excitation est répété 10 fois et une moyenne est opérée sur les mesures dans le domaine fréquentiel. Ce signal est pré-enregistré dans une session de la station de travail audio-numérique Ardour¹⁴ sous Linux/JACK. Dépendant du type de mesure, le signal est éventuellement traité en temps-réel par les outils de la suite logicielle *ambitools*, avant d’être transmis aux 50 haut-parleurs de la SpherBedev.

Reproduction de champs sonores : mesures 1, 4, 9 et 12. Pour ces mesures, le signal est directement transmis au haut-parleur.

Reproduction de champs sonores : mesures 2, 5, 10 et 13. les 50 signaux mesurés par le MemsBedev lors de la mesure 1 et 5 sont matricés à l’aide de l’outil

¹⁴<https://ardour.org/>

`hoa_mic_encoder_lebedev50` pour obtenir la TFS du champ de pression capté. Les 36 composantes obtenues sont ensuite filtrées à l'aide de `Jconvolver` suivant les filtres décrits à la section [6.3.5](#) pour obtenir les signaux ambisoniques. Enfin, ces signaux ambisoniques sont décodés à l'aide du décodeur `hoa_decoder_lebedev50` avant d'être transmis aux haut-parleurs.

Reproduction de champs sonores : mesures 3, 6, 11 et 14. Le signal d'excitation est envoyé en entrée de l'outil

`hoa_panning_lebedev50` qui calcule les signaux des haut-parleurs suivant l'équation [\(2.48\)](#).

Filtrage directionnel : mesure 7. Le signal d'excitation est envoyé simultanément aux haut-parleurs 2 et 3.

Filtrage directionnel : mesure 8. et 15. Les 50 signaux mesurés par le `MemsBedev` lors de la mesure 7 sont matricés à l'aide de l'outil `hoa_mic_encoder_lebedev50` pour obtenir la TFS du champ de pression capté à l'ordre $M = 4$. Les 25 composantes obtenues sont ensuite filtrées à l'aide de `Jconvolver` suivant les filtres décrits à la section [6.3.5](#) pour obtenir les signaux ambisoniques. Enfin, ces signaux ambisoniques sont décodés à l'aide du décodeur `hoa_decoder_lebedev50` avant d'être transmis aux 50 haut-parleurs de la `SpherBedev`.

Filtrage directionnel : mesure 9. et 16. Les 50 signaux mesurés par le `MemsBedev` lors de la mesure 7 sont matricés à l'aide de l'outil `hoa_mic_encoder_lebedev50` pour obtenir la TFS du champ de pression capté à l'ordre $M = 4$. Les 25 composantes obtenues sont ensuite filtrées à l'aide de `Jconvolver` suivant les filtres décrits à la section [6.3.5](#) pour obtenir les signaux ambisoniques. Les 25 composantes ambisoniques à l'ordre $M = 4$ sont ensuite envoyées dans l'outil de filtrage directionnel

`hoa_beamforming_hypercardioid_to_hoa`. L'outil est réglé pour utiliser une fonction de directivité hyper-cardioïde d'ordre $M' = 1$ avec un angle de pointage de $(0^\circ, 0^\circ)$. Les composantes ambisoniques filtrées sont au nombre de 36, d'après l'équation [\(5.14\)](#) du chapitre [5](#). Finalement, ces signaux ambisoniques filtrés sont décodés à l'aide du décodeur `hoa_decoder_lebedev50` avant d'être transmis aux 50 haut-parleurs de la `SpherBedev`.

6.4.5 Positionnement

Le positionnement des antennes dans la SpherBedev 2.0 est effectué à l'aide de 3 lasers. Le principe est le suivant : un cylindre en aluminium permet de centrer un laser sur un des anneaux de la SpherBedev 2.0. À l'aide des trois vis à 120° , le centrage est réglé de telle sorte que le laser vise le centre de la membrane du haut-parleur diamétralement opposé. Cette opération est répétée pour les trois directions x , y et z à l'aide des noeuds 2 et 4, 3 et 5, 1 et 6 respectivement. Il est ainsi obtenu une visualisation approximative (aux erreurs de visée, centrage, et dimensions de la sphère) des trois directions x , y et z de la SpherBedev 2.0.

Antenne plane. L'antenne plane est placée sur un support en bois fixé au cadre de la SpherBedev. Elle est ensuite positionnée en faisant affleurer les lasers des axes x et y avec les rangées de microphones correspondantes. Le microphone 29 est placé au croisement des deux directions, c'est à dire à l'origine O du repère de la SpherBedev2.0.

Antenne MemsBedev. L'antenne MemsBedev possède le même maillage de Lebedev que la SpherBedev 2.0. Elle est alors placée de telle sorte que le laser au noeud 2 de la SpherBedev vise le centre de la capsule 2 du MemBedev. De même, le laser au noeud 3 vise le centre de la capsule 3 et le laser au noeud 6 le centre de la capsule 6. La figure [6.17](#) illustre le positionnement des deux réseaux de microphone utilisés.

A noter que le positionnement précis de l'antenne MemsBedev est important pour l'appréciation des résultats, dans le cas d'une captation encodée puis décodée sur la SpherBedev. En effet, le positionnement du champ sonore est décrit par rapport au repère de l'antenne MemsBedev lorsque la captation est encodée au format ambisonique. Cependant, le décodage repositionne le champ sonore par rapport au repère de la SpherBedev. Ainsi, si les deux repères ne sont pas parfaitement alignés, le champ sonore décodé sur la SpherBedev aura subi les transformations d'un repère à l'autre lorsqu'il est de nouveau capté par l'antenne MemsBedev. Pour bien comprendre ceci, considérons seulement une erreur de rotation dans l'alignement des repères : le repère de l'antenne MemsBedev est tourné d'un angle d'azimut de $+\alpha$ par rapport au repère de la SpherBedev. Une source positionnée dans la direction $(0^\circ, 0^\circ)$ dans le repère de la SpherBedev sera encodé à $(-\alpha, 0^\circ)$ dans le repère de l'antenne MemsBedev. Au décodage, cette source sera reproduite à la direction $(-\alpha, 0^\circ)$ dans le repère de la SpherBedev et re-captée à la direction $(-2\alpha, 0^\circ)$ dans le repère de l'antenne MemsBedev. Finalement, en comparant les résultats de la captation initiale

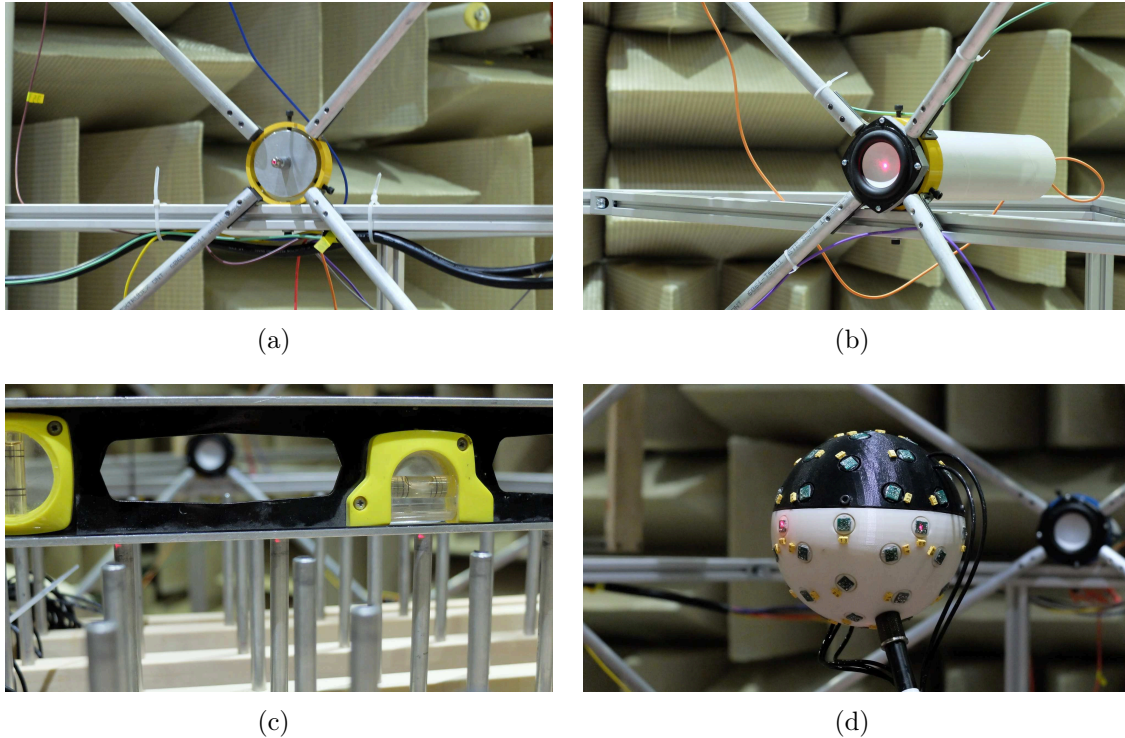


Figure 6.17 Positionnement des antennes de microphones dans la SpherBedev 2.0. (a) Un laser dans sa pièce cylindrique, centré dans un des anneaux de la SpherBedev 2.0. (b) Le laser vise le centre de la membrane du haut-parleur diamétralement opposé. (c) Positionnement de l'antenne plane : le laser affleure une rangée de microphone. (d) Positionnement du MemsBedev : les trois lasers visent le centre des trois capsules en vis-à-vis.

de l'antenne MemsBedev avec la captation du champ décodé, une rotation d'azimut $-\alpha$ aura été appliquée.

6.4.6 Calibration

L'antenne MemsBedev ainsi que l'antenne plane sont calibrées en utilisant un calibre de type piston-phone émettant 1 Pa à 1 kHz. La FRF de chacun des microphones est supposée plate dans la bande de fréquence d'intérêt. Ainsi, la valeur de calibration obtenue à 1 kHz est utilisée pour toutes les autres fréquences. La figure 6.18 montre le piston-phone en place sur un micro pour les deux types de réseau, au cours de la calibration. Le signal de calibration enregistré par les microphones est filtré à l'aide d'une filtre passe-bande centré à 1 KHz. La valeur RMS (Root Mean Square) du signal est ensuite extraite pour donner la facteur de calibration en Pa/UN où UN signifie Unité Numérique.

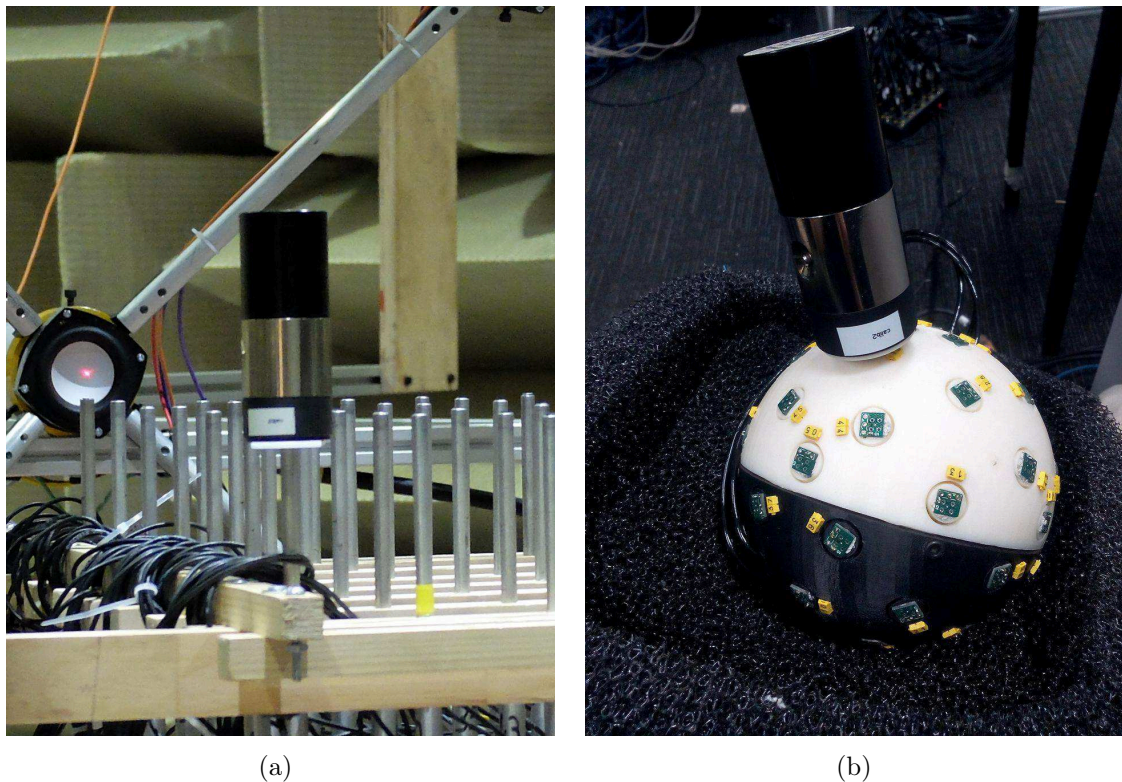


Figure 6.18 Calibration des réseaux de microphones. (a) Antenne plane. (b) Antenne sphérique MemsBedev.

6.4.7 Post-traitement des signaux de mesures

Toute la campagne de mesure utilise la station de travail audio-numérique Ardour¹⁵ sous Linux/JACK pour l'acquisition des signaux. Une fois les signaux enregistrés, ils sont exportés et traités sous MATLABTM ou MathematicaTM.

Calibration

Les signaux étant exprimés en UN, une première étape consiste à appliquer les valeurs de sensibilités des microphones mesurées à la section 6.4.2 pour obtenir les signaux en Pa.

Réponses impulsionnelles

Les réponses impulsionnelles sont systématiquement calculées entre les signaux microphones et le signal d'excitation présenté à la section 6.4.4, en utilisant la méthode proposée par Farina [Farina, 2000]. Elles sont tronquées à l'aide d'une demi fenêtre de Tukey à

¹⁵<https://ardour.org/>

$2L_w = 16384$ échantillons. La fenêtre de Tukey utilisée est définie par l'équation suivante :

$$w(n) = \begin{cases} 1 & 0 \leq n < \frac{L_w}{2} \\ \sin\left(\frac{2\pi n}{L_w}\right)^2 & \frac{L_w}{2} \leq n \leq L_w \\ 0 & n > L_w \end{cases} \quad (6.4)$$

avec $n \in \mathbb{N}$.

A ce stade, il est observé que les réponses impulsionnelles mesurées par l'antenne Mem-sBedev sont en opposition de phase par rapport à celles mesurées pour l'antenne plane, pour un même champ sonore. Ceci vient probablement du fait que l'alimentation du microphone est à polarité inversée. Il est alors choisi à ce stade d'inverser la polarité des réponses impulsionnelles de l'antenne Mem-sBedev.

Ajustement du retard des réponses impulsionnelles

Le système audio-numérique possède une latence de traitement des signaux et de conversion analogique-numérique. De plus, pour la synthèse de source virtuelle, le retard de propagation relatif entre la position de la source virtuelle et les haut-parleurs de restitution n'est pas implémenté dans *ambitools* [Lecomte et Gauthier, 2015]. Ainsi, les champs de pressions synthétisés ne sont pas en phase avec les champs cibles. En conséquence, pour apprécier au mieux les cartes de champs de pression et les comparer qualitativement, un retard global, identique sur toutes les réponses impulsionnelles aux microphones, est introduit pour aligner les champs de pression cible et reproduits.

Le retard entre la réponse impulsionnelle cible d'un microphone et la réponse impulsionnelle reproduite au même microphone est estimé en utilisant la corrélation croisée [Jacovitti et Scarano, 1993].

Antenne plane. Pour l'antenne plane, il est choisi de prendre comme retard global le retard estimé pour le microphone 29 qui se situe à l'origine O . En effet, la synthèse par Ambisonie est décrite par rapport à l'origine, et ce point est le point où le champ sonore est le mieux reproduit sur la bande fréquentielle la plus large.

Antenne Mem-sbedev. Pour l'antenne Mem-sBedev, le retard entre la réponse impulsionnelle cible et reproduite est estimé pour chacune des capsules de l'antenne. Finale-

ment, le retard global est obtenu en prenant la moyenne arithmétique des retards estimés à chaque microphone.

Pour recaler les réponses impulsionnelles mesurées avec l'antenne MemsBedev et celles mesurées avec l'antenne plane, on estime le retard entre le microphone 29, de l'antenne plane (à l'origine O) et la réponse impulsionnelle d'ordre 0 (pression à l'origine O) encodée avec l'antenne MemsBedev.

Affichage des champs

Antenne plane. Pour l'antenne plane, l'affichage des champs de pressions à une fréquence est directement fait à partir des FRFs à chaque microphone.

Antenne MemsBedev. En revanche, pour l'antenne sphérique MemsBedev, l'affichage des champs de pression requiert d'encoder dans un premier temps les FRFs des capsules du microphone, puis d'extrapoler le champ mesuré à l'aide d'une série de Fourier-Bessel, selon l'équation (2.18). Les filtres d'égalisation utilisés pour les microphones sont ceux décrits à la section 6.3.5.

Calcul de l'intensité acoustique

Pour toutes les figures présentées ci-après, un champ de vecteur d'intensité acoustique dans le plan horizontal est calculé à chaque fréquence. L'intensité acoustique est un indicateur du flux d'énergie acoustique et de la direction de propagation du champ sonore. Elle est définie par l'équation (2.55), rappelée ici :

$$\mathbf{I}(f, \mathbf{x}) = \frac{1}{2} \Re(p(f, \mathbf{x})\mathbf{v}(f, \mathbf{x})^*). \quad (6.5)$$

Dans cette équation \mathbf{I} représente le vecteur d'intensité au point \mathbf{x} , \Re la partie réelle et \mathbf{v} le vecteur vitesse au point \mathbf{x} . Le vecteur vitesse \mathbf{v} est donné par l'équation d'Euler :

$$\mathbf{v}(f, \mathbf{x}) = -\frac{\nabla p(f, \mathbf{x})}{ik\rho_0 c}, \quad (6.6)$$

où ∇ représente le gradient, ρ_0 la densité de l'air, prise à $\rho_0 = 1.2 \text{ kg/m}^3$, et c la vitesse du son dans l'air, $c = 340 \text{ m.s}^{-1}$. Les composantes du gradient de pression sont calculées par différences finies de la pression aux positions de chaque microphone de l'antenne plane, divisée par leur distance. Ce calcul par différences finies devient de plus en plus approximatif à mesure que la longueur d'onde diminue.

A noter que le champ de pression de l'antenne MemsBedev est extrapolé à partir des composantes ambisoniques : il peut alors être calculé en tout point de l'espace. Cependant à des fins de comparaisons entre l'antenne plane et l'antenne MemsBedev, le champ de vecteur d'intensité acoustique pour le microphone MemsBedev est calculé en utilisant les valeurs de pressions aux positions de microphones de l'antenne plane.

6.4.8 Source réelles et virtuelles

Les cartes des champs de pression issues des mesures 10 à 15 (antenne plane) et 1 à 6 (antenne MemsBedev) sont présentées de la figure 6.19 à la figure 6.31 pour différentes fréquences allant de 500 Hz à 2000 Hz. Le champ de pression cible est choisi comme celui généré par le haut-parleur 3 de la SpherBedev 2.0 (mesure 1 et 10) ou celui généré par le haut-parleur 51 situé à l'extérieur de la SpherBedev 2.0 (mesure 4 et 13). Comme il est précisé à la section 6.4.7, les réponses impulsionnelles de la mesure 1 sont recalées par rapport à celles de la mesure 10. Puis, les réponses impulsionnelles des mesures 2 et 3 sont recalées par rapport à celles de la mesure 1. Enfin, les réponses impulsionnelles des mesures 11 et 12 sont recalées par rapport à celles la mesure 10. Les mêmes opérations sont appliquées pour le cas d'une source à l'extérieur de la SpherBedev 2.0.

A première vue, les champs synthétisés ainsi que les champs captés par l'antenne MemsBedev puis décodés sur la SpherBedev sont en concordance avec le champ de pression cible. En effet, les fronts d'ondes sont bien reproduits et d'après le champ de vecteur d'intensité, la direction de propagation des champs sonores reconstruits par la SpherBedev est en concordance avec le champ cible.

On remarque que la phase des fronts d'ondes semble changer avec la fréquence pour les champs cibles, ainsi que pour les champs reconstruits. Ceci peut être expliqué par la réponse en phase non-linéaire des haut-parleurs utilisés pour la SpherBedev (voir figure 6.6(b)). Cependant, bien que chacun des haut-parleurs ait sa propre réponse, qui n'est pas égalisée, la reconstruction des champs cibles est tout de même effective.

Les réponses impulsionnelles n'ont pas été égalisées en gain. Ainsi l'échelle de couleur à droite de chaque figure représente l'amplitude en Pa après calibration. Pour les sources virtuelles et les captations de l'antenne MemsBedev décodées, aucune précaution particulière n'a été prise pour ajuster le gain de la source simulée ou de l'enregistrement décodé pour respecter le gain du champ cible, puisque nous sommes ici principalement intéressé par la forme du front d'ondes. Cependant, ceci ne compromet pas l'effectivité de reproduction, qui se fait alors à un gain près. Par ailleurs, pour les mêmes champs sonores mesurés,

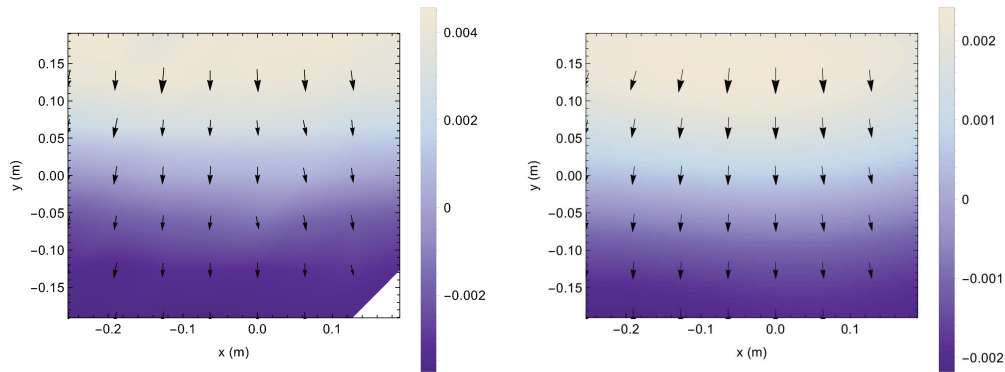
il apparaît des différences de gain entre le champ extrapolé de l’antenne MemsBedev et le champ mesuré par l’antenne plane. Ceci peut s’expliquer par plusieurs raisons. Tout d’abord, les filtres d’égalisations utilisés pour le calcul des composantes ambisoniques sont régularisés et peuvent atténuer l’amplitude du champ extrapolé obtenu. Ensuite, l’extrapolation, qui utilise une série de Fourier-Bessel tronquée à l’ordre $M = 5$, s’écarte de la réalité en dehors de la zone de reconstruction. En effet, comme il a été montré au chapitre 3 (équation (3.40)) le champ reconstruit théorique, en dehors de la zone de reconstruction est la somme du champ tronqué à l’ordre M et des termes de repliement spatial. Finalement, le dispositif expérimental lui même peut introduire des écarts par la différente géométrie des deux antennes qui a une influence acoustique différente dans les champ sonores.

Dans le cas des sources simulées aux positions des haut-parleurs 3 et 51, la différence entre les deux situations est l’activation ou non des filtres de champ proche. La source virtuelle située à la position du haut-parleur 3 est au même rayon que le rayon de la SpherBedev. Ainsi, dans l’équation (2.48) le terme $F_m(kr_s)$ traduisant le champ proche est compensé identiquement par le terme $1/F_m(kr_{\text{spk}})$. Autrement dit la synthèse d’un monopole à un rayon r_s avec un réseau de monopoles au même rayon $r_{\text{spk}} = r_s$ ne nécessite pas de compenser la courbure des fronts d’ondes. La concordance des résultats de sources virtuelles avec les sources réelles, en particulier pour la source à la position du haut-parleur 51 montre l’importance des précautions à apporter à la simulation du champ proche des sources.

Les méthodes et techniques présentées aux chapitres 2 et 3 sont alors validées expérimentalement. La mise en oeuvre en temps-réel, telle que présentée à l’annexe C et D offre la possibilité de synthétiser des sources virtuelles ou de reproduire efficacement des champs sonores captés par une antenne sphérique de microphones. La quantification de l’erreur de reproduction serait alors la prochaine étape de la validation expérimentale. Cependant, si une minimisation de l’erreur quadratique, tel que celle de l’équation (2.60), est visée, de plus grandes précautions sont à prendre pour le calcul de l’amplitude et des retards des signaux de commande des haut-parleurs. En effet, il est alors primordial que le champ reproduit soit de même amplitude et reconstruit avec un retard nul par rapport au champ cible. De plus, l’erreur quadratique est très sensible aux défauts du système, et une égalisation, telle que présentée au chapitre 4 est alors requise.

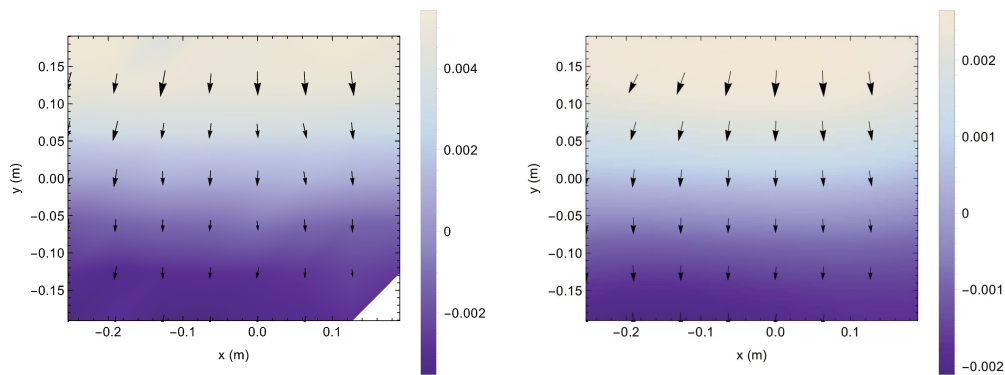
6.4.9 Champ sonore avec filtrage directionnel

Les cartes des champs de pression issues des mesures 15 à 16 (antenne plane) et 8 et 9 (antenne MemsBedev) sont présentées de la figure 6.32 à la figure 6.38 pour différentes



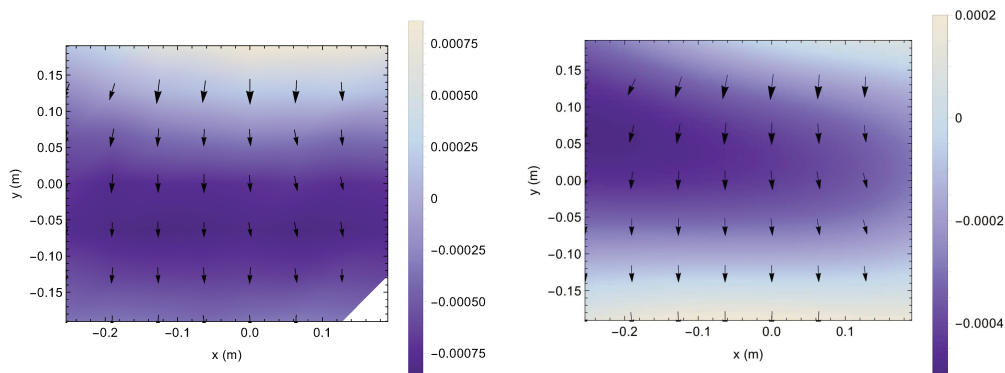
(a) Mesure 10 : antenne plane

(b) Mesure 1 : antenne MemsBedev



(c) Mesure 11 : antenne plane

(d) Mesure 2 : antenne MemsBedev



(e) Mesure 12 : antenne plane

(f) Mesure 3 : antenne MemsBedev

Figure 6.19 Partie réelle des champs de pression à 500 Hz pour les mesures 10, 11, 12 (colonne de gauche, antenne plane) et 1, 2, 3 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 3 (1,07 m, 90°, 0°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 3. (e) et (f) Haut-parleur 3 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

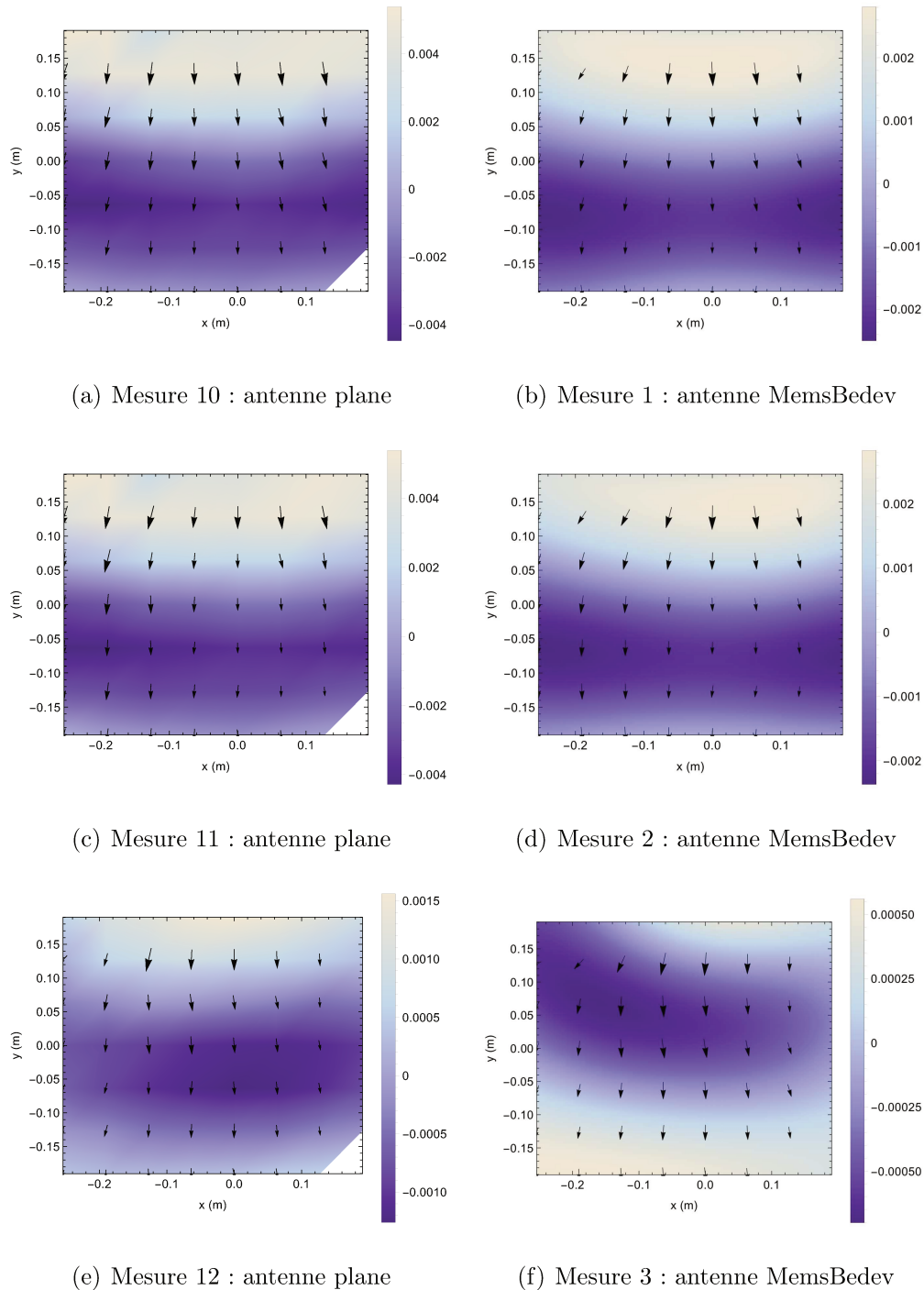


Figure 6.20 Partie réelle des champs de pression à 750 Hz pour les mesures 10, 11, 12 (colonne de gauche, antenne plane) et 1, 2, 3 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 3 (1,07 m, 90° , 0°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 3. (e) et (f) Haut-parleur 3 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

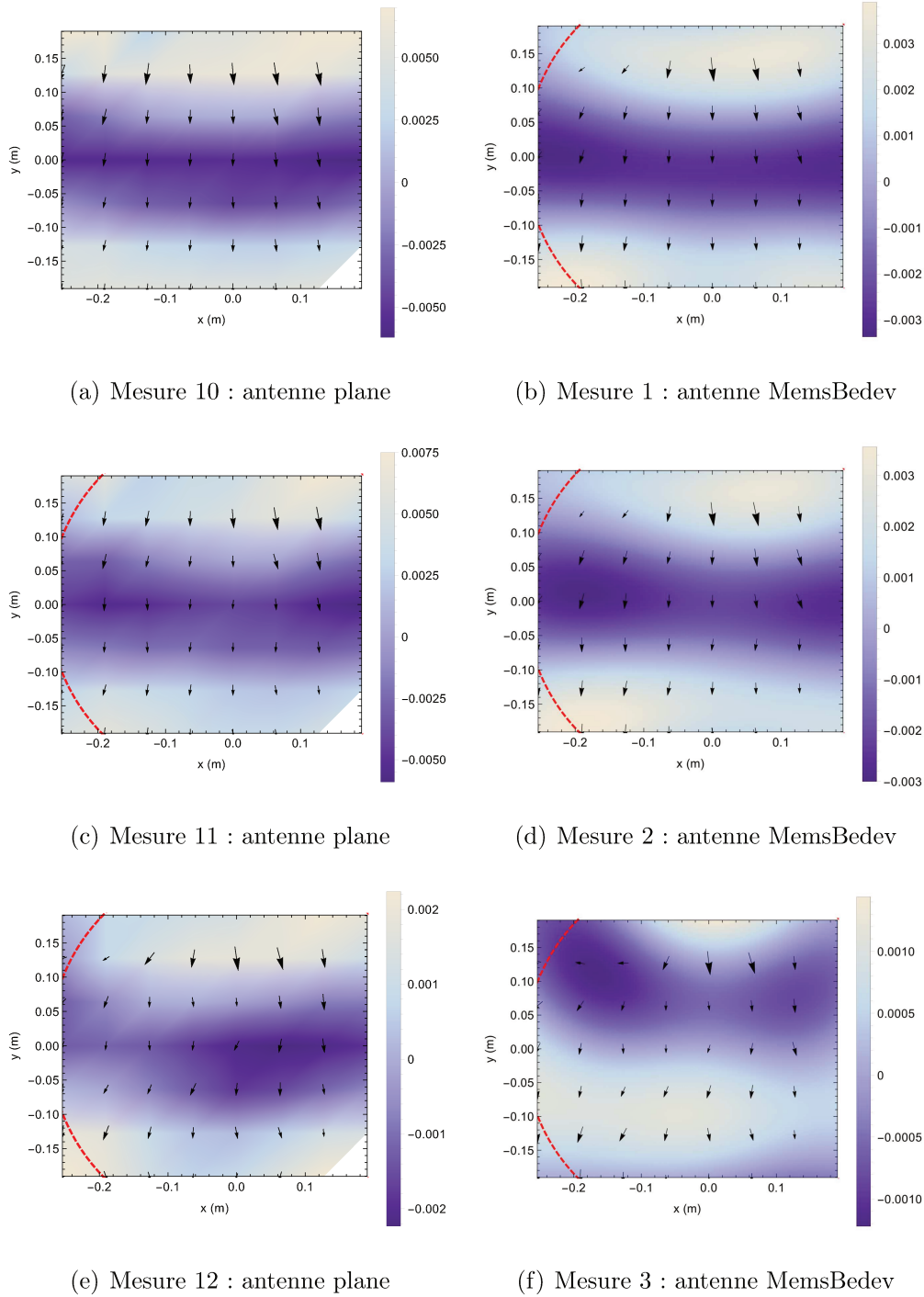


Figure 6.21 Partie réelle des champs de pression à 1000 Hz pour les mesures 10, 11, 12 (colonne de gauche, antenne plane) et 1, 2, 3 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 3 ($1,07\text{ m}$, 90° , 0°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 3. (e) et (f) Haut-parleur 3 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

