



**HAL**  
open science

# LA REPRÉSENTATION INTERMÉDIAIRE ET ABSTRAITE DE L'ESPACE COMME OUTIL DE SPATIALISATION DU SON

Pierre Guillot

► **To cite this version:**

Pierre Guillot. LA REPRÉSENTATION INTERMÉDIAIRE ET ABSTRAITE DE L'ESPACE COMME OUTIL DE SPATIALISATION DU SON: Enjeux et conséquences de l'appropriation musicale de l'ambisonie et des expérimentations dans le domaine des harmoniques sphériques. Musique, musicologie et arts de la scène. Université Paris 8 Vincennes Saint-Denis, 2017. Français. NNT: . tel-01903017

**HAL Id: tel-01903017**

**<https://hal.science/tel-01903017>**

Submitted on 5 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

**UNIVERSITÉ PARIS 8  
VINCENNES – SAINT-DENIS**

**ÉCOLE DOCTORALE ESTHÉTIQUE, SCIENCES ET TECHNOLOGIES DES ARTS**

**Doctorat**

**Musique**

Pierre GUILLOT

LA REPRÉSENTATION INTERMÉDIAIRE ET ABSTRAITE DE L'ESPACE COMME  
OUTIL DE SPATIALISATION DU SON

Enjeux et conséquences de l'appropriation musicale de l'ambisonie et des expérimentations  
dans le domaine des harmoniques sphériques

*Thèse dirigée par Anne SÈDES et Alain BONARDI*

Soutenue le 20 décembre 2017

Jury :

Myriam DESAINTE-CATHERINE (Présidente)

Pierre COUPRIE (Rapporteur)

Sylvain MARCHAND (Rapporteur)

Miller PUCKETTE (Examineur)

Anne SÈDES (Directrice)

Alain BONARDI (Co-directeur)



*À Franquin*



## Résumé

Penser les traitements du son spatialisés en ambisonie permet de mettre en valeur le potentiel musical de la décomposition du champ sonore en harmoniques sphériques, et amène à redéfinir la représentation de l'espace sonore. Cette thèse défend que les représentations abstraites et intermédiaires de l'espace sonore permettent d'élaborer de nouvelles approches originales de la mise en espace du son. Le raisonnement amenant à cette affirmation commence par l'appropriation musicale de l'approche ambisonique. La création de nouveaux traitements de l'espace et du son amène à utiliser de manière originale les signaux associés aux harmoniques sphériques, et à concevoir différemment les relations qui les régissent, ainsi que leur hiérarchisation. La particularité de ces approches expérimentales et les caractéristiques singulières des champs sonores générés, nécessitent de concevoir de nouveaux outils théoriques et pratiques pour leur analyse et leur restitution. Les changements opérés permettent alors de libérer cette approche des enjeux techniques et matériels initiaux en ambisonie. Mais ils permettent surtout de s'émanciper des modèles psychoacoustiques et acoustiques sur lesquels ces techniques reposent originellement. Dans ce contexte, les signaux associés aux harmoniques sphériques ne sont plus nécessairement une représentation rationnelle du champ sonore, mais deviennent une représentation abstraite de l'espace sonore possédant en soi, un potentiel musical. Cette thèse propose alors un nouveau modèle de spatialisation fondé sur une décomposition matricielle de l'espace sonore permettant de valider les hypothèses.

Mots-clés : ambisonie, espace sonore, source sonore, champ diffus, espace matriciel, harmonique sphérique

# Abstract

## THE INTERMEDIATE AND ABSTRACT REPRESENTATION OF SPACE AS A TOOL FOR SOUND SPATIALIZATION

*Challenges and consequences of the musical appropriation of Ambisonics and of experiments in the spherical harmonics domain.*

The creation of sound effects in space with Ambisonics highlights the musical potential of sound field decomposition by spherical harmonics, and redefines the representation of the sound space. This thesis defends that the abstract and intermediate representations of the sound space make it possible to develop new original approaches to sound spatialization. The reasoning that leads to this affirmation begins with the musical appropriation of the ambisonic approach. The creation of new space and sound processing patterns in Ambisonics leads to an original way of using signals associated with spherical harmonics, and to a different conception of the relations between them, and their hierarchization. The specificities of these experimental approaches and the singular characteristics of the sound fields generated call for the design of new theoretical and practical tools, for their analysis and restitution. The performed changes make it possible to free this approach from the initial technical and material issues of Ambisonics. But above all, it emancipates this approach from the psychoacoustic and acoustic models on which ambisonic techniques are originally defined. In this context, the signals associated with spherical harmonics are no longer necessarily a rational representation of the sound field but become an abstract representation of the sound space possessing in itself a musical potential. To validate the hypotheses, this thesis then proposes a new spatialization model based on a matrix decomposition of the sound space.

Keywords: Ambisonics, sound space, sound source, diffuse field, matrix space, spherical harmonic

# Remerciements

Je souhaite remercier Anne Sèdes pour son soutien depuis maintenant neuf ans, pour la qualité de son encadrement, pour m'avoir rassuré quand je doutais de moi-même, et de m'avoir offert de nombreuses opportunités sans lesquelles je ne serais jamais parvenu là où j'en suis aujourd'hui.

Je souhaite également remercier Alain Bonardi pour son accompagnement depuis pas moins de sept ans, pour son soutien dans les moments difficiles, pour ses conseils et pour m'avoir aussi fait confiance dans de nombreux projets.

Parce qu'au-delà d'être mes directeurs et co-directeurs, ils m'ont amené à faire de la recherche, ils m'ont fait confiance, m'ont laissé pleine liberté tout en m'encadrant, et m'ont permis d'enseigner, ce qui est autant une source de plaisir que l'est la recherche. Je leur dois sincèrement beaucoup !

Merci à l'équipe du LABEX Arts H2H de m'avoir offert pendant trois ans un cadre idéal à l'élaboration de ma thèse et au déroulement de mes recherches, et surtout de m'avoir permis de me confronter à l'enseignement.

Merci à mes collègues du département Musique de l'Université Paris 8 pour leur accueil, de m'avoir laissé passer du fond des salles de cours au-devant en tant qu'ATER, pendant un an de plus.

Je remercie aussi toute ma famille, mon père, ma mère, Dominique, Camille, Maxime, Charles pour m'avoir donné goût à l'art et à la culture, pour m'avoir toujours soutenu et aidé dans mes choix. Et pour m'avoir écouté dans les moments difficiles de cette thèse, quand j'avais besoin de maugréer et de râler. Et un grand merci en plus à Dominique de m'avoir autant aidé pendant la rédaction.



Je souhaite aussi remercier mes amis, je ne vais pas tous les citer mais je pense à tous ceux qui ont gentiment fait semblant de m'écouter quand j'essayais d'expliquer mon travail et mon sujet de thèse. Ça m'a fait chaud au cœur !

Un grand merci tout particulier à Eliott, en tant que collègue et ami, il est la seule personne qui puisse me comprendre quand je parle des recherches tout en étant ivre. Et je garderai toujours de très bons souvenirs de notre travail ensemble, surtout à la plateforme un peu délabrée de l'ancienne MSH.

Et forcément un très grand merci à Juliette, parce qu'elle m'a réconforté et soutenu dans les pires moments quand d'autres seraient partis en courant. Avec raison d'ailleurs, parce qu'elle m'engueulait quand je disais « Tant-pis, je claque tout ! », parce qu'elle m'a énormément aidé à l'écriture de cette thèse. Et surtout parce qu'elle considère tout comme moi que Gaston Lagaffe est un génie.

Résumé .....	5
Abstract.....	6
Remerciements.....	7
Introduction.....	14
Partie I – Modèles et outils ambisoniques .....	20
1. État de l’art .....	21
1.1. Origines et émergence de l’approche .....	21
1.1.1. Origines.....	22
1.1.2. Émergence.....	26
1.2. Évolutions et extensions .....	31
1.2.1. Outils d’analyse critiques.....	32
1.2.2. Encodage directionnel .....	34
1.2.3. Adaptation à des systèmes de restitution variés .....	36
1.2.4. Traitements du champ sonore .....	36
1.2.5. Extension aux ordres supérieurs.....	37
1.3. Perspectives et bilan .....	40
1.3.1. Enjeux actuels .....	40
1.3.2. Bilan.....	42
2. Modèles théoriques et pratiques .....	49
2.1. Outils mathématiques et géométriques.....	54
2.1.1. Harmoniques sphériques.....	54
2.1.2. Représentation spatiale .....	57
2.1.3. Harmoniques circulaires .....	58
2.1.4. Décomposition en harmoniques sphériques.....	60
2.2. Construction du modèle ambisonique .....	63
2.2.1. Modèles psychoacoustiques .....	64
2.2.2. Indices de localisation.....	65
2.2.3. Approche fréquentielle .....	69

2.2.4.	Décomposition en harmoniques sphériques du champ sonore .....	73
3.	Outils Logiciels.....	76
3.1.	Développements et livrables.....	77
3.1.1.	Origine et première version.....	77
3.1.2.	Extensions et déploiement multiplateforme.....	79
3.1.3.	Bilan et perspectives.....	81
3.2.	Présentation des outils .....	83
3.2.1.	Les objets de traitement du signal .....	86
3.2.2.	Les interfaces graphiques.....	96
3.2.3.	Vectorisation des traitements .....	103
3.3.	Bilan .....	106
	Partie II – Appropriation musicale de l’ambisonie .....	108
4.	Approche originale.....	109
4.1.	Variation de la résolution angulaire .....	113
4.1.1.	Spécifications .....	113
4.1.2.	Choix.....	114
4.1.3.	Mise en œuvre .....	117
4.2.	Synthèse par décorrélation temporelle .....	121
4.2.1.	Approche aléatoire.....	121
4.2.2.	Concevoir le traitement dans l’espace .....	122
4.2.3.	Extension de l’approche aux ordres élevés .....	123
4.2.4.	Renversement de l’approche .....	127
4.2.5.	Analyse de l’approche .....	129
4.3.	Synthèse par granulation .....	134
4.3.1.	Synthèse granulaire quasi-synchrone .....	134
4.3.2.	Masse sonore et ordre de décomposition.....	136
4.3.3.	Adaptation aux harmoniques sphériques .....	136
4.4.	Synthèse par pondération aléatoire .....	141

4.4.1.	Génération de gains aléatoires .....	141
4.4.2.	L'encodage aléatoire.....	144
4.4.3.	Le facteur d'aléatoire .....	147
4.4.4.	Généralisation de l'approche.....	152
4.5.	Flanging ambisonique .....	154
4.5.1.	Flanging monophonique .....	155
4.5.2.	Flanging stéréophonique.....	156
4.5.3.	Flanging enveloppant .....	157
4.5.4.	Flanging ambisonique.....	158
4.5.5.	Amélioration de l'approche .....	162
4.6.	Bilan .....	164
5.	Adaptation des outils.....	167
5.1.	Analyse du champ sonore.....	168
5.1.1.	Vecteurs énergie et vitesse .....	170
5.1.2.	Valeurs significatives .....	173
5.1.3.	Analyse complète .....	176
5.1.4.	Analyse temporelle.....	180
5.2.	Restitution du champ sonore.....	184
5.2.1.	Optimisations des sources directionnelles .....	185
5.2.2.	Optimisations des champs diffus.....	190
5.2.3.	Contraintes du décodage et solutions .....	193
5.2.4.	Approche pratique de la projection .....	195
5.3.	Analyse de l'approche .....	204
Partie III – Généralisation de l'approche ambisonique .....		209
6.	De nouveaux outils .....	210
6.1.	Spécifications.....	210
6.1.1.	Traitement du signal .....	211
6.1.2.	Interfaces graphiques .....	213

6.1.3.	Outils d'expérimentation .....	214
6.2.	La bibliothèque <i>spam</i> .....	216
6.2.1.	Gestion des arguments .....	217
6.2.2.	Gestion des entrées et sorties spécifiques .....	218
6.2.3.	Gestion des entrées et sorties partagées .....	219
6.3.	Mise en contexte .....	221
6.3.1.	Approche matricielle .....	222
6.3.2.	Approche graphique .....	225
6.3.3.	Bilan.....	229
7.	Introduction à la représentation matricielle .....	231
7.1.	Origines, principes fonctionnels et techniques .....	231
7.1.1.	Origines et l'approche.....	232
7.1.2.	Spécificité de l'approche .....	232
7.1.3.	Les matrices mathématiques .....	235
7.1.4.	L'espace matriciel .....	236
7.1.5.	Traitements audionumériques par <i>soxel</i> .....	237
7.1.6.	Fonctionnement et aspects génériques des outils .....	238
7.2.	Soxel Art.....	240
7.2.1.	Principes.....	241
7.2.2.	Approche manuelle de la synthèse sonore .....	242
7.2.3.	Approche dynamique de la synthèse sonore .....	243
7.2.4.	Manipulation de matrices.....	245
7.2.5.	Visualisation et édition graphique .....	247
7.2.6.	Édition par motifs .....	249
7.3.	Utilisation, analyse et perspectives .....	253
7.3.1.	Gestion temporelle des motifs.....	253
7.3.2.	Gestion temporelle des matrices .....	257
7.3.3.	Bilan et perspectives .....	259

8.	Développement de la représentation matricielle .....	260
8.1.	Projection des soxels .....	260
8.1.1.	Choix conceptuels .....	262
8.1.2.	Choix techniques .....	266
8.1.3.	Mise en œuvre .....	271
8.2.	Filtres spatiaux .....	274
8.2.1.	Produit de convolution.....	275
8.2.2.	Mise en œuvre .....	277
8.2.3.	Noyaux et opérations .....	280
8.3.	Synthèse de l'espace sonore .....	292
8.3.1.	Le pinceau de l'espace sonore.....	293
8.3.2.	Le pot de peinture de l'espace sonore .....	294
8.3.3.	L'aérosol de l'espace sonore.....	295
8.4.	Perspectives .....	297
	Conclusion .....	299
	Bibliographie.....	306

# Introduction

Apparue dans le courant des années 1970 dans le cadre des programmes de recherches du *National Research Development Corporation* (N.R.D.C) au Royaume-Uni – notamment sous l’impulsion de M. A. Gerzon [Gerzon, 1975c], et P. B. Fellgett [Fellgett, 1975] – l’ambisonie est un ensemble de techniques permettant l’enregistrement, la synthèse, la transformation et la restitution de champs sonores. Ces techniques visent à restituer aux oreilles de l’auditeur, les effets directionnels d’un champ sonore pouvant résulter d’un enregistrement, via l’utilisation d’un système de microphones spécifiquement destiné à cet usage ou résultant de la synthèse artificielle de sources sonores directionnelles via l’opération dite d’encodage. Ces développements se situent alors dans la continuité des techniques stéréophoniques. Mais ils se placent surtout en réponse face aux techniques quadriphoniques émergentes<sup>1</sup>, qui proposent des solutions pratiques mais non-optimales, aux dispositifs de haut-parleurs enveloppant l’auditeur pour la restitution de champs sonores en deux ou trois dimensions. L’objectif de l’approche ambisonique est alors de corriger les défauts de ces techniques de spatialisation en permettant une reproduction des champs sonores qui soit à la fois homogène, précise, fiable et reproductible. La réponse apportée par l’ambisonie se trouve notamment dans le format-B. Ce format correspond à une représentation intermédiaire du champ sonore, c’est à dire indépendante de tous systèmes de restitution. Ce qui permet donc d’adapter la génération des signaux destinés haut-parleurs, pour différentes configurations via l’opération de décodage. Mais au-delà, cette représentation équivaut à une décomposition en harmoniques sphériques d’ordre 1 du champ sonore, tel qu’il est défini localement aux oreilles de l’auditeur sous la forme d’un gradient de pression. De ce fait, elle offre les outils mathématiques pour concevoir les opérations de décodage, ainsi que des informations directionnelles claires pour optimiser ces opérations selon différents critères, notamment psychoacoustiques car liés aux mécanismes humain de localisation des sons.

Bien qu’au début, l’engouement des industries sonores pour cet ensemble de techniques ait été relativement restreint, au cours des décennies suivantes, l’ambisonie a continué d’être au centre de nombreux travaux, que ce soit dans le domaine de l’acoustique avec notamment les recherches de J. Daniel ou de F. Zotter ou dans des approches plus musicales avec, par exemple les travaux de D. G. Malham ou G. Wakefield. Aussi, les techniques ambisoniques n’ont eu de

---

<sup>1</sup> Dans ce contexte, les techniques quadriphoniques ne sont pas clairement définies sur un plan formel et sont prises au sens de quatre canaux ou quatre sources. Ce que P. B. Fellgett définit comme *quadrifrontal* [Fellgett, 1975], p. 2.

cesse de s'améliorer et de s'étendre, intégrant aujourd'hui des connaissances venant de multiples disciplines telles que les mathématiques, l'acoustique, la psychoacoustique, l'ingénierie, la musique, etc. L'ambisonie définit d'ailleurs, selon les domaines et les champs d'applications, des objectifs variés qui ont amené à créer des outils spécifiques et des approches tant particulières qu'hétéroclites. C'est dans cette multitude de pratiques originales que le CICM a entrepris depuis 2012 – dans le cadre des programmes de recherches du LABEX Arts H2H – différents projets sous la direction scientifique d'Anne Sèdes et d'Alain Bonardi, autour de l'approche musicale des techniques de spatialisation ambisoniques. Ces recherches ont amené à l'élaboration de la bibliothèque HOA (*High Order Ambisonics*) – sous la forme d'un ensemble de code C++ – initialement développée par Julien Colafrancesco, Elliott Paris et moi-même, et dont les mises en œuvre ont été réalisées pour de nombreuses plateformes logicielles<sup>2</sup>. C'est dans le cadre de ces développements et de ces recherches que mon travail de thèse commence. L'enjeu premier de ces outils logiciels est de faciliter la prise en main du modèle ambisonique et de permettre l'élaboration d'opérations de synthèse et de traitement de champs sonores dans le domaine des harmoniques sphériques. Cet enjeu répond aux recherches antérieures de Graham Wakefield [Wakefield, 2006], dans lesquelles il explore et envisage les possibilités musicales de la synthèse des signaux associés aux harmoniques sphériques, selon différentes approches originales telles que l'utilisation de filtres en peigne ou encore la modulation de fréquence. Ces expérimentations ne font qu'esquisser le réel potentiel de telles approches. Un des axes majeurs de mon travail a alors été de poursuivre ces expérimentations, et de tenter d'approfondir ces recherches.

Afin de répondre à ces exigences, la bibliothèque HOA offre, notamment à travers ses mises en œuvre pour les logiciels Max et Pure Data, de nombreux modules de traitements du signal, des interfaces graphiques de contrôle et de représentations du champ sonore. Mais mon apport le plus original a été d'élaborer un système permettant l'instanciation dynamique d'un patch pour chaque signal associé à un harmonique<sup>3</sup>. Ce système permet de mettre en œuvre des opérations en fonction des caractéristiques spatiales des signaux associés aux harmoniques sphériques et de leurs manières de contribuer respectivement et globalement au champ sonore. En facilitant l'accès au terrain d'expérimentation que sont les signaux associés aux harmoniques sphériques,

---

<sup>2</sup> Les contributions au fil des projets ont été nombreuses et résultent de travaux plus ou moins importants de plusieurs personnes. Je tenterai au mieux de tous les citer au fil de cette thèse.

<sup>3</sup> Ce système qui est sous la forme d'une série d'objets gravitant autour de l'outil central *hoa.2/3d.process~* a, par la suite, été particulièrement bien intégré au logiciel Max par Elliott Paris, tandis que je m'occupais de son développement pour le logiciel Pure Data.



ce système m'a permis de réaliser un grand nombre d'expériences et de concevoir de nouvelles approches de l'espace et du son. Certaines de ces mises en œuvre font à présent parties de la bibliothèque HOA, et ont pu être prises en main par les compositeurs et musiciens. Comme c'est le cas par exemple pour la synthèse de champ sonore par granulation, ou encore par décorrélation temporelle des signaux associés aux harmoniques. D'autres quant-à-elles, sont restées confinées au cadre expérimental ou purement didactique, comme la synthèse chaotique de champ sonore ou le *flanging* ambisonique. Mais l'ensemble de ces travaux m'a amené à concevoir différemment le modèle ambisonique et à porter un regard autre sur la représentation de l'espace, via la décomposition en harmoniques sphériques.

Les différentes expérimentations réalisées dans le cadre de ces travaux de thèse ont permis de mettre en place des approches originales de la création et la transformation de champ sonore via le modèle ambisonique. Et ces approches ont été réalisées selon des méthodologies variées afin de répondre à différents enjeux. Pour la synthèse de champ sonore par décorrélation temporelle par exemple, l'enjeu principal a été de mettre en lumière les caractéristiques spatiales des signaux associés aux harmoniques sphériques. Pour cela, la méthodologie employée a consisté à prendre comme point de départ cette opération audionumérique classique. Puis il s'agissait de la mettre en œuvre directement dans le domaine des harmoniques sphériques pour, par la suite, analyser le champ sonore synthétisé et critiquer les résultats. Au fur et à mesure des expérimentations – en passant notamment par la synthèse de champ sonore par décorrélation et la synthèse de champ sonore chaotique – les enjeux se sont alors transformés afin de prendre de plus en plus en compte les caractéristiques spatiales des harmoniques sphériques, ainsi que les propriétés de la décomposition du champ sonore lors de la mise en œuvre des opérations<sup>4</sup>. Et ce cheminement expérimental s'est conclu par la création d'un *flanging* ambisonique. La méthodologie a consisté alors à définir ce que doit être l'opération originale à mettre en œuvre sur les signaux associés aux harmoniques sphériques, pour concevoir un champ sonore possédant les caractéristiques sonores du *flanging* mis en espace par l'ambisonie. En somme, il s'agissait de définir comment, de manière analogue à l'extension stéréophonique de la technique monophonique de *flanging*, il est possible d'étendre une approche à l'ambisonie. La mise en œuvre de ces différentes approches a donc permis de proposer des méthodologies pouvant être généralisées à d'autres traitements. Mais elle a surtout

---

<sup>4</sup> D'autres approches ont, bien sûr, été mises en œuvre, comme des modules de synthèse ou de transformation de champ sonore reposant sur la modulation en anneau, la distorsion et la saturation, le *phasing*, etc. Mais les opérations énoncées dans cette thèse semblent suffire à la compréhension du raisonnement proposé et elles offrent un relativement bon aperçu du potentiel musical de l'ambisonie.

fait émerger des champs sonores originaux où les propriétés spatiales et les caractéristiques intrinsèques des sons se mélangent.

Ces champs sonores ne répondent alors plus à la même définition que ceux pour lesquels les techniques ambisoniques ont été élaborées. L'objectif est donc d'essayer de caractériser ces nouveaux champs sonores, et d'en définir les propriétés importantes à préserver. Les différentes mises en œuvre et expérimentations révèlent que les résultats ne ressemblent en rien à ce qui se trouve dans la réalité, et d'ailleurs les approches ne visent aucunement à simuler la réalité. Par ailleurs, un des aspects récurrents du champ sonore est son caractère diffus ou chaotique, en opposition avec un caractère directionnel. De ce fait, la manière de considérer la représentation du champ sonore change. En effet, les relations qui régissent les signaux associés aux harmoniques sphériques – permettant de préserver ou de consolider la directionnalité des sons – sont remplacées par un système moins hiérarchique, où chaque signal associé à un harmonique possède en soi un intérêt musical propre dans le champ sonore. Les enjeux consistant à offrir une restitution homogène, précise, fiable et reproductible des effets directionnels du champ sonore paraissent alors obsolètes ou du moins non satisfaisantes. Il est donc nécessaire de proposer de nouveaux outils théoriques et pratiques permettant de répondre aux problèmes de l'analyse et de la restitution de ces champs sonores. Pour cela, je propose de reprendre les outils et les concepts clés proposés originalement par M. A. Gerzon – notamment dans sa métathéorie générale de la localisation auditive<sup>5</sup> [Gerzon, 1992a] – afin de les adapter aux nouvelles approches de la mise en espace du son, proposées dans le cadre de cette thèse. Réaliser cette démarche amène à élaborer des outils permettant de critiquer les opérations d'optimisation et de décodage, au regard des nouvelles problématiques soulevées. Cela permet aussi de proposer des alternatives plus adaptées à la restitution des champs sonores diffus et chaotiques.

Si les nouveaux enjeux diffèrent de ceux auxquels tentent de répondre originellement l'approche ambisonique. Si la théorie sur laquelle est construit le modèle ambisonique ne suffit plus. Et si les opérations et les outils pratiques des techniques ambisoniques sont renouvelés. Alors que reste-t-il de la spatialisation du son ambisonique au sein de ces approches originales ? Ou plutôt qu'est ce qui caractérise cette nouvelle approche de la mise en espace du son ? Mon raisonnement est que les expérimentations et les mises en œuvre sont réalisées sur les signaux

---

<sup>5</sup> *General Metatheory of Auditory Localization.*

associés aux harmoniques sphériques. Cependant, étant donné que le modèle qui a défini que ces signaux offrent une représentation locale du champ sonore n'est plus valide – ou du moins ne suffit plus à embrasser toutes les approches qui en découlent – alors, dans ce contexte, il n'est plus nécessaire d'associer ces signaux au champ sonore. Néanmoins, il s'agit de préserver leur association au modèle mathématique et donc à leurs représentations et leurs caractéristiques spatiales. Aussi, les signaux associés aux harmoniques sphériques offrent une représentation abstraite de l'espace sonore. Et comme celle-ci préserve son indépendance vis-à-vis du dispositif de restitution ou d'enregistrement, elle reste une représentation intermédiaire. Partant de cette analyse, est-il possible d'envisager d'autres modèles de représentations de l'espace, fondés sur une généralisation des expérimentations établies en ambisonie et permettant de concevoir de nouvelles approches de la mise en espace du son ?

*Cette thèse montre comment l'appropriation musicale de l'ambisonie a permis de définir la représentation abstraite et intermédiaire de l'espace comme un outil de spatialisation du son ouvrant de nouvelles perspectives pour la création d'espaces sonores.*

Pour répondre à la problématique, ce travail s'articule autour de trois axes. Dans la première partie de cette thèse, l'enjeu est de présenter le modèle et les outils ambisoniques. L'objectif est double. Il s'agit d'une part de présenter les outils et les connaissances nécessaires à l'appropriation musicale de l'ambisonie et à la mise en œuvre d'opérations dans le domaine des harmoniques sphériques. D'autre part, l'objectif est de dégager les fondements des techniques ambisoniques et leurs implications, afin de concevoir comment les travaux originaux présentés dans cette thèse les remettent en question et les redéfinissent. Pour cela, le premier chapitre reviendra sur les grandes étapes de l'élaboration de ces techniques. Puis le deuxième portera un regard sur les modèles mathématiques, acoustiques et psychoacoustiques sur lesquels reposent ces techniques ambisoniques. Enfin, dans un troisième chapitre, un aperçu détaillé de la bibliothèque HOA et de son fonctionnement sera réalisé.

Dans la seconde partie, cette thèse présentera comment il est possible de s'approprier le modèle mathématique des harmoniques sphériques et la représentation de l'espace qui y est attachée, afin de concevoir et de mettre en œuvre des traitements originaux du son et de l'espace. Pour cela, dans le quatrième chapitre, les expérimentations suggérées précédemment seront

présentées en détails. Elles tenteront de mettre en avant le potentiel musical de l'ambisonie, tout en exposant les mécanismes en jeu dans l'élaboration des traitements de l'espace et du son, dans le domaine des harmoniques sphériques. Dans le cinquième chapitre seront alors présentés les changements et les conséquences de ces expérimentations – notamment liés à la décomposition du champ sonore en harmoniques sphériques – vis-à-vis des enjeux intrinsèques du système de spatialisation ambisonique. Les nouveaux outils et opérations permettant de répondre aux problématiques actuelles seront donc exposés et il sera alors possible de confronter le modèle ambisonique aux nouvelles problématiques. Enfin, de nouveaux critères théoriques et pratiques, fondés sur une représentation abstraite et intermédiaire de l'espace, seront proposés. De nouvelles perspectives de spatialisation du son seront alors envisagées.

A partir de ces conclusions, la troisième partie s'attèlera à valider cette thèse sur un plan pratique, en proposant l'élaboration d'un système de spatialisation fondé sur une représentation originale, abstraite et intermédiaire de l'espace. Pour cela, dans le sixième chapitre, un ensemble d'outils logiciels sera présenté. L'enjeu sera d'offrir un système généralisant l'approche proposée par la bibliothèque HOA, qui permette de mettre en œuvre et d'expérimenter de nouvelles approches de la spatialisation du son, fondées sur de multiples représentations. Sur ces outils reposera alors la création d'une première approche construite sur une représentation matricielle de l'espace sonore. Les origines, les concepts et les enjeux originaux liés à cette représentation de l'espace seront présentés dans le septième chapitre, notamment via la proposition d'un ensemble de pratiques originales de l'espace, sous le nom de Soxel Art. Enfin dans un huitième chapitre, l'approche sera développée pour permettre la restitution de l'espace sonore sur un système de haut-parleurs. Mais aussi afin de considérer les avantages mathématiques qu'offre cette représentation – notamment via les filtres spatiaux – et d'envisager des approches systématiques de synthèse de champ sonore.

# Partie I – Modèles et outils ambisoniques

Les travaux originaux présentés dans le cadre de cette thèse abordent la question de la spatialisation du son, au travers des techniques ambisoniques. Ces techniques sont les outils résultant de l'élaboration d'un modèle de la spatialisation du son appelé ambisonie, qui recoupe plusieurs domaines. De l'ingénierie du son aux sciences de l'informatique en passant par les mathématiques, l'acoustique et la psychoacoustique, l'ambisonie forme un réseau complexe où les connaissances s'entremêlent autour d'un ensemble d'idées majeures. L'objectif principal de cette partie est d'éclairer les principes et les théories sur lesquels repose l'ambisonie. L'enjeu est de pouvoir, par la suite, confronter les résultats apportés par les recherches et les expérimentations, au regard des idées présentées dans cette première partie de thèse. Afin d'en apprécier dans un premier temps la portée puis afin d'élaborer un raisonnement qui aboutira en de nouvelles recherches et de nombreuses perspectives.

La démarche scientifique proposée dans cette partie consiste à discerner dans chacun de ces domaines auxquels touche l'ambisonie, les éléments fondamentaux qui ont amené à concevoir son modèle. Le premier chapitre commence par présenter les aspects qui ont attiré à l'ingénierie du son, afin de définir les spécifications techniques et fonctionnelles sur lesquelles l'ambisonie a été élaborée. Pour cela, l'approche propose notamment de revenir de manière chronologique sur les problèmes techniques, matériels et conceptuels, en partant des origines de cette approche jusqu'à ses développements récents. Dans un deuxième chapitre, sont présentés les modèles théoriques et pratiques qui ont attirés aux domaines des mathématiques, avec la décomposition en harmoniques sphériques. Sont aussi présentés dans ce chapitre, les modèles liés aux domaines de l'acoustique et surtout de la psychoacoustique, avec les théories de localisation des sons. Enfin, le dernier chapitre de cette partie revient sur les outils logiciels utilisés dans cette thèse, en présentant leurs fonctionnalités et leurs mises en œuvre.

Malgré la relative complexité des éléments abordés dans le cadre de cette thèse, les travaux se veulent compréhensibles pour des musiciens et des compositeurs. Aussi, les présentations reviennent sur certaines notions qui peuvent paraître élémentaires chez les mathématiciens et acousticiens en tentant d'y apporter une approche plus adaptée aux néophytes. Enfin, les enjeux spécifiques à chaque domaine, dans le cadre de cette thèse, seront présentés en début de chaque chapitre.

# 1. État de l'art

Ce chapitre revient de manière chronologique sur les enjeux, les techniques, les pratiques et de manière plus générale les notions importantes liées aux développements de l'ambisonie. L'enjeu est de mieux envisager les standards et les usages des techniques ambisoniques actuelles pour permettre la compréhension des travaux originaux – présentés dans la suite de cette thèse – notamment en ce qui concerne l'appropriation musicale du domaine des harmoniques sphériques et ses répercussions. Aux vues des développements et recherches en ambisonie engagés dès les années 1970 et émanant de nombreux chercheurs, il apparaît redondant et peu pertinent d'énumérer toutes les contributions de tout un chacun. De fait, l'une des principales complexités des techniques ambisoniques réside en la confusion qu'il peut y avoir entre les différents formats et standards, et les différentes normalisations. Ceux-ci peuvent aujourd'hui paraître totalement obsolètes et répondre à des problématiques caduques mais restent cependant présent dans la littérature et dans les usages. Le choix a été de procéder à une sélection afin de rendre l'énoncé intelligible, et de mettre en avant les sujets considérés comme les plus pertinents, au regard des recherches de cette thèse, en conservant autant que possible la linéarité chronologique.

Enfin, avant d'introduire les principes théoriques et fonctionnels des techniques ambisoniques, il est nécessaire de présenter les origines et les objectifs de cette approche. L'objectif n'est pas de soumettre l'ensemble des travaux et recherches qui ont amené à la création des techniques ambisoniques<sup>6</sup>, mais de présenter les intentions et les méthodologies employées de même que les contraintes rencontrées afin de contextualiser les approches. Et de pouvoir, par la suite, mettre en perspective le caractère particulier et les spécificités de l'appropriation musicale de cette technique de spatialisation du son ainsi que de l'utilisation de la représentation de l'espace qui lui est associée.

## 1.1. Origines et émergence de l'approche

Les techniques ambisoniques ne sont bien entendu pas apparues d'elles-mêmes au milieu de nulle part. Elles sont le résultat d'expérimentations variées qui sont le fruit de recherches menées notamment dans le domaine de l'industrie musicale, audiovisuelle et sonore, comme le montrent les nombreux brevets, dont les exemples sont [Blumlein, 1933], [Craven & al. 1977]

---

<sup>6</sup> Pour un introduction strictement chronologique, plus objective et énumérative, des travaux et avancées dans le domaine de la spatialisation du son, le lecteur peut se référer à [Davis, 2003] et [Steinke, 1996].

ou [Gerzon, 1981], les rapports de recherches [Gaskell, 1977] et les magazines spécialisés comme l'*Audio Magazine* [Stark, 1971], [Feldman, 1971]. Cependant bien que les objectifs soient moindres, et les contraintes matérielles restrictives, les enjeux centraux de la spatialisation sonore apparaissent alors et des premières méthodologies émergent, liées à l'industrie sonore.

### 1.1.1. Origines

Dans les années 1930, apparaît le désir d'améliorer l'expérience de l'auditeur, lors de la reproduction musicale par un système de haut-parleurs – notamment la musique orchestrale – en prenant en compte l'espace au moment de l'enregistrement et de la transmission [Fletcher, 1934]. A ce propos, J. C. Steinberg et W. B. Snow affirment que « ce caractère spatial du son [la relation spatiale des instruments de l'orchestre] donne à la musique un sentiment de profondeur et une ampleur, qui pour une parfaite reproduction, doit être préservée. En d'autres termes, les sons devraient être reproduits dans une *perspective auditive* réelle »<sup>7</sup>. L'objectif présenté consiste à pouvoir reproduire une scène sonore à destination d'un auditoire. Et l'enjeu est de reconstituer les positions des sources sonores avec comme nécessité que celles-ci correspondent exactement à celles utilisées lors de l'enregistrement. En écartant bien sûr tous les artefacts pouvant engendrer des variations ou des modifications de la scène sonore. Ce but, qui émerge aux débuts de travaux sur la spatialisation sonore, perdurera et sera repris par la suite dans de nombreuses approches de l'espace sonore.

Sur le plan des spécificités techniques, les auteurs définissent par la suite que cet objectif peut être atteint soit via la reproduction « dans les oreilles d'un auditeur distant, au moyen de récepteurs placés sur la tête, les copies exactes des vibrations du son telles qu'elles auraient été s'il les avait écoutées directement »<sup>8</sup>, soit en utilisant « des haut-parleurs [...] pour reproduire dans une salle distante, une exacte copie du motif de la vibration du son qui existe dans la salle originale »<sup>9</sup>. Ces deux modes opératoires définissant la manière d'aborder et concevoir la reproduction de scènes sonores pour un auditeur sont aujourd'hui toujours d'actualité. La

---

<sup>7</sup> [Steinberg & al., 1934], p. 12, traduit de l'anglais : « *This spatial character of the sounds gives to the music a sense of depth and of extensiveness, and for perfect reproduction should be preserved. In other words, the sounds should be reproduced in true auditory perspective.* »

<sup>8</sup> [Steinberg & al., 1934], p. 12, traduit de l'anglais : « *One is binaural reproduction which aims to reproduce in a distant listener's ears, by means of head receivers, exact copies of the sound vibrations that would exist in his ears if he were listening directly.* »

<sup>9</sup> [Steinberg & al., 1934], p. 12, traduit de l'anglais : « *[...] uses loud speakers and aims to reproduce in a distant hall an exact copy of the pattern of sound vibration that exists in the original hall.* »

première approche dite binaurale<sup>10</sup> consiste à recréer les informations sonores aux oreilles – les signaux binauraux. Au moyen d'un casque ou d'écouteurs, la technique est dite de *présentation binaurale* et au moyen d'un système de haut-parleurs, celle-ci est dite de *présentation transaurale*. La deuxième approche vise en quelque sorte à recréer les sources des phénomènes psychoacoustiques. Elle se rapproche des techniques dites holophoniques telles que définies par [Jessel, 1973], [Berkhout, 1988], équivalent acoustique de l'holographie, et amenant à la *Wave Field Synthesis*<sup>11</sup> (WFS) [Berkhout & al., 1993], [Boone & al., 1995]. Que ce soit l'une ou l'autre des approches, l'enjeu consiste à définir des techniques et des outils permettant de recréer des *copies exactes* des phénomènes acoustiques ou des phénomènes psychoacoustiques tels qu'ils puissent être mesurés via des instruments ou représentés via des modèles mathématiques.

