

# Caractérisation objective et subjective d'une chaîne de traitement HOA

---

Par Clément Cerles

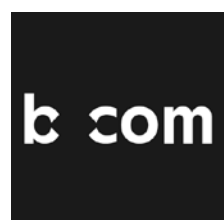
Mémoire de fin d'études

Promotion 2012-2015, Section Son Ecole Nationale Supérieure Louis Lumière

Directeur Interne : M. Frank Gillardeaux

Directeur Externe : M. Jérôme Daniel

Rapporteur : M. Alan Blum



# REMERCIEMENTS

Je tiens d'abord à remercier toute l'équipe de BCOM, l'équipe de Rennes (Jean-Yves Aubie, Danièle Cloatre, David Pesce, Arnaud Longeanie, Marc Emerit, ...) pour m'avoir accordé un équipement, des moyens techniques et un soutien pour mener ce mémoire à bien.

Egalement un grand merci à toute l'équipe de l'antenne de Lannion : Vincent Catros, Fabrice Guillemain, Imad Alawe, Olivier Thorey, Rozenn Nicol, Sergio Morant, Isabelle Hamchaoui, Xavier Priem. Ça n'a pas toujours été évident de remplir ce grand open space.

Un grand merci à Jérôme Daniel pour m'avoir fait découvrir tous les aspects de l'Ambisonie, et fait confiance tout au long de ce mémoire, ça a été un plaisir de pouvoir échanger et d'enregistrer ton quartet !

Je remercie également Frank Gillardeaux, pour son soutien indéfectible pendant ces 5 mois, malgré la distance (et un petit détour par Pleyber-Christ).

Ce mémoire m'a fait découvrir les bars de Lannion, mais aussi la Bretagne grâce aux gens rencontrés ici, aussi merci aux colocs : Antoine, Ulrich, Peter, et aux autres que j'ai croisé : Loic, Cécile, Guillaume, Aurélie, Jean-Marc. Une pensée spéciale pour Manon, aux discussions sur Voltaire autour d'une bière sur une toile cirée jaune, et à l'écoute des messes du XVI<sup>ème</sup> siècle.

Merci à ma famille et spécialement à mon père d'avoir relu et corrigé ce mémoire à la vitesse express.

Merci à Thibaud pour le délai supplémentaire, et merci à Eléonore Mallo...

## Résumé

La technologie HOA (pour Higher Order Ambisonics) offre un large éventail d'outils avancés pour l'enregistrement, la transmission et la diffusion de contenu audio 3D. Elle repose sur un format audio 3D générique, capable de représenter n'importe quel champ sonore, indépendamment de la disposition de prise de son et du système de diffusion (i.e. la configuration des haut-parleurs). Cependant, cela ne veut pas dire que l'écoute finale est strictement indépendante de l'encodage du champ sonore (via le dispositif microphonique) et de son décodage (via les haut-parleurs).

En effet, l'encodage microphonique génère des signaux HOA comprenant des artefacts comparés aux signaux théoriques. Ce traitement, qui repose sur un paramètre d'amplification maximal, demeure un compromis entre la préservation d'un rapport signal sur bruit (SNR) maximal et la largeur de la bande passante fréquentielle et spatiale. De plus, le design des matrices de décodage pour la diffusion donne également différents résultats en fonction du critère privilégié : précision spatiale, stabilité de l'image, enveloppement, distance à la source... Et ceci sur une même configuration de haut-parleurs.

Ce mémoire présente donc un bref retour sur l'histoire de l'Ambisonics et de son évolution HOA, ainsi qu'une étude comparative entre différents encodeurs et décodeurs, appliqués à un réseau de microphones particulier (Eigenmike®) et un dispositif hémisphérique de 14 haut-parleurs et d'un subwoofer. Deux tests subjectifs sont mis en place pour la partie pratique : un premier où l'on testera différents décodeurs issus du commerce et de la littérature, avec des trajectoires de bruit rose dans l'espace. Un deuxième où nous testerons à la fois les qualités audio et spatiales de l'encodeur associé à l'Eigenmike pour différents paramètres d'amplification, et des différents décodeurs. Les signaux utilisés pour ce test seront issus d'enregistrements réalisés avec l'Eigenmike.

Mots clés : HOA, Ambisonics, Audio 3D, Eigenmike, réseau de microphones, harmoniques sphériques, synthèse de champ de pression

## Abstract

HOA technology (for Higher Order Ambisonics) offers a very wide panel of possibilities for 3D audio recording, transmission and diffusion. It is centered on a generic 3D audio format that can represent any soundscape while being agnostic of the microphone array and the rendering setup (e.g. loudspeaker layout). Nevertheless it doesn't mean that the final listening experience is strictly independent from the way the sound scene is spatially encoded (via the microphone) and decoded (on loudspeakers).

On one hand, the microphone array processing generates HOA signals that comprise artifacts compared with theoretical spherical harmonics. The processing design, which is typically tuned through a “maximal amplification parameter”, is a tradeoff between preserving SNR and extending the frequency-spatial bandwidth. On the other hand, the design of decoding matrices for the rendering can imply variable results even over a given loudspeaker configuration, according to which criteria one chooses to focus on: spatial accuracy, stability, envelopment, distance, etc.

This present paper presents a brief review on Ambisonics history and its evolution HOA. It contains a comparative study of different encoding and decoding designs applied to a particular microphone (Eigenmike®) and a hemispherical array of 14 loudspeakers (and one subwoofer). Two tests are conducted: a first one, where several decoders coming from existing solutions and from bibliography are tested with spatial pink noise trajectories. A second one, where we test both spatial and audio qualities of Eigenmike encoder (with several values for the maximum amplification parameter) and HOA decoders, with natural sound recordings.

Keywords : HOA, Ambisonics, 3D Audio, Eigenmike, dense microphone array, spherical harmonics, sound-field synthesis

# Table des matières

1	Introduction.....	7
2	Aspect Théorique.....	10
2.1	Utilisation historique et actuelle de la technologie Ambisonics.....	10
2.2	Le formalisme HOA (2D & 3D) .....	18
2.2.1	Harmonique sphérique.....	18
2.2.2	Convention de normalisation des harmoniques sphériques.....	21
2.3	Encodage spatial .....	24
2.3.1	Encodage spatial d'une source élémentaire l'onde plane.....	24
2.3.2	Encodage microphonique.....	25
2.3.3	Echantillonnage spatial .....	32
2.3.4	Minimisation et définition du paramètre d'amplitude.....	34
2.3.5	Application et caractérisation d'un réseau de microphones.....	36
2.4	Traitements.....	43
2.4.1	Introduction .....	43
2.4.2	Espace .....	43
2.4.3	EQ & Dynamique.....	47
2.4.4	Autres .....	48
2.5	Décodage spatial.....	50
2.5.1	Introduction .....	50
2.5.2	Outils de caractérisation : vecteur vitesse et vecteur énergie.....	54
2.5.3	Caractérisation spatiale d'un décodeur.....	57
3	Elaboration des tests subjectifs.....	62
3.1	Introduction .....	62
3.2	Elaboration d'un protocole de test.....	63
3.3	Calibration de la salle d'écoute.....	65
3.3.1	Description.....	65
3.3.2	Ecoutes.....	67
3.3.3	Calibration et mesures.....	71

3.4	Elaboration du premier test.....	77
4	Analyse des résultats.....	81
5	Conclusion.....	82
6	Lexique.....	84
7	Annexe :.....	87
7.1	Courbes encodage spatial.....	87
7.2	Courbes décodage spatial.....	94
7.3	Caractéristiques techniques.....	98
7.4	Questionnaire fourni lors du premier test.....	107

# 1 INTRODUCTION

Ce mémoire a été effectué au sein de l'institut de recherche B<>COM, dans leur pôle Media Immersif, département audio 3D. Le fait qu'un institut de recherche se penche sur les enjeux de la captation du champ sonore et de sa transmission nous renvoie à l'histoire de l'enregistrement audio, vieille d'à peine plus d'un siècle. En effet les possibilités offertes par l'enregistrement et la diffusion de sons naturels sont infinies, de la musique jusqu'à l'art de la guerre (l'armée utilisait des enceintes diffusant des bruits de chars blindés pour faire croire à une supériorité numéraire factice pendant la deuxième guerre mondiale), aussi est-il important de définir un champ d'activité lors de l'utilisation d'une technologie appliqué à l'audio. Les nouvelles approches de captation et de restitution, notamment par synthèse de champ sonore, ouvrent assurément le champ des possibles. Elles appellent à la création de nouveaux outils, de nouvelles manières de penser la diffusion sonore dans un espace donné, afin de restituer non plus un point d'écoute, mais bien une zone d'écoute étendue dans lequel l'auditeur peut se mouvoir. Contrairement à l'approche « *channel-based* » classique, où l'environnement est capté par portions (correspondant aux directivités des microphones utilisés), l'enjeu ici est d'enregistrer celui-ci dans sa globalité. On peut imaginer pour exemple la captation d'une conférence enregistrée dans une salle, où les participants se trouvent autour d'une table, et les replacer virtuellement dans une salle similaire, où des auditeurs pourront écouter la conversion telle qu'elle est menée, et avoir le loisir de s'installer à la place d'un des participants pour profiter de son point de vue « sonore ».

Nous nous intéressons donc ici à une de ces technologies : HOA, pour Higher Order Ambisonics, et l'un des buts de ce mémoire est d'entrevoir les nouvelles possibilités qu'elle offre au traitement du son dans les différents média audiovisuels (cinéma, musique, radio, jeux vidéo, ...). La question du format se pose également. à l'heure de la profusion de format multicanal (Auro3D, Dolby Atmos, 7.1 ITU, 7.1 Surround, ...), comment conserver les différents enregistrements et pouvoir les réécouter sans avoir plusieurs salles d'écoute ? HOA offre une proposition sérieuse, en tant que format indépendant de l'encodage et du dispositif de restitution : cela signifie qu'un même fichier HOA peut être lu sur différents dispositifs sans avoir à modifier le fichier ou à opérer de « up/downmix ». Il serait donc une solution pour conserver des enregistrements de manière pérenne, capable d'être réécouté en 2D/3D sur des configurations existantes ou à venir. De plus le format MPEG-H

nouvellement standardisé permettrait de faire cohabiter toutes ses approches, et apporte de plus une solution de compression audionumérique de qualité pour HOA.

Ce mémoire s'intéresse donc à l'intégralité de la chaîne audio (enregistrement → traitements → diffusion) et comment HOA s'y insère, aux outils et plugins qui ont été développés autour, et bien évidemment au rendu sonore et spatial. Bien évidemment au vu de la contrainte du format, nous ne pouvons pas espérer donner un état de l'art de tout ce qui entoure HOA et ses applications. Nous proposons donc à un lectorat muni de quelques notions en techniques audio (et mathématiques), de faire un tour d'horizon pour chaque étape, et lui proposons des renvois bibliographiques s'il souhaite approfondir un aspect en particulier.

Ce mémoire s'axera autour de deux grands pôles : la partie théorique et la partie pratique, qui se réunissent autour de la caractérisation d'une installation type utilisée au cours de ce stage : un réseau sphérique de 32 capsules (Eigenmike) pour la captation microphonique, une suite de plugins et un studio d'écoute. L'idée sera de pouvoir s'intéresser à chaque étape d'un point de vue théorique, relié à un aspect perceptif présenté dans la partie pratique.

Concernant la partie théorique, un bref rappel de l'utilisation d'Ambisonics dans les différents secteurs de l'audiovisuel sera d'abord présenté, afin de contextualiser les apports de la technologie HOA aujourd'hui. Puis nous passerons à la partie encodage, qui comprendra l'encodage de sources préexistantes, et bien sûr la captation microphonique, comment celle-ci est rendue possible, quelles en sont les contraintes.

Nous nous intéresserons ensuite aux traitements spatiaux, fréquentiels et dynamiques du format HOA, en s'intéressant aux différents plugins existants (ils sont soit issus de solutions commerciales, ou de solutions open source (à l'instar de la suite Ambix). Nous donnons les liens de téléchargement en annexe, et invitons le lecteur à faire l'essai chez lui.

Enfin nous nous intéresserons au dernier maillon de la chaîne du son : la diffusion d'un fichier HOA sur haut-parleurs. Comme évoqué plus haut, un fichier HOA peut être diffusé sur n'importe quel type de configuration d'enceintes (même exotiques), aussi nous nous efforcerons de montrer les avantages et les défauts associés, et quelles peuvent être les effets associés à l'utilisation de différents décodeurs.



La partie pratique quant à elle, se penchera sur la perception des qualités audio et spatiales à chaque étape de la chaîne. Néanmoins au vu du temps imparti dans le cadre de ce mémoire, nous nous concentrerons sur la diffusion.

## 2 ASPECT THEORIQUE

### 2.1 Utilisation historique et actuelle de la technologie Ambisonics

La technologie Ambisonics, initiée et développée par Michael Gerzon au début des années 70 (puis développée par lui-même, Malham, Craven...) est une première approche de l'utilisation des harmoniques sphériques dans le son spatialisé. L'idée repose sur la recomposition du champ sonore au voisinage de la tête et de ses caractéristiques de propagation de proche en proche. Elle a mené au développement de dispositifs de captation (microphone Calrec Ambisonic, Soundfield ST450, SPS200 ; plus récemment le microphone Brahma, premier micro ambisonique en dessous de 1000€), d'encodage, traitement et manipulation : citons Harpex-B, la suite Blue Ripple Sound, la suite Open Source Ambix (ces 2 derniers permettent aussi des traitements aux ordres supérieurs, i.e. HOA).

La technologie Ambisonics encore présente aujourd'hui n'a pas connu de vraie évolution (mis à part de meilleurs composants et l'adaptation aux normes multicanales modernes (ITU 5.1, 7.1, ...)) depuis sa création (premier micro Calrec Ambisonic commercialisé en 1978), son échec commercial à l'époque peut être imputé (cf. article soundonsound) au raté de la quadriphonie (1969), chantre déchu de la diffusion multicanale à son époque (cf [0]). Il a en effet fallu attendre le dolby surround (1982) pour que le son multicanal soit de nouveau d'actualité commercialement.



*Richard Gilderoy avec le prototype de microphone Ambisonics (1974)*

Ces produits se sont néanmoins implantés sur le marché de l'audiovisuel, chez les principaux loueurs de matériel audio au cinéma (en France : DCA, Tapages), au sein des équipes TV pour le contenu 5.1 et utilisés par des ingénieurs du son pour la captation d'ambiances multicanales sur un tournage (citons Bruno Guéraçague (les Lascars), Gaël Nicolas (la Môme)) et dans la musique classique, en tant que micro principal ou micro d'ambiance (de nombreux exemples existent sur le label Chesky Records).

Quelques exemples jalonnent également les musiques actuelles, par exemple the trinity session des cowboy junkies, enregistré dans une église avec un seul micro : le calrec ambisonic. Plus précisément, la chanteuse avait tout de même un micro qui diffusait sa voix dans une enceinte, donnant le volume « référence » sur laquelle tous les musiciens se sont ajusté, jouant autour du micro. Ce disque est une référence pour son son intimiste jouant pleinement avec l'acoustique du lieu ([Sweet Jane - Cowboy junkies](#)). L'effet de proximité et l'acoustique du lieu ont en général primé pour l'utilisation de ce micro, mais, à ma connaissance, il n'a pas été utilisé spécifiquement en prise de son de proximité.

De par leur design (4 capsules cardioïdes montée sur les quatre faces d'un tétraèdre régulier) les micros Ambisonics peuvent être considérés comme lié à la famille des couples coïncidents (à l'instar du couple XY, MS, Stereosonic (dit Couple Blumlein)). En effet même si les capsules ne sont pas strictement coïncidentes, le déphasage créé est pris en compte dans le traitement et finalement n'apparaît pas dans les signaux en sortie (B-Format, voir Partie 2.2.1). Plus généralement, on observera que la spatialisation dans Ambisonics (et plus généralement dans HOA) relève du panning d'amplitude.



*Microphone tétraédrique Soundfield*

En effet l'approche ambisonics se concentre en un point, où l'on réalise la mesure du champ de pression et de ses dérivées spatiales, pour approcher la réalité du champ de pression dans le voisinage du point. Le panning d'amplitude nous permet

donc de contrôler les dérivées spatiales de façon cohérente pour placer une source dans l'espace.

Malgré son faible succès, la technologie Ambisonics offre des avantages indéniables qu'elle est la seule à posséder : elle est aisément manipulable, un seul microphone est nécessaire comparé aux autres dispositifs de captation multicanaux (exemple MMAD, micros WFS) qui peuvent vite s'avérer encombrants et difficiles voire impossibles à utiliser en nomade ou en situation de tournage. Il permet ainsi de capter un champ sonore en 3D, retranscrivant l'écoute naturelle.

De plus, depuis l'arrivée des formats Auro3D, 22.2 (NHK), Dolby atmos,... la composante verticale est d'autant plus à considérer dans la captation d'ambiances pour faciliter le travail au montage son et au mixage, d'où l'importance accrue de la considérer dès la captation.

Il faut bien saisir que le format Ambisonics repose sur un encodage intermédiaire entre l'enregistrement (ou la synthèse) et la diffusion, contrairement à un fichier stéréo pouvant être lu directement sur 2 enceintes (ou tout autre fichier comprenant un certain nombre de canaux et devant être lu sur le même nombre d'enceintes).

Ainsi ce format est totalement indépendant de la captation et de la diffusion. Pour Ambisonics, ce format intermédiaire est appelé B-Format et comporte 4 canaux (W, X, Y, Z).

De là, cette technologie tire son grand avantage (que dis-je immense !) : sa grande flexibilité, de par la capacité de traitement de ce format intermédiaire : il est tout à fait possible, à partir du même fichier encodé en Ambisonics, de simuler plusieurs microphones virtuels (par exemple pour une prise de son instrumentale, simuler au choix un couple AB, ou ORTF, ou de simuler un microphone très directif pour cibler un élément précis du champ sonore) qui peuvent se déplacer dans le temps ! On peut ainsi imaginer percher plusieurs acteurs en même temps avec un seul microphone (voir les travaux d'Angelo Farina à ce sujet) en étant capable de suivre à distance leur déplacement dans le temps.

L'espace de la scène sonore est également manipulable, via des opérations de rotation (la scène peut ainsi suivre une image en mouvement), de beamforming (« zoom acoustique »), filtrage spatial, élargissement de source (proposé dans la suite BRS).

Enfin, la restitution est indépendante du nombre de haut-parleurs, c'est-à-dire que le même fichier (au format Ambisonic) peut être à la fois diffusé sur une paire stéréo, un système 5.1 ou toute autre configuration de haut-parleurs, moyennant la configuration du dispositif de haut-parleur dans le décodeur, ce qui rend pérenne les ambiances encodés en B-Format pour les futurs formats de diffusions.

Bien sûr la qualité de rendu diffère entre différentes configurations, il est évident que pour un même fichier Ambisonics, la perception de l'espace ne sera pas la même entre le rendu stéréo et le rendu 5.1. Certaines configurations sont ainsi recommandées pour la diffusion, mais ne sont pas obligatoires.

Aussi les soucis de downmix (réduction du nombre de canaux pour un mixage (par exemple, le passage d'un mixage 5.1 à un mixage stéréo)), courants en post-production cinéma, n'ont plus de sens ici : en présence d'ambiances (ou de musique) en multicanal, il suffit d'exporter le fichier correspondant à chaque cas (2.0, 5.1, ...) à partir du format Ambisonics, ce qui n'est pas très fastidieux.

Enfin à l'instar des techniques utilisant la synthèse du champ sonore (comme l'holophonie, WFS,..), Ambisonics affranchit la notion de sweet spot en reconstruisant un espace homogène pour quiconque se trouve à l'intérieur de cette zone. Il offre ainsi la possibilité d'une spatialisation « robuste » pour différentes sources dans l'espace, application autant utile au cinéma qu'en sonorisation, spectacles vivants, ...

En dépit de tous ces avantages, l'Ambisonic n'a pas eu de franc succès commercial ni su susciter l'intérêt de la part de l'ensemble de la profession. En effet, plusieurs aspects peuvent s'avérer compliqués à insérer dans une chaîne de production classique :

- Concernant le tournage et la prise de son en extérieur : enregistrer à l'aide d'un microphone Ambisonics pose la question du monitoring. Pour l'instant, la plupart des enregistreurs professionnels (AATON Cantar X-2 (décodage prévue pour le X-3), Sounddevices 788, Zaxcom Deva,...) proposent une solution de décodage A/B-Format vers stéréo pour sortie casque, donc ne rendant pas compte de l'aspect 360° de l'enregistrement ni de la dimension verticale. Néanmoins, des solutions de monitoring binaural sont possibles, l'Ambisonics pouvant également être décodé sous ce format. Cela pourrait apporter la solution (c'est même ce qui a été fait pour les prises de son effectués dans le cadre de ce mémoire).

- Le micro Ambisonics consomme 4 pistes en B-Format, cela peut également s'avérer problématique pour des enregistreurs nomades légers (type Tascam HDP2), mais peu contraignants pour les enregistreurs nomades de 8 à 24 pistes présents sur le commerce.
- La taille du microphone et son poids peuvent également lui être reproché, comparé à un microphone simple. Néanmoins, cela reste tout à fait acceptable au regard des autres solutions concernant la captation multicanal. Il demeure aussi encombrant qu'un couple stéréophonique classique.
- Quant aux caractéristiques du micro (souffle, timbre, précision), je renvoie le lecteur au test de Jean Marc L'Hotel pour l'AFSI, amenant le regard critique d'un preneur de son professionnel sur des captations réalisés avec le ST450 de Soundfield. (cf [A])

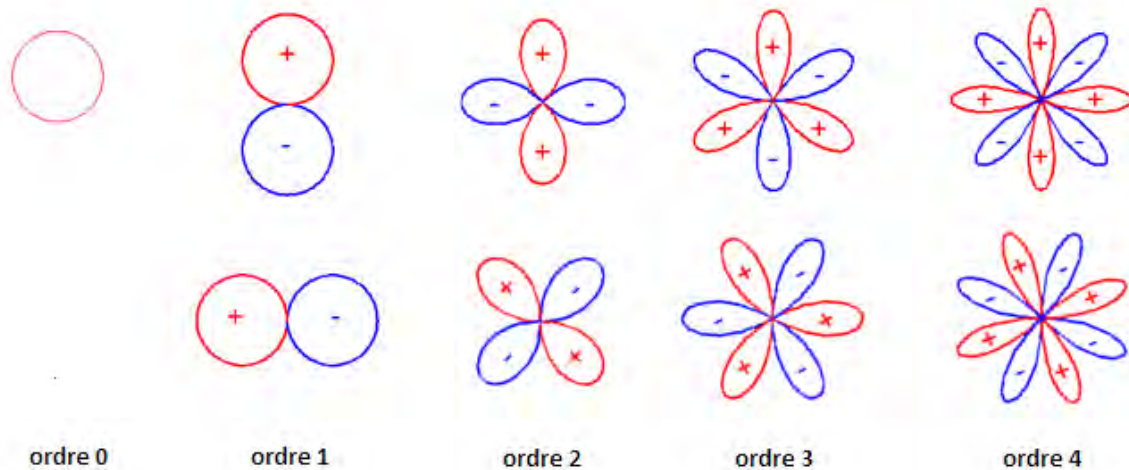


*Prise de son d'ambiance en Ambisonics (Wayne Pashley pour le film Australia avec un ST350)*

On peut s'étonner, au vu des avantages techniques et sonores de cette technologie, que peu d'échos en soit fait, même au sein des formations d'ingénieurs du son.

Cependant, depuis le début des années 2000, un regain d'intérêt de la part de la communauté scientifique est venu pour la technologie Ambisonics, dû à un nouveau formalisme : HOA (pour Higher Order Ambisonics). Pour les détails mathématiques, un bref rappel est présent en Annexe 1 (pour plus de détails se référer à la thèse de Jérôme D). L'idée consiste à formuler une décomposition de l'espace selon la série des harmoniques sphériques (en 3D) ou cylindriques (en 2D). Aussi, plus l'ordre de décomposition est élevé, plus la précision spatiale est précise et conforme au rendu naturel.

On retombe ainsi sur le formalisme Ambisonics de Gerzon en tronquant la série des harmoniques sphériques à l'ordre 1. HOA peut donc être vu grossièrement comme une extension d'Ambisonics avec une résolution spatiale accrue.



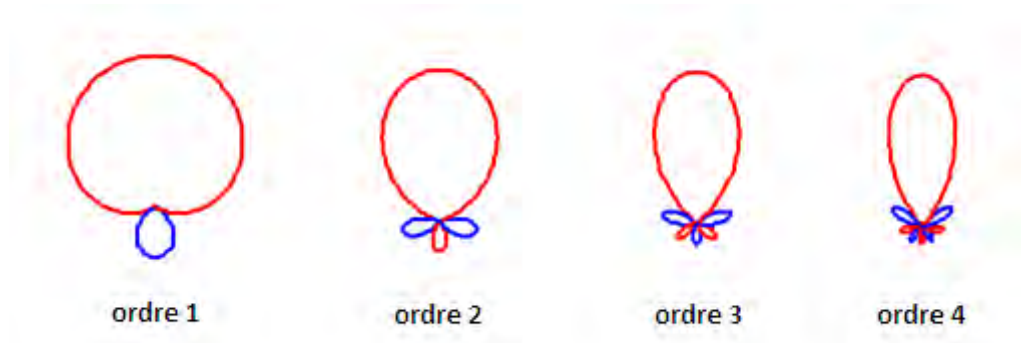
### *Harmoniques cylindriques jusqu'à l'ordre 4*

Les apports des ordres supérieurs comparativement à Ambisonics (ordre 1) sont nombreux, la précision de localisation est notamment accrue, s'approchant des performances de l'oreille humaine pour un ordre suffisamment élevé (i.e. le champ sonore décodé est reconstitué avec suffisamment de précision pour imiter un environnement naturel (cf [13]).

On arrive ainsi à augmenter la discrimination angulaire, deux sources proches seront plus facile à discriminer, permettant une meilleure « qualité spatiale ». Le champ sonore sera plus immersif, plus crédible, car se rapprochant de l'écoute naturelle. Un des exemples d'application de cette amélioration est la conférence immersive : un seul micro capte une réunion autour d'une table, le signal est encodé en HOA et rediffusé dans une autre salle en remplaçant chacun des protagonistes sur leur chaise virtuelle respective. On comprend bien ici l'enjeu de pouvoir facilement isoler chaque locuteur, ce qui demande la possibilité d'isoler une source dans l'espace avec un lobe de directivité très fin. Autrement dit, on veut pouvoir synthétiser un microphone virtuel très directif, au sens que l'entend A. Farina, à savoir que « n'importe quel signal présent au début, au milieu, ou à la fin d'une chaîne de traitement audio, doit toujours être interprété comme un signal venant d'un microphone, placé dans une position unique, dans une direction unique, avec un diagramme polaire spécifique dans une pièce particulière, aussi tout signal audio est un signal provenant d'un microphone virtuel ».

Cette possibilité est offerte par HOA, car le nombre de lobe de directivité (et donc leur précision) augmente avec l'ordre des harmoniques spatiales. Ce qui permet par sommation de créer des microphones virtuels avec des diagrammes polaires bien

plus directifs que des microphones canon classiques. Aussi un microphone HOA d'ordre  $x$ , pourra simuler un microphone d'ordre de directivité équivalent.

