

CFA/VISHNO 2016

Filtrage directionnel dans un scène sonore 3D par une utilisation conjointe de Beamforming et d'Ambisonie d'ordre élevé

P. Lecomte^{a,b}, P.-A. Gauthier^b, C. Langrenne^a, A. Berry^b et A. Garcia^a

^aLMSSC, CNAM, 2 rue Conté, 75003 Paris, France

^bGroupe d'Acoustique de l'Université de Sherbrooke, 2500 boul. de l'Université,
Sherbrooke, QC, Canada J1K 2R1

pierre.lecomte@cnam.fr



LE MANS

L'ambisonie d'ordre élevé (*higher order ambisonics*) est une technique de représentation d'un champ de pression sur une base tronquée d'harmoniques sphériques. La capture d'un champ sonore naturel est traditionnellement faite à l'aide d'un réseau sphérique de microphones pour obtenir les composantes ambisoniques. Cependant, un microphone sphérique offre également des possibilités de formation de voies (*beamforming*) dans le domaine des harmoniques sphériques : en combinant les différentes composantes ambisoniques, un signal monophonique est obtenu équivalent à celui issu d'un microphone virtuel respectant la directivité choisie pour la voie. Cet article présente une technique de filtrage directionnel d'une scène sonore tridimensionnelle encodée dans un formalisme ambisonique. En alliant les techniques de *beamforming* et d'ambisonie d'ordre élevé, il résulte une scène sonore tridimensionnelle filtrée spatialement avec la fonction de directivité et toujours encodée dans un formalisme ambisonique.

1 Introduction

L'ambisonie d'ordre élevée est une approche flexible d'encodage de scène sonore tridimensionnelle [1]. Le formalisme permet d'effectuer des transformations de la scène sonore dans le domaine des harmoniques sphériques avant décodage sur un réseau de haut-parleurs [2]. Parallèlement, les techniques de formation de voies sur réseau de microphones sphérique permettent d'extraire un signal monophonique dans une scène sonore tridimensionnelle. Ce signal reflète l'information spatiale d'une ou de plusieurs directions privilégiées au moyen d'une fonction de directivité [3]. En exploitant ces directivités dans le domaine des harmoniques sphériques, il est cependant possible de filtrer la scène sonore dans le domaine angulaire et d'obtenir une nouvelle scène sonore tridimensionnelle, pondérée par la fonction de directivité choisie. Cette approche se résume à une opération matricielle dans le cas d'une scène sonore tronquée. L'article s'articule de la manière suivante : après l'introduction des définitions et notations utilisées (Section 2), une rapide présentation de l'ambisonie d'ordre élevé est faite à la Section 3. De la même manière, une présentation succincte de la formation de voies sur réseau de microphones sphérique est faite à la Section 4. Puis, l'opération de filtrage de la scène sonore est faite à la Section 5, avec des résultats de simulations à l'appui (Section 6). Finalement, une implémentation en temps-réel du filtrage spatial est présentée à la Section 7.

2 Définitions et notations

Cet article présente le cas à trois dimensions. Le cas en deux dimensions peut être déduit directement.

2.1 Système de coordonnées

Le système de coordonnées sphériques utilisé est montré par la Figure 1 avec

$$x = r \cos(\theta) \cos(\delta), \quad y = r \sin(\theta) \cos(\delta), \quad z = r \sin(\delta) \quad (1)$$

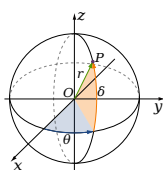


FIGURE 1 – Système de coordonnées sphériques. Un point $P(x, y, z)$ est décrit par le rayon r , l'azimut θ et l'élévation δ .

2.2 Harmoniques sphériques

Les harmoniques sphériques utilisés respectent la normalisation N3D [1], ils sont donnés par :

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

où $P_{m|n|}$ sont les fonctions de Legendre associées d'ordre m et de degré $|n|$. On note que $(n, m) \in (\mathbb{Z}, \mathbb{N})$ avec $|n| \leq m$, et $\epsilon_n = 1$ si $n = 0$, $\epsilon_n = 2$ si $|n| > 0$. Ainsi, l'Eq. (2) dénote un harmonique sphérique d'ordre m et de degré n . La famille d'harmoniques sphériques définit une base orthonormée de \mathbb{L}^2 , l'espace des fonctions de carré intégrable sur la sphère unité $\{(x, y, z) \in \mathbb{R}^3, \|x^2 + y^2 + z^2\| = 1\}$. À chaque ordre m on dénombre $(2m+1)$ harmoniques sphériques. Ainsi, jusqu'à l'ordre M il y a $(M+1)^2$ harmoniques sphériques.