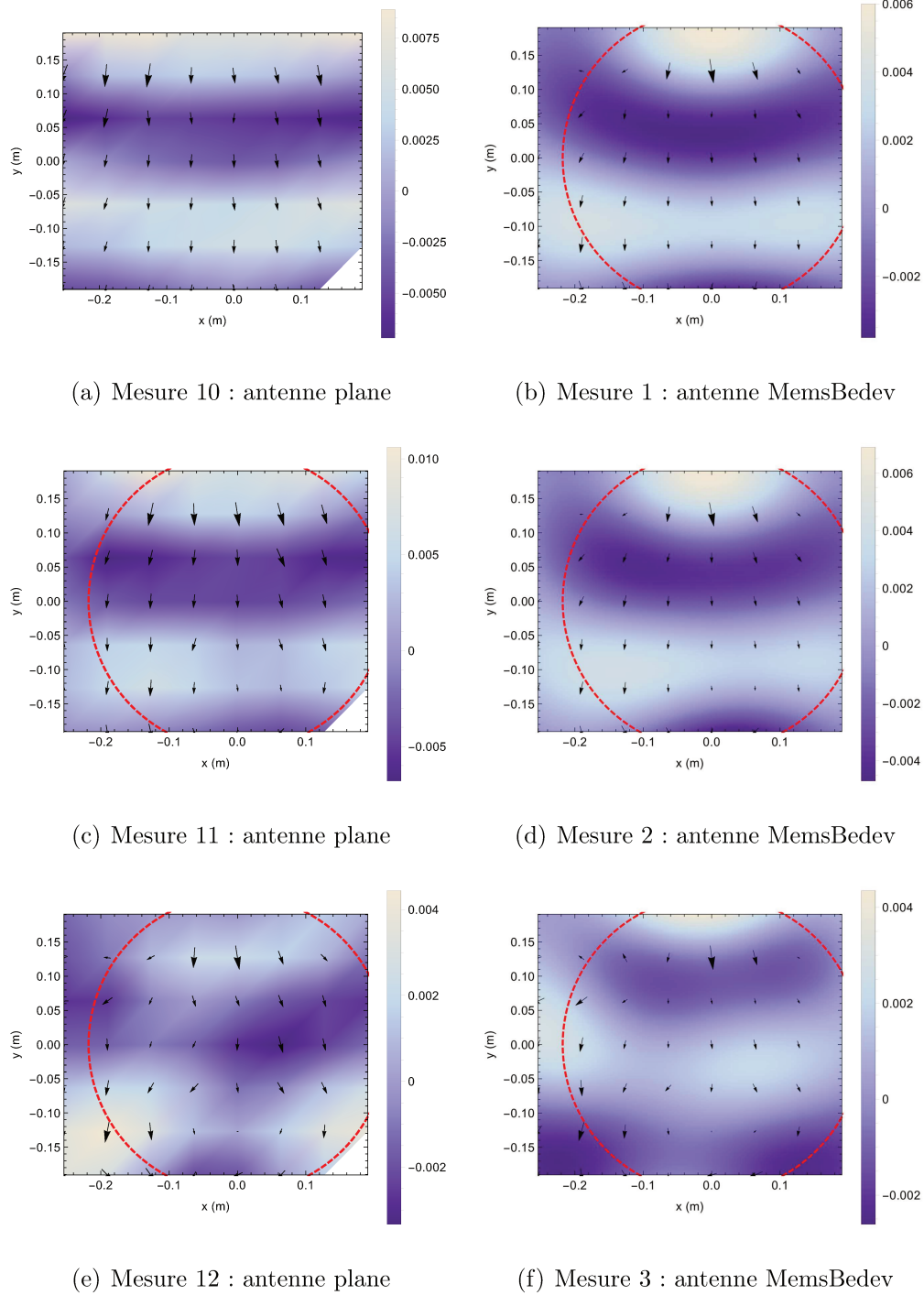
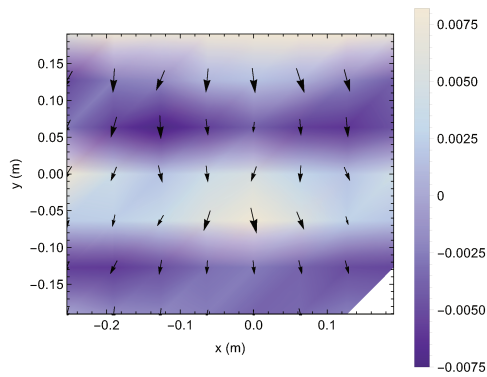
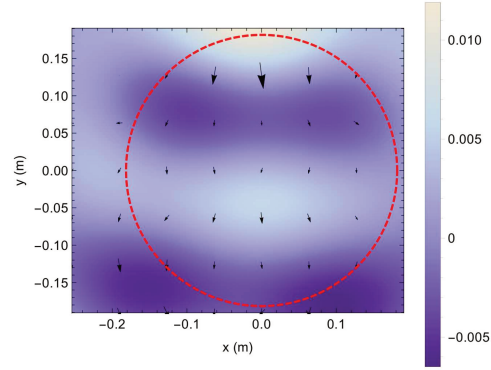


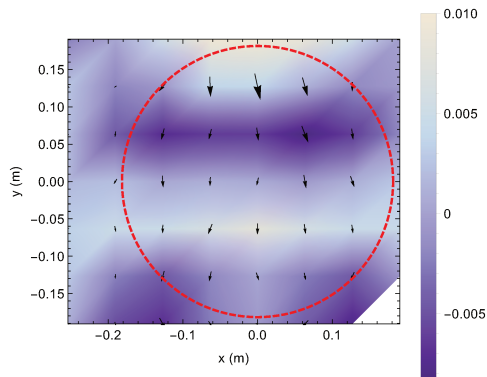
Figure 6.22 Partie réelle des champs de pression à 1250 Hz pour les mesures 10, 11, 12 (colonne de gauche, antenne plane) et 1, 2, 3 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 3 ($1,07\text{ m}$, 90° , 0°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 3. (e) et (f) Haut-parleur 3 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.



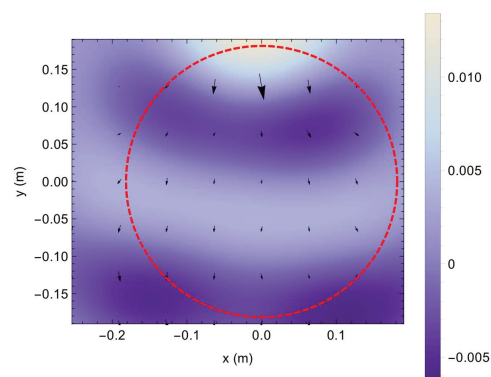
(a) Mesure 10 : antenne plane



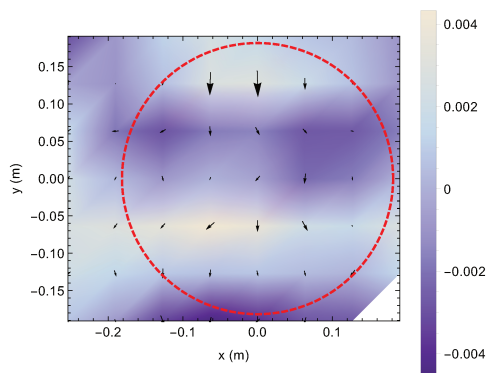
(b) Mesure 1 : antenne MemsBedev



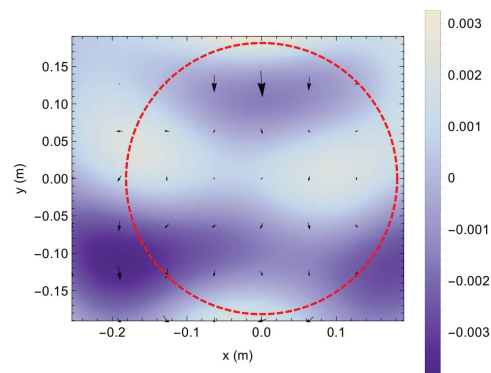
(c) Mesure 11 : antenne plane



(d) Mesure 2 : antenne MemsBedev



(e) Mesure 12 : antenne plane



(f) Mesure 3 : antenne MemsBedev

Figure 6.23 Partie réelle des champs de pression à 1500 Hz pour les mesures 10, 11, 12 (colonne de gauche, antenne plane) et 1, 2, 3 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 3 (1,07 m, 90° , 0°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 3. (e) et (f) Haut-parleur 3 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

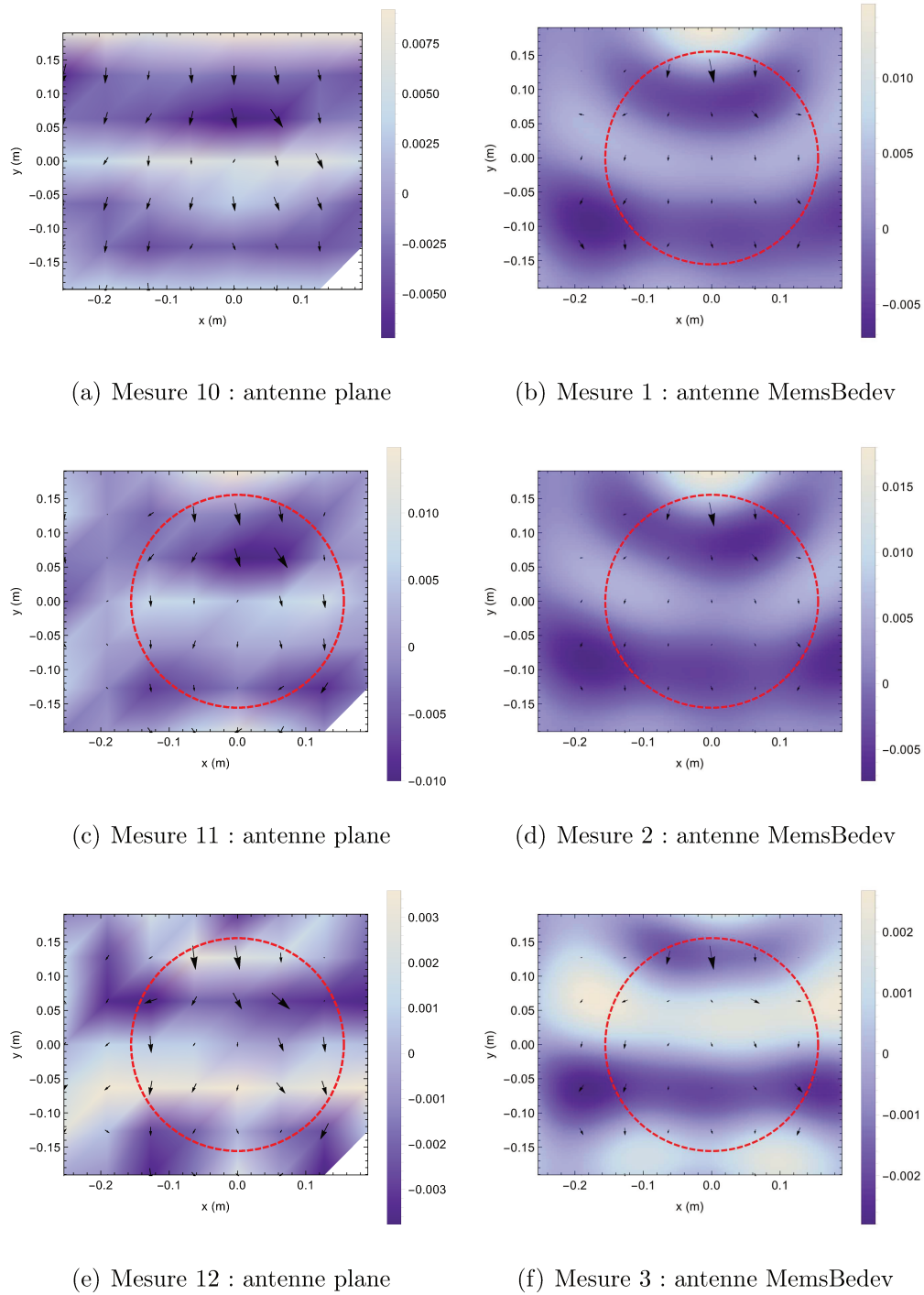


Figure 6.24 Partie réelle des champs de pression à 1750 Hz pour les mesures 10, 11, 12 (colonne de gauche, antenne plane) et 1, 2, 3 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 3 (1,07 m, 90°, 0°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 3. (e) et (f) Haut-parleur 3 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

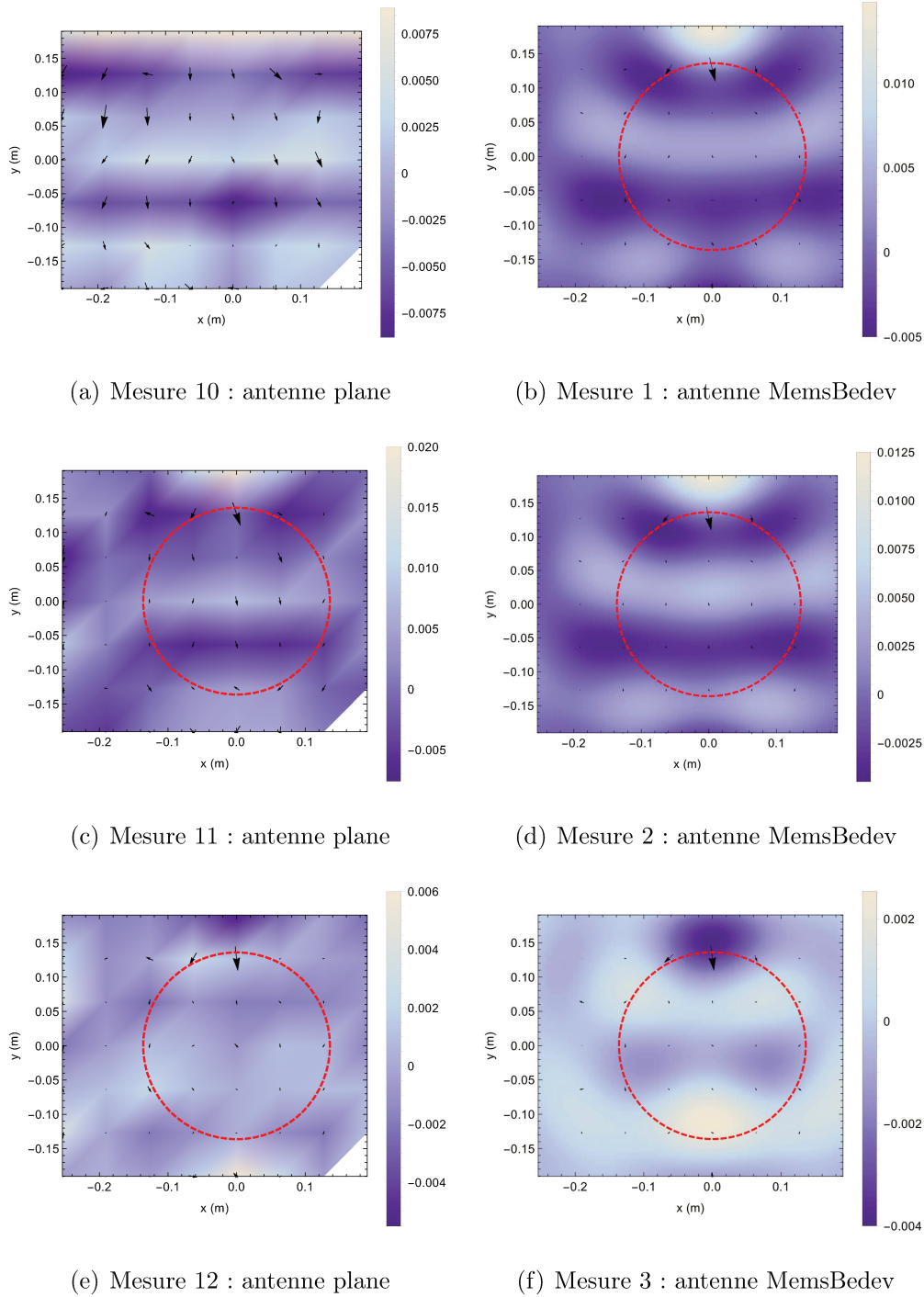


Figure 6.25 Partie réelle des champs de pression à 2000 Hz pour les mesures 10, 11, 12 (colonne de gauche, antenne plane) et 1, 2, 3 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 3 ($1,07\text{ m}$, 90° , 0°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 3. (e) et (f) Haut-parleur 3 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

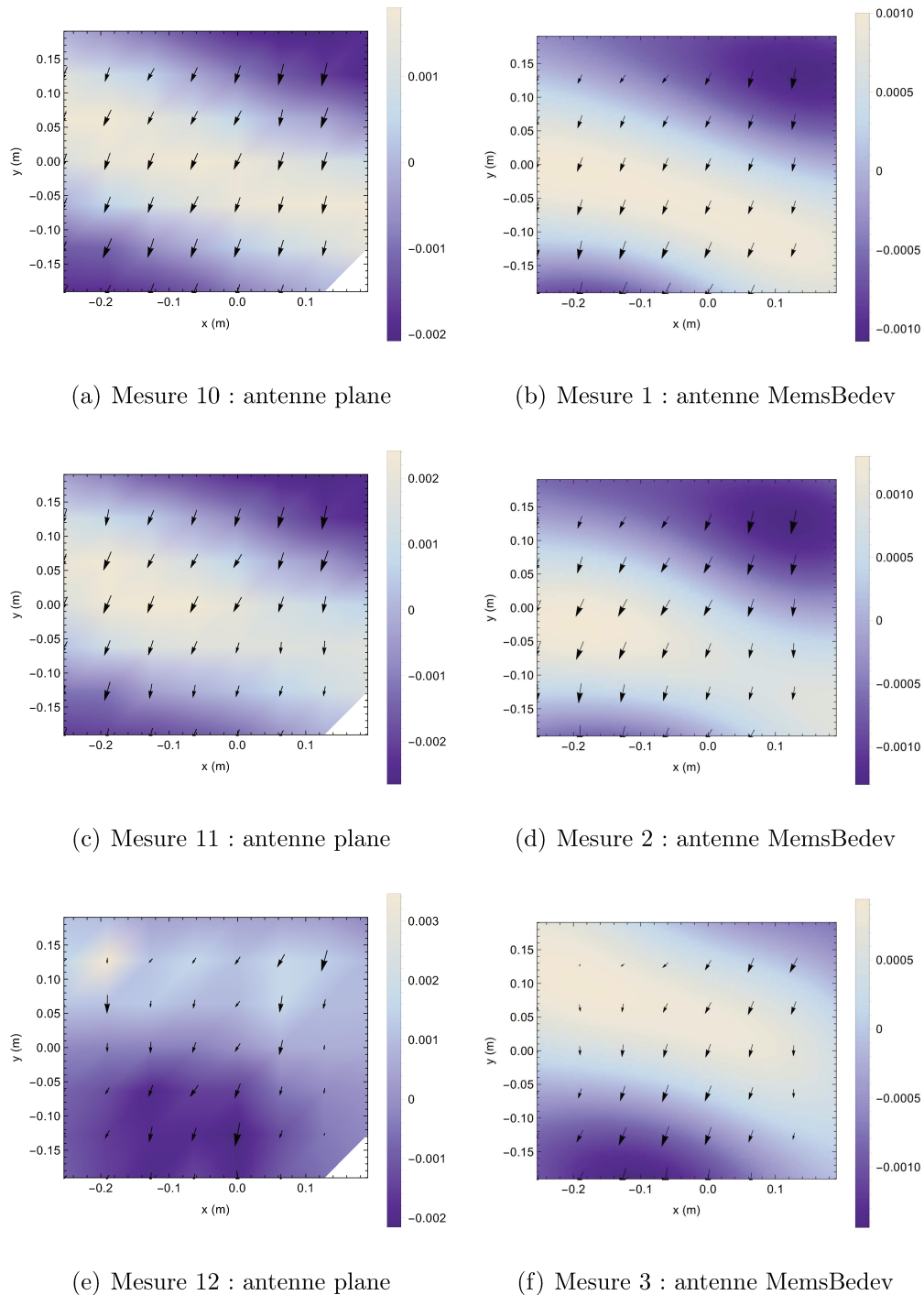
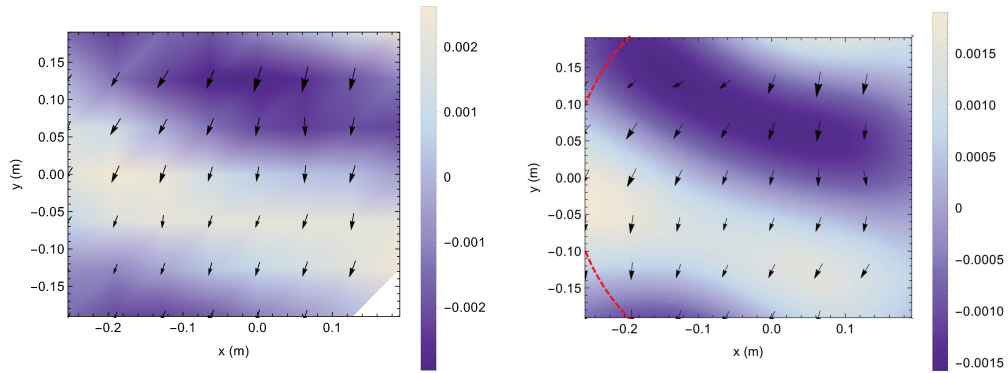
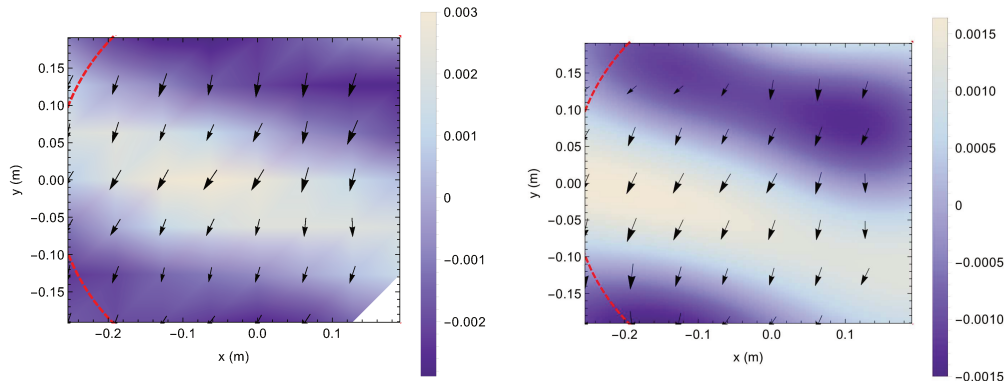


Figure 6.26 Partie réelle des champs de pression à 750 Hz pour les mesures 13, 14, 15 (colonne de gauche, antenne plane) et 4, 5, 6 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 51 (2.23 m, 73° , 12.8°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 51. (e) et (f) Haut-parleur 51 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.



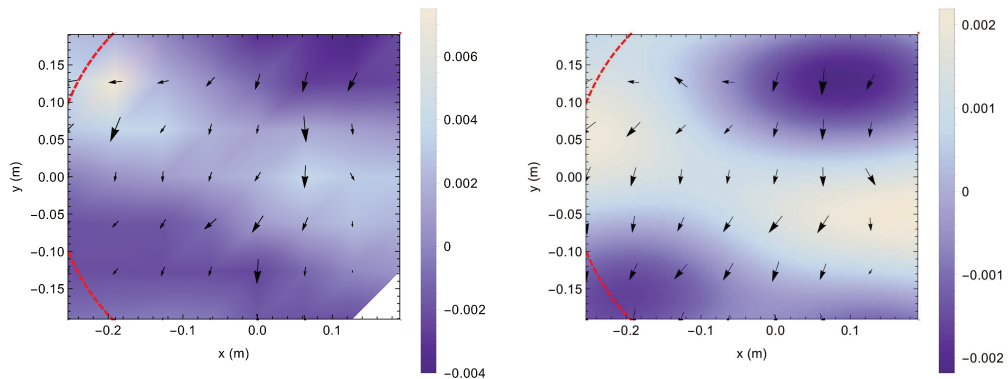
(a) Mesure 10 : antenne plane

(b) Mesure 1 : antenne MemsBedev



(c) Mesure 11 : antenne plane

(d) Mesure 2 : antenne MemsBedev



(e) Mesure 12 : antenne plane

(f) Mesure 3 : antenne MemsBedev

Figure 6.27 Partie réelle des champs de pression à 1000 Hz pour les mesures 13, 14, 15 (colonne de gauche, antenne plane) et 4, 5, 6 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 51 (2.23 m, 73°, 12.8°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 51. (e) et (f) Haut-parleur 51 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

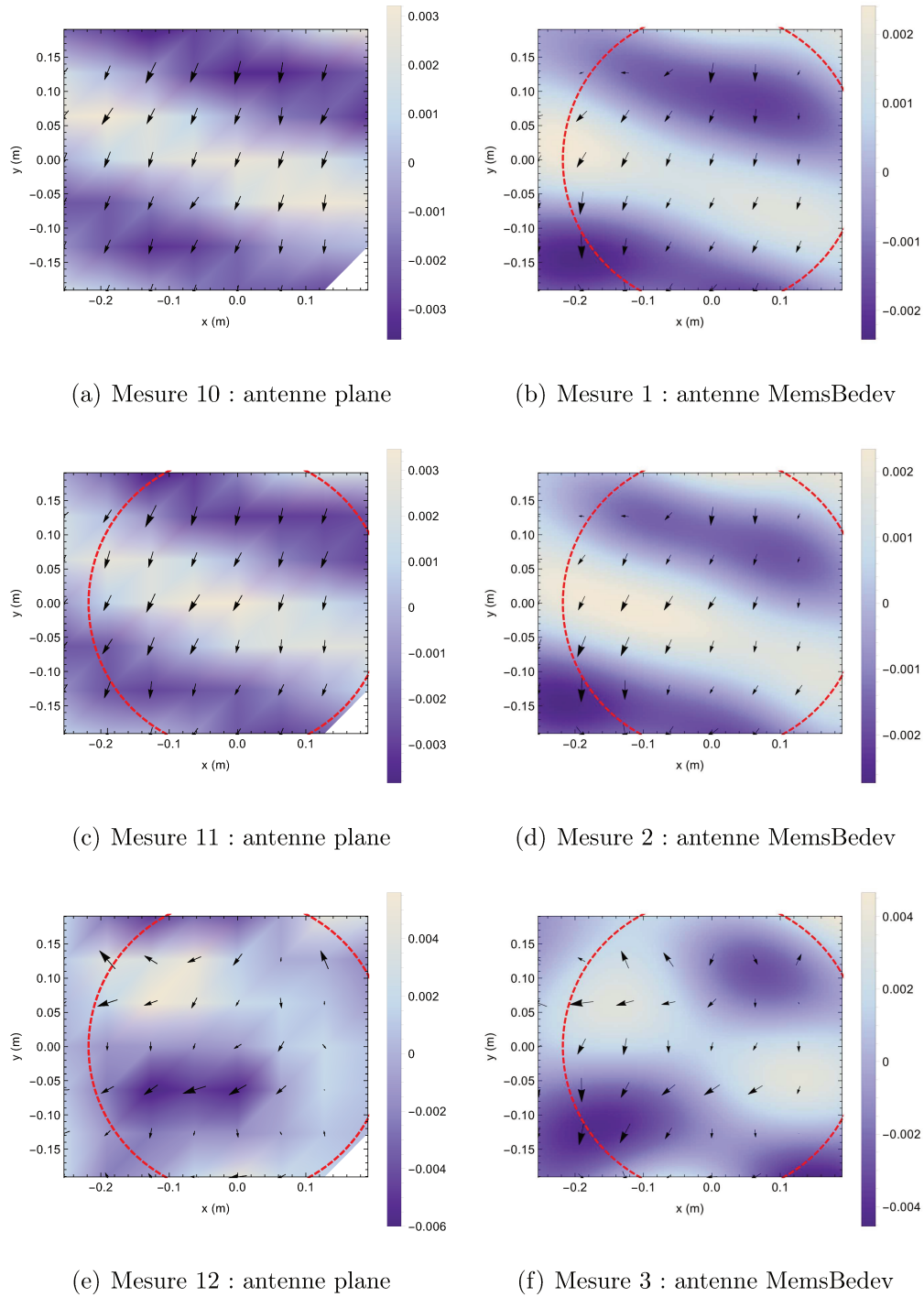


Figure 6.28 Partie réelle des champs de pression à 1250 Hz pour les mesures 13, 14, 15 (colonne de gauche, antenne plane) et 4, 5, 6 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 51 (2.23 m, 73° , 12.8°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 51. (e) et (f) Haut-parleur 51 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

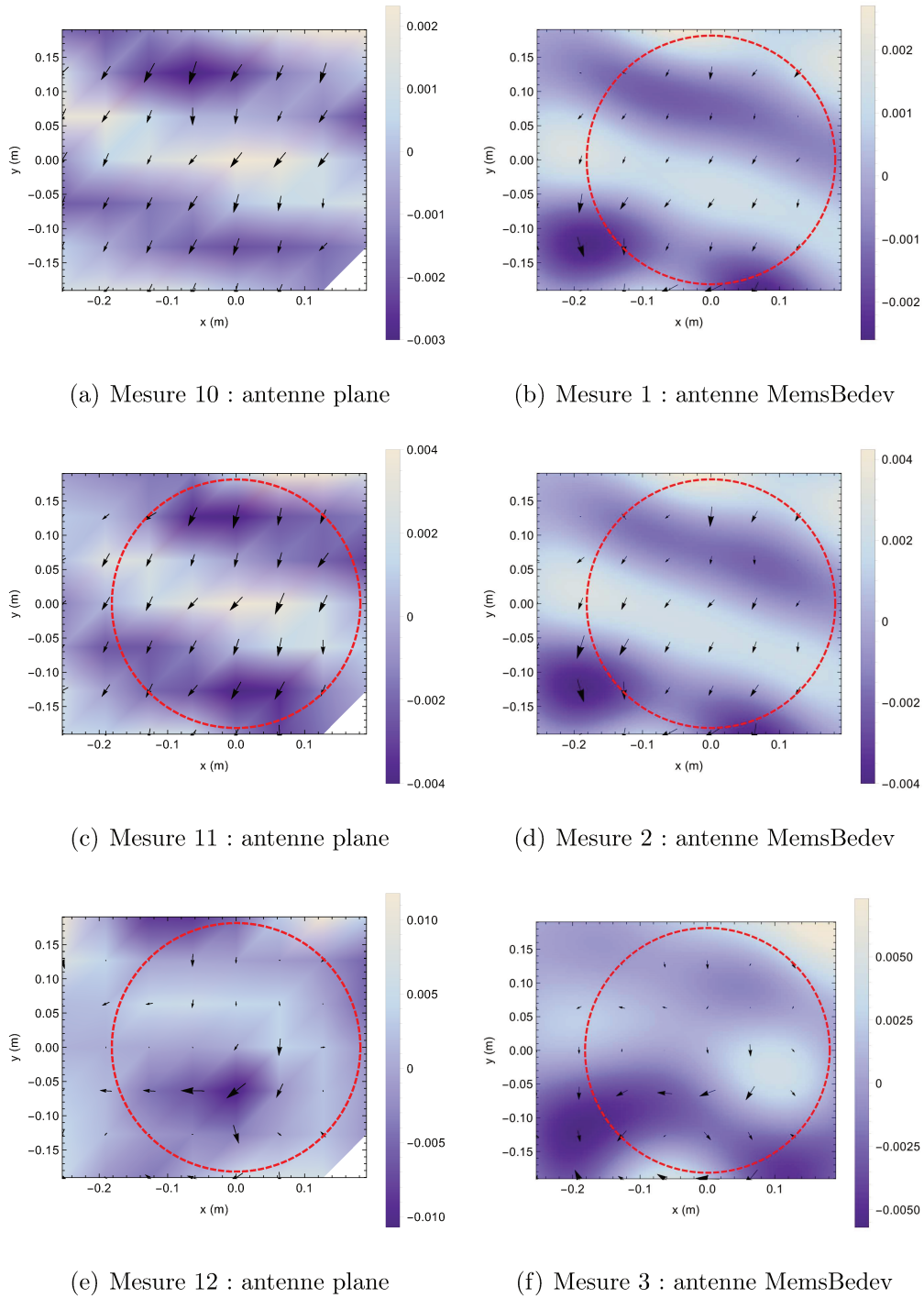


Figure 6.29 Partie réelle des champs de pression à 1500 Hz pour les mesures 13, 14, 15 (colonne de gauche, antenne plane) et 4, 5, 6 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 51 (2.23 m, 73° , 12.8°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 51. (e) et (f) Haut-parleur 51 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

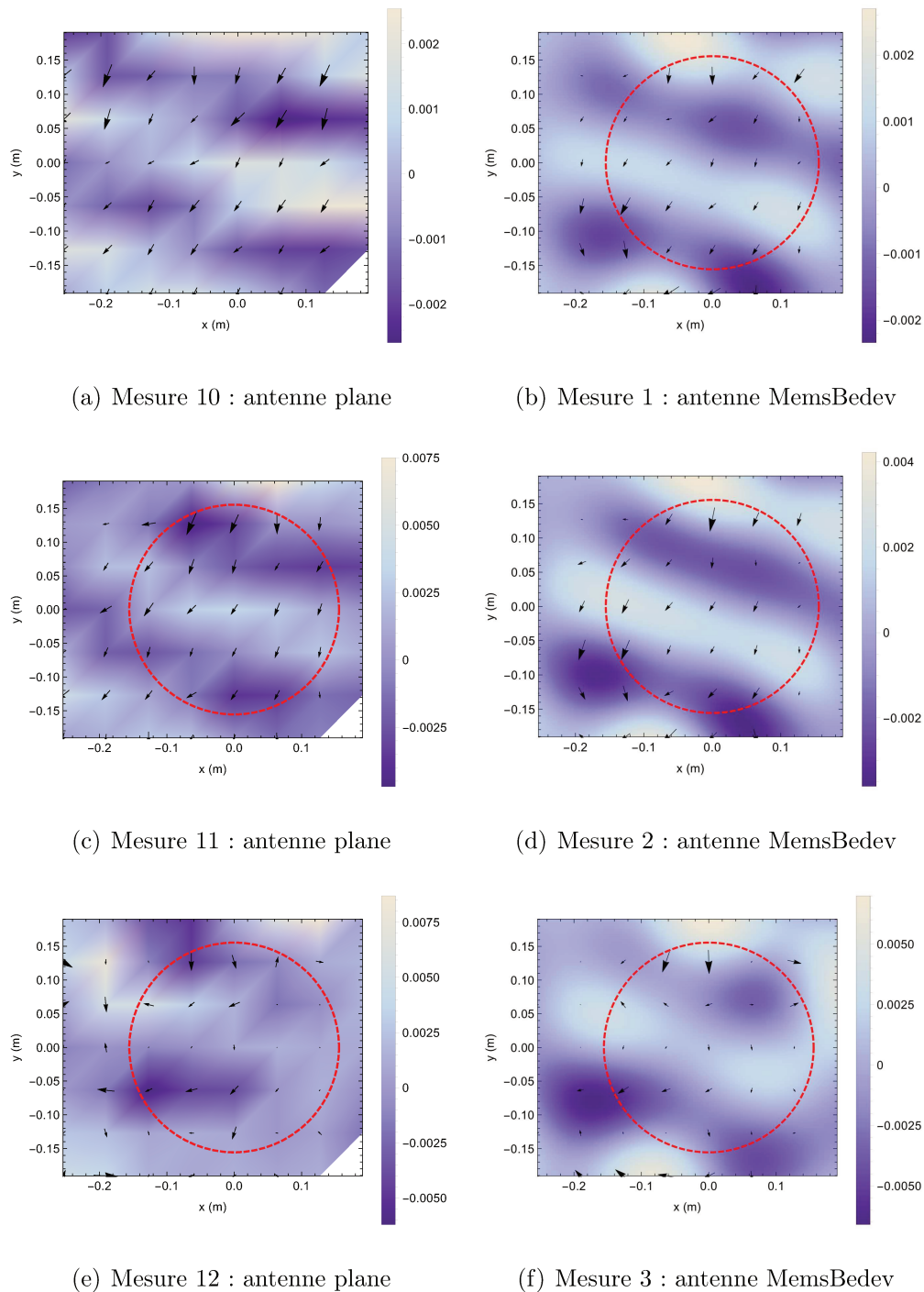
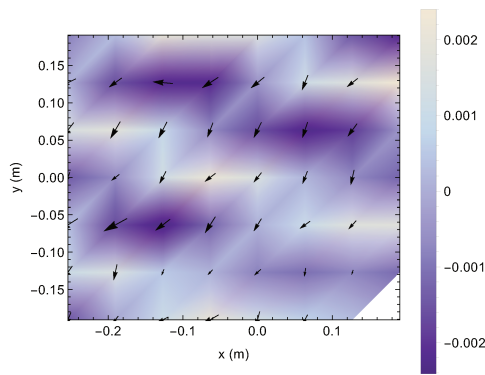
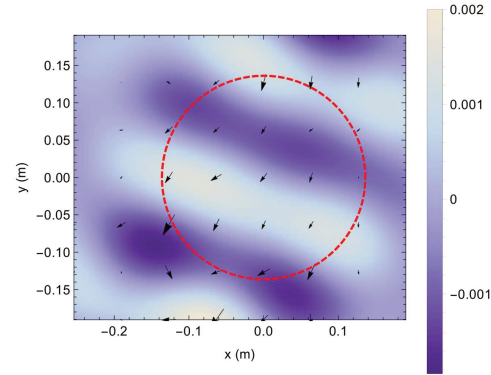


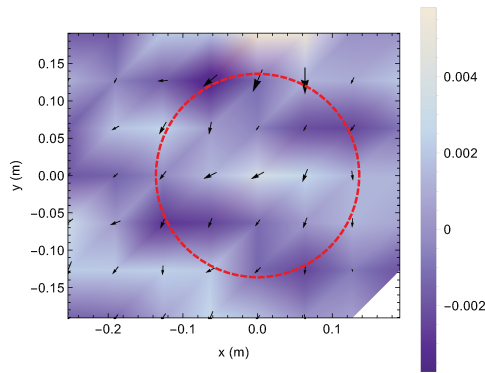
Figure 6.30 Partie réelle des champs de pression à 1750 Hz pour les mesures 13, 14, 15 (colonne de gauche, antenne plane) et 4, 5, 6 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 51 (2.23 m, 73° , 12.8°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 51. (e) et (f) Haut-parleur 51 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.