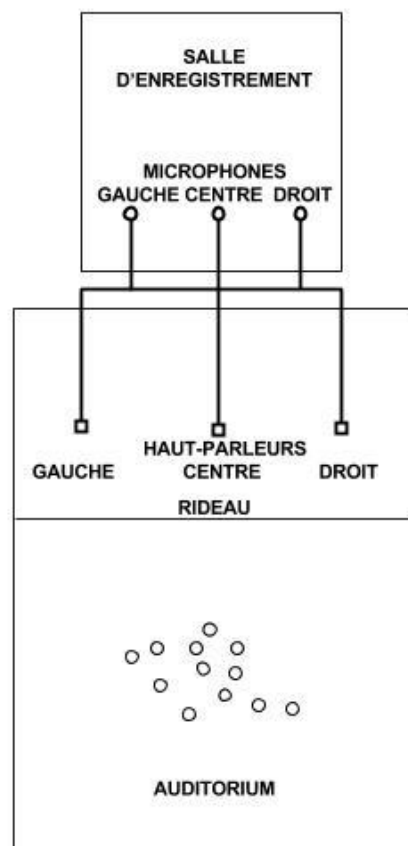


Figure 1.1. Représentation d'un système de tests pour la localisation de sons lors de la reproduction d'une scène sonore frontale de type orchestrale et inspiré du système proposé par J. C. Steinberg et W. B. Snow où le positionnement et la distance entre les microphones et les haut-parleurs ainsi que le matricage des canaux peut varier et par lequel il est possible de comparer la scène sonore virtuelle telle que perçue par l'auditoire avec la scène sonore réelle.

<sup>10</sup> La perception de la direction du son par les indices binauraux est présentée de manière approfondie par H. Møller dans [Møller, 1992] et par B. C. Moore dans le chapitre *Space Perception* [Moore, 2013], p. 245-281.

<sup>11</sup> La synthèse de champ acoustique.



Confronter cet enjeu aux restrictions techniques et matérielles amène à le considérer sous un angle différent. Par exemple, la reproduction exacte des ondes acoustiques d'une approche de type holophonique nécessite une discrétisation continue de l'espace, que ce soit lors de la captation ou de la restitution de la scène sonore. Cela implique donc une infinité de microphones et de haut-parleurs aux dimensions infinitésimales<sup>12</sup>. Dès lors, l'objectif apparaît en soi irréalisable. Les contraintes matérielles doivent donc être contournées par le biais d'une simplification des phénomènes acoustiques. Dans les conditions énoncées par J. C. Steinberg et W. B. Snow, l'approximation – considérée *in fine* comme bonne – est réalisée à l'aide d'ensembles de deux ou trois microphones et deux ou trois haut-parleurs<sup>13</sup> [

Figure 1.1]. L'enjeu n'est alors plus foncièrement de reproduire à l'identique les vibrations du son, mais de définir des systèmes et des techniques optimales pour restituer une scène sonore virtuelle cohérente avec la scène sonore réelle du point de vue de l'auditeur, notamment via des jeux sur la configuration des systèmes d'enregistrement, de reproduction mais aussi des systèmes de matriçage. Il ne s'agit nullement par cette affirmation de discréditer cette approche qui, pour de multiples raisons, reste valide. L'idée étant de jouer sur certaines caractéristiques sonores en relation avec certains éléments clefs des mécanismes de perception spatiale du son chez l'auditeur. Un exemple marquant qui découle de cette approche est le panoramique d'amplitude tel que défini par [Pulkki, 1997], [Pulkki & al., 2001] ou [Roads, 1996], p. 558-460. Cette approche fondée sur un jeu de gain et d'intensité entre des canaux audionumériques, permet d'obtenir des images spatiales pertinentes alors que l'approche ignore nombre de phénomènes acoustiques<sup>14</sup>. La mise en œuvre de techniques de spatialisation du son – présentées précédemment – nécessite de définir un cadre formel et d'adopter un point de vue tel que celui « omniscient » de l'approche holophonique ou que celui de l'auditeur dans une approche binaurale. Ces techniques font le choix de simplifier la conception de l'espace sonore et sa perception par l'auditeur, en faisant abstraction de certains phénomènes physiques en jeu, ou certaines contraintes matérielles. Ce type de procédé est d'ailleurs le propre même des sciences physiques<sup>15</sup> et une technique de spatialisation ne peut offrir qu'une approximation

---

<sup>12</sup> Evidemment d'autres contraintes techniques et matériels existent et il en est de même pour les techniques dites binaurales [Gerzon, 1975b], [Cooper & al., 1989], [Larcher, 2001], [Busson, 2006].

<sup>13</sup> [Steinberg & al., 1934], p. 12 : « Des configurations bien moins idéales, telles que composées seulement de deux ensembles de microphones-haut-parleurs ont permis d'offrir une bonne perspective auditive. », traduit de l'anglais : « *Far less ideal arrangements, consisting of as few as 2 microphone-loudspeaker sets, have been found to give good auditory perspective* ».

<sup>14</sup> Les approches présentées ne prennent pas en compte les différences entre les mécanismes de perception pour les basses et hautes fréquences présentées par la suite [Makita, 1960]

<sup>15</sup> Les sciences physiques partant de l'observation de monde et Einstein disait à ce propos « La conviction de la « vérité » des propositions géométrique en ce sens repose naturellement sur des expériences assez imparfaites. » [Einstein, 2001], p.14.

d'une scène sonore. L'élément important est alors l'introduction du terme de scène virtuelle<sup>16</sup> qui est relative à la scène sonore reproduite par le système de haut-parleurs, en opposition à la scène réelle qui est enregistrée par le système de microphones. Et dans ces premières explorations, une bonne technique – et possiblement un bon système – de spatialisation est alors définie lorsque pour l'auditeur, la scène virtuelle est la plus en adéquation possible avec la scène réelle.

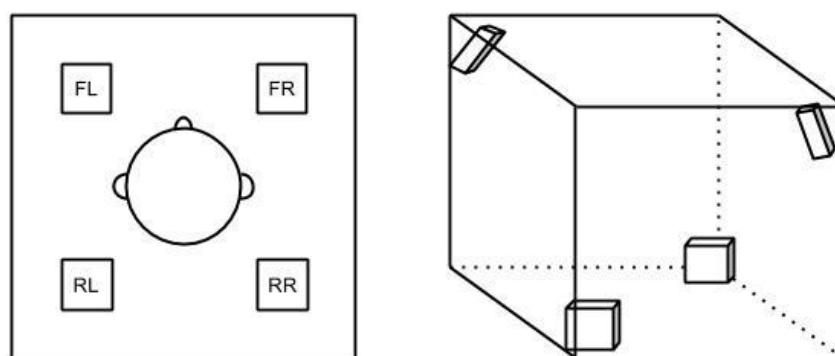


Figure 1.2. Représentation d'un système de haut-parleurs quadraphonique à gauche pour une reproduction sonore sur le plan horizontal seulement et d'un système de haut-parleurs tétraédral tel que défini par M. A. Gerzon [Gerzon, 1973] pour une reproduction sonore périphonique à droite intégrant la dimension verticale.

De ces premières démarches, émergent les techniques *stéréophoniques* visant à reproduire une scène sonore généralement de manière frontale [Blumlein, 1933], [de Boer, 1940], [Fletcher, 1941], [Gerzon, 1976], [Chowning, 1977], [Bennett & al., 1985] et qui seront commercialisées dans les années 1950 [Gerzon, 1974], [Davis, 2003]. Par la suite, ces techniques élargissent les dispositifs, afin d'offrir une scène sonore immersive enveloppant l'auditoire [Figure 1.2]. Les approches se focalisent d'abord sur le plan horizontal et la restitution dite *pantophonique* – notamment avec la *quadriphonie*<sup>17</sup> – dont les présents ouvrages couvrent une large étendue des approches et des outils disponibles à l'époque [Gerzon, 1970a], [Bauer & al., 1971], [Eargle, 1971], [Feldman, 1971], [Scheiber, 1971a], [Scheiber, 1971b], [Stark, 1971], [Itoh, 1971], [Cooper & al., 1972]. Puis les systèmes introduisent la dimension verticale et la restitution dite *périphonique* [Cooper, 1970], [Gerzon, 1970b], [Gerzon, 1971a,b,c], [Gerzon, 1973] et

<sup>16</sup> [Steinberg & al., 1934], p. 12, traduit de l'anglais : « *virtual stage* ».

<sup>17</sup> Ce terme n'est pas précisément défini et les différents usages ne sont pas forcément cohérents [Felgett, 1975] mais ils impliquent généralement quatre haut-parleurs tels que représentés en [Figure 1.2].

amènent aux techniques ambisoniques<sup>18</sup> [Felgett, 1974], [Gerzon, 1974], [Felgett, 1975], [Gerzon, 1975c] qui ont pour but de « reproduire à l'oreille de l'auditeur avec précision, fiabilité et répétition, le champ sonore directionnel créé dans le studio, soit par un ensemble de microphones permettant l'enregistrement de champ sonore, soit par des dispositifs d'encodage de directionnalité artificielle (*pan-pots*) ou des dispositifs artificiels de réverbération immersifs »<sup>19</sup>. L'ambisonie vise un objectif similaire à ces techniques et approches parentes. L'objet de ces recherches est de pouvoir reproduire une scène sonore virtuelle la plus similaire possible à une scène sonore réelle ou artificielle – créée via un encodage de directionnalité artificielle. L'idée est alors de définir un système permettant de répondre au mieux ce problème [Gerzon, 1977].

### 1.1.2. Émergence

Le projet ambisonique est originalement développé au National Research Development Corporation (N.R.D.C.) du Royaume-Uni notamment par M. A. Gerzon et P. B. Fellgett. Dans la continuité des différentes techniques et expérimentations dont ils découlent, les objectifs principaux des développements initiaux des systèmes ambisoniques restent de pouvoir enregistrer, transmettre et reproduire de façon répétée, un effet directionnel précis d'une scène sonore enveloppant l'auditeur [Gerzon, 1975c]. En 1972, D. H. Cooper et T. Shiga proposent une méthode de décomposition du champ sonore bidimensionnelle par les harmoniques azimutales<sup>20</sup> [Cooper & al., 1972] permettant notamment de décrire toutes les directions de manière homogène et offrant de nombreux avantages, tels qu'une résolution de la représentation spatiale proportionnelle au nombre de canaux, et adaptable à de nombreux systèmes existant. M. A. Gerzon étend l'approche aux systèmes périphoniques et propose une approche théorique fondée sur une représentation tridimensionnelle du champ sonore par les harmoniques sphériques [Gerzon, 1973]. En 1975, M. A. Gerzon et P. B. Fellgett proposent une mise en œuvre de cette approche, permettant notamment de répondre aux problématiques évoquées

---

<sup>18</sup> Evidemment, d'autres recherches et développements se déroulent parallèlement amenant aux systèmes 5.1 ou 7.1 et aux formats tels que Dolby ou DTS qui servent principalement des enjeux commerciaux et notamment pour des productions cinématographiques tels *Star Wars* en 1977, *Superman* en 1978 ou encore *Return of the Jedi* en 1982.

<sup>19</sup> [Gerzon, 1975c], p. 1, traduit de l'anglais : « *The aim of a surround sound system is to reproduce at the listener's ear accurately, reliably and repeatably, the directional sound field created in the studio either by a sound field encoding microphone array, or by artificial directionality encoding devices (pan-pots) or artificial surround-reverberation devices.* »

<sup>20</sup> Le terme utilisé en anglais est « *azimuthal harmonic* », ce terme est préservé ici pour correspondre à la publication mais l'usage mathématique est harmonique circulaire et afin de préserver une cohérence avec les travaux cette dernière dénomination sera utilisée par la suite.

précédemment. Ce système de spatialisation est défini tel celui de D. H. Cooper et T. Shiga, autour de deux opérations principales appelées l'encodage et le décodage :

- « L'encodage est le processus consistant à affecter à chaque direction du son une façon d'inclure ce son parmi les  $n$  canaux d'enregistrement ou de transmission disponibles pour transmettre l'information »<sup>21</sup>.
- « Le décodage est le processus consistant à dériver des canaux appropriés pour alimenter un dispositif de haut-parleurs défini à partir des signaux de canaux de transmission encodés de manière à produire une approximation de l'illusion de l'effet de direction encodé prévu »<sup>22</sup>.

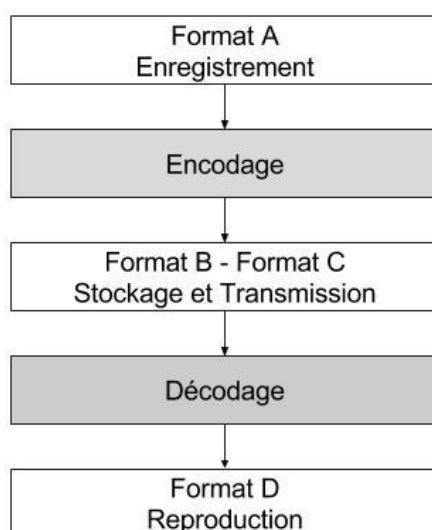


Figure 1.3. Représentation de la chaîne de traitement ambisonique avec les deux opérations d'encodage et de décodage et les formats ambisoniques originaux et leurs principales fonctionnalités.

---

<sup>21</sup> [Gerzon, 1977], p. 400, traduit de l'anglais : « *Encoding is the process of assigning to every given direction of sound a way of including this sound among the  $n$  available recording or transmission channels used to convey the information.* »

<sup>22</sup> [Gerzon, 1977], p. 400, traduit de l'anglais : « *Decoding is the process of deriving from the encoded transmission channels signals suitable for feeding a stated loudspeaker layout so as to produce an approximation to the illusion of the intended encoded directional effect.* »

Cette approche de la spatialisation en deux opérations distinctes considère trois domaines d'interactions associés à l'enregistrement, la transmission (et le stockage) et à la reproduction auxquels M. A. Gerzon et P. B. Fellgett ont associé quatre formats A, B, C et D<sup>23</sup> [Figure 1.3] :

- Le format A est associé au domaine du pré-encodage. Il possède quatre canaux  $L_b$ ,  $L_f$ ,  $R_f$  et  $R_b$  qui peuvent s'apparenter au résultat de l'enregistrement d'une scène sonore avec quatre microphones hypercardioïdes – avec une figure de directivité nulle au-delà de 120 degrés et respectivement orientés vers les quatre coins de l'espace à 45, 135, 225 et 315 degrés. Initialement horizontal donc, en considérant la hauteur, ce format correspond à l'enregistrement d'une scène sonore avec quatre microphones hypercardioïdes – avec une figure de directivité nulle au-delà de 114,1 degrés – et respectivement orientés vers les quatre coins de l'espace à 45, 135, 225 et 315 degrés d'azimut et 35,3 degrés d'élévation vers le haut, vers le bas, vers le haut et vers le bas [Figure 1.4]. Ce format peut donc être directement obtenu via l'enregistrement, à l'aide d'un microphone de champ sonore dont les spécifications ont été prédéfinies par M. A. Gerzon<sup>24</sup> [Gerzon, 1975a], [Gerzon, 1975c].

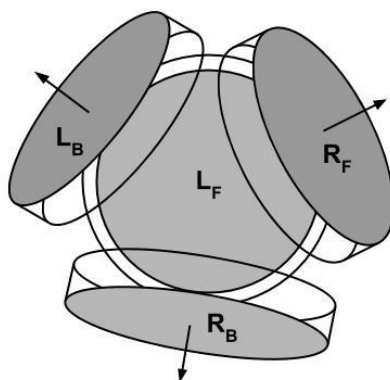


Figure 1.4: Représentation du microphone de champ sonore présenté par P. G. Craven et M. A. Gerzon [Craven & al., 1977] comportant quatre capsules cardioïdes captant quatre canaux associés au format A.

- Le format B est associé au domaine post-encodage ou pré-décodage. Il possède un canal omnidirectionnel W et trois canaux X, Y et Z associés à des figures de huit orientés respectivement sur les axes de coordonnées devant-arrière, droite-gauche et haut-bas

<sup>23</sup> Les termes anglais *A-format*, *B-format*, *C-format* et *D-format* sont plus généralement utilisés dans la littérature même francophone.

<sup>24</sup> Le terme anglais *sound field microphone* est plus généralement utilisé dans la littérature même francophone.

normalisés par un facteur de  $\sqrt{2}$  et correspondant, comme évoqué précédemment, aux quatre premières harmoniques sphériques. Permettant ainsi de considérer cette représentation de l'espace sonore comme une décomposition du champ sonore en harmonique sphérique d'ordre 1. Ce format peut être obtenu à partir d'une opération de matricage du format A<sup>25</sup> via un module électronique de type AB [Gerzon, 1975c] tel que représenté par la [Figure 1.5]<sup>26</sup>.

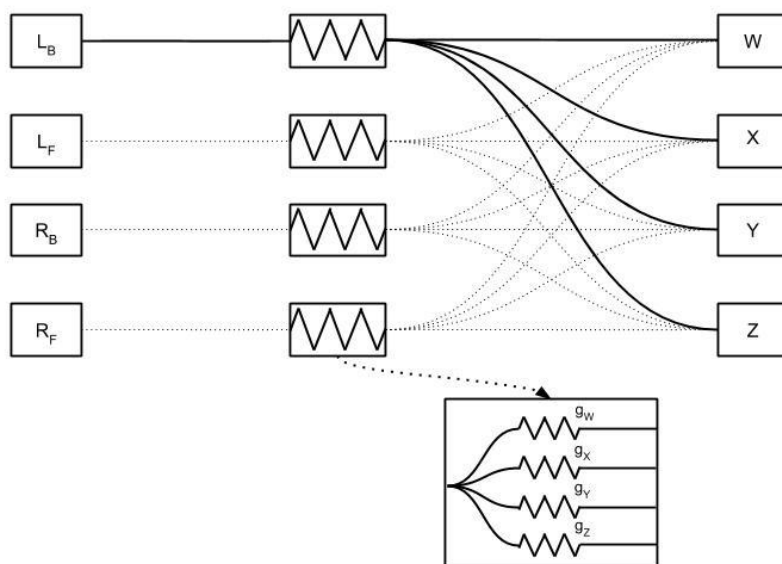


Figure 1.5. Représentation schématique du module électronique de matricage du format A vers le format B tel que défini dans [Gerzon, 1975c]. Les signaux X, W, Y et Z du format B sont les résultats de combinaisons des signaux  $L_B$ ,  $L_F$ ,  $R_F$  et  $R_B$ . Chaque signal du format A contribue à des facteurs de gains près à l'ensemble des signaux du format B. Les gains sont appliqués initialement via l'utilisation de résistances, comme représenté schématiquement en bas de la figure. Dans la pratique, le système optimise ces opérations car certaines contributions d'un signal du format A et certains gains sont identiques.

M. A. Gerzon soumet une autre approche permettant de réaliser un encodage au format B d'un signal audio, selon une direction donnée par manche de contrôle. L'opération consiste alors à utiliser les signaux  $x$  et  $y$  représentant la position cartésienne du manche, dans un système de matricage similaire au module AB [Gerzon, 1975c]<sup>27</sup>. Ce format

<sup>25</sup> Il est à noter qu'une opération d'encodage est déjà formulée mathématiquement sans pour autant définir de manière explicite une norme ambisonique dans [Gerzon, 1973].

<sup>26</sup> Les premières mises en œuvre telles que présentées dans [Gerzon, 1975c] ignorent la composante de hauteur Z et ne prennent pas en compte l'élévation mais pour des raisons de concision, la forme complète est d'ores et déjà présentée.

<sup>27</sup> Il est d'ores et déjà intéressant de noter pour la suite que les signaux  $x$  et  $y$  sont filtrés de telle sorte que  $x^2 + y^2 \leq 1$ . Cela implique que cette approche offre, en plus de couvrir les 360 degrés d'azimut, un effet d'intérieur ou de rapprochement de la source sonore par l'augmentation de la contribution du canal W par rapport aux contributions des autres canaux. Cette question sera traitée par la suite.

caractéristique de l'ambisonie est utilisé principalement pour stocker et transmettre des scènes sonores car les informations spatiales y sont clairement définies sans confusion. De plus, sur le plan de la qualité, contrairement aux techniques quadripophoniques en usage, le format offre une représentation homogène sur l'ensemble de la sphère – ou du cercle – représentant la scène sonore. Enfin, le format est aussi avantageux pour les traitements du son et les usages en studio car peu sujet aux erreurs de phase, aux effets de pompage de réduction de bruit, etc.

- Le format C est analogue au format B à la différence qu'il prend en compte certaines caractéristiques de systèmes particuliers tel le système Matrix HJ développé par la BBC [Gaskell & al., 1977] – système permettant notamment de transmettre des signaux quadripophoniques par modulation de fréquence – et assurant la compatibilité avec des systèmes stéréophoniques et monophoniques chez les particuliers. Ce format C sera généralisé par la suite pour donner le système ambisonique UHJ [N.R.D.C., 1977], [Gerzon, 1985]<sup>28</sup>.
- Enfin, suite au décodage, le format D correspond aux signaux destinés au système de haut-parleurs pour la reproduction de la scène sonore. Fellgett indique que ce format « ne peut pas être réellement standardisé car il dépend nécessairement du nombre et de la disposition des haut-parleurs de l'auditeur »<sup>29</sup>. Néanmoins M. A. Gerzon propose dans un premier temps une approche permettant de convertir le format B pour un système de haut-parleurs rectangulaire, en appliquant un système de matriçage AB pour récupérer des signaux correspondant à un système quadripophonique<sup>30</sup>. Cette approche intègre certains principes psychoacoustiques, déjà présentés dans [Gerzon, 1974]<sup>31</sup>, tels que les différents mécanismes de localisation du son dans les hautes fréquences, théorie du vecteur d'énergie, et les basses fréquences, théorie de localisation des sons de Makita [Makita, 1960], fondés sur les travaux de Lord Rayleigh [Rayleigh, 1876] en augmentant la contribution de 1,76 dB de la composante omnidirectionnelle W et en

---

<sup>28</sup> Le format C, répondant principalement à des contraintes matérielles et techniques, son étude n'offre que peu d'intérêt dans un contexte d'expérimentations artistiques et musicales comparée au format B.

<sup>29</sup> [Fellgett, 1975], p. 8, traduit de l'anglais : « *This D-format cannot be precisely standardised since it necessarily depends on the number and layout of the listener's loudspeakers.* »

<sup>30</sup> Le système intègre un paramètre de contrôle pour compenser le caractère rectangulaire du système de haut-parleurs comparé aux spécifications du format A en utilisant notamment la pseudo-inverse de la matrice d'encodage.

<sup>31</sup> Cette approche sera approfondie et généralisée par la suite dans [Gerzon, 1992a].

diminuant les contributions de 1,25 dB des composantes directionnelles X et Y dans les hautes fréquences, supérieure à 700 Hz<sup>32</sup>.

L'ambisonie, lors des premières recherches, consiste alors en un format principal, le format B, construit autour deux opérations de base, l'encodage et le décodage. Ce format a pour caractéristiques de posséder trois canaux pour une représentation bidimensionnelle de l'espace et quatre canaux pour une représentation tridimensionnelle associée à des figures de directivité et des directions précises en relation avec des systèmes de microphones clairement définis<sup>33</sup>. J. Daniel décrit alors à propos du format B qu'il offre « [...] une représentation rationnelle du champ acoustique, décrivant sans équivoque les informations directionnelles de l'espace sonore, et qui n'est pas assujettie à une configuration particulière du dispositif de restitution »<sup>34</sup>. L'approche ambisonique offre via le format B et l'opération de décodage la possibilité de s'adapter à différents systèmes de haut-parleurs et bien que les propositions ne répondent pour le moment qu'à un nombre limité de configurations, celles-ci seront complétées par la suite avec des systèmes de haut-parleurs réguliers<sup>35</sup> ou pseudo-réguliers [Gerzon, 1992a]. Enfin, la décomposition du champ sonore par les harmoniques sphériques sur laquelle l'approche est fondée [Gerzon, 1973], [Gerzon, 1975c], permet d'établir le cadre mathématique nécessaire à l'extension et la généralisation de l'approche ambisonique.

## 1.2. Évolutions et extensions

Malgré les nombreux avantages que semble posséder l'approche ambisonique, celle-ci ne reçut pas le succès escompté<sup>36</sup> et cela jusqu'à très récemment<sup>37</sup>. Mais bien que cet ensemble de techniques est resté confiné à un petit public, l'ambisonie est demeurée le sujet de nombreuses recherches. Certaines, plus scientifiques, reformulent l'approche afin de la consolider sur un

---

<sup>32</sup> Ce que J. Daniel appelle l'optimisation *max-r<sub>e</sub>* (maximisation du vecteur d'énergie) en opposition à l'optimisation basses fréquence – ou l'absence d'optimisation – qu'il dénomme *basic* [Daniel, 2001], p. 159.

<sup>33</sup> Bien que ce système de microphone associé au format B soit formalisé en théorie, il est important de noter que la mise en œuvre et l'usage d'un système de microphone ne sont alors pas réalisés au profit d'un système plus simple tel que défini par le format A.

<sup>34</sup> [Daniel, 2001], p. 147.

<sup>35</sup> Les systèmes permettent des discrétisations parfaites d'un cercle ou d'une sphère.

<sup>36</sup> La principale cause semble être des problèmes de brevets et licences commerciales. Une présentation plus complète du sujet est présentée par [Elen, 1991].

<sup>37</sup> Comme présenté par la suite, l'ambisonie se développe dans de nombreuses applications notamment la plateforme vidéo *Youtube*.



plan formel et d'autres, plus musicales, ouvrent des perspectives nouvelles dont les conséquences et avancées majeures sont présentées dans cette thèse.

### 1.2.1. Outils d'analyse critiques

Afin de consolider l'approche ambisonique, M. A. Gerzon propose une métathéorie de la localisation auditive [Gerzon, 1992a], qui tente de rassembler un ensemble de théories dites *primitives* sur la localisation de sons au sein d'une seule théorie. De cette tentative, deux outils psychoacoustiques émergent, permettant d'analyser et de prédire dans une certaine mesure, le rendu d'un système de spatialisation : le vecteur vitesse<sup>38</sup> et le vecteur énergie [Gerzon, 1992a]<sup>39</sup>. Ces vecteurs sont calculés à partir des signaux transmis par les haut-parleurs. Le vecteur vitesse, fondé sur une théorie de localisation des sons et proposé par Y. Makita [Makita, 1960], est un outil d'analyse valable dans les basses fréquences, en dessous de 700 hertz. La direction du vecteur vitesse définit la direction apparente du son, appelée localisation de Makita, et sa norme définit la précision relative de rendu ou réciproquement le flou sonore, la valeur 1 indiquant le rendu optimal correspondant à une onde plane et 0 une source sonore omnidirectionnelle<sup>40</sup>. L'onde plane est un modèle d'onde dont les fronts d'ondes sont des plans parallèles et infinis. Bien que les ondes planes n'existent pas dans la réalité, en acoustique une onde est considérée comme plane lorsque la source sonore est suffisamment lointaine pour que sa courbure puisse être négligée [Figure 1.6]. Le vecteur énergie est l'analogue du vecteur vitesse dans les hautes fréquences, entre 500 hertz et 5000 hertz. Au-delà de 5000 hertz, les mécanismes ne sont pas assez connus par M. A. Gerzon pour qu'il puisse les intégrer<sup>41</sup>.

---

<sup>38</sup> Les vecteurs vitesse et énergie sont introduits dès [Gerzon, 1974] mais ne sont réellement formalisés dans le cadre d'une théorie de localisation des sons, qu'en 1992.

<sup>39</sup> Il est important de noter que ces indices ne sont en théorie valables que pour un auditeur idéalement centré au sein du dispositif de haut-parleurs.

<sup>40</sup> Il est dès à présent nécessaire de noter que ces valeurs sont dépendantes des systèmes de haut-parleurs et ne peuvent être considérées que relativement à ces systèmes.

<sup>41</sup> L'ensemble des hypothèses seront présentées et justifiées dans le chapitre suivant consacré aux modèles psychoacoustiques.

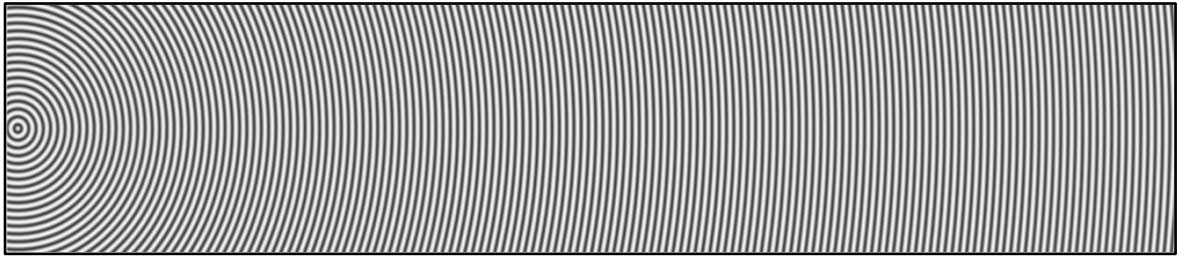


Figure 1.6. Représentation en coupe de l'approximation d'une onde plane en acoustique. La source sonore à gauche génère une onde considérée comme sphérique qui devient progressivement une onde plane à droite où les fronts d'onde sont considérés comme parallèles car la courbure de l'onde est négligeable.

Ce modèle implique de considérer un champ libre, c'est à dire un champ sonore s'étendant sans contrainte à l'infini, seul cas possible dans lequel une source sonore puisse se propager continuellement sans obstacles. L'onde plane possède des caractéristiques directionnelles et des informations d'amplitude relativement simple à définir en faisant un modèle bien plus facile à utiliser que celui des ondes sphériques. C'est pourquoi en ambisonie l'approche assume un champ libre dans lequel les sources sonores sont suffisamment lointaines pour générer des ondes planes.

Le vecteur énergie est l'analogue du vecteur vitesse dans les hautes fréquences, entre 500 hertz et 5000 hertz. Au-delà de 5000 hertz, les mécanismes ne sont pas assez connus par M. A. Gerzon pour qu'il puisse les intégrer <sup>42</sup>.

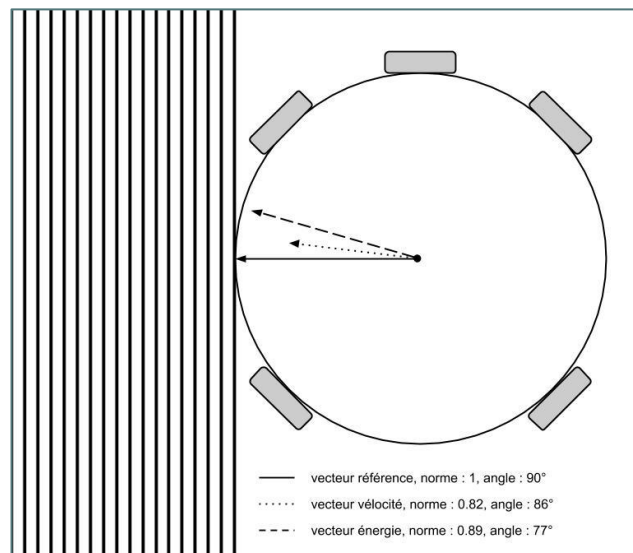


Figure 1.7. Représentation des caractéristiques du vecteur vitesse et du vecteur énergie après la restitution ambisonique optimisée pour un système de haut-parleurs 5.1, suite à l'encodage d'un signal unitaire à 90 degrés sur la gauche. Il est possible de remarquer que, compte tenu de l'irrégularité du système, les vecteurs sont légèrement décalés et de norme inférieure à 1. Il est possible de prévoir un effet de flou relativement important surtout dans les basses fréquences – car la norme du vecteur vitesse est la plus faible – et une distorsion de la

<sup>42</sup> L'ensemble des hypothèses seront présentées et justifiées dans le chapitre suivant consacré aux modèles psychoacoustiques.

scène sonore vers l'avant notamment dans les moyenne-hautes fréquences – car la différence entre l'angle de référence et l'angle du vecteur énergie est la plus importante.

L'enjeu est alors dans le cas de l'encodage, généralement d'un signal unitaire de valeur 1, et au décodage de calculer le vecteur vitesse et le vecteur énergie avec les signaux destinés aux haut-parleurs pour pouvoir critiquer et au besoin optimiser l'opération de décodage selon le domaine fréquentiel. Un rendu optimal étant lorsque les vecteurs vitesse et énergie sont colinéaires avec la direction d'encodage et de normes égales à 1. Une variation de la direction de l'un ou des deux vecteurs par rapport à la direction d'encodage, implique une distorsion directionnelle de la scène sonore dans le ou les domaines de fréquences correspondant. Une norme de l'un ou des deux vecteurs inférieurs à 1, implique une diminution de résolution spatiale [Figure 1.7]. Compte tenu des hypothèses établies sur les mécanismes de localisation du son par l'oreille humaine, ces vecteurs sont des outils puissants d'analyse et de critique des systèmes de restitution et des opérations de décodage, et peuvent servir de modèle de référence pour l'élaboration, la critique et l'évaluation de modèles de spatialisation.

### 1.2.2. Encodage directionnel

Parallèlement à ces travaux, et bien que l'approche ne trouve pas immédiatement de grand écho dans l'industrie cinématographique et musicale, à partir des années 1990, les développements de l'informatique et notamment de l'audio numérique sont propices à l'expansion et au déploiement des techniques ambisoniques. A ce titre, il est possible de noter une mise en œuvre ambisonique entièrement numérique intégrant l'utilisation de norme MIDI – restant néanmoins en temps différé – par D. G. Malham et R. Orton [Malham & al., 1991], un logiciel de décodage ambisonique développé par D. G. Malham, [Malham, 1992] ou encore une mise en œuvre de l'ambisonie dans le logiciel Csound sous la forme de Opcodes par D. G. Malham et A. Myatt [Malham & al., 1995] [Figure 1.8]. De plus, l'une des plus importantes améliorations est la formulation de l'encodage ambisonique, notamment par la démocratisation de l'usage des fonctions trigonométriques par l'informatique. Comme présentée précédemment, la « méthode d'inclusion est généralement obtenue en assignant le son à chacun des  $n$  canaux avec un gain fixe et différent (pouvant être réel ou complexe), qui est fonction de la direction d'encodage prévue »<sup>43</sup>. Le module d'encodage AB présenté par M. A. Gerzon est un *système matriciel* n'offrant que peu de flexibilité. Par opposition à cette approche, M. A. Gerzon soumet

---

<sup>43</sup> [Gerzon, 1977], p. 400, traduit de l'anglais : « *The method of inclusion is usually achieved by assigning the sound to each of the  $n$  channels with a different and stated gain (which may be real or complex) which is a function of the intended encoded direction.* »

l'avantage d'un *système noyau* « [...] dans lesquels les gains des canaux sont exprimés sous la forme d'une fonction lisse et continue de la direction, car ces systèmes sont décrits mathématiquement par « noyaux » de la même manière que les « systèmes matriciels » sont décrits par des matrices »<sup>44</sup>. D. G. Malham, de l'Université de York au Royaume-Uni, généralise l'usage de la formulation de l'encodage ambisonique d'un signal en fonction d'un azimut et d'une élévation, fondé notamment sur les équations mathématiques de la décomposition de l'espace en harmoniques sphériques de M. A. Gerzon [Gerzon, 1973], pour le premier ordre de décomposition (FOA<sup>45</sup>) [Figure 1.8].

```

; Calcul des coefficients d'encodage
; selon l'azimut et l'élévation en radian
kca = cos(azimut)
ksa = sin(azimut)
kce = cos(elevation)
kse = sin(elevation)

; Encodage vers le format B du signal
aw = sig * 0.707
ax = sig * kca * kce
ay = sig * ksa * kce
az = sig * kse

; Sortie du format B
outq aw, ax, ay, az

```

Figure 1.8. Exemple d'un code Csound adapté de [Malham & al., 1995] où l'encodage ambisonique directionnel d'une source sonore est réalisée selon l'équation des harmoniques sphériques pour le premier ordre de décomposition en fonction d'un azimut et d'une élévation donnés en radians. Les signaux du format B, noté *aw*, *ax*, *ay* et *az* sont obtenus en pondérant le signal de la source sonore est noté *sig* par les coefficients d'encodage ambisoniques fonctions de l'azimut et de l'élévation.

Cette opération d'encodage directionnel d'une source sonore en ambisonie, bien que n'étant qu'une simple reformulation autour de l'encodage du format A vers le format B [Équation 2.12], n'est pas si anodine car elle coïncide avec une réelle volonté – et surtout une réelle possibilité – de créer des espaces sonores artificiels et offre donc un véritable potentiel musical. En effet, alors que le système proposé par M. A. Gerzon de manche directionnel ne reste en réalité qu'à l'état de prototype et que les résultats spatiaux ne sont pas garantis<sup>46</sup>, l'utilisation

---

<sup>44</sup> [Gerzon, 1977], p. 400, traduit de l'anglais : « *It is convenient to term systems in which the gains of the channels are expressed as a smooth and continuous function of direction, "kernel system," since such systems are mathematically described by "kernels" in the same way as so-called "matrix systems" are described by matrices (see Appendix).* »

<sup>45</sup> *First Order Ambisonics*.

<sup>46</sup> [Gerzon, 1975], p. 5, « Étant donné que la plupart des paires de manches directionnels pour le panoramique ne sont pas bien conçus en ce qui concerne la constance du niveau sonore avec la direction, aucune garantie ne peut

de l'informatique et de contrôleurs plus performants [Malham, 1992] permet de définir un système suffisamment consistant pour une utilisation effective.

### 1.2.3. Adaptation à des systèmes de restitution variés

Ce type d'approche plus artistique concorde aussi avec l'usage de systèmes de haut-parleurs plus hétéroclites et moins conventionnels, tels que la salle de concert de l'université de York [Figure 1.9], avec des auditoires plus élargis – comparé à un auditeur utilisant un système de haut-parleurs enveloppant domestique tel que le format 5.1 le propose – nécessitant alors d'adapter l'opération de décodage. D. G. Malham propose alors une nouvelle approche complétant celle déjà proposée par M. A. Gerzon, destinée à une restitution du champ sonore pour un auditoire étendu, en augmentant la composante omnidirectionnelle  $W$  suffisamment par rapport aux composantes directionnelles  $X$ ,  $Y$  et  $Z$  – une extrapolation de l'approche hautes fréquences proposée par M. A. Gerzon – permettant de redéfinir ce qu'il appelle la réponse directionnelle des sources<sup>47</sup>. Ce afin d'annuler les contributions opposées à la direction de la source<sup>48</sup> et réciproquement renforcer l'énergie sonore dans cette direction.