2.3 Notation ACN

Dans le cas d'une notation vectorielle, les harmoniques sphériques peuvent être rangés de plusieurs manières suivant les indices (m, n) . Dans cet article, les harmoniques sphériques seront ordonnés suivant leur ordre *ACN* (*Ambisonic Channel Number*) croissant [4]

$$ACN = m^2 + m + n \quad (3)$$

Les indices m et n peuvent être retrouvés à l'aide des formules suivantes :

$$\begin{aligned} m &= \lfloor \sqrt{ACN} \rfloor \\ n &= ACN - m^2 - m \end{aligned} \quad (4)$$

où $\lfloor \cdot \rfloor$ représente la partie entière.

2.4 Transformée de Fourier sphérique

La transformée de Fourier sphérique permet d'exprimer une fonction angulaire $f(\theta, \delta)$ sur la base des harmoniques sphériques. Les coefficients f_{mn} issus de la projection de cette fonction sur la base orthonormée sont donnés par :

$$f_{mn} = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} f(\theta, \delta) Y_{mn}(\theta, \delta) \cos(\delta) d\theta d\delta \quad (5)$$

La transformée de Fourier sphérique inverse est donnée par :

$$f(\theta, \delta) = \sum_{m=0}^{\infty} \sum_{n=-m}^m f_{mn} Y_{mn}(\theta, \delta) \quad (6)$$

3 Ambisonie d'ordre élevé

Le formalisme de l'ambisonie d'ordre élevé est brièvement rappelé dans cette section. Plus de détails sont donnés par les références [1, 5].

3.1 Représentation d'un champ sonore

L'ambisonie a été introduite par Gerzon dans les années 70 [6] au premier ordre, puis étendue aux ordres supérieurs par Daniel notamment [1]. La formulation repose sur la solution de l'équation d'Helmholtz en coordonnées sphériques. Dans un domaine exempt de sources acoustiques, avec la dépendance temporelle $e^{+i\omega t}$, l'équation de Helmholtz pour la pression acoustique p s'exprime par $\Delta p + k^2 p = 0$ (ω représente la pulsation, $k = \omega/c$ le nombre d'onde où c représente la célérité du son). La solution en coordonnées sphériques donne l'expression de Fourier-Bessel du champ de pression acoustique [7] :

$$p(kr, \theta, \delta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=-m}^m B_{mn} Y_{mn}(\theta, \delta) \quad (7)$$

Dans cette solution, la partie radiale et angulaire sont séparées. La dépendance radiale est exprimée à l'aide des fonctions de Bessel $j_m(kr)$ et la dépendance angulaire est donnée par les harmoniques sphériques $Y_{mn}(\theta, \delta)$. Cependant l'expression de ces fonctions étant la même pour tout champ de pression acoustique, ce dernier est entièrement exprimé à l'aide des composantes B_{mn} , appelées composantes ambisoniques. Les composantes ambisoniques pour une onde plane harmonique d'amplitude S et de direction de propagation θ_s, δ_s sont données par :

$$B_{mn} = S Y_{mn}(\theta_s, \delta_s) \quad (8)$$

3.2 Troncature

En pratique, la représentation d'un champ de pression acoustique à l'aide d'une série de Fourier-Bessel de l'Eq. (7) est tronquée à l'ordre M . Le champ est alors représenté à l'aide de $(M+1)^2$ composantes ambisoniques. L'analyse de l'erreur de troncature est menée par Ward et al. [8] pour le cas bidimensionnel et par Poletti [5] pour le cas tridimensionnel. De ces références, une règle générale et approximative est établie : pour une erreur relative de 4 % entre le champ de pression sans troncature et le champ de pression tronqué à l'ordre M , la taille de la zone de reconstruction est donnée par une sphère de rayon $r_{4\%}$. Ce rayon est proportionnel à l'ordre de troncature M et inversement proportionnel à la fréquence, via le nombre d'onde k :

$$r_{4\%} = M/k \quad (9)$$

La Figure 2(a) montre le champ de pression acoustique d'une onde plane d'incidence $(20^\circ, 0^\circ)$, d'amplitude $S = 1$ et de fréquence égale à 500 Hz. La Figure 2(b) montre sa représentation ambisonique tronquée à l'ordre $M = 4$ suivant l'Eq. (7). Les coefficients B_{mn} sont obtenus suivant l'Eq. (8). Le cercle délimitant la zone de reconstruction selon l'Eq. (9) est montré en rouge par la Figure 2(b).