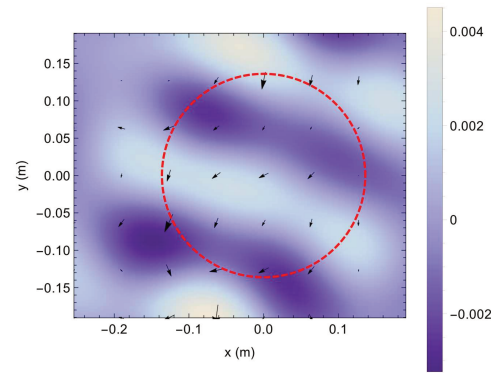
(a) Mesure 10 : antenne plane



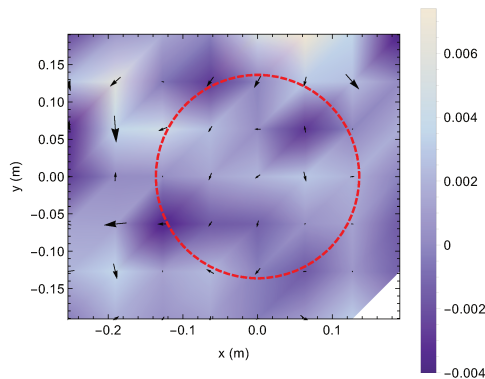
(b) Mesure 1 : antenne MemsBedev



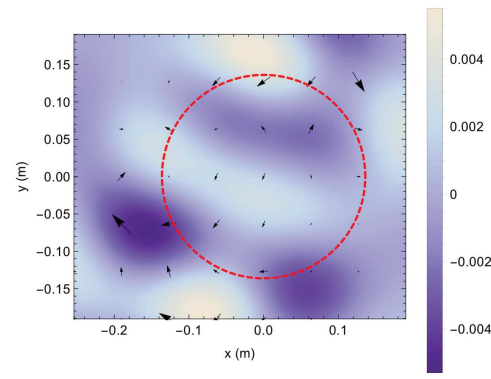
(c) Mesure 11 : antenne plane



(d) Mesure 2 : antenne MemsBedev



(e) Mesure 12 : antenne plane



(f) Mesure 3 : antenne MemsBedev

Figure 6.31 Partie réelle des champs de pression à 2000 Hz pour les mesures 13, 14, 15 (colonne de gauche, antenne plane) et 4, 5, 6 (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 51 (2.23 m, 73° , 12.8°) seul. (c) et (d) Source virtuelle à la position du haut-parleur 51. (e) et (f) Haut-parleur 51 capté par l'antenne MemsBedev puis décodé sur la SpherBedev 2.0. En rouge pointillé est montré le cercle de rayon $r = 5/k$.

fréquences allant de 500 Hz à 2000 Hz. L'objectif est de valider expérimentalement l'approche de filtrage directionnel présenté au chapitre 5. Un champ sonore avec deux sources perpendiculaires et capté à l'aide de l'antenne MemsBedev, encodé, puis filtré dans le domaine des harmoniques sphériques pour ne garder que la contribution d'une seule des deux sources. Deux situations sont présentées pour l'antenne plane et l'antenne MemsBedev :

- La première correspond à un champ sonore capté par l'antenne MemsBedev, encodé à l'ordre $M = 4$ et décodé sur la SpherBedev 2.0. (figures du haut). Le champ sonore original est créé en alimentant deux haut-parleurs simultanément avec le signal d'excitation (mesure 7). Les deux haut-parleurs sont les haut-parleurs 2 ($1,07\text{ m}$, 0° , 0°) et 3 de la SpherBedev 2.0. Ils sont dans des directions perpendiculaires dans le plan horizontal.
- La seconde situation correspond au champ sonore précédent, issu de la mesure 7, mais dont les composantes ambisoniques encodées sont filtrées avant le décodage, à l'aide de l'outil `hoa_beamforming_hypercardioid_to_hoa`. Cet outil réalise un filtrage directionnel opéré sur le champ sonore capté par l'antenne MemsBedev suivant la méthode décrite dans le chapitre 5. La fonction de directivité utilisée est une hypercardioïde d'ordre $M' = 1$ et d'angle de pointage $(0^\circ, 0^\circ)$ (direction du haut-parleur 2). La scène sonore une fois filtrée est décrite à l'ordre $M = 5$, d'après l'équation (5.14). Elle est finalement décodée sur la SpherBedev 2.0. (figures du bas).

Comme il est précisé à la section 6.4.7, les réponses impulsionnelles de la mesure 8 sont recalées par rapport à celle de la mesure 15. Puis, les réponses impulsionnelles de la mesure 9 sont recalées par rapport à celles de la mesure 8 pour l'antenne MemsBedev. De même, les réponses impulsionnelles de la mesure 16 sont recalées par rapport à celles de la mesure 15 pour l'antenne plane.

L'observation des figures correspondants aux mesures 8 et 15 (figures du haut) montre les interférences issues des deux sources émettant en phase mais dans des directions perpendiculaires. Ainsi, le champ d'intensité est en concordance avec ce phénomène et la direction de propagation moyenne semble être de $(45^\circ, 0^\circ)$. Le champ sonore est encodé à l'ordre $M = 4$. La zone de reconstruction théorique correspondante est représentée par le cercle en rouge pointillé.

Dès lors que le filtrage directionnel est actif, mesures 9 et 16, (figures du bas), le champ sonore obtenu après décodage est majoritairement dû à la contribution du haut-parleur 2. En effet, la fonction de directivité utilisée est une hypercardioïde d'ordre $M' = 1$ et son angle de pointage est choisi dans le direction du haut-parleur 2 $(0^\circ, 0^\circ)$. Ainsi, les

directions du champ sonore sont atténuées suivant le patron de directivité de l'hyper-cardioïde, visible à la figure 5.1 du chapitre 5. Le haut-parleur 3, qui se trouve dans une direction perpendiculaire à l'angle de pointage de l'hyper-cardioïde est fortement atténué. Il en résulte un champ sonore dont la direction de propagation est de $(0^\circ, 0^\circ)$, telle qu'indiquée par le champ de vecteur intensité. On rappelle que l'opération de filtrage directionnel requiert un développement du champ sonore filtré à un ordre supérieur par rapport au champ sonore original, tel qu'indiqué par l'équation (5.11). Ainsi, le champ sonore une fois filtré est à l'ordre $\tilde{M} = 5$, la zone de reconstruction correspondant à cet ordre est indiqué par le cercle en pointillé sur les figures.

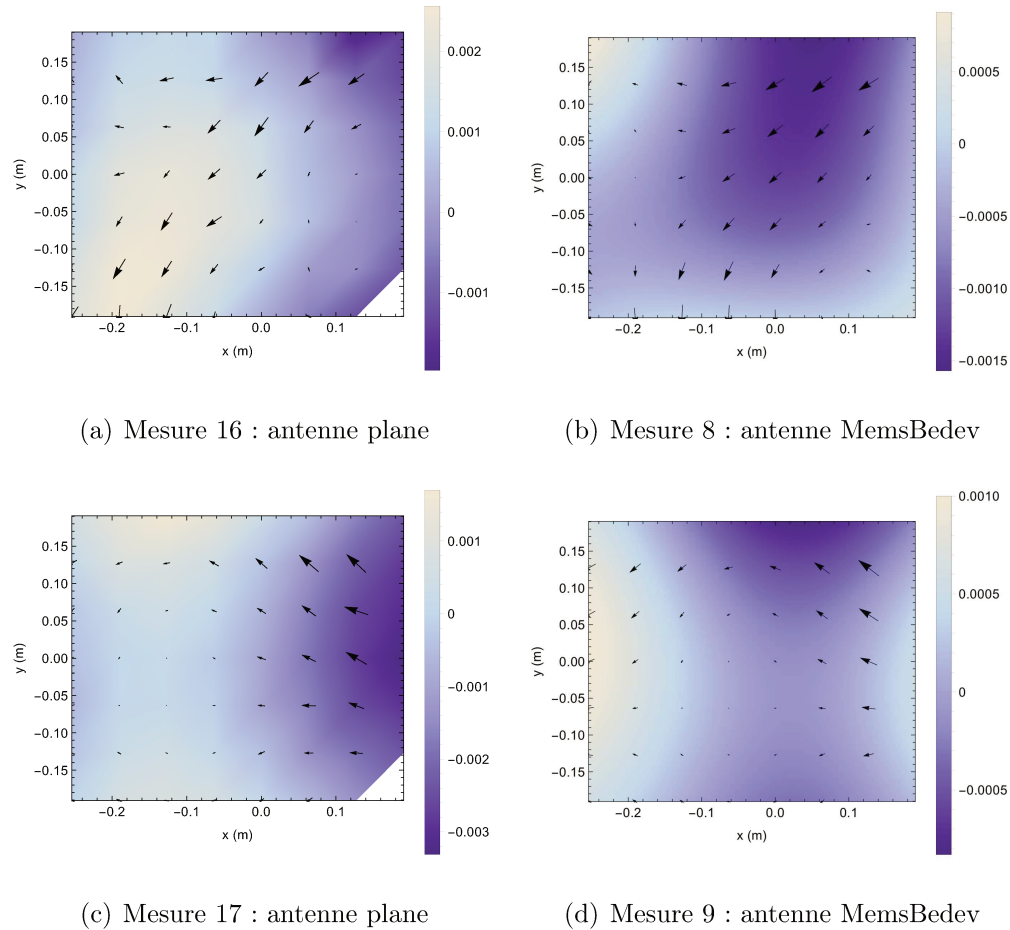


Figure 6.32 Partie réelle des champs de pression à 500 Hz pour les mesures 15, 16, (colonne de gauche, antenne plane) et 8, 9, (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 2 ($1,07\text{ m}$, 0° , 0°) et haut-parleur 3 ($1,07\text{ m}$, 90° , 0°) alimenté simultanément, captés par l'antenne MemsBedev, encodé à l'ordre $M = 4$ puis décodé sur la SpherBedev. (c) et (d) Même situation, mais le champ sonore a subi un filtrage directionnel à l'aide de l'outil `hoa_beamforming_to_hoa` avant décodage. La fonction de directivité utilisée est une hyper-cardioïde d'ordre $M' = 1$ et d'angle de pointage $(0^\circ, 0^\circ)$. En rouge pointillé est montré le cercle de rayon $r = 4/k$ ((a) et (b)) et $5/k$ ((c) et (d)).

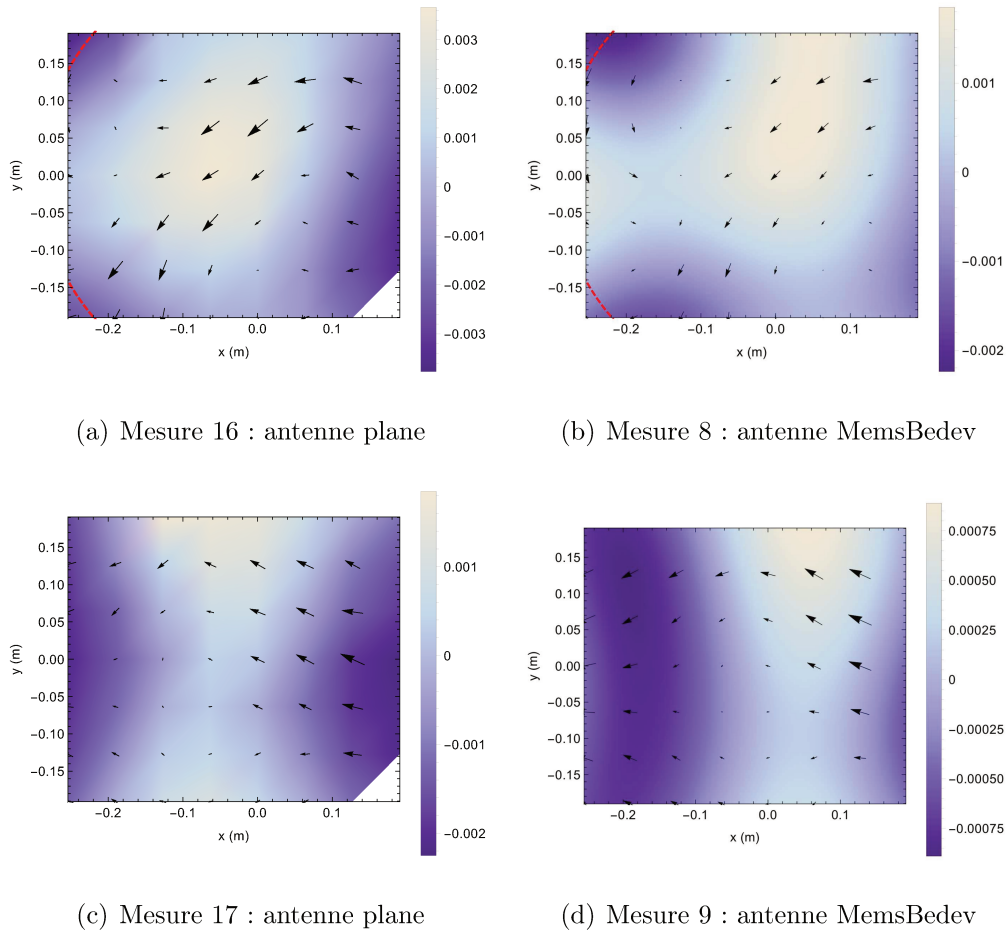


Figure 6.33 Partie réelle des champs de pression à 750 Hz pour les mesures 15, 16, (colonne de gauche, antenne plane) et 8, 9, (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 2 ($1,07 \text{ m}$, 0° , 0°) et haut-parleur 3 ($1,07 \text{ m}$, 90° , 0°) alimenté simultanément, captés par l'antenne MemsBedev, encodé à l'ordre $M = 4$ puis décodé sur la SpherBedev. (c) et (d) Même situation, mais le champ sonore a subi un filtrage directionnel à l'aide de l'outil `hoa_beamforming_to_hoa` avant décodage. La fonction de directivité utilisée est une hyper-cardioïde d'ordre $M' = 1$ et d'angle de pointage $(0^\circ, 0^\circ)$. En rouge pointillé est montré le cercle de rayon $r = 4/k$ ((a) et (b)) et $5/k$ ((c) et (d)).

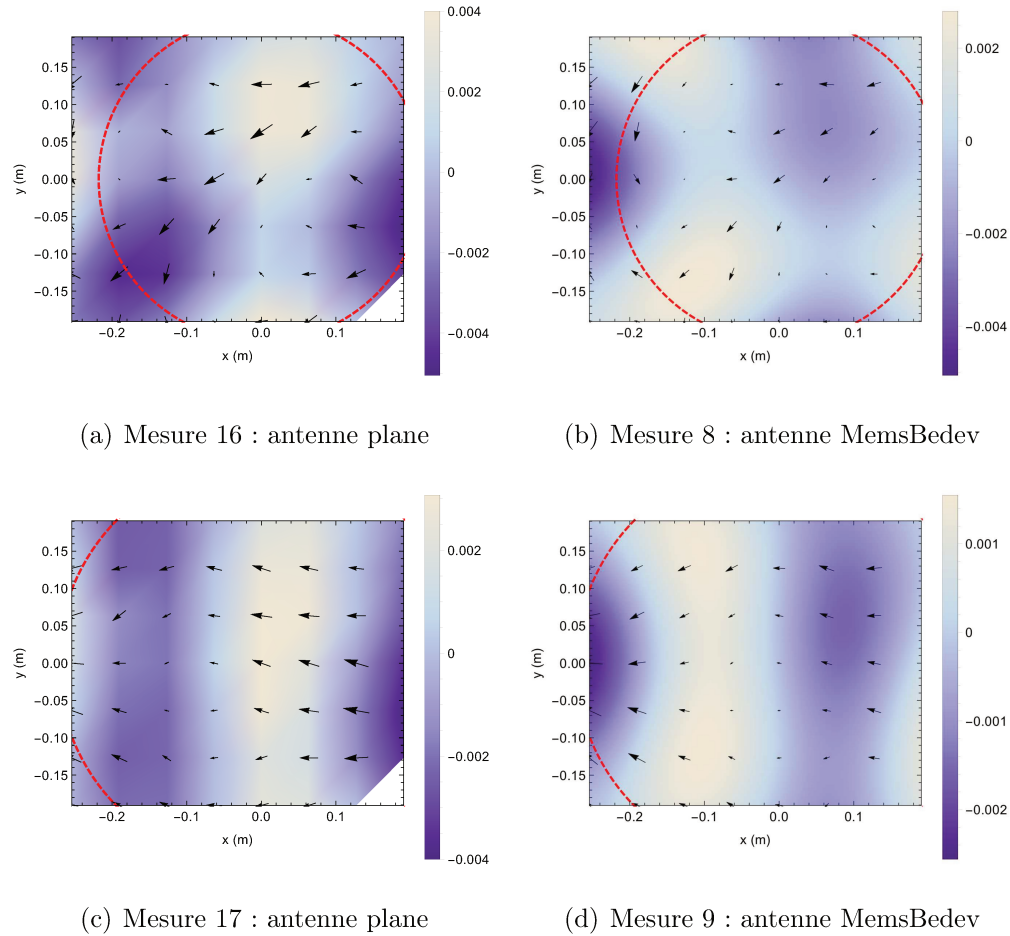


Figure 6.34 Partie réelle des champs de pression à 1000 Hz pour les mesures 15, 16, (colonne de gauche, antenne plane) et 8, 9, (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 2 ($1,07 \text{ m}$, 0° , 0°) et haut-parleur 3 ($1,07 \text{ m}$, 90° , 0°) alimenté simultanément, captés par l'antenne MemsBedev, encodé à l'ordre $M = 4$ puis décodé sur la SpherBedev. (c) et (d) Même situation, mais le champ sonore a subi un filtrage directionnel à l'aide de l'outil `hoa_beamforming_to_hoa` avant décodage. La fonction de directivité utilisée est une hyper-cardioïde d'ordre $M' = 1$ et d'angle de pointage $(0^\circ, 0^\circ)$. En rouge pointillé est montré le cercle de rayon $r = 4/k$ ((a) et (b)) et $5/k$ ((c) et (d)).

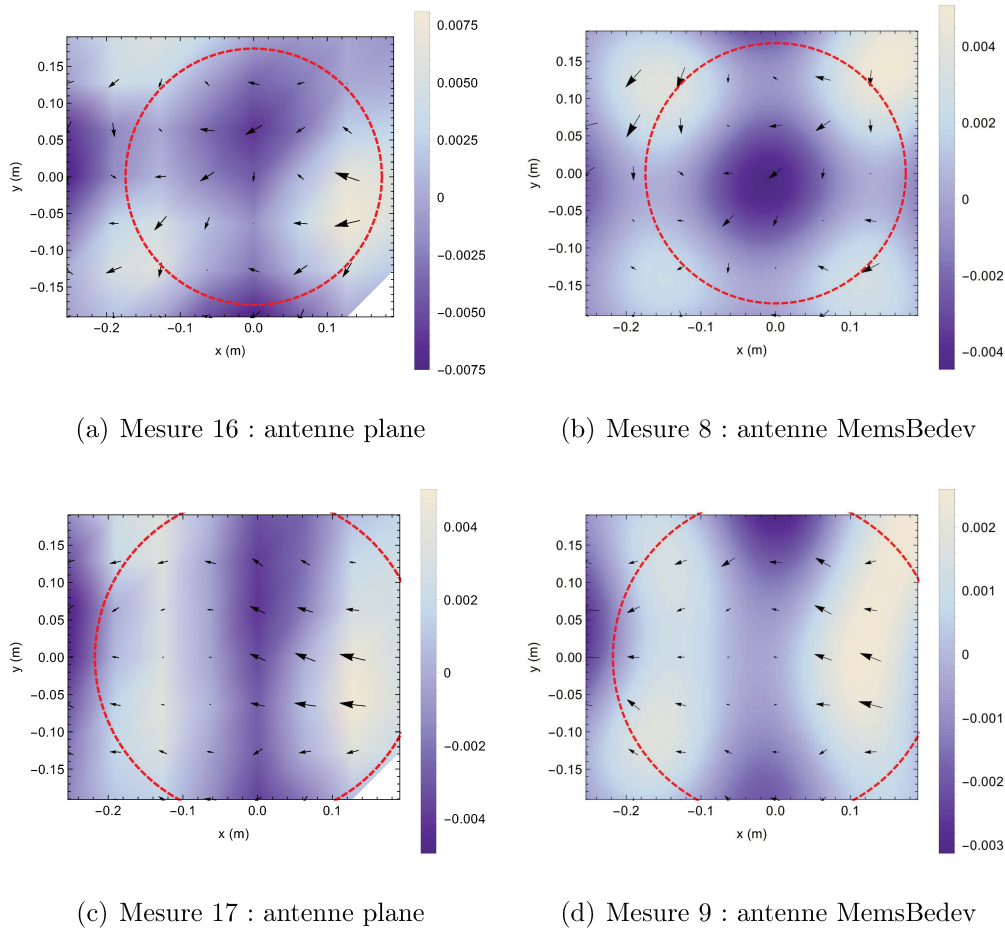


Figure 6.35 Partie réelle des champs de pression à 1250 Hz pour les mesures 15, 16, (colonne de gauche, antenne plane) et 8, 9, (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 2 ($1,07\text{ m}$, 0° , 0°) et haut-parleur 3 ($1,07\text{ m}$, 90° , 0°) alimenté simultanément, captés par l’antenne MemsBedev, encodé à l’ordre $M = 4$ puis décodé sur la SpherBedev. (c) et (d) Même situation, mais le champ sonore a subi un filtrage directionnel à l’aide de l’outil `hoa_beamforming_to_hoa` avant décodage. La fonction de directivité utilisée est une hyper-cardioïde d’ordre $M' = 1$ et d’angle de pointage $(0^\circ, 0^\circ)$. En rouge pointillé est montré le cercle de rayon $r = 4/k$ ((a) et (b)) et $5/k$ ((c) et (d)).

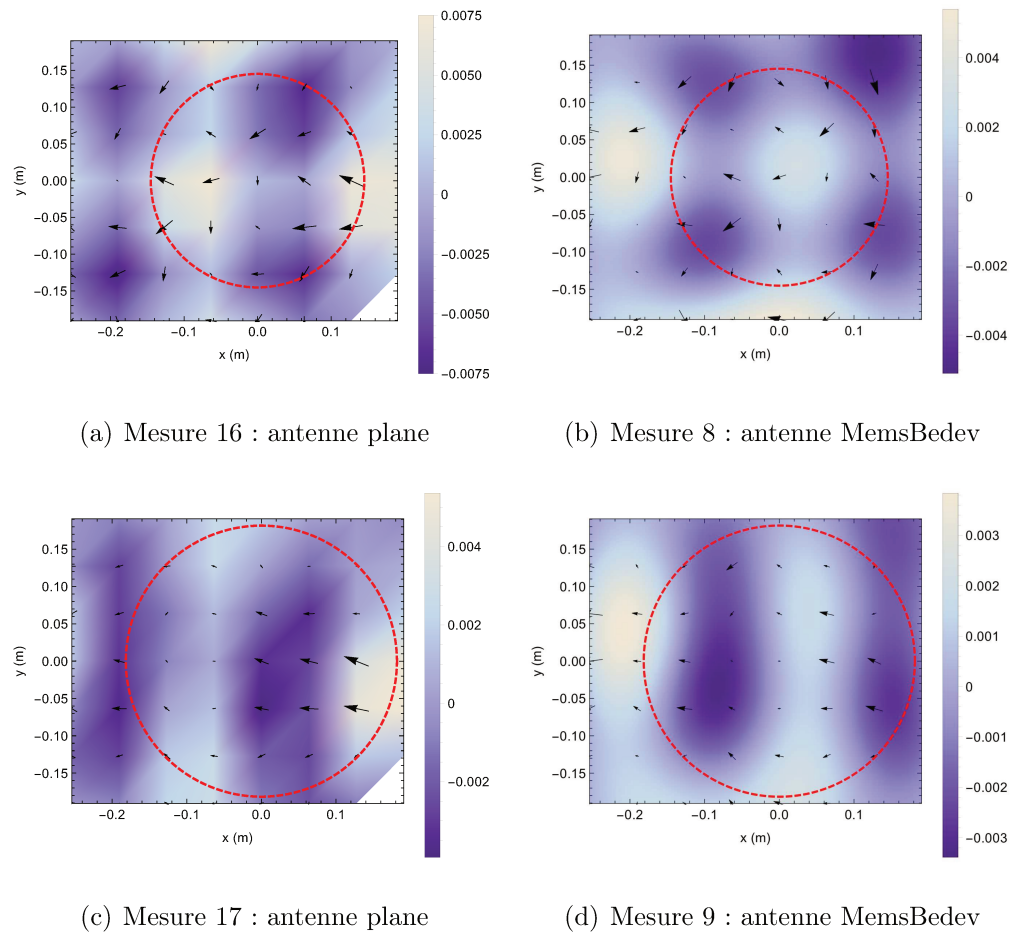


Figure 6.36 Partie réelle des champs de pression à 1500 Hz pour les mesures 15, 16, (colonne de gauche, antenne plane) et 8, 9, (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 2 ($1,07\text{ m}$, 0° , 0°) et haut-parleur 3 ($1,07\text{ m}$, 90° , 0°) alimenté simultanément, captés par l'antenne MemsBedev, encodé à l'ordre $M = 4$ puis décodé sur la SpherBedev. (c) et (d) Même situation, mais le champ sonore a subi un filtrage directionnel à l'aide de l'outil `hoa_beamforming_to_hoa` avant décodage. La fonction de directivité utilisée est une hyper-cardioïde d'ordre $M' = 1$ et d'angle de pointage $(0^\circ, 0^\circ)$. En rouge pointillé est montré le cercle de rayon $r = 4/k$ ((a) et (b)) et $5/k$ ((c) et (d)).

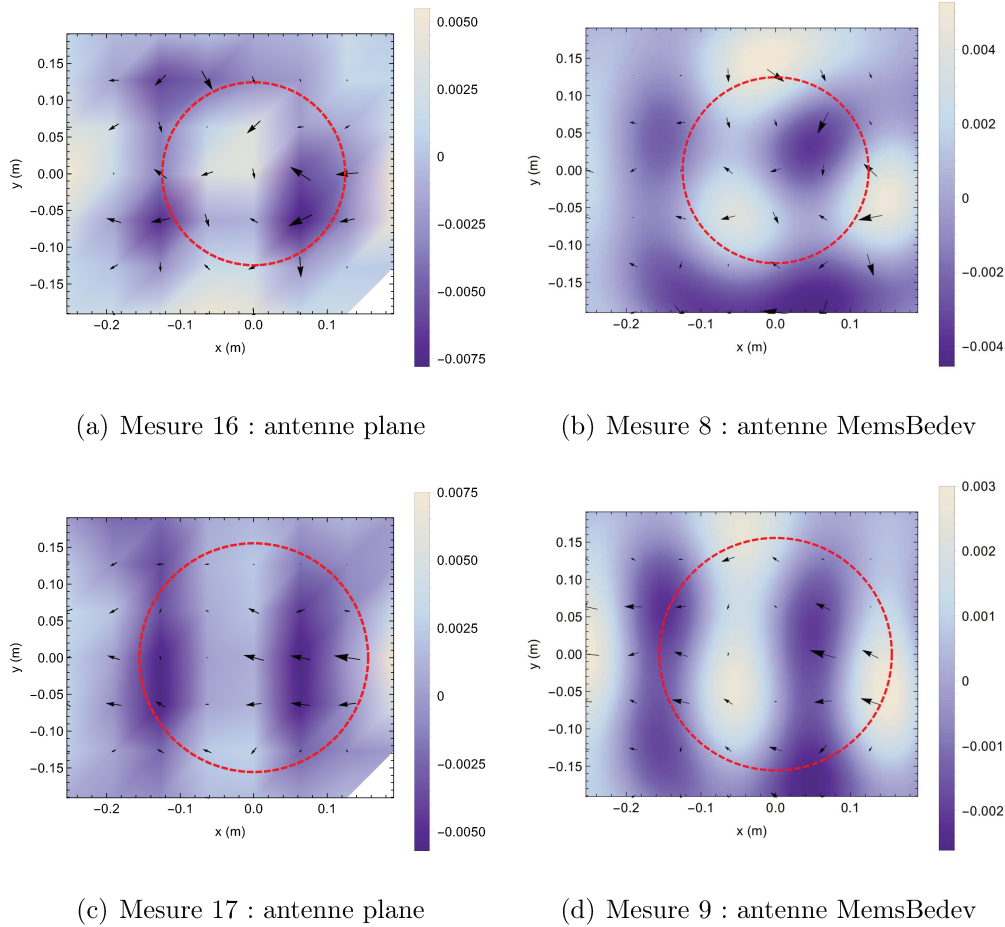


Figure 6.37 Partie réelle des champs de pression à 1750 Hz pour les mesures 15, 16, (colonne de gauche, antenne plane) et 8, 9, (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 2 (1,07 m, 0° , 0°) et haut-parleur 3 (1,07 m, 90° , 0°) alimenté simultanément, captés par l’antenne MemsBedev, encodé à l’ordre $M = 4$ puis décodé sur la SpherBedev. (c) et (d) Même situation, mais le champ sonore a subi un filtrage directionnel à l’aide de l’outil `hoa_beamforming_to_hoa` avant décodage. La fonction de directivité utilisée est une hyper-cardioïde d’ordre $M' = 1$ et d’angle de pointage (0° , 0°). En rouge pointillé est montré le cercle de rayon $r = 4/k$ ((a) et (b)) et $5/k$ ((c) et (d)).

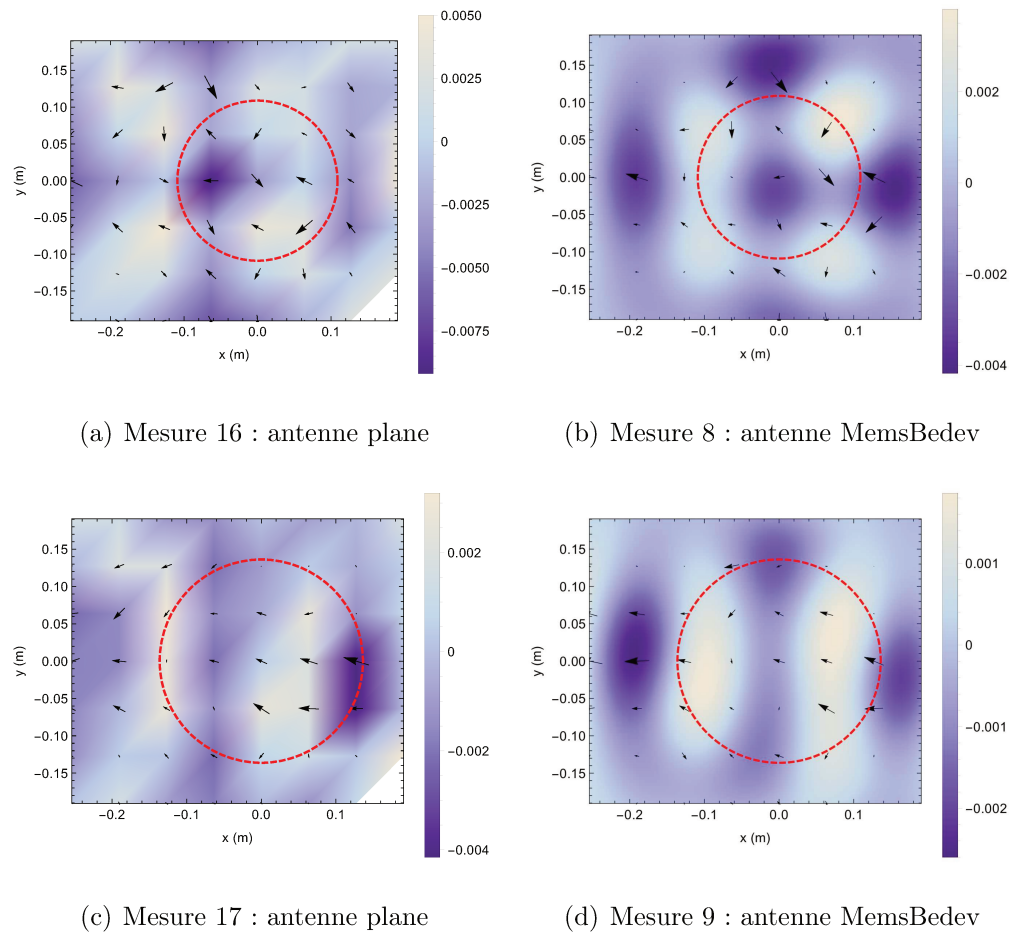


Figure 6.38 Partie réelle des champs de pression à 2000 Hz pour les mesures 15, 16, (colonne de gauche, antenne plane) et 8, 9, (colonne de droite, antenne MemsBedev, champ extrapolé). (a) et (b) Haut-parleur 2 ($1,07\text{ m}$, 0° , 0°) et haut-parleur 3 ($1,07\text{ m}$, 90° , 0°) alimenté simultanément, captés par l'antenne MemsBedev, encodé à l'ordre $M = 4$ puis décodé sur la SpherBedev. (c) et (d) Même situation, mais le champ sonore a subi un filtrage directionnel à l'aide de l'outil `hoa_beamforming_to_hoa` avant décodage. La fonction de directivité utilisée est une hyper-cardioïde d'ordre $M' = 1$ et d'angle de pointage $(0^\circ, 0^\circ)$. En rouge pointillé est montré le cercle de rayon $r = 4/k$ ((a) et (b)) et $5/k$ ((c) et (d)).