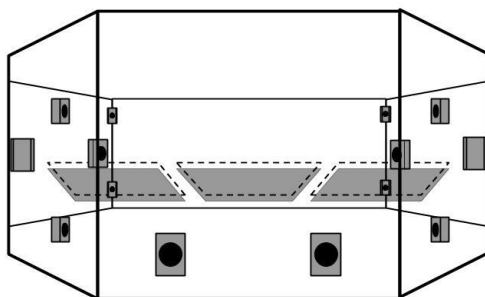


Figure 1.9: Vue de l'avant de la salle de concert Sir Jack Lyons à l'université de York, montrant les positions des haut-parleurs proposées par D. G. Malham [Malham, 1992].

---

être donnée que la modification ambisonique sera également bonne à cet égard. », traduit de l'anglais :  
« *Because most joystick pairwise pan pots are not well designed as regards constancy of sound level with direction, no guarantee can be given that the ambisonic modification will be good in this respect either.* »

<sup>47</sup> En anglais « *Source Directional Response (SDR)* », [Malham, 1992], p. 5.

<sup>48</sup> Nouvelle approche que J. Daniel dénomme *in-phase* car ne préservant que les contributions en phase [Daniel, 2001], p. 129.

## 1.2.4. Traitements du champ sonore

Le format B et la représentation du champ sonore qui lui est associée n'y est aussi presque plus exclusivement destiné au stockage et à l'enregistrement de champ sonore<sup>49</sup>. Les premières opérations pertinentes sur un plan spatial sont alors établies clairement dans la littérature telles que :

- La *rotation du champ sonore*, proposée par J. Daniel [Daniel, 2001], p. 164, permettant de changer de manière identique et homogène l'ensemble des sources composant le champ sonore. L'opération, qui peut être décomposée entre trois rotations primaires autour des axes x, y et z associés aux composantes ambisoniques, est aujourd'hui fortement ancrée dans les usages et est notamment un des atouts majeurs de l'approche ambisonique. Car elle permet en effet de s'adapter aux mouvements de la tête pour des restitutions au casque<sup>50</sup>.
- La *distorsion de la perspective*, proposée initialement par M. A. Gerzon afin de corriger les déformations de la perspective spatiale lors de décodage pour des systèmes de restitution irréguliers associés à la télévision [Gerzon & al., 1992], permet de resserrer ou élargir la scène frontale<sup>51</sup>. Cette opération a par la suite été redéfinie sous le terme d'effet *fish-eye*, au sein de la bibliothèque HOA [Colafrancesco & al., 2013] et extrapoler pour aboutir à une approche plus générale de déformation de la perspective [Guillot, 2013], p. 57, [Paris, 2013], p. 54.
- La *focalisation*, proposée par J. Daniel [Daniel, 2001], p. 167, consiste à simuler une prise de son directive dans le champ sonore, enregistrée à l'aide d'un microphone cardioïde virtuel. L'opération peut être assimilée à l'opération de décodage<sup>52</sup> vers un seul haut-parleur et a été de même développée de façon musicale au sein de la bibliothèque HOA sous la notion de filtrage de champ sonore [Guillot, 2013], p. 53, [Sèdes & al., 2014].

---

<sup>49</sup> En considérant le faible succès commercial de l'approche ambisonique, il est difficile de juger du réel avantage quant aux traitements de studio évoqués par M. A. Gerzon, et stipulés précédemment dans cette thèse.

<sup>50</sup> Restitution hybride via la synthèse binaurale qui est un des enjeux essentiels de la réalité virtuelle.

<sup>51</sup> En anglais « *forward dominance* » ou réciproquement, bien que non défini dans la littérature « *backward dominance* ».

<sup>52</sup> Présentée mathématiquement dans le chapitre suivante, l'opération consiste en une projection en direction d'un seul point de l'espace.

### 1.2.5. Extension aux ordres supérieurs

Au-delà des possibilités musicales, en reprenant le cadre mathématique formel tel que défini par M. A. Gerzon, proposant une décomposition du champ sonore en harmoniques sphériques [Gerzon, 1973] et la représentation locale du champ sonore sous la forme d'une série d'ondes planes [Gerzon, 1992a], J. Daniel transforme et augmente le format B, fondé sur une décomposition à l'ordre 1, en étendant le modèle ambisonique à des ordres de décomposition élevés<sup>53</sup> (HOA<sup>54</sup>). Offrant ainsi de nombreux avantages :

« Tout en préservant les qualités de restitution propres à *Ambisonics* énoncées plus haut, il est attendu que les systèmes d'ordres supérieurs reculent les limites ou les faiblesses des systèmes traditionnels d'ordre 1, en termes de précision des images, de stabilité hors du *sweet-spot* et de préservation des impressions spatiales (séparation latérale). [...] C'est alors une représentation *scalable* qu'il est possible de transmettre. En fonction des ressources disponibles à la transmission (débit) ou à la restitution (nombre de haut-parleurs et capacité du décodeur), l'ensemble ou seulement une partie des composantes ambisoniques peut être exploitée, donnant lieu à divers degrés de résolution de l'espace sonore restitué, mais avec toujours avec la garantie de sa cohérence. »<sup>55</sup>

---

<sup>53</sup> Il est néanmoins important de noter que la décomposition du champ sonore en harmonique sphérique était déjà établie jusqu'à l'ordre 3 par M. A. Gerzon [Gerzon, 1973] mais ne faisait pas partie d'un cadre strict tel que défini en ambisonie. Parallèlement au travail de J. Daniel, D. G. Malham propose une extension du format B à l'ordre 2 suivant la *max-normalisation (maxN)* [Malham, 1999] puis suite au travail de J. Daniel une extension à l'ordre 3 donnant le format *Furse-Malham* (ou *FuMa*) [Malham, 2003].

<sup>54</sup> Higher Order Ambisonics

<sup>55</sup> [Daniel, 2001], p. 147.

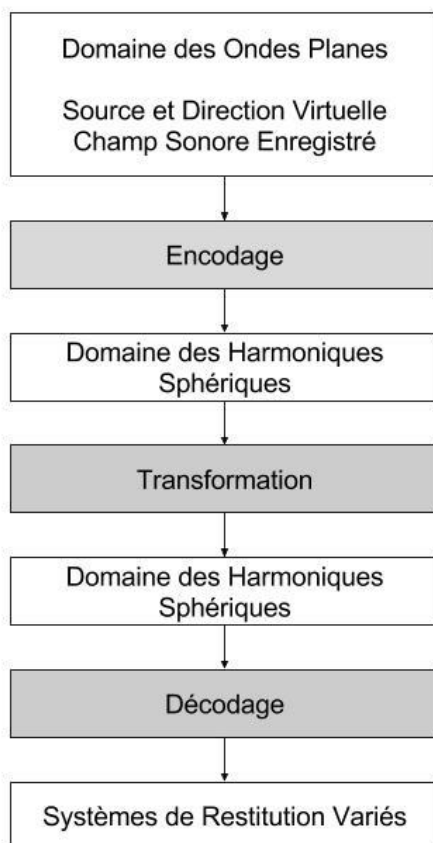


Figure 1.10. Représentation de la chaîne de traitement ambisonique avec les trois opérations redéfinies, l’encodage, la transformation et le décodage et les formats ambisoniques originaux redéfinis.

Suite à ces évolutions, l’ensemble de l’approche ambisonique se retrouve redéfinie autour d’un modèle mathématique dynamique et solide – la décomposition en harmoniques – associé à un modèle acoustique cohérent – les ondes planes – et des outils acoustiques psychoacoustiques pertinents et dans le contexte d’un système de spatialisation directionnel – les vecteurs énergie et vitesse. Cette évolution se déroule dans un contexte technique de plus en plus libre, que ce soit sur un plan computationnel avec les avancées informatiques ou sur un plan matériel avec les contrôleurs et les systèmes de restitution. Les techniques ambisoniques s’organisent selon une chaîne d’opérations et des domaines d’interactions reformulés [Figure 1.10].



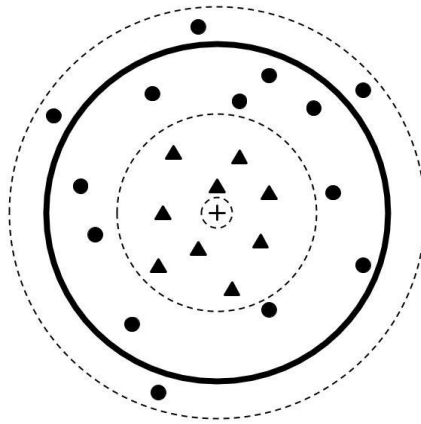


Figure 1.11. Représentation des trois dispositions possibles de l'auditoire et les optimisations du champ sonore associées pour un système de haut-parleurs enveloppant, représenté par le cercle continu. La croix centrale représente un auditeur idéalement centré, les triangles représentent un auditoire limité et confiné au centre du cercle et les petits cercles représentent un auditoire réparti dans l'ensemble du cercle et à l'extérieur. Les cercles en pointillés représentent les zones dans lesquels les optimisations valides – mais pas forcément idéales – avec du centre vers l'extérieure les optimisations dites *basic*, *max-r<sub>e</sub>* et *in-phase*<sup>56</sup>.

L'opération d'encodage peut être réalisée indifféremment pour une onde plane, associée à une source sonore et une direction donnée, ou un ensemble d'ondes planes, définissant un champ sonore. L'encodage ambisonique conduit au domaine dit des *harmoniques sphériques* qui peut être assimilé à une extension du format B et qui permet des opérations de transformations originales, telles la rotation du champ sonore, la distorsion de la perspective ou la focalisation. Le champ sonore dans le domaine des harmoniques sphériques permet, via l'opération de décodage, d'être restitué à destination de systèmes variés, allant de systèmes de haut-parleurs domestiques relativement standardisés – tels que des systèmes quadriphoniques ou le système 5.1, [Gerzon, 1985], [Gerzon & al., 1992] – aux salles de concert comportant des systèmes de haut-parleurs plus imposants et variés, comme la salle de concert Sir Jack Lyons à l'université de York, [Malham, 1992], [Daniel & al., 1998]<sup>57</sup>, en passant par la restitution au casque via la synthèse binaurale [Daniel & al., 1998], [Daniel, 2001]. Cette opération permet aussi d'adapter le champ sonore à des auditoires différents, allant d'un seul auditeur au centre du dispositif [Gerzon, 1975c] à un auditoire pouvant recouvrir ou être à l'extérieur du dispositif de restitution [Malham, 1992] ou toute disposition intermédiaire ou hybride. Cela est possible notamment

<sup>56</sup> Il est important de noter que la zone associée à l'optimisation *max-r<sub>e</sub>* recouvre la zone associée à l'optimisation *basic* et la zone associée à l'optimisation *in-phase* recouvre l'ensemble de l'espace. En cela, les zones sont considérées comme valides mais pas forcément idéales de plus le schéma ne tient pas compte du domaine fréquentiel. Le sujet est traité plus en profondeur par la suite.

<sup>57</sup> Bien que dans la pratique les opérations de décodage ne sont pas optimisées pour toutes les configurations possibles de systèmes immersifs.

grâce à ce que J. Daniel appelle les optimisations *basic*, *max-r<sub>e</sub>* et *in-phase* du champ sonore [Daniel, 2001], p. 159, [Figure 1.11].

## 1.3. Perspectives et bilan

L'approche ambisonique bien que clairement formalisée, autant d'un point de vue mathématique qu'acoustique et psychoacoustique, est néanmoins toujours au centre de nouvelles recherches et de nouveaux développements. Il est alors intéressant de montrer les enjeux actuels et améliorations envisagées, liées aux pratiques contemporaines qui trouvent notamment résonance dans l'approche expérimentale et artistique présentée dans la suite de cette thèse. Sur ces faits, il sera possible de procéder à une présentation objective des différentes spécifications techniques et fonctionnelles de l'approche, ainsi que la manière dont le modèle ambisonique répond à ces problèmes. L'enjeu étant par la suite de pouvoir les mettre en relief avec l'appropriation musicale présentée dans ce travail.

### 1.3.1. Enjeux actuels

Les développements ambisoniques se poursuivent notamment autour des deux axes centraux que sont l'encodage et le décodage :

- L'inclusion de la distance dans le système ambisonique – à savoir offrir la possibilité de donner l'impression que la source s'éloigne ou se rapproche de l'auditoire – est un des principaux obstacles à dépasser. Le modèle ambisonique étant construit autour d'une représentation du champ sonore sous la forme d'une série d'ondes planes [Annexe A.1], donc supposant des sources à des distances infinies, la distance n'est tout simplement pas intégrée dans le modèle. Cependant, certaines approches tentent de répondre au problème tel que celle présentée par J. Daniel, qui vise à simuler les effets du champ sonore proche à l'encodage [Daniel, 2003] et dont une mise en œuvre a été réalisée en C++ par F. Adriaensen [Adriaensen, 2006]. Ou encore l'approche de A. Sontacchi et R. Höldrich qui tente de combiner l'approche ambisonique à la WFS [Sontacchi & al., 2001].
- Offrir un système qui puisse s'adapter de manière universelle et optimale à l'ensemble des systèmes de restitution, quel que soit le placement des haut-parleurs tant sur le plan angulaire que de la distance, est l'un des enjeux primordiaux que tente de soulever le système. Les recherches consistent donc à optimiser l'opération de décodage pour des

systèmes de restitution toujours plus variés, notamment via les approches prometteuses d'hybridation. Ces hybridations, de manière identique à celles qu'il y a eu entre l'ambisonie et la synthèse binaurale – pour une virtualisation au casque du champ sonore – utilisent les avantages du *vector base amplitude panning* [Pulkki, 1997], combinés à ceux de l'ambisonie, afin de compenser certaines restrictions de cette dernière<sup>58</sup>. Ces recherches ont été menées parallèlement par l'Institut de Musique Electronique et d'Acoustique (IEM) à Graz en Autriche [Zotter & al., 2010], [Zotter & al., 2012], [Zotter & al., 2013] et Technicolor [Batke & al., 2010] et initialement introduites dans un contexte similaire mais parallèle à l'ambisonie [Poletti, 2005], [Poletti, 2007].

Ces deux enjeux sont au cœur d'une appropriation musicale de l'ambisonie. Tout musicien, compositeur travaillant l'espace sonore souhaite, en effet, pouvoir créer du mouvement et notamment des trajectoires, sans pour autant être limité aux seules dimensions angulaires. Lorsque en 1977, J. Chowning présente comment simuler des sources sonores se déplaçant dans l'espace [Chowning, 1977], la notion de distance y est d'ailleurs déjà intégrée. Elle continuera, au regard des usages et outils de spatialisation, d'être une requête majeure des systèmes mis en œuvre. Il en est de même pour l'adaptabilité des systèmes de spatialisation aux différents dispositifs de restitutions. Bon nombre de pièces électroacoustiques demandent d'être réécrites presque intégralement sur un plan logiciel, voire repensées, car elles ne peuvent tout simplement pas s'adapter, ou sont terriblement mal reproduites, lorsque les dispositifs de restitution ou la disposition de l'auditoire s'éloignent trop du cadre pour lequel elles étaient initialement conçues. Or, les espaces de restitution, de jeu et les systèmes de haut-parleurs peuvent être extrêmement différents et variés, ce qui est habituel dans le milieu de la musique contemporaine électroacoustique<sup>59</sup>. Ces questions alimentent un axe important de recherches dans le contexte musical et seront discutées et confrontées à l'approche particulière de l'ambisonie.

Suite à la généralisation et à la diffusion, bien que relativement tardive, du modèle ambisonique, un des enjeux majeurs reste encore la question de la standardisation de l'approche et notamment la notion du format. Au-delà des considérations extrêmement techniques des codecs de codage, des fichiers audio et de leurs descriptions dans les en-têtes des fichiers, les formats soulèvent

---

<sup>59</sup> Il est possible de faire référence au grand dôme du ZKM à Karlsruhe en Allemagne, à la salle de projection de l'IRCAM à Paris, à l'AlloSphere à Santa Barbara en Californie ou encore le studio du CICM à St-Denis.

les questions de la normalisation des canaux, notamment les uns par rapport aux autres, et du classement des canaux. De nombreux problèmes apparaissent, dus aux différentes normalisations comme la *max-normalisation* (*maxN*) [Malham, 1990], la *semi-normalisation 3D* (*SN3D*) et la *normalisation 3D* (*N3D*) [Daniel, 2001], p. 156. Une confusion dans les procédés entraîne des déformations du champ sonore, comme une prédominance de la composante omnidirectionnelle ou inversement un déséquilibre dû à l'accentuation de composantes directionnelles. Le classement des canaux ambisoniques, qui peut être réalisé selon le *Furse-Malham Set* (*FuMa*) [Daniel, 2001], p. 156, le *Single Index Designation* (*SID*) [Daniel, 2003] ou l'*Ambisonic Channel Number* (*ACN*), peut dans le cas d'une erreur, créer des rotations ou des renversements du champ sonore involontaires<sup>60</sup>. A ce sujet, le classement *ACN* utilisé dans le format *ambiX* [Nachbar & al., 2016] semble devenir la norme, notamment en raison de son utilisation sur la plateforme de vidéo Youtube de Google<sup>61</sup>. Ces questions qui peuvent *a priori* paraître absconses ou du moins très techniques, notamment lors d'une approche créative de l'ambisonie – où la question de format et de standardisation n'est clairement pas au centre des préoccupations – implique néanmoins, comme cela sera présenté, de nombreuses répercussions lors de l'appropriation musicale du modèle ambisonique et spécifiquement pour la mise en œuvre de traitements sonores dans le domaine des harmoniques sphériques.

### 1.3.2. Bilan

De ses origines à sa première mise en œuvre, proposée par M. A. Gerzon et P. B. Fellgett, à sa diffusion et sa démocratisation actuelle, en passant par les recherches et développements intermédiaires, notamment proposés par J. Daniel et D. G. Malham, le modèle ambisonique n'a cessé de s'étendre et de se diversifier. Ces changements ont eu lieu tout en gardant des constantes – des buts, des enjeux, des spécifications fonctionnelles et techniques – qui se retrouvent de manière répétée tout au long des développements. De par ses origines, le but décisif des techniques de spatialisation est de pouvoir reproduire pour un auditeur un espace sonore virtuel le plus proche, et ressemblant le plus possible à un espace sonore réel. L'approche ambisonique, et la plupart des techniques de spatialisation de la stéréophonie<sup>62</sup>, de la *wave field*

---

<sup>60</sup> Ce type d'artefact apparaît dans le meilleur des cas. Dans la plupart des situations, une complète déformation du champ sonore est générée.

<sup>61</sup> Spécification définie dans la documentation du répertoire Github *spatial-media* de Google accessible en septembre 2017 en ligne <https://github.com/google/spatial-media/blob/master/docs/spatial-audio-rfc.md>.

<sup>62</sup> Au sens « classique » du terme définie par la Commission Électrotechnique Internationale (IEC) dans EC 60050 comme une « technique de prise de son, d'enregistrement, de transmission et de reproduction des sons permettant à l'auditeur de reconstituer la répartition dans l'espace des sources sonores », accessible en ligne en septembre 2017 à l'adresse <http://www.electropedia.org/iev/iev.nsf/display?openform&ievref=723-03-43>.

*synthesis* en passant par la synthèse binaurale, intègrent une notion d'artificielle en offrant la possibilité de « créer de toute pièce » un espace sonore, notamment source par source, n'ayant aucune réalité préalable. M. A. Gerzon définit que « le but premier de tout système de son enveloppant, est de produire dans les oreilles et le cerveau d'un auditeur l'illusion d'un motif de son directionnel prévu. [...] En outre, de tels systèmes peuvent être conçus soit pour reproduire des sons à une seule distance apparente de l'auditeur, soit pour produire l'illusion de sons à partir de n'importe quelle distance de l'auditeur »<sup>63</sup>. Ainsi la constante reste la même. Que l'espace sonore soit initialement réel ou artificiel, l'objectif principal est de donner à l'auditeur l'impression ou l'illusion de cet espace sonore.

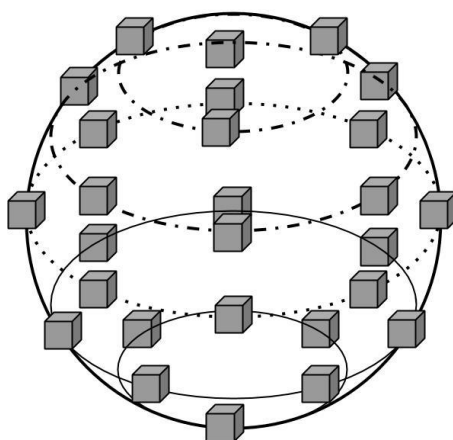


Figure 1.12. Représentation de plusieurs systèmes immersifs des lignes découpant l'espace et une disposition possible de haut-parleurs où l'ensemble des lignes pour un système associé à l'ensemble de la sphère, l'ensemble des lignes intégrant des points pour un système associé à un hémisphère et la ligne pointillée pour un système horizontal. Des systèmes exclusivement frontaux peuvent être envisagés en ne préservant que les haut-parleurs se trouvant devant les lignes.

À partir de cet objectif, un nombre important de spécifications techniques et fonctionnelles ont été établies, qu'il est intéressant de rappeler de manière concise :

1. La principale spécification est bien entendue de pouvoir offrir une restitution immersive à l'auditeur que ce soit sur un plan horizontal [Fellgett, 1975] ou sur une sphère complète [Gerzon, 1977]. Par extension de l'approche, le système devrait pouvoir être

<sup>63</sup> [Gerzon, 1977], p. 400, traduit de l'anglais : « We start by assuming that the primary the primary aim of any surround-sound system is to produce in the ears and brain of a listener the illusion of an intended pattern of directional sound. [...] In addition, such systems may be designed either to reproduce sounds at a single apparent distance from the listener, or to produce the illusion of sounds from any distance from the listener. »

envisagé pour des espaces intermédiaires tels qu'un hémisphère ou encore un hémisphère tronqué de sa scène arrière [Figure 1.12].

2. Le système doit pouvoir reproduire des sources directionnelles, limitées à la surface du cercle ou de la sphère enveloppant l'auditeur, mais idéalement le système devrait permettre de reproduire des sources ponctuelles, c'est à dire en incluant la notion de distance [Gerzon, 1977]. De plus, le système devrait aussi pouvoir reproduire des champs diffus au sens acoustique du terme. Le système devrait, dans le cas d'un enregistrement d'un champ sonore naturel – comme une salle de concert – pouvoir préserver et reproduire les phénomènes de réverbération offrant suffisamment d'informations à l'auditeur pour qu'il puisse concevoir la taille et la forme de l'espace original [Fellgett, 1975]
3. Les haut-parleurs ne doivent pas être considérés comme des sources sonores apparentes [Fellgett, 1975]. En d'autres termes, il s'agit que la restitution soit homogène sur l'ensemble de l'espace environnant l'auditeur : l'effet d'une source sonore dans la direction d'un haut-parleur doit être identique, à une rotation près, de l'effet si la source sonore se trouve dans la direction intermédiaire à deux haut-parleurs. Mais selon un aspect encore plus évident, en considérant une source sonore reproduite par quatre haut-parleurs, l'objectif est de simuler pour l'auditeur la perception d'une seule source sonore et non de quatre sources distinctes associées aux haut-parleurs. Leurs singularités doivent être inaudibles.
4. Le système ne doit pas aboutir à une « coloration » du champ sonore. Idéalement, le système devrait offrir un signal sonore aux oreilles de l'auditeur qui soit l'exacte copie du signal, tel qu'il aurait été s'il se trouvait dans cette espace [Fellgett, 1975]. En d'autres termes, le système doit pouvoir préserver l'ensemble des informations nécessaires à la reproduction identique d'un champ sonore, ou définir l'ensemble des informations nécessaires à la simulation d'un champ sonore créé artificiellement. Mais il doit aussi tenir compte du dispositif et du lieu de restitution, afin de pouvoir compenser les artefacts de ceux-ci et obtenir une reproduction exacte.
5. Attaché à cette notion de dégradation du champ sonore, M. A. Gerzon insiste sur la nécessité de pouvoir reproduire le champ sonore de manière fiable et répété. Il s'agit là

aussi de préserver les informations du champ sonore avec précision, ou de pouvoir les définir correctement dans le cas d'une approche artificielle. Cette spécification se pose en opposition aux approches quadriphoniques qui impliquent des artefacts et déformations dès l'enregistrement et le mixage de par la représentation de l'espace qui leur est induite. Ces artefacts et déformations se répercutant conséquemment au moment de la restitution chez l'auditeur [Gerzon, 1975c].

6. M. A. Gerzon complète cette spécification en y ajoutant l'idée d'adaptabilité du système de spatialisation à différents dispositifs de restitution. Il définit que « de manière conventionnelle, l'encodage sur la bande principale attribue les quatre signaux d'alimentation des haut-parleurs à quatre pistes de la bande, mais il est évident qu'il s'agit d'une procédure sous-optimale, même avec des microphones et des techniques de panoramique correctement conçus. Afin de recréer l'illusion d'un champ directionnel donné, il est évident que les signaux d'alimentation des haut-parleurs ne seront pas les mêmes pour une disposition de haut-parleur rectangulaire que pour un carré. Ainsi, si nous voulons tenir compte d'une variété de formes de disposition des haut-parleurs appropriées à des salles d'écoute de formes différentes, nous devons utiliser un décodeur pour dériver les signaux d'alimentation des haut-parleurs appropriés à la mise en page utilisée »<sup>64</sup>.
  
7. M. A. Gerzon complète la spécification en y intégrant la notion d'auditoire et en définissant qu'« idéalement, on aimerait qu'un système de son immersif présente de manière exacte sur une zone d'écoute raisonnable le champ sonore original de la salle de concert, ou dans le cas de la musique populaire ou électronique, un champ sonore envisagé par le producteur de disques, avec de nombreux sons différents dans différentes directions à différentes distances »<sup>65</sup>. Cette spécification est d'ailleurs particulièrement approfondie par J. Daniel à propos des optimisations de champ sonore.

---

<sup>64</sup> [Gerzon, 1975c], p. 2., traduit de l'anglais : « *Conventional mastertape encoding assigns the four speaker feed signals to four tracks of a tape, but it is obvious that this is a sub-optimal procedure, even with correctly designed microphones and pan-pots. In order to recreate the illusion of a given directional field, it is obvious that the speaker feed signals will not be the same for a rectangular speaker layout as for a square one. Thus if we are to accommodate a variety of shapes of speaker layout appropriate to differently shaped listening rooms, we need to use a decoder to derive the speaker feed signals appropriate to the layout used.* »

<sup>65</sup> [Gerzon, 1974b], p. 2, traduit de l'anglais : « *Ideally, one would like a surround-sound system to recreate exactly over a reasonable listening area the original sound field of the concert hall, or in the case of popular or electronic music, a sound field envisaged by the record producer, with many different sounds in different directions at different distances.* »

8. Le système doit tirer parti au mieux du nombre de canaux qu'il lui est imparti [Fellgett, 1975]. Que ce soit à l'émergence des techniques ambisoniques où M. A. Gerzon définit qu'un « système [d'enregistrement] pratique ne peut se permettre pas de gaspiller les canaux en dupliquant l'information et ne perdra pas d'information en portant une résolution directionnelle grossière au lieu de transmettre des informations de bonne résolution»<sup>66</sup> ou aujourd'hui où les questions de diffusion en continu nécessitent d'optimiser les informations distribuées, la question du nombre de canaux et plus généralement de la quantité d'informations disponibles est une des spécifications les importantes.

Afin de répondre à cet ensemble de spécifications plutôt fonctionnelles telles que définies de 1 à 4, les techniques ambisoniques proposent une approche quelque peu analogue aux techniques holophoniques, consistant à restituer le champ sonore tel qu'il a été ou eût été à l'aide d'un système de haut-parleurs<sup>67</sup>. Cependant un tel objectif soulève de nombreux problèmes une fois confronté au monde réel. Tout comme les approches dont elle est issue, l'approche ambisonique n'est pas en reste et doit faire face à des contraintes physiques et matérielles empêchant la mise en œuvre concrète d'un système idéal et parfait sur le plan conceptuel. M. A. Gerzon affirme alors au sujet de l'approche visant à définir un système immersif et enveloppant que « pour reproduire parfaitement toutes les directions, il faudrait un nombre infini de haut-parleurs minuscules sur une sphère autour de l'auditeur et un nombre infini de microphones à « fonction-delta » (c'est-à-dire infiniment directionnels) pointant dans toutes les directions pour enregistrer les sons. En pratique, le nombre restreint de canaux réduira la résolution directionnelle obtenue »<sup>68</sup>. En somme, l'approche absolue est en soi irréalisable. Malgré des améliorations du matériel et des dispositifs tant sur la qualité – notamment d'enregistrement et restitution – que sur plan de l'ergonomie – en tentant de pallier les contraintes spatiales, de telles spécifications techniques sont irréalisables.

---

<sup>66</sup> [Gerzon, 1973], p. 3, traduit de l'anglais : « *A practical recording system will not allow channels to be wasted by duplicating information, and will not omit information conveying coarse directional resolution while transmitting fine-resolution information.* »

<sup>67</sup> C'est pour cela que J. Daniel considère l'approche ambisonique comme étant holophonique [Daniel, 2001], p. 116 et p. 233.

<sup>68</sup> [Gerzon, 1973], p. 2, traduit de l'anglais : « *To reproduce all directions perfectly, one would need an infinite number of minute speakers on a sphere around the listener, and an infinite number of "delta-function" (i.e., infinitely directional) microphones pointing in all directions to record the sounds. In practice, the restricted number of channels will reduce the obtainable directional resolution.* »



En pratique, il est nécessaire de contourner les problèmes et tout comme J. C. Steinberg et W. B. Snow le réalise bien avant, en limitant leur approche à des systèmes comportant deux-trois microphones et deux-trois haut-parleurs. Il s'agit, comme cela a été dit, de transformer et optimiser au mieux l'approche en changeant le point de vue, en simplifiant les principes sur lesquels elle repose en enlevant des variables. Afin de satisfaire dans la limite du possible les spécifications établies, la direction prise par l'approche ambisonique consiste alors à tenter de « satisfaire des critères psychoacoustiques pertinents aux oreilles de l'auditeur [...] selon la connaissance des stimuli acoustiques qui sont les mieux susceptibles de produire l'illusion souhaitée »<sup>69</sup>. Il ne s'agit donc plus de créer une exacte copie du champ sonore mais de s'intéresser à comment celui-ci perçoit le champ sonore pour se focaliser et ne préserver au possible que les informations qui semblent pertinentes. En un sens, le champ sonore dans son entièreté se retrouve limité localement aux environs de l'auditeur et l'ensemble des composantes complexes qui le définissent sont appréhendés au regard des mécanismes de l'oreille humaine ou de manière plus générale, de son système auditif [Figure 1.13]. Sur ces principes, M. A. Gerzon proposera une théorie de la localisation des sons – théorie qu'il définit comme métathéorie car englobant un ensemble de théories déjà présentes dans la littérature [Gerzon, 1992]. Cette théorie psychoacoustique est donc le support primordial de l'édification de l'ensemble des techniques ambisoniques permettant notamment de valider la totalité de l'approche et donc les modèles acoustiques et mathématiques. Pour l'utilisateur final, ces considérations psychoacoustiques peuvent *a priori* paraître absconses mais elles sont essentielles, étant la partie immergée de l'approche ambisonique. Les outils d'optimisation de champ sonore proposés par J. Daniel et les outils de critique du champ sonore, que sont les vecteurs vitesse et énergie, proposés par M. A. Gerzon en sont d'ailleurs directement issus.

---

<sup>69</sup> [Gerzon, 1977], p. 400, traduit de l'anglais : « [...] *satisfying relevant psychoacoustic criteria at the ears of the listener [, so that the design of decoders is primarily an engineering task] dependent on a knowledge of what acoustic stimuli are best capable of producing the desired illusion.* »

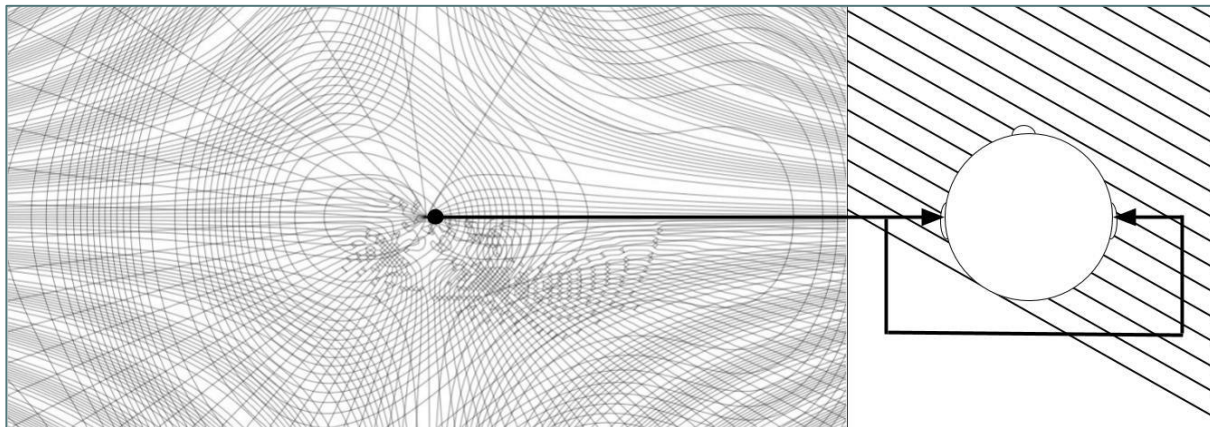


Figure 1.13. Simplification du modèle acoustique représentant le champ sonore complexe à gauche en ne considérant que localement aux environs de l'auditeur les phénomènes selon les mécanismes perceptifs de l'auditeur vibratoires à droite.

Afin de répondre aux spécifications plus techniques qui vont de la 4<sup>ème</sup> à la 8<sup>ème</sup>, une méta-spécification est établie définissant que le système doit comporter deux étapes, l'encodage et le décodage, offrant une représentation intermédiaire du champ sonore qui soit indépendante à la fois du dispositif d'enregistrement et du ou des dispositifs de restitution [Gerzon, 1977]. Il s'agit donc implicitement de posséder une représentation abstraite de l'espace sonore, comportant l'ensemble les informations directionnelles mais indépendante de toute réalité physique ou matérielle. L'enjeu étant par la suite projeter cette représentation du champ sonore sur un dispositif de diffusion de la meilleure manière possible afin de satisfaire les spécifications. Les 4<sup>ème</sup> et 7<sup>ème</sup> spécifications peuvent être satisfaites en associant au décodage les considérations psychoacoustiques et notamment les optimisations du champ sonore. Les 5<sup>ème</sup> et 6<sup>ème</sup> spécifications sont validées en adaptant les opérations de décodages aux dispositifs.

L'originalité de l'approche est alors de fonder la représentation intermédiaire du champ sonore sur une décomposition en harmoniques sphériques. De par leurs propriétés, qui seront présentés par la suite, cette approche est en adéquation avec la plupart des spécifications. La 8<sup>ème</sup> spécification est validée par le caractère « extensible » permettant d'obtenir une représentation à des degrés de précision variables. De même, il est important de noter que la décomposition en harmoniques sphériques satisfait relativement bien la 1<sup>ère</sup> car offrant une représentation sphérique de l'espace sonore mais pouvant se limiter à une représentation horizontale – dont l'approche est en accord avec la 8<sup>ème</sup> spécification car réduisant le nombre de canaux nécessaires. Enfin l'approche ne valorise pas une partie de l'espace par rapport à une autre permettant de satisfaire la 3<sup>ème</sup> spécification ce que remarque J. Daniel : « Outre l'avantage de fournir, grâce à la composante verticale Z, une description de la scène sonore en trois

*dimensions*, pouvant donner lieu à une restitution « *périphonique* », ce format a la propriété de proposer une *représentation homogène* des événements sonores, c'est-à-dire qui ne privilégie pas une direction plus qu'une autre »<sup>70</sup>.

L'enjeu est à présent de définir comment concrètement le système permet de répondre de la 1ère à la 4<sup>ème</sup> spécification. L'approche consiste à concevoir l'outil mathématique qu'est la décomposition en harmoniques sphériques et de les mettre en relation avec les modèles acoustiques et psychoacoustiques afin de comprendre comment l'approche ambisonique est structurée dans la pratique – et possiblement la critiquer. Cependant le but de cette recherche n'est pas purement intellectuel et curieux. Il s'agit, comme évoquer, d'appréhender le modèle pour pouvoir se l'approprier dans un contexte musicale et artistique. Alors il sera possible de revenir sur les objectifs initiaux et les spécifications présentées et de les envisager sous un angle nouveau pour les redéfinir et proposer au besoin de nouvelles formulations plus en adéquation avec les expérimentations et les approches présentées.

## 2. Modèles théoriques et pratiques

Ce chapitre présente les concepts clés nécessaires à la création et la mise en œuvre du modèle ambisonique, puis l'ensemble des mécanismes mis en jeu dans cet ensemble de techniques. Cette présentation est réalisée en deux étapes, qui recoupent deux domaines scientifiques.

Premièrement, les principes mathématiques dédiés à la spatialisation ambisonique sont exposés. De la formulation des équations des harmoniques sphériques à leur représentation dans l'espace, le but est d'offrir les outils nécessaires afin de pouvoir appréhender le modèle ambisonique. Cette première étape est nécessaire pour comprendre les mécanismes fondamentaux de l'ambisonie. Elle permettra notamment de visualiser dans l'espace les harmoniques sphériques, afin de comprendre leurs caractéristiques et les relations et mécanismes qui les régissent. Mais cette étape est surtout primordiale dans la suite de thèse, afin de pouvoir concrètement envisager la création de traitements et leur utilisation dans le domaine de la spatialisation du son ambisonique. Ainsi les outils mathématiques utilisés en ambisonie sont traités, de manière pleinement assumée, selon un angle didactique.