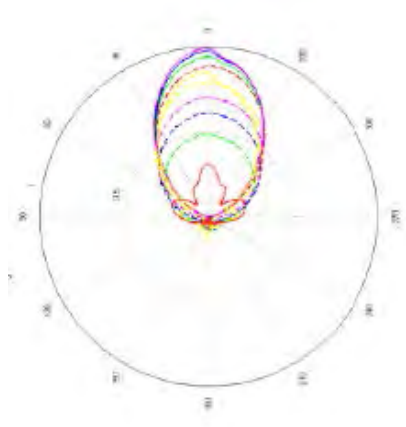
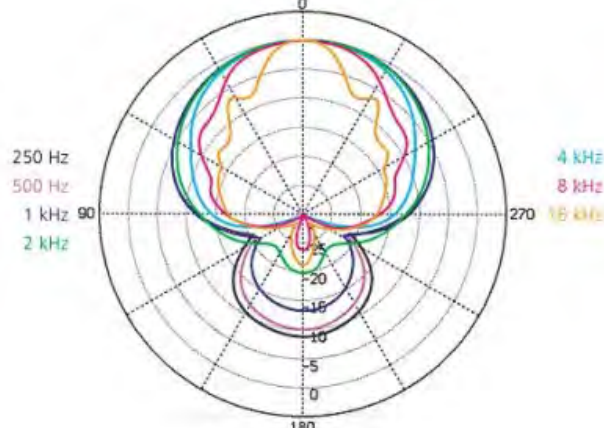
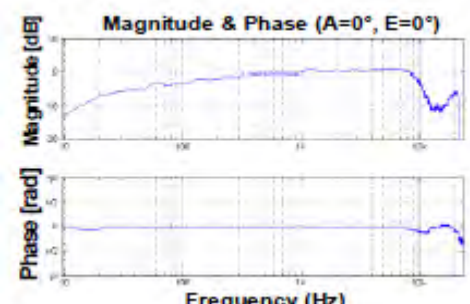
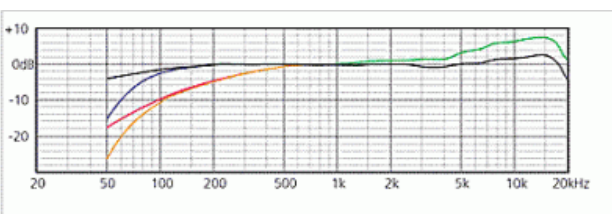




Atténuation (-3dB)	$\theta = 60^\circ$	$\theta = 43^\circ$	$\theta = 35,2^\circ$	$\theta = 30,5^\circ$
-----------------------	---------------------	---------------------	-----------------------	-----------------------

*Différents ordres de directivité pour un microphone cardioïde*

Par extension, Angelo Farina, dans ses travaux avec le laboratoire de la Rai, présente la possibilité de simuler un microphone virtuel cardioïde d'ordre 6 (soit une atténuation de -3dB pour  $\theta = 25^\circ$ , contre  $\theta = 60^\circ$  pour un microphone cardioïde d'ordre 1), avec un microphone Eigenmike commercialisé par MHAcoustics (comportant 32 capsules réparties régulièrement à la surface d'une sphère), avec la possibilité de déplacer sa directivité sans bouger physiquement le microphone. Cependant il utilise une méthode différente de HOA, consistant en une série de matrices de filtre FIR (i.e à réponse impulsionnelle finie), dépendante du dispositif de captation et de restitution. On notera qu'il arrive même avec les 32 capsules présentes dans l'Eigenmike, mais agencées différemment (sur un plan), à recréer un microphone virtuel cardioïde d'ordre 16. Malheureusement, la question du timbre n'est pas posée dans l'article. Il serait en effet très intéressant de pouvoir comparer les différents signaux récupérés et le timbre associé, comparativement à des microphones canons dits « classique ». Quoiqu'il en soit, les nombreuses possibilités de traitement spatial qui en découlent constituent un terrain d'investigation fertile.



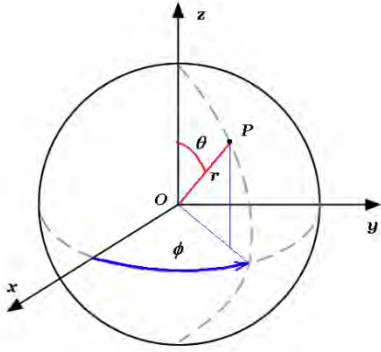
	
	
	
<p>Eigenmike (MHAcoustics)</p>	<p>CMT 5-U (Schoeps)</p>

*Comparatif entre un microphone virtuel obtenu à partir de l'Eigenmike et un microphone canon standard*

## 2.2 Le formalisme HOA (2D & 3D)

### 2.2.1 Harmonique sphérique

On introduit dans une courte notice les notions essentielles pour la compréhension des notations et des expressions développées dans les parties suivantes.



Pour rappel :

$$x = r \cos \theta \sin \delta$$

$$y = r \sin \theta \sin \delta$$

$$z = r \cos \delta$$

La représentation d'une scène sonore 3D avec la technologie HOA se base sur l'expression du champ de pression acoustique sous la forme d'une série de Fourier-Bessel, c'est-à-dire un développement en harmoniques sphériques. En effet, l'écriture de l'équation des ondes en champ libre en coordonnées sphériques permet d'exprimer ce champ sous forme d'une série de Fourier Bessel, i.e :

$$p(kr, \theta, \delta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=0}^m \sum_{\sigma=\pm 1} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\theta, \delta) \quad (1A)$$

où,  $J_m(kr)$  sont les fonctions de Bessel Sphériques à dépendance radiale et fréquentielle (cf. Figure 1),  $Y_{mn}^{\sigma}$  représentent les harmoniques sphériques (voir Figure 2), et  $B_{mn}^{\sigma}$  les fameux signaux HOA associés à chaque harmonique sphérique et  $k$  le nombre d'onde ( $k = \frac{2\pi f}{c}$ ).

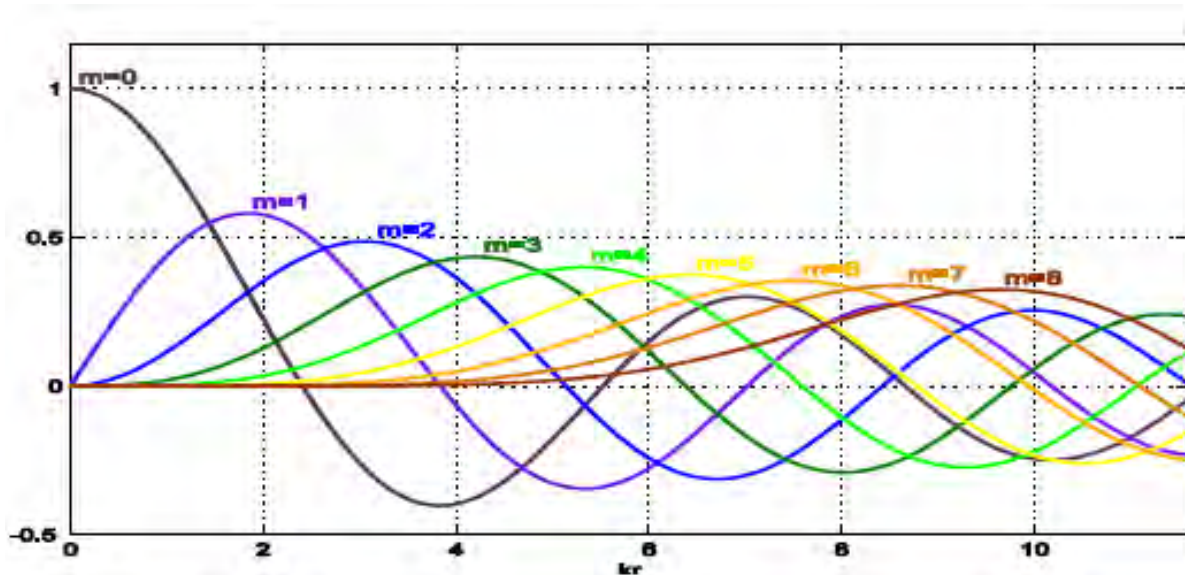


Figure 1 : fonctions de Bessel Sphériques  $J_m(kr)$  pour  $m \in \llbracket 0; 8 \rrbracket$

On remarque que plus l'ordre des fonctions de Bessel sphériques augmente, plus la valeur de  $kr$  à laquelle les fonctions s'annulent augmente, on peut donc en déduire que plus l'ordre est élevé, plus le champ sera reconstitué sur une grande distance et sur un spectre fréquentiel élargi. On aura donc une meilleure résolution spatiale et fréquentielle.

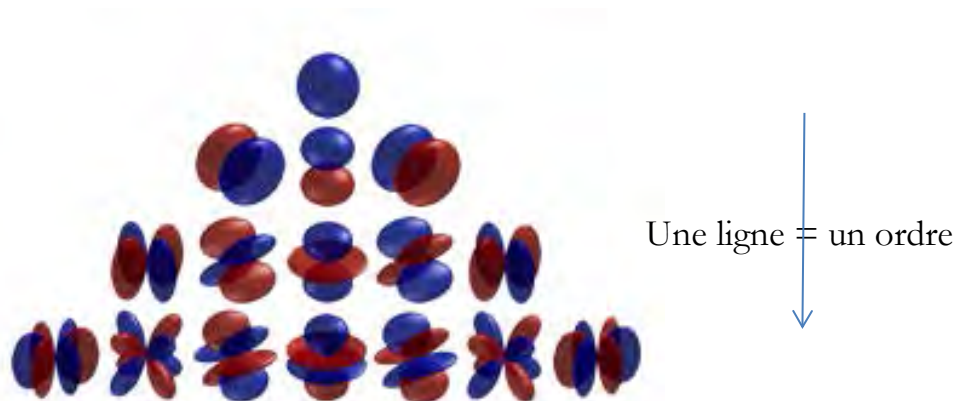


Figure 2 : représentation des harmoniques sphériques (de l'ordre 0 à l'ordre 3)

On reconnaît dans les 2 premiers ordres (0 et 1) des directivités de microphone connues (omnidirectionnel et bidirectionnel), elles définissent les 4 composantes du B-Format pour Ambisonics (nommées W X, Y et Z). Aussi s'ils existaient des microphones avec des directivités identiques aux différentes harmoniques sphériques pour les ordres supérieurs, on pourrait avoir accès directement aux

signaux HOA. Malheureusement, il n'existe pas de moyens physiques de construire de capteurs avec de telles figures de directivité.

La représentation d'une harmonique sphérique s'obtient avec la formule suivante :

$$Y_{mn}^{\sigma}(\theta, \delta) = N_{mn} P_{mn}(\sin(\delta)) * \begin{cases} \cos(n\theta) & \text{si } \sigma = 1 \\ \sin(n\theta) & \text{si } \sigma = -1 \end{cases} \quad (2)$$

Où  $P_{mn}$  est le polynôme de Legendre associée à  $m$  et  $n$ , et  $N_{mn}$  le paramètre de semi-normalisation de Schmidt.

Les harmoniques sphériques constituent une base orthogonale de l'espace sur la sphère unité, et orthonormale en fonction du paramètre de normalisation (N3D).

Tout ceci est également valable en 2D, où l'équation (1) devient :

$$p(kr, \theta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{\sigma=\pm 1} B_m^{\sigma} Y_m^{\sigma}(\theta) \quad (1B)$$

On remarque que le nombre d'harmoniques en 2D  $Y_m^{\sigma}$  (dites harmoniques cylindriques) par ordre diminue.

	Ordre	Nombre de signaux HOA	Total
2D	0	1	1
	1	2	3
	2	2	5
	m	2	2m + 1
3D	0	1	1
	1	3	4
	2	5	9
	m	2m + 1	(M + 1) <sup>2</sup>

Tableau 1 : Nombre de signaux HOA en fonction de l'ordre et de la dimension.

Ainsi il est possible de projeter une représentation 3D sur la base des harmoniques cylindriques en ne gardant que les signaux HOA pour lesquels  $m = n$  et en appliquant un facteur de conversion égal à :

$$\begin{cases} Y_m^\sigma(\theta) = Y_{00}^1(\theta, \delta = 0) = 1; & m = 0 \\ Y_m^\sigma(\theta) = \sqrt{\frac{2^{2m} m!^2}{(2m+1)!}} Y_{mm}^\sigma(\theta, \delta = 0); & m \geq 1 \end{cases} \quad \sigma = \pm 1$$

## 2.2.2 Convention de normalisation des harmoniques sphériques

Différentes conventions existent pour la catégorisation et la normalisation des harmoniques sphériques, il convient d'y faire très attention sous peine d'obtenir des aberrations au moment du décodage.

### 2.2.2.1 Catégorisation

Actuellement 3 grandes conventions existent pour la catégorisation, soit Furse-Malham, ACN et SID. Pour information, l'annexe 1 a été rédigé avec la convention SID présenté par J.Daniel dans sa thèse [2].

Convention	Ecriture Harmonique Sphérique	Ordre de classement HOA	Logique
SID	$Y_{mn}^\sigma(\theta, \delta)$	[W X Y Z U V S T R ...]	Les composantes horizontales ( $m = n$ ) en premier et la composante verticale ( $n = 0$ ) à la fin
Furse-Malham	$Y_l^k(\theta, \delta)$	[W Z X Y R S T U V ...]	La composante verticale ( $k = 0$ ) en premier et les composantes horizontales à la fin ( $l = \pm k$ )
ACN	$Y_l^k(\theta, \delta)$	[W Y Z X V T R S U]	On classe selon $l^2 + l + k$ croissant

La correspondance entre les indices de l'écriture SID et ceux utilisés par ACN et Furse-Malham est la suivante :

$$\begin{cases} m = l \\ n = |k| \\ \sigma = \text{sgn}(k) \end{cases} \text{ ou } \begin{cases} l = m \\ k = \sigma * n \end{cases}$$

Exemple :

Ordre 0 :  $W = \begin{cases} Y_{00}^1 (SID) \\ Y_0^0 (Furse, ACN) \end{cases}$



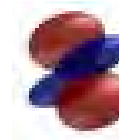
Ordre 1 :

- $X = \begin{cases} Y_{11}^1 (SID) \\ Y_1^1 (Furse, ACN) \end{cases}$
- $Y = \begin{cases} Y_{11}^{-1} (SID) \\ Y_1^{-1} (Furse, ACN) \end{cases}$
- $Z = \begin{cases} Y_{10}^{-1} (SID) \\ Y_1^0 (Furse, ACN) \end{cases}$



Ordre 2 :

- $R = \begin{cases} Y_{20}^{-1} (SID) \\ Y_2^0 (Furse, ACN) \end{cases}$
- $S = \begin{cases} Y_{21}^{-1} (SID) \\ Y_2^{-1} (Furse, ACN) \end{cases}$
- .... etc.



### 2.2.2.2 Normalisation

La normalisation concerne les rapports d'amplitude entre les composantes HOA. Elle diffère en fonction des auteurs et peut induire des problèmes de compatibilité si elle n'est pas respectée.

Historiquement Gerzon en a fourni une pour les ordres 0 et 1 [3], puis Malham et Furse en ont introduit une nouvelle valable aux ordres supérieurs, et enfin Jérôme Daniel [2] amène deux nouvelles conventions (SN3D (pour semi-normalisées 3D) & N3D (SN3D normalisées en énergie)) valables également pour n'importe quel ordre.

Nom	Ordre	Convention
Gerzon	0,1	$\begin{cases} W = 1 \\ X = \sqrt{2} \cos\theta \cos\delta \\ Y = \sqrt{2} \sin\theta \cos\delta \\ Z = \sqrt{2} \sin\delta \end{cases}$
Furse Malham	Tous	$\begin{cases} W = 1/\sqrt{2} \\ \text{Les autres composantes sont normalisées t.q} \\ \text{leur valeur maximale soit égale à 1} \end{cases}$
Daniel SN3D	Tous	$N_{mn} = \sqrt{\varepsilon_n \frac{(m-n)!}{(m+n)!}} \quad \text{où } \varepsilon_n = \begin{cases} 1 \text{ si } n = 0 \\ 2 \text{ si } n \geq 1 \end{cases}$
Daniel N3D	Tous	$N_{mn}^{(N3D)} = \sqrt{2m+1} N_{mn}$

## 2.3 Encodage spatial

Le formalisme introduit précédemment nous fournit une description de l'espace comme décomposition sur la base orthonormale des harmoniques sphériques. Aussi une conséquence logique de ce formalisme est que l'encodage d'une source synthétique (nous nous intéresserons au cas élémentaire de l'onde plane) se résume à une projection sur cette base, tronquée à l'ordre  $M$ . Le cas d'une source microphonique est plus délicat et nous nous y attarderons un peu plus longuement.

### 2.3.1 Encodage spatial d'une source élémentaire l'onde plane

On s'intéresse au cas élémentaire d'une onde plane d'incidence  $\vec{u}$  correspondant aux angles d'incidence  $(\theta_0, \delta_0)$ , d'amplitude  $S$ . Les signaux HOA  $B_{mn}^\sigma(\theta_0, \delta_0)$  s'obtiennent de la manière suivante :

$$B_{mn}^\sigma = S \cdot Y_{mn}^\sigma(\theta_0, \delta_0)$$

Aussi l'encodage d'une source complexe (somme de plusieurs ondes planes) sera relativement aisé. On remarque que cette représentation permet, de manière évidente, des modifications de champ comme la rotation. Aussi, vu que chaque signal HOA est simplement relié par un terme d'amplitude et 2 termes angulaires, le rapprochement avec un panpot d'amplitude, bien connu en mixage stéréophonique, est naturel et mène tout naturellement à un « panpot » 3D.

Aussi il est naturel de retrouver cette idée de panpot 3D dans les plugins d'encodage HOA, où la position de la source audio peut être déplacée au cours du temps (et la trajectoire résultante peut tout naturellement être automatisée). L'extension au mouvement de sources multicanal est simple, il suffit de placer la position respective de chaque canal au préalable (comme pour un 5.1 selon la norme ITU par exemple), puis il nous sera permis d'appliquer respectivement la même trajectoire à chaque canal (rien ne nous empêche non plus de déplacer chaque canal indépendamment les uns des autres).

Néanmoins l'encodage HOA implique un ordre fini  $M$  et cette troncature à l'ordre  $M$  amène avec elle des soucis liés à l'échantillonnage spatial. En effet, la série de Fourier-Bessel à un ordre fini ne donnera qu'une approximation du champ sonore idéal (cf. Figure 3).



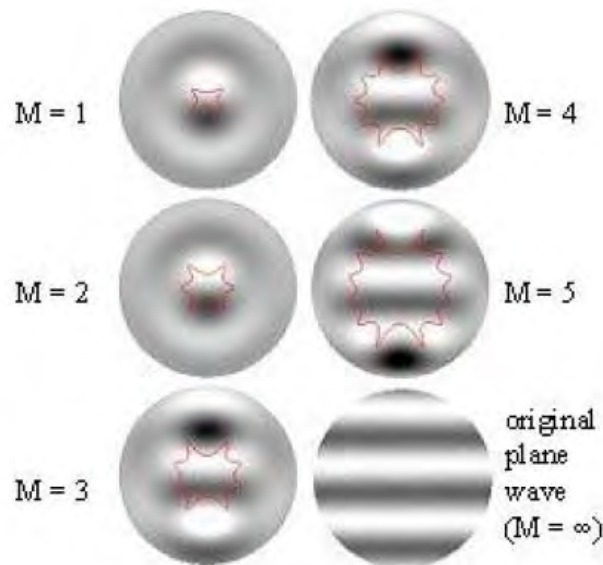


Figure 3 : Champ sonore d'une onde plane monochromatique et ses approximations à différents ordres par la série de Fourier-Bessel (source [1])

Le but ici n'est pas de retracer toute la littérature faite sur le sujet. En effet, l'encodage de source de synthèse est un thème toujours d'actualité dans la recherche sur HOA, et ce n'est pas le propos de ce mémoire. On pourra néanmoins se référer au travail de Pierre Lecomte pour la prise en compte de l'effet de salle [4], les défauts liés à l'échantillonnage spatial abordés par Spoors et Ahrens [5] ...

### 2.3.2 Encodage microphonique

L'encodage microphonique est un enjeu majeur du développement de la technologie HOA. En effet, les travaux de Gerzon avaient mené, suite à la réalisation de différents prototypes, à une solution commercialisable avec le micro Calrec Ambisonic, composé d'une capsule de directivité omnidirectionnelle, et de deux capsules de directivités bidirectionnelles (trois pour un enregistrement en 3D), centré au même point.

Néanmoins, comme vu précédemment, la forme des harmoniques sphériques permet d'accéder directement aux signaux HOA aux ordres 0 et 1 en utilisant des microphones coïncidents avec des figures de directivité similaires. Ceci n'est plus possible à partir de l'ordre 2, car de tels microphones n'existent pas.

Une évolution du microphone Ambisonics (ordre 1) a été l'utilisation de 4 capsules cardioïdes placées sur les 4 faces d'un tétraèdre.



*Micro Soundfield SPS200 : 4 capsules directives*

Ces 4 capsules sont décalées afin de les placer virtuellement sur un même pont central, et permettent d'obtenir l'information de pression, ainsi que les 3 composantes de son gradient (donc les dérivées de la pression selon les 3 axes (Ox, Oy, Oz)). On retombe sur la décomposition en harmoniques sphériques aux ordres 0 et 1, et on accède donc aux 4 premiers signaux HOA (W, X, Y, Z). On utilise donc une approche différentielle pour accéder aux signaux HOA.

Comment réussir alors à exprimer les signaux HOA relatifs à chaque harmonique sphérique pour les ordres supérieurs ?

Une première approche est l'utilisation de microphones différentiels : un microphone différentiel d'ordre  $m$  se comporte comme un filtre passe-haut avec une pente de  $-m \times 6$  dB/octave ; il mesure ainsi les différences de pressions entre différents capteurs très rapprochés de l'espace, c'est-à-dire que la distance entre capteurs est négligeable devant la longueur d'onde (pour rappel : à 100 Hz, la longueur d'onde vaut 3,4m ; 34cm à 1000 Hz, et 3,4 cm à 10000 Hz), il enregistre donc une approximation du différentiel de pression à l'aide des dérivées inférieures à l'ordre  $m$ .

$$f^{(m)}(x) \cong \frac{f^{(m-1)}(x + \varepsilon) - f^{(m-1)}(x)}{\varepsilon}$$

Néanmoins, cette différence  $\varepsilon$  correspondant à l'espace entre deux capteurs ne doit pas être trop grande vis-à-vis des longueurs d'onde qu'on cherche à estimer, ni trop

petite pour que la différence entre les 2 valeurs de capteurs soit assez significative pour ne pas être noyé dans le bruit de mesure (cf. figure 4).

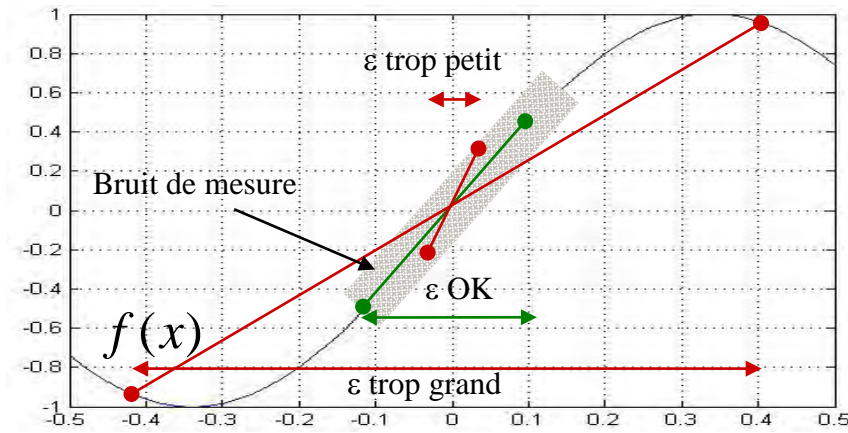


Figure 4 : Schéma explicatif de calcul de la dérivée

La littérature (cf. [6]) nous apprend que la pression acoustique peut être représentée comme une série de Taylor, c'est-à-dire comme une somme des différentes dérivées par rapport au point initial  $(x_0, y_0, z_0)$ .

$$p(x, y, z, \omega) = \sum_{i=0}^m \sum_{j=0}^m \sum_{k=0}^m p_{j,k,l}(\omega) (x - x_0)^i (y - y_0)^j (z - z_0)^k. \quad (3)$$

Avec  $p_{j,k,l}(\omega)$  coefficient correspondant aux différentes dérivées spatiales par rapport aux 3 coordonnées au point de mesure  $(x_0, y_0, z_0)$

Cette représentation alternative du champ sonore correspond à la même idée que la décomposition du champ en harmonique sphérique, où l'on déduit le champ sonore à partir de la mesure en un point. De fait, des parallèles existent entre les coefficients HOA de l'équation (1) et ceux vu précédemment (cf. [7]).

Malheureusement, les microphones différentiels nécessitent une amplification en basse fréquence trop importante, qui les rend inutilisables passé l'ordre 2, car le bruit de fond ramené par l'amplification est bien trop conséquent. De plus, cette méthode de reconstruction du champ sonore impliquerait un réseau de capteurs extrêmement proche du champ sonore, ce qui est très difficile à mettre en œuvre dans la pratique. Cette approche est donc théoriquement possible, mais seulement pratiquement réalisable à l'ordre 1 (comme vu précédemment).

Une autre approche a donc été proposée : un micro HOA sphérique.

En reprenant la formule (1A), on remarque que les signaux HOA peuvent être estimés pour une valeur de  $r$  fixe. Aussi, si on fixe  $r = R$  (constante), la surface en question est donc une sphère centrée en un point. Si on mesure la pression à la surface de cette sphère continue, on peut estimer les signaux HOA par la formule suivante :

$$B_{mn}^{\sigma} = \frac{1}{i^m J_m(kR)} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} p(kR, \theta, \delta) Y_{mn}^{\sigma}(\theta, \delta) \cos\delta d\vartheta d\delta \quad (4)$$

On intègre donc sur la surface de la sphère la pression multipliée par les harmoniques sphériques, et on multiplie par un terme d'égalisation qu'on notera  $EQ_m(kR)$  égal à :

$$EQ_m(kR) = \frac{1}{i^m J_m(kR)}$$

Ce paramètre est relié seulement à l'ordre des signaux HOA à la fréquence (de par le nombre d'onde  $k = \frac{2\pi f}{c}$ ) et au rayon de la sphère mesurée. L'analogie à un filtre fréquentiel est donc pertinente.

Néanmoins ce paramètre implique des valeurs problématiques. En effet, les fonctions de Bessel s'annulent régulièrement (cf. Figure 1) et oscillent autour de zéro, il diverge donc pour un certain nombre de valeurs dépendantes de la fréquence et de l'ordre des signaux HOA. Les signaux HOA aux fréquences correspondantes ne pourront donc pas être évalués.

De plus, lorsque les valeurs des fonctions de Bessel sont très petites, elles imposent une forte amplification, difficile voire impossible à réaliser.

En résumé, le micro HOA sphérique ne peut pas fonctionner en l'état dû à l'instabilité du paramètre d'EQ.

Une solution naïve pourrait être de réaliser les mesures sur différentes valeurs de rayons pour récupérer de l'information spatiale là où la sphère principale ne le permet pas. Ainsi on obtiendrait un microphone similaire à l'Atomium où des réseaux de microphone sphérique seraient reliés entre eux. Cette solution, bien que correcte, paraît difficilement envisageable dans un cadre autre que la recherche, de par son encombrement et de sa structure peu pratique pour l'enregistrement.



*Exemple de réseau de microphones avec plusieurs valeurs de rayons (40 microphones), 1,8m de diamètre*

Une alternative est de rendre la mesure de la pression à la surface de la sphère directive. Pour cela, deux choix ont été proposés, soit l'utilisation de capsules directives, soit la considération de la mesure de pression à la surface d'une sphère acoustiquement rigide, c'est-à-dire une sphère dont l'impédance acoustique est supérieure à celle du milieu qui l'entoure (pour rappel  $Z_{\text{air}} = 410 \text{ Pa.s/m}$ ). Des solutions alternatives ont également été proposées. Citons par exemple un cercle de microphones omnidirectionnels (cf. [8], [9]), ou les brevets relatifs à l'Eigenmike déposés par Gary Elko et Jens Meyer ([10]) gérants de MHAcoustics, la structure qui fabrique et commercialise ce microphone (appelé microphone par abus de langage, au lieu de réseau sphérique de microphones), utilisée dans le cadre de ce mémoire).

Nous allons aborder les deux solutions précédemment citées et les paramètres d'égalisation fréquentielle qu'elles impliquent, sachant que le microphone utilisé pendant le mémoire (Eigenmike de MHAcoustics) a été construit en suivant ces formulations. Ainsi, la caractérisation objective de l'encodage microphonique visée dans ce mémoire passe par la compréhension du dimensionnement des filtres associés à l'encodage microphonique.

### **Utilisation de microphones différentiels à la surface de la sphère**

Comme vu précédemment, les microphones différentiels dépassant l'ordre 2 étant très compliqué à réaliser, on restera à l'ordre 1.