6.4.10 Conclusions

Le but de ce chapitre était de démontrer expérimentalement la validité des propositions originales présentées théoriquement aux chapitre 3 et 5. Conséquemment, ce chapitre a présenté les dispositifs expérimentaux réalisés pour ce projet de doctorat et a décrit précisément une campagne de mesure effectuée à l'aide de ces dispositifs. Deux sphères de restitution « SpherBedev » ont été réalisées en se basant sur un maillage de Lededev. Elles permettent la reconstruction de champs sonores en trois dimensions par approche ambisonique jusqu'à l'ordre 5. Un prototype d'antenne de microphone sphérique « MemsBedev » a également été réalisé pour permettre la captation et l'encodage d'un champ sonore en trois dimensions jusqu'à l'ordre 5.

Une campagne de mesures en condition anéchoïque a ensuite été réalisée pour valider expérimentalement certains points développés dans ce doctorat. Ainsi, la synthèse de sources virtuelles s'est montrée efficace sur une zone étendue de l'espace en comparaison avec les champs sonores issus de sources réelles. Les fronts d'ondes sont bien reproduits ainsi que les directions de propagation, dans la zone de reconstruction. Par ailleurs, le microphone MemsBedev s'est montré apte à la captation et à l'encodage d'un champ sonore en format ambisonique. En effet, les cartes extrapolées sont en concordance avec les cartes mesurées. De même, les champs sonores décodés, captés initialement avec le microphone MemsBedev restent également similaires aux champs sonores initiaux.

Finalement, une validation expérimentale du filtrage directionnel présenté au chapitre 5 de ce manuscrit a été réalisée. Un champ sonore composé de plusieurs sources acoustiques a été capté par le microphone MemsBedev puis décodé sur la SpherBedev. Ce même champ sonore a été filtré avant le décodage à l'aide d'une fonction de directivité de type hypercardioïde. Il en résulte un champ sonore dont les contributions des sources en dehors du lobe principal de la fonction de directivité sont fortement atténuées, validant alors les résultats de simulation présentés au chapitre 5, ainsi que l'implémentation en temps-réel dans le suite logicielle *ambitools*.

Les futures expérimentations à l'aide du réseau de haut-parleurs SpherBedev et MemsBedev pourront notamment proposer une validation expérimentale de la compensation des réflexions et de l'égalisation du système sur une zone étendue par l'approche présentée au chapitre 4 de ce manuscrit. Il deviendra alors possible de caractériser plus finement les champs sonores reproduits, en termes d'erreur quadratique notamment.

CHAPITRE 7

CONCLUSIONS

Ce projet de doctorat a permis de formuler plusieurs éléments de réponses à la problématique de synthèse physique de champs sonores en trois dimensions. Les sections suivantes rappellent les principaux résultats obtenus et envisagent les perspectives de recherche envisageables à la suite de ce doctorat.

7.1 Principaux résultats

Le chapitre 2 a permis de justifier le choix de l’approche ambisonique d’ordre élevé, comme solution analytique pour la synthèse de champs sonores en trois dimensions. Le contrôle du champ de pression acoustique dans un volume sphérique est rendu possible à l’aide d’un réseau de haut-parleurs disposés sur une surface sphérique. La captation, la manipulation et la reconstruction du champ sonore font appel à un formalisme de description du champ sonore sur une base tronquée d’harmoniques sphériques. Cette approche permet de s’affranchir du dispositif de captation et de restitution, ce qui est attrayant en termes de flexibilité d’utilisation. Par ailleurs, pour la géométrie sphérique, il a été rappelé que la formulation classique de la synthèse de fronts d’ondes se basant sur une intégrale de Rayleigh de type I avec un critère de sélection des sources secondaires est une approximation hautes-fréquences du décodage ambisonique continu.

Le choix de la méthode ambisonique étant fait, la discrétisation de la sphère a été traitée au chapitre 3. Un maillage de Lededev à cinquante noeuds a été proposé. Ce maillage permet de respecter le critère d’orthonormalité des harmoniques sphériques jusqu’à l’ordre 5 avec une erreur d’intégration nulle. Les effets de la discrétisation par rapport à une formulation uniquement tronquée ont été étudiés pour la capture et la reconstruction. Pour les trois maillages considérés, il en ressort que le maillage de Lededev présente les meilleurs résultats. La construction de la grille de Lebedev repose sur les règles de quadratures de la sphère invariantes par rotation. À la lumière de cette théorie, il apparaît que la règle de quadrature de Lebedev proposée peut offrir des maillages imbriqués les uns dans les autres. Ainsi, la grille à 50 noeuds proposée contient 2 sous-grilles de 6 et 26 noeuds, permettant une utilisation jusqu’à l’ordre 1 et 3 respectivement. Cette particularité a permis de proposer l’utilisation conjointe des grilles à 6, 26, et 50 noeuds pour concevoir

un décodeur multi-bandes en fonction de l'ordre de troncature. Ceci permet l'utilisation de haut-parleurs dont la bande passante est restreinte, et donc moins coûteux pour une réalisation pratique. Des résultats de simulations ont permis de confirmer cette approche.

Pour garantir une grande précision de reconstruction dans la pratique, le chapitre 4 a présenté une approche d'égalisation d'un système de reproduction ambisonique disposé dans une salle. Les réflexions de la salle ont été modélisées dans un formalisme en harmoniques sphériques. La formulation a mis en avant une matrice de distorsion qui traduit les transformations que subissent les composantes ambisoniques cibles une fois décodées dans la salle en utilisant un décodeur conçu en champ libre. La minimisation de l'erreur de reproduction entre les composantes ambisoniques cibles et celles reproduites a permis de concevoir une matrice de filtres de compensation des réflexions, agissant dans le domaine des harmoniques sphériques. Cette approche permet d'utiliser moins de filtres de compensation et de pouvoir utiliser un décodeur conçu pour une condition de champ libre, par rapport à la conception d'un décodeur adapté au lieu de restitution. Par ailleurs, il a été montré que le problème de minimisation est mal conditionné en basses-fréquences, et ce plus l'ordre de troncature est élevé. Une inversion régularisée par la décomposition en valeur singulière tronquée a été proposée. En utilisant la règle d'approximation de la taille de la zone de reproduction en fonction de la fréquence et de l'ordre de troncature, l'inversion ne requiert que les premiers ordres en basses fréquences, ce qui permet de mélanger les systèmes inversés à différents ordres de troncature en utilisant des filtres *crossover*. Des résultats de simulations dans les domaines temporel, spatial et fréquentiel ont montré l'efficacité de la méthode sur une zone centrée à l'origine. La minimisation de l'erreur de reconstruction ambisonique garantit une zone de reconstruction dans la salle équivalente à celle en condition de champ libre. Finalement, la méthode proposée peut permettre de concevoir des effets de réverbération en trois dimensions, en se basant sur la matrice de distorsion.

En ce qui concerne la thématique de transformation des champs sonores, une approche originale de filtrage directionnel de champs sonores a été présentée au chapitre 5. Ainsi, il a été montré qu'il est possible de filtrer le champ sonore pour amplifier ou atténuer certaines directions selon une fonction de directivité choisie, et ce indépendamment de la fréquence. La formulation proposée, basée dans le domaine des harmoniques sphériques, permet de manipuler le champ sonore indépendamment du dispositif de captation et de restitution. L'opération de filtrage a alors été exprimée à l'aide d'une opération matricielle sur les composantes ambisoniques. Elle agit dans le domaine angulaire de manière indépendante de la fréquence. Une attention particulière a été portée sur l'ordre de décomposition néces-

saire pour le champ sonore filtré. Ce dernier est la somme de l'ordre de décomposition de la fonction de directivité et celui du champ sonore original. Des résultats de simulations pour une fonction de directivité de type Dirac angulaire et hyper-cardioïde ont permis d'illustrer la formulation proposée. La mise en oeuvre du filtrage directionnel dans la suite de greffons *ambitools* a été utilisée lors de la campagne de mesure menée au cours de cette thèse et a permis de valider expérimentalement le filtrage directionnel.

La validation expérimentale des concepts présentés aux chapitres 3 et 5 a été faite lors d'une campagne de mesure en chambre anéchoïque, relatée au chapitre 6 de cette thèse. A cette fin, des dispositifs expérimentaux ont été réalisés. Il s'agit de deux réseaux sphériques de haut-parleurs de restitution ambisoniques « SpherBedev » et d'un réseau de microphones disposés sur une sphère rigide « MemsBedev ». Une suite de greffons, *ambitools*, mettant en oeuvre la majorité des concepts introduits dans cette thèse a également été développée pour la campagne de mesure expérimentale. La campagne de mesure a montré l'aptitude de captations, encodage et décodage ou synthèse d'un champ sonore jusqu'à l'ordre 5, en utilisant le maillage de Lebedev présenté au chapitre 3. Elle a également permis de valider expérimentalement le filtrage directionnel présenté au chapitre 5.

7.2 Perspectives

Plusieurs perspectives de recherches peuvent être dégagées à la lumière des résultats de ce doctorat. Tout d'abord, l'approche de décodage multi-ordres, multi-bandes pourrait être validée expérimentalement à l'aide du réseau de haut-parleurs SpherBedev. La reconstruction du champ sonore, notamment en dehors de la zone de reconstruction théorique est à considérer par rapport au cas d'un décodeur basique pleine bande. Dans ce même ordre d'idée, la spécificité de la grille de Lebedev d'avoir trois maillages imbriqués offrirait la possibilité de l'étude d'un décodage jusqu'à l'ordre 3 avec 26 et 50 haut-parleurs respectivement. La différence entre les deux situations serait en dehors de la zone de reconstruction théorique, du fait du repliement spatial différent selon les cas.

Ensuite, la validation expérimentale de la méthode de compensation des réflexions semblerait être la suite logique à ce travail doctoral. Certaines précautions sont cependant à prendre pour la mise en place d'une campagne de mesures en ce sens, notamment avec le capteur ambisonique utilisé. En effet, l'estimation précise des composantes ambisoniques du champ sonore est nécessaire pour ne corriger que les défauts du système de restitution, et non ceux de toute la chaîne de captation et de restitution. À ce jour, les premiers tests réalisés à l'aide du réseau de microphones MemsBedev et du réseau

de haut-parleurs SpherBedev demandent une régularisation trop importante pour obtenir des résultats satisfaisants sur une zone étendue de l'espace. Pour améliorer l'estimation des composantes ambisoniques, l'utilisation conjointe d'une configuration sphère ouverte autour d'une sphère rigide pourrait être une solution.

Par ailleurs, la mise en oeuvre d'effets de réverbération tridimensionnels, à partir de captations ambisoniques mériterait d'être explorée. Il serait alors possible de simuler la réverbération d'une salle dans une autre salle.

Une hypothèse forte de la compensation de salle proposée dans ce travail de doctorat est l'invariance de la réverbération dans la salle où se situe le dispositif de restitution. Dans la pratique, la réverbération peut varier sensiblement avec un déplacement mineur du mobilier dans la salle, avec la température ou encore la présence d'une antenne de microphones différente de celle utilisée pour capter l'effet de réverbération. Pour remédier à ce problème, une approche par contrôle actif serait alors envisageable, en s'appuyant sur la formulation proposée dans le domaine des harmoniques sphériques. Cette approche nécessiterait alors la présence permanente d'un capteur ambisonique pour mesurer l'erreur de reproduction et permettre la mise à jour des coefficients des filtres de compensation.

Finalement, le filtrage directionnel pourrait être couplé avec l'encodage des réflexions dans la salle en composantes ambisoniques. Il serait alors envisageable de ne corriger qu'une partie des réflexions, suivant leurs directions.

ANNEXE A

DIFFRACTION D'UNE SPHÈRE RIGIDE DANS UN CHAMP ACOUSTIQUE

Cette annexe détaille la diffraction d'une sphère rigide dans un champ acoustique. Il est alors possible d'utiliser ces propriétés de diffraction pour estimer les composantes ambisoniques d'un champ sonore, tel que présenté au chapitre 2, équation (2.28). Cette formulation est également mise en oeuvre dans le chapitre 6, section 6.3.5 pour le réseau sphérique rigide de microphones MemsBedev.

Soit une sphère rigide, de rayon r_{mic} , centrée dans un volume V exempt de sources. Le champ de pression acoustique peut alors s'écrire comme la somme du champ sans la sphère et du champ de pression acoustique diffracté par cette dernière [Morse et Ingard, 1986] :

$$p_t(k, r, \theta, \delta) = p_i(k, r, \theta, \delta) + p_d(k, r, \theta, \delta), \quad (\text{A.1})$$

où p_t représente la pression acoustique dans le volume V , p_i la pression acoustique incidente, sans la sphère, et p_d la pression acoustique diffractée par la sphère.

La décomposition de p_i et p_d sur une la base des harmoniques sphériques s'exprime alors :

$$p_i(k, r, \theta, \delta) = \sum_{m=0}^{+\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta), \quad (\text{A.2})$$

et

$$p_d(k, r, \theta, \delta) = \sum_{m=0}^{+\infty} i^m h_m^{(2)}(kr) \sum_{n=-m}^m C_{mn} Y_{mn}(\theta, \delta). \quad (\text{A.3})$$

Dans ces deux équations j_m représentent les fonctions de Bessel sphériques, traduisant le champ incident, et $h_m^{(2)}$ représentent les fonctions de Hankel sphériques de deuxième espèce, traduisant le champ diffracté, se propageant vers l'extérieur. Les coefficients B_{mn} et C_{mn} sont respectivement les coefficients de la TFS du champ incident p_i et du champ diffracté p_c . La pression p_t s'exprime alors :

$$p_t(k, r, \theta, \delta) = \sum_{m=0}^{+\infty} \sum_{n=-m}^m i^m Y_{mn}(\theta, \delta) (D_{mn} h_m^{(2)}(kr) + A_{mn} j_m(kr)) \quad (\text{A.4})$$

La condition de rigidité acoustique de la sphère se traduit par une vitesse acoustique radiale v_r nulle à sa surface $r = r_{\text{mic}}$:

$$v_r(r_{\text{mic}}) = 0 \quad (\text{A.5})$$

La vitesse acoustique est reliée à la pression acoustique par l'équation d'Euler

$$v_r = -\frac{1}{ik\rho_0 c} \frac{\partial p_t}{\partial r}. \quad (\text{A.6})$$

Ainsi, d'après l'équation (A.4) :

$$0 = \sum_{m=0}^{+\infty} \sum_{n=-m}^m \frac{i^{m+1} k}{\rho_0 c} Y_{mn}(\theta, \delta) (C_{mn} h_m^{(2)}(kr_{\text{mic}}) + B_{mn} j_m'(kr_{\text{mic}})) \quad (\text{A.7})$$

où $h_m^{(2)}$ et j_m' représentent les dérivées respectives des fonctions $h_m^{(2)}$ et j_m par rapport à la variable r . Les coefficients C_{mn} et B_{mn} sont alors reliés par :

$$C_{mn} = B_{mn} \frac{j_m'(kr_{\text{mic}})}{h_m^{(2)}(kr_{\text{mic}})}, \quad (\text{A.8})$$

Et la pression p_t s'écrit :

$$p_t(k, r, \theta, \delta) = \sum_{m=0}^{+\infty} i^m \left(j_m(kr) - \frac{j_m'(kr_{\text{mic}})}{h_m^{(2)}(kr_{\text{mic}})} h_m^{(2)}(kr) \right) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta) \quad (\text{A.9})$$

L'expression de la pression p_t à la surface de la sphère peut être simplifier en utilisant le Wronksien de l'équation (2.15) [Abramowitz et Stegun, 1972] :

$$p_t(kr_{\text{mic}}, \theta, \delta) = \sum_{m=0}^{+\infty} \frac{i^{m-1}}{(kr_{\text{mic}})^2 h_m^{(2)}(kr_{\text{mic}})} \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta). \quad (\text{A.10})$$

En effectuant la TFS de cette dernière équation, l'orthonormalité des harmoniques sphériques s'applique et les coefficients B_{mn} du champ sonore, ou coefficients ambisoniques, sont obtenus :

$$B_{mn} = i^{-(m-1)} (kr_{\text{mic}})^2 h_m^{(2)}(kr_{\text{mic}}) \langle p_t(k, r_{\text{mic}}, \theta, \delta), Y_{mn}(\theta, \delta) \rangle \quad (\text{A.11})$$

Ce résultat est alors utilisé à l'équation (2.28) du chapitre 2. Elle est également mise en oeuvre dans le chapitre 6, section 6.3.5) pour le réseau sphérique rigide de microphones MemsBedev.

ANNEXE B

FORMULATION SIMPLE SOURCE POUR UNE GÉOMÉTRIE SPHÉRIQUE

es résultats présentés dans cette annexe sont utilisés pour la synthèse de champ sonore utilisant une distribution continue de sources secondaires sur une géométrie sphérique. La formulation est alors équivalente pour l'approche ambisonique et la synthèse de front d'ondes dans ce cas, comme présenté au chapitre 2.

B.1 Formulation simple source

L'intégrale de Kirchhoff-Helmholtz de l'équation (2.49) requiert la connaissance de la distribution de pression acoustique et de la vitesse normale sur la surface ∂V entourant un volume V_i . Cependant, il est possible de trouver une formulation intégrale ne faisant intervenir que la fonction de Green en champ libre G_0 . Cette approche est dite par simple source [Copley, 2005; Fazi et Nelson, 2013; Williams, 1999]. Si on considère les domaines présentés à la figure B.1, l'idée est d'exprimer la pression $p(\mathbf{r})$ partout dans l'espace par l'équation suivante :

$$p(\mathbf{r}) = \int_{\partial V} \mu(\mathbf{r}_0) G_0(\mathbf{r}|\mathbf{r}_0) dS(\mathbf{r}_0). \quad (\text{B.1})$$

Pour trouver l'expression de $\mu(\mathbf{r}_0)$, un problème intérieur et un problème extérieur sont

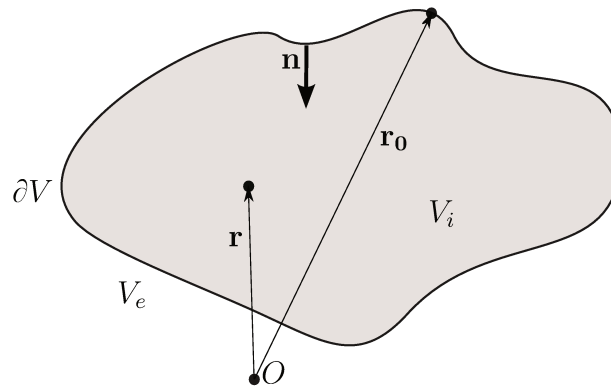


Figure B.1 Configuration du problème pour l'approche par simple source. Le volume V_e est le volume extérieur à la surface ∂V , le volume V_i est le volume intérieur à la surface ∂V . La normale intérieure à la surface ∂V est notée \mathbf{n} . Un point dans $V_i \cup V_e$ est de coordonnées $\mathbf{r} = (r, \theta, \delta)$ et un point sur la surface ∂V est de coordonnées $\mathbf{r}_0 = (r_0, \theta_0, \delta_0)$.

considérés séparément. Pour le problème intérieur, les sources acoustiques se situent dans le volume V_e , et la pression dans le volume V_i est notée $p_i(\mathbf{r})$. L'équation de Kirchhoff-

Helmholtz s'écrit alors :

$$\int_{\partial V} \left(p_i(\mathbf{r}_0) \frac{\partial G_0(\mathbf{r}|\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} - G_0(\mathbf{r}|\mathbf{r}_0) \frac{\partial p_i(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} \right) dS(\mathbf{r}_0) = \begin{cases} p_i(\mathbf{r}) & \mathbf{r} \in V_i \\ p_i(\mathbf{r})/2 & \mathbf{r} \in \partial V \\ 0 & \mathbf{r} \in V_e \end{cases} \quad (\text{B.2})$$

Pour le problème extérieur, les sources acoustiques se situent dans le volume V_i , et la pression à dans le volume V_e est notée $p_e(\mathbf{r})$. L'équation de Kirchoff-Helmholtz s'écrit alors :

$$\int_{\partial V} \left(p_e(\mathbf{r}_0) \frac{\partial G_0(\mathbf{r}|\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} - G_0(\mathbf{r}|\mathbf{r}_0) \frac{\partial p_e(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} \right) dS(\mathbf{r}_0) = \begin{cases} -p_e(\mathbf{r}) & \mathbf{r} \in V_e \\ -p_e(\mathbf{r})/2 & \mathbf{r} \in \partial V \\ 0 & \mathbf{r} \in V_i \end{cases} \quad (\text{B.3})$$

A présent, les deux problèmes sont reliés avec la condition de continuité des pressions suivante :

$$p_i(\mathbf{r}_0) = p_e(\mathbf{r}_0), \quad (\text{B.4})$$

et la discontinuité des dérivées normales $\frac{\partial p_i(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} \neq \frac{\partial p_e(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)}$ sur la surface ∂V . En faisant la différence entre l'équation (B.2) et l'équation (B.3) on obtient :

$$\int_{\partial V} \left(\frac{\partial p_e(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} - \frac{\partial p_i(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} \right) G_0(\mathbf{r}|\mathbf{r}_0) dS(\mathbf{r}_0) = \begin{cases} p_i(\mathbf{r}) & \mathbf{r} \in V_i \\ p_i(\mathbf{r}) = p_e(\mathbf{r}) & \mathbf{r} \in \partial V \\ p_e(\mathbf{r}) & \mathbf{r} \in V_e \end{cases} \quad (\text{B.5})$$

Finalement, en identifiant l'équation (B.5) avec l'équation (B.1), le terme $\mu(\mathbf{r}_0)$ s'exprime par :

$$\mu(\mathbf{r}_0) = \frac{\partial p_e(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} - \frac{\partial p_i(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} \quad (\text{B.6})$$

Ainsi, la densité $\mu(\mathbf{r}_0)$ est définie par la différence de dérivées normales de la pression définies pour le problème intérieur et extérieur. La formulation de l'équation (B.1) peut alors s'appliquer à un problème intérieur, ou extérieur : le champ de pression dans le volume V_i ou V_e , suivant le problème considéré, est dû à une distribution de sources monopolaires avec une densité $\mu(\mathbf{r}_0)$.

Connaissant le champ de pression $p(\mathbf{r})$ cible, l'étape suivante est alors de trouver la densité $\mu(\mathbf{r}_0)$ de façon explicite. Dans un contexte de reconstruction de champ sonore dans un volume fermé, le champ de pression cible est p_i et il faut alors trouver l'expression de $\frac{\partial p_e(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)}$ dans l'équation (B.6).

B.2 Application à une géométrie sphérique

Soit une surface sphérique ∂V , de rayon r_0 et de volume intérieur V_i . L'objectif est de trouver la densité $\mu(\mathbf{r}_0)$ dans l'équation (B.1), pour que la pression dans le volume V_i soit

égale à la pression cible $p(\mathbf{r})$:

$$p_i(\mathbf{r}) = p(\mathbf{r}), \quad \mathbf{r} \in V_i \cup \partial V. \quad (\text{B.7})$$

Pour cela, une approche possible est de considérer un problème de diffraction équivalent tel que proposé par Fazi [\[Fazi et Nelson, 2013\]](#). La sphère, vue comme un objet diffractant, respecte une condition de Dirichlet (pression nulle) sur la surface ∂V . Si l'on considère le champ incident noté $p(\mathbf{r})$, le champ acoustique total, noté $p_t(\mathbf{r})$, est la somme du champ incident et du champ diffracté par la sphère, noté $p_d(\mathbf{r})$:

$$p_t(\mathbf{r}) = p(\mathbf{r}) + p_d(\mathbf{r}). \quad (\text{B.8})$$

La condition de Dirichlet s'exprime par une pression nulle sur la surface ∂V , soit :

$$p_t(\mathbf{r}_0) = p(\mathbf{r}_0) + p_d(\mathbf{r}_0) = 0. \quad (\text{B.9})$$

Par ailleurs, en considérant le problème extérieur mentionné à la section précédente, le champ $p_e(\mathbf{r})$ est intégralement dû au champ diffracté par la sphère $p_d(\mathbf{r})$, par unicité de la solution du problème de Dirichlet [\[Colton et Kress, 2013a\]](#). Il respecte alors la condition de l'équation [\(B.9\)](#). De plus, pour respecter la formulation simple source, le champ $p_e(\mathbf{r})$ respecte également la condition de l'équation [\(B.4\)](#), avec $p_i(\mathbf{r}_0) = p(\mathbf{r}_0)$ d'après l'équation [\(B.7\)](#). Alors,

$$p_d(\mathbf{r}) = -p_e(\mathbf{r}) \quad \mathbf{r} \in V_e \cup \partial V. \quad (\text{B.10})$$

A présent, le champ incident $p(\mathbf{r})$ et le champ diffracté $p_d(\mathbf{r})$ sont décomposés sur la base des harmoniques sphériques [\[Williams, 1999\]](#) :

$$p(k, r, \theta, \delta) = \sum_{m=0}^{+\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta), \quad r \leq r_0. \quad (\text{B.11})$$

$$p_d(k, r, \theta, \delta) = \sum_{m=0}^{+\infty} i^m h_m^{(2)}(kr) \sum_{n=-m}^m C_{mn} Y_{mn}(\theta, \delta), \quad r \geq r_0. \quad (\text{B.12})$$

La condition de Dirichlet de l'équation [\(B.9\)](#) permet de trouver la relation entre les coefficients B_{mn} et C_{mn} des décompositions :

$$\sum_{m=0}^{+\infty} i^m j_m(kr_0) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta) = - \sum_{m=0}^{+\infty} i^m h_m^{(2)}(kr_0) \sum_{n=-m}^m C_{mn} Y_{mn}(\theta, \delta). \quad (\text{B.13})$$

En appliquant une TFS aux deux membres de l'égalité précédente, le critère d'orthonormalité des harmoniques sphériques s'applique et on obtient :

$$C_{mn} = - \frac{j_m(kr_0)}{h_m^{(2)}(kr_0)} B_{mn}. \quad (\text{B.14})$$

La pression totale de l'équation (B.8) s'exprime alors :

$$p_t(r, \theta, \delta) = \sum_{m=0}^{+\infty} i^m \left(j_m(kr) - \frac{j_m(kr_0)}{h_m^{(2)}(kr_0)} h_m^{(2)}(kr) \right) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta), \quad (\text{B.15})$$

et sa dérivée normale, par rapport à la variable r , est donnée par :

$$\frac{\partial p_t(r, \theta, \delta)}{\partial r} = \sum_{m=0}^{+\infty} i^m k \left(j'_m(kr) - \frac{j_m(kr_0)}{h_m^{(2)}(kr_0)} h_m^{(2)'}(kr) \right) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta). \quad (\text{B.16})$$

Pour $r = r_0$, la dérivée normale $\frac{\partial p_t(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)}$ peut se calculer en utilisant le Wronksien de l'équation (2.16).

$$\frac{\partial p_t(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} = - \sum_{m=0}^{+\infty} \frac{i^{m+1}}{h_m^{(2)}(kr_0) kr_0^2} \sum_{n=-m}^m B_{mn} Y_{mn}(\theta_0, \delta_0) \quad (\text{B.17})$$

À noter le signe négatif dû au sens de la normale $\mathbf{n}(\mathbf{r}_0)$ qui pointe vers l'intérieur de ∂V sur la figure B.1

Finalement, sachant que $p_t(\mathbf{r}) = p_d(\mathbf{r}) + p(\mathbf{r})$ d'après l'équation (B.8), et en utilisant les équations (B.7) et (B.10), on obtient :

$$\frac{\partial p_t(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} = - \frac{\partial p_e(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} + \frac{\partial p_i(\mathbf{r}_0)}{\partial \mathbf{n}(\mathbf{r}_0)} \quad (\text{B.18})$$

Ainsi, en identifiant avec l'équation (B.6), la densité $\mu_0(\mathbf{r}_0)$ est donné par :

$$\mu_0(\mathbf{r}_0) = \sum_{m=0}^{+\infty} \frac{i^{m+1}}{h_m(kr_0) kr_0^2} \sum_{n=-m}^m B_{mn} Y_{mn}(\theta_0, \delta_0) \quad (\text{B.19})$$

En conclusion, pour une géométrie sphérique de rayon r_0 , l'approche par simple source permet d'exprimer la pression pour tout point \mathbf{r} , à l'intérieur du volume V_i , à l'aide d'une distribution continue de sources monopolaires telles que :

$$p(\mathbf{r}) = \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} \left(\sum_{m=0}^{+\infty} \sum_{n=-m}^m \frac{i^{m+1}}{h_m(kr_0) kr_0^2} B_{mn} Y_{mn}(\theta_0, \delta_0) \right) G_0(\mathbf{r}|\mathbf{r}_0) r_0^2 \cos(\delta_0) d\theta_0 d\delta_0, \quad (\text{B.20})$$

Cette dernière équation peut se reformuler avec le produit scalaire sphérique sur la sphère unité (équation (2.2)) et les filtres de champs proches de l'équation (2.23), $F_m(kr_0) =$

$i^{-(m+1)}kh_m^{(2)}(kr_0)/(4\pi)$:

$$p(\mathbf{r}) = \left\langle \sum_{m=0}^{+\infty} \sum_{n=-m}^m \frac{B_{mn}Y_{mn}(\theta_0, \delta_0)}{F_m(kr_0)}, G_0(\mathbf{r}|\mathbf{r}_0) \right\rangle \quad (\text{B.21})$$

On retrouve alors la formulation de l'ambisonie, présentée au chapitre 2 et celle de la synthèse de front d'ondes (sans critère de sélection des sources secondaires) pour une géométrie sphérique.

ANNEXE C

SYNTHÈSE TEMPS-RÉEL DE CHAMPS SONORES EN TROIS DIMENSIONS EN UTILISANT FAUST, PROCESSING, PURE DATA ET OSC

C.1 Avant-propos

Cet article de conférence détaille l'implémentation des équations de l'ambisonie dans une collection d'outils de synthèse en temps-réel, nommée par la suite *ambitools*. Ces outils sont alors utilisés notamment dans le chapitre [6](#) pour la validation expérimentale des concepts et techniques présentés aux chapitres [2](#) et [5](#).

Auteurs et affiliations :

- Pierre Lecomte, Candidat au doctorat, Université de Sherbrooke, Conservatoire National des Arts et Métiers.
- Philippe-Aubert Gauthier, Professionnel de recherche, Université de Sherbrooke.

Conférence Digital Audio Effects 2015, (DAFx-15), Trondheim, Norvège

Date 30 Novembre 2015 - 3 Décembre 2015

Référence P. Lecomte and P.-A. Gauthier, "Real-Time 3D Ambisonics using Faust, Processing, Pure Data, And OSC," in 15th International Conference on Digital Audio Effects (DAFx-15), 2015.

Résumé français Cet article présente plusieurs outils de traitement du signal numérique pour la synthèse en temps-réelle d'un champ sonore en trois dimensions par approches ambisoniques. La spatialisation d'un signal monophonique où la reconstruction d'un champ sonore tridimensionnelle capté sont considérés. Les outils qui calculent les signaux de commandes des hauts-parleurs sont développés dans le langage FAUST¹. Ce langage permet de compiler les outils proposés suivant différents formats de greffons pour plusieurs architectures. Dans cet article, une attention particulière est faite sur l'implémentation des

¹<http://faust.grame.fr>

filtres de champs proches qui permettent l'encodage d'ondes sphériques. La variation de gain avec la distance est également prise en compte. Le contrôle de la synthèse peut être fait avec l'interface graphique utilisateur, ou bien avec une manette de contrôle, à l'aide de PURE DATA² et du protocole OPEN SOUND CONTROL³ (OSC). Un outil visuel, développé en utilisant le langage PROCESSING⁴ est également présenté.

Les pages suivantes sont la reproduction de la version éditée de l'article.

²<https://puredata.info/>

³<http://opensoundcontrol.org/>

⁴<https://processing.org/>

REAL-TIME 3D AMBISONICS USING FAUST, PROCESSING, PURE DATA, AND OSC

Pierre Lecomte,

Lab. de Mécanique des Structures et Systèmes Couplés
(LMSSC), Conservatoire National des Arts et Métiers
Paris, France
Groupe d'Acoustique de l'Université de Sherbrooke
(GAUS), Université de Sherbrooke, Québec, Canada
pierre.lecomte@gadz.org

Philippe-Aubert Gauthier,

GAUS
Université de Sherbrooke, Québec, Canada

ABSTRACT

This paper presents several digital signal processing (DSP) tools for the real-time synthesis of a 3D sound pressure field using Ambisonics technologies. The spatialization of monophonic signal or the reconstruction of natural 3D recorded sound pressure fields is considered. The DSP required to generate the loudspeaker signals is implemented using the FAUST programming language. FAUST enables and simplifies the compilation of the developed tools on several architectures and on different DSP tool format. In this paper, a focus is made on the near-field filters implementation which allows for the encoding of spherical waves with distance information. The gain variation with distance is also taken into account. The control of the synthesis can be made by software controllers or hardware controllers, such as joystick, by the mean of PURE DATA and OPEN SOUND CONTROL (OSC) messages. A visual feedback tool using the PROCESSING programming language is also presented in the most recent implementation. The aim of this research derives from a larger research project on physically-accurate sound field reproduction for simulators in engineering and industrial applications.

1. INTRODUCTION

Ambisonics technologies allow describing 3D sound pressure fields using a projection on a truncated spherical harmonics basis [1]. The resulting 3D encoded sound pressure field can later be manipulated, decoded and reconstructed over several loudspeaker-layouts or even headphones [2]. Ambisonics has two main objectives: the spatialization of virtual sound sources or the reconstruction of recorded sound pressure fields [3]. Several software solutions exist to create, transmit, manipulate, and render sound pressure fields using Ambisonics technologies. See references [4, 5, 6, 7, 8] as examples. Albeit being popular for practical applications in music, sound design and audio context, classical and common Ambisonics implementations suffer from few drawbacks that limit their use for physically-accurate sound field reproduction with applications to environment simulators (vehicles, working environments, etc.) in industrial or engineering context. Indeed, the near-field encoding [2] is rarely provided and the encoding/decoding in 3D context is limited to the first orders, hence limiting the spatial resolution and area size of physically-accurate reproduction. Indeed, if the sound field is controlled up to an order M , the reconstruction area size is frequency-dependent and given by $r = Mc/2\pi$ [9] (where r is the area size radius, c the speed of sound in air, and f the frequency). The near-field support is also

detrimental for physically-accurate sound field reproduction as it takes into account the loudspeaker distance from the origin in order to compensate for the loudspeakers spherical waves. In this trend, this work is motivated by the need to develop a practical implementation of Ambisonics for industrial applications such as laboratory reproduction of industrial sound environments, vehicles cabin noise, virtual vibroacoustics models, architectural spaces, etc. In these scenarios, the reproduced sound field must be as close as possible than the target sound field. Some recent examples of such applications potentials are found in Refs. [10, 11, 12, 13]. Typical outcomes are related to listening tests, sound quality testing, perceptual studies and other. On this matter, as mentioned by Vorländer [12] with respect to Ambisonics implementations that often include modifications inspired by auditory perception to increase the listener experience with respect to some expectations, a "generally applicable reproduction system [for sound field simulation] must not introduce any artificial auditory cue which is not part of the simulated sound." [12].