---

<sup>70</sup> [Daniel, 2001], p. 102.

Puis les modèles acoustiques, et plus spécifiquement psychoacoustiques, sur lesquels sont construits les techniques ambisoniques, sont présentés. L'enjeu ici est de pouvoir appréhender avec une oreille de musicien, la création du modèle de spatialisation ambisonique, afin d'en rendre intelligibles ses fondations. Cette approche, plus originale, propose de retracer les idées importantes et le cheminement scientifique qui ont permis de concevoir le modèle ambisonique dans son ensemble et pas seulement le champ sonore sous la forme de signaux associés aux harmoniques.

Enfin, suite à ces présentations, l'enjeu consistera à faire coïncider l'ensemble des savoirs présentés dans chaque domaine, afin de comprendre la nature des relations entre les mathématiques, la représentation de l'espace, les modèles acoustiques et les modèles psychoacoustiques. En somme, il s'agira simplement de montrer comment la représentation acoustique du champ sonore issu des théories psychoacoustiques est associable à une représentation de l'espace sous la forme d'une combinaison d'harmoniques sphériques.

Sur un plan théorique, l'ensemble des outils aura alors été présenté, permettant évidemment d'éclairer les approches pratiques de l'ambisonie – telles que notamment proposées par les opérations d'encodage ou de décodage – mais surtout de comprendre les opérations originales sur l'espace sonore, présentées dans le cadre de cette thèse, et de mettre en relief leurs conséquences.

Comme il est nécessaire à chaque fois de rappeler les enjeux de la présentation de ces modèles théoriques, il est aussi nécessaire de rappeler que les différentes présentations de ce chapitre ne visent pas à recouvrir l'état actuel des recherches dans ces domaines. Et encore moins de faire une présentation exhaustive de l'ensemble des connaissances ayant une relation plus ou moins lointaine avec la spatialisation du son. Les présentations se contraignent à fournir les concepts clefs qui semblent essentiels dans le cadre de ces travaux, et dans la limitation du savoir au moment où ils ont été utilisés. L'objectif n'est pas non plus d'éviter tout principe trop complexe, car la complexité est quoiqu'il advienne, inhérente aux notions présentées quelles qu'elles soient. Il s'agit simplement d'éviter de noyer le sujet sous un flot de considérations quelques peu hors-sujet<sup>71</sup>.

---

<sup>71</sup> Ici, la limite entre ce qui est nécessaire et ne l'est pas est bien sûr subjective mais j'espère avoir trouvé un juste milieu.

Avant d'introduire les notions mathématiques et géométriques, une brève présentation des systèmes de coordonnées et de la représentation de l'espace est indispensable, afin d'être sur la même longueur d'onde<sup>72</sup> et d'éviter toute confusion.

---

<sup>72</sup> Poum tchak !

## Espace et systèmes de coordonnées

L'espace peut être défini dans le cadre de ces travaux de deux manières : selon une approche psychoacoustique afin de décrire le rapport entre l'auditeur et les sources sonore. Et selon une approche géométrique, afin de pouvoir offrir des formulations mathématiques des mécanismes mis en jeu et des techniques utilisées. Ces deux approches se recoupent et possèdent évidemment des moyens de correspondance. De manière générale, les principes régissant les systèmes de coordonnées, suivent la représentation de l'espace définie par J. Blauert [Blauert, 1997] sont qualifiés d'acoustiques par J. C. Schacher dans la dernière version des outils ambisoniques de l'ICST [Schader, 2010] [Figure 2.1].

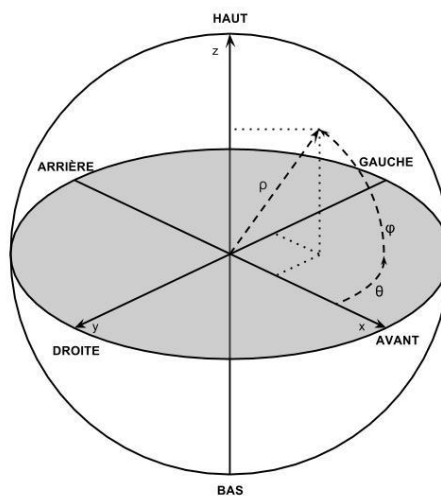


Figure 2.1. Représentation des relations dites acoustiques entre un système de coordonnées sphériques et un système de coordonnées cartésien avec des indications de directions relatives à un auditeur placé au centre du système.

L'espace, en psychoacoustique, est centré sur l'auditeur pour lequel il existe trois axes – l'axe avant-arrière, l'axe gauche-droite et l'axe haut-bas – qui composent trois plans :

- Le plan frontal qui est vertical sur l'axe droite-gauche et découpe l'espace entre l'arrière et l'avant.
- Le plan médian qui est vertical sur l'axe avant-arrière et découpe l'espace entre la gauche et la droite.
- Le plan horizontal à hauteur des oreilles qui découpe l'espace entre le bas et le haut.

En géométrie, l'espace peut être défini selon un système de coordonnées sphériques ou un système de coordonnées cartésiennes. Les coordonnées sphériques d'un point sont représentées par deux angles, un azimut – noté  $\theta$  – et qui définit l'angle sur le plan horizontal et l'élévation – noté  $\phi$  – qui définit l'angle depuis le plan horizontal, et une distance – notée  $\rho$  – depuis l'intersection des différents plans. Selon l'approche acoustique, l'azimut se déplace dans un sens antihoraire, avec 0 radian ou 0 degré à l'avant – tel que l'azimut vaut  $\frac{\pi}{2}$  radians ou 90 degrés à gauche,  $\pi$  radians ou 180 degrés à l'arrière et  $\frac{3\pi}{2}$  radians ou 270 degrés à droite. L'élévation se déplace du bas vers le haut avec  $-\frac{\pi}{2}$  radians ou  $-90$  degrés en bas,  $\frac{\pi}{2}$  radians ou 90 degrés en haut et 0 radian ou 0 degré sur le plan horizontal. Les coordonnées cartésiennes d'un point sont représentées par trois valeurs : l'abscisse notée  $x$ , l'ordonnée notée  $y$  et la hauteur  $z$ , correspondant à la projection de ce point sur les axes éponymes associés réciproquement aux axes arrière-avant, gauche-droite et bas-haut. Selon cette représentation, les coordonnées sphériques peuvent être obtenues à partir de coordonnées cartésiennes selon

$$\theta = \text{atan2}(y, x) - \frac{\pi}{2}$$

$$\phi = \text{asin} \frac{z}{\rho}$$

$$\rho = \sqrt{x^2 + y^2 + z^2}$$

Équation 2.1. Conversion de coordonnées cartésiennes vers des coordonnées sphériques.

Et réciproquement

$$x = \rho \cos\left(\theta + \frac{\pi}{2}\right) \cos \phi$$

$$y = \rho \sin\left(\theta + \frac{\pi}{2}\right) \cos \phi$$

$$z = \rho \sin \phi$$

Équation 2.2. Conversion de coordonnées sphériques vers des coordonnées cartésiennes.

## 2.1. Outils mathématiques et géométriques

La présentation mathématique des harmoniques sphériques et de leurs propriétés géométriques, peut sembler quelque peu austère et ardue dans un contexte artistique et musical, c'est pourquoi il semble nécessaire de la justifier dès à présent. D'une part, les harmoniques sphériques sont les outils élémentaires et centraux des recherches présentées dans ce travail de thèse. Il est donc primordial que le lecteur puisse non seulement avoir accès à leurs définitions mais surtout comprendre certaines caractéristiques et propriétés de ces fonctions. A cet effet, un soin particulier est pris, afin de faire preuve autant que possible de vulgarisation et de pédagogie. D'autre part, la formulation des harmoniques sphériques varie selon les domaines d'applications, amenant souvent à des confusions. Notamment d'un point de vue de la normalisation, engendrant des problèmes de compatibilités entre les différentes bibliothèques ambisoniques<sup>73</sup>. Ainsi, il semble important de préserver uniquement les informations claires et pertinentes, dans le cadre de ces recherches et en adéquation avec le modèle ambisonique.

### 2.1.1. Harmoniques sphériques

Les harmoniques sphériques  $Y$  sont un ensemble de fonctions définies sur la sphère unité. Ainsi prennent-ils en entrée deux angles : l'azimut  $\theta$  et l'élévation  $\phi$ . Ces fonctions sont organisées selon des degrés  $l$  comportant chacun  $2l + 1$  harmoniques auxquels sont associés un ordre azimutal  $m$  tel que  $-l \leq m \leq +l$  [Figure 2.2]. Les harmoniques sphériques sont alors exprimées mathématiquement dans un système de coordonnées acoustiques<sup>74</sup> par

$$Y_{l,m}(\theta, \phi) = P_{l,|m|}^{SN3D}(\sin\phi)e^{im\theta}$$

Équation 2.3. Formulation des harmoniques sphériques.

où  $P_{l,|m|}^{SN3D}$  sont les versions semi-normalisées au sens de Schmidt des polynômes associés de Legendre, correspondant à la partie liée à l'élévation de la formulation mathématique, et  $e^{im\theta}$  l'exponentielle complexe, correspondant à la partie liée à l'azimut de la formulation mathématique. Les polynômes associés de Legendre peuvent être exprimés de multiples façons, notamment en les déduisant des polynômes de Legendre qui sont eux-mêmes sujets à plusieurs

---

<sup>73</sup> C'est le cas pour la version actuelle de la bibliothèque HOA.

<sup>74</sup> Le système de coordonnées mathématiques utilise le cosinus de l'élévation entrée des polynômes associés de Legendre au lieu du sinus de l'élévation.



formulations<sup>75</sup>. La formulation la plus simple et offrant facilement des possibilités d'optimisation lors de la mise en œuvre, consiste en une approche récursive degré après degré telle que

$$\begin{aligned}
 P_{0,0}(x) &= 1 \\
 P_{l+1,l+1}(x) &= -(2l+1)\sqrt{1-x^2}P_{l,l}(x) \\
 P_{l+1,l}(x) &= x(2l+1)P_{l,l}(x) \\
 P_{l+1,m}(x) &= \frac{(2l+1)xP_{l,m}(x) - (l+m)P_{l-1,m}(x)}{l-m+1}
 \end{aligned}$$

Équation 2.4. Formulation des polynômes associés de Legendre.

Dans la pratique, notamment parce que de nombreux usages sont limités aux premiers degrés, des formulations simples et directes – n'utilisant que des opérations élémentaires, l'exponentiation ou la racine carrée – sont prédéfinies pour les polynômes des trois à cinq premiers degrés<sup>76</sup>.

La semi-normalisation au sens de Schmidt<sup>77</sup>, aussi appelée en ambisonie semi-normalisation 3D (SN3D), est définie par

$$k_{l,m}^{SN3D} = \sqrt{\frac{(l-|m|)!}{(l+|m|)!}} \sigma \text{ avec } \sigma = 1 \text{ si } m = 0 \text{ sinon } \sigma = 2$$

Équation 2.5. Formulation de la semi-normalisation 3D.

---

<sup>75</sup> Il y a par exemple la formule de récurrence de Bonnet ou encore la formule de O. Rodrigues [Askey, 2005].

<sup>76</sup> Il est aisément possible d'en définir, en développant à l'avance les formulations récursives. La dernière mise en œuvre, version 3, de la bibliothèque HOA utilise cette approche jusqu'à l'ordre 5, recouvrant les usages les plus répandus, puis une mise en œuvre optimisée des formulations récursives pour les ordres supérieurs.

<sup>77</sup> Cette normalisation est présentée notamment par J. Daniel [Daniel, 2001] et représente la semi-normalisation 3D (SN3D) utilisée dans le format ambiX.

Ainsi

$$P_{l,|m|}^{SN3D}(x) = P_{l,m}(x)k_{l,m}^{SN3D}$$

Équation 2.6. Formulation des polynômes associés de Legendre semi-normalisés.

La normalisation complète des polynômes associés de Legendre, aussi appelée normalisation 3D (N3D) en ambisonie, notamment utilisées lors des opérations présentées par la suite, telles que la projection ou encore la rotation, peut être exprimée à partir de la semi-normalisation 3D par

$$k_{l,m}^{N3D} = (2l + 1)k_{l,m}^{SN3D}$$

Équation 2.7. Formulation de la normalisation 3D.

Dans les travaux présentés, il est assumé que, par défaut, les harmoniques sont semi-normalisées. Les harmoniques normalisées – utilisant les polynômes associés de Legendre normalisés sur la sphère – pourront être notés  $Y^{N3D}$ .

Enfin l'exponentielle complexe peut être exprimée sous une forme réelle par

$$e^{im\theta} = \cos(|m|\theta) \text{ si } m \geq 0$$
$$e^{im\theta} = \sin(|m|\theta) \text{ si } m < 0$$

Équation 2.8. Formulation l'exponentiel complexe.

Pour des raisons de pérennité, il est important de remarquer que les présentes formulations suivent la norme de l'ambisonie, définie par le format ambiX [Nachbar & al., 2016]. Cela impliquant notamment l'utilisation d'un système de coordonnées acoustiques au lieu d'un système de coordonnées mathématiques, la semi-normalisation 3D, l'omission du facteur de phase de Condon-Shortley de  $(-1)^l$  ainsi que du facteur de normalisation de  $\frac{1}{\sqrt{4\pi}}$ <sup>78</sup>.

---

<sup>78</sup> La version actuelle destinée à l'utilisateur final de la bibliothèque HOA ne tient pas en compte de toutes ces spécifications, un correctif nommé *Patch SN3D* pour les logiciels Max et Pure Data, publié sur le site du projet, assure la compatibilité avec le format ambiX et présente de plus les implications de ces changements. La dernière version en ligne du code C++ intègre quant à lui directement l'ensemble de ces spécifications.

## 2.1.2. Représentation spatiale

Étant définis sur la sphère, il est possible de représenter les harmoniques sphériques dans l'espace, afin notamment de mettre en avant leurs propriétés géométriques. Il existe deux manières principales de procéder : la représentation bidimensionnelle en coupe ou la représentation tridimensionnelle<sup>79</sup>. Alors qu'une coupe est particulièrement adaptée pour appréhender localement une fonction, la représentation tridimensionnelle semble plus adaptée pour concevoir l'espace dans son entièreté – bien qu'une partie reste forcément invisible – la symétrie de ces fonctions permet de compenser cet artefact visuel. La représentation tridimensionnelle implique encore un choix. Il est possible d'associer les sorties des fonctions à un gradient de couleur sur la sphère, mais l'approche semble peu adaptée à la compréhension des harmoniques sphériques dans le domaine acoustique. Alors que le fait d'associer les sorties à la distance du tracé par rapport au centre, en utilisant la couleur pour définir la polarité de la fonction, permet d'offrir des représentations analogues aux directivités des microphones auxquelles elles sont associées en ambisonie [Figure 2.2]<sup>80</sup>.

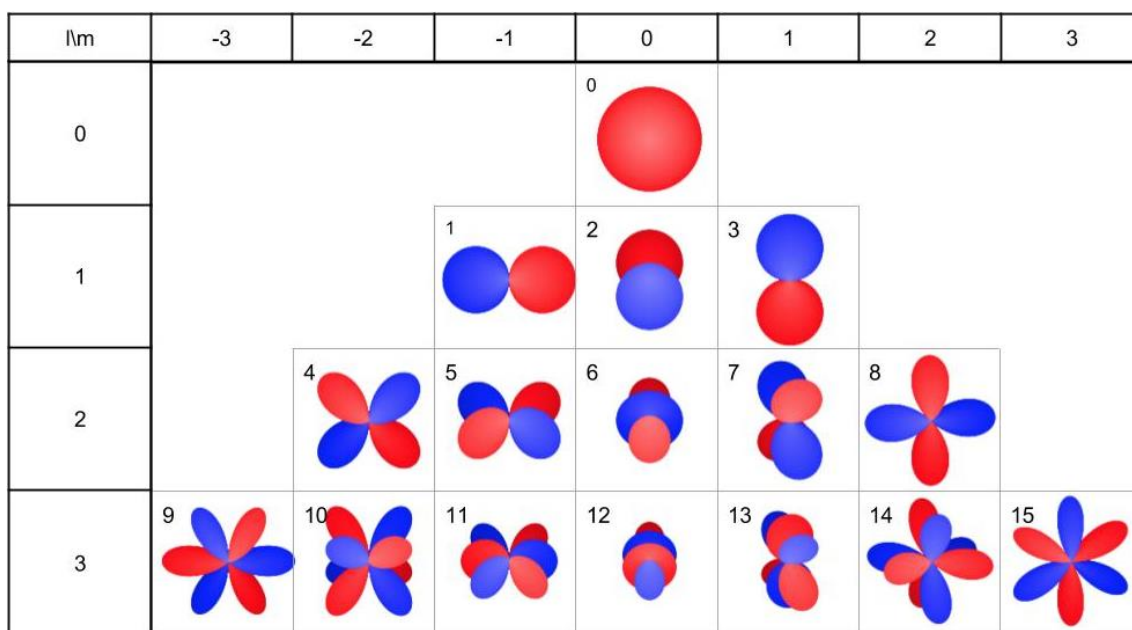


Figure 2.2. Représentation des 16 premiers harmoniques sphériques, en vue du dessus avec des légères inclinaisons, allant de haut en bas du degré 0 au degré 3 avec les ordres azimutaux croissant de gauche à droite. En rouge sont définies les parties où les fonctions sont positives et en bleu les parties où les fonctions sont négatives.

<sup>79</sup> D'autres représentations sont évidemment possible, il s'agit de présenter les possibilités généralement usitées en l'ambisonie.

<sup>80</sup> Du moins pour les harmoniques de degrés 0 et 1, les autres tracés pouvant être associés à des figures de directivités de microphones fictifs.

Grâce à leurs représentations, les harmoniques sphériques apparaissent comme des fonctions orthogonales et offrent des propriétés géométriques intéressantes<sup>81</sup>, telles que la symétrie et la rotation. Ainsi les harmoniques sphériques d'ordre azimutal 0 sont invariantes en rotation autour de l'axe z ou encore les harmoniques d'ordres azimutaux positifs sont symétriques sur l'axe y<sup>82</sup>. Mais l'enjeu principal est de pouvoir concevoir et appréhender les figures et les directions qui leurs sont associés, notamment lorsqu'il s'agira de les jouer afin de créer, moduler et transformer des espaces sonores :

- L'harmonique de degré 0 correspond à une figure omnidirectionnelle
- Les harmoniques d'ordre 1 correspondent à des figures de huit.
- Pour un ordre azimutal donné, le nombre de lobes de chaque harmonique augmente avec le degré.
- Pour un degré donné les harmoniques, les harmoniques d'ordre azimutal  $m$  possèdent un harmonique  $-m$  équivalent à une rotation près.
- Les harmoniques d'ordre azimutal 0 sont « orientés » sur l'axe z.
- Les harmoniques dont la valeur absolue de l'ordre azimutal est égale à leur degré, tel que  $|m| = l$ , sont « parallèles » au plan azimutal.

Les harmoniques sphériques possèdent bien d'autres caractéristiques spatiales notables et il est impossible de toutes les énumérer. Il convient au final à chacun de remarquer et définir celles qu'il considère pertinentes pour son approche.

### 2.1.3. Harmoniques circulaires

La représentation spatiale des harmoniques sphériques met en valeur certaines spécificités spatiales, dont l'une vient d'être citée : « les harmoniques dont la valeur absolue de l'ordre azimutal est égale à leur degré,  $|m| = l$ , sont « parallèles » au plan azimutal ». En des termes

---

<sup>81</sup> Il est évidemment possible de retrouver certaines propriétés mathématiquement mais c'est dans le cas présent non nécessaire.

<sup>82</sup> Une présentation plus exhaustive des symétries est présentée par F. Zotter en ligne <https://ambisonics.iem.at/xchange/fileformat/docs/spherical-harmonics-symmetries> (consulté dernièrement en septembre 2017).

mathématiquement plus exacts, les droites formées par les extrema des fonctions – les sommets des lobes – sont coplanaires et définissent le plan azimutal. Elles possèdent comme propriétés que pour un azimut donné, la valeur extrême est toujours obtenue lorsque l'élévation est nulle car selon l'Équation 2.4,  $P_{l,l}(0) = 1$ . Sous cette condition, la partie relative à l'élévation étant négligeable sur ce sous-ensemble d'harmoniques, il est possible de définir un modèle strictement circulaire prenant en compte uniquement la partie relative à l'azimut tel que

$$Y_m^{circulaire}(\theta) = e^{i|m|\theta}$$

Équation 2.9. Formulation des harmoniques circulaires selon la définition de l'Équation 2.8].

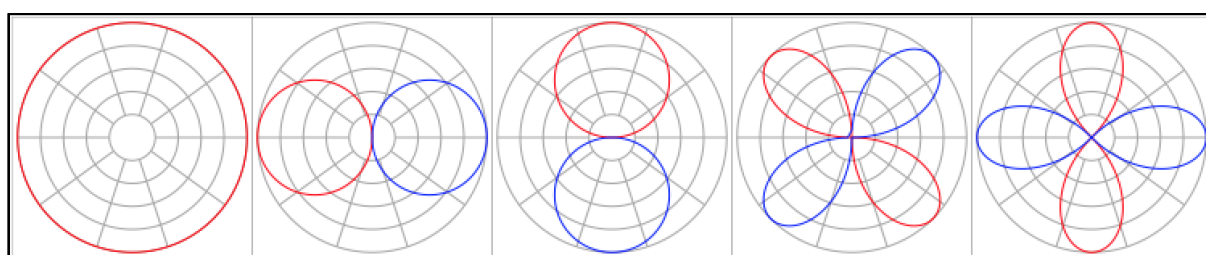


Figure 2.3. Représentation des 5 premières harmoniques circulaires d'ordre azimutal 0, -1, 1, -2 et 2. En rouge sont définies les parties où les fonctions sont positives et en bleu les parties où les fonctions sont négatives.

Ces harmoniques sont dites alors circulaires<sup>83</sup> [Figure 2.3] et peuvent être associées uniquement à un ordre azimutal  $m$  étant donné que chaque degré  $l$  ne comporte plus que deux harmoniques d'ordre azimutal  $-l$  et  $l$ , à l'exception du degré 0 que ne possède que l'harmonique d'ordre azimutal 0. Enfin, pouvant être considérés comme un sous ensemble des harmoniques sphériques, les harmoniques circulaires ne peuvent pas représenter les harmoniques sphériques sans perte d'information<sup>84</sup>. La réciproque étant quant à elle valable.

<sup>83</sup> Il est possible de trouver dans la littérature le terme d'harmoniques cylindriques tel que dans [Daniel, 2001].

<sup>84</sup> Cette propriété est importante dans le cas d'un passage d'une représentation d'une scène sonore tridimensionnelle à une représentation bidimensionnelle.

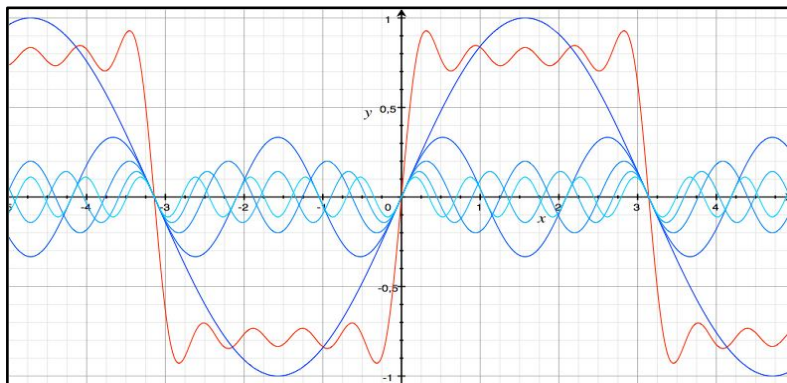


Figure 2.4. Représentation du développement en série de Fourier d'une fonction périodique en rouge, approximant un signal carré, par un ensemble de fonctions sinusoïdales en bleue.

### 2.1.4. Décomposition en harmoniques sphériques

Les harmoniques sphériques sont les analogues sphériques des fonctions sinusoïdales, permettant de représenter une fonction définie sur un cercle via l'analyse en série de Fourier. Ainsi, tout comme le développement en série de Fourier permet de représenter une fonction périodique par une somme pondérée de fonctions sinusoïdales [Figure 2.4], le développement en harmoniques sphériques permet de représenter une fonction continue sur la sphère par une somme pondérée d'harmoniques sphériques [Figure 2.5].

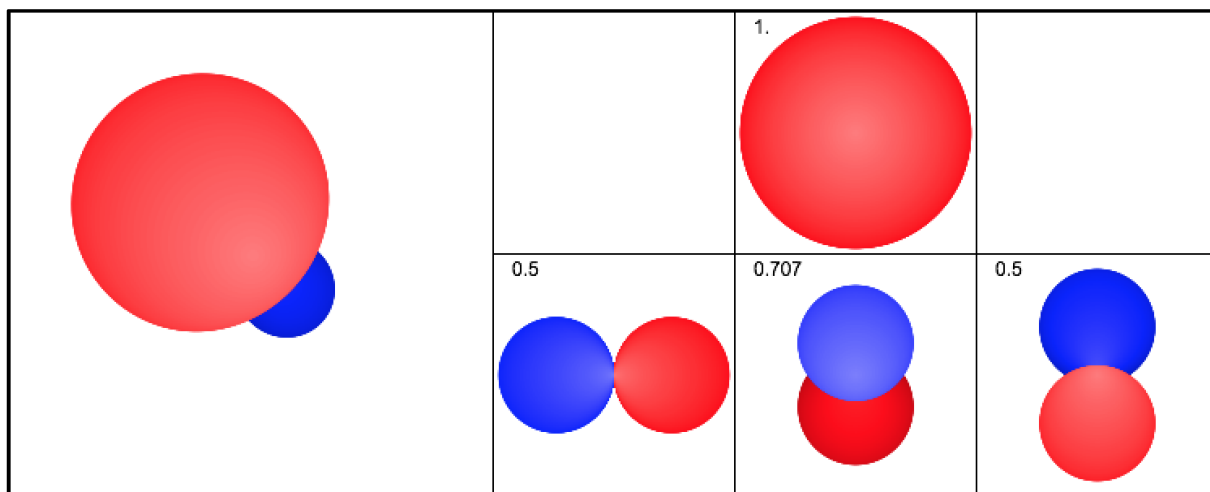


Figure 2.5. Représentation du développement en série d'harmoniques sphériques d'une fonction sphérique à gauche – définissant une figure hypercardioïde dirigée vers  $\frac{\pi}{4}$  d'azimut et  $\frac{\pi}{4}$  d'élévation – par les trois premiers harmoniques sphériques de degrés 0 et 1 à droite et dont les amplitudes sont représentées au-dessus.

De manière analogue à la transformation de Fourier discrète – où un ensemble d'échantillons d'un signal peut être représenté par un ensemble fini de fonctions trigonométriques – la décomposition en harmoniques sphériques peut, elle aussi, être réalisée avec un ensemble fini

d'éléments – ou points – représentant la fonction sur la sphère. Contrairement au développement en harmoniques sphériques, dans ce contexte la décomposition en harmoniques est une opération considérée comme s'appliquant sur un ensemble d'éléments discrets, non continus<sup>85</sup>. L'opération consiste alors à définir les amplitudes des harmoniques sphériques – équivalent du spectre fréquentiel d'une transformation de Fourier, généralement présentée par les amplitudes et les phases des fonctions sinusoïdales – en fonction de l'ensemble de ces éléments auxquels sont associés une valeur – ou une amplitude – et une direction, l'emplacement sur la sphère correspondant à l'élément. L'amplitude d'une harmonique revient alors à calculer la somme des valeurs de tous les éléments, pondérée par la sortie de l'harmonique dans la direction des éléments tel que

$$a_{l,m} = \sum_{i=1}^I Y_{l,m}(\theta_i, \phi_i) x_i$$

Équation 2.10. Calcul de l'amplitude d'un harmonique sphérique à partir des amplitudes des éléments d'une fonction sphérique.

Avec  $\theta_i$  l'azimut du  $i^{\text{ème}}$  élément,  $\phi_i$  l'élévation du  $i^{\text{ème}}$  élément et  $x_i$  l'amplitude du  $i^{\text{ème}}$  élément.

En continuant les analogies avec la transformation de Fourier discrète, le nombre d'échantillons du signal utilisé permet de définir le nombre de bandes spectrales, donc leurs largeurs et la précision spectrale. Augmenter le nombre d'harmoniques sphériques permet d'accroître la précision de la décomposition<sup>86</sup>. La décomposition en harmoniques sphériques est ainsi réalisée selon un ordre de décomposition<sup>87</sup>  $N$  définissant le nombre de degrés utilisés pour représenter la fonction sphérique. Etant donné que chaque degré possède  $2l + 1$  harmoniques sphériques, le nombre d'harmoniques sphériques utilisé est  $(N + 1)^2$ . E il est possible de définir que pour une décomposition dans un espace bidimensionnel, le nombre d'harmoniques circulaires est  $2N + 1$ . Tout comme lors d'une transformée discrète où le nombre de points représentant le signal temporel détermine le nombre de bandes fréquentielles, il est important que le nombre

---

<sup>85</sup> Il est néanmoins important de noter que cette définition n'est valable que dans ce contexte pour des raisons de commodité et d'usage, en mathématique une fonction peut être décomposée par une série infinie d'harmoniques sphériques.

<sup>86</sup> Cette caractéristique peut d'ores et déjà être mise en relation avec l'amincissement des lobes sur les représentations des harmoniques sphériques avec l'augmentation du degré.

<sup>87</sup> Afin d'éviter toutes confusions avec l'ordre azimutal présenté par la suite, le choix a été fait d'utiliser autant que possible le terme complet d'ordre de décomposition.

de point  $I$  sur la sphère représentant la fonction à décomposer, soit supérieur ou égal au nombre d'harmoniques utilisées – et donc que  $N = \sqrt{l} - 1$  pour une fonction sphérique ou  $N = \frac{l-1}{2}$  pour une fonction circulaire – sans quoi l'information associée aux harmoniques de degrés supérieurs devient redondante. Enfin, une décomposition parfaite en harmoniques sphériques assume que les points doivent discrétiser de manière parfaite la sphère – autrement dit que tous les points soient à la même distance angulaire – sans quoi des artefacts se produisent<sup>88</sup>. Principe restreignant alors les configurations envisageables de points sur une sphère aux cinq solides de Platon<sup>89</sup> [Figure 2.6], mais offrant une infinité de possibilités sur le cercle.

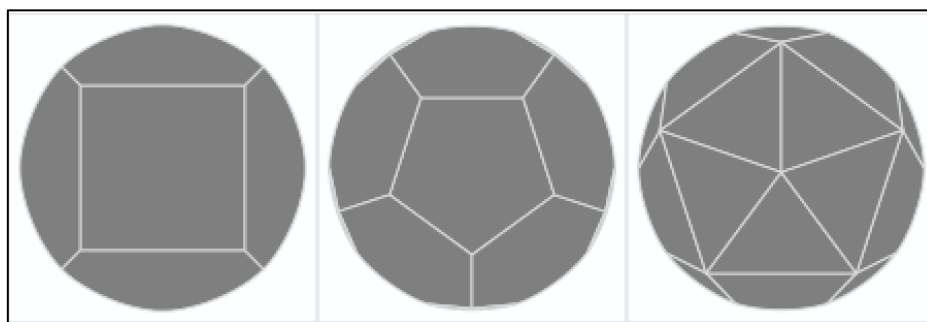


Figure 2.6. Représentation en vue du dessus de la discrétisation parfaite d'une sphère par trois des cinq solides de Platon, avec de gauche à droite : le cube, le dodécaèdre et l'icosaèdre.

Enfin il est possible de réaliser l'opération réciproque – l'analogie sphérique d'une transformation de Fourier inverse – pour récupérer à partir des coefficients d'amplitude des harmoniques sphériques, les valeurs d'amplitude des éléments discrétisant la fonction initiale. L'opération consiste alors en une projection des harmoniques sphériques dans la direction donnée telle que

$$x_i = \sum_{l=0}^N \sum_{m=-l}^l (2l+1) Y_{l,m}(\theta_i, \phi_i) a_{l,m}$$

Équation 2.11. Calcul de l'amplitude d'un élément d'une fonction à partir des amplitudes des harmoniques sphériques.

Avec  $(2l+1)$  le facteur de normalisation 3D tel que présenté [Équation 2.7].

<sup>88</sup> Par analogie, il est possible d'envisager un signal qui serait échantillonné à une fréquence variable impliquant nécessairement une déformation dans la reconstruction de la forme d'onde.

<sup>89</sup> Néanmoins au prix d'une marge d'erreur, il est possible d'approximer des configurations parfaites avec de grands nombres de points.



Projeter les harmoniques sphériques dans les directions définies par les éléments lors de la décomposition, permet d’obtenir la discrétisation initiale de la fonction. Mais il est aussi possible d’augmenter le nombre d’éléments ou d’appliquer une rotation sur l’ensemble d’entre eux, tant que la discrétisation finale respecte les principes de régularité préalablement établis.

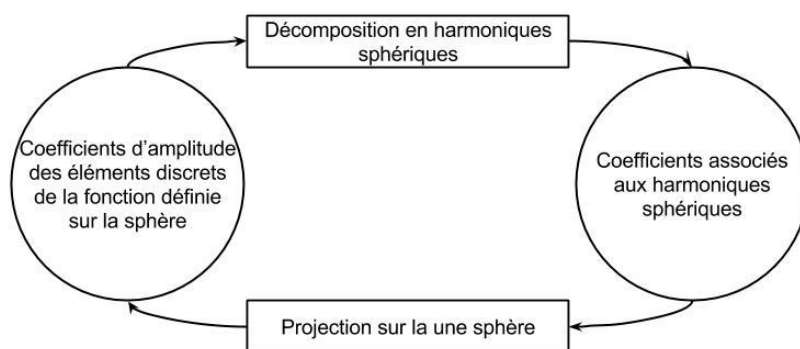


Figure 2.7. Représentation du cycle des opérations autour des harmoniques sphériques et des éléments discrets d’une fonction définie sur la sphère.

En résumé, les harmoniques sphériques offrent via l’opération de décomposition, une représentation alternative des fonctions, ou d’un ensemble d’éléments représentatifs, définis sur la sphère. Cette représentation alternative prend la forme d’une série de coefficients associés aux harmoniques sphériques permettant, tout comme le spectre fréquentiel obtenu via une transformée de Fourier, d’analyser et de modifier la fonction initiale. Cette opération est réversible via l’opération de projection qui permet de retrouver la représentation originale de la fonction [Figure 2.7].

## 2.2. Construction du modèle ambisonique

Le modèle de spatialisation ambisonique a été construit dans un contexte particulier, visant notamment à améliorer l’approche quadriphonique et proposer une approche alternative plus stable et cohérente selon les spécifications établies dans le premier chapitre de cette partie. Les modèles psychoacoustiques établis notamment par M. A. Gerzon sont à considérer dans ce contexte lié aux années 1970. Dès lors, les mécanismes en jeu peuvent paraître répondre à des problématiques obsolètes et faire usage de notions qui sont aujourd’hui bien plus approfondies. Alors que la décomposition en harmoniques sphériques aux ordres supérieurs – jusqu’à l’ordre de décomposition 3 – du champ sonore a déjà été présentée, mais cependant en dehors du cadre conceptuel ambisonique [Gerzon, 1973], les modèles psychoacoustiques établis permettent

d'associer la représentation du champ sonore associé à une décomposition en harmoniques sphériques, limitée à l'ordre 1. Notamment du fait qu'il se place dans un contexte de restitution sur quatre haut-parleurs [Gerzon, 1975c]. La raison derrière cette limitation initiale est donc *a priori* matérielle. Elle peut par là même être potentiellement esquivée, et la présentation des mécanismes pourrait se concentrer sur une approche directement associée aux harmoniques sphériques d'ordres supérieurs. Cependant, comme cela a été évoqué, la limitation du modèle mathématique pour une représentation tridimensionnelle est d'ordre 3<sup>90</sup>. De plus les limitations matérielles ne sont et ne peuvent toujours pas être résolues aujourd'hui. Sauf au moyen de solutions et d'approximations qui dérivent du modèle original. Il semble donc plus pertinent d'un point de vue didactique, mais aussi conceptuel, de suivre la même approche que M. A. Gerzon en commençant par le premier ordre de décomposition. Où sur un plan à la fois matériel et mathématique, les principes sont valides, pour par la suite étendre l'approche au moyen des ordres supérieurs en considérant les limites que cela impose.

Les outils mathématiques que sont les harmoniques sphériques, les opérations de décomposition et de projection qui leurs sont liées et leurs représentations spatiales ont déjà été présentées. L'enjeu consiste alors à revenir sur les modèles acoustiques et psychoacoustiques sur lesquels reposent les techniques ambisoniques, éléments centraux permettant d'utiliser par la suite la décomposition de l'espace en harmoniques sphériques, d'en concevoir les implications et les limites. Et permettant ensuite d'introduire le processus même de la construction du modèle ambisonique autour de modèles et théories psychoacoustiques.

### 2.2.1. Modèles psychoacoustiques

La perception spatiale des sons et de l'espace sonore environnant fait appel à de nombreux mécanismes. Comme le définit M. A. Gerzon : « On suppose que les oreilles n'ont pas une seule méthode de localisation des sons, mais que de nombreuses méthodes différentes sont utilisées. »<sup>91</sup>. En effet, selon Lord Rayleigh [Rayleigh, 1876], les systèmes auditifs semblent être fondés sur différents mécanismes, en fonction du domaine fréquentiel. Ainsi, les mécanismes de la perception auditive de l'espace ne fonctionnent pas de la même manière en dessous de 700 hertz qu'au-dessus. De plus, au-dessus de 5000 hertz, de nouveaux mécanismes apparaissent, et sont encore mal définis au moment de l'établissement de l'approche

---

<sup>90</sup> Raison notamment pour laquelle M. A. Gerzon se limite aussi à cet ordre de décomposition dans sa publication de 1973.