On rappelle la formule d'un microphone différentiel d'ordre 1 pour R fixé :

$$s(kR, \theta, \delta) = \alpha p(kR, \theta, \delta) + (1 - \alpha) \frac{\partial p(kR, \theta, \delta)}{\partial kR} \quad 0 \leq \alpha \leq 1 \quad (4A)$$

On injecte l'expression de la pression sous la forme d'une série de Fourier-Bessel tel que le présente l'équation (1A) et on obtient :

$$s(kR, \theta, \delta) = \sum_{m=0}^{\infty} i^m (\alpha j_m(kR) + (1 - \alpha) j'_m(kR)) \sum_{n=0}^m \sum_{\sigma=\pm 1} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\theta, \delta) \quad (4B)$$

Avec  $j'_m$  la dérivée de la fonction de Bessel  $j_m$  par rapport à  $kR$ .

En notant 
$$W_m(kR) = i^m (\alpha j_m(kR) + (1 - \alpha) j'_m(kR))$$

Et en utilisant la propriété d'orthonormalité des harmoniques, on retombe sur l'expression des signaux HOA.

$$B_{mn}^{\sigma} = \frac{1}{W_m(kR)} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} s(kR, \theta, \delta) Y_{mn}^{\sigma}(\theta, \delta) \cos \delta d\theta d\delta \quad (5)$$

On se retrouve donc avec un nouveau paramètre d'égalisation

$$EQ_{m(micro)}(kR) = \frac{1}{W_m(kR)}$$

Cette fois l'analyse de la fonction (illustrée en Figure 5 avec  $\alpha = 1/2$  (directivité cardioïde), nous montre que la fonction  $W_m(kR)$  ne s'annule que pour  $kR = 0$  et  $m \geq 1$ , autrement dit pour une fréquence nulle ou une sphère de rayon nulle, donc le paramètre d'égalisation est défini et continu sur tout le spectre audio. Cette solution semble donc convenir.

### Utilisation d'une sphère rigide

L'utilisation d'une sphère rigide (déjà utilisée pour certains prototypes de microphones HOA d'ordre 1, notamment 6 capsules omnidirectionnelles placées sur une sphère en nylon cf. [11]) va provoquer des effets de diffraction et amener une sommation du champ de pression en champ libre et du champ de pression diffracté par la sphère. Cela aura pour effet de rendre la mesure du champ « directive » en fonction de l'emplacement à la surface de la sphère.

On passera sur l'expression mathématique de la pression (cf. [7]), pour aboutir directement à la formulation des signaux HOA, de la même forme que (5), où

$$W_m(kR) = \frac{i^{m-1}}{(kR)^2 h_m^{-'}(kR)} \quad (6)$$

Avec  $h_m^{-'}$  la dérivée de la fonction de Hankel par rapport à  $kr$  et  $R$  diamètre de la sphère.

Cette fonction ne s'annule également qu'en  $kR = 0$  et  $m \geq 1$  ; autrement dit elle s'annule uniquement pour  $k=0$  ; elle est donc définie et continue sur tout le spectre audio. On obtient donc un paramètre d'égalisation  $EQ_{m(sphère)}(kR) = \frac{1}{W_m(kR)}$  défini quelque soit l'ordre pour  $k \neq 0$ .

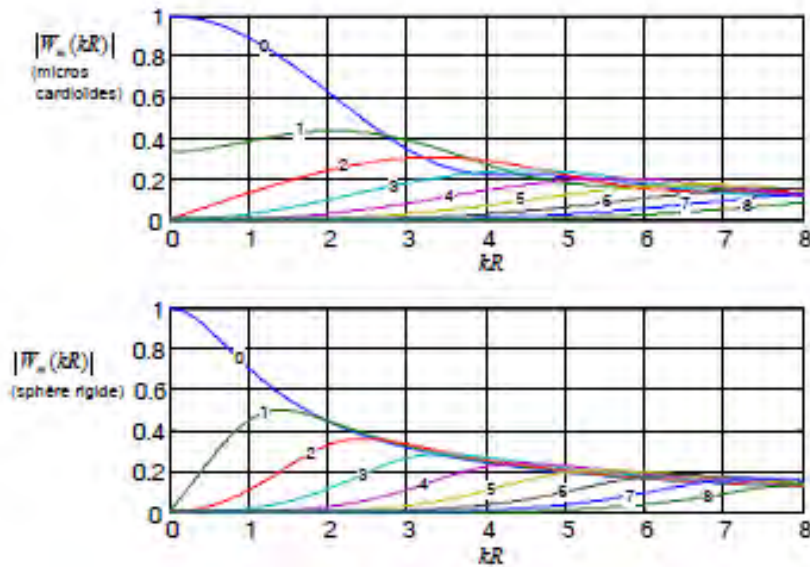


Figure 5 : Courbes d'amplitude du paramètre  $W_m(kR)$  à différents ordres

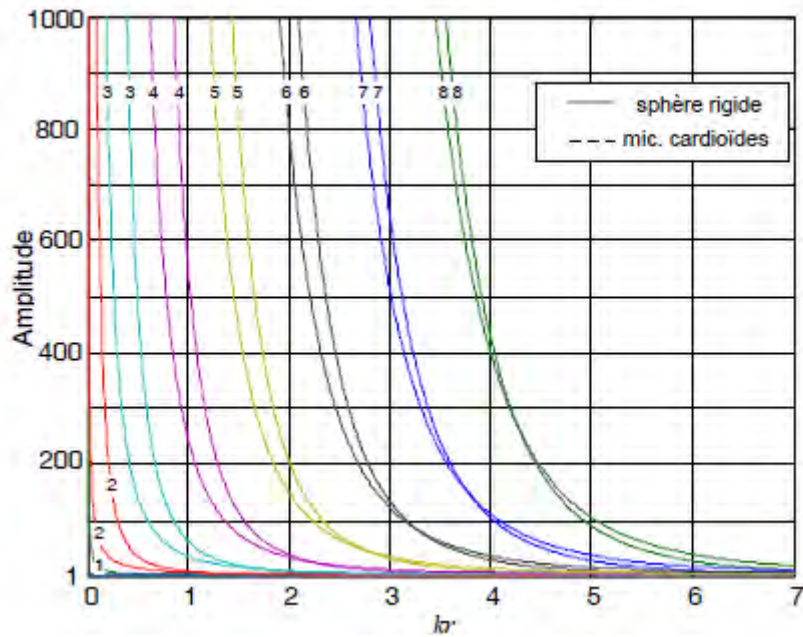


Figure 6 : Filtres  $EQ_m$  pour différents ordres

On remarque que les filtres résultants divergent vers l'infini quand  $kr$  tend vers 0. Aussi, implémenter de tels filtres imposera une augmentation énorme du bruit de fond en basses fréquences, et on risque d'avoir un signal fortement dégradé. Il faudra donc ajouter des contraintes pour que ce filtre soit applicable. De même il sera impossible dans la pratique d'avoir une capture du champ de pression sur la surface totale de la sphère.

On discerne là deux points essentiels pour que ce modèle puisse être applicable à la réalisation d'un microphone : l'échantillonnage de la sphère avec des capteurs de pression, et la minimisation des filtres à l'aide d'une méthode choisie.

### 2.3.3 Echantillonnage spatial

Tout d'abord intéressons-nous dans les grandes lignes aux problèmes qu'induit l'échantillonnage :

L'échantillonnage avec un nombre  $N$  de capteurs influe en premier lieu sur la possibilité de capter les signaux HOA. En effet comme vu précédemment, en 3D pour un ordre  $M$  donné, il y a  $(M+1)^2$  signaux HOA à capter, il faut donc un nombre de capteurs égal ou supérieur à ce nombre, d'où la condition :

$$N \geq (M + 1)^2 \quad \forall M \in \mathbb{N}$$



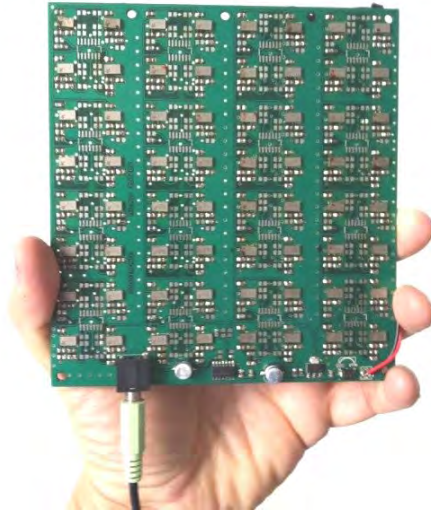
En général le nombre de capteurs est supérieur au nombre de signaux HOA pour des raisons de répartition régulière des capteurs à la surface de la sphère (problème mathématique millénaire, faisant intervenir les solides platoniciens, les solides d'Archimède, les polyèdres semi-réguliers, etc.) afin d'avoir une distance constante entre chaque capteur. Par exemple, dans le cas pratique qui nous intéresse (l'Eigenmike), une répartition suivant le pentakidodécaèdre (composé de 60 triangles isocèles et comportant 32 arêtes) a été choisie. Pour parfaire ce court paragraphe, il faudrait également s'attarder sur la question de la surdétermination du système d'équation obtenu lorsque le nombre de capteurs dépasse le nombre de signaux HOA (on obtient donc plus d'équations que d'inconnues, en général un tel système n'admet de solutions que dans des cas très particuliers), il faut donc faire appel à un appareillage mathématique sur lequel nous ne nous attarderons pas (mais pour plus d'informations à ce sujet cf. [7]).

En passant dans le domaine discret, la question de l'échantillonnage avec un nombre  $N$  de capteurs doit également respecter le critère de Shannon, aussi un phénomène d'aliasing spatial apparaîtra plus la fréquence sera élevée, cette fréquence étant directement reliée à la distance entre capteurs  $d$  par la formule suivante :

$$f_{al} = \frac{c}{2\Delta d}$$

Comme on se place sur une sphère à rayon  $R$  constant, la condition de distance minimum revient à déterminer un écart angulaire minimal  $\gamma$  entre chaque capteur. Comme vu précédemment, cet écart est directement fixé par la répartition régulière choisie pour le placement des capteurs. On comprend bien que pour augmenter cette fréquence d'aliasing, deux choix sont possibles : soit augmenter le nombre de capteurs pour diminuer cet écart angulaire, soit diminuer le rayon de la sphère, ce qui impose pour des capsules standards un compromis à trouver en fonction l'ordre de signaux HOA voulu.

Néanmoins avec l'utilisation récente des MEMS (microsystème électromécanique) comme microphones (notamment dans les téléphones portables, mais aussi dans des applications de laboratoire), la création d'un réseau de microphone miniature appliqué à HOA pourrait être une piste à creuser (cf. [17]), à condition d'assurer une réponse en fréquence suffisante. De plus, vue leur taille réduite, on peut imaginer constituer des microphones différentiels à la surface de la sphère avec plusieurs MEMS rassemblés en un seul capteur.



Réseau de 64 microphones MEMS (source dodotronic.com)

### 2.3.4 Minimisation et définition du paramètre d'amplitude

On a vu précédemment que les filtres obtenus à l'aide des 2 cas de figure (équations (4) et (5)) sont théoriquement définis, mais impossibles à mettre en pratique car ils impliquent une amplification trop importante en basse fréquence. Il faut donc définir de nouveaux filtres avec un gabarit le plus proche possible de ceux définis, mais comportant un seuil maximal d'amplification au-delà duquel l'amplification s'arrête. Plusieurs outils mathématiques existent pour résoudre ce problème de minimisation, néanmoins nous choisirons celui utilisé par l'encodeur dans le cadre de ce mémoire. Cet outil se nomme méthode de Tikhonov (défini dans [18]) et consiste en une extension de la méthode des moindres carrés, elle se définit par une minimisation de la norme quadratique avec l'ajout d'une limite supérieure  $\lambda$  appelée paramètre de régularisation.

Passée cette étape de régularisation, les filtres résultants sont égaux à :

$$EQ_{m(Tikhonov)}(kR) = \frac{\overline{W_m(kR)}}{|W_m(kR)|^2 + \lambda^2} \quad (7)$$

Où  $\overline{W_m(kR)}$  est le conjugué complexe de  $W_m(kR)$ .

C'est ce fameux paramètre  $\lambda$  qui nous intéresse ici, car pour notre application il est directement relié à une amplification maximale  $A_{\max}$  qui correspond donc au seuil maximal que l'on peut s'accorder pour la génération des filtres. Ils sont reliés par la relation suivante :

$$\lambda = \frac{1 - \sqrt{1 - 1/A_{max}^2}}{1 + \sqrt{1 - 1/A_{max}^2}} \quad A_{max} \geq 1$$

On voit donc que le paramétrage du filtre est directement lié à l'amplification maximum admissible, sachant que ce paramètre d'amplification maximum admissible peut se faire plus ou moins finement, en lui appliquant ou non une dépendance à l'ordre choisi (on se souviendra des courbes de Bessel où les seuls composantes influentes sur le bas du spectre sont celles d'ordre 0,1 voire 2).

Néanmoins, on considérera ici que ce paramètre sera le même pour tous les ordres, aussi le choix de ce paramètre dépendra d'un compromis entre l'augmentation du bruit de fond inhérent aux capsules (aussi plus le nombre de capsules est élevé plus ce bruit de fond diminue) et l'extension de la bande passante pour chaque ordre.

On touche ici directement au choix entre fidélité audio et rapport signal sur bruit. C'est pourquoi les différents filtres issus de cette méthode devront également être validés par des tests d'écoute.

Ainsi on a donc :

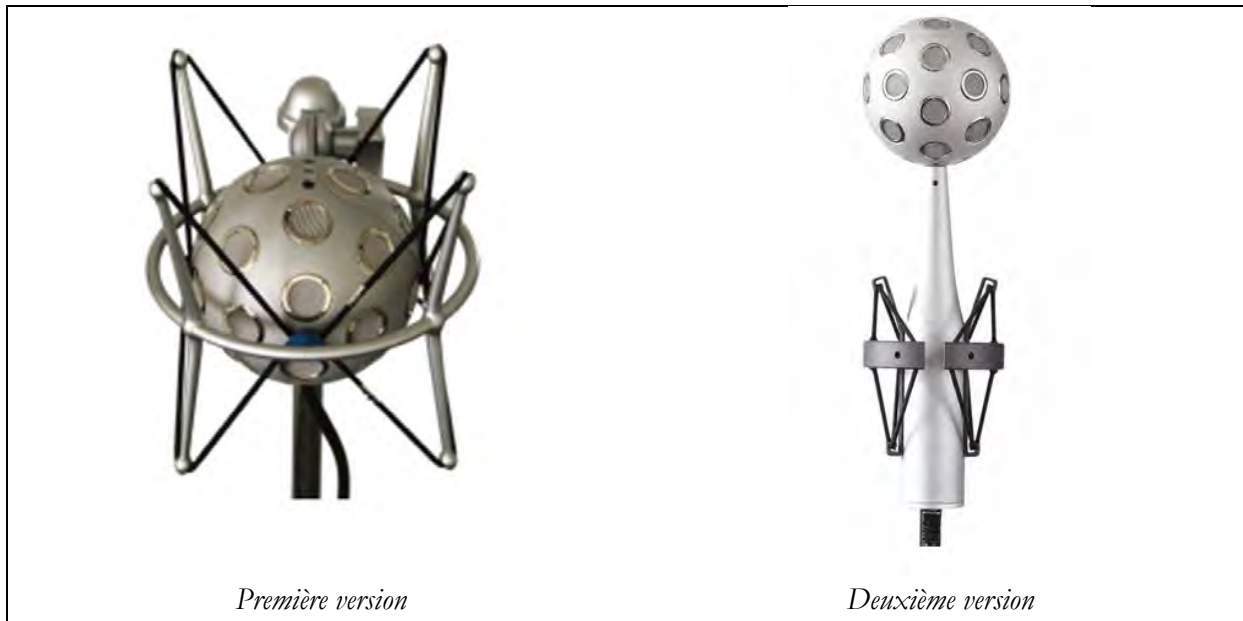
$$A_{max} = \sqrt{N} * 10^{\frac{a}{20}}$$

Avec N le nombre de capteurs, et à l'amplification maximale tolérée pour le rehaussement du bruit de fond. Ainsi, à partir du réglage d'un seul paramètre, on peut générer les matrices de filtres spécifiques à chaque ordre (définies par l'équation (7)).

Nous sommes donc en présence d'un modèle théorique satisfaisant, qu'il nous reste à appliquer à un cas concret, afin de pouvoir obtenir une caractérisation objective du dispositif testé.

### 2.3.5 Application et caractérisation d'un réseau de microphones

Le modèle que nous proposons d'étudier dans ce mémoire est le microphone Eigenmike® développé par Gary Elko et Jens Meyer au sein de MHAcoustics. Ce microphone a connu deux versions successives (cf Photos), et c'est à la deuxième que nous nous intéresserons.



Le microphone se compose de 32 microphones statiques à électret (capteur de pression, donc directivité omnidirectionnelle) répartis sur une sphère selon la figure du pentakidodécaèdre, qui assure une répartition régulière des capsules. La partie conversion analogique-numérique se fait directement dans le corps du microphone via des « convertisseurs 24 bits » (source : <http://www.mhacoustics.com/products>). Les 32 signaux sont multiplexés afin de pouvoir être transmis via un câble Ethernet (catégorie 5 minimum) avec connecteur RJ45 ou Ethercon, le protocole de transport est propriété de MHAcoustics. Ces signaux sont ensuite transmis vers une carte son propriétaire MHAcoustics : EMIB (pour Eigenmike Interface Box) qui permet de récupérer les 32 signaux en wav vers n'importe quel DAW. La communication entre l'ordinateur et l'EMIB se fait via firewire. Le gain de l'Eigenmike est piloté à distance à l'aide d'un logiciel très basique em32SetGain, qui comporte 3 indicateurs : EMIB, host, calibration pour signifier si la connexion se fait bien avec la carte son, le microphone et si la calibration est activée ou non. Celle-ci consiste en un gain appliqué à chaque capsule pour que « les signaux enregistrés via un DAW autre que celui fourni par MHAcoustics soient calibrés ». Aucune information supplémentaire

n'est fournie par la fiche technique, néanmoins les valeurs associées sont consultables via leur logiciel.

Penchons-nous sur les caractéristiques utiles pour le calcul des filtres associés, nécessaires à l'encodage microphonique. La sphère a un diamètre de 8,4 cm (ce qui est vérifié par la mesure, elle comporte 32 capteurs. Leur réponse en fréquence (supposée égale pour toutes les capsules) en champ libre a été mesurée indirectement d'après une mesure de directivité de l'eigenmike, par déduction prenant en compte la directivité de la sphère rigide. C'est donc une simulation pouvant comporter une légère marge d'erreur par rapport à la réponse de la capsule en champ libre, malheureusement non fournie par le constructeur.

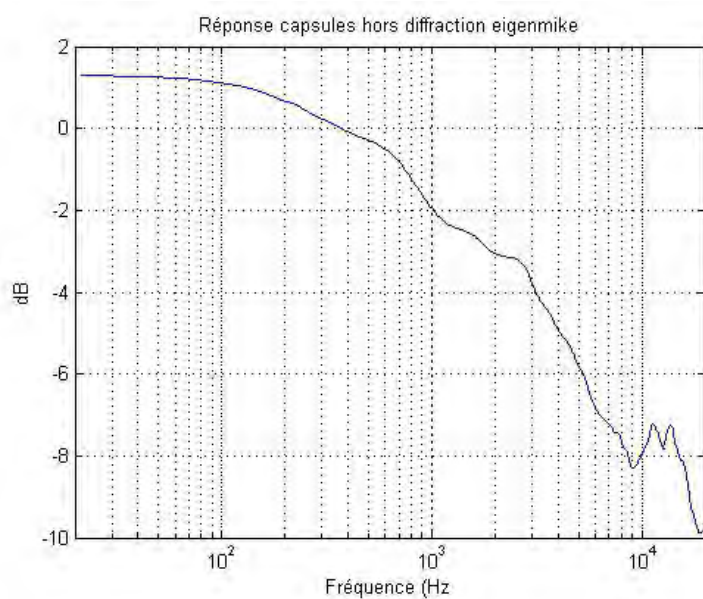


Figure 7 : Réponse en fréquence d'une capsule de l'Eigenmike

Cette réponse est compensée en amont de l'encodage microphonique, on convolue les 32 signaux par le filtre dont le gabarit est l'inverse de la réponse en fréquence d'une capsule. Ceci aura donc pour conséquence de remonter le haut du spectre.

D'après les propriétés mathématiques du pentakidodécaèdre, l'espacement entre chaque capsule est de 2,89 cm, soit une fréquence d'aliasing spatial de 5882 Hz. On exploite le modèle de la sphère rigide pour le calcul des filtres. Le calcul se fera avec différents paramétrages de  $a$ , pour 4 valeurs : +6 dB, +12 dB, +18 dB et +24 dB. Le tracé des courbes a été réalisé sous matlab et affiche à chaque fois la courbe calculée à partir du modèle idéal.

En plus du calcul des filtres jusqu'à l'ordre 4, la corrélation spatiale a aussi été représentée pour ces 4 paramètres ; celle-ci correspond au produit scalaire normalisé entre la directivité théorique fixée par les harmoniques sphériques, et celle trouvée en pratique par la simulation ; c'est cette corrélation qui nous informe sur l'aliasing spatial (cf [7]). Cet aliasing spatial est dû à l'influence des harmoniques sphériques d'un ordre sur un autre ordre (explicable comme une « diaphonie » spatiale). Pour la configuration en pentakidodécaèdre, la matrice d'erreur est présentée ci-dessous (cf Figure 8).

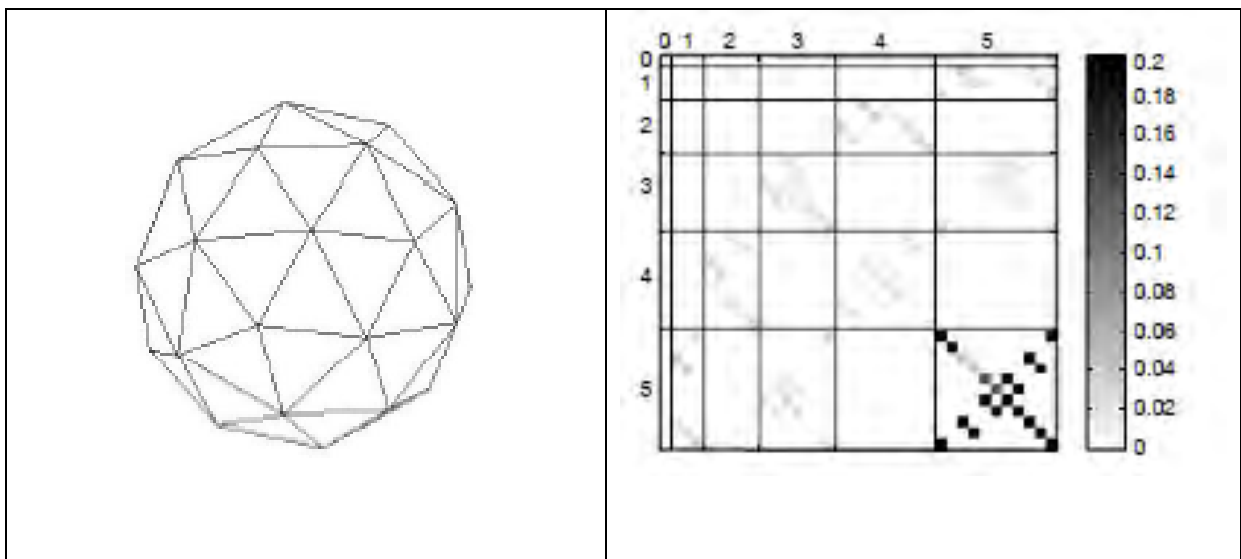


Figure 8 : Matrice d'erreur d'orthonormalité des harmoniques sphériques (source [7])

On constate par exemple, que les harmoniques sphériques d'ordre 5 vont apporter du bruit sur l'encodage des harmoniques d'ordre 1 et 3. Ainsi on devine que les harmoniques d'ordre supérieur à 5 vont elle aussi amener une erreur plus ou moins importante dans l'encodage du champ sonore. Cette erreur est insignifiante pour les fréquences en dessous de la fréquence d'aliasing spatial, mais peut devenir très importante passé celle-ci (cf Annexe 1, courbe encodage spatial : corrélation spatiale). L'apparition de lobes secondaires de plus en plus importants, va venir ainsi brouter l'information spatiale.



(a) optimal



(b) encoded,  $f = 5\text{kHz}$



(c) encoded,  $f = 10\text{kHz}$



(d) encoded,  $f = 15\text{kHz}$

*Exemple d'erreur d'encodage pour une directivité en fonction de la fréquence*

Pour le calcul de  $A_{max}$  on rappelle que sa valeur minimale est fixée par la sommation du rapport signal sur bruit des 32 capsules, soit :

$$20 \log \sqrt{32} = 15 \text{ dB}$$

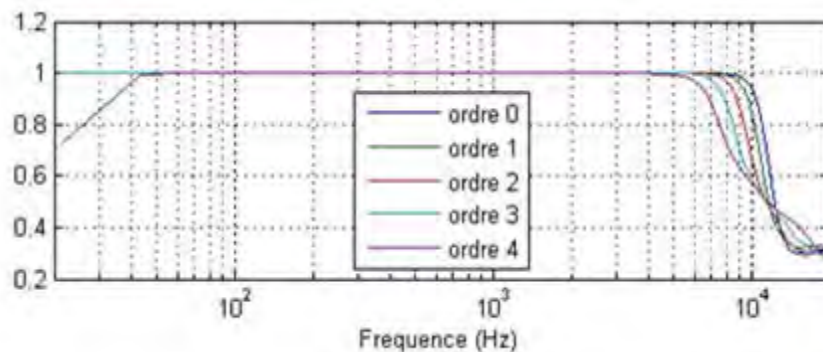
NB : on rappelle que ce chiffre ne correspond pas au rapport signal sur bruit du microphone, mais bien de la valeur minimale du seuil au-delà duquel les filtres d'égalisation cesseront d'amplifier le signal.

Pour une lecture plus aisée, une autre série de courbe appelée gain relatif montre la contribution à la description spatiale à chaque ordre d'harmonique sphérique en fonction de la fréquence.

Ces courbes sont visibles en annexe 1 : courbes encodage spatial

Pour rappel, le traitement se fait ici en 3 étapes, d'abord chaque canal est convolué avec le filtre inverse à sa réponse en fréquence (pour obtenir une réponse en fréquence résultante droite sur tout le spectre audio). Puis une matrice de gain est appliquée donnant la « forme » des directivités (i.e les harmoniques sphériques). La troisième étape est l'application des filtres  $EQ_m$ , qui corrige l'amplitude et la phase en fonction de la fréquence, avec la contrainte du paramètre d'amplitude maximale.

Tout d'abord on constate que la corrélation spatiale n'a pas de lien avec l'amplitude, autrement dit que les figures de directivité du microphone correspondent quelle que soit l'amplitude. La corrélation spatiale ne traduisant que la ressemblance de forme (sans l'échelle ni la phase), elle est donc totalement indépendante du filtrage.



*Corrélation spatiale en fonction de l'ordre, la perte de l'information spatiale débute à partir de 6kHz*

On remarque une chute de la corrélation spatiale (donc l'information spatiale encodée devient erronée) à partir d'environ 6kHz pour l'ordre 4, et au-delà par étalement pour les ordres supérieurs. Aussi, plus on descend dans l'ordre des harmoniques sphériques, plus le microphone est robuste à l'aliasing spatial. Ceci s'explique par la structure de la matrice d'erreur d'orthonormalité lié au micro (cf Figure 8), les harmoniques aux ordres supérieurs vont venir polluer la forme globale de directivité. Les figures de directivité des ordres inférieurs sont moins précises (ou directives), elles sont donc plus robustes à l'aliasing.



On constate des différences significatives entre le gabarit des filtres de l'Eigenmike et ceux du modèle idéal pour les valeurs +6 dB et + 12 dB. La réponse (en champ libre) non-plate des capsules de l'Eigenmike est en effet compensée et combinée à la correction de l'effet de diffraction par la sphère rigide, avant l'application de la méthode de régularisation (qui amène le seuil maximal d'amplification). Celle-ci n'opère donc plus de la même façon en fonction de la fréquence, et (cf Figure 7) comme la compensation de la réponse des capsules implique une remontée des hautes fréquences, il est logique de constater que les filtres  $EQ_m$  atteignent ce seuil plus rapidement en HF (donc pour les ordres plus élevés).