In this context, this paper presents an implementation of Ambisonics technologies for real-time synthesis of 3D sound field up to 5th order. The signal processing is implemented in FAUST¹ (Functional Audio Stream) programming language. This language proposes a functional approach to implement the signal processing and it compiles in efficient C++ code [14]. From this code, DSP tools are provided in several formats such as: VST, LV2, Pure Data, Max/MSP, JACK QT, and others. Thus, from the same FAUST code, one can generate tools working on various operating systems and configurations.

The focus of this paper is on physical encoding and decoding of 3D sound pressure fields with a special attention dedicated to the near field filters development, definition and implementation. After defining the notations in use in Sec. 2, the main equations of Ambisonics are recalled in Sec. 3. In Sec. 4 the implementation in FAUST programming language is addressed with a special attention on near-field filters. Section 5 presents a visual feedback tool using PROCESSING language. This tool helps visualizing in 3D the virtual sources and the loudspeaker levels. Finally, in Sec. 6, the user control interface is addressed, presenting the possibility of interfacing all elements with OSC protocol.

¹<http://faust.grame.fr/>

2. COORDINATE SYSTEM AND NOTATIONS

In this section, the different notations used throughout the paper are presented and illustrated.

Spherical coordinate system

The following spherical coordinate system is used throughout this paper and shown in Fig. 1:

$$u_1 = r \cos(\theta) \cos(\delta), \quad u_2 = r \sin(\theta) \cos(\delta), \quad u_3 = r \sin(\delta) \quad (1)$$

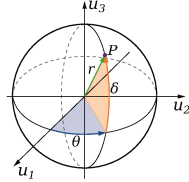


Figure 1: Spherical coordinate system. A point $P(u_1, u_2, u_3)$ is described by its radius r , azimuth θ and elevation δ .

A virtual source position is denoted with its spherical coordinates r_1, θ_1, δ_1 . The rendering loudspeakers are arranged on a sphere of radius r_0 , as shown in Fig. 2.

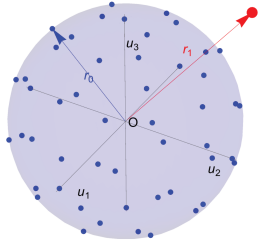


Figure 2: Ambisonics playback layout. Blue: the rendering loudspeakers disposed on a spherical layout of radius r_0 . Red: the virtual source at radius position r_1 .

Spherical harmonics

The spherical harmonics used in this paper are defined as follow in [1]:

$$Y_{mn}^\sigma(\theta, \delta) = \sqrt{(2m+1)\epsilon_n \frac{(m-n)!}{(m+n)!}} P_{mn}(\sin(\delta)) \times \begin{cases} \cos(n\theta) & \text{if } \sigma = 1 \\ \sin(n\theta) & \text{if } \sigma = -1 \end{cases} \quad (2)$$

where P_{mn} are the associated Legendre functions of order m and degree n , m and $n \in \mathbb{N}$ with $n \leq m$, $\sigma = \pm 1$ and $\epsilon_n = 1$ if $n = 0$, $\epsilon_n = 2$ if $n > 0$. The spherical harmonics order are denoted by m and its degree, by n . For each order m there are $(2m+1)$ spherical harmonics. Thus a basis truncated at order M contains $L = (M+1)^2$ functions.

Notations

The Laplace variable is denoted s and the discrete time variable z (discrete domain). A vector is denoted by lowercase bold font \mathbf{v} and a matrix by uppercase bold font \mathbf{M} . Superscript T designates the transposition operation. j is imaginary unit with $j^2 = -1$.

3. AMBISONICS

In this section, the principal Ambisonics equations are recalled. They will later be used for the real-time DSP implementation.

3.1. Encoding

In Ambisonics, the encoding step consists in deriving *B-Format* signals² from either monophonic signal (with a spatialization context) or microphone array signals (natural sound field encoding).

3.1.1. Monophonic signal

From a monophonic signal, the encoding can be done for a plane wave with amplitude $a(z)$, azimuth θ_1 and elevation δ_1 direction or a spherical wave, adding a distance information r_1

$$\begin{aligned} B_{mn}^\sigma(z) &= a(z) Y_{mn}^\sigma(\theta_1, \delta_1) && \text{Plane wave} \\ B_{mn}^\sigma(z) &= a(z) F_{m,r_1}(z) Y_{mn}^\sigma(\theta_1, \delta_1) && \text{Spherical wave} \end{aligned} \quad (3)$$

In Eq. 3, filters $F_{m,r_1}(z)$ are the forward near-field filters which take into account the finite distance of the virtual source r_1 [2].

3.1.2. Rigid spherical microphone array encoding

For a natural 3D sound pressure field recording made with a rigid spherical microphone array of radius r_a , the B_{mn}^σ are given by:

$$B_{mn}^\sigma(z) = E_{m,r_a}(z) \sum_{i=1}^N Y_{mn}^\sigma(\theta_i, \delta_i) w_i p_i(z) \quad (4)$$

$E_{m,r_a}(z)$ are the equalization filters which take into account the diffraction of the rigid sphere [3]. The sound pressure signal at the i^{th} capsule position $(r_a, \theta_i, \delta_i)$ is denoted $p_i(z)$ on the array of N microphones. The B_{mn}^σ components are guaranteed to be exact up to order M if the spatial sampling of the spherical microphone array respects the orthonormality criterion of spherical harmonics up to M [15, 16]. Thus, there is possibly a weighting factor w_i for each capsule in Eq. (4) to ensure this condition. The working bandwidth of the array without aliasing is given by: $f \leq Mc/(r_a 2\pi)$ [17], where f is the frequency, c the sound speed, and M the maximum working order for the array.

Equation (4) is for a triplet of indices (m, n, σ) . Thus, up to order M , the *B-Format* signals vector is obtained with matrix notation:

$$\mathbf{b}(z) = \mathbf{E}(z) \cdot \mathbf{Y}_{\text{mic}}^T \cdot \mathbf{W}_{\text{mic}} \cdot \mathbf{p}(z) \quad (5)$$

$\mathbf{b}(z)^{(L \times 1)} = [B_{00}^1(z) \cdots B_{mn}^\sigma(z) \cdots B_{M0}^1(z)]$. $\mathbf{E}(z)^{(L \times L)}$ is the diagonal matrix of equalization filters with diagonal terms $[E_{0,r_a}(z) \cdots E_{m,r_a}(z) \cdots E_{M,r_a}(z)]$, each $E_{m,r_a}(z)$ term being replicated $(2m+1)$ times on the main diagonal. $\mathbf{Y}_{\text{mic}}^{(N \times L)}$ is the matrix of spherical harmonics up to order M evaluated at each direction (θ_i, δ_i) . $\mathbf{W}_{\text{mic}}^{(N \times N)}$ is the diagonal matrix of weightings.

²We call here the *B-Format* the signals vector $\mathbf{b}(z)$ $[B_{00}^1(z) \cdots B_{mn}^\sigma(z) \cdots B_{M0}^1(z)]$

$\mathbf{p}^{N \times 1} = [p_1(z) \cdots p_i(z) \cdots p_N(z)]$ is the vector of capsule signals.

3.2. Decoding

In this paper, only the basic decoder (mode-matching [18]) is recalled for a full-sphere configuration of radius r_0 respecting the spherical harmonics orthonormality via weighting coefficients [16]. For decoders adapted to irregular loudspeaker layout or other decoding concerns, see for example [6].

If one considers a set of N_2 loudspeakers, the input signal of the l^{th} loudspeaker at position $(r_0, \theta_l, \delta_l)$ is obtained from the B_{mn}^σ signals by:

$$s_l(z) = w_l \sum_{m=0}^M (F_{m,r_0})^{-1}(z) \sum_{n=0}^m \sum_{\sigma=\pm 1} Y_{mn}^\sigma(\theta_l, \delta_l) B_{mn}^\sigma(z) \quad (6)$$

where $(F_{m,r_0})^{-1}(z)$ are the inverse near field filters or near field compensation filters, which take into account the finite distance r_0 of the rendering loudspeakers [2]. Since the reconstructed sound pressure field is the summation of each sound field generated by each loudspeakers, the vector of input signals \mathbf{s} is given by:

$$\mathbf{s} = \mathbf{W}_{\text{spk}} \cdot \mathbf{Y}_{\text{spk}} \cdot \mathbf{F}_{r_0}^{-1}(z) \cdot \mathbf{b}(z) \quad (7)$$

where $\mathbf{s}(z)^{(N_2 \times 1)} = [s_1(z) \cdots s_l(z) \cdots s_{N_2}(z)]$, $\mathbf{F}_{m,r_0}^{-1}(L \times L)(z)$ is the diagonal matrix of near field compensation filters with diagonal terms $[1/F_{0,r_0}(z) \cdots 1/F_{m,r_0}(z) \cdots 1/F_{M,r_0}(z)]$, each $1/F_{m,r_0}(z)$ term being replicated $(2m+1)$ times on the main diagonal. $\mathbf{Y}_{\text{spk}}^{(N_2 \times L)}$ is the matrix of L spherical harmonics up to order M evaluated at each direction (θ_l, δ_l) . $\mathbf{W}_{\text{spk}}^{(N_2 \times N_2)}$ is the diagonal matrix of weightings w_l . Note that the resulting matrix $\mathbf{M}_{\text{spk}} = \mathbf{W}_{\text{spk}} \cdot \mathbf{Y}_{\text{spk}}$ is the Ambisonics decoding matrix as defined in Ref.[6].

3.3. Equivalent panning-law

In a spatialization context, recalling Eq. (3) and Eq. (6) along with the additivity theorem of spherical harmonics [19], one can directly compute the loudspeaker input signals:

$$s_l(z) = a(z) w_l \sum_{m=0}^M \frac{(2m+1)}{F_{m,r_0}(z)} P_m(\gamma_l) \quad \text{Plane wave}$$

$$s_l(z) = a(z) w_l \sum_{m=0}^M (2m+1) \frac{F_{m,r_1}(z)}{F_{m,r_0}(z)} P_m(\gamma_l) \quad \text{Spherical wave} \quad (8)$$

where $\gamma_l = \cos(\delta_1) \cos(\delta_l) \cos(\theta_1 - \theta_l) + \sin(\delta_1) \sin(\delta_l)$ is relative to the angle between the virtual source in θ_1, δ_1 and the l^{th} loudspeaker.

4. FAUST IMPLEMENTATION

The implementation of Ambisonics tools such as an encoder, basic decoder, near-field filters, and spatialization tools using equivalent panning-law is made using the FAUST language. An overview of the current developed tools is shown in Fig.3. From left to right, the vertical branches correspond to: 1) microphone signal processing, 2) panning of a monophonic signal as virtual point source with encoding and decoding and 3) panning of a monophonic signal as

virtual point source with equivalent panning law. The rightmost branch is the control via OSC and visual feedback. Each box corresponds to a module described in the next sections. The common functions of each modules are implemented in library files, which enable the quick design of new modules by re-using existing FAUST code. In the following section, the near-field filters implementation is detailed.

4.1. Near-field filters

4.1.1. Forward filters

The forward filters are given in the Laplace domain [2]:

$$F_{m,r_1}(s) = \sum_{i=0}^m \frac{(m+i)!}{(m-i)!i!2^i} \left(\frac{c}{sr_1} \right)^i \quad (9)$$

with $s = j2\pi f$. To convert this analog filter in the digital z domain, the use of a bilinear transform requires precision arithmetic to be efficient, as pointed out by Adriaensen in [20]. In this latter reference, he proposes another s -to- z mapping scheme to obtain a digital realization of the filter which is robust with single precision floating point format:

$$\zeta^{-1} = \frac{z^{-1}}{1-z^{-1}} = \sum_{k=1}^{\infty} (z^{-1})^k \quad \text{for } |z| > 1 \quad (10)$$

$$s = \frac{2F_s}{1+2\zeta^{-1}}$$

where F_s is the sampling frequency. The implementation of an m^{th} order forward filter is made by product of sections of the form:

$$H_{1,r_1}(z) = g_1(r_1)(1 + d_{1,1}(r_1)\zeta^{-1})$$

$$H_{2,r_1}(z) = g_2(r_1)(1 + d_{2,1}(r_1)\zeta^{-1} + d_{2,2}(r_1)\zeta^{-2}) \quad (11)$$

For example a 3rd order filter F_{3,r_1} is realized with the product of a 1st order section and a 2nd order section: $F_{3,r_1}(z) = H_{1,r_1}(z) \cdot H_{2,r_1}(z)$. The computation of each coefficient $g_1, g_2, d_{1,1}, d_{2,1}, d_{2,2}$ in Eq. (11) is detailed in [20].

4.1.2. Stabilization with inverse filters

The forward filters $F_{m,r_1}(z)$ present an infinite gain at 0 Hz. Thus, they must be stabilized by multiplying with an inverse filter (or near field compensation filters) $1/F_{m,r_0}(z)$. This can be done at the encoding stage as suggested by Daniel [2]:

$$B_{mn}^\sigma(z) = a(z) \frac{F_{m,r_1}(z)}{F_{m,r_0}(z)} Y_{mn}^\sigma(\theta_1, \delta_1) \quad \text{Spherical wave} \quad (12)$$

It means that to encode a spherical wave, according to Eq. (12), one should know the rendering array radius r_0 a priori. However this is not a major concern. Indeed, if the encoded spherical wave is reconstituted on another layout of radius r_2 , one can correct by multiplying the $B_{mn}^\sigma(z)$ components by $F_{m,r_0}(z)/F_{m,r_2}(z)$.

4.1.3. Gain correction for spherical source

In [2], the near field filters are defined with the assumption that the amplitude $a(z)$ of the spherical wave in Eq. (12) is taken at the origin O in Fig. 2. Thus, the propagation term $e^{sr_1/c}/(4\pi r_1)$

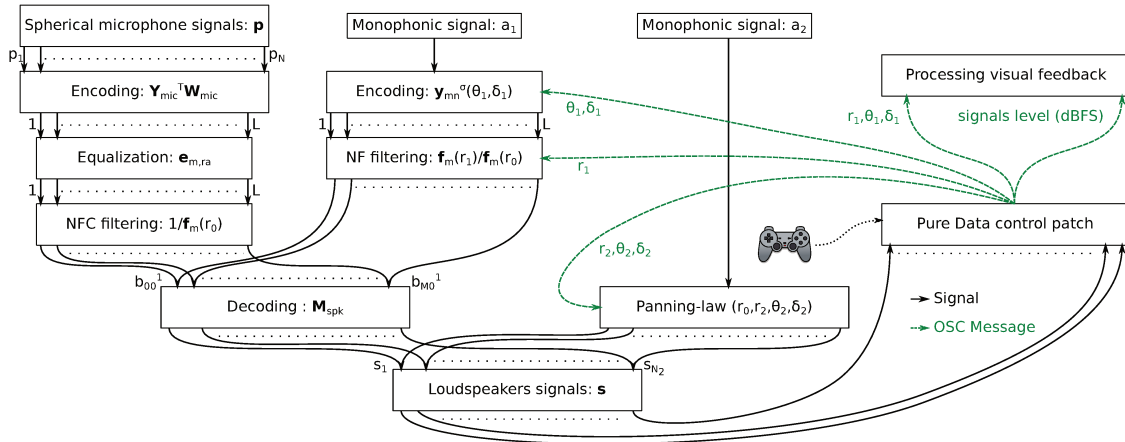


Figure 3: Main features of the current implementation: real-time synthesis of 3D recordings (leftmost vertical branch), spatialization (two centered vertical branches), and control via OSC with visual feedback (rightmost vertical branch). The digital temporal signals are represented by black arrows and the OSC instructions message in green dashed arrows.

(Laplace's domain) is not taken into account. This leads to the definition of Eq. (9) for the forward filters. In the same way, the near-field compensation filters do not take into account the propagation from rendering source to origin O : $e^{sr_0/c}/(4\pi r_0)$. The main consequence is that there is no gain variation with the distance of the virtual source. To correct this, one multiplies the amplitude $a(z)$ for spherical wave in Eq. (12) by: $\frac{r_0}{r_1} z^{-(r_1-r_0)/c}$. The corresponding delay can be fractional depending on r_1 and r_0 . In the case of a focused source (i.e. inside the rendering loudspeakers enclosure), the delay is negative and an implementation could require time-reversal method, as in Wave Field Synthesis [21]. However, since this delay is the same for all rendering loudspeakers, according to Eqs. (6) and (8), it is omitted for simplicity. As a result, the pressure sound field of a virtual point source will be reproduced physically in the reproduction zone with correct gain and with a phase shift of $z^{-(r_1-r_0)/c}$ when this latter is omitted. Finally, the encoding and the panning-law of a spherical wave becomes:

$$B_{mn}^\sigma(z) = a(z) \frac{r_0}{r_1} \frac{F_{m,r_1}}{F_{m,r_0}}(z) Y_{mn}^\sigma(\theta_1, \delta_1) \quad (13)$$

$$s_l = a(z) \frac{r_0}{r_1} w_l \sum_{m=0}^M (2m+1) \frac{F_{m,r_1}}{F_{m,r_0}}(z) P_m(\gamma_l)$$

4.1.4. FAUST implementation

The near field filters $F_{m,r_1}(z)/F_{m,r_0}(z)$ (as well as the near field compensation filters $1/F_{m,r_0}(z)$) are implemented in a FAUST library `nfc.lib` following Eq. (11) and up to 5th order. The implementation is based on a Direct-Form II [22]. As an example, the FAUST code for the second order section is given as:

```
TFNF2(b0, b1, b2, a1, a2) =
sub~sum1(a1, a2): sum2(b0, b1, b2)
with {
sum1(k1, k2, x)=
x:(+~_<:((_' :+~_), *(k1)) :*(k2), _ :+);
sum2(k0, k1, k2, x)=
x<:*(k0), +~_ , _ :_, (-<:*(k1), (_' :+~_))
*(k2) :+ :+;
```

$$\text{sub}(x, y) = y - x;$$

};

The block diagram for this code is shown in Fig. 4. The coefficients $g_2(r_1), d_{2,1}(r_1), d_{2,2}(r_1), g_2(r_0), d_{2,1}(r_0), d_{2,2}(r_0)$ were precomputed in this figure for simplicity to obtain the corresponding $b_0(r_1), b_1(r_1), b_2(r_1), a_1(r_0), a_2(r_0)$ coefficients, poles and zeroes of the filter. However, `nfc.lib` provides the real-time computation of these coefficients knowing r_1 and r_0 , according to [20]. In the current implementation, the near-field filters are provided as independent modules (see Fig. 3), or included in the equivalent panning-law module (see Sec. 4.4)

As an example, the gain frequency response of the filters $r_0/r_1 \times F_{m,r_1}/F_{m,r_0}$ up to order five are shown in Fig. 5 for $r_1 = 1$ m, $r_0 = 3$ m (solid lines) and $r_1 = 3$ m, $r_0 = 1$ m (dashed lines). $c = 340$ m/s and $F_s = 48000$ Hz. For focused sources (i.e. $r_1 \leq r_0$, solid lines in Fig. 5), as r_1 decreases, or as r_0 increases and as m increases, the gain of the filters increases at low frequencies. Thus, when encoding a focused spherical source (i.e. $r_1 \leq r_0$), one should be aware of these extreme gains. These "bass-boost" effects create strong artifacts outside the control zone and can easily damage the rendering loudspeakers. A solution could be to impose a minimum r_1 regarding to maximum a $r_0/r_1 \times F_{m,r_1}/F_{m,r_0}$ gain. This maximum gain is then related to the maximum linear excursion of the loudspeakers. Note that another approach for focused sources in Ambisonics can as well be a solution [23].

4.2. Encoding of captured sound field and decoding

This case corresponds to the leftmost vertical branch of Fig. 3. The encoding of a 3D sound field captured by a spherical microphone array as described in Eq. (5) requires a matrix operation as the decoding operation of Eq. (7). This matrix operation is done in FAUST as suggested by Heller [6]:

```
// bus with gains
gain(c) = R(c) with {
R((c, c1)) = R(c), R(c1);
R(1) = _;
```

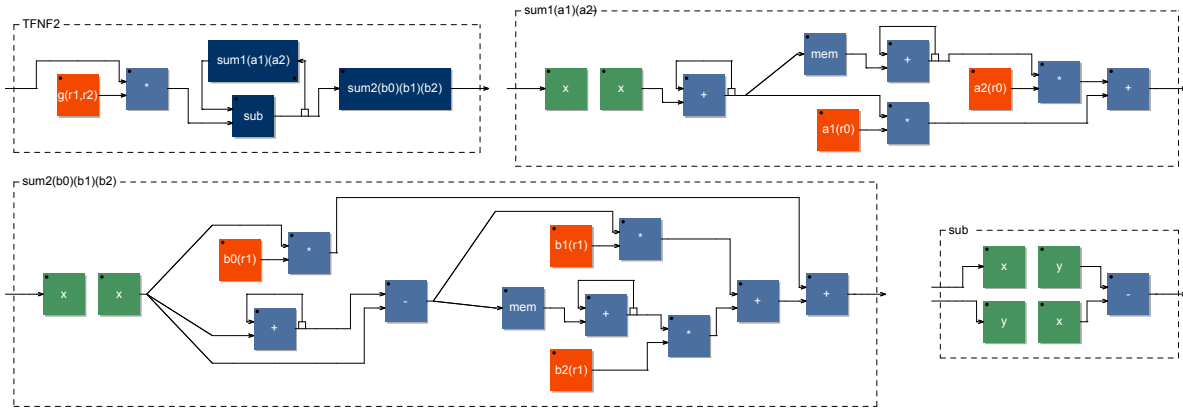


Figure 4: Block diagram representation of the TFNF2 function. The input signal to be filtered is on the main top left diagram. The others diagrams detail each block in this main diagram. x and y are input signals in a block. These diagrams were generated using `faust2svg` tool.

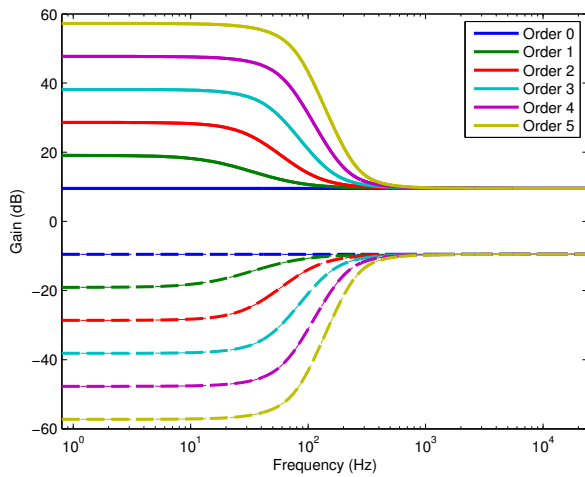


Figure 5: Gain of the frequency response function of filters $\frac{r_0}{r_1} \frac{F_{m,r_1}}{F_{m,r_0}}(f)$ for $r_1 = 1$ m, $r_0 = 3$ m (solid lines) and $r_1 = 3$ m, $r_0 = 1$ m (dashed lines), $m \in \{0, 1, 2, 3, 4, 5\}$, $F_s = 48000$ Hz.

```

R(0)      = !;
R(float(0)) = R(0);
R(float(1)) = R(1);
R(c)      = *(c);
};

matrix(n,m) = par(i,n,_);
<: par(i,m,gain(a(i))>:_);
// n: number of inputs : column number
// m: number of output : row number
// a(i): row vectors of matrix coefficients

```

In Eqs. (5) and (7) the matrix $\mathbf{Y}_{mic}^T \cdot \mathbf{W}_{mic}$ and $\mathbf{W}_{spk} \cdot \mathbf{Y}_{spk}$ are pre-computed and then implemented numerically in the FAUST code, row by row, according to the code above. In the current version,

the encoding matrix for 2 types of microphone is coded:

- Spherical microphone using Lebedev's grids working up to 1st, 3rd or 5th order as shown in Fig. 9.
- Spherical microphone using Pentakis-Dodecahedron grid working up to 4th order as shown in [3].

In Fig. 3, the obtained modules are sketched under the names "Encoding" and "Decoding".

4.3. Rigid spherical microphone filters

As mentioned in Sec. 3.1.2, the $B_{mn}^\sigma(z)$ components derived from a rigid spherical microphone array signals should be filtered by $E_m(z)$ filters to take into account the diffraction of the rigid sphere supporting the capsules. These filters present a theoretical infinite gain at 0 Hz and very large "bass-boost" for high orders [24]. They are stabilized by high-pass filters [25] or Tikhonov filters [3], which cut the low frequencies at high orders. Resulting filters are implemented as FIR filters. However, in [26, 27, 24] IIR implementations are proposed with lower orders amplification according to higher orders limitation. These approaches should be investigated in future works for a FAUST implementation since an IIR parametric filters could be interesting to monitor in real-time the performances of a spherical microphone array. For now, in the reported implementation, the filters are implemented as FIR filters. The cut-off frequencies of high-pass filters are chosen with a maximum amplification level. The real-time convolution is made with BRUTEFIR³ for Linux environment. This module is sketched in Fig. 3 under the name "Equalization".

4.4. Encoding of virtual source and panning law

In the FAUST implementation reported in this paper, the encoding of a virtual source is based on Eq. (3) for a plane wave and Eq. (13) for a spherical wave. In the current state of the implementation, the spherical harmonics are explicitly coded in a `ymn.lib` library up to order five (36 functions). However, they could be computed by recurrence if higher orders would be required. The monophonic

³<http://www.ludd.luth.se/~torger/brutefir.html>

signal a_1 is multiplied by corresponding spherical harmonics evaluated at desired direction (θ_1, δ_1) and filtered by near-field filters with desired radii (r_1, r_0) as shown in Fig. 3: Ambisonics signals are then obtained.

The equivalent panning laws of Eq. (8) circumvent the need to encode and decode with matrix operations. Indeed, the computation is reduced to a sum on the orders thanks to the additivity theorem of spherical harmonics. This is of great interest for real-time synthesis. For the moment, the panning-laws are implemented for several Lebedev spheres using $N = 6, 26,$ or 50 loudspeakers, as presented in [16]. The Legendre polynomials P_m are explicitly coded in `ymn.lib` up to order five. The weights w_l and angles θ_l, δ_l of the spheres are implemented in a `lebedev.lib` file. This spatialization module is sketched in Fig. 3 under the name "Panning-law": the loudspeakers input signals are computed from a monophonic signal a_2 and spatial parameters $(r_2, \theta_2, \delta_2, r_0)$.

5. VISUAL FEEDBACK WITH PROCESSING

FAUST provides a graphical user interface with VU-Meters. Unfortunately, those meters are organized as classical lines of VU-Meters (see Fig. 7 as an example). Thus, they do not provide an efficient visual cue of where the signal energy is distributed on the loudspeaker spherical layout. Moreover, a 3D representation of the virtual source helps to provide an objective spatial view of the source position. Such type of 3D representation is also much more useful to composers or engineers who are not familiar with under-the-hood DSP. In this context, a visual tool for 3D visual feedback was implemented using PROCESSING⁴ language. The code uses an `Atom` class to create a ball representing a loudspeaker. The loudspeakers coordinates are given in a `.cvs` file and are easily adaptable to any configuration. The virtual sources are represented as balls with fading trajectories. The RMS gain (in dBFS) of each loudspeaker given by the Ambisonics solution drives in real-time the size and the color of the loudspeaker's-balls via OSC messages. The position of each source is also received by OSC messages. The `Peasycam` library⁵ allows to zoom, move and pan the 3D view. The described visual tool is shown in Fig. 6. The sphere used here is a Lebedev sphere using $N_2 = 50$ loudspeakers as used in [16].

6. USER CONTROL INTERFACE

6.1. Controls by software

The user control interface is provided by FAUST compiled code. Thus, depending on the application it could be a standalone application, a MAX/MSP patch, VST or LV2 plugin, or others. It consists in sliders and entry boxes, to control the parameters $(r_1, \theta_1, \delta_1, r_0)$. Check-boxes are used to mute an order m . VU-Meters give the signals level in dBFS for the $B_{mn}^\sigma(z)$ components or s_l .

As an example, a JACK⁶ standalone application interface for real-time 3D spatialization using the panning law of Eq. (13) is shown in Fig. 7.

⁴<https://processing.org/>

⁵<http://mrfeinberg.com/peasycam/>

⁶<http://jackaudio.org>

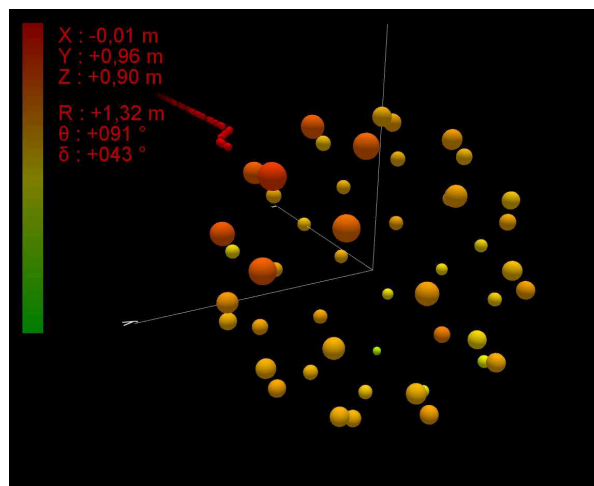


Figure 6: 3D visual feedback tool using PROCESSING language. The virtual source is shown as a red ball with fading trajectory (top left corner). Each loudspeaker is represented by a ball with radius and color driven by loudspeaker signal level (RMS, dBFS). 0 dBFS correspond to the red color on the left color scale and $-\infty$ dBFS to green. All informations are received by OSC messages generated by a joystick and a PURE DATA patch.

6.2. Controls by hardware device and OSC messages

FAUST supports OSC to control the parameters of the DSP tools. For example, with a hardware controller and a PURE DATA patch, it is possible to generate OSC messages which control the plug-in as well as the visual feedback tool of Sec. 5. This correspond to the rightmost branch of Fig. 3. The use of hardware controllers and FAUST generated plug-ins allow for the easy control and recording of the synthesis parameters. As an example, the implemented Ambisonics encoder loaded in ARDOUR⁷ Digital Audio Workstation (DAW) as a LV2 plugin is shown in Fig. 8. With the visual feedback tool described in Sec. 5, this configuration enables to record automation tracks describing "manually-created" trajectories.

7. EXPERIMENTAL SETUP

An experimental setup is presented briefly in this section as an application of the tools presented above.

7.1. 3D sound pressure field capture

The recording of natural sound pressure fields is made with a rigid spherical microphone "MemsBedev"⁸ shown in Fig. 9. This microphone was made by 3D printing and uses $N = 50$ microphone capsules on a Lebedev grid [28]. Each capsule is made of 4 MEMS⁹ microphones to reduce the background noise. The microphone works up to 5th order, as explained in Ref.[16]. The analog signals are digitalized with a commercial front-end. The tools presented in this article can provide in real-time the Ambisonics components $B_{mn}^\sigma(z)$ up to 5th order according to Eq. (5) and Fig. 3.

⁷<http://ardour.org/>

⁸<http://www.cinela.fr/>

⁹MEMS: Micro Electro-Mechanical System

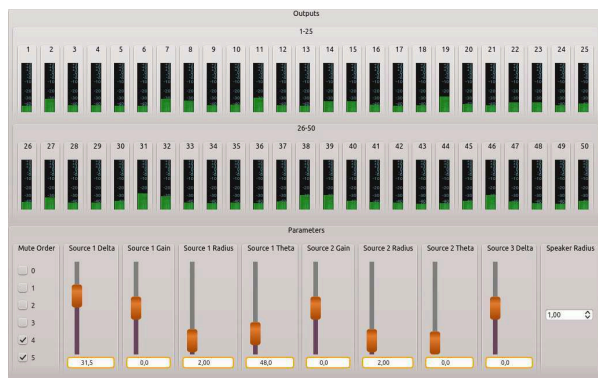


Figure 7: Standalone application interface running under JACK. The sliders allow controlling the gains and positions of two virtual sources (bottom, grey and yellow). The check-boxes (bottom-left) mute and unmute a specific order at the synthesis stage. The entry box allows giving the r_0 radius of the rendering spherical loudspeakers layout (bottom right). The VU-meters give the signals levels in dBFS for each loudspeaker. In this case there are $N_2 = 50$ loudspeakers according to a Lebedev grid.

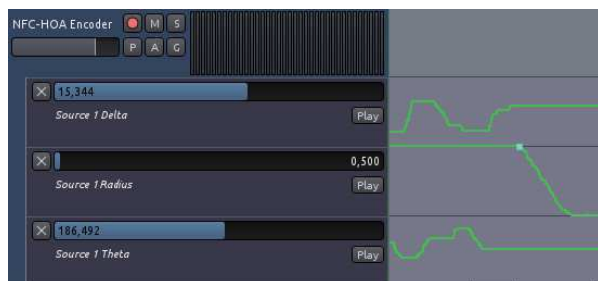


Figure 8: Ambisonics encoder loaded in ARDOUR as a LV2 plug-in. The positions parameters are controlled with the mouse or by an hardware joystick via OSC messages. Resulting automation tracks are shown in green in this figure. The LV2 plug-in is generated from FAUST code using `faust2lv2`.

7.2. 3D sound pressure field synthesis

The decoding of the pressure sound field is made in real-time with a decoder made with FAUST according to Eq. (7) and Fig. 3. The Lebedev spherical loudspeaker layout used for the sound field rendering is shown in Fig. 10. For recording rendering purposes, it is possible to record in 3D in one place and render in real-time the 3D sound pressure field with the loudspeaker sphere located at another location. For spatialization purposes, the user takes place in the sphere and drives the position of the virtual sources using an hardware controller. The hardware controller is linked to a PURE DATA patch and generates OSC messages for the synthesis using FAUST, as summarized in Fig. 3. This experimental setup can thus help to mix 3D audio contents *in situ*.



Figure 9: MemsBedeve microphone arrays with $N = 50$ MEMS microphones

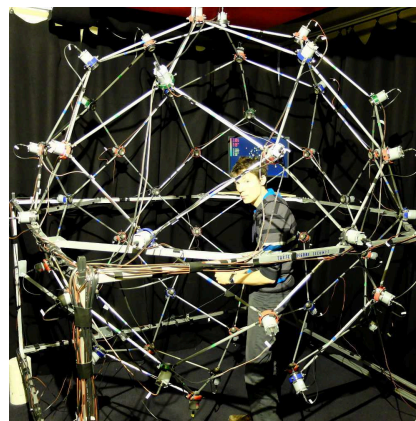


Figure 10: Lebedev spherical loudspeakers layout with $N_2 = 50$ loudspeakers.

8. CONCLUSION

Several integrated tools for real-time 3D sound pressure field synthesis using Ambisonics were developed and presented in this paper. The signal processing was implemented using FAUST, the visual feedback with PROCESSING and the communication between tools with OSC protocol. The near-field filters were implemented in a FAUST library and allow synthesizing spherical waves. The gain correction with distance is also implemented, which is of great importance for engineering applications or simulators. In the current version of the implementation, the synthesis is controlled up to 5th order in 3D. Some future ameliorations could include transformations in Ambisonics domain [29], recent advances in radial filters implementation [27] or inclusion of the Doppler effect for moving sources [30]. The code of this project is freely available online under GPL license¹⁰. In a near future, the described implementation and experimental setup will be used to reproduced various recorded environments and achieve physical evaluation of the reproduced sound fields.