<sup>91</sup> [Gerzon, 1992], p. 3., traduit de l'anglais : « It is assumed that the ears have no single method of localizing sounds, but that many different methods are used. »

ambisonique. Ce chapitre revient donc sur les différents modèles psychoacoustiques sur lesquels repose l'ambisonie, pour par la suite les associer à l'approche fréquentielle proposée par M. A. Gerzon et établir ainsi le modèle ambisonique. Le sujet de la psychoacoustique n'étant pas au cœur de ce travail de recherche, l'approche se borne à présenter les concepts essentiels à la compréhension du modèle ambisonique, sans forcément tenter d'embrasser l'ensemble des connaissances plus avancées. Néanmoins, autant que possible, la terminologie actuelle sera utilisée afin de clarifier les propos.

### 2.2.2. Indices de localisation

En psychoacoustique, des indices physiques et perceptifs issus de l'analyse combinée des sons parvenant aux deux oreilles, permettent au système auditif de déterminer la position spatiale des sons : les indices binauraux. Ces derniers sont ceux opposés aux indices monauraux, issus de l'analyse exclusive d'une seule oreille. Ces indices sont des évaluations des différences entre l'oreille gauche et l'oreille des signaux les atteignant.

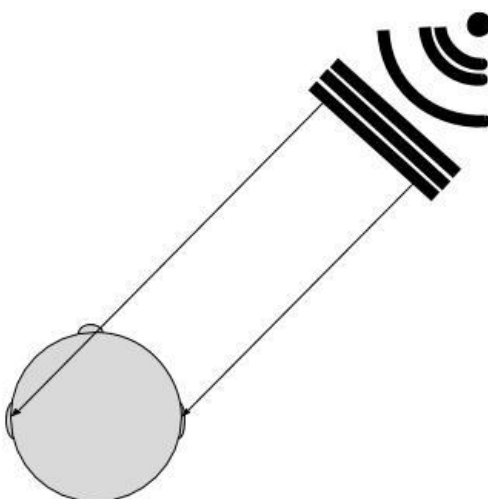


Figure 2.8 : Représentation de la différence de distances des deux oreilles par rapport à une source sonore située à 45 degrés à droite de l'auditeur, impliquant une différence de temps entre les signaux arrivant à chaque oreille.

Afin de concevoir les mécanismes en jeu, il est possible de comprendre le fonctionnement de ces indices en imaginant un son localisé vers l'avant-gauche de la tête à 45 degrés d'azimut et 0 degré d'élévation. Ce son atteindra alors l'oreille gauche légèrement avant l'oreille droite et, de même, son niveau sonore sera légèrement plus accru à l'oreille gauche qu'à l'oreille droite. Ces deux indices, qui se nomment différence de temps interaurale<sup>92</sup> (ITD) – différence de phase

---

<sup>92</sup> *Interaural Time Difference.*

interaurale si elle est définie relativement à un cycle<sup>93</sup> (IPD) – et différence d'intensité interaurale<sup>94</sup> (IID) pour la seconde – ou différence de niveau interaurale<sup>95</sup> (ILD) – lorsque spécifiée en décibel [Moore, 2013], p. 247. En assumant une deuxième position à la source sonore, cette fois proche de 0 degré d'azimut, alors la source sonore se retrouve presque à égale distance azimutale des deux oreilles et les différences de temps et d'intensité des signaux de la source sonore entre les deux oreilles sont moindres et quasiment nulles. Aussi, les différences de temps et d'intensité interaurale – et notamment leurs variations relatives – permettent au système auditif humain de localiser les sources, notamment sur un plan azimutal dans les processus de latéralisation.

Le mécanisme justifiant la différence de temps entre le signal perçu par l'oreille gauche et celui perçu par l'oreille droite, réside dans l'écart entre les deux oreilles. Cet écart implique que lorsqu'une source sonore est latéralisée vers une oreille – plus vers la gauche ou vers la droite – la distance entre la source sonore et cette oreille est plus petite que la distance entre la source sonore et l'autre oreille. Le son se déplaçant, il va mettre légèrement plus de temps à atteindre la deuxième oreille en raison de la distance supplémentaire à parcourir. En considérant la vitesse du son, généralement définie à 340,29 mètres par seconde, et en connaissant la distance entre les deux oreilles, il est possible de définir la différence de temps interaurale telle que présentée pour une onde plane sur la [Figure 2.8] permettant au système auditif de localiser les sons<sup>96</sup>.

Les différences d'intensités sont quant à elles, dues aux effets de perturbations engendrés par la tête humaine qui, comme le suggère J. Daniel, « la tête (de l'auditeur, avec ses oreilles) est un « instrument de mesure » volumineux qui perturbe fortement le champ sonore ». [Daniel, 2001], p. 105. En somme, le son traverse la tête à des endroits différents selon sa direction avant, d'atteindre chaque oreille. A chaque direction peut être associée un mode de perturbation pour chaque oreille, qui peut être associé à un filtre variant plus ou moins selon l'intensité du

---

<sup>93</sup> *Interaural Phase Difference.*

<sup>94</sup> *Interaural Intensity Difference.*

<sup>95</sup> *Interaural Level Difference.*

<sup>96</sup> Il est à noter une lacune dans ce mécanisme de localisation. Si la source sonore est composée d'ondes sinusoïdales dont la période équivaut à la différence de temps interaurale alors les deux signaux reçus par les oreilles sont identiques et empêchent la localisation via cet indice. Par analogie avec le système visuel, il est possible de considérer deux formes identiques mais de tailles différentes arrivant chacune dans un œil, tel que la plus grosse soit plus éloignée que la plus petite. Donnant ainsi l'illusion qu'elles sont à la même distance.

son. Il est alors de même possible de représenter ces différences d'intensité interaurales en fonction de la direction azimutale du son.

Néanmoins, comme le note B. C. J. Moore, les mécanismes sont complexes et dépendent de la nature des sons et ne sont pas effectifs de la même manière sur l'ensemble des fréquences. Il convient donc de traiter ces données, si ce n'est selon un domaine fréquentiel continu, au moins selon des bandes de fréquences. C'est l'approche adoptée par M. A. Gerzon en ambisonie. Mais avant de revenir dessus, il faut noter le défaut vis à vis de ces seuls indices binauraux. En considérant la tête comme exactement symétrique<sup>97</sup>, les différences interaurales se trouvent nulles sur le plan médian où les sources sont à égale distance des deux oreilles et le système auditif est alors incapable de différencier les événements se produisant le long de ce plan, à l'aide seul de ces indices. Les problèmes possibles, si seulement ces indices existaient, seraient une mauvaise reconnaissance de l'élévation des sources sonore, mais surtout une grande confusion entre l'avant et l'arrière.

C'est ici qu'entrent en jeu les indices physiques et perceptifs monauraux, issus de l'analyse exclusive de chaque oreille. De manière analogue à la différence d'intensité interaurale, ces indices sont la résultante des perturbations générées par la tête, les oreilles (les lobes) et le buste, sur les signaux sonores. L'idée étant que ces perturbations des phénomènes vibratoires sont différentes selon l'incidence de la source sonore. En raison de la forme du corps humain, les perturbations sur le signal ne sont pas du même ordre pour une source sonore provenant de derrière, de devant ou d'au-dessus de l'auditeur. De façon concise<sup>98</sup>, l'approche du système perceptif consiste à associer l'incidence de la source à une fonction de filtrage spécifique, aidant à définir la position des sources sonores sur le plan médian<sup>99</sup>. Néanmoins, les altérations du signal ne sont compréhensibles par le système auditif qu'en comparaison d'un référentiel. En schématisant, le problème pourrait se révéler simplement tel qu'en écoutant un son sans une bande de fréquences spécifique. L'auditeur ne peut savoir si cette bande de fréquences a été enlevée du son d'origine, ou si elle est initialement absente. Il en est de même pour le système auditif : les filtrages appliqués au son ne peuvent être interprétés qu'à l'écoute de leurs variations, donc pour une source sonore se déplaçant. Le problème étant, dans le cas d'une

---

<sup>97</sup> Ce qui n'est pas le cas bien sûr mais cette approximation peut être valable au regard de la précision de calcul du système auditif.

<sup>98</sup> Pour présentation approfondie des mécanismes en jeu, se référer à [Moore, 2013], p. 263-266.

<sup>99</sup> Et a priori ces indices ne sont pas exclusivement utilisés par le mécanisme de perception auditif pour la localisation sur le plan médian, mais aussi sur le plan frontal et horizontal.

source sonore statique, que le filtre associé à la position de la source ne peut être reconnu par le système perceptif.

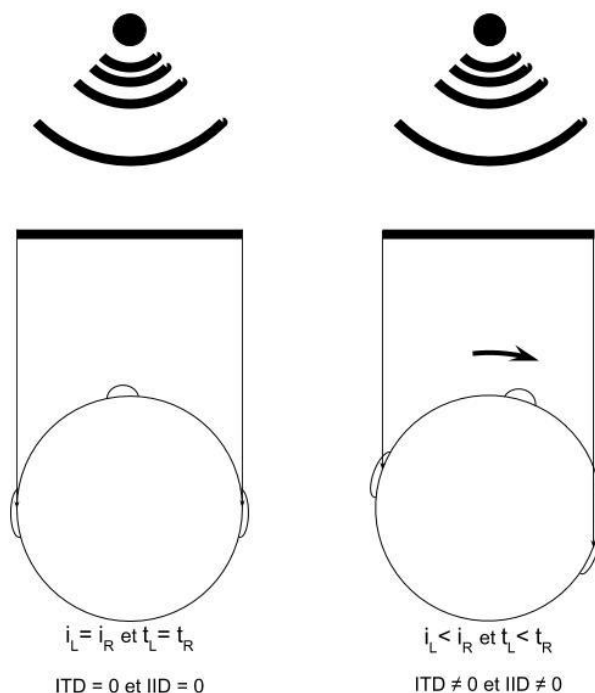


Figure 2.9 : Représentation des variations de l'ITD et l'IID selon les mouvements de la tête. Sur la figure de gauche, la tête est orientée en direction de la source sonore, le signal de la source sonore arrive en même temps aux deux oreilles. Et avec la même intensité, l'ITD et l'IID sont donc nuls empêchant la localisation correcte de la source sonore. Sur la figure de droite, la tête pivote légèrement vers la droite, le signal de la source sonore arrive légèrement avant à l'oreille gauche qu'à l'oreille droite et avec une intensité un peu plus forte. Engendrant ainsi un ITD et un IID non nul permettant la localisation correcte de la source sonore.

Néanmoins, il existe un dernier moyen utilisé par le système perceptif pour compenser cet artefact en intégrant les mouvements de la tête. Si la tête pivote autour de l'axe vertical  $z$ , alors l'ensemble des sources sonores subissent par rapport aux oreilles une rotation inverse. Aussi, les sources sonores situées sur l'axe médian et impliquant des indices binauraux nuls – notamment celles se trouvant à l'avant ou à l'arrière de la tête – subissent un décalage impliquant une différence de temps et d'intensité entre les signaux reçus respectivement par chaque oreille et donc des indices binauraux non-nuls. De même, les indices monauraux se trouvent modulés par ce mouvement de rotation, aidant de même les mécanismes de localisation [Figure 2.9]. En somme, la conséquence de ces rotations est la variation des indices binauraux et monauraux, qui permettent au système de perception auditif –en intégrant les informations de mouvement de la tête – d'analyser de manière plus robuste la direction de la source sonore. Notamment dans les directions préalablement critiquées, situées sur le plan médian.

L'ensemble de ces indices, à priori, permettent donc au système perceptif de discerner les directions des sources. Cependant, comme suggéré précédemment, ces mécanismes ne sont pas tous égaux sur le plan fréquentiel. M. A. Gerzon propose alors des domaines de validité de ces indices de localisation auditifs, permettant de définir les informations pertinentes à conserver du champ sonore.

### 2.2.3. Approche fréquentielle

Tout d'abord, M. A. Gerzon part du postulat qu'en considérant la vitesse de propagation du son, généralement de 340,29 mètres par secondes, il est possible de définir que la distance entre les deux oreilles, entre 15 et 20 centimètres, correspond approximativement à une demi-longueur d'onde de 700 hertz [Gerzon, 1974]. En utilisant notamment les travaux de Lord Rayleigh [Rayleigh, 1907] définissant qu'en dessous de cette fréquence, la tête n'est plus un obstacle aux ondes sonores, les différences d'amplitude interaurale sont alors négligeables, virtuellement identiques. Et n'offrant *a priori* que les indices binauraux temporels aux systèmes de perceptions auditives, pour localiser la direction des sources sonores. En associant ce retard de temps à un décalage de phase établi selon M. A. Gerzon, la seule information disponible pour la localisation du son à ces fréquences basses, ce sont les différences de phases des signaux atteignant chaque oreille.

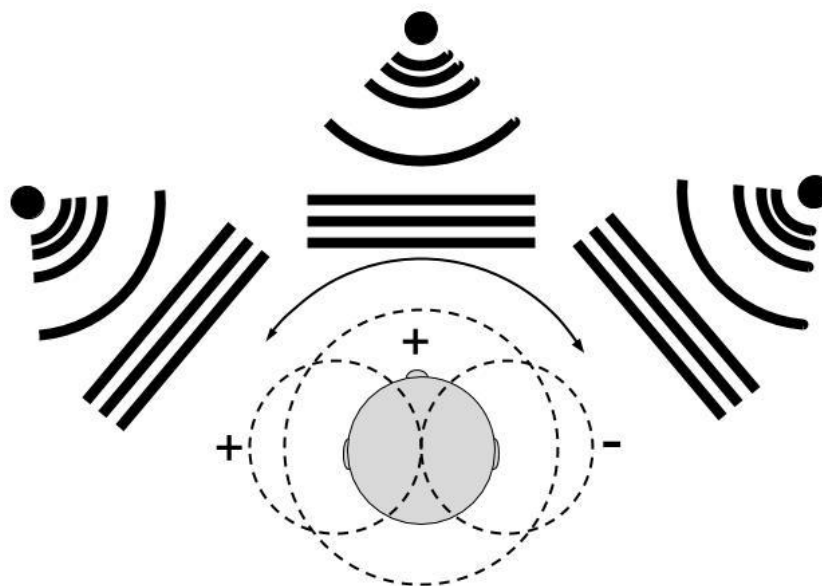


Figure 2.10. Représentation du système de perception humain dans les basses fréquences, comme un système de microphones doubles – un microphone omnidirectionnel et un microphone de vélocité en figure-de-huit sur l'axe frontal – pouvant potentiellement pivoter autour de l'axe z. Pour des raisons de commodité la représentation est limitée à la scène frontale, mais les sources sonores peuvent potentiellement arriver de toutes les directions.

L'enjeu consiste alors à déterminer les informations de phases nécessaires à la bonne localisation des sons. Pour cela, il est nécessaire de définir quels sont les réelles informations de phases utilisées par l'oreille humaine et comment les représenter et les préserver dans un système multicanal.

M. A. Gerzon remarque ensuite qu'il existe un désaccord quant à la manière dont le système auditif déduit les positions des sons grâce aux différences de phases. Un premier groupe de chercheurs, représenté par H. A. M. Clark, G. F. Dutton, P. B. Vanderlyn [Clark & al. 1958] et B. B. Bauer fonde sa théorie de perception sur la supposition que la tête de l'auditeur est immobile. Un deuxième groupe de chercheurs, représenté par Y. Makita [Makita, 1960], D. M. Leakey et P. G. Tager inclut les mouvements de la tête déjà présentés dans les mécanismes de perceptions auditives de l'espace. M. A. Gerzon propose alors ce qu'il dénomme comme une « super-théorie », permettant de regrouper ces deux approches au sein d'un même système, sous la forme de deux cas particuliers. Cette théorie assume que la somme des ondes sonores atteignant les oreilles de l'auditeur correspond à la pression sonore au centre de la tête de celui-ci. Il est à noter que la tête et le corps de l'auditeur ne sont pas considérés sur un plan acoustique et qu'il est assumé que les variations qu'ils génèrent sur le champ sonore sont nulles. Les informations directionnelles restantes aux basses fréquences atteignant l'auditeur sont les différences de phases des ondes atteignant les oreilles. Aussi associe-t-il le mécanisme de perception basse fréquence à un système de microphones doubles – un microphone omnidirectionnel et un microphone de vitesse sur l'axe frontal – correspondant à un microphone en figure-de-huit – [Figure 2.10]. La caractéristique principale est que les informations directionnelles de phases sont nulles lorsqu'elles parviennent de l'axe médian, en adéquation avec les deux théories proposées. Cela n'exclut pas pour autant la possible rotation de la tête, impliquant dans les deux cas la nécessité de reconstruire l'information de phase dans cette direction. Que le mouvement de tête soit pris en compte dans le modèle psychoacoustique ou non.

Il s'agit alors de définir un système, de concevoir un format, permettant de représenter et restituer les informations de phase d'un champ sonore, pour ce système de captation correspondant au mécanisme de perception basse fréquence de l'oreille humaine. L'enjeu étant que le résultat de son enregistrement soit identique à ce qu'il serait pour une source réelle,



quelle que soit la direction de la source par rapport au mécanisme<sup>100</sup>. Pour répondre M. A. Gerzon propose un système de représentation du champ sonore sous la forme de canaux résultant de l'enregistrement de quatre microphones associés à quatre canaux tels que présentés dans [Kohsaka & al., 1972] et [Fellgett, 1974]<sup>101</sup> et défini dans le format A présenté dans le premier chapitre de cette partie [Figure 1.4]. Les informations de localisation basses fréquences sont alors représentées selon une composante omnidirectionnelle  $W$  relative à la pression globale du champ sonore et de deux composantes directionnelles orthogonales  $X$  et  $Y$  relatives aux informations de phases selon les axes éponymes (ou trois avec la composante  $Z$  pour une représentation tridimensionnelle de l'espace) définie par

$$W = L_b + L_f + R_f + R_b$$

$$X = -L_b + L_f + R_f - R_b$$

$$Y = L_b + L_f - R_f - R_b$$

$$Z = -L_b + L_f - R_f + R_b$$

Équation 2.12. Encodage du format A vers le format B. Le facteur de  $\frac{1}{2}$  sur l'ensemble des composantes est omis pour assurer la correspondance avec les informations de localisation basses fréquences d'un système quadriphonique.

Par la suite, M. A. Gerzon définit que réciproquement, au-dessus de 700 Hz, la longueur d'onde du son est suffisamment petite pour que les relations de phase entre les haut-parleurs ne soient plus importantes dans la localisation du son. Aussi, contrairement aux basses fréquences, l'important n'est plus la phase mais le comportement directionnel de l'énergie du champ sonore autour de l'auditeur. Cependant, il note qu' « en raison de la nature positive de l'énergie (au sens mathématique), on ne peut recréer exactement le champ d'énergie d'une source sonore en direct à travers un petit nombre de haut-parleurs que si le son se trouve à la position de l'un

---

<sup>100</sup> En somme, il s'agit d'offrir les informations de phase relative à une source sonore de manière homogène et optimale pour l'ensemble des directions.

<sup>101</sup> Il est aussi possible d'utiliser un format résultant de trois microphones hypercardioïdes pointant respectivement à 60 degrés (à gauche), -60 degrés (à droite) et 180 degrés (à l'arrière) discrétisant parfaitement un cercle tel que défini en [Gerzon, 1970a]. De même le système accepte une extension en trois dimensions, avec le dispositif tel que défini dans le premier chapitre de cette partie.

d'entre eux. Ainsi, aux fréquences moyennes et hautes, tous les mécanismes de localisation de l'oreille ne peuvent être satisfaits dans un système de reproduction pratique. »<sup>102</sup>

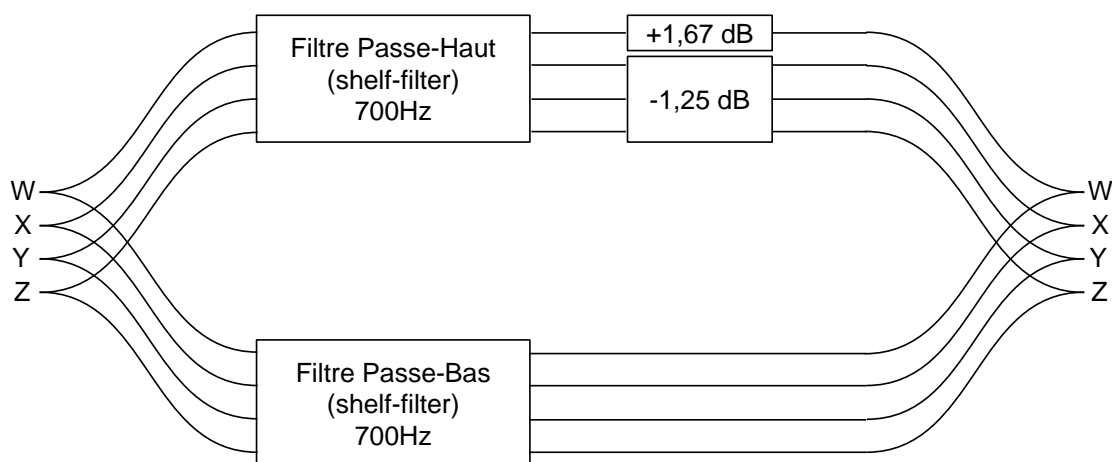


Figure 2.11. Optimisation de l'énergie dans les hautes fréquences (au-dessus de 700 hertz) en amplifiant la composante W de 1,67 dB et diminuant les autres composantes X, Y et Z de 1,25 dB.

Ce que M. A. Gerzon propose alors, est d'analyser les informations directionnelles de l'énergie du champ sonore à partir de la représentation sous la forme d'une composante omnidirectionnelle et vectorielle (les composantes directionnelles de phase), utilisée en basse fréquence. De cette analyse, il proposera de consolider la direction de l'énergie dans les hautes fréquences en augmentant la contribution de la composante omnidirectionnelle de 1,67 décibels et en baissant les contributions des composantes directionnelles de 1,25 décibels [Figure 2.11].

Enfin, lors de la conception du modèle ambisonique, les connaissances sur les mécanismes de localisation des sons dans les hautes fréquences, au-delà de 5000 hertz, restent relativement mal définies, selon M. A. Gerzon. En faisant références, sans les citer, aux indices monauraux et en suivant les travaux notamment de Rayleigh [Rayleigh, 1907], il détermine que les lobes des oreilles jouent un rôle essentiel en hautes fréquences, notamment dans la différenciation arrière-avant. Cependant les recherches sur le sujet et leur intégration dans une approche enveloppante en reste là. Mais il note cependant que « le mécanisme de localisation du lobe n'est pas bien compris, mais semble compter sur le fait que les sons de chaque direction arrivent à l'intérieur

---

<sup>102</sup> [Gerzon, 1974], p. 4, traduit de l'anglais : « [...] because of the positive nature of energy (in the mathematical sense), one can only exactly recreate the energy field of a live sound source through a small number of loudspeakers if the sound happens to be at the position of one of these. Thus at mid and high frequencies, not all of the ear's localization mechanisms can be satisfied in a practical reproduction system. »

de l'oreille de l'auditeur avec une coloration distincte. Ainsi, si nous pouvons reproduire cette coloration dans un enregistrement, nous pouvons renforcer le sens de l'orientation créée ; à la connaissance de l'auteur, cela n'a pas encore été fait dans les enregistrements sonores enveloppants. »<sup>103</sup>

#### 2.2.4. Décomposition en harmoniques sphériques du champ sonore

Les informations nécessaires à une bonne reproduction du champ sonore selon les critères psychoacoustiques établis par M. A. Gerzon, se définissent sur la forme de quatre composantes notées W, X, Y et Z. Les figures de directivités de ces quatre composantes peuvent être associées aux représentations spatiales des quatre premières harmoniques sphériques [Figure 2.5]. Cette relation était en réalité déjà exprimée par M. A. Gerzon en dehors du contexte ambisonique. L'idée est d'associer les quatre canaux quadriphoniques à des microphones dont les « [...] caractéristiques directionnelles sont des combinaisons linéaires des harmoniques de degré 0 et 1 »<sup>104</sup>. Pour comprendre alors comment associer le modèle mathématique au domaine acoustique, il s'agit de concevoir le champ sonore comme une fonction continue et circulaire (ou sphérique) enveloppant l'auditeur et les canaux quadriphoniques,  $L_f, L_b, R_f$  et  $R_b$ <sup>105</sup> – comme les éléments résultant de la discrétisation de cette fonction du champ sonore<sup>106</sup>. Aussi est-il possible d'assimiler la valeur du signal contenu dans chaque canal quadriphonique, à un instant  $t$  comme la valeur d'amplitude de la fonction du champ sonore à cet instant. En procédant à une décomposition en harmonique sphérique à l'ordre 1 de cette fonction du champ sonore, suivant les termes définis par [Équation 2.10], les signaux obtenus associés aux harmoniques  $Y_{0,0}, Y_{1,1}, Y_{1,-1}$  et  $Y_{1,0}$  sont ceux équivalent aux composantes W, X, Y et Z [Figure 2.13]. Cette approche du champ sonore proposée par le modèle ambisonique permet donc d'utiliser les outils mathématiques présentés précédemment, pour définir et formaliser de nombreuses opérations. Enfin, il est évidemment possible d'étendre l'approche aux ordres de décomposition supérieurs, impliquant une meilleure représentation du champ sonore. Comme énoncé précédemment, la décomposition exacte n'est pas valide au-delà de l'ordre de décomposition 3 et les modèles

---

<sup>103</sup> [Gerzon, 1973], p. 4, traduit de l'anglais : « *The pinnae localization mechanism is not well understood, but appears to rely on the fact that sounds from each direction arrive inside the listener's ear with a distinctive coloration. Thus, if we can reproduce that coloration in a recording, we can reinforce the sense of direction created; to the author's knowledge, this has not yet been done in surround-sound recordings.* »

<sup>104</sup> [Gerzon, 1970a], p. 2., traduit de l'anglais : « *This arises from the fact that all conventional microphone directional characteristics are linear combinations of zero and first order spherical harmonics.* »

<sup>105</sup> Pour des raisons de compréhension et de facilitation, il est préférable de rester sur une représentation bidimensionnelle. Mais l'approche est tout aussi valide pour une représentation tridimensionnelle avec une discrétisation du champ sonore sous la forme d'un tétraèdre comme proposé par M. A. Gerzon [Gerzon, 1970b].

<sup>106</sup> Cette représentation est valide notamment parce qu'il est assumé que les haut-parleurs génèrent des ondes planes équivalentes donc à une direction précise sur le cercle ou sur la sphère.

harmoniques ne correspondent plus à du matériel de captation existant. L'idée est alors de travailler directement dans le domaine des harmoniques sphériques, pour tirer parti de cette représentation d'ordre élevé, pour *a posteriori* tenter de restituer au mieux le champ sonore pour un système dit irrégulier.

Enfin, un point qu'il est intéressant de soulever est l'approche proposée par J. Daniel [Daniel, 2001] pour associer le champ sonore au modèle mathématique de la décomposition en harmoniques sphériques. Au-delà des principes reposant sur l'approche de M. A. Gerzon qui vient d'être présentée, le format B peut être vu comme une représentation locale du champ sonore autour de l'auditeur sous la forme d'un gradient de pression. Un gradient de pression est un vecteur représentant les variations de pression en un point via sa direction et sa norme, définissant en quelque sorte la direction dans laquelle se propage les variations de l'air – l'onde sonore – et l'amplitude de ces variations. En considérant le point  $O$  l'origine d'un espace tridimensionnelle, il est possible d'associer le vecteur  $\overrightarrow{OP}$  défini par le point  $O$  et le point  $P$  à un gradient de pression, les coordonnées cartésiennes  $x_p$ ,  $y_p$  et  $z_p$  définissant la direction du gradient de pression et la distance  $OP$  définissant sa norme [Figure A.2].

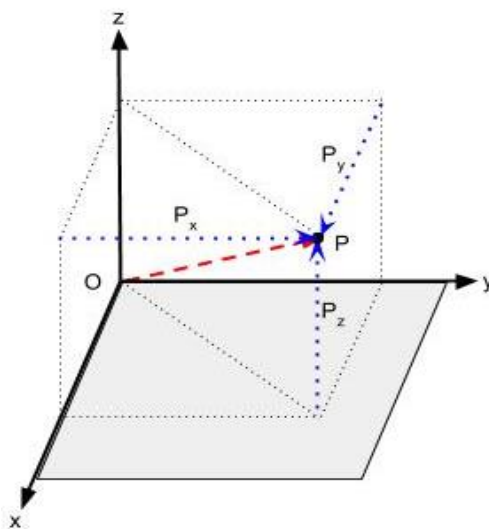


Figure 2.12. Représentation d'un gradient de pression associé à un vecteur  $(OP) \vec{}$  dans un espace tridimensionnel.

La norme du gradient de pression peut alors être associée au coefficient d'amplitude de l'harmonique omnidirectionnelle  $Y_{0,0}$  et la direction du gradient de pression sous sa forme cartésienne -respectivement aux coefficients d'amplitude des harmoniques bidirectionnelles – en figure de huit  $Y_{1,1}$ ,  $Y_{1,-1}$  et  $Y_{1,0}$ . Il est alors possible de considérer cet ensemble

d'harmoniques associés à leurs coefficients d'amplitude respectifs, comme le résultat d'une décomposition au premier ordre du champ sonore en harmoniques sphériques<sup>107</sup>.

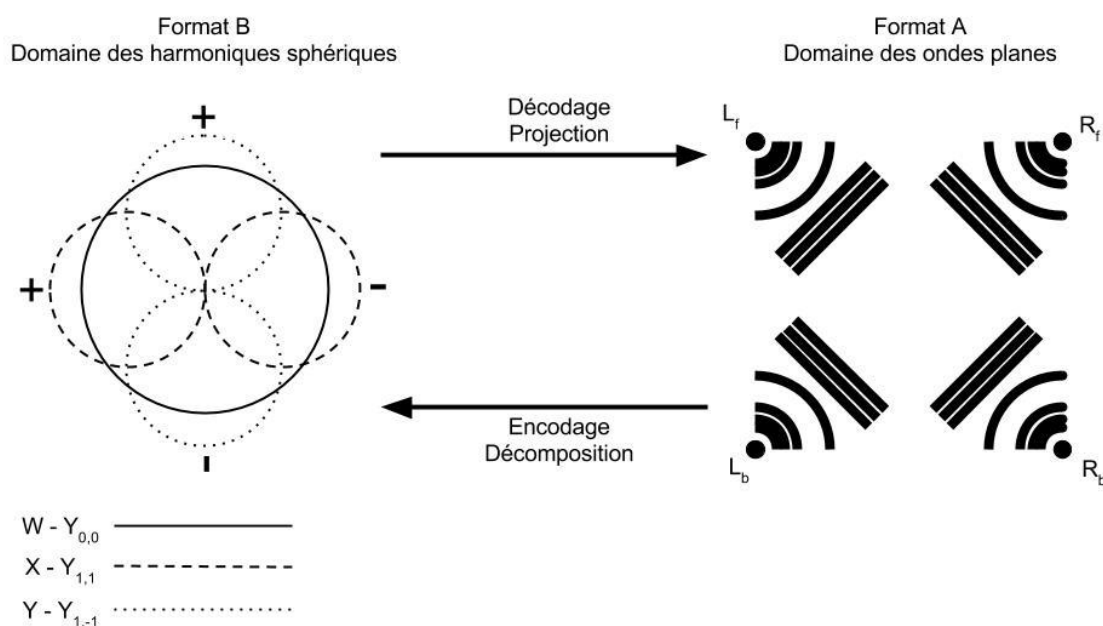


Figure 2.13. Représentation des relations entre les formats ambisoniques, A et B, et les harmoniques sphériques. Ainsi que les opérations associées, respectivement d'encodage et décodage et de décomposition et projection.

En somme, en suivant la logique de création présentée, les signaux associés aux harmoniques sphériques peuvent donc être considérés comme des signaux associés à des prises de son spécifiques. Le signal associé à l'harmonique de degré 0 étant rattaché à une figure omnidirectionnelle. Les signaux associés aux harmoniques de degré 1 étant rattachés à des figures-de-huit colinéaires aux axes x, y et z. Et les signaux des harmoniques de degré supérieurs étant rattachés à des figures directives quelques peu originales mais restant néanmoins tout à fait appréhendables grâce à leurs représentations [Figure 2.2]. Les signaux associés aux harmoniques sphériques reposant sur ces principes, il est bien plus simple pour le musicien, d'envisager les opérations « originelles » des techniques ambisonique. L'enjeu consistera en des opérations de matriçage de signaux, en considérant les figures directives liées à ces derniers. Et en relation avec les représentations spatiales des harmoniques sphériques dont

<sup>107</sup> Cette approche peut paraître plus simple *a priori*, mais ne peut néanmoins pas faire *a posteriori* abstraction des principes psychoacoustiques. Ici l'enjeu est aussi de montrer la construction du modèle ambisonique et non seulement les conséquences déterminant que l'approche est basée sur « une décomposition du champ sonore en harmoniques sphériques ». Cette phrase est souvent énoncée comme point d'appui mais ne révèle pas le modèle ambisonique dans son intégralité. De plus, alors que l'approche ambisonique se démocratise, il me semble très important d'expliquer et même de vulgariser l'approche dont les principes mathématiques sonnent comme des axiomes aux oreilles de nombreux musiciens. Je peux à ce titre sortir des pages de correspondance, où je m'efforce exclusivement à expliquer le modèle ambisonique.

ils sont associés. Ces opérations de matriçage, telles que présentées par exemple pour l’encodage du format A au format B [Équation 2.12], sont issues des formulations mathématiques des harmoniques sphériques. Les enjeux sont donc à présent de définir comment concrètement le modèle ambisonique – avec ses opérations de matriçage, ses équations mathématiques, sa représentation spécifique de l’espace, etc. – s’applique en traitement du signal. Et quels sont les outils permettant de mettre en œuvre à la fois les opérations classiques mais aussi les traitements originaux de l’espace et son.

La décomposition en harmoniques sphériques offre une représentation du champ sonore cohérente avec les modèles acoustiques et psychoacoustiques, définis notamment par M. A. Gerzon. Leur compréhension permet dès lors d’envisager concrètement la mise en œuvre des opérations tels que l’encodage, le décodage ou l’optimisation de champs sonores. Cependant l’atout majeur est de pouvoir se les approprier sur un plan artistique et musical. Et de pouvoir appréhender les répercussions de la mise en œuvre de traitements particuliers comme la granulation, le filtrage en peigne, la décorrélation temporelle, etc. Suite à ces expérimentations, il sera possible de revenir sur le modèle mathématique des harmoniques sphériques et de critiquer son rapport avec les modèles acoustiques et psychoacoustiques. Cette analyse visera à soulever la question de la nécessité même de cette relation. Et à se demander si la représentation abstraite de l’espace sous la forme d’un ensemble d’harmoniques sphériques, n’offre pas en soi un potentiel musical suffisant pour envisager d’autres approches plus libérées des règles et contraintes induites par les modèles acoustiques et psychoacoustiques. En somme cela permettra notamment de discerner les notions de champ sonore liées à l’approche acoustique et d’espace sonore liées à une approche plus ouverte, musicale et artistique.

### 3. Outils Logiciels

Après avoir dégagé les objectifs et les spécificités de l’approche ambisonique en retraçant les développements de ses origines à aujourd’hui et présenté les fondements mathématiques et acoustiques de son modèle, il convient à présent d’introduire les différents outils logiciels dédiés à cet ensemble de techniques de spatialisation : la bibliothèque HOA<sup>108</sup> (High Order Ambisonics). L’enjeu à présent est de fournir les clefs permettant de comprendre sous un aspect plus technique les développements et les approches qui seront présentées dans la deuxième

---

<sup>108</sup> La bibliothèque HOA est disponible en ligne sur le site du projet <http://www.mshparisnord.fr/hoalibrary/> ou sur le site du répertoire GitHub <https://github.com/CICM/HoaLibrary-Light> accédé en septembre 2017.

partie de cette thèse. Pour cela, ce chapitre revient sur les différentes possibilités offertes par cette bibliothèque et notamment ses mises en œuvre pour les logiciels de programmation graphique et modulaire, Pure Data et Max. Cette présentation vise donc à mettre en avant les différentes fonctionnalités offertes par ces outils logiciels. Mais au-delà, il est d'ores et déjà possible de discerner à travers elle comment les mises en œuvre des outils ont été guidées par l'approche particulière du modèle ambisonique, et le besoin de s'appropriier les domaines d'interactions permettant de créer et manipuler des espaces sonores. En somme, le but est aussi de faire valoir les spécificités de ces mises en œuvre, notamment dans leur approche modulaire et dynamique, facilitant la prise en main des signaux dans le domaine des harmoniques sphériques et leur manipulation à des fins musicales. Suite à l'expérience acquise grâce aux différentes expérimentations présentées dans la suite de cette thèse, il sera possible, ici aussi, de revenir et critiquer les choix réalisés, au regard des nouvelles spécifications techniques et fonctionnelles établies pour les systèmes de spatialisation.

### 3.1. Développements et livrables

Les projets HOA proposent de nombreux livrables offrant différentes fonctionnalités et prenant des formes variées suivant le langage de programmation utilisé et les plateformes auxquels ils sont destinés. Ces différentes bibliothèques de code et progiciels sont le résultat de plusieurs années de recherche et développement de 2012 à aujourd'hui, conséquences des projets coordonnés par Anne Sédès et Alain Bonardi, et développés au CICM dans le cadre notamment des programmes du Labex Arts H2H de 2013 à 2015<sup>109</sup>. Afin d'éclairer les choix fonctionnels et techniques et de présenter les outils mis à disposition des développeurs, des musiciens et des compositeurs, il est intéressant de revenir sur le contexte de déroulement de ces projets, sur les partenariats, sur les différentes versions publiées et leurs évolutions.