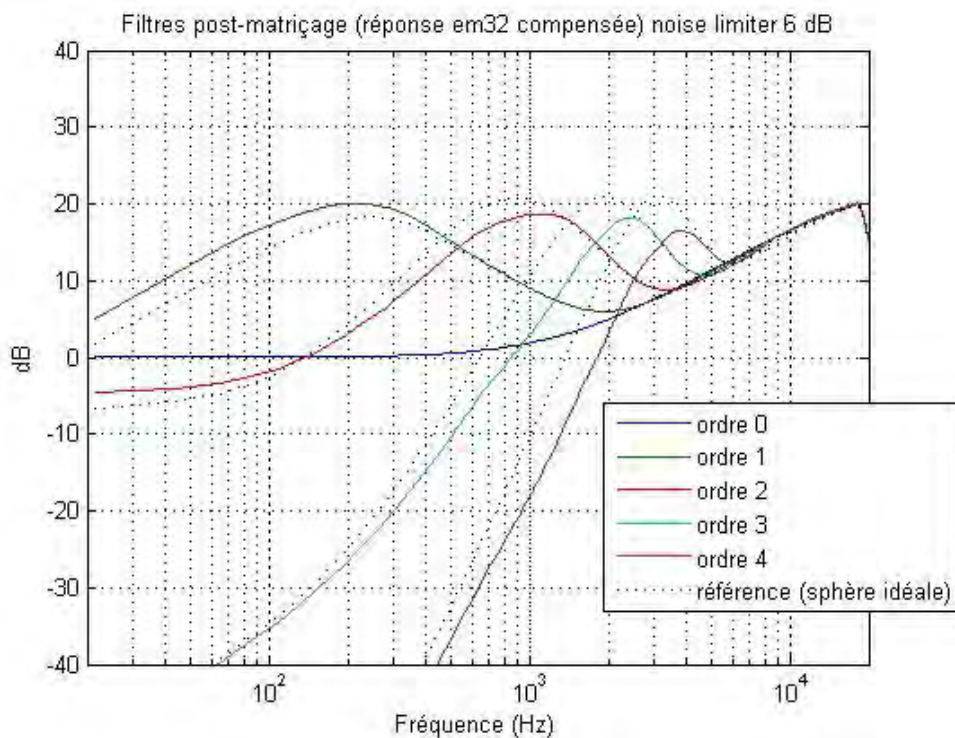


Figure 9 : Filtres post-matçage appliqués à l'encodage microphonique pour  $A_{max} = +21\text{dB}$

On constate d'ailleurs que dans le cas d'un microphone idéal (donc sans compensation de la réponse en fréquence des capsules), les filtres  $EQ_m$  idéaux, atteignent tous le même seuil d'amplitude maximal (+21 dB pour la figure 9).

Cette différence entre le seuil d'amplitude d'un microphone idéal et de l'Eigenmike s'amenuise pour les valeurs +18 dB et + 24 dB. En effet, les courbes de d'égalisation théorique, avec ou sans correction des capsules, atteignent le seuil de limitation d'amplification sur des régions de fréquence plus limitées.

Cette comparaison doit être relativisée néanmoins : il serait nécessaire de considérer le niveau de bruit réel des capsules et sa courbe d'évolution en fonction de la fréquence, lesquels n'ont pas encore pu être mesurés en chambre sourde au moment de la rédaction de ce mémoire. De même, il nous paraît nécessaire de pouvoir quantifier la dynamique du microphone, afin de pouvoir donner une valeur de son SNR, à mettre en lien avec les seuils d'amplification maximale proposés.

Enfin la courbe des gains relatifs nous montre que le choix d'un paramètre d'amplification maximale plus élevée permet d'augmenter significativement la bande passante pour chaque ordre d'harmoniques sphériques (jusqu'à plus d'une octave entre +6 dB et + 24 dB), et donc d'augmenter la qualité spatiale du champ encodé.

Une interprétation fréquentielle liée à l'audio de ces filtres post-matçage nous semble un peu hasardeuse, nous pouvons simplement estimer que plus le seuil d'amplification augmente, plus on compensera la perte d'informations en basse fréquence, renforçant ainsi cette partie du spectre et amenant un côté plus « rond ». Ce n'est toutefois qu'une hypothèse, et il faudra plutôt s'appuyer sur des tests perceptifs pour pouvoir aborder cette partie-là.

## 2.4 Traitements

### 2.4.1 Introduction

Comme évoqué précédemment, le formalisme HOA permet d'effectuer différents traitements sur le fichier encodé, sans modifier le fichier source. Nous nous proposons ici de faire un rapide tour d'horizon des traitements existants sur des fichiers HOA ainsi que les plugins associés : nous resterons sur la suite BRS, Ambix, Harpex, Soundfield BCOM, facilement trouvables dans le commerce (mis à part la suite BCOM non commercialisée pour l'instant qui a été utilisée pour ce mémoire). Les plugins Harpex-B et Soundfield Surround Zone 2 sont des solutions tout-en-un qui permettent plusieurs traitements et se limitent pour l'instant au B-Format, aussi nous nous contenterons de les nommer simplement (contrairement à Ambix ou BRS qui ont un plugin dédié pour chaque traitement). Nous répartirons ces différents traitements en deux catégories : Espace et EQ & Dynamique. Nous aborderons également les méthodes de visualisation liées aux signaux HOA.

Nous ne nous attarderons sur le formalisme mathématique associé à chaque traitement (qu'il faudrait idéalement développer), mais des références bibliographiques explicatives sont présentes à la fin de ce mémoire (partie Traitements & Visualisation).

### 2.4.2 Espace

On rappelle que l'encodage HOA d'une source audio (quel que soit le nombre de canaux) permet de lui imprimer un mouvement dans l'espace au cours du temps. On obtient donc un fichier HOA dans lequel une ou plusieurs sources audio se déplacent au cours du temps. Ainsi les modifications spatiales d'un fichier HOA concernent de manière globale toutes les trajectoires qui y ont été encodées. C'est pourquoi nous n'aborderons pas dans cette partie les plugins fonctionnant comme des encodeurs HOA (`ambix_encoder` (*Ambix*), `HOAencoder` (*BCOM*), `TOA Panner` (*BRS*), etc.).

#### 2.4.2.1 Rotation

Cette opération bien connue permet d'effectuer une rotation de la scène sonore. Néanmoins deux approches pour celle-ci sont possibles : soit la rotation concerne le point d'écoute (c'est-à-dire la « tête » au sein de la scène reproduite, ce se conçoit notamment en écoute binaurale avec *head-tracking*), soit la scène sonore.

Si l'on se place au point d'écoute, on simule un mouvement de tête de l'auditeur. Ainsi les 3 mouvements de tête élémentaires sont généralement appelés **Roll**, **Pitch** et **Yaw** et correspondent respectivement à une rotation autour de l'axe  $O_x$ ,  $O_y$  et  $O_z$ . Si on apparente cela au mouvement naturel, le **Roll** consiste à pencher la tête vers la gauche ou la droite, le **Pitch** consiste à lever la tête ou à la baisser et enfin le **Yaw** consiste à tourner la tête vers la gauche ou la droite.

Si l'on considère la scène sonore, les 3 opérations sont généralement appelées **Tumble**, **Tilt**, et **Rotate**, qui correspondent respectivement à une rotation autour de l'axe  $O_x$ ,  $O_y$  et  $O_z$ .

On comprend bien que ces deux notations se valent, ce sont simplement des indicateurs pour la conception d'un « Rotator ». Veut-on que la tête de l'auditeur virtuel pivote en écoutant la scène, ou que la scène pivote autour de l'auditeur.

Plugins existants : HoaRotator (*BCOM*), ambix\_rotator (*Ambix*), TOA Rotation (*BRJ*), TOA Zoom (*BRJ*), Surround Zone 2 (*Soundfield*), Harpex-B (*Harpex*), ...

#### 2.4.2.2 Symétrie

Même si la symétrie par rapport à un axe revient à effectuer une rotation d'angle  $\pi$  selon ce même axe, on retrouve cette possibilité dans certains plugins. Nous les mentionnons donc ici.

Plugins existants : ambix\_mirror (*Ambix*)

#### 2.4.2.3 Translation (ou “warp”)

De la même manière que la rotation, une « translation » de la scène sonore est possible, on opère ainsi un éloignement/rapprochement de la scène par rapport à la position d'écoute. Vu qu'on se place en coordonnées sphériques, la métaphore du zoom est plus parlante pour décrire cet effet (on déplace la position d'écoute vers un point ce qui aura pour effet de rabattre les éléments avoisinants sur l'espace restant).

Plugins existants : ambix\_warp (*Ambix*), TOA Move (*BRJ*), TOA Zoom (*BRJ*), Harpex-B (*Harpex*)

#### 2.4.2.4 Diffusion (ou “Widening”)

Mathias Kronlachner (créateur de la suite Ambix) définit ce traitement dans un article (cf. Bibliographie) comme « une rotation autour de l’axe z dépendante de la fréquence, qui a pour effet un élargissement des sources ou la création de premières réflexions ». On se retrouve donc avec un traitement amenant une sensation de scène plus élargie, un peu à la manière d’une réverbération « premières réflexions ».

Plugins existants : ambix\_widening (*Ambix*), TOA Diffuser (*BRS*)

#### 2.4.2.5 Masquage

Ce traitement consiste en une sorte de filtrage spatial, où il est possible de « masquer » une partie de la scène sonore en atténuant/augmentant le volume dans des portions de l’espace que l’utilisateur définit de manière plus ou moins précise.

Plugins existants : TOA Directional Mask, TOA Spatial Mask (*Blue Ripple Sound*), ambix\_directional\_loudness (*Ambix*)

#### 2.4.2.6 Réverbération

Comment ne pas parler de traitement spatial sans parler de réverbération ? A l’instar des chaînes de traitements usuels, le format HOA permet également l’ajout d’une reverb. Il existe donc des plugins de reverb artificielle, néanmoins il est tout à fait possible d’utiliser une reverb à convolution avec un fichier HOA, cela nécessite d’avoir une réponse impulsionnelle mesurée avec un dispositif microphonique correspondant à l’ordre voulu et de l’encoder en format HOA (ou d’y avoir directement accès en format HOA et de l’importer). Il suffit d’avoir un plugin qui permet une convolution avec un nombre suffisamment élevé de canaux (et un paramétrage suffisant), ou de faire appel à plusieurs instances de ce plugin si la charge CPU nous le permet.

Plugins existants :

Réverbération à convolution : Pristine Space (*Voxengo*), Altiverb (*AudioEase*), Ircam Verb (*Flux*) ...

Réverbération HOA : TOA Reverb (*BRS*) (avec possibilité d’ajout de premières réflexions, dont on peut fixer la direction de la provenance)

#### 2.4.2.7 Microphone virtuel

La synthèse de directivité de microphone virtuel est possible dans un format HOA, elle revient à sélectionner une portion de l'espace plus ou moins finement (selon un ordre de directivité plus ou moins élevé). Ainsi, on peut extraire un fichier audio correspondant à ce microphone virtuel depuis le fichier HOA. On voit bien ici que les possibilités sont donc multiples, soit on se concentre sur une source précise (« beamer ») en simulant un microphone virtuel le plus directif possible, soit plusieurs microphones virtuels à la fois pour retomber sur des cas plus classiques (comme un couple stéréo coïncident ou non, 5 cardioïdes placés selon une prise 5.0, etc.). D'autres applications plus exotiques sont également permises.

Plugins existants : ambix\_vmic (*Ambix*), TOA Virtual Microphone (stéréo seulement) (*BRS*), Eigenstudio (*MHLAcoustic*), Surround Zone 2 (*Soundfield*), Harpex-B (*Harpex*)

#### 2.4.2.8 Upmixer

Il est également possible de passer d'un fichier HOA d'un ordre quelconque à un fichier HOA d'ordre plus élevé (on utilisera le terme d'upmix faute de mieux), cela peut avoir un intérêt pour l'utilisation de certains plugins limités à un ordre particulier. Par exemple tous les plugins BRS fonctionnent avec des fichiers HOA d'ordre 3, de même les plugins Ambix ne fonctionnent que pour un ordre donné (il est néanmoins possible de recompiler chacun des plugins à l'ordre souhaité).

Plugins existants : Harpex-B (*Harpex*) (jusqu'à l'ordre 3), TOA Harpex Upsampler (*BRS*)

#### 2.4.2.9 Convertisseurs

Comme vu précédemment, les nomenclatures de normalisation et de catégorisation sont à considérer avec le plus grand soin. En effet, chaque suite de plugins correspond à une nomenclature donnée (Furse-Malham pour BRS, ACN/SN3D pour Ambix p.ex). Aussi est-il nécessaire d'avoir des convertisseurs permettant de passer de l'un à l'autre. Il n'existe pas à ma connaissance de modules permettant de détecter sous quel formalisme un fichier HOA non référencé se trouve.

Plugins existants : ambix\_convert (*Ambix*) (comprend toutes les nomenclatures), TOA N3D Injector (ne prend que du ACN pour la catégorisation, et du N3D ou

SN3D pour la normalisation en entrée, et renvoie un format adapté aux plugins BRS).

### 2.4.3 EQ & Dynamique

De même que pour n'importe quelle chaîne audio, il est nécessaire de pouvoir appliquer des traitements fréquentiels et dynamiques à notre fichier HOA. On se souvient qu'il est possible de prétraiter plusieurs sources audio (monophonique ou multicanal) avec les suites de plugins usuels, avant de les encoder en format HOA. Aussi les traitements abordés dans cette partie s'appliquent directement au fichier HOA. Nous avons construit cette partie sur le même modèle qu'une tranche de console pour plus de facilité.

#### 2.4.3.1 Gain et délai

Nous avons vu dans la partie précédente que des modifications spatiales pouvaient modifier le gain de certaines portions de l'espace. Néanmoins rien n'empêche d'appliquer un gain global au fichier HOA. D'autres subtilités sont possibles, par exemple affecter un gain pour les signaux HOA d'un ordre en particulier (si on veut renforcer les composantes d'ordre 1 par exemple). Il en va de même pour le délai, il est tout à fait possible de délayer un fichier HOA dans sa globalité. Délayer les signaux HOA d'un ordre en particulier est techniquement possible, mais n'a pas été implémenté à ma connaissance (l'intérêt technique serait en effet assez hasardeux). On ne listera pas ainsi les plugins de gain global ou de délai global (de plus certains DAW le permettent directement sur la piste) car trop nombreux.

Plugins existants : TOA Order Amplifier (BRS) (capable d'augmenter indépendamment les signaux HOA d'ordre 0, 1,2 et 3)

#### 2.4.3.2 EQ

Le format HOA permet les traitements fréquentiels classiques, aussi on retrouve les filtres paramétriques classiques (shelf, bell, notch, low pass, hi pass) qui s'appliqueront à tout le mix HOA. De même, un égaliseur graphique 31 bandes assurera le même rôle sur tout le mixage HOA. Néanmoins le format HOA permet d'égaliser différemment plusieurs portions de l'espace (utile si on a réalisé par exemple un enregistrement d'orchestre avec un microphone HOA et que l'on veut traiter fréquentiellement plusieurs pupitres séparément ; ceci n'est qu'une possibilité parmi d'autres).

Plug-ins existants: TOA Parametric EQ, TOA Graphic EQ, TOA Spatial Equalizer (BRS)

### 2.4.3.3 Dynamique

De manière analogue au traitement fréquentiel, le traitement de la dynamique en HOA peut être abordé d'un point de vue classique : on applique un compresseur/limiteur/expandeur/gate à tout le mix. Mais ce format permet également de compresser différemment plusieurs zones de l'espace (si on reprend l'analogie d'une prise de son orchestrale, on pourrait compresser uniquement les pupitres avec une forte dynamique (cuivres, voix, ...)).

Plug-ins existants : HOAPeakLimiter (BCOM), TOA Compressor (BRS)

## 2.4.4 Autres

On notera que d'autres plug-ins plus exotiques existent, nous incitons donc le lecteur si il le souhaite à se rendre dans la partie bibliographique pour y avoir accès et les tester par lui-même !

De même, les différentes suites comprennent en général des décodeurs HOA plus ou moins flexibles, adaptés soit à des cas connus (2.0, 5.1, 7.1, Auro3D, etc), soit à des configurations particulières de haut-parleurs (nécessitant dans ce cas l'édition d'un fichier de configuration). Nous y reviendrons dans la partie décodage spatial de ce mémoire.

En résumé, on constate que peu de plug-ins existent pour Ambisonics, malgré l'ancienneté de son invention. Néanmoins, le renouveau d'intérêt pour HOA et les nouvelles possibilités de développement logiciel amenées avec ce formalisme permettent pour l'instant une intégration partielle de ce format dans un workflow classique (notamment en ce qui concerne l'intégration d'ambiances HOA, les traitements spatiaux possibles).

Il nous semble que les outils sont encore un peu légers (ils ont été développés très récemment !) pour une chaîne de traitement intégralement en HOA. Néanmoins ces plug-ins amènent une première base sérieuse qui ne demande qu'à être étayée et surtout testée par des professionnels afin de pouvoir les améliorer et les intégrer à des réalisations audiovisuelles (la suite Blue Ripple Sound a par exemple un pied dans l'audio 3D appliqué au jeu vidéo et notamment aux casques audio). Nous ne sommes donc pas si loin de pouvoir travailler sur une chaîne de traitement HOA.



La seule possibilité d'obtenir un master HOA en bout de chaîne, diffusable sur n'importe quelle configuration de haut-parleur (et en haute qualité), ou les possibilités de traitement spatial devraient ainsi intéresser nombre de professionnels de l'audiovisuel.

## 2.5 Décodage spatial

### 2.5.1 Introduction

Nous voilà arrivés au dernier maillon de la chaîne de traitement HOA : sa diffusion sur une configuration donnée de haut-parleurs. Le format HOA étant indépendant du dispositif de diffusion, on comprend donc que le décodage de celui-ci sera lié au nombre de haut-parleurs ainsi qu'à leur disposition dans l'espace. Ces paramètres nous renseigneront directement sur l'ordre maximal possible pour une restitution sonore. On considère ici un décodage spatial dont la vocation est de reconstruire les composantes HOA (dont le champ de pression et ses dérivées spatiales jusqu'à l'ordre HOA considéré) au centre du dispositif. La littérature (cf. [2] p.178) nous apprend qu'une restitution optimale en 2D ou en 3D se fait pour une configuration régulière de haut-parleurs (i.e. répartis de manière régulière à la surface d'un cercle en 2D, d'une sphère en 3D) dont le nombre est égal ou juste supérieur au nombre d'harmoniques sphériques à l'ordre voulu.

Autrement dit pour une restitution à l'ordre  $M$ , soit  $N$  le nombre de haut-parleurs :

$$\begin{cases} (2D) & N \geq 2M + 1 \\ (3D) & N \geq (M + 1)^2 \end{cases}$$

Précisons que la définition d'une configuration régulière en 3D est beaucoup plus contraignante qu'en 2D où il suffit de considérer les sommets d'un polygone régulier (il en existe forcément quelque soit le nombre  $N$ ). Néanmoins, comme nous le verrons par la suite, une restitution est possible sur n'importe quelle configuration de haut-parleurs.

Formalisons un peu. Soit  $S$  notre signal encodé en HOA (sous forme de matrice, dont la dimension est égale à  $K \times t$ , où  $K$  est le nombre de signaux HOA à l'ordre  $M$ , et  $t$  la durée du fichier en échantillons) à diffuser sur une configuration de  $N$  haut-parleurs. Appelons donc  $S_N$  les signaux diffusés par les  $N$  haut-parleurs une fois notre signal décodé (sous forme de matrice également, de dimension  $N \times t$ , où  $t$  est la durée du fichier en échantillon).

Nous cherchons donc à définir la matrice de décodage  $D$  (matrice de dimension  $N \times K$ ), telle que :

$$S_N = D * S \tag{8}$$

Introduisons la matrice de ré-encodage  $C$ . En effet, le champ reconstruit par les  $N$  haut-parleurs peut lui aussi être décomposé sur une base d'harmoniques sphériques, en considérant chaque haut-parleur comme un générateur d'onde plane (si placés suffisamment loin du point d'écoute). C'est donc la matrice d'encodage fonction de la position de chaque haut-parleur tel que  $S'$  le signal HOA résultant soit égal à :

$$S' = C * S_N \quad (9)$$

Où  $S'$  est de même dimension que  $S$ , et  $C$  de dimension  $(K \times N)$ .

En couplant donc (8) et (9) on obtient :

$$S' = C * D * S \quad (10)$$

On comprend donc que, si l'on veut que le signal encodé  $S$  soit égal au signal décodé  $S'$ , il faut que  $C * D$  soit égal à  $I_K$ , la matrice identité de dimension  $K \times K$ , autrement dit que la matrice de décodage  $D$  soit l'inverse de la matrice de ré-encodage  $C$ .

On obtient donc une définition de notre matrice de décodage. Par contre, rien ne nous informe que la matrice  $C$  soit carrée et inversible quelle que soit la configuration de haut-parleurs (elle ne le sera pas la plupart du temps). Aussi travaille-t-on avec sa matrice pseudo-inverse, ce qui nous amène à redéfinir la matrice  $D$  sous la forme :

$$D = C^T (C^T * C)^{-1} \quad (11)$$

C'est la forme la plus simple pour obtenir une matrice de décodage pour n'importe quel dispositif de haut-parleurs, par le calcul de la matrice pseudo-inverse. C'est pourquoi l'on nommera cette méthode de décodage **pseudo-inverse**.

Revenons aux configurations optimales (pour  $K = N$  et une configuration régulière de haut-parleur). J.Daniel montre dans sa thèse que la matrice de décodage  $D$  s'exprime sous la forme simplifiée suivante :

$$D = \frac{1}{N} * C^T$$

Plus généralement, on comprend bien pourquoi le fichier HOA est indépendant de la configuration de haut-parleurs, mais pas sa restitution. Aussi, si un fichier a été encodé en HOA à l'ordre  $M$ , il pourra également être décodé à l'ordre  $M$ , ainsi qu'à tous les ordres inférieurs (car ses informations sont également présentes dans le fichier). Néanmoins, on ne pourra pas décodé à un ordre supérieur à  $M$  car les informations relatives aux ordres supérieurs ne sont tout simplement pas présentes dans le fichier encodé.

L'intérêt d'un décodage à un ordre le plus élevé possible relève des mêmes arguments que pour l'encodage. En effet, plus l'ordre est élevé, plus le champ sonore pourra être reconstruit avec une fidélité spatiale accrue. Le formalisme HOA se construit à partir d'un point central (le centre de la sphère et donc le point d'une écoute individuelle), aussi un ordre plus élevé assure une zone de reconstruction optimale plus large, c'est-à-dire où le champ est reproduit sans artefact. Si on revient au cas de l'onde plane, pour une zone d'écoute de la taille d'une tête humaine (à peu près 20 cm), on obtient les fréquences limites d'aliasing spatial suivantes en fonction de l'ordre (cf [13], p.32).

$f_{\text{lim}}$ (Hz)	700	1300	1900	2500
Ordre HOA	1	2	3	4

*Fréquence limite de reconstruction pour une onde plane sur une zone d'écoute de 20 cm au centre du dispositif*

Ces fréquences ne veulent pas dire que la description s'arrête au-delà de ces fréquences, mais simplement qu'elle sera de plus en plus erronée. On comprend que le design des matrices de décodage dépend aussi de l'utilisation du dispositif, s'il est destiné à un nombre réduit d'utilisateurs placés au centre, ou à une audience élargie comme dans le cadre d'une diffusion en public.

De plus l'impact d'un plus grand nombre de haut-parleurs sur la restitution n'est pas clair : Gerzon recommande un plus grand nombre de haut-parleurs pour la restitution en Ambisonics (ordre 1) pour éviter le *detent effect* ou effet du son « collé au haut-parleur ». Néanmoins, des tests récents menés par Solvang (cf. [19]) ont montré qu'utiliser un nombre plus élevé de haut-parleurs provoque un filtrage en peigne au « sweet spot » et donc une perception dégradée, et conclut qu'une amélioration technique à l'aliasing spatial mène nécessairement à une dégradation de la qualité audio.

On privilégiera ainsi toujours un encodage à l'ordre le plus haut possible afin de pouvoir restituer le signal le mieux possible sur la plupart des dispositifs de haut-parleurs.

Par exemple, pour un fichier encodé en HOA à l'ordre 5, le nombre optimal de haut-parleurs sera de 36. Il pourra donc être restitué dans de bonnes conditions sur à peu près n'importe quel système d'écoute.

Enfin, il est quand même utile de noter que la qualité des haut-parleurs a bien évidemment un impact déterminant sur la qualité audio du dispositif. Il est donc essentiel pour une application professionnelle d'avoir des enceintes de qualité.

Pour en revenir au design des décodeurs, il n'existe donc pas de meilleure solution, mais des designs adaptés aux caractéristiques du dispositif de restitution (nombre d'enceintes utilisées, taille de la salle), de la taille de la zone d'écoute, du nombre d'auditeurs et de leur placement, de la précision recherchée, des signaux utilisés (musique, parole, ...), etc. Par conséquent, le design des décodeurs demeure un compromis entre leur robustesse pour tout type de configuration et les performances du point de vue des qualités audio et spatiales.

Ce mémoire n'a donc pas pour but de détailler les nombreuses approches existantes (nous renverrons le lecteur à la bibliographie pour cela, notamment à [2], [16], [20], [21]), mais de définir une méthode de caractérisation objective pour différents décodeurs appliqués à une salle d'écoute préexistante. Nous devons donc déterminer des outils pour nous permettre de comparer les performances de ceux-ci.

## 2.5.2 Outils de caractérisation : vecteur vitesse et vecteur énergie

« Les oreilles prises comme un système peuvent être considérées comme une fonction dépendante de la direction » (cf. [22]). Michael Gerzon introduit entre autres dans sa métathéorie de la localisation audio les vecteurs vitesse et énergie comme indice de localisation d'un son dans l'espace. Il les définit de la manière suivante :

Soit  $N$  onde plane de pression, d'amplitude  $P_i$  et d'angle d'incidence  $\vec{U}_i$   $\begin{cases} x_i \\ y_i \\ z_i \end{cases}$

Il définit les vecteurs et quantités suivantes relatives à ces signaux de pression :

$$\vec{V}' \begin{cases} x'_V = \sum_{i=1}^N x_i P_i \\ y'_V = \sum_{i=1}^N y_i P_i \\ z'_V = \sum_{i=1}^N z_i P_i \\ w_V = \sum_{i=1}^N P_i \end{cases} \quad \text{où } w_V \text{ est la somme de toutes les amplitudes.}$$

Il divise les 3 composantes  $x_V, y_V$  et  $z_V$  par le facteur  $w_V$ , afin de normaliser l'amplitude de  $\vec{V}'$  entre 0 et 1 (on notera  $X_V, Y_V$  et  $Z_V$ ) ses composantes normalisées). Ce sont des nombres complexes, aussi il distingue leur partie réelle qui constitue le vecteur vitesse, et leur partie imaginaire qui constitue selon Gerzon, un indice de « phasing » ou d'effet de phase qu'on ressent en bougeant la tête (similaire au hors phase en stéréophonie).

Il aboutit donc à  $\begin{cases} x_V = r_v * Re (X_V) \\ y_V = r_v * Re (Y_V) \\ z_V = r_v * Re (Z_V) \end{cases}$  qui définissent les composantes du vecteur  
vitesse  $\vec{V}$ .

Avec  $r_v = \sqrt{Re (X_V)^2 + Re (Y_V)^2 + Re (Z_V)^2}$  , soit la norme du vecteur  
vitesse, comprise entre 0 et 1.

On notera que pour une seule onde plane ( $N = 1$ ),  $r_v = 1$ .

Gerzon préconise que, pour qu'une source reconstitué par un ensemble de haut-  
parleurs soit bien localisé,  $r_v$  doit être le plus proche possible de la valeur unité. Il  
précise que le vecteur vitesse est l'indicateur qui prévaut pour la localisation pour  
les fréquences en dessous de 1000 Hz.