¹⁰<https://github.com/sekisushai/ambitools>

9. REFERENCES

- [1] Jérôme Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*, Ph.D. thesis, Université Paris 6, Paris, 2000.
- [2] Jérôme Daniel, “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*, Helsingør, 2003, pp. 1–15, AES.
- [3] Sébastien Moreau, Jérôme Daniel, and Stéphanie Bertet, “3d sound field recording with higher order ambisonics-objective measurements and validation of spherical microphone,” in *Audio Engineering Society Convention 120*, Paris, 2006, pp. 1–24, AES.
- [4] Matthias Kronlachner, “Plug-in Suite for mastering the production and playback in surround sound and ambisonics,” in *Linux Audio Conference*, 2013.
- [5] Christian Nachbar, Franz Zotter, Etienne Deleflie, and Alois Sontacchi, “Ambix - A suggested ambisonics format,” in *Ambisonics Symposium*, Lexington, 2011.
- [6] Aaron J. Heller and Eric M. Benjamin, “The Ambisonic Decoder Toolbox: Extensions for Partial-Coverage Loudspeaker Arrays,” in *Linux Audio Conference*, 2014.
- [7] Julien Colafrancesco, Pierre Guillot, Elliott Paris, Anne Sèdes, and Alain Bonardi, “La bibliothèque HOA, bilan et perspectives,” in *Journées d’Informatique Musicale*, 2013, pp. 189–197.
- [8] Matthias Geier and Sascha Spors, “Spatial Audio with the SoundScape Renderer,” in *27th Tonmeistertagung–VDT International Convention*, 2012.
- [9] Darren B. Ward and Thushara D. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 697–707, 2001.
- [10] Philippe-Aubert Gauthier, Cédric Camier, Thomas Padois, Yann Pasco, and Alain Berry, “Sound Field Reproduction of Real Flight Recordings in Aircraft Cabin Mock-Up,” *J. Audio Eng. Soc.*, vol. 63, no. 1/2, pp. 6–20, 2015.
- [11] Anthony Bolduc, Philippe-Aubert Gauthier, Telina Ramanana, and Alain Berry, “Sound Field Reproduction of Vibroacoustic Models: Application to a Plate with Wave Field Synthesis,” in *Audio Engineering Society Conference: 55th International Conference: Spatial Audio*, 2014.
- [12] Michael Vorländer, *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*, Springer Science & Business Media, 2007.
- [13] Estelle Bongini, *Modèle acoustique global et synthèse sonore du bruit d’un véhicule: application aux véhicules ferroviaires*, Ph.D. thesis, Université de Provence-Aix-Marseille I, 2008.
- [14] Yann Orlarey, Dominique Fober, and Stéphane Letz, “FAUST: an efficient functional approach to DSP programming,” *New Computational Paradigms for Computer Music*, vol. 290, 2009.
- [15] Boaz Rafaely, “Analysis and design of spherical microphone arrays,” *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 135–143, Jan. 2005.
- [16] Pierre Lecomte, Philippe-Aubert Gauthier, Christophe Langrenne, Alexandre Garcia, and Alain Berry, “On the use of a Lebedev grid for Ambisonics,” in *Audio Engineering Society Convention 139*, New York, 2015.
- [17] Boaz Rafaely, Barak Weiss, and Eitan Bachmat, “Spatial aliasing in spherical microphone arrays,” *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 1003–1010, 2007.
- [18] Mark A. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” *Journal of the Audio Engineering Society*, vol. 53, no. 11, pp. 1004–1025, 2005.
- [19] George B. Arfken and Hans J. Weber, *Mathematical methods for physicists - sixth edition*, Elsevier, 6th edition, 2005.
- [20] Fons Adriaensen, “Near field filters for higher order Ambisonics,” <http://kokkinizita.linuxaudio.org/papers/hoafilt.pdf>, 2006, Accessed: 2013-10-28.
- [21] Sascha Spors, Rudolf Rabenstein, and Jens Ahrens, “The theory of wave field synthesis revisited,” in *Audio Engineering Society Convention 124*, Amsterdam, 2008, pp. 1–19, AES.
- [22] Alan V. Oppenheim and Ronald W. Schaffer, *Digital signal processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
- [23] Jens Ahrens and Sascha Spors, “Focusing of virtual sound sources in higher order Ambisonics,” in *Audio Engineering Society Convention 124*, Amsterdam, 2008, pp. 1–9, AES.
- [24] Stefan Lösler and Franz Zotter, “Comprehensive radial filter design for practical higher-order Ambisonic recording,” *Fortschritte der Akustik, DAGA*, 2015.
- [25] Jérôme Daniel and Sébastien Moreau, “Further study of sound field coding with higher order ambisonics,” in *Audio Engineering Society Convention 116*, Berlin, 2004, pp. 1–14, AES.
- [26] Hannes Pomberger, “Angular and radial directivity control for spherical loudspeaker arrays,” 2008.
- [27] Robert Baumgartner, Hannes Pomberger, and Matthias Frank, “Practical implementation of radial filters for ambisonic recordings,” *Proc. of ICSA, Detmold*, 2011.
- [28] V.I. Lebedev, “Values of the nodes and weights of quadrature formulas of Gauss-Markov type for a sphere from the ninth to seventeenth order of accuracy that are invariant with respect to an octahedron,” *USSR Computational Mathematics and Mathematical Physics*, vol. 15, no. 1, pp. 44–51, 1975.
- [29] Matthias Kronlachner and Franz Zotter, “Spatial transformations for the enhancement of Ambisonic recordings,” in *Proceedings of the 2nd International Conference on Spatial Audio, Erlangen*, 2014.
- [30] Gergely Firtha and Peter Fiala, “Sound Field Synthesis of Uniformly Moving Virtual Monopoles,” *Journal of the Audio Engineering Society*, vol. 63, no. 1/2, pp. 46–53, 2015.

ANNEXE D

EXTRAIT DE LA DOCUMENTATION D'AMBITOOLS

D.1 Avant-propos

Cette annexe présente un extrait de la documentation en anglais de la boîte à outils *ambitools*¹, développés dans le cadre de la thèse. Les différents outils disponibles dans *ambitools* sont présentés, notamment ceux utilisés au chapitre 6, pour la validation expérimentale.

D.2 Introduction

D.2.1 Goals of *ambitools*

ambitools is a collection of tools for sound field synthesis using Near-Field-Compensated Higher Orders Ambisonics (NFC-HOA). For the rest of this document, the denomination Ambisonics will be used for convenience. *ambitools* is developed in the context of my PhD on 3D sound field synthesis. The audio processing is coded in FAUST² (Functional AUdio Stream) which allows to produce efficient C++ code and exports in various DSP tools format: VST, standalone applications, LV2, etc. Thus, *ambitools* is multi-platform (although conceived under Linux/Jack).

The goal of *ambitools* is mainly to produce several modules to encode, decode and transform 3D synthesized sound fields or 3D recordings in a context of physical sound field synthesis.

The project is open-source under GPL licence. *ambitools* is thought to be flexible enough to adapt to the user configuration. The FAUST code for each tool has an header with several variable that the user can change, such as Ambisonic order M , number of sources N , etc. Thus, if the user needs an encoder up to order $M = 4$, for $N = 5$ sources, one just has to change these two parameters in `hoa_encoder_N_sources.dsp` and produce the required plug-in using FAUSTLIVE² for instance.

So far, these tools are used to drive the several spherical loudspeakers arrays at Université de Sherbrooke, Québec, Canada and Conservatoire National des Arts et Métiers, Paris, France.

Hopefully, since Ambisonics are flexible, the user can use *ambitools* with binaural rendering and listen to 3D sound without the need for a spherical loudspeaker array!

¹*ambitools* est disponible en ligne : <http://www.sekisushai.net/ambitools/>

²<http://faust.grame.fr>

D.3 Installing *ambitools*

D.3.1 Retrieve *ambitools* repository

To install *ambitools*, simply go on the github repository³ and clone it. To do so, open a terminal in the directory one would like to clone the repository and type the following command:

```
$ git clone https://github.com/sekisushai/ambitools
```

To keep the repository up to date, type the following command at the root of the directory `ambitools/`:

```
$ git pull
```

One can also download a `.zip` file from github⁴

The resulting `ambitools/` folder should have the following structure:

- **Documentation/** Everything concerning the documentation (pdfs, including some scientific papers).
- **Faust/** Everything written in FAUST language (all the *ambitools* plug-ins with some utilities).
 - `bin/` Compiled plug-ins in various formats.
 - `src/` Source code of the plug-ins.
 - `src/lib/` Shared libraries (spherical harmonics, gui, etc.).
- **FIR/** Finite Impulse Response (FIR) filters banks for binaural rendering and spherical microphone equalization filters, to use with Jconvolver⁵, fast convolution software.
 - `spherical_microphones/` Equalization filters for rigid spherical microphone, such as mh acoustics EigenMikeTM⁶.
 - `hrir/` Head Related Impulses Responses (HRIR) of several people to use with binaural rendering over headphones.
- **Processing/** Everything written in PROCESSING language, namely the spherical VU-Meter (see Sec. [D.4.2](#)).
 - `bin/` Compiled spherical VU-Meter in Java for various architectures.
 - `src/` PROCESSING source code.

³<https://github.com/sekisushai/ambitools>

⁴<https://github.com/sekisushai/ambitools/archive/master.zip>

⁵<http://kokkinizita.linuxaudio.org/linuxaudio/>

⁶<http://www.mhacoustics.com>

- **PureData/** Everything written in Pure Data (a few sounds generator patches, head-tracking patch and PlayStation-like remote patch to drive FAUST plug-ins with Open Sound Control protocol, OSC).

D.3.2 Dependencies

Depending on what one is looking for with the code, the following dependencies should be provided:

Dependency	Usage
OSC support	Drive the tools standalone tools with OSC.
JAVA	Run the Spherical VU-Meter (see Sec. D.4.2).
JCONVolver	Real-time convolution for binaural rendering or spherical microphone signals filtering.
PROCESSING	Edit and compile the Processing code.
FAUST, FAUSTLIVE	Edit and compile the FAUST code.

Table D.1 Dependencies of *ambitools*.

D.3.3 Compile the Faust plug-ins

The FAUST plug-ins source codes are in the sub-folder `Faust/src/`.

D.3.4 Local Faust installation

If one has FAUST installed on ones machine with the required dependencies, one can run the scripts collection `faust2*` to produce the plug-in of ones choice in the desired format.

For example, to compile `hoa_encoder_N_sources.dsp` into a standalone Linux jack-qt application with OSC support, type the following command in a terminal in the folder `\Faust/src`

```
$ faust2jaqt hoa_encoder_N_sources.dsp -osc
```

D.3.5 FaustLive

To compile the plug-ins to ones requirements, load the chosen plug-in in FAUSTLIVE² and choose `Window/Export As...` (see Fig. [D.1](#)).

D.3.6 Compile the Processing VU-Meter

The PROCESSING source code is in the folder `Processing/scr`. One should open the file `Spherical_VU_Meter.pde` in the PROCESSING editor and select "File/Export..." to produce a binary application.



Figure D.1 FAUSTLIVE Export Manager.

D.4 The different tools

This section gives a quick presentation of each tool contained in *ambitools*.

D.4.1 Faust

The core of the sound processing is written in FAUST language. Note that the majority of the figures presented in the following section will be screenshots of the tools compiled as standalone JACK applications for Linux, using `faust2jaqt` script. However, thanks to the versatility of FAUST language, each of these tools can be compiled for various architectures, using FAUSTLIVE² for example.

hoa_N_sources_encoder

- Inputs: N
- Outputs: $(M + 1)^2$

This first tool allows to encode N signal inputs to an 3D Ambisonics scene described with Ambisonics components signals up to order M (the B -Format). One can choose N and M at the compilation time in file `hoa_encoder_N_sources.dsp`. The resulting graphical user interface when using `faust2jaqt` script is shown in Fig. D.2 for $N = 3, M = 5$. Each input signal can be encoded in B -Format as a plane wave or a spherical wave.

For the plane wave case, the check-box **Spherical Wave** should be unchecked. Consequently, the knob **Radius** and input entry **Speakers Radius** are without effect in this case (a plane wave has no distance information).

For the spherical wave case, the knob **Radius** is used to set the source radius to origin. Note that to represent a spherical wave in Ambisonic domain, near-field filters are activated [Daniel, 2003; Lecomte et Gauthier, 2015]. Those filters, unstable by nature, are stabilized using the near-field compensation filters from decoding step. Thus in this case, the radius of the spherical loudspeakers grid should be known prior to decoding and given in **Speakers Radius** input entry.

Finally, the outputs of this tool are *B*-Format signals up to order M . Each order can be muted or each individual Ambisonic components individually.



Figure D.2 `hoa_encoder_N_sources` plug-in compiled using `faust2jaqt` script with $N = 2$ and $M = 5$. The **Gain** knob is used to adjust the input level of the input. **Radius** knob is used to set the source distance to origin in case of spherical wave encoding. **Azimuth** and **Elevation** knobs are used to set the source direction. The **Spherical Wave** check-box enables the encoding of spherical wave, using near-field filters. **Speakers Radius** sets the radius for the spherical array of loudspeakers at decoding stage in case of spherical wave encoding. Finally, the VU-Meters show the levels of 3D *B*-Format signals in dBFS up to order M . Each meter as a **Mute** check-box and each order as well.

`hoa_mic_encoder*`

- Inputs: 06, 26, 32, or 50 depending on the configuration.
- Outputs: $(M + 1)^2$

It is possible to encode a natural sound field captured by a rigid spherical microphone. To do so, the signals from each capsule on the spherical microphone are mixed to operate the Discrete Spherical Fourier Transform (DSFT) [Lecomte *et al.*, 2015]. Then, the resulting $(M + 1)^2$ components are filtered to obtain the Ambisonics components. In *ambitools*, the DSFT projection is done for several spherical rigid spherical microphone configurations using `hoa_mic_encoder*` tools. The filtering of the resulting components is done using *Jconvolver*, a real-time fast convolution software. See Sec. D.4.3 for further details. Note that the filtering of the outputs signals of `hoa_mic_encoder*` tools is mandatory in order to obtain the Ambisonics components. The different microphone configurations available in *ambitools* are 06, 26 or 50 capsules spherical microphone arranged on a Lebedev grid

[Lecomte *et al.*, 2015] and a 32 capsules spherical microphone arranged on a pentakis-dodecahedron grid, i.e. the mh acoustics Eigenmike™ microphone [Elko *et al.*, 2009]. Those configuration can operate the Spherical Fourier Transform up to order $M = 1$, $M = 3$, $M = 5$ and $M = 4$ respectively [Lecomte *et al.*, 2015; Moreau *et al.*, 2006]. The corresponding tools are named `hoa_mic_encoder_lebedev06`, `hoa_mic_encoder_lebedev26`, `hoa_mic_encoder_lebedev50` and `hoa_mic_encoder_eigenmike32` respectively. The graphical user interface when using `faust2jaqt` script is shown in Fig. D.3 for the tool `hoa_mic_encoder_lebedev26` compiled with $M = 3$.

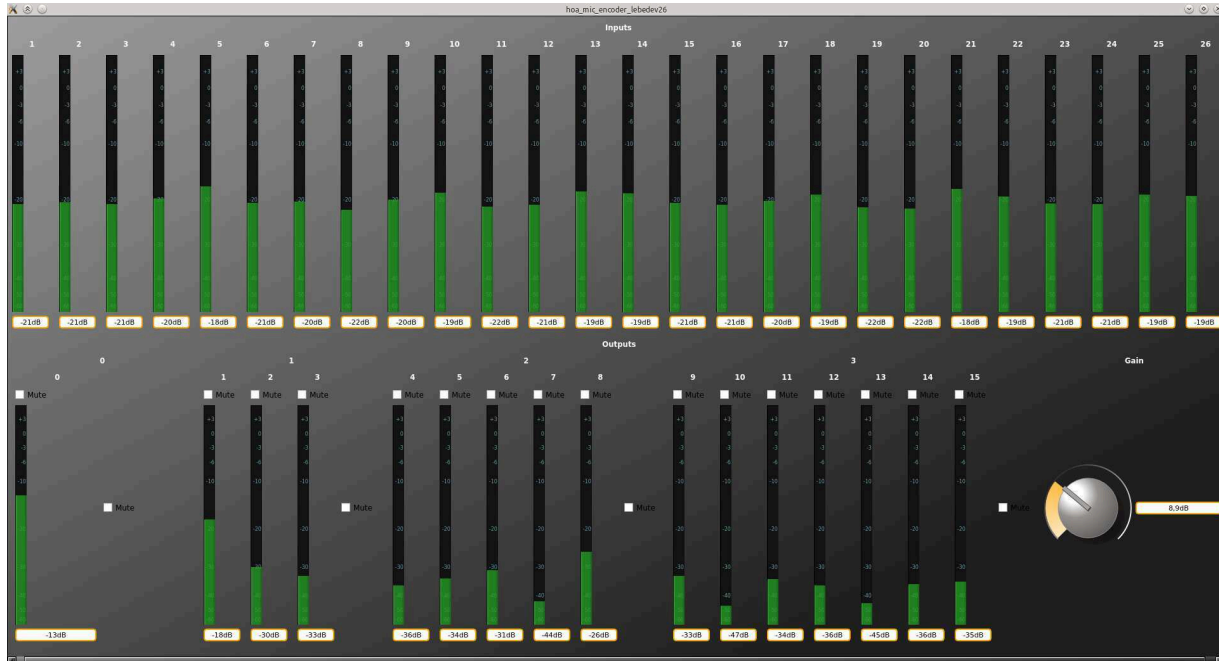


Figure D.3 `hoa_mic_encoder_lebedev26` compiled using `faust2jaqt` script with $M = 3$. The VU-Meters at the top show the signal level in dBFS of each capsule on the spherical microphone. The VU-Meters at the bottom show the signal level of each component after the DSFT operation. The check-boxes **Mute** are used to mute some components or all components of an order. The knob **Gain** applies a global gain to the output signals.

Fig. D.4 shows the connection in order to obtain the Ambisonics components from the signals of a rigid spherical microphone. The raw signals issued from the spherical microphone are played using a Digital Audio Workstation, Ardour 4 in this case. These signals are then connected to the input of the tool `hoa_mic_encoder_lebedev50` in the middle of Fig. D.4. This tool mixes the signals to operate the DSFT. The resulting output signals are finally filtered using `Jconvolver`, on the right of Fig. D.4. The outputs of `Jconvolver` are the Ambisonics components, here up to order $M = 5$.

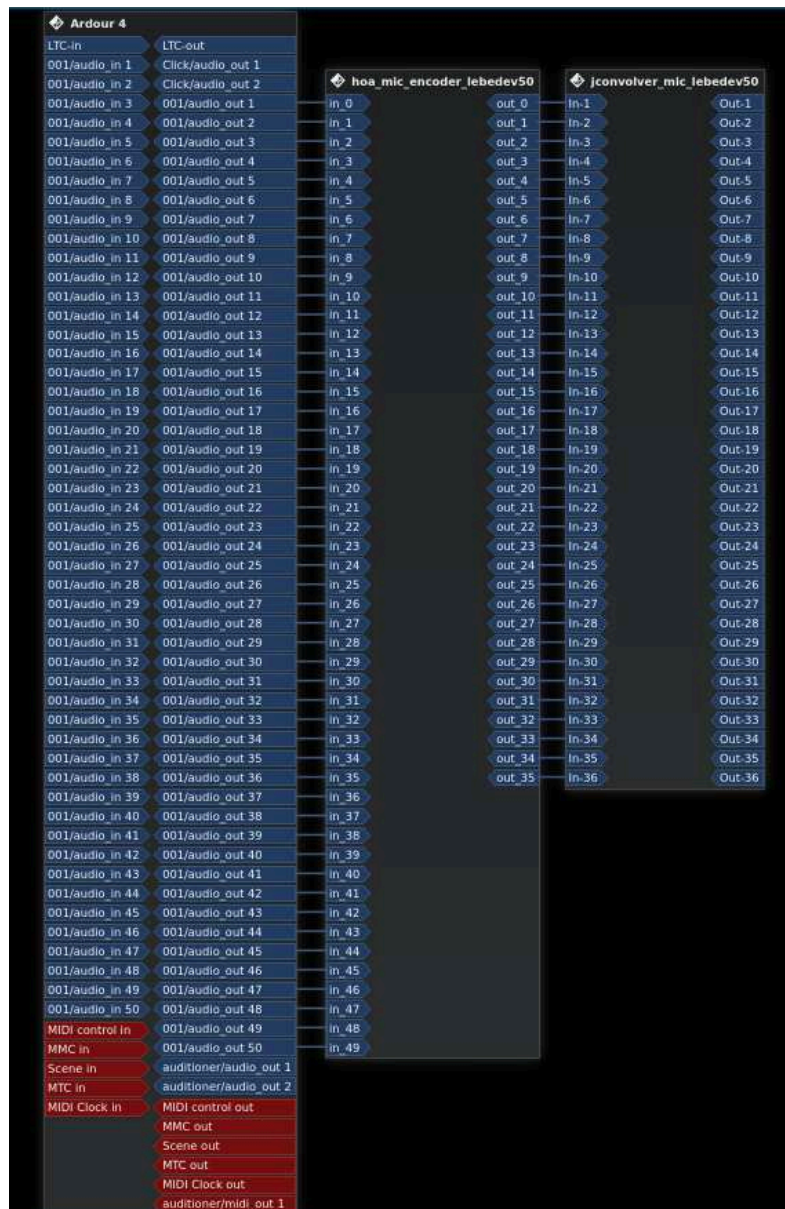


Figure D.4 Connections in order to obtain the Ambisonics components from the signals of a rigid spherical microphone. On the left, the 50 outputs of Ardour 4: the 50 signals of the rigid spherical microphone with a 50-node Lebedev configuration. On the middle, the tool `hoa_mic_encoder_lebedev50` which operates the DSFT. On the right, the filtering of the signals with `Jconvolver`, the outputs of this box are the Ambisonics components up to order $M = 5$. Screenshot taken from Claudia under KXStudio Linux distribution.

hoa_decoder_*

- Inputs: $(M + 1)^2$
- Outputs: 06, 26 or 50 depending on the decoder.

Three basic decoders by mode-matching [Daniel, 2000; Poletti, 2005] are currently available in *ambitools*: `hoa_decoder_lebedev06`, `hoa_decoder_lebedev26` and `hoa_decoder_lebedev50`. Those decoders allow to decode $(M + 1)^2$ Ambisonics signals on Lebedev grids with respectively 6, 26 and 50 loudspeakers [Lebedev, 1975; Lecomte *et al.*, 2015]. Those grids are able to reconstruct the sound field up to order $M = 1$, $M = 3$ and $M = 5$ respectively. If other decoders are required, one should have a look at the ambisonic decoder toolbox from Aaron Heller [Heller *et al.*, 2012], or contact the author. The graphical user interface when using `faust2jaqt` script is shown in Fig. D.5 for the `hoa_decoder_lebedev50` decoder compiled with $M = 5$. The decoder can be with or without near-field compensation (NFC) filters [Daniel *et al.*, 2003; Lecomte *et al.*, 2015]. Those filters allow to take into account the finite distance of the loudspeakers: In other words, if they are disabled, the loudspeakers are modeled as plane-wave generators. In case of spherical wave encoding using the `hoa_N_sources_encoder` plug-in, the `NFC` check-box should be unchecked as the near-field compensation filters are already used at encoding step (see Sec. D.4.1).

Usage with the Spherical VU-Meter Note that when using these tools to drive the Spherical VU-Meter (Sec. D.4.2) with OSC, one needs to activate OSC transmission mode 2. This is done by launching the tool with the argument `-xmit 2`. For example, with the decoder `hoa_decoder_lebedev50`, the command should be:

```
$ ./hoa_decoder_lebedev50 -xmit 2
```

hoa_decoder_lebedev50_binaural

- Inputs: $(M + 1)^2$
- Outputs: 2

`hoa_decoder_lebedev50_binaural` is a decoder which computes directly the binaural signals for left and right ears from the Ambisonic components. It is constructed using a basic decoder for 50-node Lebedev grid [Lecomte *et al.*, 2015], and Head Related Impulse Response (HRIR) of a dummy head Neumann KU-100 from [Bernschütz, 2013]. There are $(M + 1)^2 \times 2$ Finite Impulse Responses (FIR) filters involved in the process, with 256 taps each. Thus, this tool can be quite CPU consuming as it applies $(M + 1)^2 \times 2$ linear convolution operations. The graphical user interface when using `faust2jaqt` script is shown in Fig. D.6 for the `hoa_decoder_lebedev50_binaural` binaural compiled with $M = 2$.

hoa_panning_*

- Inputs: N

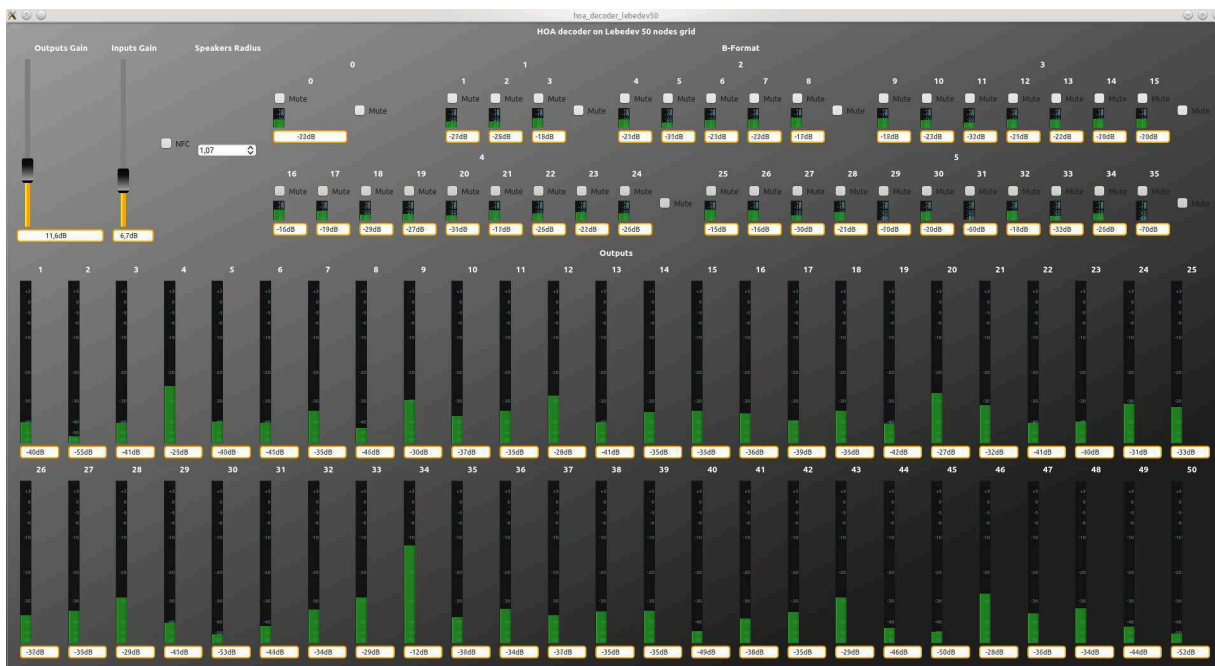


Figure D.5 `hoa_decoder_lebedev50` compiled using `faust2jaqt` script with $M = 5$. The slider **Outputs Gain** applies a global gain on all outputs (loudspeaker signals). The slider **Inputs Gain** applies a global gain on all inputs (B -Format signals). The VU-Meters **Inputs** and **Outputs** give the signals level in dBFS. The check-box **NFC** activate or deactivate the near-field compensation filters. The input entry **Speakers Radius** are used to set the spherical grid radius. Finally, the check-boxes **Mute** above all inputs VU-meters are used to mute some specific B -Format signals. Or, all the signal from an order with **Mute** check-boxes on side of a VU-Meter group.

- Outputs: 06, 26 or 50 depending on the configuration.

It is possible to compute directly the signals of the loudspeakers without relying on encoding and decoding processes by an encoding and decoding process [Lecomte *et al.*, 2015]. This equivalent 3D-panning is implemented with the tools `hoa_panning_N_sources_lebedev06`, `hoa_panning_N_sources_lebedev26` and `hoa_panning_N_sources_lebedev50` for Lebedev grids with 6, 26 and 50 loudspeakers, respectively [Lebedev, 1975; Lecomte *et al.*, 2015]. These grids allow to control the sound field up to order $M = 1$, $M = 3$ and $M = 5$ respectively [Lecomte *et al.*, 2015]. The graphical user interface when using `faust2jaqt` script is shown in Fig. D.7 for the `hoa_panning_N_sources_lebedev50` panning tool compiled with $M = 5$ and $N = 2$.

Use with the Spherical VU-Meter When using these tools to drive the Spherical VU-Meter (Sec. D.4.2) with OSC, one needs to activate OSC transmission mode 2. This is achieved by launching the tool with the argument `-xmit 2`. For example, with the panning tool `hoa_panning_N_sources_lebedev50`, the command should be:



Figure D.6 `hoa_decoder_lebedev50_binaural` compiled using `faust2jaqt` script with $M = 2$. The VU-Meters on top show the signal level of inputs Ambisonics components in dBFS. The check-boxes `Mute` are used to mute certain component or all components of an order. The slider `Inputs Gain` applies a global gain on all inputs. The slider `Outputs Gain` applies a global gain on all outputs. Finally, the horizontal VU-Meters at the bottom show the signal level in dBFS of Left and Right channels for headphone signal.

```
$ ./hoa_panning_N_sources_lebedev50 -xmit 2
```

Note about sound field transformation tools in Ambisonic domain

The tools presented in the next sections perform sound field transformations in Ambisonic domain. Thus, they should be used on Ambisonic signals and inserted after encoding step and before decoding step. For instance, the connections for the tool `hoa_mirroring` are shown in Fig. D.8. It is inserted after `hoa_encoder_N_sources` and before `hoa_decoder_lebedev50`, acting on Ambisonic signals.

`hoa_mirroring`

- Inputs: $(M + 1)^2$
- Outputs: $(M + 1)^2$

`hoa_mirroring` performs mirroring operations on the sound field. The directions front and back can be inverted, as well as left and right, up and down, or any combinations of the above. In fact, the transformation changes the sign of particular Ambisonic components [Kronlachner, 2014]. The graphical user interface when using `faust2jaqt` script is shown



Figure D.7 `hoa_panning_N_sources_lebedev50` compiled using `faust2jaqt` script with $M = 5$ and $N = 2$. The slider **Outputs Gain** applies a global gain on all outputs (loudspeakers signals). The check-boxes **Mute Order** are used to mute some Ambisonic order in the computation of the driving signals. The sliders **Gain Radius Azimuth Elevation** control the position and gain of the sources. The check-box **Spherical Wave** is used to switch between the synthesis of a spherical wave or a plane wave. In the case of spherical wave synthesis, the input entry **Speakers Radius** sets the spherical loudspeakers system radius.

in Fig. D.9. To visualize the transformation, Fig. D.10 shows the Spherical VU-Meter (see Sec. D.4.2) for a 50-node Lebedev grid on which a sound field is decoded. The sound field is composed of one source positioned at $(1.07 \text{ m}, 0^\circ, 90^\circ)$, i.e. above the head of the listener. One can note the signal level for each loudspeaker after decoding step. Thus, as shown in Fig. D.10(a), when the `hoa_mirroring` tool is inactive, the maximum sound energy is mostly coming from above the listener head. When the **up-down** is checked, as shown in Fig. D.10(b), one notes that the sound field has been mirrored according to the up-down inversion. Thus, the sound energy is mostly coming from underneath the listener.

`hoa_azimuth_rotator`

- Inputs: $(M + 1)^2$
- Outputs: $(M + 1)^2$

`hoa_azimuth_rotator` performs a rotation of the sound field around the z -axis (yaw angle). The trigonometric sens is selected (i.e. anti-clockwise). The transformation is done by a rotation matrix operation on the Ambisonic components [Daniel, 2000; Kronlachner, 2014; Moreau, 2006]. The graphical user interface when using `faust2jaqt` script is shown in Fig. D.11. It is basically a slider to select the azimuthal angle of rotation. To

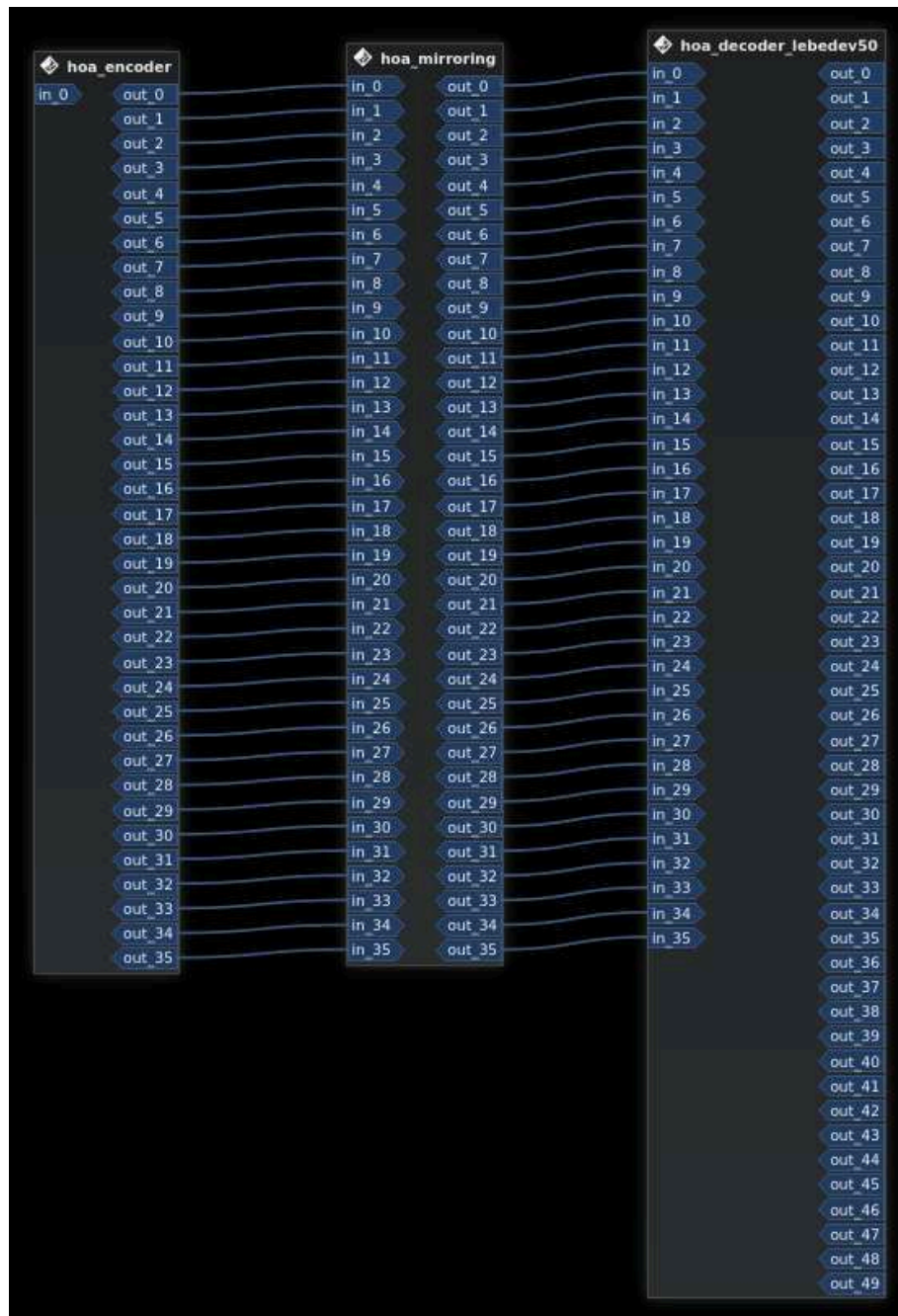


Figure D.8 Connections between plug-ins `hoa_encoder_N_sources`, `hoa_mirroring` and `hoa_decoder_lebedev50` compiled with `faust2jaqt` script with $M = 5$. `hoa_mirroring` acts on Ambisonics signals. Screenshot taken from Claudia under KXStudio Linux distribution.



Figure D.9 `hoa_mirroring` compiled using `faust2jaqt` script. The checkboxes are used to select the mirror transformations among: `left-right`, `front-back` or `up-down`.