#### 3.1.1. Origine et première version

Les techniques de spatialisation et de manière plus générale la mise en espace des sons, est un des axes principaux de recherche au CICM [Vaggione, 2001], [Vaggione, 2002], [Sédès & al., 2003a], [Sédès & al., 2003b]. Dans le cadre de l'ACI jeunes chercheurs dirigé par A. Sédès à la Maison des Sciences de l'Homme Paris Nord (MSH Paris Nord), R. Mignot et B. Courribet proposent un premier ensemble d'objets pour les logiciels Max et Pure Data de moteurs de

---

<sup>109</sup> Comme il sera présenté par la suite, les développements continuent encore aujourd'hui.

spatialisation ambisonique et de VBAP [Pulkki, 1997] sous le nom des CICM Tools<sup>110</sup>. Par la suite, de 2007 à 2010, dans le cadre du projet HD3D2, le CICM développe une maquette d'un plugin VST de panoramique, d'amplitude destiné à l'industrie audiovisuelle. Et en 2012, J. Colafrancesco propose un premier ensemble d'objets ambisonique pour le logiciel Max sous le nom de *Hoa.lib*, permettant les opérations classiques d'encodage de source sonore et de décodage de champ sonore [Colafrancesco, 2012]<sup>111</sup>. Ces objets posent les bases de ce qui sera par la suite la bibliothèque HOA, notamment en orientant les outils vers une approche modulaire en laissant accès aux signaux associés aux harmoniques sphériques – contrairement aux approches plus « opaques » telles que proposées par les CICM Tools où l'outil donne directement les signaux destinés aux haut-parleurs –. Et en élaborant une architecture de code en C++ partiellement indépendante du logiciel Max, permettant une réutilisation plus simple sur d'autres plateformes. Il est néanmoins important de noter que des approches similaires étaient déjà proposées par les bibliothèques ambisoniques pour le logiciel Max telles que la bibliothèque de l'ICST (Institute for Computer Music and Sound Technology à Zurich en Suisse)<sup>112</sup> développée par an C. Schacher et P. Kocher depuis 2006 [Schader & al., 2006], [Schader, 2010], celle développée par G. Wakefield [Wakefield, 2006] ou encore celle développée par D. Malham et Matthew Paradis de 2002 à 2010 à l'Université de York<sup>113</sup>. Et pour le logiciel Pure Data comme la bibliothèque de l'IEM (Institute of Electronic Music and Acoustics à Graz en Autriche) développée par Thomas Musil depuis 2006<sup>114</sup>.

Par la suite, les développements ont rapidement pris de l'ampleur autour des différents projets du Labex Arts H2H sous la responsabilité scientifique d'A. Sèdes et A. Bonardi. C'est au cours du premier projet « La spatialisation du son par les musiciens pour les musiciens » à l'été 2012 que J. Colafrancesco, E. Paris et moi-même avons développé la première ébauche de la

---

<sup>110</sup> La bibliothèque est encore accessible en septembre 2017 sur le site du CICM à l'adresse <http://cicm.mshparisnord.org/>. Une version mise à jour et augmentée pour Max 6 a été réalisée par E. Paris en 2013.

<sup>111</sup> Bien que ces outils n'aient jamais été publiés, leur code source est disponible en ligne en septembre 2017 sur le répertoire du projet <https://github.com/jcolafrancesco/HoaLib>.

<sup>4</sup> La bibliothèque est disponible en ligne et accessible en septembre 2017 à l'adresse <https://www.zhdk.ch/downloads-ambisonics-externals-for-maxmsp-5381>.

<sup>113</sup> La bibliothèque ne semble plus disponible en ligne, elle l'était cependant jusqu'à encore un an. L'université en fait référence en ligne à cette adresse consultée en septembre 2017 <https://www.york.ac.uk/music/staff/honorary/dave-malham/>.

<sup>114</sup> La bibliothèque est disponible en ligne à l'adresse [https://puredata.info/downloads/iem\\_ambi](https://puredata.info/downloads/iem_ambi) consultée en septembre 2017.



bibliothèque de code C++ destinée à l'ambisonie d'ordre supérieur HOA<sup>115</sup>, aboutissant en octobre à la première version disponible pour le logiciel Max v1.0 [Colafrancesco & al., 2013].

### 3.1.2. Extensions et déploiement multiplateforme

Les résultats convaincants de ces développements ont amené à un deuxième projet dès 2013 intitulé « Interfaces pour la mise en espace du son ». Ce projet pour lesquels E. Paris et moi-même avons continué les développements en tant qu'ingénieurs de recherche a été l'occasion d'étendre notamment les interfaces graphiques, mais aussi de proposer de nouvelles mises en œuvre de traitements où l'aspect modulaire et dynamique a commencé à être vraiment élaboré. Ces développements coïncident avec les versions 1.1 et 1.2 publiées en mars et avril 2013. Ce projet a aussi été l'occasion de proposer une version alternative en langage FAUST<sup>116</sup> [Trausmuth & al., 2006], [Orlarey & al., 2006] avec notamment la contribution de Yann Orlarey, Dominique Fober et Stéphane Letz du GRAME (Centre national de création musicale à Lyon). L'occasion aussi de proposer un moteur de spatialisation de sources ponctuelles sous la forme d'un plugin VST<sup>117</sup> avec notamment la contribution João Svidzinski ou encore la première mise en œuvre de la première version pour Pure Data avec la contribution de Manuel Deneu. Une partie de ces travaux de portage ont été présentés dans [Guillot & al., 2013]. Alors que la première mise en œuvre pour le logiciel Max était relativement simple, grâce à une API très adaptée aux besoins – tant pour le développement du traitement multicanal que pour la création d'interfaces graphiques – de nombreuses complications sont apparues lors des mises en œuvre pour le logiciel Pure Data. Ces difficultés ont obligé à repenser la structure et le fonctionnement global des objets graphiques et des objets de traitement du signal dans Pure Data. Ainsi, ces spécifications ont amené à la création d'une API en C et Tcl/Tk : le CICM Wrapper<sup>118</sup> [Guillot, 2014]. Un ensemble de codes qui fournit de nouvelles fonctionnalités et facilite la mise en œuvre d'objets a permis de développer parallèlement les versions pour le logiciel Max et Pure Data<sup>119</sup>. Au cours des développements, la bibliothèque HOA opère de nombreux changements, des fonctionnalités sont enlevées et remplacées par de nouvelles. La

---

<sup>115</sup> Le code de cette version est disponible en ligne sur le premier répertoire du projet à l'adresse consultée en septembre 2017 <https://github.com/CICM/HoaLibrary>.

<sup>116</sup> La bibliothèque HOA en FAUST est incluse dans la distribution officielle du langage.

<sup>117</sup> Le plugin n'est aujourd'hui plus fonctionnel mais le livrable est néanmoins téléchargeable sur le site du projet HOA et le code source est disponible sur le répertoire Github. Une mise à jour du code ou une alternative en embarquant Pure Data dans un plugin sont envisagées pour de futurs développements.

<sup>118</sup> La bibliothèque CICM Wrapper est disponible sur le répertoire Github du CICM : <https://github.com/CICM/CicmWrapper>.

<sup>119</sup> Ce projet de développement m'a aussi amené à concevoir une bibliothèque d'objet graphique nommée Cream destinée à étendre les fonctionnalités de Pure Data.

bibliothèque subit notamment plusieurs refontes intégrales, il serait donc vain et laborieux de vouloir présenter toutes les modifications<sup>120</sup>. Néanmoins il est à noter qu'en novembre 2013, est publiée la version 1.3 [Sèdes & al, 2014] de la bibliothèque, à la fois pour Max et Pure Data. Cette version offre de nombreux nouveaux traitements originaux de l'espace et du son. Un des objectifs importants a été de réduire les dépendances à des codes sources fermés ou payants, afin de créer une version plus accessible et en adéquation avec des recherches publiques.

Enfin en 2014 commence le projet « HOA 3D » sur lequel E. Paris et moi-même continuons les développements. Comme son nom l'indique, le projet est principalement centré autour de l'extension de la bibliothèque, à une représentation tridimensionnelle de l'espace sonore. Néanmoins les travaux continuent sur de nombreux aspects : création de traitements, améliorations des interfaces, optimisation du code, correction des bugs, etc. et amènent en juillet 2014 à la version 2 de la bibliothèque. Les travaux de développement se poursuivent en 2015, amenant à une refonte totale de la bibliothèque. Pour ces travaux, nous recevons entre autres l'aide de Thomas Le Meur amenant à une version dite « light »<sup>121</sup>. Sans pour autant enlever de fonctionnalités, cette nouvelle version propose un code source sans aucune dépendance externe et uniquement sous la forme de fichiers d'en-tête<sup>122</sup>, facilitant l'utilisation et le portage de la bibliothèque, ainsi qu'une complète documentation en ligne<sup>123</sup>. A ce propos, il est possible de noter l'intégration de la bibliothèque en tant que module open-framework<sup>124</sup> et la mise en œuvre pour le logiciel SuperCollider proposés par Claudio Cabral. En juin 2015 pour le logiciel Pure Data puis en juillet pour le logiciel Max, est publiée la dernière version 2.2 destinée aux utilisateurs de la bibliothèque HOA. Cette nouvelle version est toujours l'occasion de proposer de nouveaux traitements et possède la particularité d'intégrer une série de tutoriels pour le logiciel Pure Data<sup>125</sup>, que j'ai réalisés afin de faciliter aux utilisateurs la prise en main du modèle ambisonique.

---

<sup>120</sup> Il est néanmoins possible de se référer aux fichiers de « *changelog* » des distributions pour les logiciels Max et Pure Data ainsi qu'à l'historique des répertoire Github.

<sup>121</sup> Cette refonte coïncide avec un changement de répertoire pour la distribution du code source toujours sur Github à l'adresse <https://github.com/CICM/HoaLibrary-Light>.

<sup>122</sup> Cette approche facilite notamment l'étape de compilation en évitant de définir les fichiers sources à inclure au compilateur.

<sup>123</sup> La documentation est disponible à l'adresse <http://cicm.github.io/HoaLibrary-Light/>.

<sup>124</sup> Le module est disponible sur le projet Github <https://github.com/CICM/ofxHoa/>.

<sup>125</sup> Eliott Paris en a réalisé le portage pour le logiciel Max.

### 3.1.3. Bilan et perspectives

La bibliothèque HOA est avant tout un ensemble de classes C++, libre et ouvert<sup>126</sup>, optimisées pour le traitement du signal multicanal en temps-réel et dont l'architecture facilite la portabilité et le déploiement sur de nombreux systèmes d'exploitation – tels que Linux, MacOS, Windows – et pour de nombreux environnements logiciels. Les deux principales mises en œuvre de la bibliothèque ont été réalisées pour les logiciels musicaux de programmation graphique et modulaire Max et Pure Data. Et dont les dernières versions (v2.2) ont été téléchargées plus 6000 fois, créant ainsi un des outils de référence dans le domaine de la spatialisation ambisonique.

Cette bibliothèque a par ailleurs été présentée dans de nombreux contextes. En juin 2013, la présentation aux journées d'informatique musicale à St-Denis, nous a valu le Prix du jeune chercheur, décerné par l'association d'informatique musicale, à Eliott Paris, Julien Colafrancesco et moi-même [Figure 3.1]. Par la suite, le projet « Interfaces pour la mise en espace du son » a été présenté en octobre 2013 à l'université Paris 13 dans le cadre du festival Savante Banlieue et en juin de la même année au 104 à Paris dans le cadre du Festival du numérique Futur en Seine. La version Pure Data a quant à elle été présentée à distance en novembre 2013 pour le PdWeekend à Los Angeles. La bibliothèque dans son ensemble a été le centre d'expérimentations et de présentations à la Muse en Circuit<sup>127</sup> en 2014, pour l'ICMC | SMC<sup>128</sup>, en septembre 2014 à Athènes [Sèdes & al., 2014], au ZKM<sup>129</sup> à Karlsruhe en octobre 2014, à l'Université de Montréal en décembre 2014, aux Journées d'Informatique Musicale en mai 2015 toujours à Montréal [Sèdes, 2015], [Bonardi & al., 2015], ou encore à la Pure Data Convention qui s'est déroulée au Stevens Institute of Technology et à la New York University en 2016.

---

<sup>126</sup> La bibliothèque est notamment sous licence GPL-3.0.

<sup>127</sup> Centre National de Création Musicale d'Alfortville.

<sup>128</sup> Conférence combinée de l'*International Computer Music Conference* (ICMC) et de la *Science and Music Conference* (SMC).

<sup>129</sup> Centre d'art et de technologie des médias de Karlsruhe.



Figure 3.1. Dispositif de présentation de la bibliothèque HOA aux JIM 2013 à l'Université Paris 8 à St-Denis.

Enfin sur un plan plus musical, la bibliothèque a été utilisée dans de nombreuses œuvres et installations et il est évidemment impossible de toutes les citer. Il y a eu en effet au cours de ces années de recherche plus de 19 concerts<sup>130</sup> avec des œuvres utilisant la bibliothèque HOA rien que dans le cadre de ceux organisés par le CICM<sup>131</sup>. A titre indicatif, il est possible de citer les installations « Transduction » réalisée par Anne Sèdes, Eliott Paris et moi-même dans le cadre des vitrines du Labex arts H2H en octobre 2014 à la Gaîté lyrique, puis au Salon de la valorisation de la recherche en sciences sociales Dionys'Hum en avril 2015. Ou encore les pièces de *Pianotronics 1 à 4* pour piano et électronique temps-réel composées respectivement en 2012 et 2016 par Alain Bonardi, *Immersion* pour violoncelle et électronique temps-réel composée en 2012 par Anne Sèdes, *Ouverture*, notamment pour percussion et électronique temps-réel, mais qui se décline sous plusieurs versions, composée en 2015 toujours par Anne Sèdes. Mais aussi les pièces *Cantos de Viento* pour flûte à bec et électronique temps-réel et *Broken Dance*, pièce acousmatique, de Sabina Covarrubias ou encore les nombreuses pièces de João Svidzinski : *Turdus* pour flûte et électronique (2013), *Sigel* pour flûte et électronique (2014), *M1910* pièce acousmatique (2014), *Les âmes remerciées* pour percussion et électronique (2014) et *Aphorismes* pièce acousmatique (2015).

---

<sup>130</sup> Nous avons tenté d'en regrouper un aperçu avec João Svidzinski sur le site du projet HOA dans la rubrique Concerts des événements.

<sup>131</sup> Et bien entendu il y a sûrement d'autant plus en dehors du cadre du CICM et de l'Université Paris 8.

Fort de ses nombreuses années de recherches la bibliothèque HOA semble être un projet établi. Néanmoins depuis 2015, je poursuis les développements. Les objectifs principaux sont à présent d'une part d'assurer une maintenance et de corriger les erreurs du code C++ et de ses mises en œuvre pour les différentes plateformes<sup>132</sup> et d'autre part l'amélioration de l'approche didactique et pédagogique. À ce sujet, les séminaires de Master « Créations d'Espaces Sonores » que j'ai pu dispenser à l'université Paris 8 de 2015 à 2017 ont été une source d'inspiration notoire pour de nouveaux tutoriels qui, je l'espère, pourront être publiés avec la prochaine version de la bibliothèque.

Suite à cette présentation de l'historique et du contexte dans lequel la bibliothèque HOA a été développé, il est nécessaire de présenter concrètement son contenu.

## 3.2. Présentation des outils

La bibliothèque HOA présentée dans le cadre de cette thèse est la version 2.2 disponible à l'utilisateur et finalisée depuis juillet 2015 pour les logiciels Max et Pure Data. Les deux mises en œuvre comportent quelques différences notables qui sont les résultats des différences intrinsèques des logiciels. C'est le cas par exemple de la gestion des préréglages dans le logiciel Max, qui ne se retrouve pas dans le logiciel Pure Data. Il y aussi des différences liées aux questions de rétrocompatibilité, notamment au sein du logiciel Max pour lequel la mise en œuvre est plus ancienne. Cela nécessite de préserver certaines fonctionnalités même si elles paraissent obsolètes. C'est le cas par exemple de la configuration des haut-parleurs dans le décodeur de Max, qui est définissable à l'aide d'attributs dynamiques et non d'arguments. La présentation suivante se focalise majoritairement sur la version Pure Data<sup>133</sup> néanmoins, autant que possible et dans la limite des informations pertinentes, les différences seront notifiées. Aussi, mentions spéciales, les objets sont disponibles pour une représentation et une restitution de l'espace sonore en deux dimensions et en trois dimensions. Enfin, de nombreux développements ont été réalisés depuis juillet 2015, impliquant beaucoup de changements et d'améliorations. Des versions dites beta sont en effet d'ores et déjà fonctionnelles pour le

---

<sup>132</sup> Les mises à jour majeures des logiciels Pure Data ou Max nécessitent d'adapter les distributions de la bibliothèque HOA.

<sup>133</sup> Le logiciel Pure Data me servant de terrain d'expérimentation depuis la mise en œuvre de la bibliothèque pour cette plateforme. C'est dès lors au sein de cette plateforme que la grande majorité des objets et des abstractions ont été réalisés et testés avant leur adaptation au logiciel Max.

logiciel Pure Data, bien que non-publiée<sup>134</sup>. Afin de préserver une cohérence avec les outils disponibles en ligne, ces évolutions seront donc néanmoins majoritairement ignorées. Lorsque cela semble particulièrement nécessaire, ces changements seront présentés sous forme prévisionnelle ou de simple perspective de développement.

Un petit ensemble d'outils destinés à faciliter l'usage de la bibliothèque HOA n'est pas présenté en détail dans cette thèse, car leurs fonctionnalités s'éloignent quelque peu du sujet principal de cette thèse et que E. Paris l'a déjà fait [Paris, 2013]. Il s'agit de l'objet *hoa.connect* permettant de raccorder plus facilement les objets de la bibliothèque entre eux<sup>135</sup>, l'objet *hoa.dac~* qui permet de créer rapidement une sortie audio multicanal<sup>136</sup> et les objets *hoa.pi* et *hoa.pi~* qui génèrent le nombre *pi*. Enfin, avant de rentrer dans une présentation plus détaillée des outils, il est nécessaire de présenter certains aspects génériques.

### Nomenclature des objets

La bibliothèque suit une syntaxe particulière pour le nom des objets et des abstractions, offrant des avantages pour leur organisation, qui est construite autour de morphèmes. Tous les objets et toutes les abstractions commencent par le terme *hoa* faisant référence à la bibliothèque. Pour les objets dont le comportement varie selon le nombre de dimensions spatiales, le second terme permet de définir à quel espace il est destiné, à savoir *2d* ou *3d*. Dans le cadre de cette présentation, afin de marquer le fait que le nombre de dimensions spatiales est indifférent, le morphème *2/3d* est utilisé. Et suite aux dimensions, se trouve le nom de l'objet définissant sa fonction [Figure 3.2]. Sauf cas particuliers, les objets sont destinés à être utilisés dans le domaine des harmoniques sphériques ou circulaires, ou du moins en relation avec ce domaine. Ainsi, sauf exception le premier argument de chaque objet est donc l'ordre de décomposition.

---

<sup>134</sup> Avant de la publier officiellement, je souhaite ajouter quelques traitements, dont certains sont présentés dans cette thèse. Afin d'améliorer la documentation et inclure une série de nouveaux tutoriels tirés à la fois du séminaire de « Créations d'Espaces Sonores » dispensé à l'Université Paris 8 et de présentations de l'approche ambisonique telle que celle réalisée à la Pd Convention 2016. De plus, certains outils sont aujourd'hui obsolètes et ne correspondent pas avec les enjeux et la direction générale des approches, ils sont donc progressivement enlevés de la bibliothèque. Enfin je souhaite améliorer la compatibilité avec les dernières versions de Pure Data et de Max.

<sup>135</sup> Il s'agit néanmoins d'un des premiers objets mis en œuvre et qui améliore énormément le confort lors des expérimentations.

<sup>136</sup> De plus cet objet tend à disparaître en raison de problèmes de pérennité du code.

```
hoa.3d.encoder~ 3
```

Figure 3.2. Représentation de la syntaxe d'un objet type dans la bibliothèque HOA. Le terme *hoa* fait référence à la bibliothèque elle-même, le terme *3d* définit l'utilisation dans un contexte tridimensionnel et le terme *encoder* définit la fonction de l'objet. Le nombre 3 est le premier argument et définit l'ordre de décomposition.

### Numérotation des canaux

Afin d'assurer le bon fonctionnement des différents objets, il a été défini que l'agencement des canaux suivent la norme définie par l'ACN<sup>137</sup> telle qu'énoncée dans le précédent chapitre. Et qui peut être exprimée suivant le degré et l'ordre azimutal des harmoniques

En 2D par

$$i = 2l + \frac{m}{l}$$

En 3D par

$$i = l \times (l + 1) + m$$

Équation 3.1. ACN en fonction du degré et de l'ordre azimutal d'un harmonique.

Réciproquement, le degré et l'ordre azimutal des harmoniques peuvent être exprimés suivant l'ACN

En 2D par

$$l = \frac{i + i\%2}{2} \text{ et } m = l \times (1 - 2(i\%2))$$

En 3D par

$$l = \sqrt{i} \text{ et } m = i - (l(l + 1))$$

Équation 3.2. Degré et ordre azimutal d'un harmonique en fonction de l'ACN.

### Information générale

Avant de présenter les fonctionnalités des objets, il est important de définir que dans les logiciels Max et Pure Data, les usages veulent que la dénomination de « signal » se réfère de manière générale aux signaux audio numériques, dont les valeurs sont traitées de manière

---

<sup>137</sup> *Ambisonic Channel Number* (Numéro du canal ambisoniques) définie par *The Ambisonics Association* disponible consultée en septembre 2017 à l'adresse <http://ambisonics.ch/standards/channels>.

continue à une fréquence haute<sup>138</sup>. Alors que la dénomination de « contrôle », ou de « message », se réfère de manière générale aux signaux d'informations traités de façon intermittente à des fréquences relativement faibles<sup>139</sup>.

### 3.2.1. Les objets de traitement du signal

Les mises en œuvre de la bibliothèque HOA pour les logiciels Max et Pure Data offrent un ensemble d'objets pour le traitement du signal en ambisonie. Ces objets, apparaissant sous forme de boîtes, sont compilés à partir du code C++, donc *a priori* relativement optimisés. Mais les mécanismes inhérents aux opérations qu'ils proposent restent opaques à l'utilisateur. Celui-ci ne peut en effet concevoir et appréhender l'utilisation des objets, que d'après les fichiers d'aides et des tutoriels. A ce titre un soin tout particulier y a été apporté lors de leurs conceptions<sup>140</sup>. Néanmoins et ce pour les mêmes raisons, ces objets tendent à disparaître des pratiques présentées dans le cadre de cette thèse pour des raisons à la fois didactiques, notamment liées à l'intelligibilité des mécanismes mis en jeu. Mais surtout pour le manque de flexibilité et de souplesse dans la prise leur main<sup>141</sup>.

Les deux objets les plus usuels de la bibliothèque – se trouvant au cœur même des principes ambisoniques et offerts dans la plupart des bibliothèques – sont destinés à l'encodage directionnel de sources sonores et au décodage de champ sonore pour un système de restitution, et sont respectivement dénommés *hoa.2/3d.encoder~* et *hoa.2/3d.decoder~*. Entre l'encodage et le décodage, la bibliothèque HOA permet deux opérations génériques et prédéfinies : la rotation et la variation de la résolution angulaire grâce aux objets *hoa.2d.rotate~* et *hoa.2/3d.wider~*. Sur ce dernier objet, une approche originale est construite, permettant la synthèse de source sonore ponctuelle disponible via l'objet *hoa.2/3d.map~*. Associé au décodage, l'opération d'optimisation de champ sonore telle que présentée par J. Daniel est disponible via l'objet *hoa.2/3d.optim~*. Enfin, suite à ces opérations, il est possible d'analyser la restitution du champ sonore suivant les outils psychoacoustiques définis par M. A. Gerzon avec l'objet *hoa.2/3d.vector~*. Trois objets ne sont pas présentés plus en détail : les objets

---

<sup>138</sup> Fréquences allant généralement de 44100 Hz à 192000 Hz en fonction du logiciel, des paramètres utilisateurs et du matériel.

<sup>139</sup> Fréquences autour généralement de 50 Hz en fonction aussi du logiciel, des paramètres utilisateurs et du matériel.

<sup>140</sup> Une brève présentation et critique du contenu des tutoriels dans les logiciels Max et Pure Data est réalisée par la suite.

<sup>141</sup> La mise en œuvre de traitement dans le domaine des harmoniques sphériques montrera l'aspect restrictif des objets compilés de type boîte.



*hoa.2d.projector~*, *hoa.2d.recomposer~* et *hoa.2d.exchanger~*. L'objet *hoa.2d.exchanger~* permet de réordonner les canaux suivant différents formats *Furse-Malham*, *SID* ou *ACN* et/ou de normaliser les canaux selon les normalisations *MaxN*, *SN3D* ou *N3D*. La question de la normalisation est intéressante et est discutée dans le cadre de cette thèse. Néanmoins, l'objet en lui-même n'offre que peu d'intérêt à l'étude. Les objets *hoa.2d.projector~* et *hoa.2d.recomposer~* ne sont ni plus ni moins qu'un décodeur et un encodeur de champ sonore. En somme, l'ensemble des opérations réalisables avec ces objets fonctionnent avec les objets *hoa.2/3d.encoder~*, *hoa.2/3d.map~* et *hoa.2/3d.decoder~*<sup>142</sup>, qui en prime, ne sont pas limités à une représentation bidimensionnelle de l'espace. Selon moi, le choix de mettre en œuvre ces outils était à l'origine une erreur, car elle ne fait qu'apporter de la confusion aux mécanismes mis en jeu en ambisonie, en donnant des différents pour des opérations similaires.

### L'encodage de sources directionnelles

L'objet *hoa.2/3d.encoder~* permet de synthétiser les signaux associés aux harmoniques sphériques ou circulaires, selon une direction donnée par un azimut et potentiellement une élévation dans le cas d'une restitution tridimensionnelle [Figure 3.3].

La mise en œuvre permet de définir ces angles en contrôle ou signal, permettant alors d'atteindre des vitesses de déplacement suffisamment élevées pour générer des effets de distorsion tels que la modulation en anneau le permet. Il est à noter qu'aucune interpolation n'est réalisée lorsque les angles sont définis en contrôle, impliquant potentiellement des craquements lors de leur variation. C'est pourquoi il est préconisé que cette utilisation ne soit restreinte qu'à des sources statiques. Le choix a été fait de définir les angles suivant la formulation mathématique en radians<sup>143</sup>. Cependant les usages musicaux favorisent souvent une formulation en degrés, telle que pratiquée pour le positionnement de haut-parleurs dans la bibliothèque. De plus, les usages initiaux des logiciels favorisent l'utilisation de signaux entre 0 et 1<sup>144</sup>. Une amélioration future pourrait donc tenir compte de ces usages, tout en préservant aisément la compatibilité avec les versions précédentes via l'utilisation d'arguments pour définir le système.

---

<sup>142</sup> A l'origine l'objet *hoa.2/3d.recomposer~* facilitait la mise en place des opérations de déformation de perspective telles que présentées par M. A. Gerzon [Gerzon & al., 1992] puis formalisées aux ordres élevés par J. Daniel [Daniel, 2001], p. 164.

<sup>143</sup> Les valeurs ne sont néanmoins pas bornées (entre 0 et  $2\pi$  radians par exemple) afin faciliter les mises en œuvre.

<sup>144</sup> Les logiciels Max et Pure Data proposent des opérations trigonométriques de traitement du signal où les angles définis entre 0 et 1 sont équivalents à des angles entre 0 et  $2\pi$  radians.

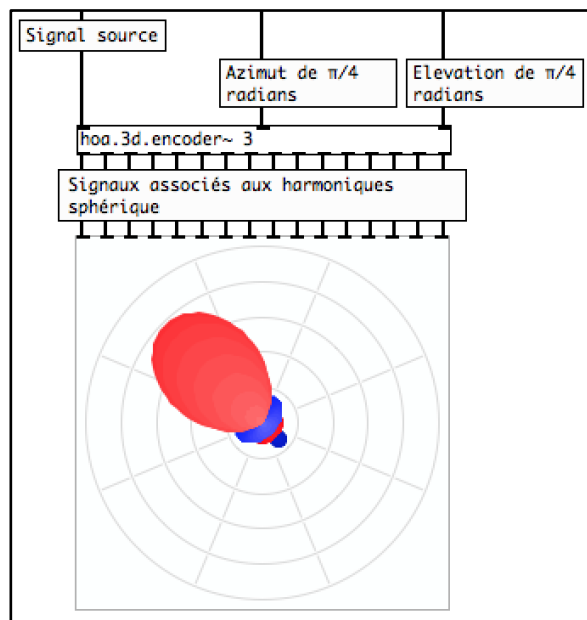


Figure 3.3. Opération d'encodage dans le logiciel Pure Data avec l'objet *hoa.3d.encoder~ 3* synthétisant les signaux associés aux harmoniques sphériques pour un ordre de décomposition 3 d'une source, reçu dans l'entrée signal de gauche, et selon la direction définie par l'azimut, reçu dans l'entrée signal du milieu, et l'élévation, reçu dans l'entrée signal de droite.

### La variation de la résolution angulaire

L'objet *hoa.2/3d.wider~* permet de faire varier la résolution angulaire du champ. La résolution spatiale permet de définir la précision avec laquelle les images sonores sont restituées [Daniel, 2001]. La résolution angulaire étant l'équivalent limité au domaine angulaire, celle-ci ne prend pas en considération la distance. Si la résolution angulaire d'un champ sonore est faible, l'auditeur ne pourra définir avec précision la provenance des sources sonores ponctuelles qui composent le champ sonore. À l'opposé, si la résolution spatiale est importante, l'auditeur pourra définir leurs directions avec plus de précision – dans la limite de notre perception spatiale auditive [Blauert, 1997]. Cette propriété du champ sonore est bien sûr difficilement quantifiable sur le plan perceptif, elle dépend de l'auditeur, de son placement dans l'espace de restitution et de nombreux autres facteurs tels que la salle, le système de restitution, etc. Cependant, la résolution angulaire n'en reste pas moins un élément des plus notables sur le plan auditif. La variation de cette propriété peut, en outre, devenir un paramètre musical. L'un des atouts principaux de l'ambisonie est donc de pouvoir varier cette propriété, et cela grâce à l'ordre de décomposition. En effet, plus l'ordre de décomposition augmente, plus la résolution angulaire du champ sonore augmente. En partant de cette propriété, il a été envisagé de pouvoir faire varier dynamiquement l'ordre de décomposition, en diminuant progressivement les contributions des harmoniques en partant des degrés élevés pour ne laisser que l'harmonique omnidirectionnelle de degré 0 à la fin.

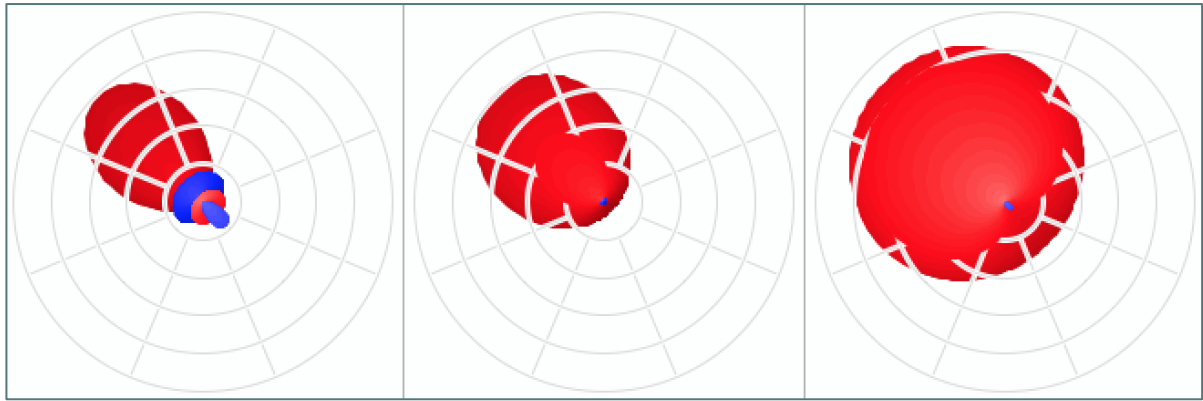


Figure 3.4. Représentation des figures directionnelles d'une source sonore associée à une onde plane pour trois valeurs de résolution angulaire : 1 à gauche, 0.75 au centre et 0.5 à droite.

Pour offrir un paramètre jouable, la valeur de résolution angulaire est définie entre 0 et 1 quel que soit l'ordre de décomposition tel qu'à 0 les sources sont omnidirectionnelles et à 1 leur directionnalité est optimale [Figure 3.4]. L'enjeu de l'approche étant de préserver autant que possible un effet linéaire ainsi qu'une énergie constante dans le champ sonore. De nombreuses propositions et mises en œuvre ont été réalisées afin d'aboutir à cet objectif. Ce travail est présenté et critiqué plus en profondeur dans la suite de cette thèse.

### L'encodage de sources ponctuelles

Une version alternative à l'encodage de sources directionnelles prenant en compte la notion de distance, est disponible via l'objet *hoa.2/3d.map~*. Dans ce cas, l'usage veut que la source soit définie comme ponctuelle, venant d'un point de l'espace, plutôt que directionnelle et venant d'une direction donnée. Afin d'offrir un système adaptable et indépendant des dispositifs de restitution, la distance est considérée de manière relative et abstraite où 0 correspond au « centre » de la sphère ambisonique. Ou plus exactement au centre du dispositif de restitution quel qu'il soit, 1 correspond à la surface de la sphère ambisonique, ou à la distance définie par le système de restitution, et au-delà, la source s'éloigne vers l'infini. L'approche est ici différente de l'encodage NFC proposé par J. Daniel [Daniel, 2003]. En somme, elle se résume seulement à diminuer le gain de la source sonore lorsqu'elle s'éloigne et à diminuer la résolution angulaire lorsque la source sonore se rapproche jusqu'à atteindre une source sonore omnidirectionnelle lorsqu'elle est au centre. Cette approche n'est évidemment pas valide sur un plan acoustique. Mais ce choix de mise en œuvre peut néanmoins être justifié au regard de plusieurs aspects tels que le contexte musical, l'optimisation des calculs et des approximations des modèles acoustiques déjà établies en ambisonie. Malheureusement en raison de la problématique quelque peu en dehors du sujet de cette thèse, cette question ne peut pas être

traitée dans ce contexte<sup>145</sup>. Il est d'ailleurs intéressant de remarquer qu'une approche sensiblement similaire était déjà pratiquée dans la mise en œuvre de l'objet *ambipan~* des CICM Tools, et que la mise en œuvre présentée par M. A. Gerzon de l'encodage directionnel d'une source sonore avec un manche directionnel, utilise aussi cette approche [Gerzon, 1975c]. Enfin, l'un des facteurs les plus importants de l'impression de distance est la réverbération [Chowning, 1977]. Cependant dans une approche musicale, ce type de procédé peut difficilement être imposé par l'outil, car le choix de la réverbération fait partie de l'acte compositionnel. Une approche juste amène donc à laisser le compositeur ou le musicien gérer lui-même cet aspect de la simulation de la distance. Sur le plan des fonctionnalités, il faut remarquer que cet objet est plus dynamique que l'objet *hoa.2/3d.encoder~*. Il offre la possibilité d'encoder soit une seule source sonore, auquel cas trois entrées permettent de contrôler la position de la source en signal, ou deux pour une représentation bidimensionnelle de l'espace. Ou plusieurs sources sonores, auquel cas il y a autant d'entrées que de sources et les positions sont définies par des messages. De plus, cet objet permet de définir la position en coordonnées cartésiennes<sup>146</sup>, aussi bien qu'en coordonnées sphériques ou polaires. Du point de vue des perspectives de développement, étant donné que l'ensemble des opérations réalisables avec l'objet *hoa.2/3d.encoder~* est réalisable avec l'objet *hoa.2/3d.map~*, il peut être envisagé de ne conserver que ce dernier dans de futures versions de la bibliothèque<sup>147</sup>.

### La rotation de champ sonore

L'objet *hoa.2d.rotate~* permet de d'appliquer une rotation du champ sonore uniquement dans le domaine des harmoniques circulaires. Si la rotation est relativement simple à mettre en œuvre autour de l'axe z, effet *yaw* ou *rotate*, elle ne l'est pas tout autant autour des axes x, effet *roll* ou *tilt*, et y, effet *pitch* ou *tumble*. Dans la pratique, la rotation d'un champ sonore dans le domaine des harmoniques sphériques consiste à synthétiser de nouveau chaque signal associé à une harmonique sphérique, à partir de versions pondérées de certains des autres signaux selon l'angle donné. L'opération peut être représentée sous forme matricielle. Dans la pratique, par exemple, le signal résultant d'une rotation autour de l'axe z de l'harmonique de degré 1 et d'ordre azimutal -1 s'obtient par l'addition du signal de cette harmonique pondérée par le cosinus de l'angle de rotation avec le signal de l'harmonique de degré 1 et d'ordre azimutal 1

---

<sup>145</sup> Néanmoins, une première justification a déjà été proposée [Guillot, 2013] et il est envisagé de compléter cette démarche au regard des récentes connaissances acquises.

<sup>146</sup> Approche uniquement valide si la notion de distance est intégrée.

<sup>147</sup> Il est possible de concevoir un objet où les deux noms seraient valides afin de préserver le bon fonctionnement des patches antérieurs. Néanmoins l'utilisation d'une entrée dédiée à la distance avant les angles directionnels implique, a priori, des problèmes de compatibilité sur lesquels il est nécessaire de réfléchir davantage.

pondérée par le sinus de l'angle de rotation. Cependant, alors que l'opération reste relativement simple au premier degré pour l'ensemble des axes de rotation – ou pour tous les degrés mais uniquement pour l'axe z comme suggéré – celle-ci se complique d'autant plus que le degré augmente. En somme, les calculs nécessaires pour définir les opérations sont très complexes et ne sont généralement définis que jusqu'au troisième degré. Le choix a donc été fait de ne pas mettre en œuvre ce traitement dans la version tridimensionnelle de la bibliothèque, plutôt que d'offrir une version restreinte. Cependant, force est de constater qu'il existe une forte demande de cet outil, notamment pour une utilisation dans un contexte de restitution binaurale. La rotation de la scène sonore, en fonction des rotations de la tête et afin de recréer de manière virtuelles les variations des indices de perception, semble être très importante pour la solidité du rendu. Aussi, il est envisagé de réaliser un compromis pour les versions futures, en publiant une version tridimensionnelle de l'outil, limité à la rotation autour de l'axe z ou en limitant au 3<sup>ème</sup> degré les rotations. Enfin, un ensemble de travaux tels que ceux de M. A. Blanco, M. Florez et M. Bermejo [Blanco & al, 1997] visent à définir – ou plutôt évaluer – des matrices de rotation pour les ordres élevés. Cependant aucune mise en œuvre n'a été réalisée dans le contexte de la spatialisation ambisonique<sup>148</sup>.