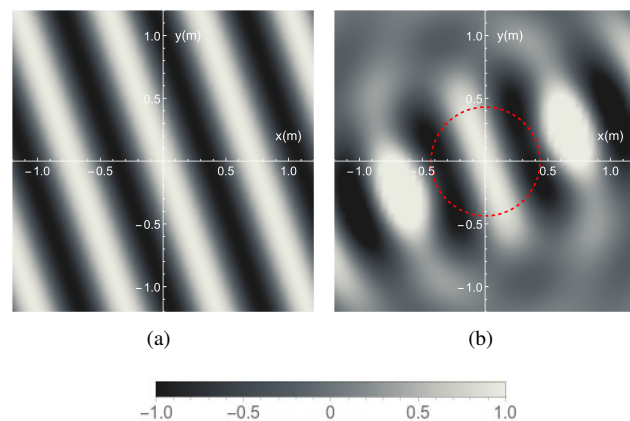


FIGURE 2 – (a) Partie réelle du champ de pression acoustique d'une onde plane. (b) Représentation tronquée de la même onde plane à l'ordre $M = 4$. Le cercle en rouge donne la taille de la zone de reconstruction suivant l'Eq. (9).

3.3 Enregistrement d'un champ sonore

L'estimation des composantes ambisoniques B_{mn} pour un champ sonore mesuré se fait à l'aide d'un réseau de microphones, le plus souvent répartis sur la surface d'une sphère rigide [9]. La géométrie sphérique permet d'effectuer une transformée de Fourier sphérique du champ de pression et d'obtenir les coefficients p_{mn} issus de cette transformée (Eq. (5)), avec p et p_{mn} qui remplacent f et f_{mn} , respectivement. La condition de rigidité de la sphère permet d'exploiter les propriétés de la diffraction et donne des filtres radiaux sans singularités [9]. Finalement, les composantes ambisoniques sont obtenues à l'aide de l'équation :

$$B_{mn} = i^{-(m-1)} (ka)^2 h_m^{(2)}(ka) p_{mn} \quad (10)$$

où a représente le rayon du microphone sphérique et $h_m^{(2)}$ la dérivée première des fonctions de Hankel de seconde espèce d'ordre m . Puisque le réseau de microphones contient un nombre fini de capteurs, la répartition et le nombre de ces derniers conditionnent l'estimation des composantes ambisoniques jusqu'à un ordre M . À titre d'exemple, le maillage de Lebedev est un bon candidat pour limiter le repliement et avoir une bonne estimation des composantes harmoniques sphériques jusqu'à l'ordre $M = 5$ [10]. Finalement, le rayon a du microphone fixe le compromis entre l'estimation des composantes ambisoniques en basses fréquences et le repliement en hautes fréquences [9].

4 Beamforming avec un réseau de microphones sphérique

Les réseaux de microphones sphériques sont typiquement utilisés pour extraire un signal monophonique résultant d'une écoute dans une ou plusieurs directions privilégiées. En combinant les signaux des différents microphones il est possible de simuler une prise de son directionnelle suivant une fonction de directivité déterminée (*beam pattern*). Cependant la géométrie sphérique permet de travailler directement dans le domaine des harmoniques sphériques après encodage du champ, suivant l'Eq. (10). Cette approche est notamment proposée dans [3]. La combinaison des signaux se fait alors sur les composantes ambisoniques.

De manière générale, le signal monophonique y , extrait du réseau de microphones sphérique est donné par [11] :

$$y = \sum_{m=0}^{\infty} \sum_{n=-m}^m w_{mn} B_{mn} \quad (11)$$

où w_{mn} représente les poids associés à la directivité désirée et B_{mn} , les composantes ambisoniques du champ de pression enregistré. Plusieurs approches pour la conception des fonctions de directivité (c.-à.d. la détermination des poids w_{mn}) sont proposées dans la littérature parmi lesquels la *delay-and-sum*, par approche Dolph-Chebyshev [12], ou par approches régulières [3, 13]. Une bonne revue des différentes techniques est faite dans [14, Chapitre 11]. À titre d'exemple, la Section 5.3.1 présente le cas d'une directivité prescrite simple, soit la fonction hyper-cardioïde.