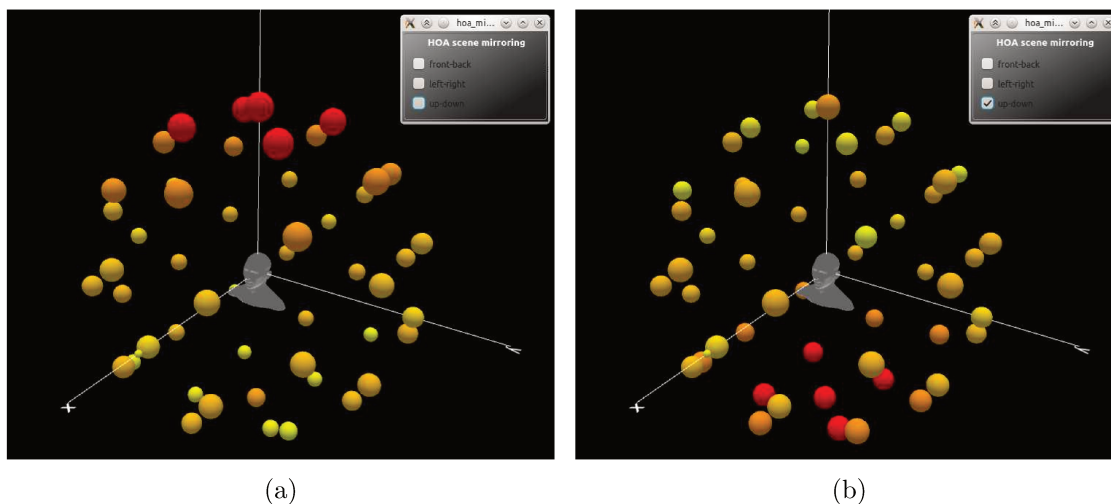


Figure D.10 `hoa_mirroring` effects on a sound field. (a): loudspeakers energy with the original sound field decoded on a 50 loudspeaker Lebedev grid. (b) the sound field has been transformed according to a mirroring up-down with `hoa_mirroring` tool.

visualize the transformation, Fig. D.12 shows the Spherical VU-Meter (see Sec. D.4.2) for a 50-node Lebedev grid on which a sound field is decoded. The sound field is made by one source located at $(1.07 \text{ m}, 0^\circ, 0^\circ)$, i.e. in front of the listener. One can observe the signal level for each loudspeaker after the decoding step. Thus, in Fig. D.10(a), when the `hoa_azimuth_rotator` slider `Azimuth` is set to 0° , i.e. no rotation, the maximum sound energy is coming from the front of the listener. When the slider `Azimuth` is set to 90° , in Fig. D.10(b), one observes that the sound field has been rotated 90° counter-clockwise around the z -axis. Thus, the sound energy is coming from the left of the listener in this case.

Head Tracking One possible use of the tool `hoa_azimuth_rotator` is to compensate for listener's head rotation around the z -axis. To this end, an head-tracking system is required [Noisternig *et al.*, 2003]. Thus, when using binaural rendering over headphone, the sound field is rotated of minus the head rotation angle in order to compensate the



Figure D.11 `hoa_azimuth_rotator` compiled using `faust2jaqt` script. The slider `Azimuth` is used for the setting of the the azimuthal angle of rotation, around the z -axis.

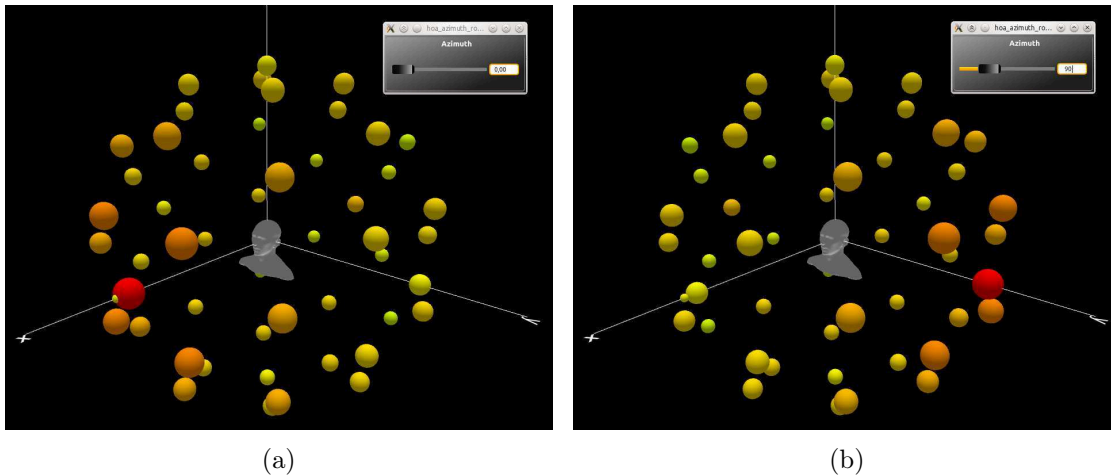


Figure D.12 `hoa_azimuth_rotator` effects on a sound field. (a): loudspeakers energy for the original sound field decoded on a 50 loudspeaker Lebedev grid. (b) the sound field has been rotated according to a rotation around z -axis of $+90^\circ$ using the `hoa_azimuth_rotator` tool.

head rotation and leave the virtual source, as perceived by the listener, at the same spatial location. *ambitools* provides a *Pure-Data* patch to drive `hoa_azimuth_rotator` with OSC and to perform an head-tracking system with a smartphone. See Sec. D.4.4 for further informations.

`hoa_beamforming_hypercardioid_to_mono`

- Inputs: $(M + 1)^2$
- Output: 1

`hoa_beamforming_hypercardioid_to_mono` is a tool to perform modal beamforming and extract a monophonic signal as if it was recorded by a directional microphone in the sound field. In fact, the Ambisonic signals are weighted and summed to give the monophonic output signal. The weights of the combination traduce the beampattern of the directional microphone which is used Meyer et Elko, 2002. Thus, this tool allows for the listening in a given direction in the 3D sound field, as if one was recording in the middle of the sound field with a directional microphone with a specific beampattern. `hoa_beamforming_hypercardioid_to_mono` provides the weights corresponding to regu-

lar hyper-cardioid microphone up to order 3 [Lecomte *et al.*, 2016c; Meyer et Elko, 2002]. In Fig. D.14, the available beampatterns are shown: the higher the order of the hyper-cardioid, the more directional the corresponding virtual microphone. The graphical user interface when using `faust2jaqt` script is shown in Fig. D.13 for $M = 3$.

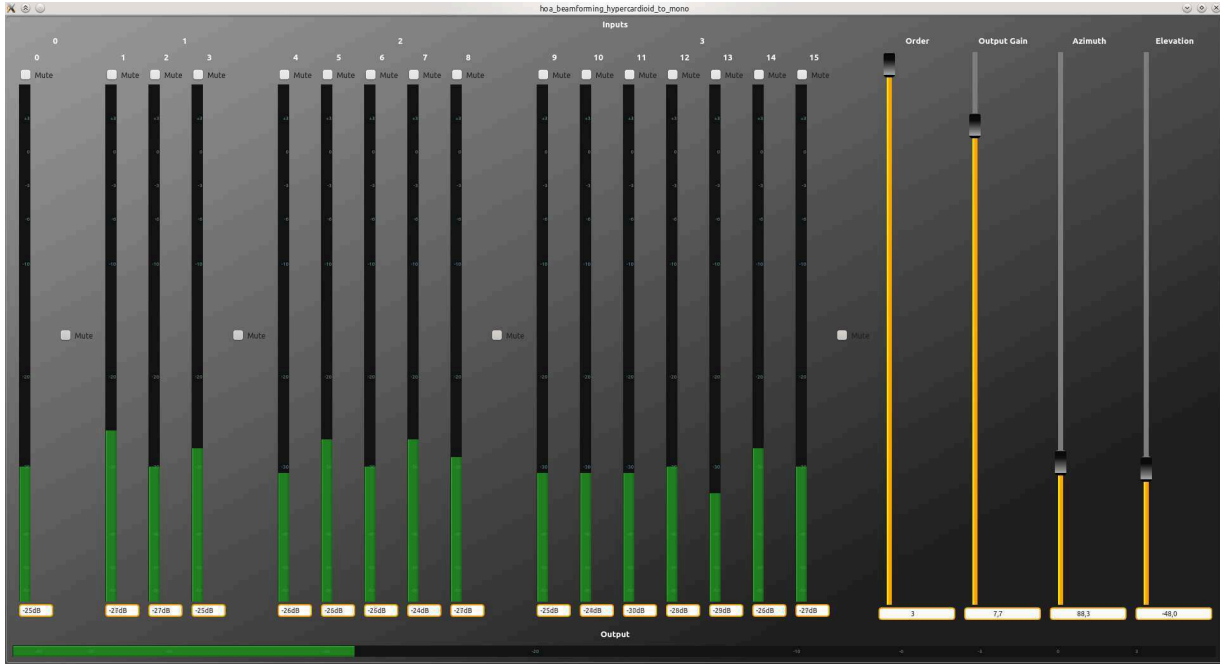


Figure D.13 `hoa_beamforming_hypercardioid_to_mono` compiled using `faust2jaqt` script. The VU-Meters on the left give the signal levels for each Ambisonic component. The check-boxes **Mute** are used to mute some components or all the components of a given order. On the right, the slider **Order** is used for the selection of the regular hyper-cardioid order. The slider **Output Gain** applies a gain on the output signal. The sliders **Azimuth** and **Elevation** are used for the setting of the steering angles (θ_0, δ_0) of the hyper-cardioid beampattern. The horizontal VU-Meter at the bottom of the GUI shows the output signal level.

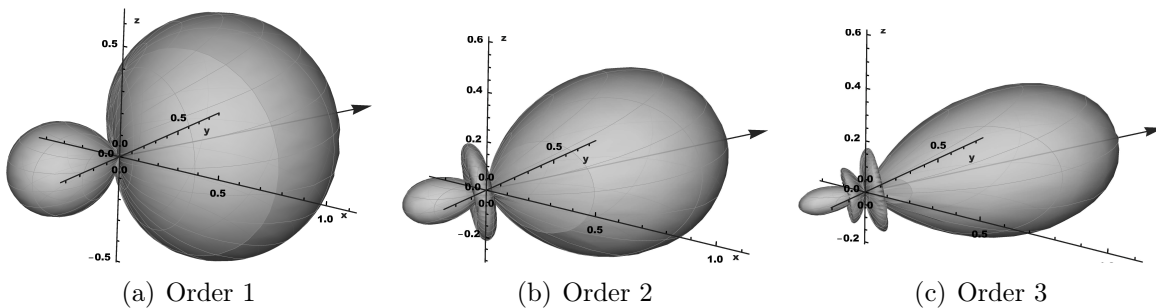


Figure D.14 Regular hyper-cardioid beampatterns up to order 3. The steering angles are $(\theta_0 = 45^\circ, \delta_0 = 10^\circ)$.

hoa_beamforming_hypercardioid_to_hoa

- Inputs: $(M + 1)^2$
- Outputs: $(M_1 + M + 1)^2$

`hoa_beamforming_hypercardioid_to_hoa` performs a directional filtering with directional pattern described by regular hyper-cardioid beampatterns of order M_1 , as shown in Fig. D.14. Thus, unlike the tool `hoa_beamforming_hypercardioid_to_mono` where the output is a monophonic signal, the outputs are, in this case, Ambisonic signals [Lecomte *et al.*, 2016c]. The transformation applies a beampattern on the sound field to enhance some directions and reject other directions. The output HOA scene should be of higher order than the input HOA scene to avoid the lost of spatial informations after the filtering [Lecomte *et al.*, 2016c]. In fact, the input scene has $(M + 1)^2$ components, and the output scene should have $(M + M_1 + 1)^2$ components with a beampattern of order M_1 . The graphical user interface when using `faust2jaqt` script is shown in Fig. D.15 for $M = 3$ and $M_1 = 2$. In this figure, one observes that the input scene is of order $M = 3$ (16 Ambisonic signals), and the output scene is of order $M_1 + M = 4$ (25 active Ambisonic components). In fact, the slider **Order** is set to 1, thus $M_1 = 1$ and $M_1 + M = 4$. Note that when the slider **Order** is set to 0, the effect is bypassed as a beampattern of order $M_1 = 0$ is omnidirectional.



Figure D.15 `hoa_beamforming_hypercardioid_to_hoa` compiled using `faust2jaqt` script with $M = 3$ and $M_1 = 2$. The VU-Meters at the top show the inputs signal levels in dBFS. The VU-Meters at the bottom show the outputs signal levels in dBFS. The knob **Azimuth** and **Elevation** are used to set the steering angles of the beampattern. Finally, the slider **Order** is used to set the hyper-cardioid order, from $M_1 = 0$ to $M_1 = 2$ in this case.

To visualize the effect of the transformation on the loudspeaker signals,, Fig. D.16 shows the Spherical VU-Meter (see Sec. D.4.2) for a 50-nodes Lebdev grid on which a sound field is decoded. The sound field is an combination of two plane waves coming from direction $(0^\circ, 0^\circ)$

and $(90^\circ, 0^\circ)$ and was produced using tool `hoa_encoder_N_sources` (see Sec. D.4.1). Figure D.16(a) shows the spherical VU-meter when the knob `Order` in Fig. D.15 is set to 0, i.e the effect is by-passed. Fig. D.16(b) shows the same meter when knob `Order` is set to 1 and knobs `Azimuth` and `Elevation` are set to 0° and 0° respectively. This means the use of a hypercardioid beampattern of order $M_1 = 1$ (Fig. D.14(a)) with steering angles in front of the listener. One notes that the sound energy now comes more from the frontal direction. This is confirmed when the slider `Order` is set to 2 (with same steering angle) in Fig. D.16(c), i.e, with a hypercardioid of order $M_1 = 2$. In conclusion, the tool `hoa_beamforming_hypercardioid_to_hoa` is useful for the directional exploration of a sound scene [Lecomte *et al.*, 2016c].

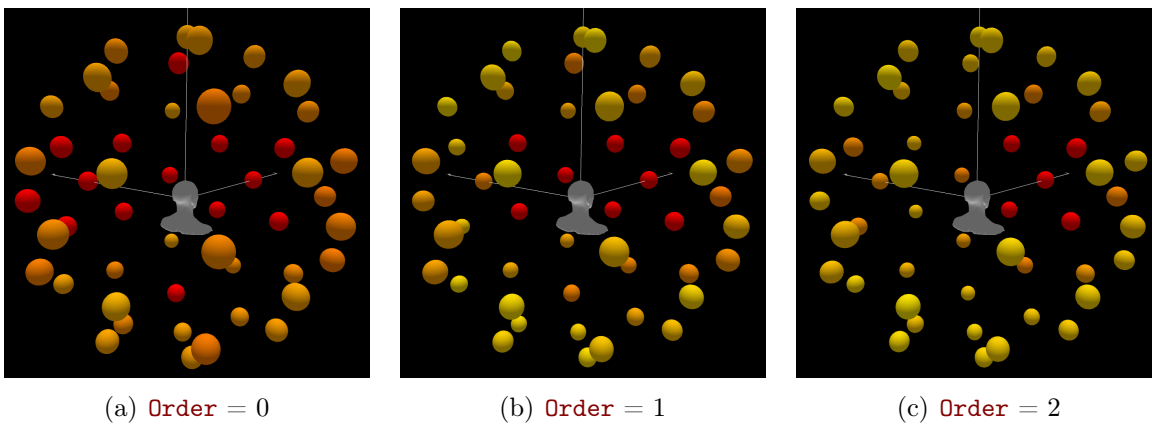


Figure D.16 The effects of `hoa_beamforming_hypercardioid_to_hoa` on the loudspeaker signals for a given sound field. (a) Loudspeakers signal level for the original sound field decoded on a 50-loudspeaker Lebedev grid. (b) the loudspeakers signal level when the knob `Order` is set to 1. (c) the loudspeakers signal level when the knob `Order` is set to 2.

`hoa_beamforming_dirac_to_hoa`

- Inputs: $(M + 1)^2$
- Outputs: $(M + 1)^2$

`hoa_beamforming_dirac_to_hoa` performs a directional filtering of the sound field in order to listen to only one given direction. The transformation is similar to `hoa_beamforming_hypercardioid_to_hoa`, however, the beampattern is a truncated directional Dirac [Lecomte *et al.*, 2016c]. That is to say a function which is null everywhere except for the steering angles (θ_0, δ_0) . Therefore, this tool is useful to listen to a single direction from the sound field. The graphical user interface when using `faust2jaqt` script is shown in Fig. D.17 for $M = 3$. Note that this tool provides a crossfade between the states without and with filtering. This means that one can smoothly change between a sound field without any directional filtering and a sound field where just one direction is selected and decoded accordingly.

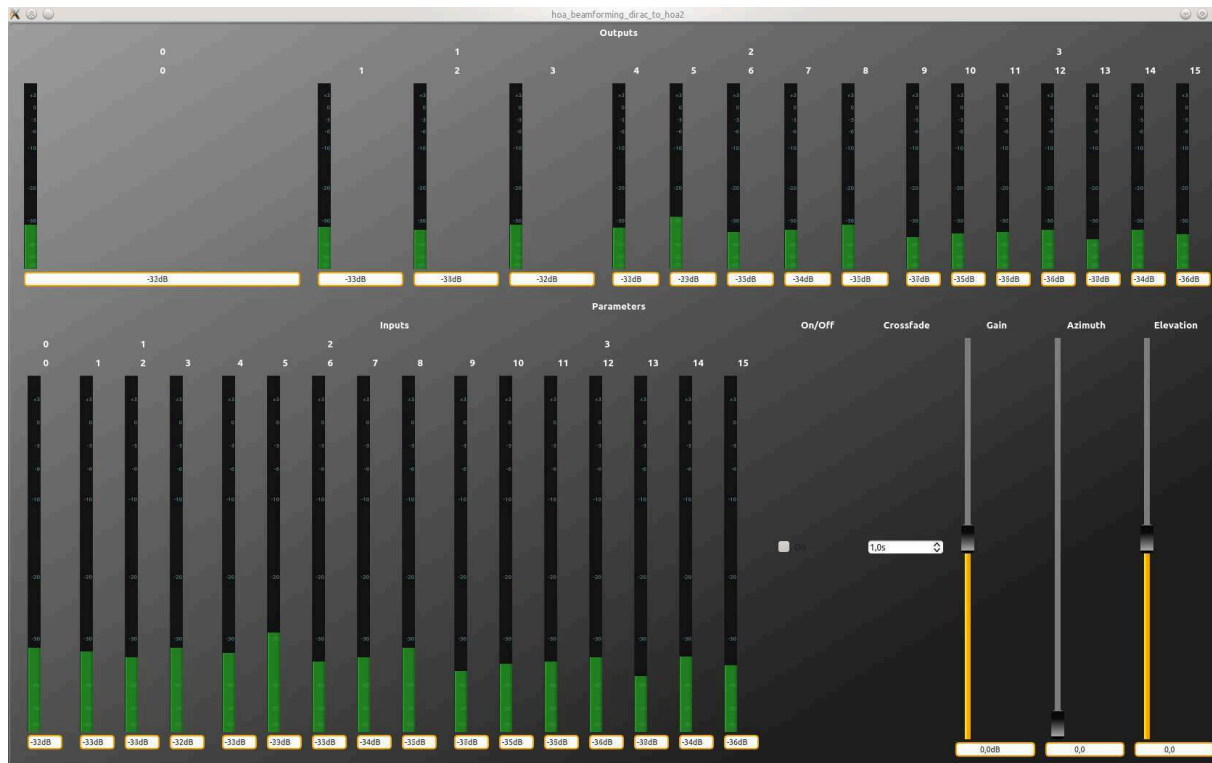


Figure D.17 `hoa_beamforming_dirac_to_hoa2` compiled using `faust2jaqt` script with $M = 3$. The VU-Meters at the top show the inputs signals levels in dBFS. The VU-Meters at the bottom show the output signals levels in dBFS. The check-box **On/Off** is used to bypass the effect with a crossfade. The input **Crossfade** sets the durations of the crossfade in seconds. The slider **Gain** applies a global gain on the output signals. The slider **Azimuth** and **Elevation** set the steering angles of the directional Dirac, i.e the direction to listen to.

To visualize the effect of the transformation on the loudspeaker signals, Fig. D.18 shows the Spherical VU-Meter (see Sec. D.4.2) for a 50-loudspeaker Lebedev grid on which a sound field is decoded. The original sound field was obtained from a sound capture with a rigid spherical array of microphones using a 50-node Lebedev grid [Lecomte *et al.*, 2015]. It is encoded using the tool `hoa_mic_encoder_lebedev50` (see Sec. D.4.1) and `Jconvolver` filters (see Sec. D.4.3). Figure D.18(a) shows the original sound field. No particular direction predominant the signals. However, once the spatial filtering is active (**On/Off** checked in Fig. D.17), a predominant direction is visible for the signals as shown in Fig. D.18(b). The steering angles are set to $(\theta_0 = 0^\circ, \delta_0 = -45^\circ)$. One then notes that this direction corresponds to the main sound energy direction in the resulting filtered and decoded sound field.

D.4.2 Processing

Spherical_VU_Meter

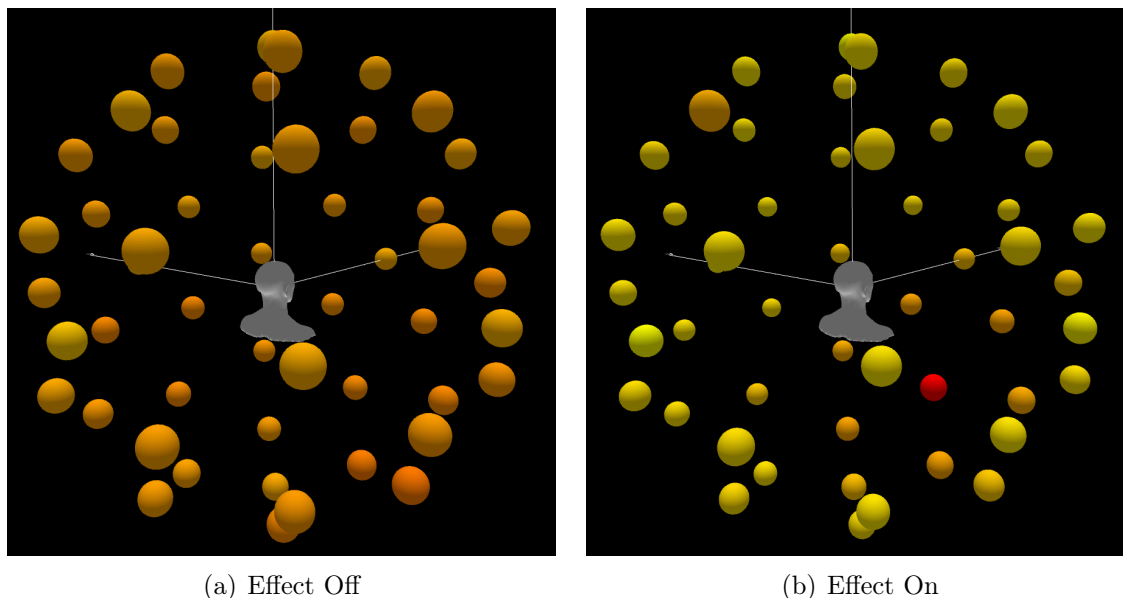


Figure D.18 `hoa_beamforming_dirac_to_hoa` effects on a sound field. (a) loudspeakers energy for the original sound field decoded on a Lebedev grid. (a) the sound field after activation of the directional filtering.

ambitools provides a Spherical VU-Meter developed with PROCESSING language. A screenshot is shown in Fig. D.19 for a 50-node Lebedev grid [Lecomte *et al.*, 2015]. This tool is intended for the visualization of the energy of the loudspeaker signals as function of space. With this tool, it is possible to zoom, pan and rotate with the mouse. The FAUST tools `hoa_panning*` and `hoa_decoder*` emit OSC messages on port UDP⁷ 5511. These messages drive the spherical VU-Meter rendering in real-time: loudspeakers size and color for the meter. They both correspond to moving Root-Mean Square (RMS) values in dBFS. On the left, a color scale in dBFS is provided. Furthermore, this tool is used for the visualization of virtual sources in 3D space. The virtual sources are shown as small color balls with a fading trail when they move. Spherical and Cartesian coordinates of the virtual sources are also shown on the left. The tools `hoa_encoder*` and `hoa_panning*` emit OSC messages on port UDP 5511 with values of virtual source position to drive the spherical VU-Meter rendering.

Note that the meter is here presented for a 50-node Lebedev grid and for two virtual sources. However, the code is flexible and a meter can be compiled for other loudspeaker configurations and number of virtual sources. Indeed, the loudspeaker Cartesian coordinates are provided as a `.csv` file in folder `/Processing/src/Spherical_VU_Meter/data/`. Therefore, to change the meter configuration, the user must simply provide the Cartesian coordinates of the loudspeakers as a `.csv` file and make appropriate reference to this file in the code.

⁷UDP : User Datagram Protocol

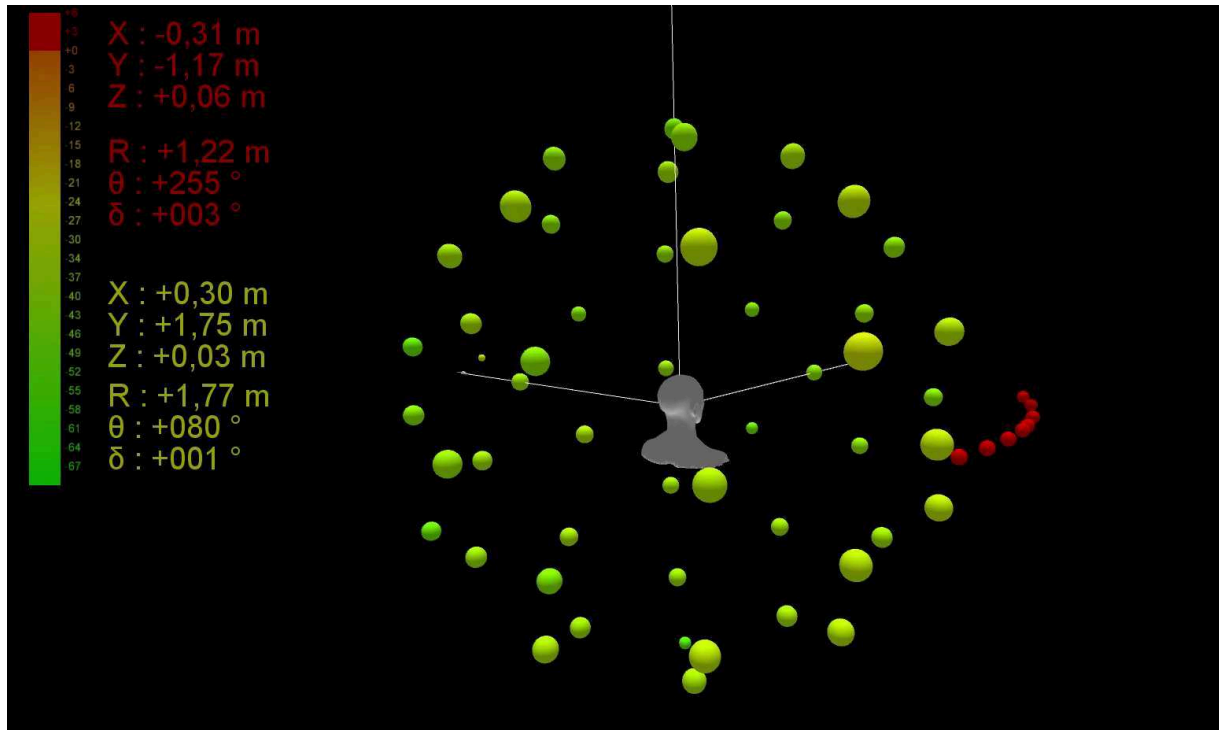


Figure D.19 **Spherical_VU_Meter** for a 50-loudspeaker Lebedev array and two virtual sources. Each loudspeaker is represented as a color ball with size and color proportional to RMS level in dBFS. A scale in dBFS is shown on the left. The virtual sources are represented as red and yellow dots with fading trails when they move. The spherical and Cartesian coordinates of the virtual sources are also shown on the left. A fictitious grey manikin head and torso is placed at the origin to indicate frontal direction.

D.4.3 Jconvolver

Jconvolver is a real-time fast convolution software written by Fons Adriaensen^[5]. It allows to do convolution operation with FIR filters. *ambitools* uses FIR filters for two main tasks:

- Binaural rendering with Head Related Impulse Responses (HRIRs).
- Radial equalization filters for rigid spherical microphone.

The filters are provided as **.wav** files containing the impulse responses. They are located in the folder **/FIR** as well as the *Jconvolver* configuration files. The latter are text files defining the vector or matrix of filters.

jconvolver_mic*

The folders **FIR/spherical_microphones/jconvolver_mic*/** contain the configuration files and impulse responses of radial equalization filters for several rigid microphone arrays. These filters are unstable by nature and are stabilized by regularization, as in [Moreau et al., 2006](#). They are applied after the DSFT performed by **hoa_mic_encoder*** tool (see Sec. [D.4.1](#)). They are mandatory to obtain the Ambisonic components in the context of

sound field capture by rigid spherical microphone arrays. Thus, the connections should be as in Fig. [D.4](#). Without providing more mathematical details (see [Daniel, 2000](#); [Lecomte et al., 2015](#); [Moreau et al., 2006](#) for instance), it is reminded that there is one filter by order. The filter at order m is replicated $(2m + 1)$ times, for each m -th order component from the DSFT. For example, the rigid spherical microphone array by mh acoustics, the EigenMike™, works up to order $M = 4$ [Elko et al., 2009](#), there are 4 filters (i.e. 4 `.wav` files) in folder `/FIR/jconvolver_mic_eigenmike32`.

To start the convolution under JACK, the first thing to do is to edit the configuration file, for example `mic_eigenmike32.conf`. Edit and uncomment the following line

```
#!/cd /put/here/your/local/directory/adress/
```

with the current absolute path address on the local installation. Then, open a terminal in the mentioned folder and type the following command:

```
$jconvolver mic_eigenmike.conf
```

One should now have a JACK client named `Jconvolver` with $(M + 1)^2$ inputs and $(M + 1)^2$ outputs, as in Fig. [D.4](#).

hrir*

The folder `FIR/hrir/` contains several set of Head Related Impulses Responses (HRIRs) computed on various grids. Using these HRIRs, it is possible to achieve binaural rendering of Ambisonics decoding [Noisternig et al., 2003](#). The signals are decoded on a virtual grid of L loudspeakers, and each virtual loudspeaker signal is convolved with the corresponding HRIR for the left and right ears. The summation of all speakers signals for the left and right ears gives the binaural signals. Thus, this operation is computed based on a matrix of filters with L inputs and 2 outputs. For binaural rendering, the `Jconvolver` JACK client should be inserted after the decoder as shown in Fig. [D.20](#). To start the convolution under JACK, the first thing to do is to edit the configuration file, for example `hrir_ku100_lebedev50.conf`. Edit and uncomment the line

```
#!/cd /put/here/your/local/directory/adress/
```

with the current absolute path address on ones local installation. Then, open a terminal in this folder and type the following command:

```
$jconvolver hrir_lebedev50.conf
```

One should now have a JACK client named `Jconvolver` with L inputs and 2 outputs, as shown in Fig. [D.20](#).

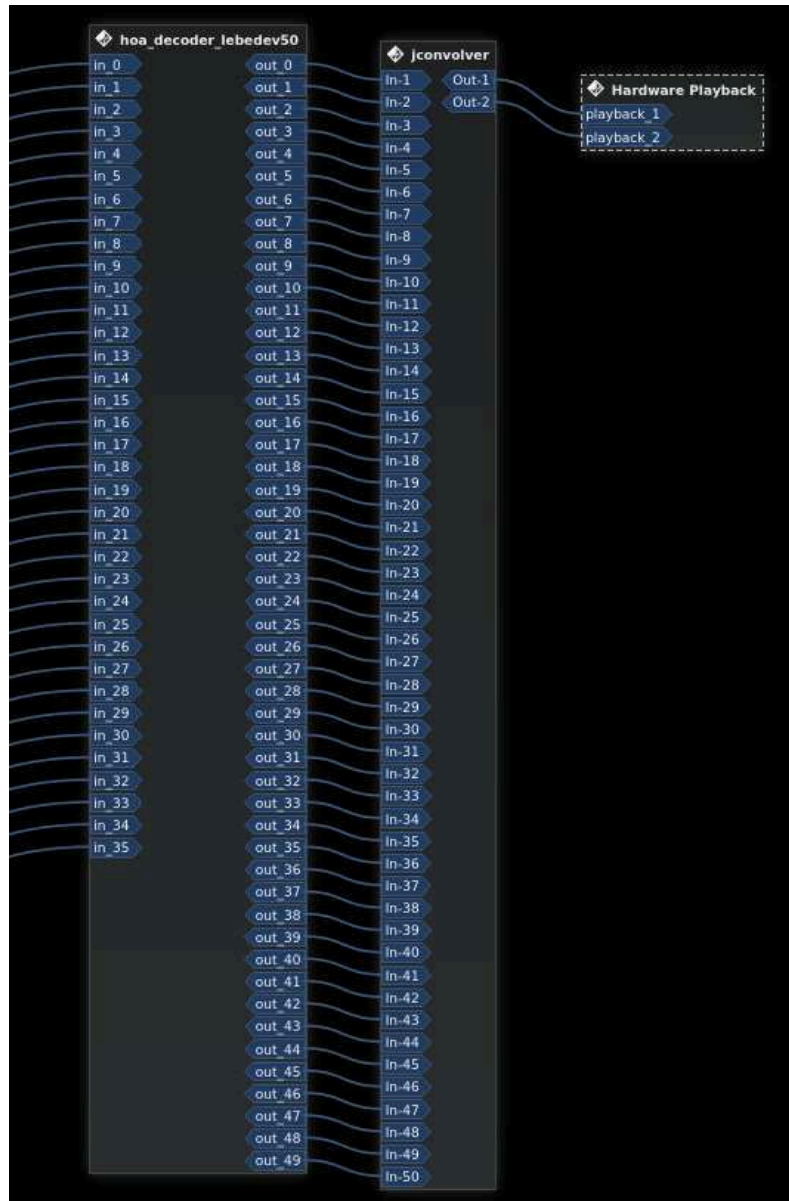


Figure D.20 Connection between `hoa_decoder_lebedev50` tool and *Jconvolver* for binaural rendering. Screenshot taken from Claudia under KXStudio Linux distribution.

D.4.4 Pure Data

ambitools uses *Pure-Data* to generate some sounds but also to emit or convert OSC messages. Particularly the libraries [mrpeach](https://puredata.info/downloads/mrpeach)⁸ and [hid](https://puredata.info/downloads/hid)⁹ are used to manipulate OSC instructions signals of a USB joystick respectively.

andOSC.pd for Head Tracking

The patch [andOSC.pd](#) is used to convert the OSC instruction send from the free Android application [andOSC](#) running on a smartphone. This application transmits on local network OSC message with the values of accelerometer and gyroscope of a smartphone. The GUI of [andOSC](#) is shown in Fig. [D.21](#). The value of interest in this figure is the first value of orientation, which correspond to the yaw angle. Basically, the Pure Data patch [andOSC.pd](#) listen to the information sent by the [andOSC](#) application and extract the yaw angle to drive the tool [hoa_azimuth_rotator](#) (see Sec. [D.4.1](#)). The graphical user interface of the patch [andOSC.pd](#) is shown in Fig. [D.22](#). The yaw angle value is extracted and lowpass filtered to avoid abrupt change of rotation values for the [hoa_azimuth_rotator](#) tool and to limit audible click when abrupt changes in rotation are requested. The positive sens of rotation for [andOSC](#) is clockwise, and it is anti-clockwise for *ambitools*. Therefore, the value of yaw angle transmitted by [andOSC](#) is the head rotation compensation angle needed for head-tracking with [hoa_azimuth_rotator](#). In order to run the transmission of OSC message, the DSP button should be engaged in *Pure-Data*. Also, the tool [hoa_azimuth_rotator](#) should listen to OSC messages on port 6000. To do so, launch [hoa_azimuth_rotator](#) with the following command:

```
$. /hoa_azimuth_rotator -port 6000
```

D.5 Examples

This section presents examples of required connections for various practical scenarios.