### Le décodage de champ sonore

L'objet *hoa.2/3d.decoder*~ permet de synthétiser les signaux pour un système de restitution. L'objet offre trois modes de restitution<sup>149</sup> :

Le premier dit « régulier » réalise une simple projection des signaux associés aux harmoniques dans la direction des haut-parleurs. Cette approche n'est donc adéquate que pour un système de haut-parleurs représentant une discrétisation parfaite de la sphère ou du cercle ambisonique [Figure 3.5]. Il est néanmoins possible de définir manuellement les angles des haut-parleurs dans des configurations non idéales<sup>150</sup>, la restitution étant alors sujette à des artefacts<sup>151</sup>. Mais cette fonctionnalité a son utilité pour expérimenter et comparer les approches.

---

<sup>148</sup> A vrai dire, la complexité mécanismes mis en jeu me dépasse et ce sujet m'a déjà demandé et me demanderai encore trop de recherches et d'investissement.

<sup>149</sup> La présentation actuelle se veut relativement brève, les questions du décodage du champ sonore seront discutées dans la suite de cette thèse de manière plus approfondie.

<sup>150</sup> L'opération est réalisable en 3D dans Max mais pas en 2D, contrairement à la mise en œuvre pour Pure Data qui offre la fonctionnalité pour les deux cas – ce problème devrait être corrigé dans des versions ultérieures.

<sup>151</sup> Dans la pratique cela ne semble pas toujours le cas, ou du moins négligeable sous certaines conditions. La question des effets indésirables est traitée par la suite dans cette thèse.

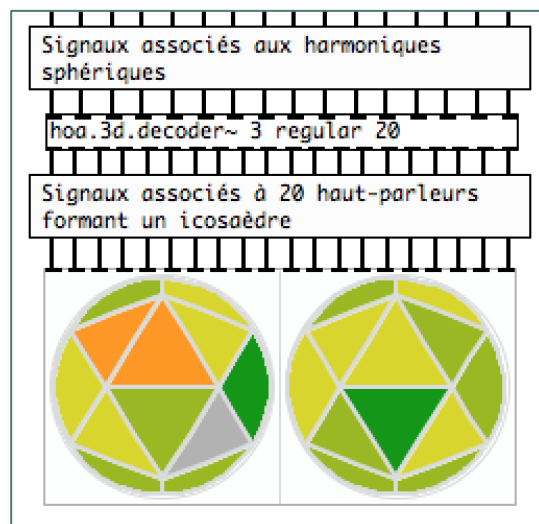


Figure 3.5. Opération de décodage dans le logiciel Pure Data avec l'objet *hoa.3d.decoder~* en projetant les signaux associés aux harmoniques sphériques à ordre de décomposition 3 pour un système de 20 haut-parleurs formant un polyèdre régulier, un icosaèdre.

Le deuxième mode permet une restitution via la synthèse binaurale offrant deux signaux associés aux deux oreilles. La version actuelle utilise la base de données d'HRIR (*Head-Related Impulses Responses*) LISTEN de l'IRCAM<sup>152</sup> dont la mise en œuvre actuelle permet d'atteindre l'ordre de décomposition 5 pour une synthèse binaurale bidimensionnelle et l'ordre de décomposition 3 pour une synthèse binaurale tridimensionnelle. Alors que les premières versions de la synthèse binaurale [Colafrancesco & al., 2013], avec la base de données CIPIC [Algazi & al., 2001] permettaient d'atteindre l'ordre de décomposition 35 pour une synthèse binaurale bidimensionnelle, la mise en œuvre construite sur une convolution dans le domaine temporel permettant d'atteindre cet ordre s'est rapidement révélée peu pérenne et peu en accord avec les critères d'un code libre et ouvert, principalement dû à la dépendance à la bibliothèque MKL pour réaliser cette opération<sup>153</sup>. Par la suite, le choix a donc été fait de réaliser les calculs matriciels directement dans le code sans dépendance<sup>154</sup> ainsi que d'utiliser la base de données de l'IRCAM pour des raisons de commodité lors du passage à la restitution tridimensionnelle. Il faut aussi noter que le choix d'une approche temporelle se révèle finalement peu adapté,

<sup>152</sup> La base de donnée est disponible en ligne à l'adresse <http://recherche.ircam.fr/equipes/salles/listen/index.html> (consulté en septembre 2017).

<sup>153</sup> Bibliothèque mathématique de Intel disponible en ligne à l'adresse <https://software.intel.com/en-us/mkl> (consulté en septembre 2017).

<sup>154</sup> L'approche offre des résultats satisfaisants compte tenu des calculs en jeu dont les performances sont proches des bibliothèques tels que la mise en œuvre de BLAS de Netlib disponible en ligne à l'adresse <http://www.netlib.org/blas/> (consulté en septembre 2017) déjà utilisée dans une version antérieure de la bibliothèque [Sèdes & al, 2014]. Néanmoins sur certaine machine il convient de réduire virtuellement la taille des HRIR (fonctionnalité possible dans l'objet) afin de revenir à une taille similaire à celle des HRIR utilisées dans la première version de la bibliothèque (256 échantillons au lieu de 512).

notamment lorsque la taille des HRIR devient importante, mettant en cause certaines des affirmations définies lors de précédentes publications [Colafrancesco & al., 2013], [Sèdes & al., 2014]. Enfin, pour de futurs développements et afin de répondre à de nombreuses demandes utilisateurs, il est envisagé d'offrir la possibilité à ces derniers de changer de banque d'HRIR. Pour cela plusieurs possibilités ont été envisagées. Le cadre de cette thèse ne permettant pas de les présenter en intégralité, une réponse satisfaisante sera tout de même apportée plus loin, dans la deuxième partie de ce travail.

Le dernier mode est le décodage irrégulier. Ce dernier, seulement fonctionnel en deux dimensions pour le moment, est basé sur une sur-discrétisation parfaite du cercle ambisonique permettant par la suite d'appliquer un panoramique d'amplitude, afin de compenser les artefacts de la configuration irrégulière de haut-parleurs. Cette approche, qui se voulait originale [Guillot, 2013], a déjà été formulée à quelques variantes près en dehors du contexte ambisonique en 2D par M. Poletti [Poletti, 2007] puis étendue à un système tridimensionnel et intégrée au modèle ambisonique par J-M. Batke et F. Keiler [Batke & al, 2010]. Et enfin généralisée sous le terme de « *All-round Ambisonic Decoding* » par F. Zotter, M. Frank et A. Sontacchi [Zotter & al., 2010], [Zotter & al., 2012]. L'approche sera aussi plus amplement discutée dans la suite de cette thèse. Bien que ne répondant pas encore à toutes les situations d'écoute et tous les systèmes de restitution circulaires, hémisphériques ou sphériques, l'objet *hoa.2/3d.decoder* offre déjà une grande flexibilité quant aux configurations et aux systèmes de restitutions.

### L'optimisation de champ sonore

Associée à l'opération de décodage ambisonique, la bibliothèque offre la possibilité d'optimiser le champ sonore selon les critères définis par M. A. Gerzon [Gerzon, 1974] et ceux définis par D. G. Malham [Malham, 1992] ainsi que J. Daniel, J-B. Rault et J-D. Polack, qu'ils nomment respectivement *max-r<sub>e</sub>* et *in-phase* [Daniel & al., 1998], [Daniel, 2001], p.311-316, et auxquels vient s'ajouter l'absence d'optimisation ou l'approche dite *basic*.

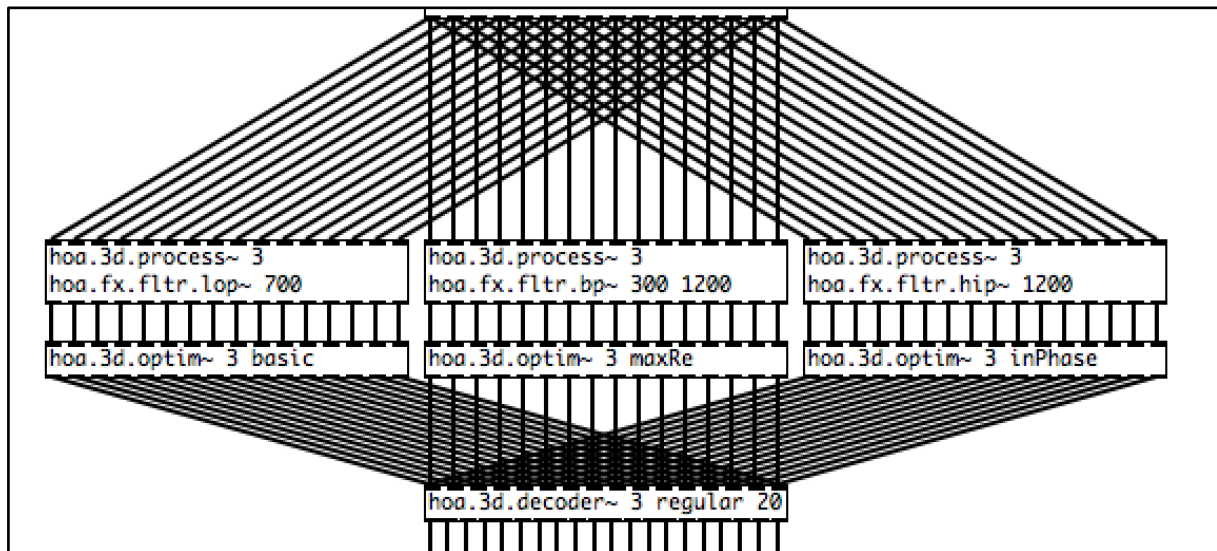


Figure 3.6. Représentation de l'application de différentes optimisations à un ordre de décomposition 3 selon le domaine fréquentiel, tel qu'inspiré de l'approche présentée par J. Daniel [Daniel, 2001]. Dans cet exemple, l'optimisation *basic* est appliquée en dessous de 700 Hz, l'optimisation *max-re* est appliquée entre 700 Hz et 1200 Hz et l'optimisation *in-phase* est appliquée au-delà de 1200 Hz. Les filtrages sont respectivement réalisés dans les abstractions *hoa.fx.fltr.lop~*, *hoa.fx.fltr.bp~* et *hoa.fx.fltr.hip~* utilisés dans l'objet *hoa.2d.process~* et les optimisations sont appliquées par les objets *hoa.3d.optim~*.

L'optimisation *basic* permet d'assurer, pour une configuration régulière de haut-parleurs et pour une position d'écoute centrale, que le vecteur vitesse soit colinéaire à la direction de la source sonore. Ce qui est une condition significative selon la théorie de localisation des sons de M. A. Gerzon [Gerzon, 1992], d'une bonne restitution dans les basses fréquences (fréquences inférieures à 700 Hz).

L'optimisation *max-re* consiste à concentrer l'énergie dans la ou les directions des sources sonores, en diminuant les contributions associées aux harmoniques d'ordres élevés comparé à la contribution de la composante omnidirectionnelle. Elle permet d'assurer que le vecteur d'énergie soit colinéaire à la direction de la source sonore, condition significative selon la théorie de localisation des sons de M. A. Gerzon [Gerzon, 1992] d'une bonne restitution dans les hautes fréquences (fréquences supérieures à 700 Hz).

L'optimisation *in-phase* consiste à concentrer encore plus l'énergie dans la ou les directions des sources sonores, en annulant notamment toute contribution en opposition de phase des haut-parleurs, d'où le nom. L'approche consiste à diminuer d'autant plus les contributions associées aux harmoniques d'ordres élevés, comparé à la contribution de la composante omnidirectionnelle permettant, selon D. G. Malham, d'assurer une meilleure stabilité de la scène sonore pour des auditeurs excentrés.



En somme, l'approche *basic* implique un équilibre égal entre les signaux associés aux harmoniques, quel que soit le degré. Alors que l'optimisation *max-r<sub>e</sub>*, et d'autant plus l'optimisation *in-phase*, diminuent les contributions des signaux associés aux harmoniques en fonction que les degrés augmentent. Cela implique, comme le suggère A. Farina, que l'énergie des champs sonores doit être rééquilibrée dans ces deux derniers cas [Farina, 1998]. Enfin, l'approche modulaire des logiciels Max et Pure Data, offre l'avantage de combiner les optimisations afin d'expérimenter des solutions avec des critères combinés. Que ce soit le décodage psychoacoustique de M. A. Gerzon ou des combinaisons des optimisations *max-r<sub>e</sub>* et *in-phase*, tels que présentés par J. Daniel, [Daniel & al., 1998], [Daniel, 2001], p.311-316 afin de satisfaire des conditions d'écoute variées [Figure 3.6]. Seront présentées par la suite les améliorations possibles de cette approche, et notamment dans quelles conditions et comment mettre dans un contexte d'une approche musicale, des traitements dans le domaine des harmoniques sphériques.

### L'analyse de champ sonore

L'objet *hoa.2/3d.vector~* permet de calculer le vecteur vitesse et le vecteur énergie dans un champ sonore, à partir des signaux destinés aux haut-parleurs et de la configuration du système de haut-parleurs. Le vecteur vitesse permettant de représenter la qualité de restitution basses fréquences, et le vecteur énergie permettant de représenter la qualité de restitution hautes fréquences, ces outils sont d'une grande importance pour analyser l'optimisation de l'opération de décodage. Notamment pour des sources sonores directionnelles. L'opération consiste en général à encoder un signal unitaire et, suite au décodage et aux optimisations du champ sonore, d'analyser la norme et l'angle des vecteurs en fonction des directions d'encodage données.

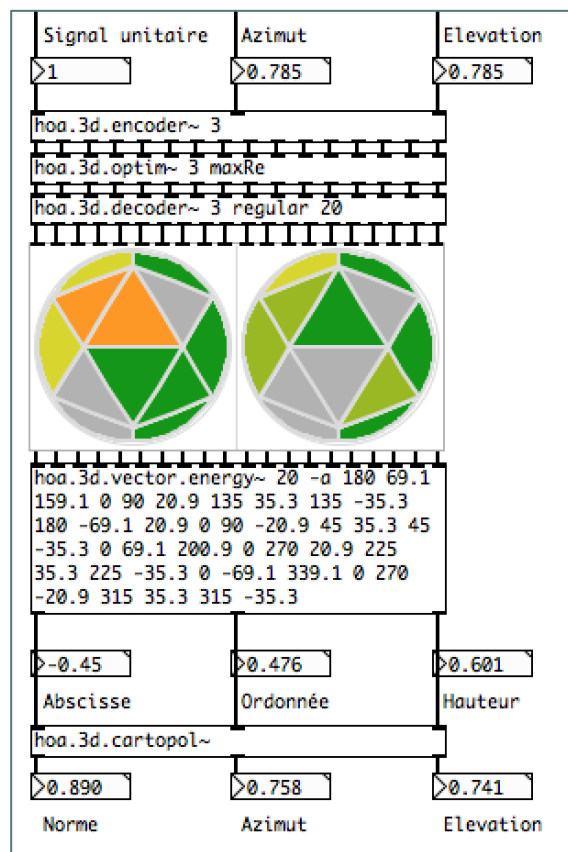


Figure 3.7. Patch Pure Data permettant de calculer le vecteur énergie d'un champ sonore. Le champ sonore tridimensionnel est composé d'un signal unitaire encodé à l'ordre 3. Une optimisation *max-r<sub>e</sub>* est opérée avant le décodage pour un système de haut-parleur formant un icosaèdre. Le champ sonore est alors analysé pour les haut-fréquences. L'enjeu est de comparer la valeur du signal unitaire et la direction d'encodage, affichés en haut, avec la norme et la direction du vecteur énergie, affichés en bas. Une différence entre la norme du vecteur énergie et la valeur du signal unitaire est représentative d'une diminution de la résolution angulaire. Et une différence entre les angles d'encodage et les angles du vecteur unitaire est représentative d'une déformation directionnelle du champ sonore.

Cet objet n'est disponible que dans la mise en œuvre pour le logiciel Max de la bibliothèque HOA. Néanmoins une mise en œuvre sous la forme d'objets compilés et d'une série d'abstractions est disponible pour le logiciel Pure Data de façon indépendante [Figure 3.7]. Ils font partie d'un ensemble d'outils plus développés et spécifiquement destinés à l'analyse de champ sonore<sup>155</sup>, qui sera discuté dans la deuxième partie de cette thèse.

### 3.2.2. Les interfaces graphiques

Un autre ensemble d'outils important de la bibliothèque HOA implique de faciliter le contrôle et l'appropriation des traitements, par l'utilisation d'interfaces graphiques. Ces interfaces graphiques peuvent être séparées en deux groupes : le premier visant à contrôler ou transformer

<sup>155</sup> Ces outils sont donc encore en développement mais d'ores et déjà fonctionnels et disponibles en ligne sur un répertoire Github spécifique à l'adresse accédé en septembre 2017 <https://github.com/pierreguillot/HoaLibrary-Analyzer>.

le champ sonore et le deuxième visant à le représenter. Le premier groupe comporte une interface de contrôle de sources sonores ponctuelles via l'objet *hoa.2/3d.map*, une multi-glissière circulaire via l'objet *hoa.2d.space*, un objet spécifique au moteur sonore *hoa.2d.recomposer* et une glissière de contrôle de gain multicanal *hoa.gain~*. De même que certains objets du moteur sonore ont été exclus de la présentation, ces deux dernières interfaces graphiques ne seront pas traitées en profondeur. La première pour la même raison que son analogue en traitement du signal *hoa.2d.recomposer~* et le second car, malgré son utilité notoire, l'objet est extrêmement similaire à l'objet *live.gain~* du logiciel Max et n'est pas réellement spécifique à l'approche ambisonique, mais plutôt au multicanal de manière général. C'est pour ces raisons que ces objets n'ont pas été inclus dans la version Pure Data de la bibliothèque. Néanmoins une description détaillée de leurs fonctionnalités est présentée par E. Paris dans [Paris, 2013]. Cette présentation complète le travail accompli par E. Paris, en y intégrant notamment les réponses aux questions de l'adaptation tridimensionnelle des interfaces.

### Le contrôle de sources sonores

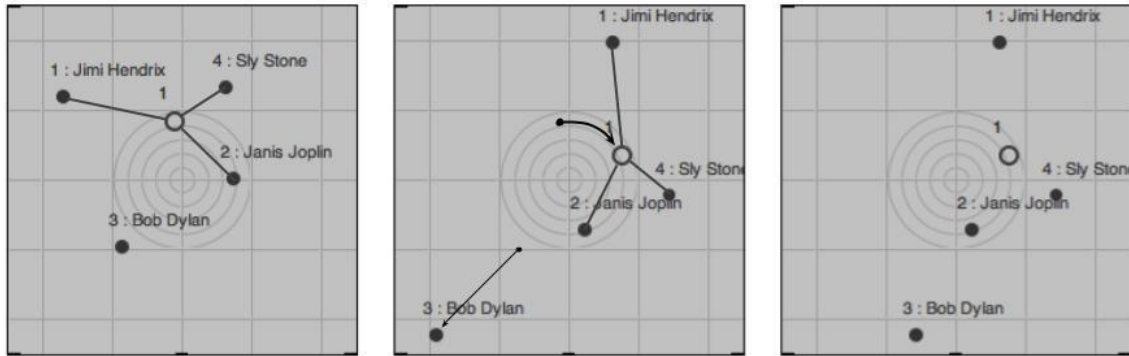
Les coordonnées des sources sonores au sein de l'objet *hoa.map~* sont généralement définies à l'aide de messages<sup>156</sup>. La spatialisation de multiples sources sonores nécessite de générer de multiples lots de coordonnées associées aux indices des sources, rendant l'opération peu ergonomique. Dès lors, le contrôle des coordonnées spatiales des sources sonores peut se révéler relativement abstrait, ou du moins peu maniable pour les utilisateurs. Tout comme l'ICST le propose avec l'interface graphique *ambimonitor* [Schacher, 2010], [Schacher & al., 2006] ou l'IRCAM avec le *spat.viewer*<sup>157</sup>, la bibliothèque HOA offre elle aussi une interface de contrôle de position et d'édition de trajectoire via l'objet *hoa2/3d.map*. Cela afin d'éviter toute dépendance logicielle externe, mais surtout afin d'être adaptée à l'approche spécifique de la bibliothèque.

---

<sup>156</sup> Dans la pratique, les signaux étant utilisés dans des cas relativement spécifiques tels qu'une rotation cyclique de la source où les coordonnées peuvent être générées à l'aide d'oscillateurs sinusoidaux ou en dents-de-scie.

<sup>157</sup> La bibliothèque de spatialisation *Spat* de l'IRCAM est disponible notamment pour le logiciel Max en ligne <http://forumnet.ircam.fr/fr/produit/spat/> en septembre 2017.

Vue du dessus (axes x et y)



Vue frontale (axes x et z)



Figure 3.8. Représentation d'interactions de l'utilisateur sur des interfaces graphiques *hoa2/3d.map* dans le logiciel Pure Data. Les interactions se lisent de gauche à droite. Les interfaces d'en haut représentent une vue du dessus (les axes x et y) et sont liées aux interfaces d'en bas qui représentent une vue frontale (les axes x et z), afin d'avoir un aperçu tridimensionnel. Les points noirs numérotés de 1 à 4 représentent les sources sonores. Le cercle numéroté 1 représente un groupe comprenant les sources 1, 2 et 4. Des interfaces de gauche aux interfaces du centre, une opération de rotation (où la distance est contrainte) sur le groupe, et donc sur ses sources, est réalisée via la vue du dessus. La source 3 subit quant à elle une translation. Ces déplacements se répercutent automatiquement sur la vue frontale par une modification des sources sur l'axe x. De l'interface du centre aux interfaces de droite, le groupe subit une translation vers le bas et la source 3 subit une translation vers le haut via la vue frontale en contraignant l'axe des x. Ces modifications impliquent que la hauteur ne soit pas visible sur la vue du dessus.

L'enjeu derrière cette interface graphique est d'offrir un outil simple, ergonomique et adapté aux particularités de la bibliothèque HOA. Les sources sont représentées par des points noirs, qu'il est possible de déplacer à la souris sur un système quadrillé. Le système est centré sur un auditeur virtuel et la sphère unité, ou le cercle unité, est représenté par une série de cinq cercles concentriques [Figure 3.8]. Pour une représentation tridimensionnelle, il est possible d'utiliser conjointement deux ou trois interfaces liées entre-elles par un nom de référence et associées à des plans différents – horizontal, médian ou frontal – afin que le déplacement sur une interface apparaisse sur les autres interfaces. Lors du déplacement des sources, l'interface génère automatiquement les coordonnées cartésiennes ou sphériques destinées à l'objet de traitement du signal *hoa.2/3d.map*. Ce afin d'éviter toute opération intermédiaire à l'utilisateur final. L'interface graphique propose un nombre important d'opérations :

- Créer, supprimer, rendre muet ou audible une source via un menu contextuel
- Contraindre les déplacements selon une approche cartésienne sur les axes x, y ou z, selon une approche sphérique en appliquant des rotations autour de l'axe ou en modifiant exclusivement le rayon
- Sélectionner des sources pour créer un groupe permettant d'appliquer les opérations précédentes à l'ensemble des sources qui le compose

Cette interface offre de plus la possibilité de sauvegarder des états, avec le mécanisme natif de pré-réglages de logiciel Max ou l'interface graphique *c.preset* de la bibliothèque *Cream* [Guillot, 2014] offrant la possibilité d'interpoler entre eux, pour générer des trajectoires. Enfin l'ensemble des opérations peuvent être réalisées en envoyant des messages à l'interface, permettant de s'en servir uniquement pour représenter des états. Dans le cas d'un contrôle avec une interface externe, MIDI ou tablette graphique, il peut être nécessaire d'avoir un retour graphique des positions effectives des sources sonores.

### La multi-glissières circulaire

L'interface graphique *hoa.2d.space* est un objet relativement simple, mais offrant un fort potentiel musical. Il s'agit d'une interface qui pourrait être caractérisée de type multi-glissières circulaire (*circular multi-sliders*) [Figure 3.9]. Elle offre la possibilité de contrôler une série de valeurs représentées par des points, qu'il est possible d'associer à des curseurs sur un disque, en jouant sur leurs distances par rapport au centre. Chaque point étant contraint par son angle. En modifiant la distance d'un point, l'interface génère la liste des valeurs correspondant aux distances – initialement entre 0 et 1 – mais il est possible de modifier cette plage de valeur.

Ces valeurs sont utilisables pour contrôler les paramètres d'un traitement associé à un dispositif circulaire, tel que défini par le domaine des ondes planes. L'une des approches les plus musicales est d'associer chaque valeur de gain à un signal associé à une onde plane, ou d'un signal associé à un haut-parleur, et de réaliser un filtrage spatial sur certaines parties de l'espace [Guillot, 2013].

L'interface offre la possibilité de gérer les curseurs de manière groupée en augmentant ou diminuant de manière globale leurs valeurs respectives des curseurs, ou en appliquant une rotation. Chaque valeur d'un curseur se trouvant alors être en une combinaison linéaire entre son état précédant et la valeur d'un curseur adjacent (avant ou après suivant le sens de rotation). En soi, l'interface n'offre pas beaucoup plus que l'objet *multislider* du logiciel Max, mais son approche circulaire en fait un outil très ergonomique dans le domaine de la spatialisation, et notamment pour les systèmes enveloppant. Il faut cependant noter certaines limites : dès que l'on souhaite adapter cette approche à une représentation tridimensionnelle, aucune solution satisfaisante n'a pour le moment été trouvée.

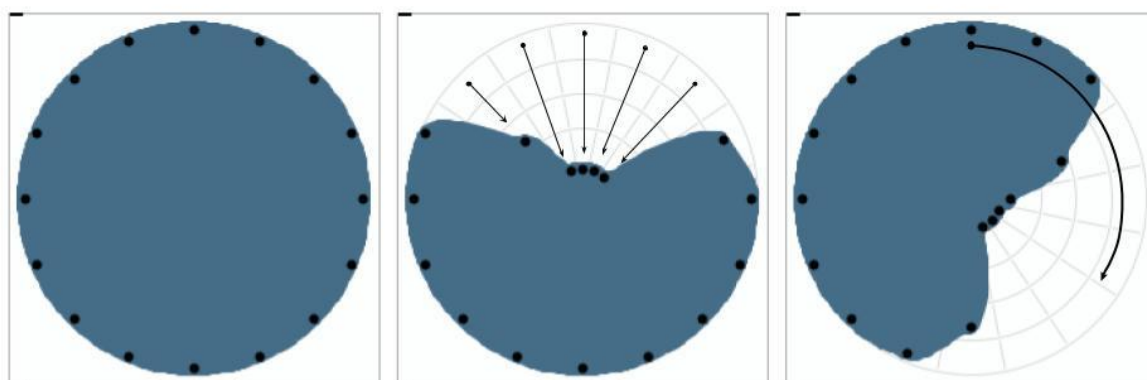


Figure 3.9. Interface *hoa.2d.space* dans le logiciel Pure Data. Les points noirs servent de curseur permettant de définir une série de valeurs associés à un espace circulaire. La masse bleue est une représentation spatiale de cette série de valeur. Les interfaces se lisent de gauche à droite. Sur l'interface de gauche les curseurs sont sur la position maximum. Sur l'interface au centre les curseurs pouvant être associés à une zone frontale de l'espace sont baissés enfin sur l'interface de droite une rotation est appliquée dans le sens horaire modifiant globalement l'ensemble des curseurs.

La bibliothèque HOA offre aussi deux outils graphiques particulièrement utiles pour la compréhension des mécanismes en jeu : l'interface *hoa.2/3d.scope~* offre une représentation du champ sonore dans le domaine intermédiaire des harmoniques spatiaux et l'interface *hoa.2/3d.meter~* offre une représentation du champ sonore dans le domaine « définitive » associé aux haut-parleurs.

### Représentation intermédiaire du champ sonore

L'interface *hoa.2/3d.scope~* réalise une projection des signaux associés aux harmoniques sphériques ou circulaires sur un ensemble très important de points. Elle représente les valeurs d'amplitude de ces points dans un espace tridimensionnel ou bidimensionnel avec la distance par rapport au centre définissant la valeur absolue des amplitudes, associé à la norme des

harmoniques. Et le signe de l'amplitude, associé à la polarité des harmoniques, par des couleurs, en général bleu pour les valeurs négatives et rouge pour les valeurs positives<sup>158</sup>. Ces interfaces permettent donc d'offrir une représentation du champ sonore sous la forme de signaux associés aux harmoniques spatiaux à un instant donné.

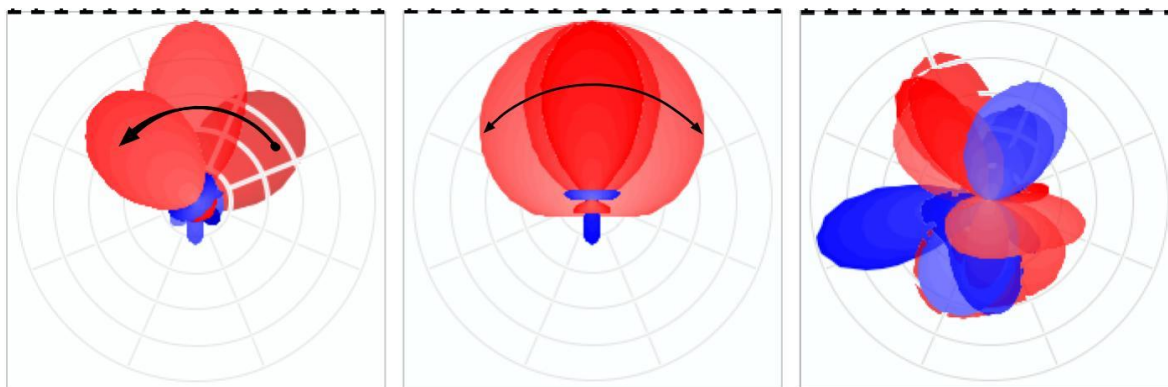


Figure 3.10. Interface `hoa.3d.scope~` dans Pure Data représentant la projection des signaux associés aux harmoniques sphériques dans l'espace. Chaque figure représente une superposition des figures dessinées par l'interface à trois instants différents. A gauche, il est possible de reconnaître les variations de l'azimut et de l'élévation d'une source directionnelle. Au centre, la superposition des figures représente une variation de la résolution angulaire. A droite, les figures chaotiques pointant dans toutes les directions sont représentatives d'un champ sonore diffus.

De manière pratique les figures dessinées par l'interface peuvent être associées aux figures directives du champ sonore, à des instants donnés. Avec un peu d'expérience, il est possible d'interpréter certains motifs caractéristiques comme la direction d'une source sonore, une variation de la résolution angulaire ou encore un champ diffus [Figure 3.10]. Enfin, cet outil peut être utilisé simplement pour représenter dans l'espace un harmonique spécifique, comme c'est le cas dans le deuxième chapitre de cette partie [Figure 2.2]. Il est néanmoins important de noter que malgré le potentiel de cette interface, son utilisation dans un contexte final de représentation ou d'enregistrement, est fortement déconseillé en raison de la complexité des calculs, demandant énormément de ressources processeurs<sup>159</sup>. Aussi cet outil est principalement destiné aux phases de prototypage et d'expérimentation, ainsi qu'à des fins didactiques et pédagogiques.

<sup>158</sup> Certains paramètres de représentation comme les couleurs peuvent être modifiés par l'utilisateur.

<sup>159</sup> L'opération de projection est réalisée sur des centaines de points, équivalent à une opération de décodage pour plus de cent haut-parleurs. Et ni le logiciel Max ni le logiciel Pure Data ne sont adaptés à une représentation tridimensionnelle aussi l'ensemble des effets 3D sont réalisés « manuellement » en représentant la hauteur couche après couche, décuplant les calculs pour le rendu graphique.

## Représentation du champ sonore restitué

L'interface *hoa.2/3d.meter~* offre une série de modulomètres de type *PPM* (*Peak Programme Meter*), répartis de manière circulaire ou sphérique et dirigés vers le centre, permettant chacun de représenter le niveau de crête d'un signal audio dont l'amplitude varie entre -1 et 1. Cet outil a été élaboré afin d'avoir un retour visuel des signaux du champ sonore visant à être restitués plus en adéquation avec les dispositifs de haut-parleurs utilisés.

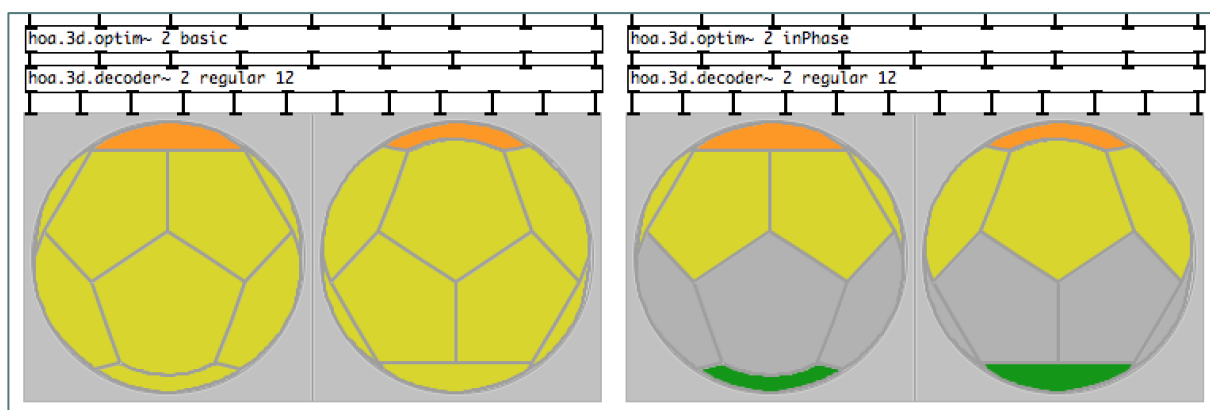


Figure 3.11. Interfaces graphiques *hoa.3d.meter~* dans Pure Data. Les interfaces représentent le niveau de crête pour un ensemble de 12 haut-parleurs formant un dodécaèdre sur la sphère. Pour chaque figure, à gauche l'interface représente une vue du dessus et à droite une vue du dessous. Le champ sonore composé d'une source directionnelle orientée vers 0 degré d'azimut et d'élévation n'est pas optimisé à gauche avant d'être décodé, résultant en une contribution de l'ensemble des signaux associés aux haut-parleurs. A droite, la chaîne de traitement opère une optimisation *in-phase* sur le champ sonore avant le décodage, résultant une annulation presque complète des contributions des haut-parleurs opposés à la source.

La version bidimensionnelle de l'interface représente chaque canal audio par un ensemble de LED<sup>160</sup> orientées vers le centre d'un cercle formant une « part de camembert » ou une « pointe de pizza ». Le signal de chaque haut-parleur est alors représenté par une série de LED s'illuminant en fonction du niveau de crête du signal. La version tridimensionnelle de l'interface discrétise la sphère en un ensemble de surface associées aux haut-parleurs. Ces surfaces sont définies suite à l'élaboration d'un diagramme de Voronoï de la sphère à partir des coordonnées des haut-parleurs. Aussi, le signal de chaque haut-parleur est associé à une surface dont la couleur varie en fonction du niveau de crête du signal. Afin d'offrir une représentation de l'ensemble de la sphère, il est possible d'offrir à la fois une vue du dessus et une vue du dessous [Figure 3.11]. Cette représentation semble relativement adaptée au vu des contraintes techniques de dessin tridimensionnelle dans les logiciels Max et Pure Data. Elle permet notamment d'appréhender et d'analyser rapidement les signaux envoyés aux haut-parleurs.

<sup>160</sup> Il s'agit graphiquement de petit arcs de cercles associés à une plage de décibel.



D'autres approches sont néanmoins envisagées afin d'améliorer la lisibilité. Cependant, au regard de la complexité de la mise en œuvre d'un diagramme de Voronoï sphérique et du temps de mise en place nécessaire à le réaliser, il m'a été impossible d'explorer plus en profondeur d'autres approches.

### 3.2.3. Vectorisation des traitements

L'enjeu principal de cette bibliothèque est de pouvoir offrir un système facilitant l'accès et la mise en œuvre de traitement de l'espace et du son sur les signaux associés aux harmoniques sphériques. Comme cela a été suggéré par leurs représentations dans l'espace [Figure 2.2], les éléments essentiels sur lesquels il est possible d'élaborer des traitements et des opérations dans les domaines des harmoniques sphériques, consistent à fonder l'approche sur les figures de directivité associées aux harmoniques, ainsi que leurs relations spatiales et notamment de symétrie. Ces caractéristiques dépendent de l'ordre de décomposition, du degré et de l'ordre azimutal des harmoniques. Cette relation est d'ailleurs évidente via la formulation mathématique des harmoniques sphériques [Équation 2.3], les opérations d'encodage et de décodage en sont d'ailleurs les meilleurs exemples jusqu'à présent. L'objectif au sein des logiciels de programmation modulaire tels que Max et Pure data, est alors de créer et proposer des patchs de traitements destinés au domaine des harmoniques sphériques, et dont les opérations soient fonctions des degrés et des ordres azimutaux des harmoniques. Aussi a-t-il paru nécessaire de proposer un outil permettant de charger autant d'instances d'un patch qu'il y a de signaux associés aux harmoniques sphériques. Tout en offrant un système permettant d'adapter ces patchs aux spécificités des harmoniques, via leurs degrés et leurs ordres azimutaux ainsi que l'ordre de décomposition. Le procédé nécessite de « vectoriser » les opérations au domaine des harmoniques sphériques.