5 Filtrage angulaire d'une scène sonore ambisonique

Dans cette section, la scène sonore captée et décrite par les signaux ambisoniques B_{mn} est filtrée à l'aide d'une fonction angulaire $g(\theta, \delta)$ (une directivité choisie) pour obtenir, non pas un signal monophonique, mais une scène sonore tridimensionnelle complète pour laquelle certaines directions sont atténuées. En d'autres termes, à l'instar d'une approche par formation de voies où le signal de sortie est monophonique, (l'information spatiale autre que dans la direction de pointage θ_0, δ_0 est perdue), l'approche décrite ci-après permet de rester dans le domaine ambisonique après combinaison, et d'obtenir une scène sonore spatialement filtrée qui pourrait être par la suite reproduite. Cet aspect représente la contribution principale de cet article. Afin d'alléger les notations, la notation *ACN* est utilisée dans cette section et les harmoniques sphériques sont identifiés à l'aide d'un seul indice.

5.1 Fonction angulaire explicite

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

$$\tilde{b}(\theta, \delta) = g(\theta, \delta)b(\theta, \delta) \quad (12)$$

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

$$b(\theta, \delta) = \sum_{j=0}^{\infty} b_j Y_j(\theta, \delta) \quad (13)$$

Chaque composante \tilde{b}_i de la scène sonore filtrée est obtenue en appliquant une transformée de Fourier sphérique sur l'Eq. (13) :

$$\tilde{b}_i = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} g(\theta, \delta) \left(\sum_{j=0}^{\infty} b_j Y_j(\theta, \delta) \right) Y_i(\theta, \delta) \cos(\delta) d\theta d\delta \quad (14)$$

soit :

$$\tilde{b}_i = \frac{1}{4\pi} \sum_{j=0}^{\infty} b_j \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} g(\theta, \delta) Y_j(\theta, \delta) Y_i(\theta, \delta) \cos(\delta) d\theta d\delta \quad (15)$$

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

$$\tilde{\mathbf{b}} = \mathbf{M}\mathbf{b} \quad (16)$$

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

$$\mathbf{M}_{ij} = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} g(\theta, \delta) Y_j(\theta, \delta) Y_i(\theta, \delta) \cos(\delta) d\theta d\delta \quad (17)$$

Il est important de noter le fait que l'Eq. (17) représente un filtrage mais aussi un changement d'ordre de part la taille potentiellement rectangulaire de \mathbf{M} . Le calcul de \mathbf{M}_{ij} peut poser des difficultés. Il suppose notamment de connaître de manière analytique la fonction angulaire g . Néanmoins, il est possible de calculer l'intégrale en utilisant une règle de quadrature pour la sphère, qui respecte l'orthonormalité des harmoniques sphériques [10]. Cette approche est proposée par Kronlachner et Zotter [2] à l'aide de maillages de type *t-design* [16]. Il devient alors possible de définir une fonction angulaire g arbitraire. L'intégration utilisant une quadrature donnera un résultat exact jusqu'à un ordre M' pour laquelle elle fonctionne, mais si g possède dans sa transformée de Fourier sphérique des coefficients non-nuls pour les ordres plus grands que M' , ces derniers peuvent se replier sur les ordres inférieurs selon le comportement de la règle de quadrature [17]. La matrice \mathbf{M} étant de dimension finie, le champ sonore filtré $\tilde{\mathbf{b}}$ est tronqué à l'ordre \tilde{M} . Le contenu du spectre de Fourier sphérique de g peut nécessiter un ordre de redécomposition \tilde{M} assez important pour obtenir le résultat escompté.

5.2 Dirac angulaire

Un cas particulier de l'Eq. (17), $g(\theta, \delta) = \delta(\theta - \theta_0, \delta - \delta_0)$, est un Dirac angulaire dans la direction (θ_0, δ_0) , et le terme \mathbf{M}_{ij} de l'Eq. (17) se simplifie alors comme suit :

$$\mathbf{M}_{ij} = \frac{1}{4\pi} Y_i(\theta_0, \delta_0) Y_j(\theta_0, \delta_0) \quad (18)$$

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

$$\tilde{b}_i = \frac{1}{4\pi} Y_i(\theta_0, \delta_0) \sum_{j=0}^{(M+1)^2-1} b_j Y_j(\theta_0, \delta_0) \quad (19)$$

soit

$$\tilde{b}_i = \frac{1}{4\pi} Y_i(\theta_0, \delta_0) b(\theta_0, \delta_0) \quad (20)$$

5.3 Approche utilisant la décomposition de la fonction de directivité g

En repartant de l'Eq. (14), la fonction g peut également être décrite par sa transformée de Fourier sphérique inverse :

$$g(\theta, \delta) = \sum_{k=0}^{\infty} g_k Y_k(\theta, \delta) \quad (21)$$

Si la fonction g est décrite jusqu'à l'ordre M' , le terme M_{ij} de l'Eq. (16) devient :

$$M_{ij} = \sum_{k=0}^{(M'+1)^2-1} g_k C_{ijk} \quad (22)$$

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

$$C_{ijk} = \frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\delta=-\pi/2}^{\pi/2} Y_i(\theta, \delta) Y_j(\theta, \delta) Y_k(\theta, \delta) \cos(\delta) d\theta d\delta \quad (23)$$