Il définit de manière analogue pour la localisation en haute fréquence (où la tête ne  
devient plus négligeable pour la longueur d'onde), le vecteur énergie, indicateur de  
la direction du son. Il n'a pas de réalité physique, mais représente la moyenne des  
énergies associées à chaque source.

$$\vec{E}' \begin{cases} x'_E = \sum_{i=1}^N x_i |P_i|^2 \\ y'_E = \sum_{i=1}^N y_i |P_i|^2 \\ z'_E = \sum_{i=1}^N z_i |P_i|^2 \\ w_E = \sum_{i=1}^N |P_i|^2 \end{cases}$$

On normalise ce vecteur en divisant les 3 composantes par  $w_E$ . On note  
 $X_E, Y_E$  et  $Z_E$  ses composantes normalisées, qui sont des nombres complexes. Il  
définit donc le vecteur énergie par :

$$\vec{E} \begin{cases} x_E = r_E * Re (X_E) \\ y_E = r_E * Re (Y_E) \\ z_E = r_E * Re (Z_E) \end{cases}$$

Avec  $Re$  la partie réelle, et  $r_E = \sqrt{Re (X_E)^2 + Re (Y_E)^2 + Re (Z_E)^2}$  .

On remarque que  $r_E$  est égal à 1 pour une seule onde plane ( $N = 1$ ).

Cet indicateur  $r_E$  pour Gerzon doit également être toujours le plus proche possible de 1 pour un son reconstitué afin d'assurer une bonne localisation en haute fréquences (au-dessus de 1000 Hz). Il faut noter que Gerzon préconisait également l'utilisation d'un troisième indice (avec des termes de degré 3), mais qui n'a pas été retenu.

Ces deux indicateurs ont été généralisés pour HOA par J.Daniel (cf. [2]) et appliqués pour estimer la direction d'une source reconstruite après décodage HOA sur un dispositif de haut-parleurs.

Appliquons ces deux indices à un dispositif de N haut-parleurs, disposés en configuration régulière. On veut estimer la direction d'une source S située virtuellement en dehors du dispositif, avec  $\vec{u}_n$  le vecteur unitaire qui pointe vers le n-ième haut-parleur. On obtient donc après décodage HOA de cette source S, les vecteurs vitesse et énergie de la manière suivante :

$$\begin{cases} \vec{V} = \frac{\sum_{i=1}^N g_n \vec{u}_n}{\sum_{i=1}^N g_n} \\ \vec{E} = \frac{\sum_{i=1}^N |g_n|^2 \vec{u}_n}{\sum_{i=1}^N |g_n|^2} \end{cases}$$

$$\text{Avec } \begin{cases} \vec{V} = r_v \vec{u}_v \\ \vec{E} = r_E \vec{u}_E \end{cases} \quad (12)$$

Ces deux indicateurs nous fournissent donc une manière de caractériser spatialement un décodeur, grâce à l'étude de la direction et de la norme de ces deux vecteurs.

Nous nous appuyerons donc sur eux pour tenter une description spatiale de différentes matrices de décodage.

Nous n'avons malheureusement pas eu le temps d'implémenter un outil adéquat pour caractériser l'impact sur le spectre audio des différentes matrices de décodage. Un tel outil doit prendre en compte le cas échéant, l'impact d'un encodage microphonique non idéal, et dans tous les cas l'ensemble de la chaîne de reproduction, jusqu'à la construction des signaux binauraux. L'impact spectral est alors à apprécier en fonction de la direction de la source sonore. Cette piste reste donc à creuser et demeure une bonne piste pour des tests subjectifs.



### 2.5.3 Caractérisation spatiale d'un décodeur

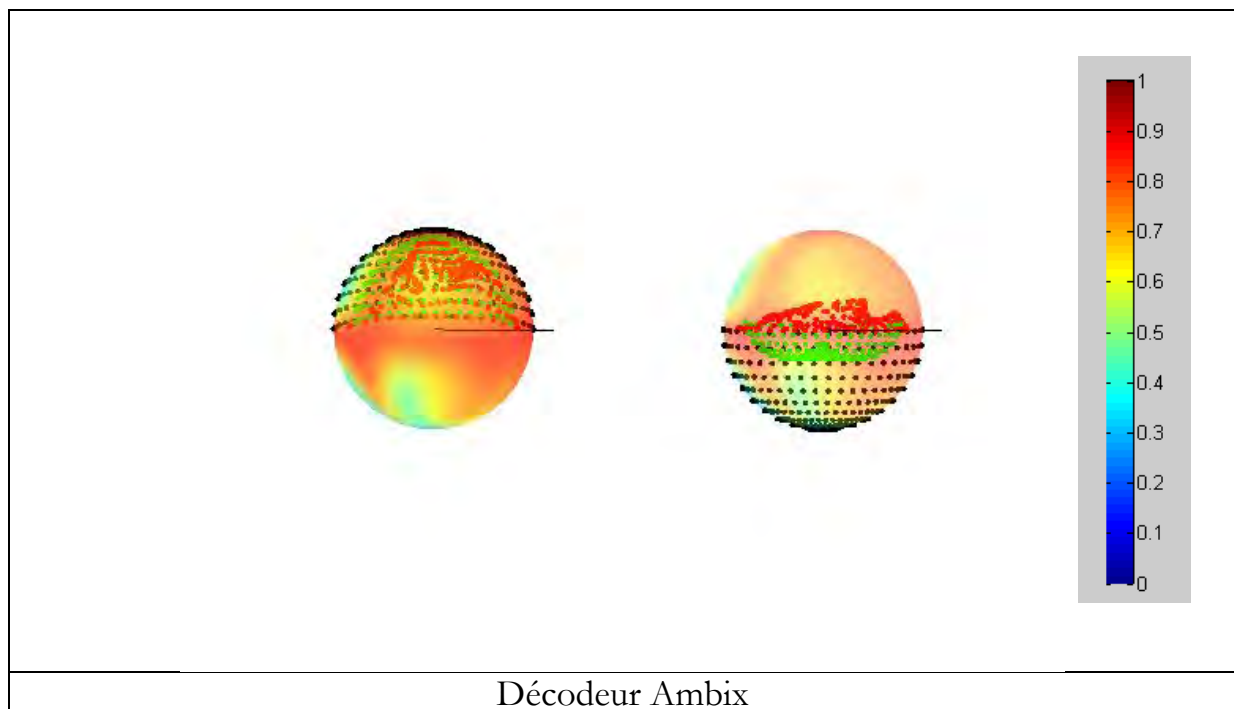
Nous nous proposons donc d'utiliser les indicateurs introduits dans la partie précédente pour tenter une première caractérisation de différents décodeurs. Cette partie est à mettre en rapport avec la partie 3.3, qui présente les décodeurs choisis et les trajectoires abordées pour un test subjectif mené dans le cadre de la partie pratique.

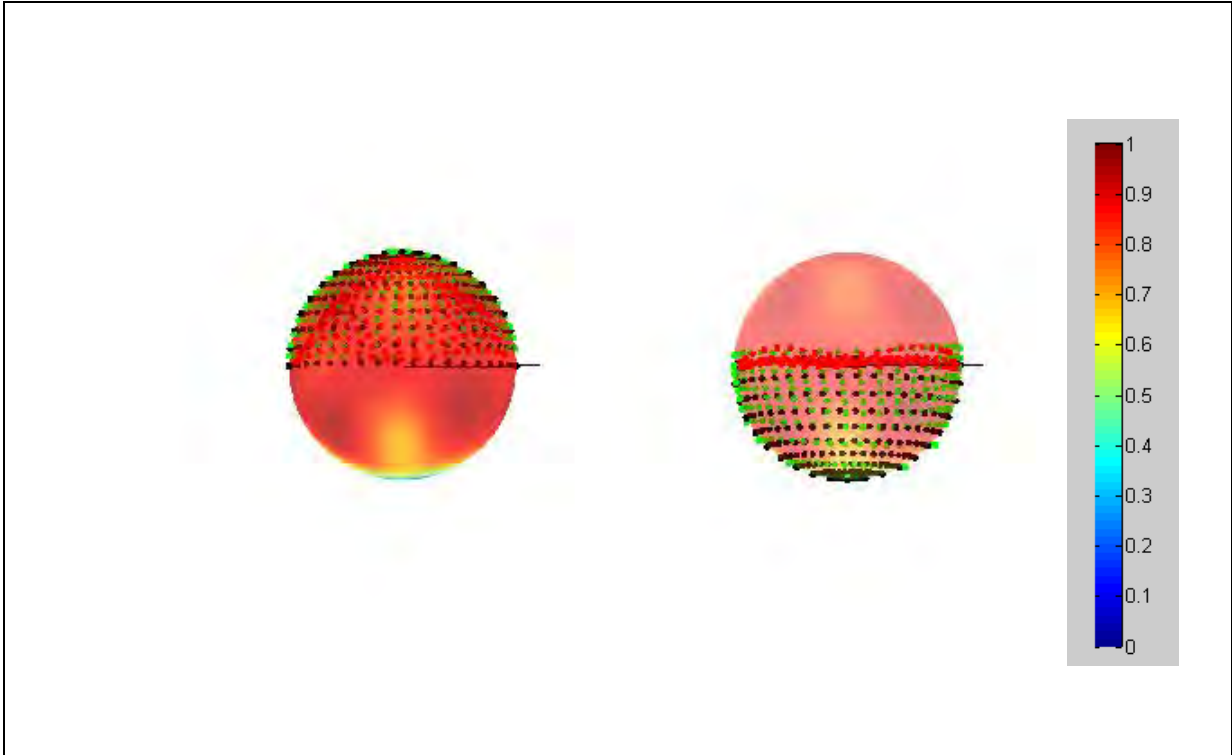
L'étude ici des différents décodeurs se fera donc sur une configuration de 14 haut-parleurs répartis régulièrement sur une structure hémisphérique, les positions de chaque enceinte sont rappelées dans la partie 3.2.

La partie visualisation a été faite sous le logiciel Matlab. La méthodologie utilisée est la suivante :

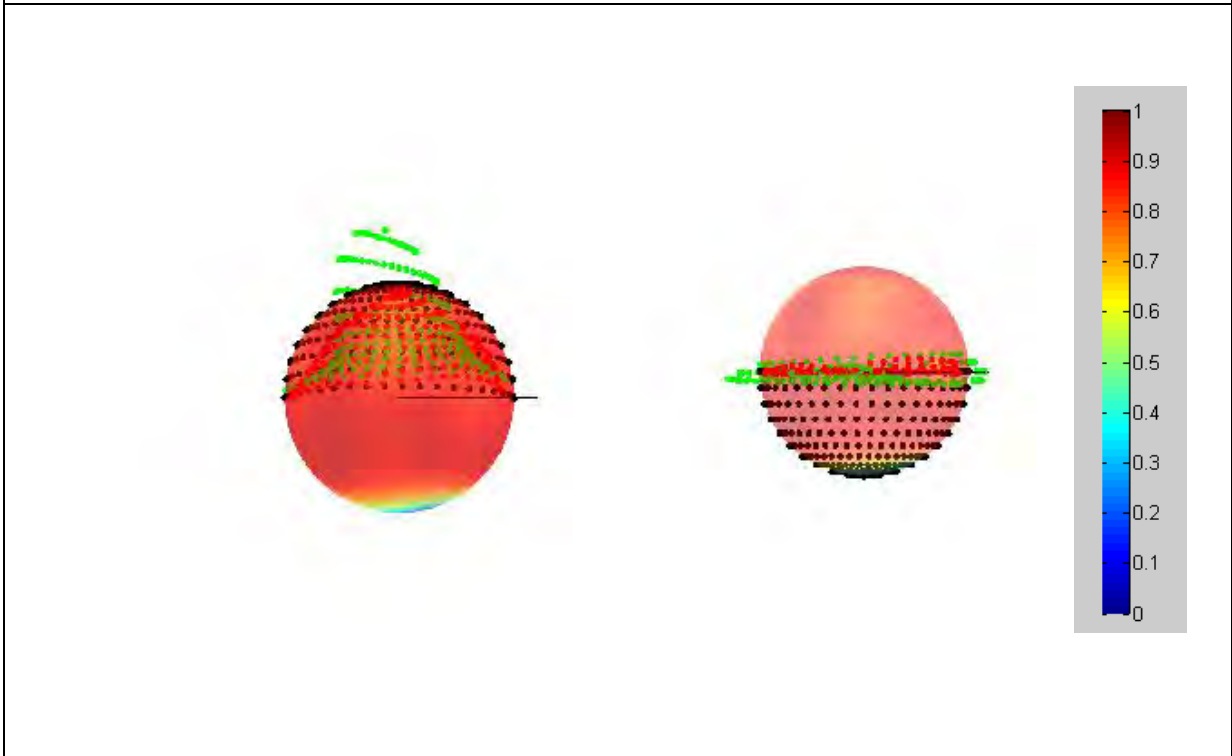
On échantillonne une sphère avec un grand nombre de points, qu'on encode comme des sources sonores (i.e générateur d'ondes planes placé à l'infini) en HOA. On fournit les positions des haut-parleurs en coordonnées sphériques (azimut, élévation) aux différents décodeurs, et on obtient nos 14 canaux de sortie qui nous fournissent les positions des vecteurs énergie et vitesse pour chaque source (avec la formule (12)).

On obtient les graphes suivants pour les décodeurs utilisés dans la partie 3.3

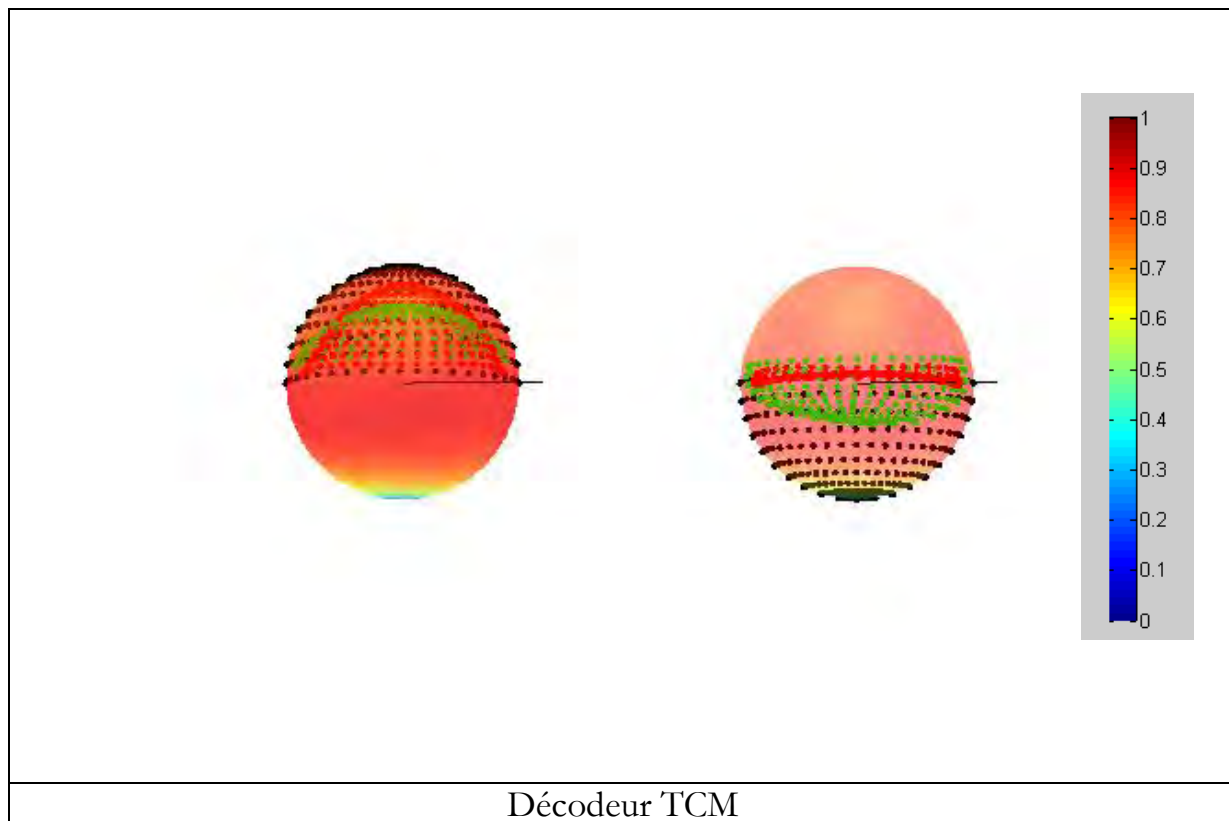




Décodeur BCOM



Décodeur BRS



Le vecteur énergie est représenté en **rouge**, le vecteur vitesse en **vert**, et les sources en **noir**. Nous avons séparé les deux hémisphères (supérieur et inférieur) pour plus de lisibilité. La sphère a un rayon unitaire, c'est-à-dire que les points présents à sa surface, traduisent que les vecteurs ont une norme égale à 1, inférieure à 1 si le point est à l'intérieur, et supérieur à 1 si il est à l'extérieur.

L'échelle de couleur de la sphère représente l'énergie totale diffusée en fonction de l'incidence de la source, on cherche à ce qu'elle soit le plus homogène possible. C'est la moyenne quadratique de tous les gains des haut-parleurs pour une source donnée. L'échelle de couleur est présentée dans la barre de couleur associé à chaque figure.

Les figures présentées ci-dessus, sont des coupes de représentation en 3D, aussi l'analyse se porte sur ces représentations, les coupes sont présentées pour aider à la visualisation.

A première vue, ces graphes corroborent le fait que pour un même dispositif de haut-parleurs, on peut obtenir des résultats dissemblables avec différents décodeurs.

Gerzon préconise une colinéarité des vecteurs énergie et vitesse dans le design du décodeur, il suppose que le vecteur énergie assure la localisation de la source, et le vecteur vitesse sa stabilité. Ainsi si le vecteur énergie et vitesse ont des directivités très différentes, on aura une source plus ou moins localisée (en fonction de la norme du vecteur énergie) mais instable.

Pour l'analyse, on regarde si la norme du vecteur énergie est proche de 1 (donc proche du bord de la sphère), car plus elle l'est, plus la source sera localisée, un accident d'énergie veut dire que la source va se coller aux haut-parleurs. Si la norme du vecteur vitesse est supérieure à 1 (elle sort de la sphère) on va avoir un effet de « surlatéralisation », c'est-à-dire un effet s'approchant du hors-phase en stéréophonie, où la source sera perçue de manière déformée.

On remarque que sur l'hémisphère supérieur, les 4 décodeurs ont une répartition homogène du vecteur énergie (sauf le décodeur Ambix où la figure en triangle laisse penser que le haut-parleur placé au-dessus de la tête a trop de « poids » dans le décodage). On voit également que la norme du vecteur vitesse dépasse 1 pour les sources en élévation dans le décodeur BRS. On peut donc s'attendre à des effets de phasing sur les sources en élévation.

La gestion de l'hémisphère inférieure est plus problématique comme on pouvait s'y attendre (vu qu'il n'y a physiquement pas de haut-parleurs). Logiquement le vecteur énergie ne peut pas pointer en dessous de l'équateur (voir (12)), car aucune source réelles ne se trouve en dessous. Encore une fois 3 décodeurs sur 4 ont donc leurs vecteurs énergie bien répartis sur la périphérie de l'équateur, les sources placées dans l'hémisphère sud seront donc « rabattues » sur la ceinture horizontale des haut-parleurs. Le décodeur Ambix montre une répartition plus disparate avec des remontées du vecteur énergie pour les sources placées aux alentours du pôle sud. On peut donc s'attendre à avoir un rabattement des sources placées en bas sur le haut-parleur du haut, ce qui peut s'avérer gênant.

On remarque à ce sujet une grande différence entre les décodeurs BRS et BCOM, en effet dans le décodeur BRS les vecteurs énergie et vitesse sont colinéaires et de norme assez proche, les sources situées dans l'hémisphère inférieur seront donc floues, mais assez stables. Au contraire dans le décodeur BCOM, les vecteurs énergie ont une norme proche de 1 (sauf pour les sources placées au niveau du pôle) et les vecteurs vitesse associés sont colinéaires à la direction de la source. Pour que cela soit possible, cela veut dire que les haut-parleurs situés à l'opposé de la source vont générer des signaux hors-phase, on aura donc des sources rabattues sur les haut-parleurs (car norme énergie proche de 1), mais très floues.

Ces considérations sont une interprétation possible de phénomènes empiriques, néanmoins elles demandent à être associés à un test perceptif (cf partie 3.3) pour être vérifiées psycho-acoustiquement.

## 3 ELABORATION DES TESTS SUBJECTIFS

### 3.1 Introduction

« [...] les déterminations géométriques et mathématiques dont font usage les sciences de la nature supposent l'opération subjective qui les produit et sans laquelle elles ne seraient pas : il n'y a dans la nature ni nombre ni calcul, ni addition ni soustraction, ni droite ni courbe : ce sont là des significations idéales qui trouvent leur origine absolue dans la conscience qui les crée au sens strict du mot et qu'on doit appeler à leur égard une conscience transcendantale. Si donc les idéalizations géométriques et mathématiques proviennent de la subjectivité c'est que, loin de réduire celle-ci à n'être qu'une apparence, le monde de la science trouve au contraire en elle le principe qui l'engendre continuellement comme la condition permanente de sa propre possibilité. » (Michel Henry, *La barbarie*, 1987)

La question de l'évaluation subjective dans le contenu audio est prépondérante. En effet, nous ne pouvons pas nous satisfaire d'un modèle théorique satisfaisant pour valider une technologie. La fidélité du timbre et de l'espace doit à tout prix être conservée, ou du moins être le moins dégradée possible à travers les différentes étapes de la chaîne de traitement. On remarque également dans le design des encodeurs et des décodeurs, qu'un véritable compromis est effectué entre la préservation du spectre audio (où l'on cherche à avoir une chaîne « transparente » pour le signal) et la qualité spatiale (entre la précision de localisation, l'enveloppement, etc.).

Aussi la part subjective et le domaine d'application doivent être pris en compte dans la comparaison des différents paramètres. On ne cherchera pas à mettre en avant les mêmes qualités en fonction du domaine d'application (par exemple une préservation des timbres sera préféré à un rendu plus immersif pour une application au mixage musique). Enfin, de par la structure de la chaîne de traitement, la possibilité d'un signal référence est tout simplement impossible, car la technologie HOA implique un format intermédiaire dans la synthèse du champ sonore, et il n'y pas d'équivalent possible à un format laissé intact (comme une référence non compressée pour un test subjectif soumis aux méthodes de l'ITU). Cette partie comportera donc des pistes de réflexion sur l'élaboration d'un test subjectif concernant la qualité de l'audio spatialisé en HOA, s'inspirant en partie du travail de Sarah Le Bagousse dans le cadre de sa thèse ([12]). Nous aborderons

également toute la partie en amont des tests concernant la synoptique de l'installation, ainsi que les différentes calibrations effectuées.

Enfin nous présenterons les deux tests subjectifs qui constituent la partie pratique de ce mémoire, et l'analyse des résultats.

## 3.2 Elaboration d'un protocole de test

La technologie HOA et plus largement les technologies de synthèse de front d'onde (WFS p.ex) ont déjà été le sujet de nombreux tests subjectifs, concernant notamment la localisation (cf. [13], [14]) pour différents ordres, différentes configurations de haut-parleurs et différents signaux.

Cependant, ces tests permettent l'utilisation d'une référence haute à l'aide d'un haut-parleur mono représentant la direction idéal du son voulu. Ce que nous nous proposons dans ce mémoire est de mener un test sur l'aspect qualitatif du son, et non pas mesurer la précision d'un paramètre (aspect quantitatif). La norme ITU-R BS.1116 (1997) définit une méthodologie de test de type MUSHRA pour du contenu sonore multicanal, elle juge de la fidélité de différents codecs comparativement à une référence haute (son non traité) et de différentes ancrs basses qui dépendent des attributs testés (ex : si l'on veut comparer le timbre, l'ancre basse subira en général un filtrage très destructif (passe bas à 3,5 kHz en gl)). Cette méthode permet d'avoir un ambitus entre la référence et l'ancre dans lequel vont s'intégrer les différents extraits testés. Malheureusement cette méthode n'est pas adaptée à ce que nous recherchons. En effet, cette recommandation n'intègre pas la qualité spatiale comme critère, notamment au niveau du choix de l'ancre. De plus, elle impose une référence haute, concept impossible à transposer en HOA pour comparer différents décodeurs/encodeurs.

Sarah le Bagousse a donc développé, en partant du même constat (mais pour la comparaison de contenu spatial en 5.1 et en binaural), un test de préférence inspiré du test MUSHRA, mais sans référence haute, où l'on intègre une ancre basse pour les critères recherchées, qui peut tous les combiner (donc satisfaire à la condition d'une ancre pour tous les attributs testés). Cette méthode est « *destinée à l'évaluation de qualités dites intermédiaires. En effet l'évaluation de codages haute qualité ne permet de faire de discrimination selon les attributs proposés dans ce test (NDR : dans son cas Défauts, Espace, Timbre, Qualité Globale) réalisé sans référence explicite. Il faut donc faire appel à la recommandation ITU-BS1116 pour ces degrés de qualité* ». (Source [12], p.61)

Elle définit également une ancre spatiale pour un test en binaural, validée par un test subjectif sur un panel d'experts, qui consiste en : « *La dégradation appliquée pour l'ancrage spatial consiste à créer un mouvement de balancement entre le canal droit et le canal gauche. Ceci engendre une incohérence dynamique dans la perception de l'espace ainsi que des modifications des caractéristiques spatiales telles que la profondeur, l'enveloppement...* » (Source [12], p.90)

Elle est implémentée en agissant sur la loi de pan L-R à l'aide d'une fonction trapézoïdale et de son opposé, de fréquence 0,175 Hz, créant ainsi ce mouvement de balancier entre la voix gauche et la voix droite. Sa thèse conclue par l'affirmation que : « *La méthode d'évaluation subjective de la qualité pour les sons spatialisés a été utilisée pour deux systèmes de restitution spatialisée, le 5.1 et le binaural. Dans le but d'avoir une méthode recommandée valide pour tout système spatialisé, la méthode pourra être utilisée avec d'autres techniques de spatialisation tel que Ambisonics ou Wave Field Synthesis. L'ancrage unique notamment pour la dégradation de l'Espace pourra alors être adapté.* ».

Nous nous baserons donc sur cette méthode d'évaluation pour cette partie pratique, en prenant bien soin d'adapter l'ancre spatiale dans le domaine HOA.



## 3.3 Calibration de la salle d'écoute

### 3.3.1 Description

Avant de commencer les phases préliminaires de l'élaboration d'un protocole de test, il a d'abord fallu prendre en main le studio 14.0 de l'institut de recherche BCom. Le laboratoire de recherche est intégré dans un bâtiment de bureau, aussi la pièce dans laquelle il a été installé n'a pas été traitée, ni prévue acoustiquement pour l'installation d'un système audio (en résumé c'est un bureau). Le dispositif se compose d'un cube en tube métalliques dont les parois (sauf le sol et le plafond) ont été recouvertes de mousse acoustique. Un premier cercle de haut-parleurs a été fixé à 1m35 de hauteurs environ et un deuxième cercle a été fixé à 2m10 de hauteur. Ces cercles servent à la fixation de chacune des 14 enceintes.