Binaural listening of 3D HOA recording

```
Ardour -> hoa_mic_encoder_lebedev50 -> jconvolver_mic_lebedev50 ->
hoa_azimuth_rotator -> hoa_decoder_lebedev50 ->
jconvolver_hrir_ku100_lebedev50 -> outputs
```

Spatial audio with virtual sources

```
hoa_encoder_N_sources -> hoa_decoder* -> outputs,
or
hoa_panning_N_sources_* -> outputs
```

⁸<https://puredata.info/downloads/mrpeach>

⁹<https://puredata.info/downloads/hid>

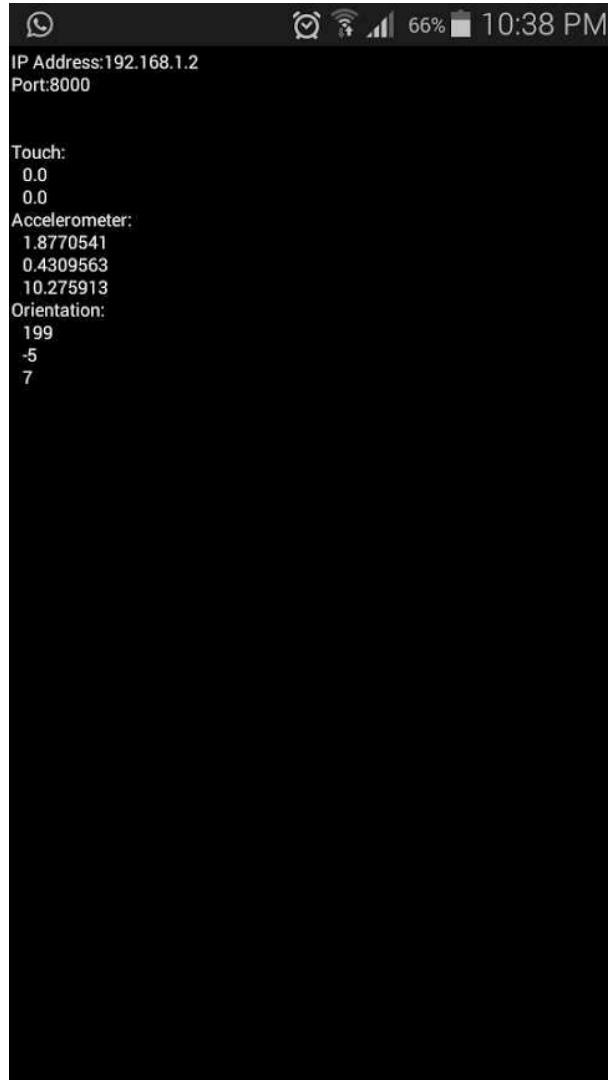


Figure D.21 Screenshot of application [andOSC](#) on Android. The value of accelerometer and gyroscope are sent by OSC on the local network.

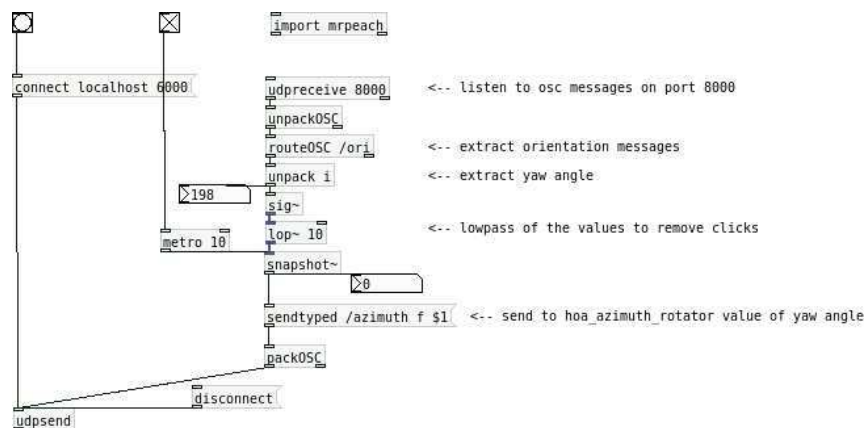


Figure D.22 Graphical user interface of patch [andOSC.pd](#). By default, the listening port on is UDP 8000 and transmission port is UDP 6000.

LISTE DES RÉFÉRENCES

- Abramowitz, M. et Stegun, I. A. (1972). *Handbook of mathematical functions*, 10^e édition. Dover Publications, 1046 p.
- AES (1991). AES Recommended Practice for Digital Audio Engineering - Serial Multi-channel Audio Digital Interface (MADI).
- Ahrens, C. et Beylkin, G. (2009). Rotationally invariant quadratures for the sphere. *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences*, volume 465, numéro 2110, p. 3103–3125.
- Ahrens, J. (2012). *Analytic Methods of Sound Field Synthesis*. Springer, 299 p.
- Ahrens, J. et Spors, S. (2008a). Analytical driving functions for higher order ambisonics. Dans *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, Las Vegas, p. 373–376.
- Ahrens, J. et Spors, S. (2008b). Focusing of virtual sound sources in higher order Ambisonics. Dans *Audio Engineering Society Convention 124*. AES, Amsterdam, p. 1–9.
- Ahrens, J. et Spors, S. (2012). Wave field synthesis of a sound field described by spherical harmonics expansion coefficients. *The Journal of the Acoustical Society of America*, volume 131, numéro 3, p. 2190–2199.
- Ahrens, J., Spors, S. et Wierstorf, H. (2010). Comparison of higher order ambisonics and wave field synthesis with respect to spatial discretization artifacts in time domain. Dans *Audio Engineering Society Conference : 40th International Conference : Spatial Audio : Sense the Sound of Space*. Audio Engineering Society, p. 1–8.
- Ajdler, T., Sbaiz, L. et Vetterli, M. (2006). The plenacoustic function and its sampling. *IEEE Transactions on Signal Processing*, volume 54, numéro 10, p. 3790–3804.
- Allen, J. B. et Berkley, D. A. (1979). Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, volume 65, numéro 4, p. 943–950.
- António, J., Godinho, L. et Tadeu, A. (2002). Reverberation times obtained using a numerical model versus those given by simplified formulas and measurements. *Acta Acustica United with Acustica*, volume 88, numéro 2, p. 252–261.
- Arfken, G. (1985). *Mathematical Methods for Physicists Academic*, 3^e édition. Academic press, New York, 985 p.
- Arfken, G. B. et Weber, H. J. (2005). *Mathematical methods for physicists*, 6^e édition. Elsevier, 1200 p.
- Atalla, N. et Sgard, F. (2015). *Finite Element and Boundary Methods in Structural Acoustics and Vibration*. CRC Press, 470 p.

- Bamford, J. S. (1995). *An Analysis of Ambisonic Sound Systems of First and Second Order*. Thèse de doctorat, University of Waterloo, Waterloo, 270 p.
- Bauck, J. et Cooper, D. H. (1996). Generalized transaural stereo and applications. *Journal of the Audio Engineering Society*, volume 44, numéro 9, p. 683–705.
- Baumgartner, R., Pomberger, H. et Frank, M. (2011). Practical implementation of radial filters for ambisonic recordings. Dans *1st International Conference on Spatial Audio*. ICSA, Detmold, p. 1–6.
- Beranek, L. L. (1993). *Acoustics*. Acoustical Society of America, Cambridge, 491 p.
- Berkhout, A. J. (1988). A Holographic Approach to Acoustic Control. *Journal of the Audio Engineering Society*, volume 36, numéro 12, p. 977–995.
- Berkhout, A. J., de Vries, D. et Vogel, P. (1993). Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America*, volume 93, numéro 5, p. 2764–2778.
- Bernschütz, B. (2013). A spherical far field hrir/hrtf compilation of the neumann ku 100. Dans *40th Italian Annual Conference on Acoustics (AIA) and the 39th German Annual Conference on Acoustics (DAGA) Conference on Acoustics*. DAGA, Merano, p. 1–4.
- Betlehem, T. et Abhayapala, T. D. (2005). Theory and design of sound field reproduction in reverberant rooms. *The Journal of the Acoustical Society of America*, volume 117, numéro 4, p. 2100–2111.
- Betlehem, T. et Poletti, M. A. (2014). Two dimensional sound field reproduction using higher order sources to exploit room reflections. *The Journal of the Acoustical Society of America*, volume 135, numéro 4, p. 1820–1833.
- Blauert, J. (1997). *Spatial hearing : the psychophysics of human sound localization*. MIT press, Cambridge, 508 p.
- Blauert, J. (2013). *The technology of binaural listening*. Springer, Heidelberg, 526 p.
- Boone, M. M., Verheijen, E. N. et van Tol, P. F. (1995). Spatial sound-field reproduction by wave-field synthesis. *Journal of the Audio Engineering Society*, volume 43, numéro 12, p. 1003–1012.
- Braikia, Y., Melon, M., Langrenne, C., Bavu, É. et Garcia, A. (2013). Evaluation of a separation method for source identification in small spaces. *The Journal of the Acoustical Society of America*, volume 134, numéro 1, p. 323–31.
- Bruneau, M. (1998). *Manuel d'acoustique fondamentale*. Hermes, 576 p.
- Camier, C., Gauthier, P.-A., Pasco, Y. et Berry, A. (2010). Sound Field Reproduction Applied to Flight Vehicles Sound Environments. Dans *Audio Engineering Society Conference : 40th International Conference : Spatial Audio : Sense the Sound of Space*. AES, Tokyo, p. 1–10.
- Chapman, M. (2009). Symmetries of Spherical Harmonics : applications to ambisonics. Dans *Ambisonics Symposium*. IEM, Graz, p. 1–14.

- Chapman, M., Ritsch, W., Musil, T., Zmölzig, J., Pomberger, H., Zotter, F. et Sontacchi, A. (2009). A Standard for Interchange of Ambisonic Signal Sets. Including a file standard with metadata. Dans *Ambisonics Symposium*. IEM, Graz, p. 1–6.
- Choi, J.-W. et Kim, Y.-H. (2002). Generation of an acoustically bright zone with an illuminated region using multiple sources. *The Journal of the Acoustical Society of America*, volume 111, numéro 4, p. 1695–1700.
- Choi, J.-W. et Kim, Y.-H. (2012). Theories and signal processing techniques for the implementation of sound ball in space using loudspeaker array. Dans *Signal Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*. IEEE, Hollywood, p. 1–6.
- Chouquet, E. et Hamoui, A. (2014). *Design of a wide band speaker to use with High Order Ambisonics soundfield synthesis*. Thèse de doctorat, Université du Maine, Le Mans, 25 p.
- Cohen, I., Benesty, J. et Gannot, S. (2009). *Speech processing in modern communication : challenges and perspectives, volume 3*. Springer, Heidelberg, 342 p.
- Colton, D. et Kress, R. (2013a). *Integral equation methods in scattering theory, volume 72*. SIAM, Philadelphia, 271 p.
- Colton, D. et Kress, R. (2013b). *Inverse acoustic and electromagnetic scattering theory*, 3^e édition. Springer, New York, 406 p.
- Copley, L. G. (2005). Fundamental results concerning integral representations in acoustic radiation. *The journal of the acoustical society of America*, volume 44, numéro 1, p. 28–32.
- Corteel, É. (2006). Equalization in an extended area using multichannel inversion and wave field synthesis. *Journal of the audio Engineering Society*, volume 54, numéro 12, p. 1140–1161.
- Daniel, J. (2000). *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. Thèse de doctorat, Université Paris 6, Paris, 319 p.
- Daniel, J. (2003). Spatial sound encoding including near field effect : Introducing distance coding filters and a viable, new ambisonic format. Dans *Audio Engineering Society Conference : 23rd International Conference : Signal Processing in Audio Recording and Reproduction*. AES, Helsingør, p. 1–15.
- Daniel, J. et Moreau, S. (2004). Further study of sound field coding with higher order ambisonics. Dans *Audio Engineering Society Convention 116*. AES, Berlin, p. 1–14.
- Daniel, J., Moreau, S. et Nicol, R. (2003). Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging. Dans *Audio Engineering Society Convention 114*. AES, Amsterdam, p. 1–18.

- Daniel, J., Rault, J.-B. et Polack, J.-D. (1998). Ambisonics encoding of other audio formats for multiple listening conditions. Dans *Audio Engineering Society Convention 105*. AES, San Francisco, p. 1–29.
- Deprez, R., Nicol, R. et Friot, E. (2010). Validation théorique de la correction des réflexions sur la base d’une représentation en harmoniques sphériques. Dans *10ème Congrès Français d’Acoustique*. CFA, Lyon, p. 1–6.
- Driscoll, J. R. et Healy, D. M. (1994). Computing Fourier transforms and convolutions on the 2-sphere. *Advances in applied mathematics*, volume 15, p. 202–250.
- Dupont, S. (2016). *Sound Field Synthesis With Higher-Order Ambisonics*. Thèse de doctorat, Université du Maine, Le Mans, 39 p.
- Elko, G., Kubli, R. A. et Meyer, J. (2009). Audio system based on at least second-order eigenbeams.
- Elliott, S. J. (2001). *Signal Processing For Active Control*. Academic Press, London, 511 p.
- Farina, A. (2000). Simultaneous measurement of impulse response and distortion with a swept-sine technique. Dans *Audio Engineering Society Convention 108*. AES, p. 23.
- Fazi, F. M. (2010). *Sound field reproduction*. Thèse de doctorat, University of Southampton, Southampton, 313 p.
- Fazi, F. M. et Nelson, P. A. (2013). Sound field reproduction as an equivalent acoustical scattering problem. *The Journal of the Acoustical Society of America*, volume 134, numéro 5, p. 3721–3729.
- Fazi, F. M., Nelson, P. A. et Potthast, R. (2009). Analogies and differences between three methods for sound field reproduction. Dans *Ambisonics Symposium*. IEM, Graz, p. 1–9.
- Fielder, L. D. (2001). Practical limits for room equalization. Dans *Audio Engineering Society Convention 111*. AES, New York, p. 1–20.
- Fliege, J. et Maier, U. (1996). A two-stage approach for computing cubature formulae for the sphere. Dans *Mathematik 139T, Universität Dortmund*. p. 1–31.
- Gauthier, P.-A. (2007). *Synthèse de champs sonores adaptative*. Thèse de doctorat, Université de Sherbrooke, Sherbrooke, 420 p.
- Gauthier, P.-A. et Berry, A. (2006). Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction : Theory. *The Journal of the Acoustical Society of America*, volume 119, p. 2721.
- Gauthier, P.-A. et Berry, A. (2007). Adaptive Wave Field Synthesis for Sound Field Reproduction : Theory, Experiments, and Future Perspectives. *J. Audio Eng. Soc.*, volume 55, numéro 12, p. 1107–1124.
- Gauthier, P.-A., Camier, C., Lebel, F.-A., Pasco, Y. et Berry, A. (2012). Experiments of sound field reproduction inside aircraft cabin mock-up. Dans *Audio Engineering Society Convention 133*. AES, San Francisco, p. 1–18.

- Gauthier, P.-A., Camier, C., Padois, T., Pasco, Y. et Berry, A. (2015). Sound Field Reproduction of Real Flight Recordings in Aircraft Cabin Mock-Up. *J. Audio Eng. Soc.*, volume 63, numéro 1/2, p. 6–20.
- Gauthier, P.-A., Chambatte, É., Camier, C., Pasco, Y. et Berry, A. (2011). Derivation of ambisonics signals and plane wave description of measured sound field using irregular microphone arrays and inverse problem theory. Dans *Ambisonics Symposium*. IEM, Lexington, p. 1–17.
- Gerzon, M. A. (1973). Periphony : With-height sound reproduction. *Journal of the Audio Engineering Society*, volume 21, numéro 1, p. 2–10.
- Gerzon, M. A. (1975). The design of precisely coincident microphone arrays for stereo and surround sound. Dans *Audio Engineering Society Convention 50*, Audio Engineering Society. AES, Londres, p. 1–5.
- Gerzon, M. A. (1985). Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society*, volume 33, numéro 11, p. 859–871.
- Gerzon, M. A. (1992). General metatheory of auditory localisation. Dans *Audio Engineering Society Convention 92*. AES, Vienne, p. 2–63.
- Gerzon, M. A. et Barton, G. J. (1992). Ambisonic decoders for HDTV. Dans *Audio Engineering Society Convention 92*, Audio Engineering Society. AES, p. 42.
- Golub, G. H. et Van Loan, C. F. (2012). *Matrix computations, volume 3*. Johns Hopkins University Press, Baltimore, 728 p.
- Grasgruber, P., Cacek, J., Kalina, T. et Sebera, M. (2014). The role of nutrition and genetics as key determinants of the positive height trend. *Economics & Human Biology*, volume 15, p. 81–100.
- Guastavino, C. et Katz, B. F. G. (2004). Perceptual evaluation of multi-dimensional spatial audio reproduction. *The Journal of the Acoustical Society of America*, volume 116, p. 1105.
- Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems : Numerical aspects of linear inversion*. Society For Industrial and Applied Mathematics, Philadelphia, 259 p.
- Hardin, R. H. et Sloane, N. J. A. (1996). McLaren’s improved snub cube and other new spherical designs in three dimensions. *Discrete & Computational Geometry*, volume 15, numéro 4, p. 429–441.
- Hartmann, W. M. (1997). *Signals, sound, and sensation*. Springer, New York, 647 p.
- Heller, A. J., Benjamin, E. M. et Lee, R. (2012). A toolkit for the design of ambisonic decoders. Dans *Linux Audio Conference*. LAC, Stanford, p. 1–12.
- ISO8000-2 (2009). *Quantities and units - Part 2 : Mathematical signs and symbols to be used in the natural sciences and technology*. International Organization for Standardization, Geneva.

- Ivanic, J. et Ruedenberg, K. (1996). Rotation matrices for real spherical harmonics. Direct determination by recursion. *The Journal of Physical Chemistry*, volume 100, numéro 15, p. 6342–6347.
- Jacovitti, G. et Scarano, G. (1993). Discrete time techniques for time delay estimation. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, volume 41, numéro 2, p. 525–533.
- Kirkeby, O. et Nelson, P. A. (1993). Reproduction of plane wave sound fields. *The Journal of the Acoustical Society of America*, volume 94, numéro 5, p. 2992–3000.
- Kirkeby, O. et Nelson, P. A. (1999). Digital Filter Design for Inversion Problems in Sound Reproduction. *Journal of the Audio Engineering Society*, volume 47, numéro 7/8, p. 583–595.
- Kirkeby, O., Nelson, P. A., Hamada, H. et Orduna-Bustamante, F. (1998). Fast deconvolution of multichannel systems using regularization. *IEEE Transactions on Speech and Audio Processing*, volume 6, numéro 2, p. 189–194.
- Kolundzija, M., Faller, C. et Vetterli, M. (2011). Reproducing sound fields using MIMO acoustic channel inversion. *Journal of the Audio Engineering Society*, volume 59, numéro 10, p. 721–734.
- Koretz, A. et Rafaely, B. (2009). Dolph–Chebyshev beam pattern design for spherical arrays. *IEEE Transactions on Signal Processing*, volume 57, numéro 6, p. 2417–2420.
- Kronlachner, M. (2014). *Spatial Transformations for the Alteration of Ambisonic Recordings*. Thèse de doctorat, Graz University Of Technology, Graz, 73 p.
- Kuttruff, H. (2009). *Room acoustics*, 5^e édition. Spon Press, Abingdon, 392 p.
- Laakso, T. et Välimäki, V. (1996). Splitting the unit delay. *IEEE Signal Processing Magazine*, volume 13, p. 30–60.
- Lebedev, V. (1975). Values of the nodes and weights of quadrature formulas of Gauss–Markov type for a sphere from the ninth to seventeenth order of accuracy that are invariant with respect to an octahedron. *USSR Computational Mathematics and Mathematical Physics*, volume 15, numéro 1, p. 44–51.
- Lebedev, V. I. (1976). Quadratures on a sphere. *USSR Computational Mathematics and Mathematical Physics*, volume 16, numéro 2, p. 10–24.
- Lecomte, P. et Gauthier, P.-A. (2015). Real-Time 3D Ambisonics using Faust, Processing, Pure Data, And OSC. Dans *15th International Conference on Digital Audio Effects (DAFx-15)*. DAFx, Trondheim, p. 1–8.
- Lecomte, P., Gauthier, P.-A., Langrenne, C., Berry, A. et Garcia, A. (2016a). A Fifty-Node Lebedev Grid And Its Applications To Ambisonics. *Journal of the Audio Engineering Society*, volume 64, numéro 11, p. 868–881.

- Lecomte, P., Gauthier, P.-A., Langrenne, C., Berry, A. et Garcia, A. (2016b). Cancellation of Room Reflexions Over an Extended Area using Ambisonics. *Soumis à The Journal of the Acoustical Society of America (en cours de révision)*.
- Lecomte, P., Gauthier, P.-A., Langrenne, C., Berry, A. et Garcia, A. (2016c). Filtrage directionnel dans un scène sonore 3D par une utilisation conjointe de Beamforming et d'Ambisonie d'ordre élevés. Dans *CFA / VISHNO 2016*. SFA, Le Mans, p. 169–175.
- Lecomte, P., Gauthier, P.-A., Langrenne, C., Garcia, A. et Berry, A. (2015). On the use of a Lebedev grid for Ambisonics. Dans *Audio Engineering Society Convention 139*. AES, New York, p. 1–12.
- Lehmann, E. A. et Johansson, A. M. (2008). Prediction of energy decay in room impulse responses simulated with an image-source model. *The Journal of the Acoustical Society of America*, volume 124, numéro 1, p. 269–277.
- Li, Z. et Duraiswami, R. (2007). Flexible and Optimal Design of Spherical Microphone Arrays for Beamforming. *IEEE Transactions on Audio, Speech and Language Processing*, volume 15, numéro 2, p. 702–714.
- Li, Z., Duraiswami, R., Grassi, E. et Davis, L. S. (2004). Flexible layout and optimal cancellation of the orthonormality error for spherical microphone arrays. Dans *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, Montréal, p. 1–4.
- Linkwitz, S. H. (1976). Active crossover networks for noncoincident drivers. *Journal of the Audio Engineering Society*, volume 24, numéro 1, p. 2–8.
- Lösler, S. et Zotter, F. (2015). Comprehensive radial filter design for practical higher-order Ambisonic recording. *Fortschritte der Akustik, DAGA*, p. 4.
- Malham, D. (1992). Experience with a large area 3d ambisonic sound systems. Dans *Proceedings-Institute of Acoustics*. Institute of Acoustics, p. 1–7.
- Malham, D. (1999). Higher order Ambisonic systems for the spatialisation of sound. Dans *Proceedings, ICMC99, Beijing*. ICMC, p. 1–4.
- Malham, D. (2003). *Higher order ambisonic systems*. Thèse de doctorat, University Of York, York.
- Malham, D. G. (1990). Ambisonics-A Technique for Low Cost, High Precision Three-dimensional Sound Diffusion.
- Malham, D. G. et Myatt, A. (1995). 3-D sound spatialization using ambisonic techniques. *Computer Music Journal*, volume 19, numéro 4, p. 58–70.
- Meyer, J. et Elko, G. (2002). A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing*. volume 2. IEEE, p. 1781–1784.
- Møller, H. (1992). Fundamentals of binaural technology. *Applied Acoustics*, volume 36, p. 171–218.

- 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, Le Mans, 213 p.
- Moreau, S., Daniel, J. et Bertet, S. (2006). 3d sound field recording with higher order ambisonics-objective measurements and validation of spherical microphone. Dans *Audio Engineering Society Convention 120*. AES, Paris, p. 1–24.
- Morse, P. M. et Ingard, K. U. (1986). *Theoretical acoustics*. Princeton University Press, New York, 950 p.
- Nachbar, C., Zotter, F., Deleflie, E. et Sontacchi, A. (2011). Ambix - A suggested ambisonics format. Dans *Ambisonics Symposium*. IEM, Lexington, p. 11.
- Nelson, P. A., Orduna-Bustamante, F. et Hamada, H. (1992). Multi-channel signal processing techniques in the reproduction of sound. *J. Audio Eng. Soc*, volume 44, numéro 11, p. 973–989.
- Nicol, R. (1999). *Restitution sonore spatialisée sur une zone étendue : application à la téléprésence*. Thèse de doctorat, Université du Maine, Le Mans, 480 p.
- Nicol, R. et Emerit, M. (1999). 3D-sound reproduction over an extensive listening area : A hybrid method derived from holophony and ambisonic. Dans *Audio Engineering Society Conference : 16th International Conference : Spatial Sound Reproduction*. AES, p. 436–453.
- 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*, volume 71, p. 30–39.
- Noisternig, M., Sontacchi, A., Musil, T. et Holdrich, R. (2003). A 3d ambisonic based binaural sound reproduction system. Dans *Audio Engineering Society Conference : 24th International Conference : Multichannel Audio, The New Reality*. AES, p. 1–5.
- Oppenheim, A. V. et Schaffer, R. W. (2010). *Discrete-Time Signal Processing*, 3^e édition. Pearson Higher Education, Upper Saddle River, 1120 p.
- Plessas, P. (2009). *Rigid sphere microphone arrays for spatial recording and holography*. Thèse de doctorat, Graz University of Technology, Graz, 71 p.
- Poletti, M. A. (1996). The design of encoding functions for stereophonic and polyphonic sound systems. *Journal of the audio Engineering Society*, volume 44, numéro 11, p. 948–963.
- Poletti, M. A. (2000). A unified theory of horizontal holographic sound systems. *Journal of the Audio Engineering Society*, volume 48, numéro 12, p. 1155–1182.
- Poletti, M. A. (2005). Three-dimensional surround sound systems based on spherical harmonics. *Journal of the Audio Engineering Society*, volume 53, numéro 11, p. 1004–1025.

- Poletti, M. A., Fazi, F. M. et Nelson, P. A. (2010). Sound-field reproduction systems using fixed-directivity loudspeakers. *The Journal of the Acoustical Society of America*, volume 127, p. 3590–3601.
- Pulkki, V. (1997). Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, volume 45, numéro 6, p. 456–466.
- Rafaely, B. (2005). Analysis and design of spherical microphone arrays. *IEEE Transactions on Speech and Audio Processing*, volume 13, numéro 1, p. 135–143.
- Rafaely, B. (2015). *Fundamentals of Spherical Array Processing*. Springer, Heidelberg, 193 p.
- Rafaely, B., Weiss, B. et Bachmat, E. (2007). Spatial aliasing in spherical microphone arrays. *IEEE Transactions on Signal Processing*, volume 55, numéro 3, p. 1003–1010.
- Sabine, W. C. et Egan, M. D. (1994). Collected papers on acoustics. *The Journal of the Acoustical Society of America*, volume 95, numéro 6, p. 3679–3680.
- Sergi, G. (2013). Knocking at the door of cinematic artifice : Dolby Atmos, challenges and opportunities. *The New Soundtrack*, volume 3, numéro 2, p. 107–121.
- Sobolev, S. L. (2006). *Selected Works of S.L. Sobolev - Volume I : Mathematical Physics, Computational Mathematics, and Cubature Formulas*. Springer, New York, 643 p.
- Sobolev, S. L. et Vaskevich, V. L. (1997). *The Theory Of Cubature Formulas*. Kluwer Academic Publishers, Dordrecht, 418 p.
- Spors, S. (2007). Extension of an analytic secondary source selection criterion for wave field synthesis. Dans *Audio Engineering Society Convention 123*. AES, New York, p. 1–15.
- Spors, S. et Ahrens, J. (2008). A comparison of wave field synthesis and higher-order Ambisonics with respect to physical properties and spatial sampling. Dans *Audio Engineering Society Convention 125*. AES, San Francisco, p. 1–17.
- Spors, S., Kuntz, A. et Rabenstein, R. (2003). An approach to listening room compensation with wave field synthesis. Dans *Audio Engineering Society Conference : 24th International Conference : Multichannel Audio, The New Reality*. AES, Banff, p. 1–13.
- Spors, S. et Rabenstein, R. (2006). Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis. Dans *Audio Engineering Society Convention 120*. AES, Paris, p. 1–14.
- Spors, S., Rabenstein, R. et Ahrens, J. (2008). The theory of wave field synthesis revisited. Dans *Audio Engineering Society Convention 124*. AES, Amsterdam, p. 1–19.
- Spors, S., Wierstorf, H., Geier, M. et Ahrens, J. (2009). Physical and perceptual properties of focused sources in wave field synthesis. Dans *Audio Engineering Society Convention 127*. AES, New York, p. 1–19.

- Stefanakis, N., Jacobsen, F. et Sarris, J. (2010). Effort variation regularization in sound field reproduction. *The Journal of the Acoustical Society of America*, volume 128, numéro 2, p. 740–750.
- Steinke, G. (1996). Surround sound-The new phase. Dans *Audio Engineering Society Convention 100*, Audio Engineering Society. AES, Copenhagen, p. 1–44.
- Teutsch, H. (2007). *Modal Array Signal Processing : Principles and Applications of Acoustic Wavefield Decomposition*. Springer, Heidelberg, 253 p.
- Theile, G. (1993). The new sound format : 3/2-stereo. Dans *Audio Engineering Society Convention 94*, Audio Engineering Society. AES, Berlin, p. 1–17.
- 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*, volume 80, numéro 5, p. 809–820.
- Verheijen, E. N. (1998). *Sound reproduction by wave field synthesis*. Thèse de doctorat, Delft University of Technology, Delft, 189 p.
- Vorländer, M. (2007). *Auralization : fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Springer, Heidelberg, 335 p.
- Ward, D. B. et Abhayapala, T. D. (2001). Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on Speech and Audio Processing*, volume 9, numéro 6, p. 697–707.
- Wierstorf, H., Raake, A., Geier, M. et Spors, S. (2013). Perception of Focused Sources in Wave Field Synthesis. *Journal of the Audio Engineering Society*, volume 61, p. 5–16.
- Williams, E. G. (1999). *Fourier Acoustics : Sound Radiation And Nearfield Acoustical Holography*. Academic Press, London, 306 p.
- Williams, E. G., Valdivia, N., Herdic, P. C. et Klos, J. (2006). Volumetric acoustic vector intensity imager. *The Journal of the Acoustical Society of America*, volume 120, numéro 4, p. 1887–1897.
- Yon, S., Tanter, M. et Fink, M. (2003). Sound focusing in rooms. II. The spatio-temporal inverse filter. *The Journal of the Acoustical Society of America*, volume 114, numéro March, p. 3044–3052.
- Zotkin, D. N., Duraiswami, R. et Gumerov, N. A. (2008). Sound field decomposition using spherical microphone arrays. Dans *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, p. 277–280.
- Zotter, F. (2009a). *Analysis and synthesis of sound-radiation with spherical arrays*. Thèse de doctorat, University of Music and Performing Arts, Graz, 182 p.
- Zotter, F. (2009b). Sampling strategies for acoustic holography/holophony on the sphere. Dans *NAG/DAGA 2009 International Conference on Acoustics*. DAGA, Rotterdam, p. 1–4.

- Zotter, F. et Frank, M. (2012). All-round ambisonic panning and decoding. *Journal of the Audio Engineering Society*, volume 60, numéro 10, p. 807–820.
- Zotter, F. et Pomberger, H. (2011). Warping of the recording angle in Ambisonics. Dans *1st International Conference on Spatial Audio*. ICSA, Detmold, p. 1–3.
- Zotter, F., Pomberger, H. et Frank, M. (2009). An alternative ambisonics formulation : Modal source strength matching and the effect of spatial aliasing. Dans *Audio Engineering Society Convention 126*. AES, Munich, p. 1–12.
- Zotter, F., Pomberger, H. et Noisternig, M. (2010). Ambisonic decoding with and without mode-matching : A case study using the hemisphere. Dans *2nd international symposium on ambisonics and spherical acoustics*. IRCAM, Paris, p. 1–11.
- Zotter, F., Pomberger, H. et Noisternig, M. (2012). Energy-preserving ambisonic decoding. *Acta Acustica united with Acustica*, volume 98, numéro 1, p. 37–47.

Abstract:

Sound field synthesis is an active research domain with various musical, multimedia or industrial applications. In the latter case, the accurate reconstruction of the sound field is targeted, which involves answering several scientific questions. Using arrays of microphones and loudspeakers, the capture, synthesis and accurate reconstruction of sound fields are theoretically possible. However, for practical applications, the arrangement of the loudspeakers and the acoustic influence of the restitution room are critical factors to consider in order to ensure the accurate reconstruction of the sound field. In this context, this thesis proposes methods and techniques for the capture, transformations and accurate reconstruction of sound fields in three dimensions based on the Higher Order Ambisonics (HOA) method. A spherical configuration for the array of microphones and loudspeakers is proposed. It follows a fifty-node Lebedev grid that enables the capture and reconstruction of the sound field up to order 5 with HOA formalism. The limitations of this approach, such as the spatial aliasing, are studied in detail. A transformation operation of the sound field is also proposed. The formulation is established in the spherical harmonics domain and enables a directional filtering on the sound field prior to the decoding step. For the reconstruction of the sound field, an original approach, also established in the spherical harmonics domain, can take into account the acoustic influence of the restitution room and the defects of the playback system. This treatment then adapts the synthesis of sound fields to the restitution room, maintaining the theoretical formalism established in free field. Finally, an experimental validation of methods and techniques developed in the thesis is made. In this context, a digital signal processing toolkit is developed. It processes in real-time the microphones, ambisonics, and loudspeaker signals for the sound field capture, transformations, and decoding.

Keywords:

acoustic, spatial audio, higher order ambisonics, multichannel signal processing.

Résumé :

La synthèse de champs sonores est un domaine de recherche actif trouvant de nombreuses applications musicales, multimédias ou encore industrielles. Dans ce dernier cas, la reconstruction précise du champ sonore est souhaitée, ce qui implique de répondre à un certain nombre de questionnements scientifiques. À l'aide de réseaux de microphones et de haut-parleurs, la captation, la synthèse et la reconstruction précise de champs sonores sont théoriquement possibles. Seulement, pour des applications pratiques, la disposition des haut-parleurs et l'influence acoustique du lieu de restitution sont des facteurs cruciaux à prendre en compte pour s'assurer de la bonne reconstruction du champ sonore.

Dans ce contexte, cette thèse de doctorat propose des méthodes et des techniques pour la captation, la transformation et la reconstruction précise de champs sonores en trois dimensions en se basant sur la méthode ambisonique d'ordre élevé. Une configuration sphérique pour le réseau de microphones et de haut-parleurs est proposée. Elle suit un maillage de Lebedev à cinquante points qui permet la captation et la reconstruction du champ sonore jusqu'à l'ordre 5 avec le formalisme ambisonique. Les limitations de cette approche, tel le repliement spatial, sont étudiés en détails. De plus, une opération de transformation du champ sonore est présentée. Elle est établie dans le domaine des harmoniques sphériques et permet d'effectuer un filtrage directionnel avant le décodage pour privilégier certaines directions dans le champ sonore, suivant une fonction de directivité choisie. Pour la reconstruction, une approche originale, également établie dans le domaine des harmoniques sphériques, permet de prendre en compte l'influence acoustique du lieu de restitution, ainsi que les défauts du système de restitution. Ce traitement permet alors d'adapter la synthèse de champs sonores au lieu de restitution, en conservant le formalisme théorique établi en champ libre. Finalement, une validation expérimentale des méthodes et des techniques développées au cours de la thèse est faite. Dans ce contexte, une suite logicielle de synthèse et traitement en temps-réel des champs sonores est développée.

Mots clés :

acoustique, audio spatialisée, ambisonie d'ordre élevé, traitement du signal multicanaux.