Cette intégrale est indépendante des fonctions b et g et est reliée aux coefficients de Clebsch-Gordan qui peuvent être pré-calculés pour une implémentation efficace [18, pp. 803-805]. L'ordre de redécomposition \tilde{M} nécessaire pour ne pas omettre de composantes ambisoniques dans le résultat filtré est tel que $\tilde{M} = M + M'$. En effet, en remplaçant g dans l'Eq. (14) par son expression Eq. (21), il est plus évident que la fonction à projeter sur les harmoniques sphériques est d'un ordre maximal $M + M'$.

Avec cette approche, la littérature du *beamforming* pour réseau de microphones sphérique devient un atout pour calculer la matrice \mathbf{M} de l'Eq. (16). En effet, les coefficients g_k de l'Eq. 21 sont fournis dans la littérature pour des fonctions directionnelles optimisées, par exemple par Meyer et Elko [3] pour une directivité hyper-cardioïde, présentée dans la section suivante.

5.3.1 Fonction de directivité régulière de type hyper-cardioïde

La directivité régulière de type hyper-cardioïde avec un angle de pointage donné par (θ_0, δ_0) est décrite suivant l'Eq. (21) avec les poids g_{mn} donnés par :

$$g_{mn} = c_m \sqrt{\frac{1}{2m+1}} Y_{mn}(\theta_0, \delta_0) \quad (24)$$

Les coefficients c_m sont récapitulés par le Tableau (1), selon [3], à ceci près qu'une normalisation par la valeur maximale du lobe principal est effectuée ici. L'obtention de ces coefficients est décrite par Teutsch [15, Chapitre 5].

TABLEAU 1 – Coefficients c_m pour un directivité de type hyper-cardioïde

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

La Figure 3 montre une fonction de directivité résultant des coefficients choisis pour le cas considéré.

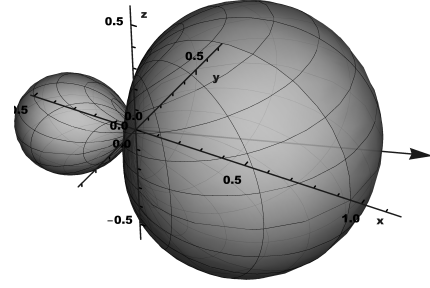


FIGURE 3 – Directivité obtenue pour le cas hyper-cardioïde d'ordre 1, pointant vers la direction $(20^\circ, 0^\circ)$ indiquée par la flèche.

6 Simulations

Dans cette section, des résultats de simulation sont présentés pour un filtrage angulaire par deux fonctions de directivité : d'une part la fonction de directivité hyper-cardioïde d'ordre 1, telle que présentée à la Section 5.3.1, et, d'autre part, un Dirac angulaire tel que présenté à la section 5.2.

La scène sonore à filtrer est composée de 2 ondes planes harmoniques à 500 Hz, d'amplitudes $S_1 = S_2 = 1$ et de directions de propagation respectives de $(\theta_1 = 20^\circ, \delta_1 = 0^\circ)$ et $(\theta_2 = 110^\circ, \delta_2 = 0^\circ)$. La scène sonore est décrite jusqu'à l'ordre $M = 4$. Le champ de pression acoustique tel que représenté par la série tronquée à l'ordre $M = 4$ est montré par la Figure 4(a). Le diagramme de directivité de la scène sonore est montré par la Figure 5 en rouge pointillée d'après l'Eq. (13).

6.1 Fonction hyper-cardioïde d'ordre 1

La matrice \mathbf{M} est construite pour une fonction de directivité hyper-cardioïde d'ordre 1, telle qu'indiquée par l'Eq. (22). L'hyper-cardioïde en question est montré par la Figure 3 pour angle de pointage de $(20^\circ, 0^\circ)$. Elle est également visible en noir pointillé à la Figure 5. L'ordre de redécomposition choisi pour le filtrage est de $\tilde{M} = 5$. Ceci dans le but d'éviter d'omettre de l'énergie dans les composantes ambisoniques après filtrage, tel que mentionné à la Section 5.1. Ainsi, la matrice \mathbf{M} de l'Eq. (17) est de dimensions $(5 + 1)^2 \times (4 + 1)^2$. Le résultat après filtrage est montré par la Figure 4(b). Le diagramme de directivité correspondant est montré par la Figure 5 en bleu.