*Studio 14.0 de Pleumeur-Bodou*

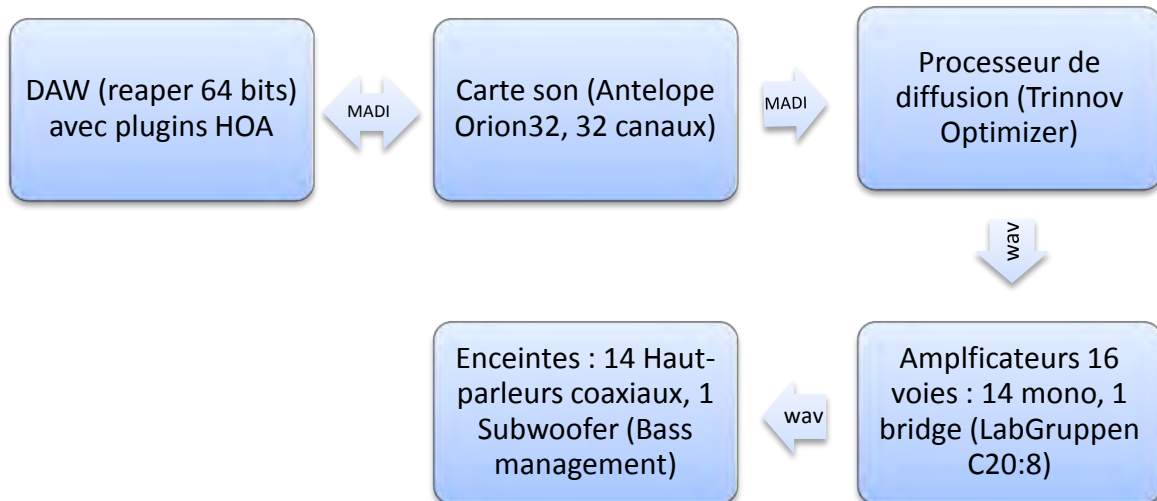
La partie logicielle se fait sur un iMac 25 pouces, sur un Windows 7 installé via une machine virtuelle VMWare Fusion, relié à un rack 10U qui comporte la carte son Antelope Orion32, le processeur de diffusion Trinnov Optimizer et les deux amplificateurs LabGruppen 20:8.



*Photo du rack audio du studio 14.0 (du haut vers le bas : Carte son, processeur, amplificateurs)*

Il a été très vite décidé de déporter ce rack dans la pièce d'à côté afin de se passer du bruit qu'il amenait. Il s'en est suivi une brève remise à neuf du câblage afin de pouvoir faire passer les différents modules par le faux plafond. Ceci a également nécessité l'installation d'un VNC (pilotage à distance) pour le processeur afin d'éviter les allers-retours superflus.

Une synoptique du matériel est présentée :



*Schéma fonctionnel de l'installation du studio 14.0 de Pleumeur*

### 3.3.2 Ecoutes

Un des premiers enjeux a été de se familiariser avec cette salle un peu particulière comparativement à un studio d'écoute professionnel, où le traitement acoustique et l'intégration des enceintes ont été pensés en amont.

Ici l'installation a été pensée sous forme modulaire, afin de pouvoir être montée et démontée en une journée. Ce qui représente un avantage ergonomique certain, mais peut poser quelques problèmes gênants pour une écoute critique. Une première solution proposée pour améliorer l'installation est le processeur de diffusion Trinnov Optimizer, qui intègre un mode Optimization (d'où son nom) comprenant : un alignement des retards, un alignement des niveaux, ainsi que de la phase. Il comprend également une compensation de l'effet de salle, mais le manuel ne précise pas son fonctionnement. Il est également possible de simuler des haut-parleurs virtuels, mais nous ne nous sommes pas servis de cette fonction. Nous reviendrons plus tard sur le fonctionnement de cette calibration. Ce processeur avait déjà été configuré avant mon arrivée, j'ai donc utilisé les presets existants pour une première écoute.

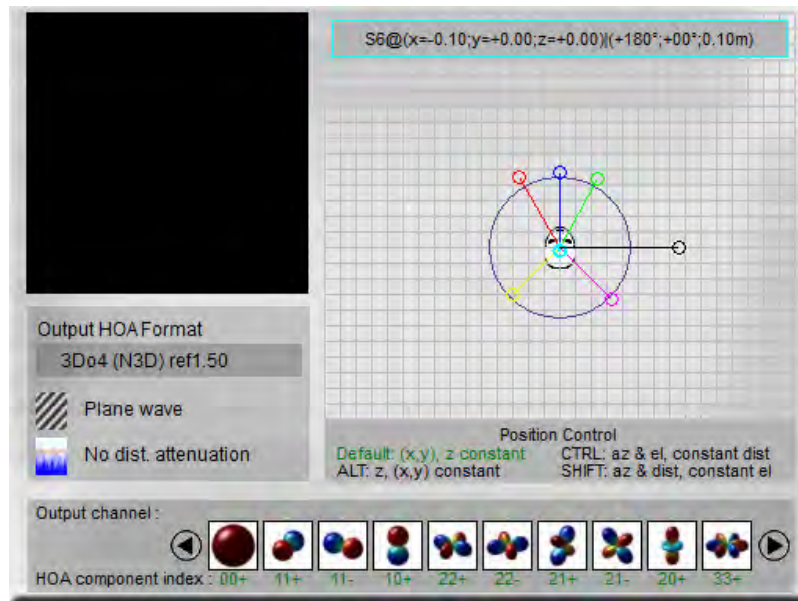
Peu familier avec l'écoute HOA, et ne disposant ni de contenu 14.0 ni Ambisonics, j'ai opté pour une première écoute comparative avec des extraits de différents CD (en stéréo) et 5.1 (DVD Audio), entre une écoute classique (1 canal par enceinte) et une écoute HOA (où chaque canal est encodé puis décodé par HOA, il est donc possible de manipuler les positions des enceintes virtuelles correspondantes). Cette écoute comparative a eu un intérêt double : découvrir les enceintes Amadeus et pouvoir me faire une première idée de la coloration amenée par HOA.

Le panel de morceaux utilisés pour cette écoute est listé en dessous (signaux audio non compressés, 44,1kHz/16 bits pour les CD, 48kHz/24 bits pour le DVD Audio) :

- John Zorn – Meow Meow (The Dreamers) Jazz – Easy Listening (stéréo)
- Beach Boys – Caroline, No (Pet Sounds) Pop 60's (mono)
- My Bloody Valentine – Soon (Loveless) Noise rock (stéréo)
- Andy Stott – Passed me By (Passed me By) Electro-dub (stéréo)
- Bjork – t's not up to you (Vespertine) Pop 00's (5.1)

Une première écoute a révélé deux problèmes inhérents à l'installation : premièrement, sur une écoute à fort volume avec des morceaux chargés en basse fréquence (notamment Passed me by avec beaucoup d'informations dans l'extrême grave), les lampes du plafonnier vibraient par sympathie, provoquant une vibration métallique assez désagréable ; ensuite les tubes constituant la structure métallique étaient également excités pour des signaux à transitoires très prononcés (une écoute de bursts de bruits roses à fort volume a confirmé ce constat) : une vibration tonale pouvait être entendue comme si l'on frappait sur le tube à l'aide d'un petit objet contondant.

Les enceintes Amadeus PMX4 sont quant à elles assez droites, même si elles ne descendent pas très bas. L'espace sonore était bien respecté, et elles apportent une très bonne définition aux morceaux (notamment sur des mixes comportant beaucoup de relief, comme le morceau des Beach Boys ou de My Bloody Valentine).



*Encodeur HOA en configuration stéréo/ 5.1 (L Rouge, R Vert, C Bleu, Ls Jaune, Rs Rose, LFE Cyan)*

La comparaison entre une écoute stéréo classique et reconstruite en HOA est assez surprenante : on se retrouve en effet à écouter l'image stéréophonique entre 2 enceintes fantômes, ce qui peut être perturbant de prime abord, la position physique de l'enceinte en écoute stéréophonique classique demeurant un ancrage visuel appuyant la perception spatiale. Néanmoins, sur les différents extraits, j'ai constaté une coloration assez medium, les hautes fréquences sont moins bien définies (l'aliasing spatial présent en hautes fréquences en HOA peut en être la cause) comparativement à l'écoute standard. L'espace sonore est également légèrement modifié, l'écoute en 5.1 le révèle particulièrement : on perd un peu en profondeur de mix. Néanmoins il faut nuancer ce propos en rappelant que chaque enceinte contribue à la restitution d'une paire stéréophonique virtuelle, aussi, sans compensation, les enceintes vont se sommer dans le domaine basse fréquence et donner une coloration plus ronde, défaut que l'on retrouve également en musique dans l'upmix d'un morceau stéréo vers un 5.1 sans rééquilibrage de la balance tonale.

Cette piste de la coloration amenée par HOA est donc une bonne piste à creuser car elle pourrait être réglée par un jeu de filtrage validé par un test perceptif.

Passé cette première écoute, j'ai également voulu confronter l'aspect localisation entre les positions respectives de chaque enceinte (rappelé dans le tableau ci-dessous) et simuler grâce à HOA (à l'ordre 4) une enceinte virtuelle qui se déplacerait pour chaque position. L'idée est donc de diffuser un signal mono pour chaque enceinte (encapsulé dans un fichier wav 14 canaux) et de le confronter à ce même signal mono spatialisé à la position théorique de l'enceinte.

J'ai donc créé plusieurs signaux test :

- un décompte de 1 à 14 par une voix de synthèse masculine
- un décompte de 1 à 14 par une voix de synthèse féminine
- un sinus de fréquence 1000 Hz répartis en 14 courtes impulsions
- un bruit rose répartis en 14 courtes impulsions

N°	Azimut ( $\theta$ )	Élévation ( $\delta$ )	Hauteur (m)
1	0	0	1.35
2	45	0	1.35
3	90	0	1.35
4	135	0	1.35
5	180	0	1.35
6	-135	0	1.35
7	-90	0	1.35
8	-45	0	1.35
9	0	35	2.10
10	45	35	2.10
11	135	35	2.10
12	-135	35	2.10
13	-45	35	2.10
14	0	90	2.10
(SUB)	0	0	0

*Tableau récapitulatif de la position théorique des enceintes*

Cette deuxième écoute a confirmé ce qu'avait esquissé la première écoute : on a des modifications de timbre, donnant une couleur plus ronde aux signaux écoutés (décelables sur la voix notamment) qui apparaissent dans des zones où toutes les enceintes sont particulièrement sollicitées (au centre par exemple). D'autres choses sont apparues : la précision de localisation pour le signal HOA était variable en fonction du signal utilisé : elle était excellente pour les signaux vocaux, bonne pour le bruit rose, et moyenne pour le 1000 Hz.

Cette deuxième écoute a amené une idée en écoutant les impulsions de bruit rose : pourquoi ne pas créer des trajectoires continues de bruit rose dans l'espace ?

Cette idée s'est avérée fructueuse car elle mettait en exergue plusieurs caractéristiques inhérentes à notre système d'écoute : la gestion de la spatialisation dans les espaces entre chaque haut-parleur, l'absence de haut-parleurs dans l'hémisphère inférieur, et de manière plus générale des informations sur le traitement de l'information spatiale par la méthode de décodage HOA. Nous nous sommes contentés de synthétiser des trajectoires simples (voir annexe) : des cercles et des spirales, se déplaçant selon une coordonnée constante  $(x,y,z)$  et comparer différents décodeurs à l'écoute. Celle-ci a révélé des disparités suffisamment intéressantes pour que nous décidions d'en faire l'objet d'un test plus approfondi, à mettre en lien avec la simulation des trajectoires théoriques entendues.

Nous avons également jugé intéressants d'aborder une deuxième piste concernant l'impact sur la qualité spatiale (enveloppement, distance à la source, ...) et le spectre audio des différents décodeurs. Aussi, tester ceux-ci à l'aide d'enregistrements réalisés avec le microphone Eigenmike, qui devront présenter un panel assez large (voix, musique, ambiances) pour juger au mieux des modifications apportées par chaque décodeur.

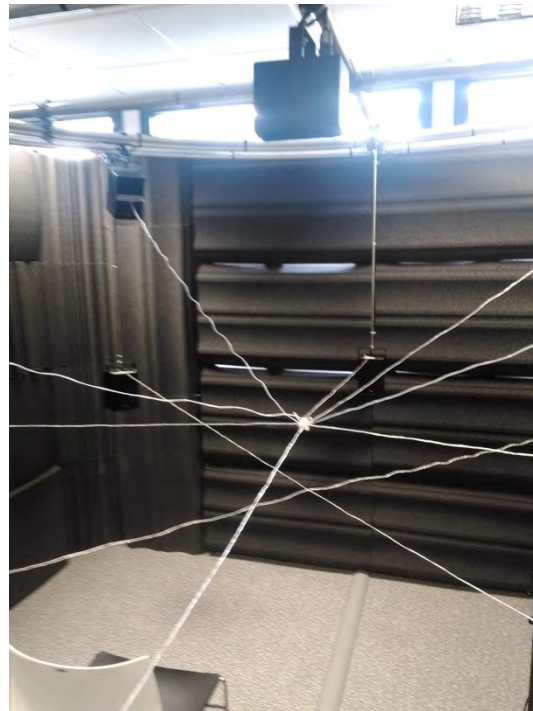
### 3.3.3 Calibration et mesures

Après les écoutes test, et malgré une première calibration réalisée par TMS sur la salle d'écoute, il nous a semblé utile de repasser en vue la chaîne audio et de refaire quelques vérifications et modifications si besoin. La première étape a été faite en réaction avec les vibrations gênantes du plafonnier : nous avons placé des rectangles de mousse pour découpler les lampes du faux plafond, empêchant ainsi la transmission solidienne. Le résultat a été probant car les vibrations ont été fortement diminuées. Nous avons également découplé l'enceinte de subwoofer du sol à l'aide d'un tapis anti-vibration, pour empêcher la transmission des basses fréquences par le plancher.

Aucune solution simple ne nous est venue pour les résonances des tubes métalliques, seul un projet avorté d'insérer de la mousse à l'intérieur des tubes a été évoquée, mais laissé sans suite. Ce problème nous a paru négligeable au vu des extraits proposés dans nos tests. En effet, les tubes ne résonnaient pas à un volume raisonnable, ce problème a donc été mis de côté.

Ensuite, il a fallu mesurer la distance de chaque enceinte par rapport au centre de la pièce afin de valider le point d'écoute pour un sujet test (être au centre du dispositif garantit la meilleure écoute possible). Cette tâche a été réalisée avec une

pelote de ficelle, en découpant différents bouts de même longueur, chacun scotché au centre de l'enceinte et tendu en direction de l'enceinte opposée. Cette première approximation nous a permis de voir les petites disparités entre la distance par rapport au centre et de les corriger, et de fixer un point central pour y placer le micro de mesure.



*Etablissement du point de mesure*

La prochaine étape était la configuration de l'Optimizer Trinnov, qui assure l'optimisation du système via une égalisation des niveaux, du spectre fréquentiel et une compensation de l'effet de salle. Trinnov ne fournit pas les méthodes de calcul de leur processeur, notamment pour l'effet de salle, aussi nous ne pourrions pas fournir un point de vue théorique sur les traitements mis en œuvre, mais seulement les résultats qu'il fournit.

La calibration se fait à l'aide d'une sonde Trinnov, soit un micro à 4 capsules, chacune placée aux quatre sommets d'un tétraèdre régulier.





*Sonde Trinnov*

Ainsi l'optimisation se calcule au centre du point d'écoute qu'on aura pris soin de déterminer au préalable. La mesure consiste en un burst de bruit rose envoyé successivement aux 14 enceintes, la comparaison entre le bruit rose et sa version captée par la sonde est faite directement dans l'optimizer. Une fois la mesure faite, on peut choisir d'appliquer le calcul à cette mesure afin de calculer les différents paramètres nécessaires aux différents traitements. Il est également possible de réaliser plusieurs mesures à la fois, puis d'accorder une valeur de pondération à chacune lors du calcul (aussi peut-on réaliser la mesure plusieurs fois aux mêmes points pour assurer un calcul fiable).

La mesure nous renvoie donc un tableau comprenant plusieurs informations relatives à chaque enceinte : position, angle d'azimut et d'élévation, délai de groupe, etc. (cf. Image). On peut ainsi corriger avec plus de finesse la position des enceintes, en s'assurant que la sonde est bien dirigée (un mauvais placement induit que les positions seront calculées à partir d'un point différent du centre, les valeurs ne correspondront donc pas à ce que l'on cherche).

-30.0 dB 10: 150415 LabGruppen TMS 0.5 reset presetEQ Test															
Optimizer Settings															
Top View															
Speaker	'1'	'2'	'3'	'4'	'5'	'6'	'7'	'8'	'9'	'10'	'11'	'12'	'13'	'14'	
Runtime	Distance (m)	2.39	2.40	2.35	2.42	2.37	2.49	2.42	2.49	2.37	2.40	2.42	2.43	2.42	1.60
	Elevation (°)	1.5	-0.5	-0.2	-2.9	-2.6	-3.2	-0.8	-0.2	35.2	32.9	29.1	27.9	32.2	87.4
	Azimuth (°)	6.3	51.3	98.7	141.5	-170.8	-128.2	-82.9	-37.5	5.1	52.5	141.3	-129.0	-38.4	-168.3
Settings	Level A (dBFS)	-51.3	-51.3	-52.4	-51.6	-51.2	-52.6	-51.6	-52.5	-51.8	-52.4	-51.6	-52.2	-52.6	-45.5
	Level C (dB SPL)	-25.4	-24.9	-25.7	-25.0	-24.6	-25.6	-25.2	-24.6	-25.4	-25.4	-24.9	-25.8	-25.1	-19.6
	Level A compensated (dB SPL)	-25.6	-25.3	-25.1	-25.4	-25.2	-25.4	-25.4	-24.7	-25.3	-25.3	-25.1	-25.4	-25.2	-25.8
	Level C compensated (dB SPL)	-24.6	-24.7	-24.0	-24.5	-24.3	-24.6	-24.6	-23.8	-24.4	-24.6	-24.6	-24.6	-24.4	-24.9
Positions	Delay (ms)	6.98	7.00	6.63	7.06	6.90	7.29	7.04	7.23	6.90	6.98	7.08	7.08	7.02	4.89
	EM Delay (ms)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Delay compensation (ms)	0.31	0.25	0.46	0.23	0.40	0.00	0.25	0.06	0.40	0.31	0.21	0.21	0.27	2.60
Calibration	Crest factor (dB)	34.7	34.2	34.4	34.3	34.4	34.3	34.2	33.7	34.1	33.9	34.6	34.3	34.2	34.2
	8dB Bandwidth (Hz)	101 - 23.3k	135 - 24k	103 - 24k	113 - 21.5k	106 - 23.2k	99.1 - 24k	111 - 21.7k	133 - 21.5k	104 - 22.9k	144 - 22k	151 - 24k	148 - 21.8k	144 - 22.7k	113 - 22.9k

Tableau de valeurs rendues par le Trinnov une fois la mesure effectuée

Une fois ce calcul effectué, on accède aux différents traitements (alignement du niveau, phase, délai, compensation effet de salle) qu'on peut à loisir désactiver ou activer. A chaque fois, Trinnov fournit, via son interface, la courbe de réponse en fréquence/phase mesurée pour chaque enceinte, le filtre inverse appliquée et la courbe de fréquence résultante. En plus de ces différents traitements automatisés, on peut accéder à la gestion du bass management (choisir la fréquence du crossover via un filtrage actif), le routing entrée/sortie, ainsi que d'autres fonctionnalités (pour plus d'info se référer à la notice).

Les résultats issus de cette optimisation sont présentés sous forme de graphique à la page suivante. Les 15 voies sont présentes (14 enceintes + 1 subwoofer), la fréquence du filtre cross over a été réglée sur 150 Hz, et un léger boost a été opéré à cette fréquence sur tous les canaux à l'aide d'un égaliseur graphique. Pour une lecture plus aisée, on notera que la voie du subwoofer est représentée par la couleur rouge.

Enfin une mesure du bruit de fond au point d'écoute a été réalisée à l'aide d'un sonomètre avec une pondération en dBA. Une VMC est présente dans la pièce,

nous avons donc effectué deux mesures (avec et sans VMC) afin de décider si elle pouvait s'avérer gênante à l'écoute.

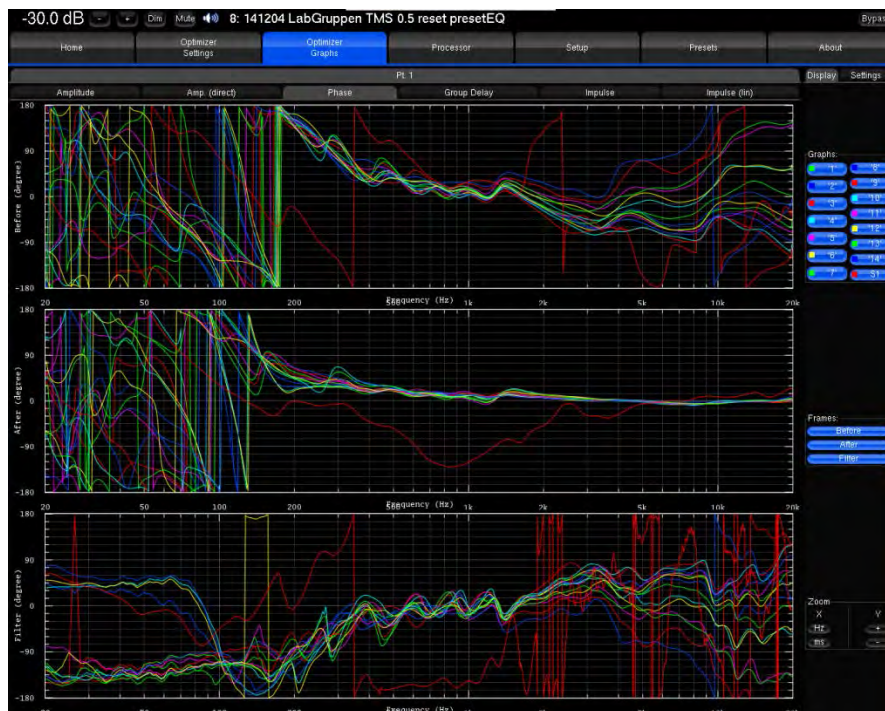
Les valeurs mesurées sont :

- Avec VMC : 30.5 dBA
- Sans VMC : 23,2 dBA

L'écart entre les deux valeurs est donc significatif, il a donc été décidé de réaliser tous les tests en coupant la VMC afin d'avoir un niveau de bruit de fond tout à fait raisonnable au vu du niveau de diffusion des tests. Nous avons fixé le niveau moyen d'écoute à environ 70 dB à l'aide d'un bruit rose décorrélé entre tous les canaux.



*Egalisation effectuée par l'Optimizer (les 15 courbes d'entrée sont en haut, au milieu les courbes de sortie, en bas les filtres appliqués)*



*Correction de phase effectuée par l'Optimizer (les 15 courbes d'entrée sont en haut, au milieu les courbes de sortie, en bas les filtres appliqués)*

### 3.4 Elaboration du premier test

Une fois la calibration effectuée, il a fallu se pencher sur l'établissement d'un protocole de test subjectif. Comme évoqué dans la partie 3, la technologie HOA comporte quelques particularités l'empêchant de rentrer dans une recommandation ITU pour un protocole de test. Aussi nous nous sommes inspirés du travail de S. Le Bagousse pour établir un test préliminaire.

La première étape a été de choisir un nombre réduit de décodeurs parmi ceux existants, et parmi ce nombre réduit de décodeurs choisir un paramétrage particulier. Les conditions de choix ont été : la possibilité d'un décodage à l'ordre 3 minimum, un rendu satisfaisant, la possibilité de configurer le décodeur avec notre configuration de haut-parleurs. Nous présentons donc les quatre décodeurs choisis, en rappelant tout de même qu'il existe d'autres solutions de décodage existantes.

**Décodeur Technicolor** (cf [15]) : la solution de décodage de Technicolor est celle comprise dans le standard MPEG-H, elle repose sur le calcul de la matrice pseudo-inverse, couplé à un algorithme issu du VBAP.

**Décodeur Blue Ripple Sound** : la solution de décodage de BRS est une solution payante et propriétaire, sa version commercialisée va jusqu'à l'ordre 3. Nous avons choisi le preset Basic après discussion avec son créateur Richard Furse. La méthode de décodage repose sur la maximisation du vecteur énergie lors du calcul de la matrice de gain.

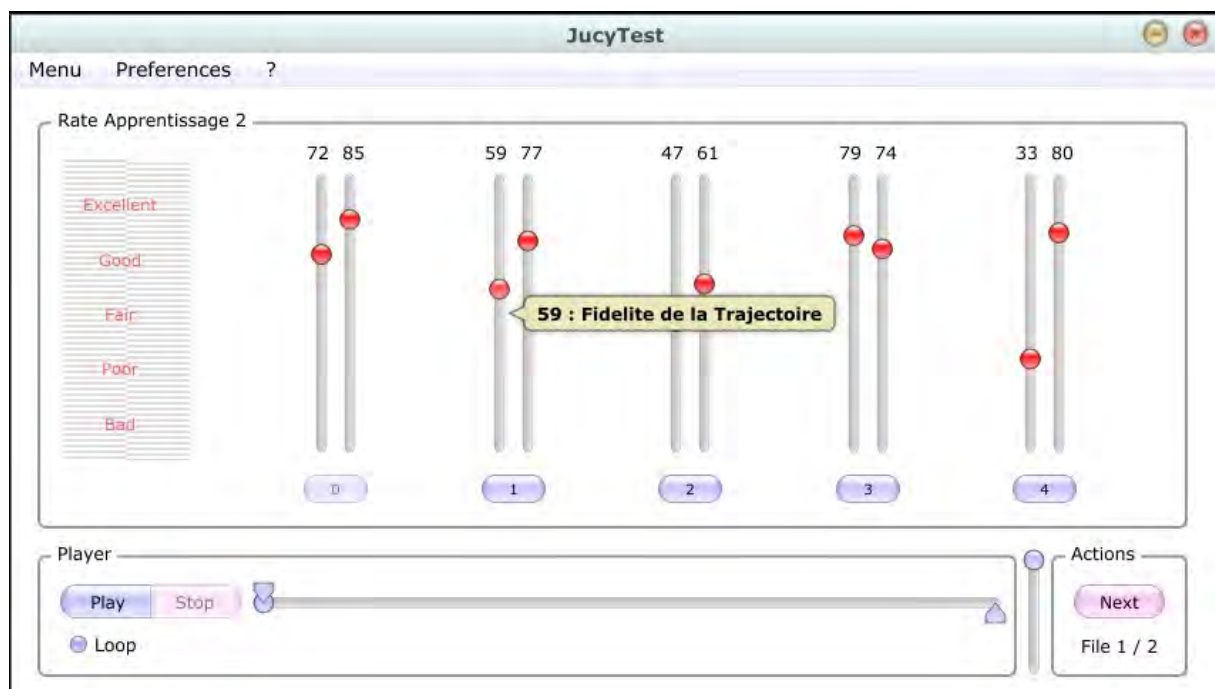
**Décodeur Ambix** : cette solution de décodage HOA fait partie de la suite open source Ambix développée par Mathias Kronlachner, qui permet plusieurs traitements, l'encodage et le décodage en HOA jusqu'à l'ordre 7. La méthode de décodage choisie (P-INV) repose sur le calcul simple de la matrice pseudo-inverse, cf [16]). Ambix propose également deux autres solutions de décodage (Allrad et SSF) qui n'ont pas été retenues pour ce test, car les résultats étaient incohérents. Ce n'est pas la méthode la plus optimisée fournie par Ambix, et représente la solution la plus « pauvre » des 4 décodeurs.

**Décodeur BCom** : cette solution de décodage est le fruit des travaux réalisés chez Orange Labs et poursuivi par l'institut de recherche BCom. Elle repose sur le calcul de la matrice pseudo-inverse, puis sur l'optimisation de cette matrice via un paramètre reposant sur l'optimisation du vecteur énergie, du vecteur vitesse, ainsi que sur les paramètres  $R_v$  et  $R_e$ .

Une fois ces décodeurs choisis, il a fallu se pencher sur une interface de test correspondant à nos besoins, à savoir :

- lire des fichiers multicanaux (ici 14 canaux)
- la gestion d'un pilote ASIO pour communiquer avec la carte son
- évaluer plusieurs paramètres en même temps (ici 2)
- avoir une interface de type MUSHRA, mais sans référence.

Notre choix s'est porté sur le logiciel JucyTest (fourni par Orange Labs), qui est une interface modulaire adaptée à des tests subjectifs. Le logiciel présente ainsi une suite de paramètres réglables concernant les différents attributs, leur notation, l'ordre de passage des fichiers audio. Une fois tous les paramètres réglables, on compile l'interface graphique, qui nous renvoie un exécutable contenant une coquille de test, dans lequel on vient créer une session test (c'est un fichier texte .ses), ajouter les fichiers audio en .wav, et configurer les périphériques audio.



*Interface de test JucyTest (5 extraits et 2 attributs à noter)*

Nous avons tous les outils à notre disposition, définissons maintenant le contenu du test.