Il est possible d'observer à la Figure 4(b) que le résultat retrouvé ressemble à celui d'une seule onde plane provenant de la direction $(20^\circ, 0^\circ)$, montré par la Figure 2(b). Cependant, d'après les diagrammes de directivité Figure 5, l'onde plane provenant de la direction $(110^\circ, 0^\circ)$ est encore présente dans la scène sonore filtrée. Ceci est dû au fait que la fonction de directivité utilisée ici n'est pas très directive. Même si la fonction de directivité g choisie était plus directive, elle ne permettrait pas d'isoler totalement l'onde plane provenant de $(20^\circ, 0^\circ)$, car des lobes secondaires de l'onde plane $(110^\circ, 0^\circ)$ sont également présent vers $(20^\circ, 0^\circ)$. Ceci est une conséquence du fait que la scène sonore est tronquée à un ordre trop faible pour limiter les lobes secondaires associés à une onde plane d'une direction de propagation donnée (ici, $M = 4$).

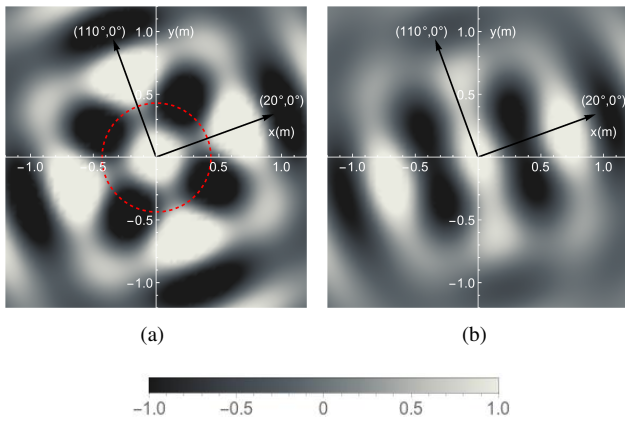


FIGURE 4 – (a) Partie réelle du champ de pression acoustique, tel que représenté par une série de Fourier-Bessel tronquée à l'ordre $M = 4$ pour le cas de deux ondes planes harmoniques. Le cercle rouge montre la taille de la zone efficace de reconstruction suivant l'Eq. (9). (b) Partie réelle du champ de pression acoustique montré en (a) filtré par \mathbf{M} , lorsque défini pour une hyper-cardioïde d'ordre 1 visant à $(20^\circ, 0^\circ)$. Représentation tronquée à l'ordre $\tilde{M} = 5$. Les flèches indiquent la direction de propagation des deux ondes planes qui composent la scène sonore **b**.

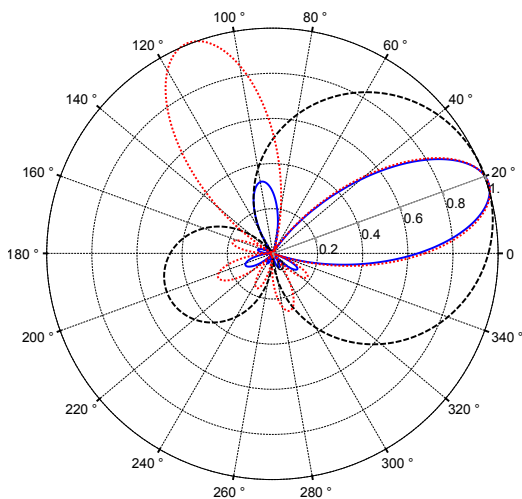


FIGURE 5 – Diagrammes de directivité dans le plan horizontal ($\delta = 0$) de la scène sonore à filtrer $b(\theta, 0)$ (en rouge pointillé), de la fonction de directivité utilisée pour le filtrage $g(\theta, 0)$ (en noir pointillé) et le résultat après filtrage $\tilde{b}(\theta, 0)$ (en bleu).

6.2 Dirac angulaire

Un deuxième cas de simulation utilisant un Dirac angulaire pour g , comme décrit à la section 5.2, est montré par la Figure 6 pour une représentation tronquée à l'ordre $\tilde{M} = 5$. Le Dirac vise dans la direction $(20^\circ, 0^\circ)$.