Ce premier test se définit comme un test de préférence entre 4 décodeurs et une ancre spatiale. Ils seront testés à l'aide de trajectoires de bruit rose dans l'espace. Nous pensions d'abord proposer un report graphique au sujet, mais nous sommes rapidement rendu compte qu'il nous a été impossible de réaliser une interface dans le temps imparti pour que l'utilisateur puisse facilement prendre en main un dessin en 3D et que les données soient retranscrites pour être traitées. Nous avons donc opté pour une méthode différente : nous proposons le dessin des différentes trajectoires sur papier au sujet, et celui-ci doit juger à partir de ce qu'il entend si la trajectoire lui semble correspondre à ce qui lui est présenté.

On a ainsi défini un premier attribut à noter : **la fidélité à la trajectoire**.

Un deuxième attribut a été défini : **la constance du timbre**. Derrière ce terme nous avons voulu juger de la manière dont la couleur du son évolue au cours du temps (donc au cours de la trajectoire). En effet, nous avons observé à l'écoute que les différents décodeurs avaient des comportements différents dans les zones de vide (sans haut-parleur) se traduisant par des modifications de timbre, ou de perte de niveau. Nous avons voulu rendre compte de ces changements à travers cet attribut.

Une phase d'entraînement a également été introduite dans le test afin que le sujet puisse se familiariser avec l'interface et la perception de trajectoires dans l'espace, qui peut être déroutante au premier abord.

Les trajectoires utilisées pour le test sont présentées dans la feuille de test présente en annexe 3. Elles ont toutes une durée fixe de 7 secondes exactement.

La dernière étape a été de créer l'ancre spatiale sur les recommandations de la thèse de S. Le Bagousse suivant la méthodologie exposée ci-après.

Nous sommes partis de chaque trajectoire ; elles ont pour point commun d'avoir un mouvement de rotation autour d'un axe (une des trois coordonnées cartésiennes reste donc constante). Aussi, pour proposer une version du mouvement d'aller et retour préconisé pour une ancre spatiale, nous avons remplacé la coordonnée constante par un terme d'erreur oscillant entre 2 valeurs, donnant ainsi une impression de va-et-vient autour de la valeur fixée. De plus cette trajectoire

dégradée est décodée avec un décodeur à l'ordre 1, afin de flouter la trajectoire comparativement aux autres décodeurs d'ordre 3.

Exemple : pour la trajectoire effectuant une rotation autour de l'axe  $Ox$ , avec  $x = 0$ , nous avons remplacé cette valeur par  $x = 0.5 * \sin(5 * \text{pitch})$  ;  $\text{pitch}$  représentant l'angle parcouru par les deux autres coordonnées pendant le temps de l'extrait audio (i.e 7s). On a donc remplacé la coordonnée fixe par un sinus oscillant 5 fois entre  $-0,5$  et  $0,5$  au cours du temps. Cette valeur  $0,5$  a été choisie à l'écoute parmi d'autres valeurs, elle nous a semblée représenter une valeur adéquate, ni trop dégradante, mais juste assez pour que la fidélité de la trajectoire et la constance du timbre soit suffisamment altérées au regard des autres extraits.

L'analyse des résultats validera ou non cette méthodologie.

Toutes les trajectoires ont été encodées directement en HOA et décodées à l'ordre 3 (sauf l'ancre spatiale décodée à l'ordre 1).



## 4 ANALYSE DES RESULTATS

Cette partie sera malheureusement écourtée, nous avons en effet fait passer à 10 sujets « experts » le premier test décrit dans la partie 3.3 du 11 au 19 mai. Il ne nous a donc pas été possible de mettre en forme et d'analyser les résultats obtenus, pour cause de rédaction de ce présent mémoire.

Néanmoins, suite à ce mémoire, les données seront traitées et abordés lors de la soutenance.

## 5 CONCLUSION

Le présent mémoire aura eu pour vocation de montrer à un lecteur peu familier avec le formalisme HOA, les nombreuses possibilités offertes par ce format appliqué au domaine des techniques audio. Cette technologie en amenant un format intermédiaire, placé entre l'étape d'encodage du champ sonore et de la restitution sur un système de haut-parleurs innove. Ces deux étapes d'ordinaire intrinsèquement liées (un fichier stéréo est fait pour être lu sur une paire stéréo) peuvent être désormais décorréliées. C'est cette flexibilité qui demeure selon nous le grand atout d'HOA comparativement à d'autres techniques d'audio spatialisé. Il permet ainsi de s'adapter aux contraintes de transport, notamment du dispositif de restitution à l'arrivée, amener un format normalisé pour le codage de voie (MPEG-H propose une standardisation depuis l'été dernier). Il peut utiliser n'importe quelle configuration de haut-parleurs existante et même future (avec quelques contraintes évoqués dans ce mémoire), il est cependant certain qu'on pourra décoder un fichier HOA sur n'importe quel système Dolby à venir.

L'ordre de l'encodage du fichier HOA amène une indication sur le degré de flexibilité permis, un fichier encodé à l'ordre 1 aura une résolution spatiale plus faible qu'un fichier encodé à l'ordre 5, et sera plus difficile à diffuser dans de bonnes conditions sur un système donné. Néanmoins un fichier HOA pourra être toujours être décodé sur n'importe quel configuration de haut-parleurs munie de sa matrice de décodage associé.

Quel autre format audio permet une telle souplesse ?

Pour nuancer ce propos, on rappelle quand même que certains paramètres viennent influencer sur la qualité de rendu :

- Le paramétrage de l'encodage microphonique
- Les conventions d'encodage
- Les différentes méthodes de calcul des matrices de décodage
- Les contraintes des systèmes de diffusion (mais indépendant du fichier HOA)

Il sera donc intéressant de pouvoir quantifier avec finesse l'impact de chacun de ces paramètres sur le rendu audio. Mais le développement de cette technologie audio 3D est assez récent (début des années 2000), c'est pourquoi les outils capables d'évaluer l'impact des différents paramètres de la chaîne de traitement sur la qualité audio et spatiale sont encore expérimentaux, il a donc été difficile dans ce mémoire d'amener un point de vue tranché sur les influences des différents paramètres sur

ces critères qualitatifs. Nous souhaitons néanmoins avoir amené des pistes de réflexion pour leur développement.

## 6 LEXIQUE

*A-Format* : Format de fichier 4 canaux où chaque canal correspond à une capsule cardioïde situé sur l'un des 4 côtés d'un tétraèdre. Ce format peut être converti en B-Format pour être traité comme de l'Ambisonics.

*B-Format* : Format de fichier Ambisonics 4 canaux, où chaque canal correspond à une des 4 composantes des harmoniques sphériques à l'ordre 0 et 1 (W, X, Y, Z).

*Ambisonics* : Décomposition du champ sonore en harmoniques sphériques à l'ordre 1

*HOA* : Décomposition du champ sonore en harmoniques sphériques aux ordres supérieurs à 1

*HF/BF* : Hautes fréquences / Basses fréquences

*Ordre de directivité* : La formule de la directivité d'un microphone d'ordre  $n$  est donné par :  $(\alpha \cos(\theta) + (1 - \alpha))^n$ , où  $\alpha \in \mathbb{N}$ . Par exemple un microphone cardioïde classique, est un microphone d'ordre de directivité 1, avec  $\alpha = 1/2$

*MUSHRA (Multiple Signals with Hidden Reference)* : Méthode de test recommandé par l'ITU pour comparer différents codeurs, à l'aide d'une référence haute (le signal source) et des ancres basses (signal fortement dégradé)

*SNR (Signal to Noise Ratio)* : Rapport signal sur bruit, il nous renseigne sur la dynamique d'amplitude d'un équipement audio. Il sert à quantifier le seuil des niveaux faibles enregistrables (avant d'être noyé dans le bruit), et la dynamique de l'équipement.

## Bibliographie :

- [0] M.Gerzon “*What’s wrong with quadrasonics?*”, source web : [http://audiosignal.co.uk/Resources/What\\_is\\_wrong\\_with\\_quadrasonics\\_A4.pdf](http://audiosignal.co.uk/Resources/What_is_wrong_with_quadrasonics_A4.pdf)
- [1] J. Daniel, R.Nicol, S.Moreau : “*Further Investigations of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging*”, AES 114<sup>th</sup> convention 2013
- [2] J.Daniel “*Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un context multimedia*”, Thèse Université Paris 6, 2001
- [3] M.Gerzon « *Periphony : With-Height Sound Reproduction* » JAES, 1973
- [4] P.Lecomte “*Synthèse de champ sonore, par ambisonie adaptative d’ordres élevés (ANFC-HOA)* », Mémoire Master ATIAM, 2013
- [5] S. Spors and J. Ahrens, “*A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling*,” AES 125<sup>th</sup> convention.,2008
- [6] G Dickins “*Towards Optimal Sound Field Representation*”. AES 106<sup>th</sup> Convention,1999.
- [7] S Moreau “*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 à l’école doctorale du Maine, 2003
- [8] C-A Engelbert « *Réallisation d’une réverbération à convolution et mise en place de méthodes de mixage pour la technique de spatialisation sonore HOA* », Mémoire fin d’étude FSMS CNSMDP, 2013
- [9] O.Rosset «*Optimisation d’un système microphonique 2D et pratique avancée du mixage pour la technologie de spatialisation sonore HOA* » Mémoire fin d’étude FSMS CNSMDP,2012
- [10] G.Elko, J ;Meyer Brevets industriels déposés disponibles sur <http://bases-brevets.inpi.fr/fr/resultats-recherche-simple/1431194712403/result.html>
- [11] G.Elko, A.T Nguyen Pong « *A steerable and variable first-order differential microphone array*”, Acoustics, Speech, and Signal Processing, 1997. ICASSP-97, p223-226
- [12] S.Le Bagousse “*Elaboration d’une méthode de test pour l’évaluation subjective des sons spatialisés*», Thèse UBO (Université de Bretagne Occidentale), 2014
- [13] S. S. Bertet « *Format audio 3D hiérarchiques : caractérisation objective et perceptive des systèmes Ambisonics d’ordres supérieurs* », Thèse INSA Lyon, 2006
- [14] P.Power, C. Dunn, B. Davies, J. Hirst “*Localisation of Elevated Sources in HOA*” BBC R&D White paper,WHP261, 2013
- [15] J.Herre, J.Hilpet, A. Kuntz, J.Plogsties “*MPEG-H Audio : The new standard for Universal Spatial/ 3D Audio coding*”, AES 137<sup>th</sup> convention, Fraunhofer IIS, 2014
- [16] A.Heller, E. Benjamin “*The Ambisonic Decoder Toolbox: Extensions for Partial-Coverage Loudspeaker Arrays*”, Linux Audio Convention, 2014
- [17] J.Beckamn “*Miniature Microphone Arrays for Multi-channel Recording*”, AES 120<sup>th</sup> Convention, 2006
- [18] Tikhonov A.N, ARSENIN V.A “*Solution of Ill-posed problems*”, Washington: Winston & Sons, 1977
- [19] A.Solvang “*Spectral impairment of two-dimensional higher order Ambisonics*” JAES, vol 56, n°4,p 267-279,

[20] S.Spoors “*Spatial sound with loudspeakers and Its perception : a Review of the current state*” IEEE Vol. 101, No. 9, September 2013

[21] N.Epain, F.Zotter “*Ambisonic decoding with constant angular spread*”, ACTA, Vol. 100 (2014) p928 – 936

[22] M.Gerzon “*Metatheory of audio localization*” AES Convention 92th, March 24-27, Vienna, 1992

## Traitements & Visualisation

NB : les sites web ont été consultés pour la dernière fois le 26/05/15

Manuel Harpex-B : <http://www.harpex.net/manual.pdf>

Manuel Surround Zone 2 : <http://www.tslproducts.com/soundfield/soundfield-surroundzone2/>

Plugins Blue Ripple Sound: <http://www.blueripplesound.com/products/>

Plugins Ambix : <http://www.matthiaskronlachner.com/?p=2015>

Papiers développant les concepts mathématiques derrière les plugins ambix :

<https://depositonce.tu-berlin.de/bitstream/11303/169/1/11.pdf>

[http://iaem.at/Members/zotter/publications/2014\\_KronlachnerZotter\\_AmbiTransformationEnhancement\\_ICSA.pdf](http://iaem.at/Members/zotter/publications/2014_KronlachnerZotter_AmbiTransformationEnhancement_ICSA.pdf)

Site web :

[A] <http://www.tapages.fr/spip.php?article508>

[B] <http://www.soundonsound.com/sos/Oct01/articles/surroundsound3.asp>

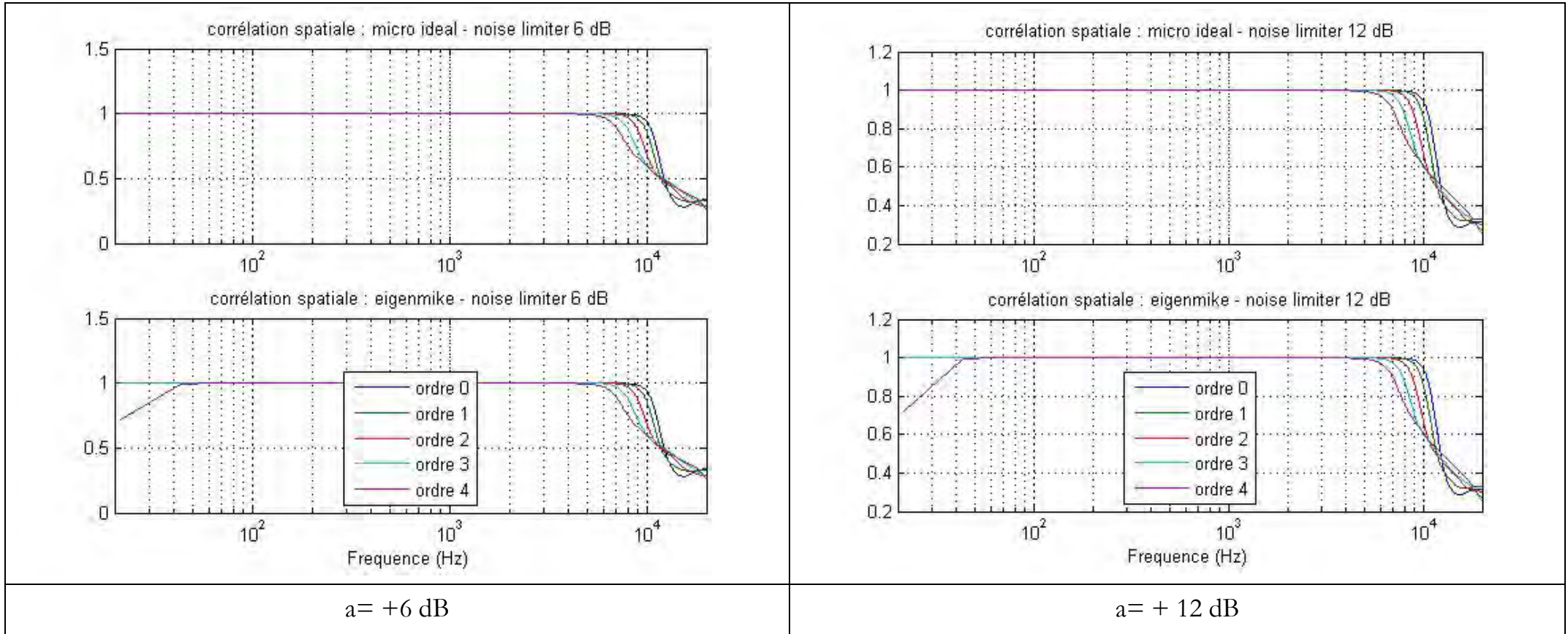
[C] <http://www.surrounddiscography.com/uhjdisc/uhjhtm.htm>

[D] <http://www.matthiaskronlachner.com/?p=2015>

[E] <http://www.surround-library.com/> (Sonothèque de son surround (5.1, B-Format, double MS, IRT...))

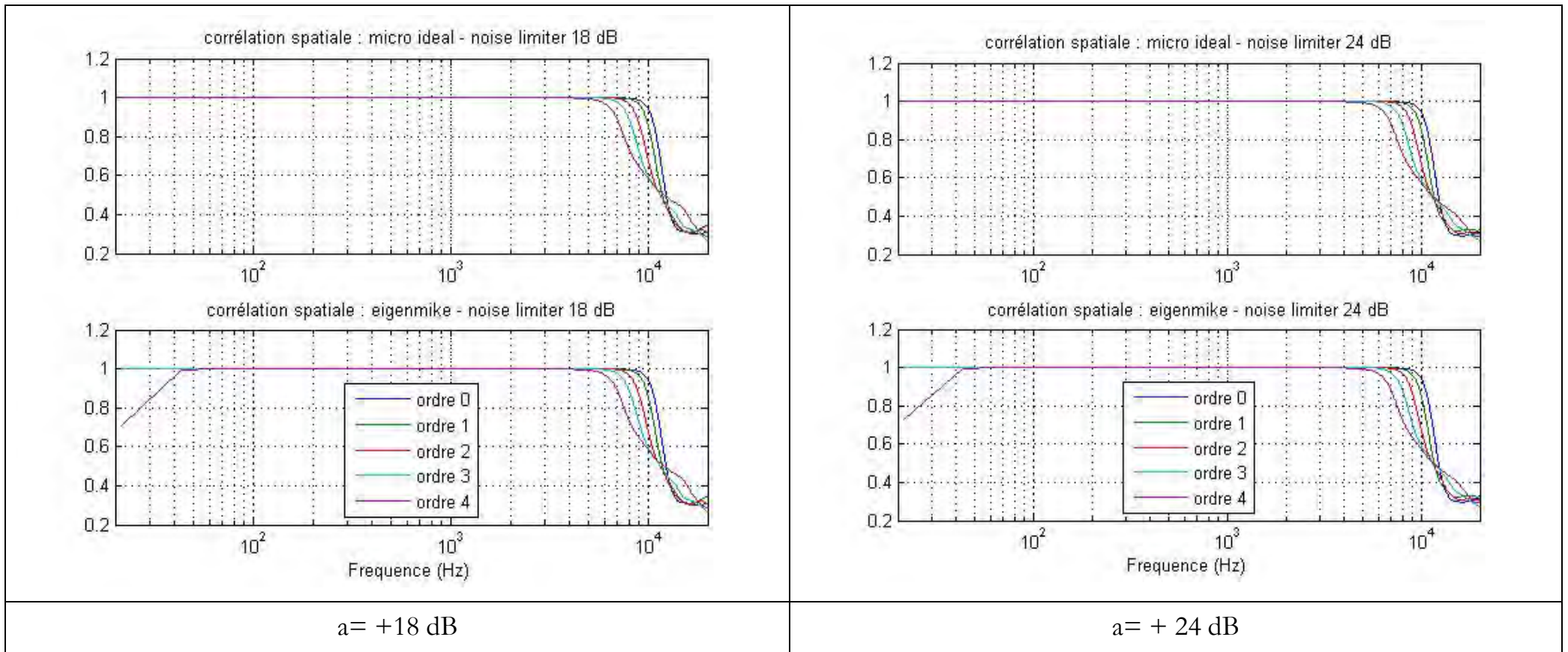
## 7 ANNEXE :

### 7.1 Courbes encodage spatial

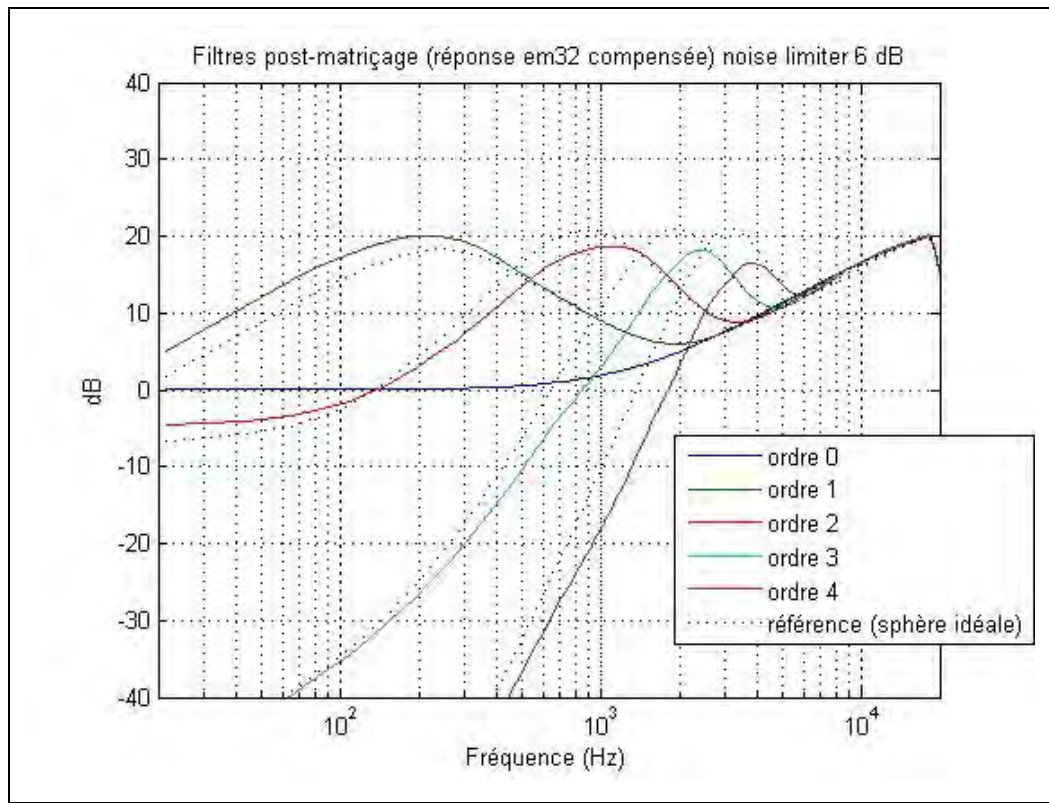


Corrélation spatiale pour  $a = +6 \text{ dB}$  et  $+12 \text{ dB}$

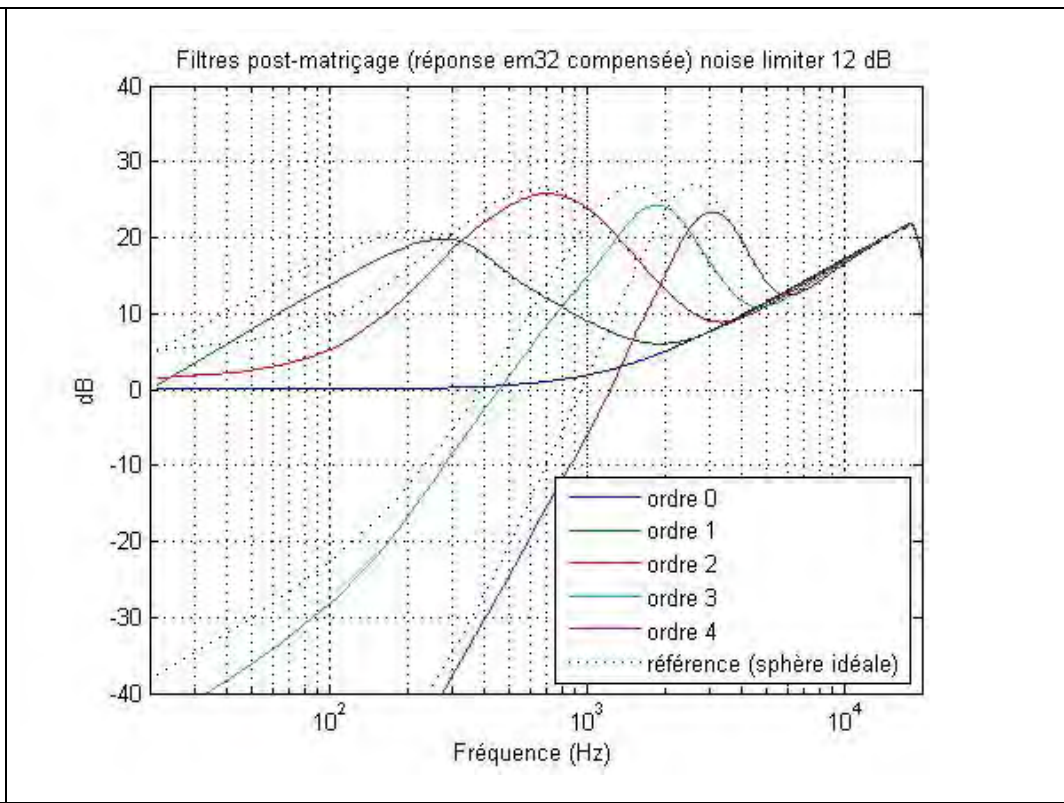




Corrélacion spatiale pour  $a = +18$  dB et  $+24$  dB

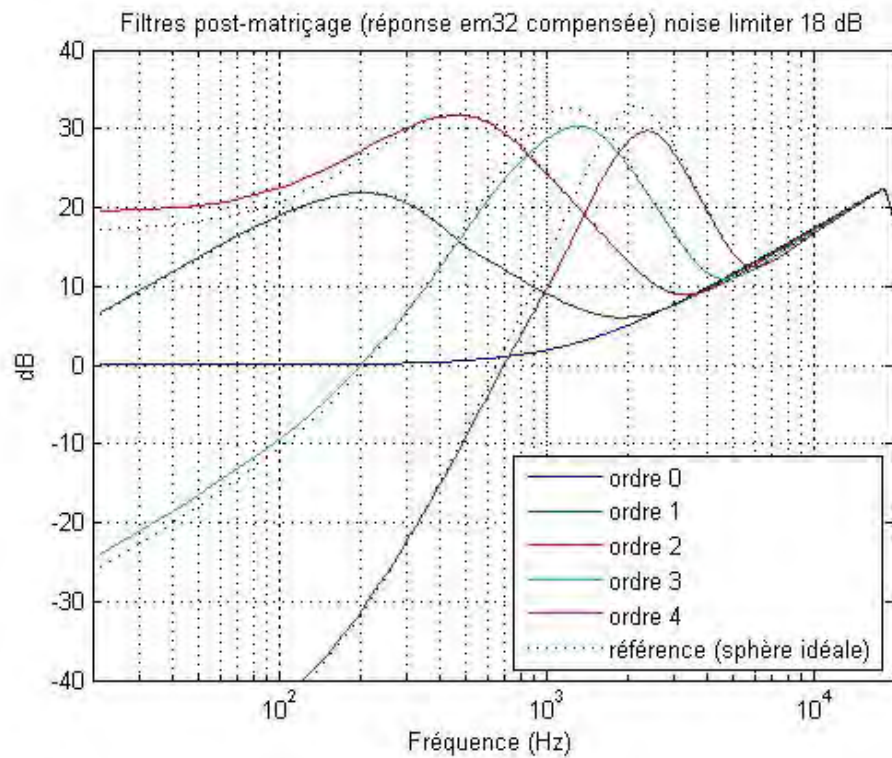


a= +6 dB

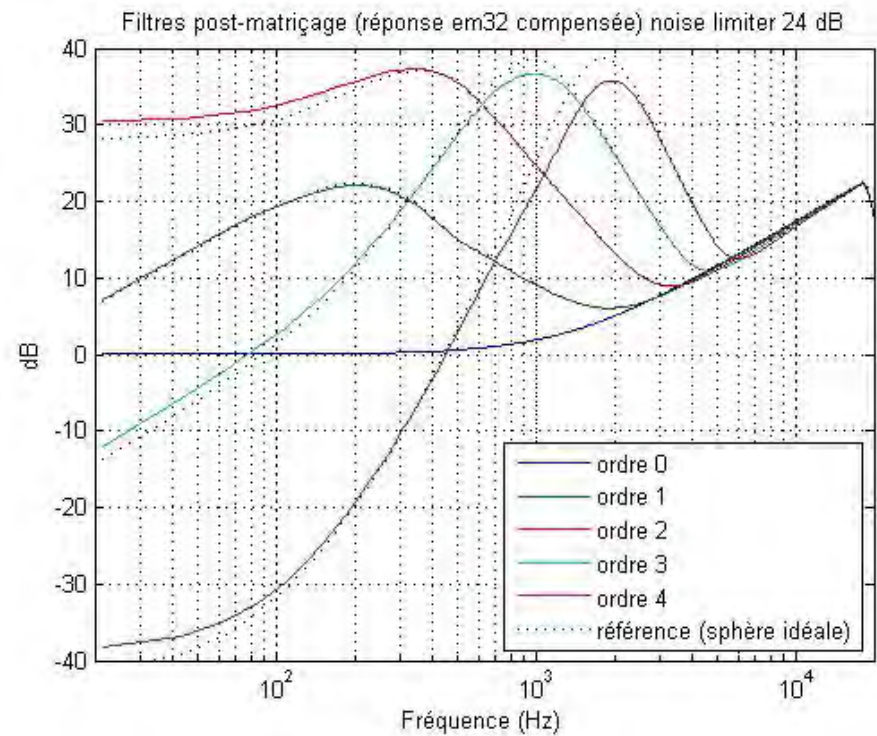


a= +12 dB

Filtres  $EQ_m(Tikhonov)$  pour a = +6 dB et +12 dB

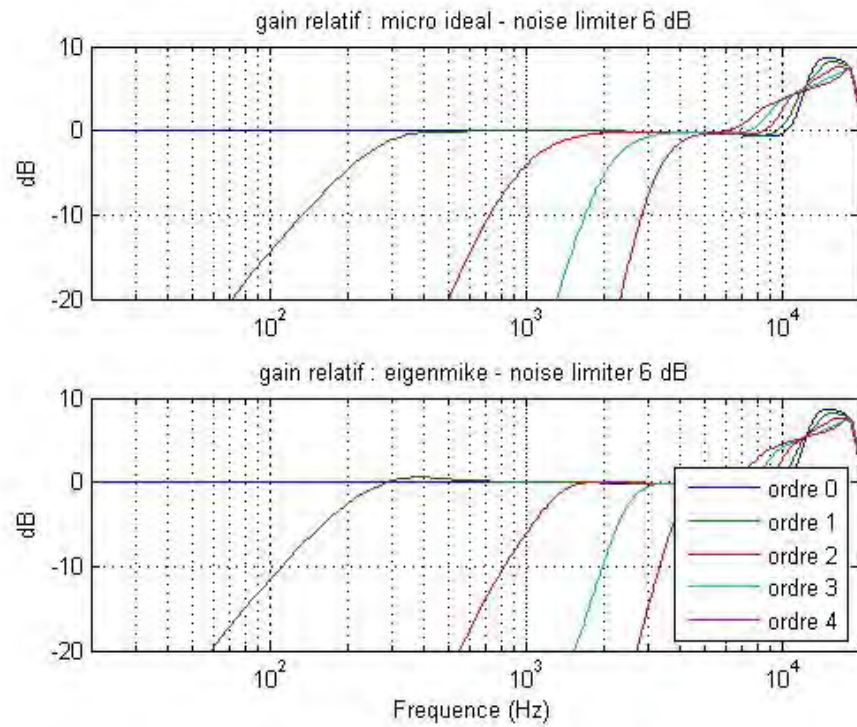


a = +18 dB

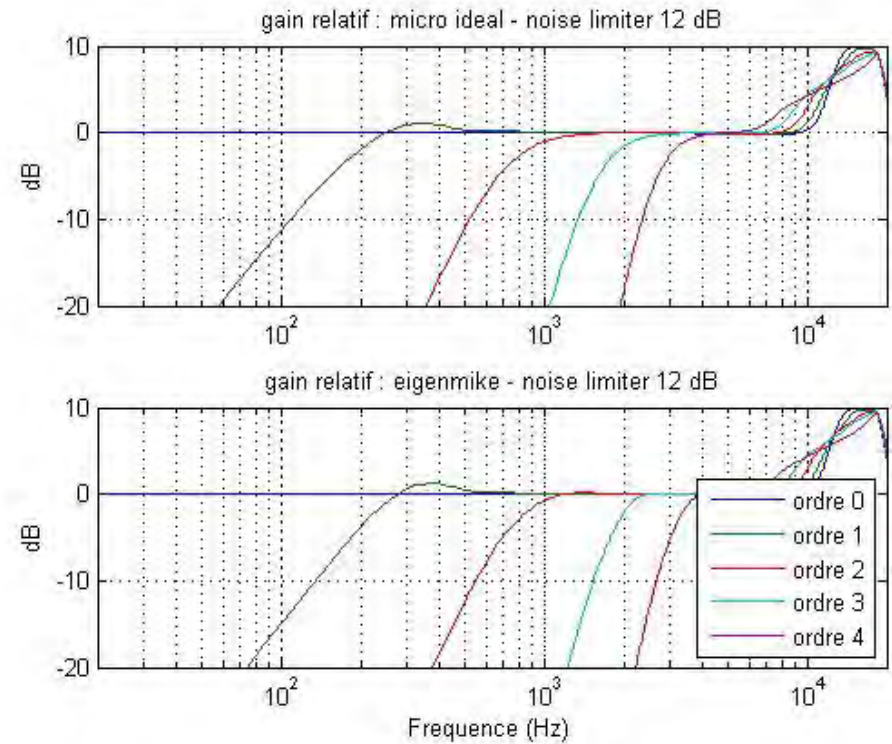


a = +24 dB

Filtres  $EQ_m(Tikhonov)$  pour a = +18 dB et +24 dB

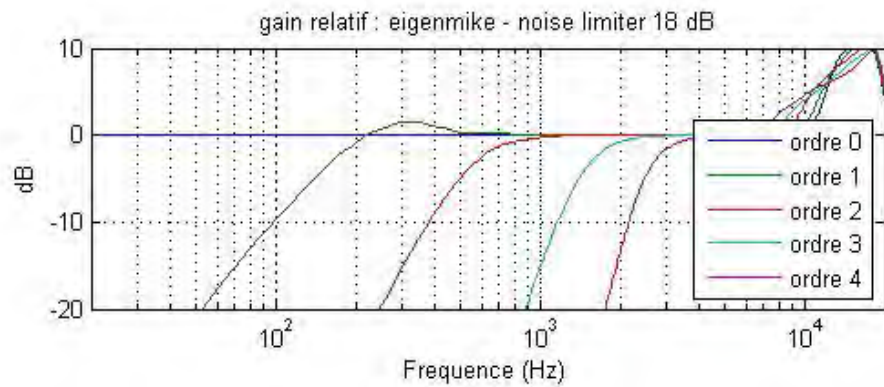
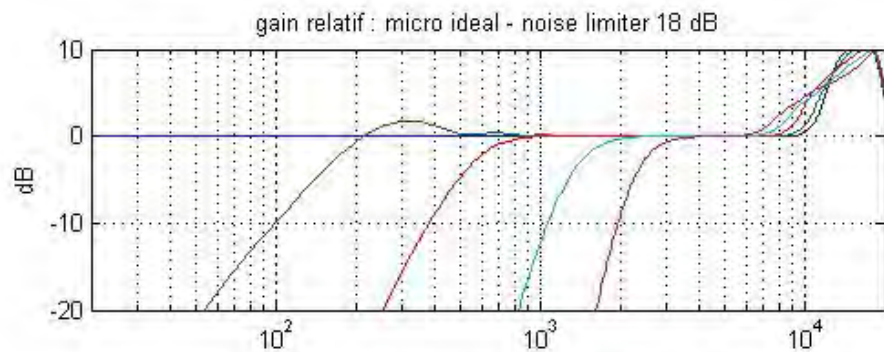


a= +6 dB

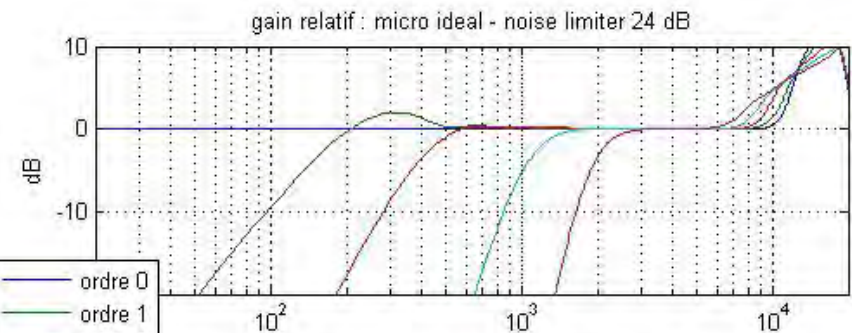


a= +12 dB

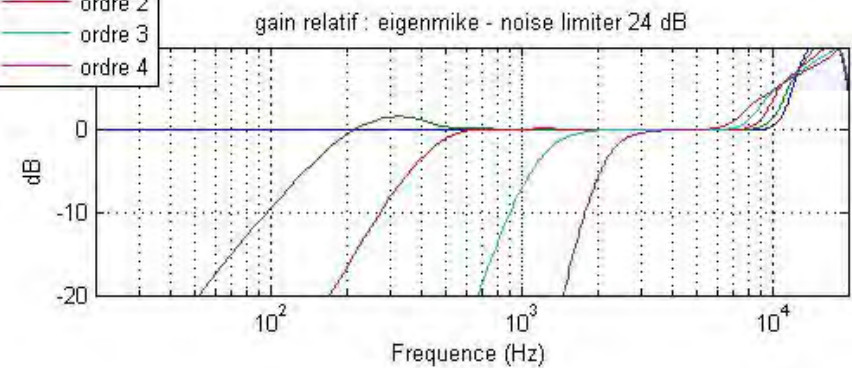
Gain relatif pour a = +6 dB et +12 dB



$a = +18$  dB



- ordre 0
- ordre 1
- ordre 2
- ordre 3
- ordre 4



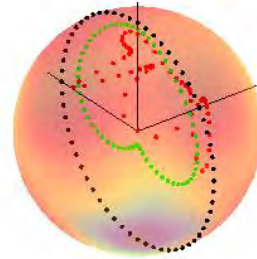
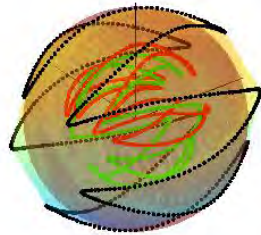
$a = +24$  dB

Gain relatif pour  $a = +18$  dB et  $+24$  dB

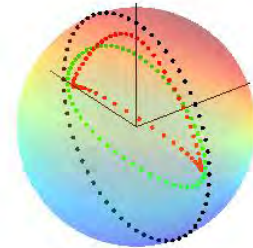
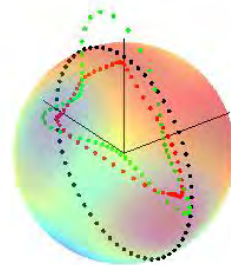
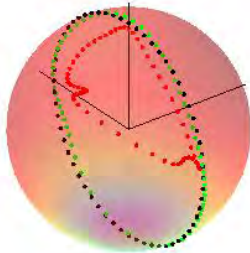
## 7.2 Courbes décodage spatial

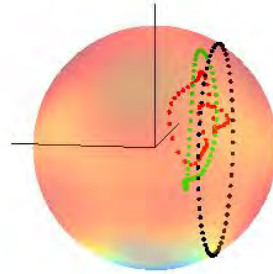
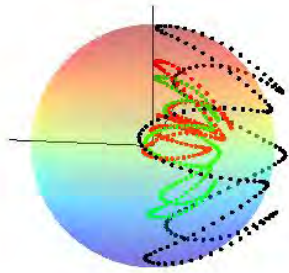
Pour chaque trajectoire, les différents décodeurs sont répartis de la façon suivante

<b>ANCRE</b>	<b>AMBIX PINV</b>	<b>NOM TRAJECTOIRE</b>
<b>BCOM</b>	<b>BRS BASIC</b>	<b>TCM</b>

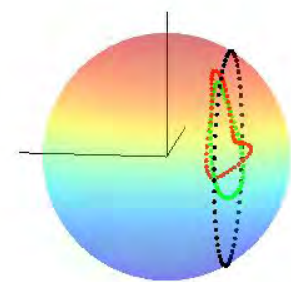
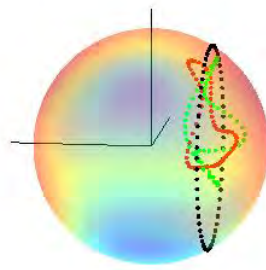
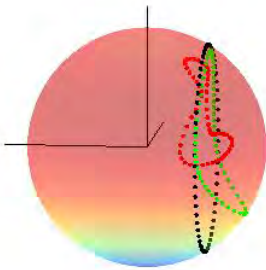


Trajectoire circulaire  $x$  fixe ( $x = 0$ )

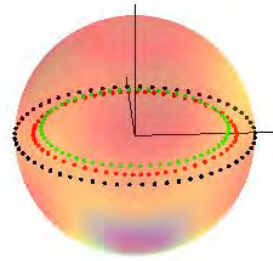
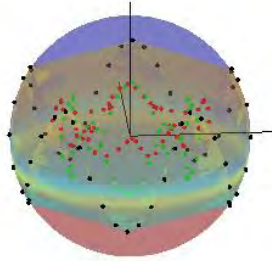




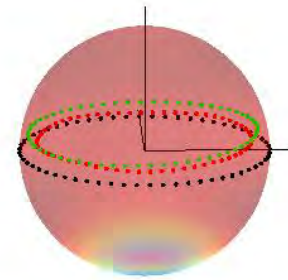
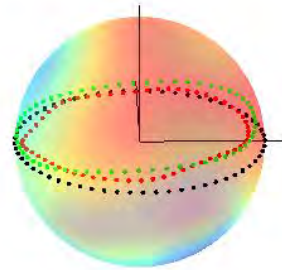
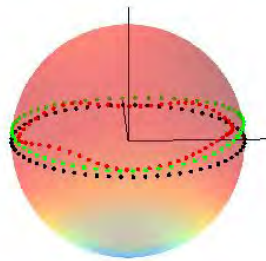
Trajectoire circulaire  $y$  fixe ( $y = -\pi/6$ )







Trajectoire circulaire  $z$  fixe ( $z = 0$ )



## 7.3 Caractéristiques techniques

Caractéristiques Enceintes

(source [http://amadeus-audio.com/fr/produits/pmx/pmx\\_4](http://amadeus-audio.com/fr/produits/pmx/pmx_4))



*PMX 4*

### Spécifications acoustiques

Bande passante 120 Hz à 22 000 Hz

Sensibilité 90 dB

Niveau SPL MAX 112 dB

Dispersion horizontale 100°

Dispersion verticale 100°

### Spécifications techniques

Impédance nominale 8  $\Omega$

Puissance AES 100 W

Puissance programme 200 W

Puissance en crête 300 W

Transducteur LMF	1 haut parleur néodyme de 4 pouces (100 mm)
Transducteur HF	1 tweeter à dome coaxial de 0,8 pouce (20 mm)
Type de filtrage	Passif
Fréquence de filtrage	3 500 Hz
Pentes de filtrage	HPF 6 dB / Octave et LPF 6 dB / Octave

### **Spécifications physiques**

Hauteur	188 mm
Largeur	125 mm
Profondeur	150 mm
Connections	2 speakON NL4
Poids	2,2 kg



### *ML 8*

#### Spécifications acoustiques

Bande passante	46 Hz à 350 Hz
Sensibilité	92 dB
Niveau SPL MAX	118 dB

#### **Spécifications techniques**

Impédance nominale	8 $\Omega$
Puissance AES	300 W
Puissance programme	600 W
Puissance en crête	600 W
Transducteur	1 haut parleur de 8 pouces ventilé haute puissance
Type de charge	Bass reflex

#### Specifications physiques

Hauteur	334 mm
Largeur	310 mm
Profondeur	370 mm
Connections	2 speakON NL4
Poids	14.8 kg
Poignées	2

Amplificateur LABGRUPPEN C 20 :8: (source <http://labgruppen.com/view-model/c-series/c20-8x?page=spec>)



## General

Number of channels	8
Peak total output all channels driven	2000 W
Peak output voltage per channel	100 V
Max output current per channel	8 Arms
Max output power all channels driven	
16 ohms	250 W
8 ohms	250 W
4 ohms	250 W
2 ohms	125 W
Hi-Z per ch. (all ch.'s driven): 70 Vrms / 100 V peak	250 W
Max output power bridged	
16 ohms	500 W
8 ohms	500 W

4 ohms	250 W
2 ohms	Not recommended <sup>1</sup>
Bridged per ch.: 140 V <sub>rms</sub> / 200 V peak	500 W
Performance	
THD 20 Hz - 20 kHz at 1W	<0.1%
THD at 1 kHz and 1 dB below clipping	<0.05%
Signal to Noise Ratio	>112 dBA
Channel separation (Crosstalk) at 1 kHz	>70 dB
Frequency response (1 W into 8 ohms) +0/-3 dB	6.8 Hz - 34 kHz
Input impedance	20 kOhm
Common Mode Rejection (CMR)	>50 dB, 20 Hz - 20 kHz
Output impedance @ 100 Hz	48 mOhm
Gain, Sensitivity and Limiters	
VPL, selectable per ch <sup>2</sup>	100, 63, 45, 32 V
VPL, when bridged <sup>2,3</sup>	200, 126, 90, 64 V
Voltage Peak Limiter mode (per ch.)	Hard / Soft
Amplifier gain selectable (all channels) – rear-panel switches <sup>3</sup>	29, 32, 35, 38 dB
Default gain	32 dB
Level adjustment (per channel)	Front-panel potentiometer, 21 position detented from -inf to

## Connectors and Buttons

Input connectors (per ch.)	3-pin Phoenix, electronically balanced
Output connectors (per ch.)	Barrier strip 2-pole screw terminals
Output bridge mode	A+B, C+D, E+F, G+H, inputs A, C, E, G are signal source
High pass filter	Fixed at 35 Hz, switchable per channel
NomadLink network	On board, 2 x RJ45 connectors IN and OUT
Intelligent fans (on/off)	Yes, depending on presence of output signal
Power	Individual switches on front panel
Cooling	Two fans, front-rear airflow, temperature controlled speed
General Purpose Outputs (GPO)	Contact Closure types, 2-pole Phoenix
General Purpose Inputs (GPI)	Contact Closure types, 2-pole Phoenix

## Power

Operating voltage, 230 V / 115 V nominal	Universal power supply 65-265 V
Minimum power-up voltage, 230 V / 115 V	80 V
Power Average Limiter (PAL) <sup>4</sup>	Yes
Power Factor Correction (PFC)	Yes
Soft-start / Inrush Current Draw	Yes / max. 5 A
Mains connector	IEC inlet



Dimensions

Dimensions

W: 483 mm (19"), H: 88 mm (2 U), D:  
343 mm (13.5")

Weight

Weight

8.5 kg (18.75 lbs.)

Carte son Antelope Orion32 (source <http://antelopeaudio.com/fr/products/Orion32-Multi-Channel-AD-DA-converter>)

<b>Inputs</b>	
Analog Inputs	4 x D-SUB 25 (32 channels)
Digital Inputs	1 x Fiber Optic MADI 2 x Fiber Optic ADAT 1 x S/PDIF
<b>Outputs</b>	
Analog Outputs	4 x D-SUB 25 (32 channels)
Digital Outputs	1 x Fiber Optic MADI 2 x Fiber Optic ADAT 1 x S/PDIF
<b>USB I/O</b>	USB 2.0 Hi-Speed; Data stream up to 480Mbps/192kHz, 32 channels I/O, Type B
<b>Word Clock</b>	1 x 10M Atomic Clock input @ 75 Ohms 3Vpp on BNC 44.1 - 192kHz 1 x Word Clock Input @ 75 Ohms 3Vpp on BNC 44.1 - 192kHz 4 x Word Clock Outputs @ 75 Ohms 3Vpp on BNC 44.1 - 192kHz square wave signal
<b>D/A Converter</b>	
Dynamic Range	118dB
THD + N	-98 dB
<b>A/D Converter</b>	
Dynamic Range	118dB
THD + N	-105 dB
<b>Clock Specs</b>	
Clocking System	4th Generation Acoustically Focused Clocking 64-bit DDS Oven Controlled Crystal Oscillator
Clocking Stability	<+/-0.02 ppm, oven controlled at at 64.5°C/ 148.1°F
Clock Aging	< 1 ppm per year
Clock Calibration	<+/-0.001 ppm
Atomic Clock Input	10MHz @ 1Vpp, BNC
Sample Rates (kHz)	32, 44.1, 48, 88.2, 96, 176.4, 192
Operating Temperature	0-50°C/32- 122°F
Weight	3 kg/6.6 lb approx.
Dimensions (approx.)	Width: 483 mm/19" Height: 44 mm/1.75" Depth: 220 mm/8.66"
Power Supply	AC Universal input 95-245VAC 20 Watts Max Consumption

## 7.4 Questionnaire fourni lors du premier test

### Feuille de renseignement préliminaire au test

#### Informations

##### A.1 Contact

nom	<input type="text"/>	H/F	<input type="text"/>
prénom	<input type="text"/>		
date de naissance	<input type="text"/>	<input type="text"/>	<input type="text"/>
	d	d	m
	<input type="text"/>	<input type="text"/>	<input type="text"/>
	m	m	y
	<input type="text"/>	<input type="text"/>	<input type="text"/>
	y	y	y
	<input type="text"/>	<input type="text"/>	<input type="text"/>
	y	y	y

Adresse mail si vous souhaitez être tenu au courant des résultats :

##### A.2 Avez-vous déjà participé à une expérience sur le son spatialisé ? (5.1, 7.1, Binaural, HOA)

oui  non

##### A.3 Quelle est votre rapport avec le son spatialisé ? (5.1, 7.1, Binaural, HOA, WFS)

Professionnel (recherche, test, audio)  Amateur (musique, cinéma, radio...)   
Pas familier

##### A.4 Etes-vous familier avec les termes Ambisonics, HOA?

oui  non

Le test que vous allez passer consistera à noter des trajectoires de bruit rose dans l'espace. Ces trajectoires vous seront présentées graphiquement sur une feuille, chacune est symbolisée par une lettre.

Le test se déroulera en deux étapes, d'abord un apprentissage pour vous familiariser avec les trajectoires dans l'espace et l'interface de test.

Puis le test en lui-même : il consistera à noter pour chaque trajectoire présentée, deux critères sur une échelle de 0 à 100

Cette notation se fera sur deux critères :

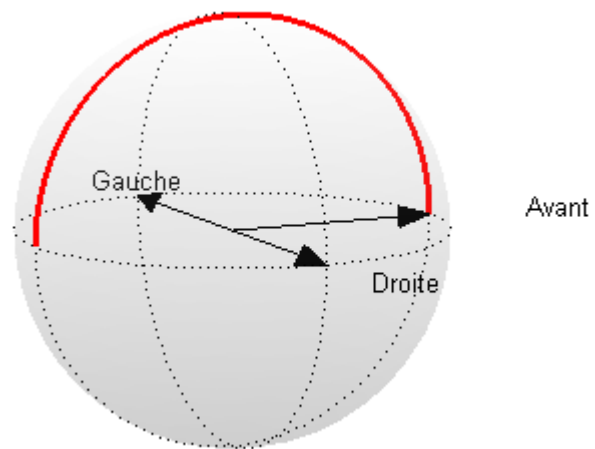
- Fidélité à la trajectoire : La trajectoire entendue vous semble-t-elle correspondre à celle présentée ?
- Constance du timbre : La couleur du son vous semble-t-elle rester constante au cours de la trajectoire ?

Vous verrez apparaître le critère en question à chaque fois que vous appuierez sur le fader correspondant.

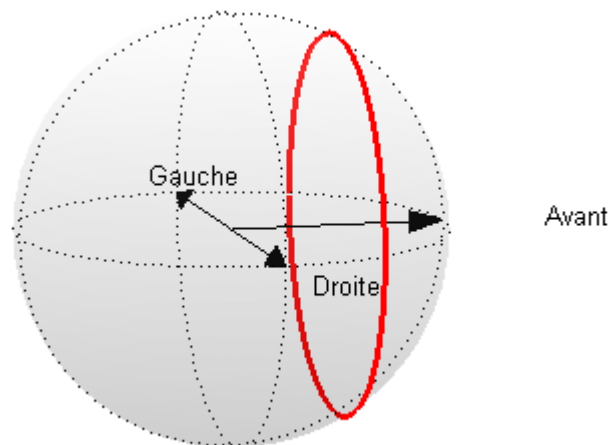
Si vous des questions, n'hésitez pas à poser des questions si les termes ne vous semblent pas clairs, ou si vous avez des problèmes avec le test.

Bon test et bonne écoute !

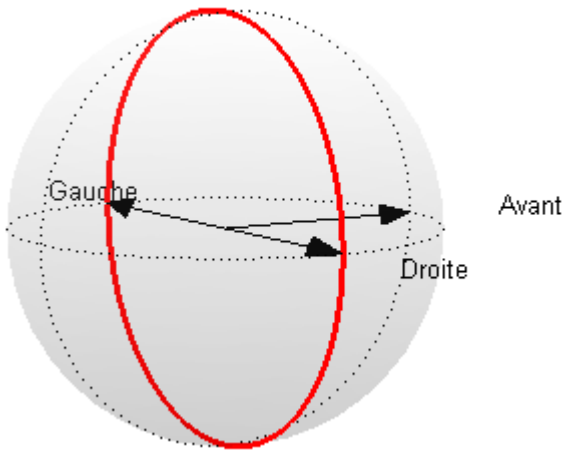
## Apprentissage :



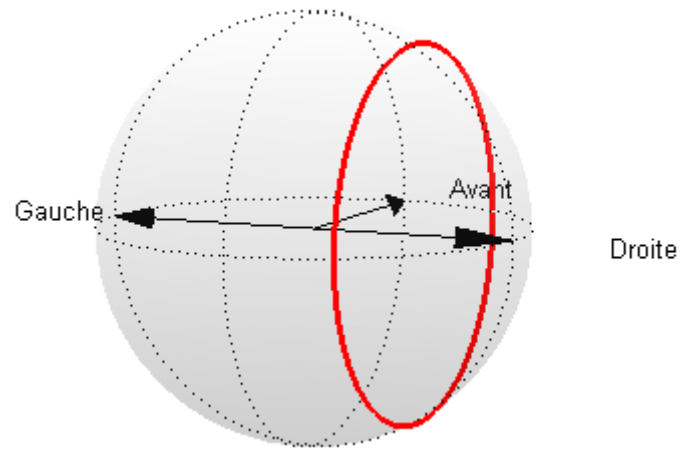
1



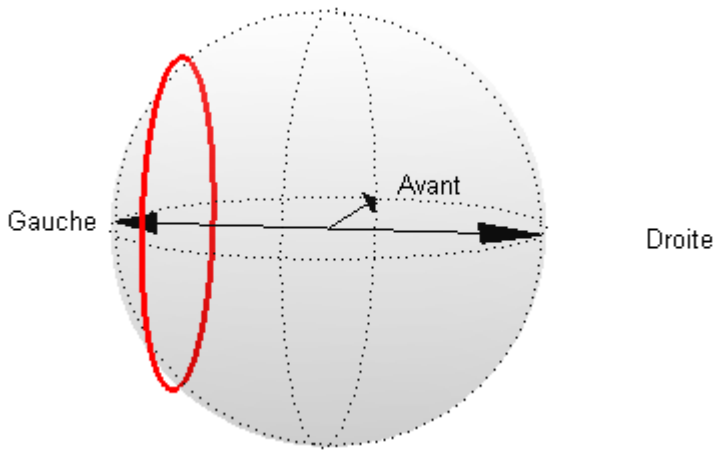
2



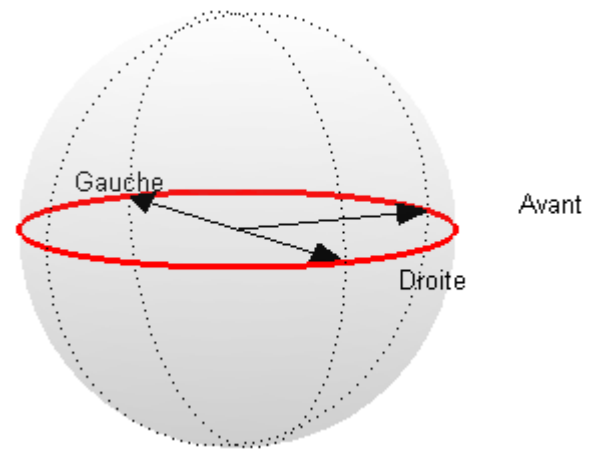
**A**



**B**



**C**



**D**