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

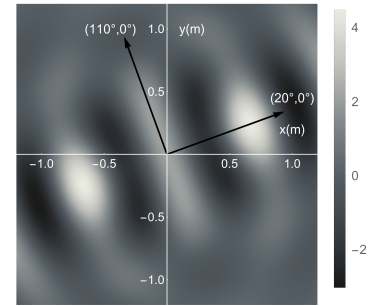


FIGURE 6 – Partie réelle du champ de pression acoustique filtré par \mathbf{M} , lorsque g est défini par un Dirac angulaire de direction $(20^\circ, 0^\circ)$. Représentation tronquée à l'ordre $\tilde{M} = 5$. Les flèches indiquent la direction de provenance des deux ondes planes qui composent la scène sonore **b**.

7 Implémentation temps-réel

La matrice \mathbf{M} est implémentée dans la suite logicielle *ambitools*¹ [19], pour un Dirac angulaire (Eq. (18)), et pour des fonctions de directivités hyper-cardioïdes jusqu'à l'ordre $M' = 3$, (Eq. (22)). Les dimensions de \mathbf{M} sont telles que $M = \tilde{M} = 5$. Cette matrice de filtrage s'insère dans le domaine ambisonique avant le décodage (c.-à-d. avant la restitution sur réseau de haut-parleurs du champ sonore filtré spatialement). Une capture d'écran de l'interface sous Linux Jack/Qt est montré par la Figure 7 pour le cas des filtres hyper-cardioïdes. L'utilisateur peut activer, ou non, l'effet ($\mathbf{M} = \mathbf{I}$ lorsque l'effet est désactivé). L'angle de pointage peut-être changé en temps-réel, ainsi que l'ordre dans le cas d'une directivité hyper-cardioïde. Au final, cet outil permet d'explorer finement la scène sonore en temps-réel.

8 Conclusions

Cet article a présenté une approche pour le filtrage spatial directif d'une scène sonore tridimensionnelle encodée dans un formalisme ambisonique d'ordre élevé. En utilisant des fonctions de directivité issues de la littérature de la formation de voies, il est possible de filtrer spatialement la scène sonore pour obtenir une nouvelle scène sonore tridimensionnelle prête à être décodée sur un réseau de haut-parleurs. Des résultats de simulations convaincants ont été présentés pour une fonction de directivité hyper-cardioïde d'ordre un et pour un Dirac angulaire. Un outil de filtrage en temps-réel a également été présenté. Il permet une exploration fine d'une scène sonore.

Références

- [1] Jérôme DANIEL : *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. Thèse de doctorat, Université Paris 6, Paris, 2000.
- [2] Matthias KRONLACHNER et FRANZ ZOTTER : Spatial transformations for the enhancement of Ambisonic recordings. *2nd International Conference on Spatial Audio*, Erlangen, 2014.

1. <https://github.com/sekisushai/ambitools>

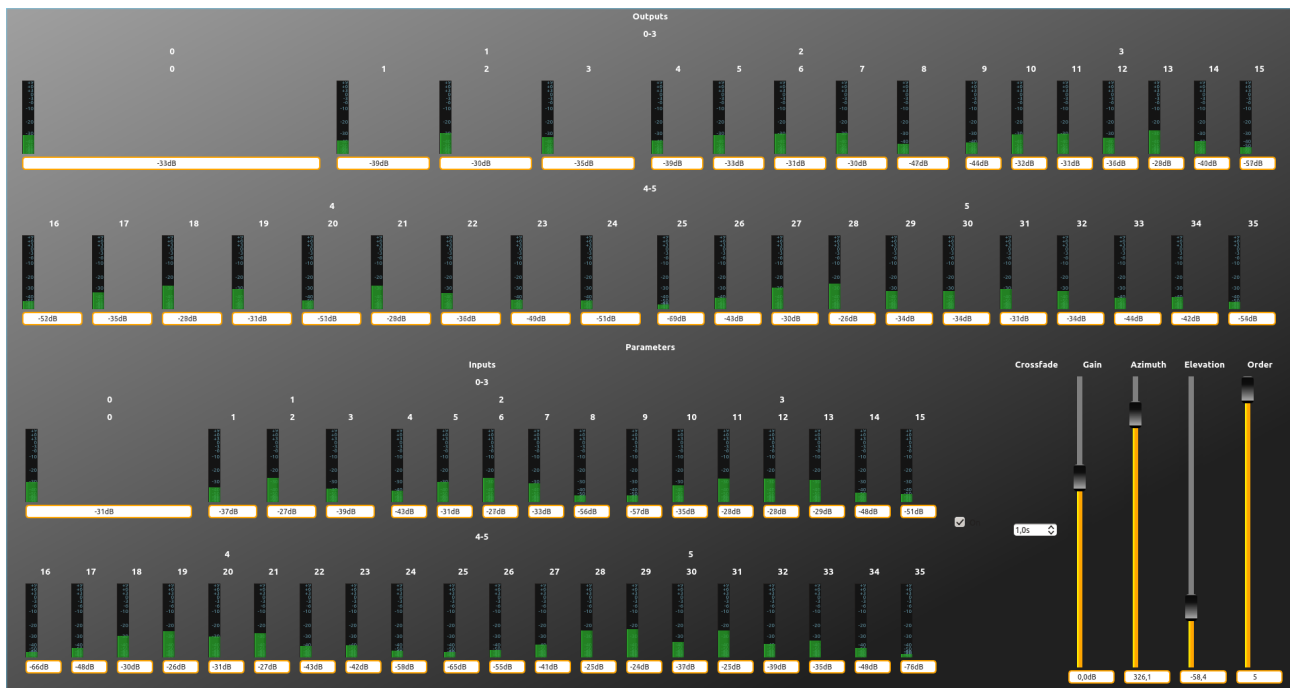


FIGURE 7 – Capture d’écran de l’interface utilisateur sous Linux Jack/Qt pour un filtrage d’une scène sonore ambisonique. Les VU-Mètres permettent de voir les signaux avant et après filtrage en $dBFS$. Les contrôleurs permettent : d’activer l’effet avec un *cross-fade*, d’ajouter un gain global sur les signaux de sortie, de régler l’angle de braquage de l’hyper-cardioïde, et de sélectionner l’ordre de cette dernière.

- [3] Jens MEYER et Gary ELKO : A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 1781–1784, 2002.
- [4] Christian NACHBAR, Franz ZOTTER, Etienne DELEFLIE et Alois SONTACCHI : Ambix - A suggested ambisonics format. *Ambisonics Symposium*, Lexington, 2011.
- [5] Mark A. POLETTI : Three-dimensional surround sound systems based on spherical harmonics. *Journal of the Audio Engineering Society*, 53(11):1004–1025, 2005.
- [6] Michael A. GERZON : Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society*, 33(11):859–871, 1985.
- [7] Philip McCord MORSE et K. Uno INGARD : *Theoretical Acoustics*. McGraw-Hill, 1986.
- [8] Darren B. WARD et Thushara D. ABHAYAPALA : Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on Speech and Audio Processing*, 9(6):697–707, 2001.
- [9] Sébastien MOREAU, Jérôme DANIEL et Stéphanie BERTET : 3d sound field recording with higher order ambisonics-objective measurements and validation of spherical microphone. *Audio Engineering Society Convention 120*, Paris, 2006.
- [10] Pierre LECOMTE, Philippe-Aubert GAUTHIER, Christophe LANGRENNE, Alexandre GARCIA et Alain BERRY : On the use of a Lebedev grid for Ambisonics. *Audio Engineering Society Convention 139*, New York, 2015.
- [11] Boaz RAFAELY : Analysis and design of spherical microphone arrays. *IEEE Transactions on Speech and Audio Processing*, 13(1):135–143, 2005.
- [12] Amitai KORETZ et Boaz RAFAELY : Dolph–Chebyshev beampattern design for spherical arrays. *IEEE Transactions on Signal Processing*, 57(6):2417–2420, 2009.
- [13] Zhiyun LI et Ramani DURAIWAMI : Flexible and Optimal Design of Spherical Microphone Arrays for Beamforming. *IEEE Transactions on Audio, Speech and Language Processing*, 15(2):702–714, 2007.
- [14] Israel COHEN, Jacob BENESTY et Sharon GANNOT : *Speech processing in modern communication : challenges and perspectives*, volume 3. Springer, 2009.
- [15] Heinz TEUTSCH : *Modal Array Signal Processing : Principles and Applications of Acoustic Wavefield Decomposition*. Springer, 2007.
- [16] Ronald H HARDIN et Neil J A SLOANE : McLaren’s improved snub cube and other new spherical designs in three dimensions. *Discrete & Computational Geometry*, 15(4):429–441, 1996.
- [17] Boaz RAFAELY, Barak WEISS et Eitan BACHMAT : Spatial aliasing in spherical microphone arrays. *IEEE Transactions on Signal Processing*, 55(3):1003–1010, 2007.
- [18] George B. ARFKEN et Hans J. WEBER : *Mathematical methods for physicists*. Elsevier, 6ème édition, 2005.
- [19] Pierre LECOMTE et Philippe-Aubert GAUTHIER : Real-Time 3D Ambisonics using Faust, Processing, Pure Data, And OSC. *15th International Conference on Digital Audio Effects (DAFx-15)*, Trondheim, 2015.