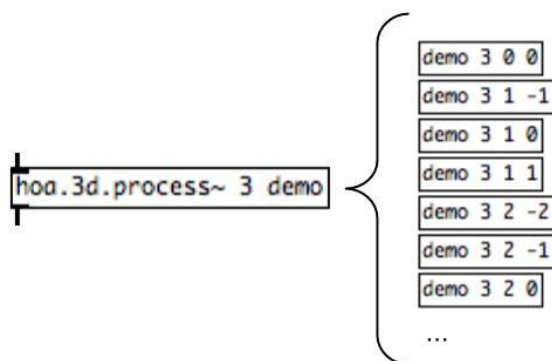


Figure 3.12. Représentation de l'instanciation dynamique d'abstractions par l'objet `hoa.3d.process~`. Dans cet exemple, le patch `demo` est instancié autant de fois qu'il y a d'harmoniques sphériques pour un ordre de décomposition 3.

Pour cela, la bibliothèque HOA propose une petite sous collection d'outils organisés autour d'un outil central, l'objet *hoa.2/3d.process~*. Cet objet peut être vu comme l'équivalent adapté à l'approche ambisonique des objets *poly~* dans Max et *clone* dans Pure Data. Cet objet prend en argument, après l'ordre de décomposition, le nom d'un patch qui charge autant de fois que nécessaire, comme défini dans la partie précédente [Figure 3.12]. Ce qui fait l'originalité de l'approche, comparé à ces homologues, est la gestion particulière des entrées et des sorties. Alors que dans l'objet *clone*<sup>161</sup> les entrées et sorties sont globales à toutes les instances, l'objet *hoa.2/3d.process~* offre la possibilité de créer des entrées et des sorties qui soient spécifiques à chaque instance.

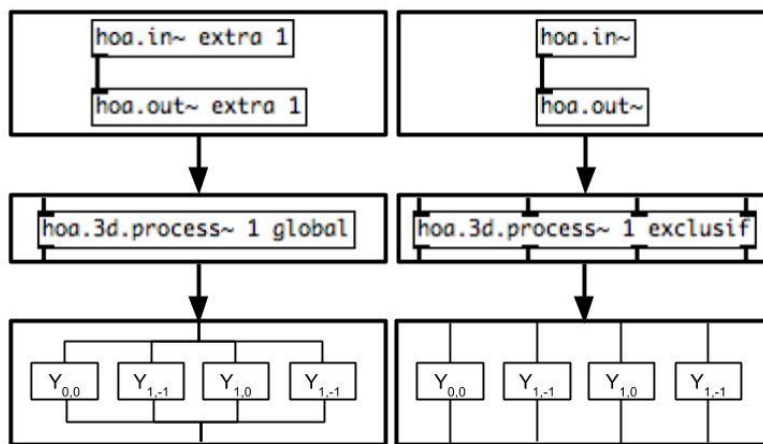


Figure 3.13. Représentation de la gestion des entrées et sorties de l'objet *hoa.3d.process~*. En haut sont représentés les patches chargés. Au milieu sont représentés les boîtes des objets *hoa.3d.process~* avec leurs entrées et sorties dépendant des patches. Et en bas sont représentés les routages des signaux. A gauche, l'argument *extra* des objets *hoa.in~* et *hoa.out~* défini que ces objets sont associés à une entrée et une sortie globale aux 4 instances de patch chargés par l'objet *hoa.3d.process~*. A droite, l'absence d'argument objets *hoa.in~* et *hoa.out~* défini que ces objets sont associés à une entrée et une sortie respectives aux 4 instances de patch chargés par l'objet *hoa.3d.process~*.

Les entrées et les sorties sont définies à l'intérieure d'un patch destiné à l'objet *hoa.2/3d.process~* avec les objets *hoa.in* et *hoa.out*, ou leurs équivalents en signal *hoa.in~* et *hoa.out~*. De base, ces entrées et sorties sont spécifiques à chaque instance. C'est à dire qu'un *hoa.in~* tel que c'est le cas dans une abstraction, apparaîtra sur la boîte de l'objet *hoa.2/3d.process~* comme autant d'entrées qu'il y a d'instances. Tout message ou tout signal passé à une de ces entrées est envoyé uniquement à l'instance à laquelle cette entrée est associée.

<sup>161</sup> Afin d'éviter des redondances, les comparaisons sont restreintes à Pure Data, cependant elles sont aussi valables dans Max.

En ajoutant un argument *extra*, puis une valeur définissant son indice, à l'objet *hoa.in~*, celle-ci apparaîtra alors sur la boîte de l'objet *hoa.2/3d.process~* sous la forme d'une seule entrée globale à l'ensemble des instances. Tout message ou tout signal passé à cette entrée est envoyé à l'ensemble des instances de manière analogue au système natif de Pure Data [Figure 3.13]. Ce fonctionnement est aussi valide pour les sorties et il est bien sûr possible de combiner les types d'entrées et de sorties.

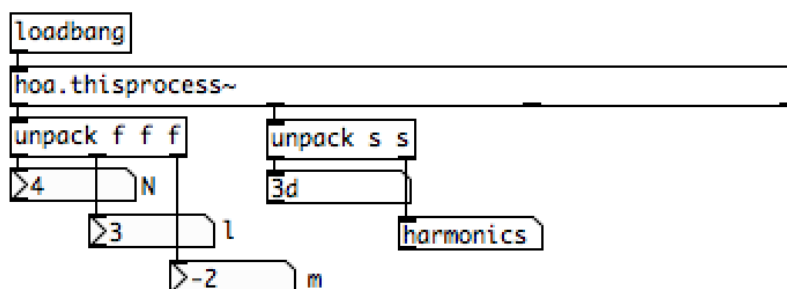


Figure 3.14. Objet *hoa.thisprocess~* dans Pure Data permettant de récupérer l'ordre de décomposition  $N$  utilisé par l'objet *hoa.2/3d.process~*, le degré  $l$  et l'ordre azimutal  $m$  de l'harmonique auquel le patch est associé, ainsi que le nombre de dimensions spatiales ( $2d$  ou  $3d$ ) et le domaine (*harmonics* ou *planewaves*).

Enfin l'objet *hoa.thisprocess~* utilisé à l'intérieur d'un patch destiné à l'objet *hoa.2/3d.process~* permet de manière analogue à l'objet *thispoly~* de Max, de recevoir des informations qui sont dans ce cas relatives à l'instance de l'harmonique auquel l'abstraction est associée, ainsi que l'ordre de décomposition et le nombre de dimensions spatiales (2D ou 3D) [Figure 3.14]. De plus, l'objet offre un système de gestion d'arguments et d'attributs identiques à l'objet *patcherargs* de Max, permettant de créer des patches plus dynamiques où les mécanismes et opérations peuvent varier en fonction d'arguments et d'attributs définis dans l'objet *hoa.2/3d.process~*.

Il est important de noter à titre indicatif que l'objet peut fonctionner aussi dans le domaine des ondes planes. Pour cela, le troisième argument défini dans l'objet *hoa.2/3d.process~* doit être *planewaves* (au lieu de *harmonics* qui est facultatif) et le premier argument ne fait plus référence à l'ordre de décomposition, mais au nombre d'ondes planes discrétisant l'espace. Les informations reçues via l'objet *hoa.thisprocess~* sont alors le nombre d'ondes planes et l'indice de l'onde plane. En pratique, l'idée de ce domaine est dans ce cas quelque peu abstraite, car le mécanisme est extrêmement similaire à l'objet *poly~* ou *clone*.

Cet objet permet donc d'expérimenter facilement dans le domaine des harmoniques sphériques, grâce à un système dynamique. Mais la réelle spécificité de cet outil est d'amener implicitement à penser les traitements comme fonction des caractéristiques des harmoniques sphériques, en offrant la possibilité d'adapter les traitements à leurs degrés, leurs ordres azimutaux et à l'ordre de décomposition. Il apparaît aussi que cet objet est un excellent outil pédagogique et didactique, car il permet de mettre à la portée de toute personne familière avec les logiciels Max ou Pure Data, de comprendre les mécanismes mis en œuvre.

### 3.3. Bilan

Ce chapitre a dans un premier temps présenté un aperçu général du développement de la bibliothèque HOA. Il a été montré que ces travaux ont émergé au CICM de l'Université Paris 8, dans un contexte scientifique où la spatialisation du son était déjà un thème central des recherches menées et des préoccupations musicales depuis plusieurs années. Les recherches autour de la bibliothèque HOA se sont déroulées sur de nombreuses années, allant de 2013 à aujourd'hui. De ce fait, elles ont recouvert de multiples aspects – la création de traitements, les interfaces graphiques, etc. – donnant lieu à de nombreux livrables sur plusieurs plateformes. Une des caractéristiques de ces recherches et de ces développements, est qu'elles ont été nourries par les retours des musiciens et des compositeurs, dont les nombreuses œuvres en sont les témoignages. Ces recherches ont également été alimentées par les retours des étudiants au travers des cours, des ateliers de compositions et des séminaires dédiés à la spatialisation ambisonique. Et plus largement encore, ces recherches ont été nourries par les apports de la grande communauté d'utilisateurs, que la bibliothèque HOA intéresse aujourd'hui : que ce soit en ligne, via notamment les forums et réseaux sociaux, ou lors de présentations dans des partenariats et des conventions.

Par la suite, les différents outils disponibles dans la bibliothèque ont été présentés de manière détaillée, afin de faire ressortir leurs spécificités. Les enjeux auxquels ils tentent de répondre ont aussi été soulevés : les questions d'optimisation et d'utilisabilité dans un contexte de musique temps-réel, les aspects d'ergonomie destinés à faciliter l'utilisation des outils ou encore l'approche pédagogique et didactique visant à donner les clés aux utilisateurs pour s'approprier le modèle ambisonique. Finalement, l'un des aspects les plus caractéristiques est la volonté d'offrir des outils permettant de nouvelles expérimentations de l'espace et du son, via la mise en œuvre de traitements originaux dans le domaine des harmoniques sphériques. Pour donner suite aux travaux présentés dans ce chapitre, il sera intéressant de revenir

confronter les approches proposées et les outils logiciels fournis, afin d'en définir les limites et, à partir de là, envisager de possibles améliorations.

Dans cette partie, l'ensemble des outils nécessaires aux expérimentations musicales dans le domaine ambisonique a été présenté. Cela comprend les outils logiciels dont les spécificités viennent d'être présentées dans ce dernier chapitre, et notamment leur approche dynamique. Mais aussi les outils mathématiques, via la décomposition en harmoniques sphériques, avec lesquels il sera possible de jouer pour créer de nouvelles approches originales de la mise en espace du son. En parallèle à ces outils, les modèles théoriques et pratiques – appliqués aux domaines acoustiques et psychoacoustiques – sur lesquels reposent les harmoniques sphériques, ont été définis. Cela afin d'analyser et critiquer les expérimentations et permettre une appropriation musicale du modèle. Enfin, grâce aux enjeux et aux spécifications fonctionnelles ainsi que techniques présentées dans le premier chapitre, il est possible de mettre en perspective les fondements même de l'approche ambisonique. Aussi est-il envisageable, à présent, de présenter les travaux qui sont au cœur de ce sujet de thèse.

## Partie II – Appropriation musicale de l’ambisonie

L’objectif de cette partie est de montrer de manière pratique comment l’appropriation musicale ambisonique permet, un fois confrontée aux différents principes et aux différentes théories présentées dans la partie précédente, de remettre en questions les enjeux même auxquels ce modèle tente de répondre. Ces raisonnements vont permettre d’envisager différemment les outils mathématiques utilisés initialement, pour répondre aux spécifications de l’approche ambisonique, et notamment de la décomposition en harmoniques sphériques. Puis, cela permettra de définir des nouveaux enjeux pour la mise en espace du son. En considérant ces nouveaux enjeux et les expérimentations qui auront permis de les définir, de nouvelles approches construites sur la notion de représentation intermédiaire et abstraite de l’espace pourront être considérées, amenant de nouvelles perspectives qui seront traitées dans la troisième partie de cette thèse.

A présent que les modèles théoriques et pratiques de l’ambisonie ont été présentés et que des outils adaptés à la mise en œuvre de traitements dans le domaine des harmoniques sphériques ont été proposés, il est possible d’envisager des nouvelles opérations de synthèse et de transformation du champ sonore et d’explorer leurs possibilités musicales. C’est donc par cette approche expérimentale, que la démarche d’appropriation commence. Le premier chapitre de la partie proposera un ensemble de mises en œuvre où le modèle ambisonique est appréhendé selon un axe musical. Les enjeux principaux seront de présenter le potentiel de ces nouvelles approches, de définir les méthodologies employées et d’en faire ressortir les notions importantes. Dans le deuxième chapitre, les répercussions de ces expérimentations sur le modèle ambisonique original seront considérées, et des solutions d’adaptation seront proposées. A partir de l’analyse et la critique de ces procédés, il sera alors possible de répondre aux objectifs présentés, et d’établir le raisonnement amenant aux nouvelles perspectives. Celles-ci seront explorées dans la dernière partie de ce travail de thèse.

## 4. Approche originale

En 2006, G. Wakefield présente une bibliothèque ambisonique de spatialisation du son, allant jusqu'à l'ordre 3 pour le logiciel Max qu'il développe à l'Université de Santa Barbara en Californie [Wakefield, 2006]. Cette bibliothèque ne semble aujourd'hui plus maintenue<sup>162</sup> mais elle est sur un plan historique des plus intéressantes, car elle offre la possibilité d'opérer sur les signaux associés aux harmoniques sphériques. A ce propos, l'auteur de cette bibliothèque remarque que « l'accès aux signaux du domaine ambisonique dans un environnement de traitement du signal flexible, amène à expérimenter de nouvelles manipulations du champ sonore, et offre la possibilité de synthétiser directement différents signaux directionnels (à la place ou en plus de l'encodage de sources virtuelles)<sup>163</sup> ». Les expérimentations présentées ne restent cependant qu'à un stade préliminaire. G. Wakefield propose, sans pour autant le mettre en œuvre, de générer les signaux associés aux harmoniques sphériques via un système de synthèse par modulation de fréquence et émet l'hypothèse que cela devrait ouvrir sur de relativement complexes et nouveaux événements spatiaux. Par la suite, il explique avoir utilisé une série de filtres en peigne et un système de matricage avec réinjection, directement sur les canaux ambisoniques produisant une image spatiale. Il qualifie d'ailleurs cette image de complexe, subtile et chaotique, amenant à réaliser de plus amples recherches sur le sujet. Ces deux approches, bien que n'étant pas formalisées et ne restant encore qu'à l'état d'ébauche, offrent un énorme potentiel musical. C'est sur ces constats que les expérimentations et les développements dans le cadre des projets bibliothèque HOA, prennent racine.

L'enjeu est de définir les cadres et méthodes permettant la création de traitements spatialisés dans le domaine des harmoniques sphériques, afin de pouvoir l'étendre par la suite à d'autres approches. L'objectif de cette démarche est de montrer comment la mise en œuvre de traitements dans le domaine spatial et la façon dont les traitements apparaissent sur le plan auditif, sont fortement liés au modèle de représentation de l'espace et du champ sonore. Pour cela, cette partie propose de revenir sur quelques-unes des expérimentations réalisées dans le cadre de cette thèse, afin de synthétiser et transformer des champs sonores dans le domaine des harmoniques sphériques et circulaires. Dans une logique de compréhension des différentes

---

<sup>162</sup> La bibliothèque est disponible à <http://www.mat.ucsb.edu/~wakefield/software.html> (consulté en mai 2017).

<sup>163</sup> [Wakefield, 2006], p. 126, traduit de l'anglais : « *Furthermore, access to the Ambisonic domain signals within a flexible DSP environment invites experimentation with new manipulations of the sound field, and the possibility of directly synthesizing the directional difference signals (rather than or in addition to encoding virtual sources).* »

étapes de recherches, il est nécessaire de revenir brièvement sur la méthodologie expérimentale employée.

### Méthodologie expérimentale

De même que certaines opérations, qu'elles soient mathématiques ou autres, prennent un sens particulier selon le contexte dans lesquelles elles sont mise en œuvres, les opérations du son appliquées aux signaux associés aux harmoniques sphériques, possèdent des propriétés sonores particulières qui diffèrent de celles offertes par des approches plus conventionnelles. A titre de comparaison, l'application d'un filtre à réponse impulsionnelle finie [Smith, 2007] peut être réalisée dans le domaine temporel, en procédant à une combinaison linéaire entre le signal original et le signal représentant les coefficients du filtre. Mais aussi dans le domaine fréquentiel en réalisant une convolution sur les représentations fréquentielles sur ces signaux, obtenues par des transformées de Fourier discrètes. Il s'agit de deux opérations distinctes, mais les résultats obtenus sont identiques car leurs domaines d'applications sont différents. Réciproquement, appliquer la même opération, au sens algorithmique du terme, dans le domaine temporel et fréquentiel, n'aboutit pas nécessairement à un résultat identique. Ainsi une multiplication terme à terme des échantillons des signaux dans le domaine fréquentiel n'offre pas le même résultat que la multiplication terme à terme des échantillons des signaux dans le domaine temporel. Partant de ce constat, il est possible de distinguer deux approches particulières afin de jouer sur les spécificités de la représentation spatiale.

La première approche tente de créer et retrouver l'analogie parfaite d'un traitement ou d'une opération d'une représentation spatiale vers une autre. En théorie, l'exercice conceptuel et mathématique peut paraître vain en soi, mais il prend tout son sens dans certains cas pratiques, car les changements de domaines et de représentations permettent de simplifier et souvent d'optimiser les opérations. A ce sujet, l'exemple précédent de possibilités d'application d'un filtre à réponse impulsionnelle finie dans deux domaines différents, offre de grands avantages relatifs à l'optimisation des mises en œuvre selon le contexte, notamment pour le traitement du signal temps-réel et sans latence [Gardner, 1995]. Il en est de même en ambisonie avec la décomposition en harmoniques circulaires ou sphériques du champ sonore. L'approche ambisonique est d'ailleurs intrinsèquement liée aux avantages apportés par son changement de représentation de l'espace sonore. A son origine, l'ambisonie tente de répondre aux contraintes des techniques multicanales liées à la prise de son, au matriçage, à la dégradation du signal lors d'opérations telles que la réduction du bruit ou encore au déséquilibre et aux erreurs de phases lors de l'étape de restitution des canaux [Gerzon, 1975]. Notamment via le format-B qui résulte



d'un modèle de représentation du champ sonore avec une information de pression acoustique relative à la composante omnidirectionnelle et des informations directives de vitesse associées aux autres canaux [Gerzon, 1974]. De ce changement de représentation découle alors directement des traitements très intéressants, pour resserrer ou élargir la scène frontale [Gerzon, 1972a] ou encore pour effectuer la rotation du champ sonore, qui se retrouve d'ailleurs dans la plupart des bibliothèques logicielles ambisoniques [Wakefield, 2006], [Kronlachner, 2013], [Kronlachner, 2014b]. J. Daniel évoque à ce sujet que le format B « [...] se prête très naturellement à des manipulations du champ, et en particulier des rotations » [Daniel, 2001]. Suite à cela, généraliser le cadre formel de l'ambisonie, en étendant la représentation du champ sonore du format B au domaine des harmoniques sphériques, permet d'ouvrir et d'envisager de nombreuses perspectives. Beaucoup de ces traitements de l'espace ont déjà abondamment été traités, tels la rotation mais aussi l'effet de miroir [Chapman & al., 2009], la déformation du champ sonore dit *warping* [Pomperger & al., 2011] ou encore le filtrage spatial [Kronlachner, 2014b]. Ces opérations permettent de générer des transformations du champ sonore reposant sur des jeux de pondérations, d'additions et de multiplications entre les signaux associés aux harmoniques sphériques. Ces opérations sont généralement représentées sous une forme matricielle. Ces formules découlent des propriétés géométriques et mathématiques de la décomposition en harmoniques sphériques mais aussi, comme brièvement suggéré par M. Kronlachner, sur des approches moins conventionnelles et plus artistiques telles que l'utilisation de matrices inspirées des paraboloides utilisés par I. Xenakis et Le Corbusier pour le Pavillon Philips de l'exposition universelle de 1958 à Bruxelles.

Ce dernier exemple d'utilisation d'une matrice saugrenue, renvoie à la deuxième approche. Il s'agit de se libérer totalement des contraintes et du cadre formel défini par le domaine utilisé, de laisser libre cours aux expérimentations et de mettre en œuvre de manière totalement directe des opérations, au sens algorithmique du terme, d'un autre domaine vers celui-ci. L'objectif final n'est plus d'obtenir une transformation similaire via une opération différente, mais de faire émerger de nouvelles transformations via une opération identique dans le nouveau domaine. Ce type d'approche est par nature peu scientifique et peut *a priori* se révéler très aléatoire voir improductif. En effet, en reprenant l'un des exemples présentés précédemment, il peut être constaté qu'un signal temporel multiplié par un autre signal temporel, aboutit à une modulation d'amplitude ou une modulation en anneau selon la nature des signaux. Transposer directement l'opération de multiplication dans le domaine fréquentiel résulte, bien sûr, en une toute autre distorsion du son, dont les principes et fonctionnements sont difficilement intelligibles à première vue. Mais, au risque d'énoncer un truisme, cette approche qui peut paraître tout à fait

impertinente et aléatoire, est à l'origine de bon nombre de pratiques des musiciens, compositeurs et, de manière encore plus générale, des personnes travaillant le son de façon artistique. Où il s'agit d'écouter comment un traitement « sonne » et d'expérimenter pour s'approprier ses paramètres de jeux et pouvoir le contrôler. Ce type de pratique n'est pas sans rappeler la création de la synthèse par modulation de fréquence par J. Chowning [Chowning & al., 1986]. Et c'est ainsi, sous l'angle du chercheur-musicien, que les expérimentations ont été conduites. D'où le nom du premier projet HOA « la spatialisation du son pour le musicien, par le musicien » qui résonne en écho à l'ouvrage de J. Chowning.

La méthodologie d'expérimentation n'est donc pas à proprement parler exclusivement empirique. Elle consiste à mettre en œuvre des traitements audionumériques, la plupart du temps familiers aux compositeurs et musiciens. De bien les installer dans les pratiques et les usages du domaine des harmoniques sphériques puis d'analyser les résultats sonores et les mettre en relation avec le modèle de représentation de l'espace. Cette méthodologie permet de comprendre comment le modèle interagit avec un type de traitement spécifique et de pouvoir en tirer des éléments pertinents sur un plan à la fois spatial et sonore. L'objectif principal de cette approche est de commencer à concevoir le domaine des harmoniques et les opérations sur les signaux qui leurs sont associés, comme un espace composable. C'est à dire de pouvoir envisager quels sont les contraintes et les fondements qui régissent le système, afin de faire émerger un ou même plusieurs cadres propices à la création. Et aussi, pour créer des paramètres de contrôle pertinents sur un plan artistique et musical. L'approche vise donc à obtenir des résultats fondés sur une approche scientifique et logique, et dont la méthodologie puisse potentiellement être généralisée. Mais l'approche en elle-même a été établie en suivant une méthodologie fortement basée sur l'expérience, avant d'être formalisée. Enfin, il va de soi que les travaux présentés dans le cadre de ce document ne suivent pas l'ordre chronologique des expérimentations, et font d'ailleurs abstraction de bon nombre d'entre-elles. En effet, comme indiqué précédemment, beaucoup d'approches peuvent bien sûr être envisagées, surtout dans un contexte d'expérimentations musicales libérées des questions de vraisemblance acoustique. Et où le dessein est justement de créer des espaces sonores les plus originaux possibles, tout en adoptant une démarche éclectique. Cependant, l'idée est de tenter de les définir afin de mettre en exergue les résultats obtenus et les conclusions qui en découlent.

## 4.1. Variation de la résolution angulaire

Comme suggéré dans le chapitre précédent, une des premières opérations qui peut être envisagée sur le champ sonore est de jouer sur sa résolution angulaire. Cette opération de l'espace consiste à faire varier virtuellement l'ordre de décomposition. De nombreuses approches peuvent être proposées afin d'obtenir cet effet. Par exemple, en considérant un ordre de décomposition  $N$  donné, une approche somme toute très rudimentaire peut envisager de simplement annuler les contributions du dernier degré pour simuler un ordre de décomposition  $N-1$ . Force est de constater qu'une telle approche n'est que peu pertinente et n'offre aucun potentiel musical mais néanmoins l'idée est présente. Il s'agit donc de définir les spécifications fonctionnelles de cette approche pour pouvoir réaliser des choix en fonction des connaissances ambisoniques, et notamment du modèle des harmoniques sphériques pour, par la suite, proposer une approche pratique répondant au problème.

### 4.1.1. Spécifications

La mise en œuvre d'un traitement permettant de faire varier la résolution angulaire d'un champ sonore, nécessite de définir un certain nombre de spécifications élémentaires afin de diriger les recherches. Le premier objectif qu'il semble nécessaire de définir est la plage de « résolution » que cette opération doit pouvoir offrir. Cela paraît évident, mais il est nécessaire de rappeler que pour un espace sonore ambisonique donné, les changements de résolution angulaire sont limités par deux extrêmes associés d'une part à un champ sonore omnidirectionnel, et d'autre part à un champ sonore plus directionnel, en adéquation avec l'ordre de décomposition. Cela implique que la plage de variations possibles de la résolution angulaire d'un champ sonore est d'autant plus importante que l'ordre de décomposition augmente. Il est en effet compréhensible que nul ne souhaite limiter les possibilités à une sous-plage de résolution, telle que l'espace sonore ne puisse varier que d'omnidirectionnel à peu directionnel ou inversement faiblement directionnel à très directionnel. En somme, il est souhaitable de pouvoir circuler sur l'ensemble des possibilités de résolution angulaire qu'un espace sonore peut offrir, et de tirer parti de toutes ses contributions. Partant de cette exigence, il est aussi envisageable de pouvoir offrir une variation linéaire de la résolution angulaire du champ sonore. L'enjeu est bien entendu d'éviter des phénomènes d'escalier, afin de pouvoir glisser d'une directionnalité à une autre. Mais une caractéristique importante de cette spécification est aussi qu'au centre de la plage de valeurs de résolution angulaire jouables, l'espace sonore paraisse moitié moins « précis ». Cette notion est difficilement quantifiable. L'idée est d'éviter d'avoir un phénomène de variation non homogène tel qu'une large plage de valeurs soit dédiée à la variation entre une figure hypercardioïde et

cardioïde mais qu'une petite plage de valeurs soit dédiée à la variation entre une figure cardioïde et omnidirectionnelle. Une autre spécification est de préserver l'énergie du champ sonore. Quelle que soit l'approche employée, l'idée est pour une source sonore de seulement percevoir une variation de sa résolution qui ne soit pas accompagnée d'une augmentation ou d'une diminution de son intensité générale, qui pourrait être perçue comme un *effet de bord*. Ceci notamment afin de pouvoir l'utiliser conjointement avec d'autres traitements, sans avoir à prendre en compte et compenser cet artefact. Enfin, comme suggéré dans le chapitre précédent, une bonne approche en terme de jouabilité de la variation de la résolution angulaire, est de définir un paramètre dont les limites ne varient pas en fonction de la plage de résolution angulaire jouable. Il s'agit en somme d'avoir, en tant qu'utilisateur, une plage de valeurs jouables prédéfinie tel que puisse l'être un pourcentage, afin de faciliter l'adaptation des traitements à différents ordres de décomposition.

#### 4.1.2. Choix

Afin de satisfaire la première spécification ainsi que la dernière, il est défini de manière arbitraire qu'au lieu de considérer un facteur qui est un nombre compris entre 0 et l'ordre de décomposition effectif, les approches utilisent un facteur compris entre 0 et 1. Une valeur de 0 équivalent à l'ordre de décomposition 0 soit un champ sonore omnidirectionnel et un facteur de 1 équivalent à l'ordre de décomposition effectif, ou à l'ordre de décomposition maximal, soit la résolution angulaire maximale.

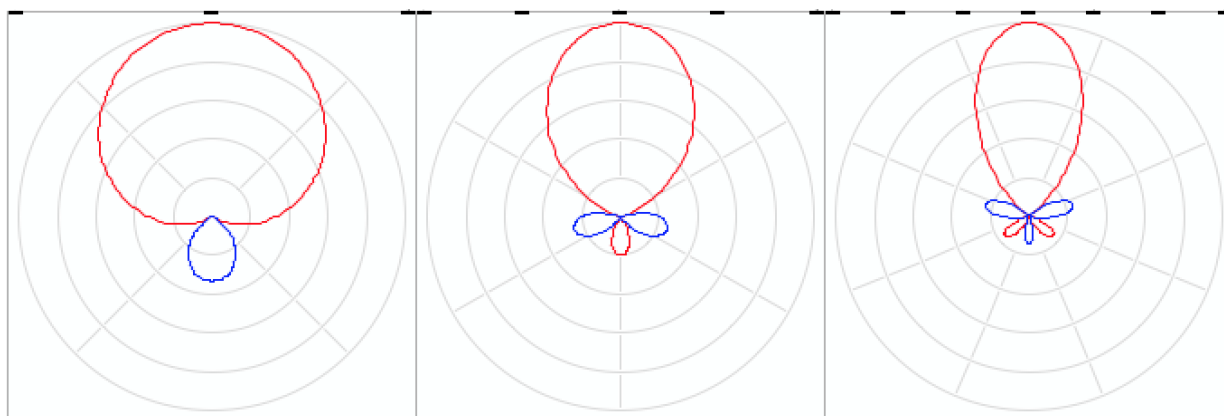


Figure 4.1. Représentation de l'augmentation de la résolution angulaire selon l'ordre de décomposition. A gauche, l'espace est discrétisé par quatre points espacés de 90 degrés correspondant au premier ordre de décomposition. Au centre, l'espace est discrétisé par six points espacés de 60 degrés correspondant au deuxième ordre de décomposition. A droite, l'espace est discrétisé par huit points espacés de 45 degrés correspondant au troisième ordre de décomposition.

Afin de satisfaire la deuxième spécification, il est défini qu'entre 0 et 1, les résolutions angulaires varient linéairement entre les deux extrêmes. La complexité de la mise en œuvre vient alors du fait que tous les harmoniques ne contribuent pas également à la précision de la décomposition de l'espace. Une manière simple de comprendre ce problème consiste à utiliser la projection des harmoniques et la relation entre le nombre associé de points permettant de discrétiser l'espace et l'ordre de décomposition. En prenant une représentation bidimensionnelle de l'espace via un cercle, ce dernier peut être discrétisé par 4 points espacés de 90 degrés pouvant correspondre au premier ordre de décomposition. En ajoutant deux points à la discrétisation du cercle, l'écart angulaire entre les points est de 60 degrés, la précision de la représentation augmente de 30 degrés, alors que l'ordre de décomposition optimal correspondant est 2. Enfin en ajoutant encore deux points à la discrétisation du cercle, l'écart angulaire est de 45 degrés, la précision de la représentation augmente d'environ 15 degrés, alors que l'ordre de décomposition optimal correspondant est 3 [Figure 4.1]. En partant de l'ordre de décomposition 0 en passant par le premier et second pour atteindre le troisième ordre de décomposition, les écarts angulaires sont de 360, 90, 60 et 45 degrés. Aussi l'écart angulaire associé à la résolution angulaire augmente de manière non linéaire. Ainsi, les contributions des signaux associés aux premiers harmoniques sphériques ou circulaires de degrés 0 et 1 sont essentielles car elles construisent la caractéristique directionnelle du son. Tandis que les contributions des signaux associés aux harmoniques de degrés élevés permettent d'améliorer la résolution angulaire du champ sonore de manière inversement proportionnelle à leurs degrés. Les harmoniques de degré 7 contribuent moins à la précision de la décomposition de l'espace que les harmoniques de degré 3 qui, de même, contribuent moins que les harmoniques de degré 1.

L'algorithme proposé se doit alors de prendre en compte ces caractéristiques des contributions des signaux, en fonction des degrés des harmoniques auxquels ils sont associés. Le facteur d'entrée doit permettre de faire varier les contributions des harmoniques, en considérant leurs degrés sur une échelle non linéaire. Un facteur d'un demi ne correspond pas à l'ordre de décomposition maximal divisé par deux mais à un ordre de décomposition sensiblement moins élevé. L'idée est que les haut-parleurs, séparés de plus d'environ 90° de l'angle d'incidence de la source sonore, contribuent déjà de manière relativement conséquente à la restitution de la source sonore. De même, à moindre mesure, pour les haut-parleurs diamétralement opposés à la direction de la source. En prenant le cas concret d'un ordre de décomposition maximal de 6, sans prendre en compte cette échelle non linéaire de contribution des harmoniques à la restitution du champs sonore, l'ordre de décomposition pour un facteur d'un demi serait de 3.

Or un ordre de décomposition de 3 n'offrirait pas l'impression d'un semi recouvrement de l'espace par la source sonore. Les haut-parleurs opposés contribuant d'ailleurs de manière sensiblement similaire à un ordre de décomposition 3 ou 4 pour une source sonore directionnelle [Figure 4.2]. Il faut aussi prendre en compte l'échelle non linéaire, qui amène à définir un ordre de décomposition virtuel bien plus proche du premier ordre de décomposition ou même de l'ordre de décomposition zéro. Cette échelle donne alors une part importante au premier harmonique omnidirectionnel et offre alors le recouvrement escompté de l'espace par la source sonore.

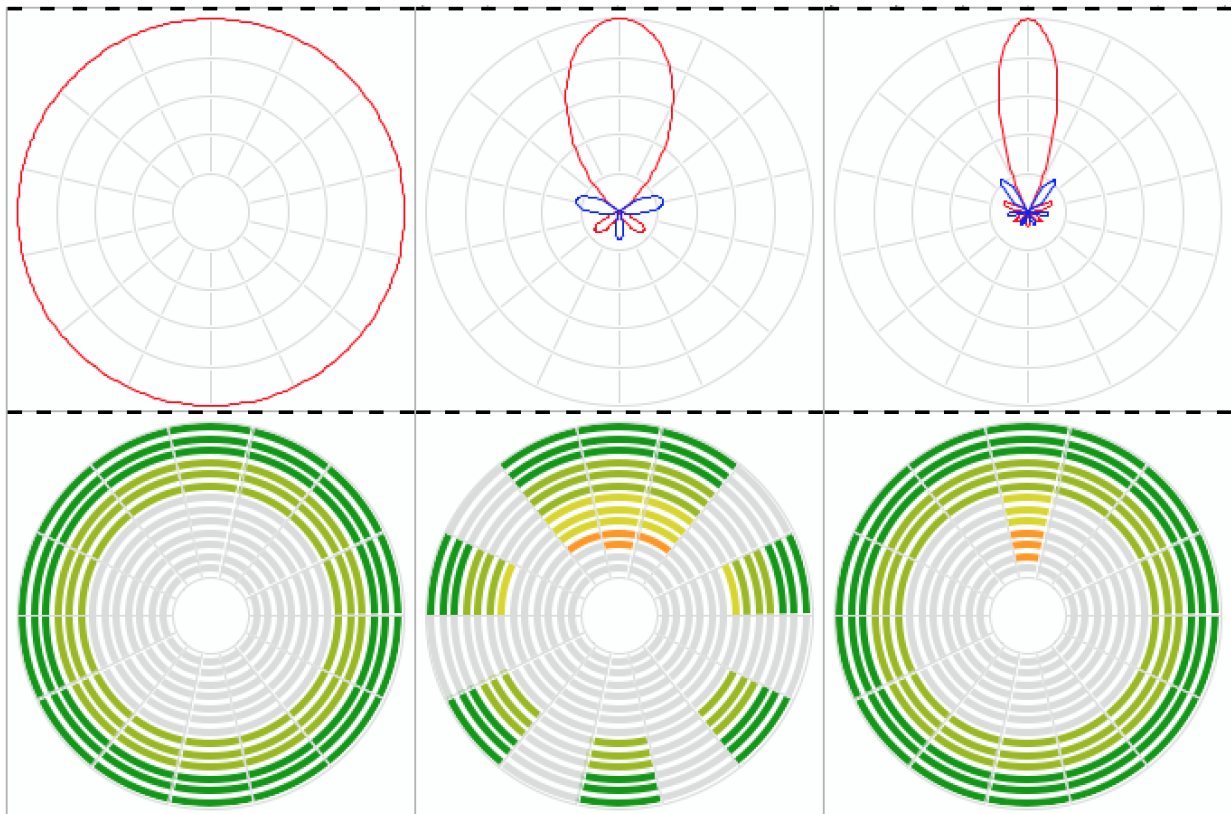


Figure 4.2. Représentation de la résolution angulaire d'un champ sonore pour les ordres de décomposition 0 à gauche, 3 au centre et 6 à droite avec en dessous les contributions associées sur système de seize haut-parleurs régulièrement répartis sur un cercle.

Enfin, pour satisfaire la troisième spécification nécessitant de préserver l'énergie du champ sonore, il faut considérer les contributions des signaux de chaque degré en terme d'énergie dans le champ sonore. Une des propriétés importantes des harmoniques sphériques est que pour une direction donnée, la somme du carré des harmoniques d'un degré soit toujours égale à 1 tel que

$$E_l = \sum_{m=-l}^l Y_{l,m}^2 = 1.$$

Équation 4.1. Loi de l'énergie d'un degré d'harmoniques sphériques.

Cette propriété implique que l'énergie – au sens défini par M. A. Gerzon – de l'ensemble des contributions associées à un degré suite à un encodage directionnel, soient normalisées à 1<sup>164</sup>. Cela implique aussi que, suite à l'encodage d'un signal unitaire de valeur 1, l'énergie globale du champ sonore soit toujours égale au nombre de degrés utilisés pour la décomposition du champ sonore tel que

$$E = \sum_{l=0}^N E_l = \sum_{l=0}^N \sum_{m=-l}^l Y_{l,m}^2 = N + 1.$$

Équation 4.2. Loi de l'énergie de la décomposition en harmoniques sphériques.

Aussi, l'objectif consiste à préserver l'énergie globale du champ sonore pour toute valeur de variation de la résolution angulaire selon cette dernière loi. Pour toute annulation des contributions d'un degré spécifique, il est nécessaire d'augmenter les contributions des autres degrés en considérant ces contributions définies par [Équation 4.1].

### 4.1.3. Mise en œuvre

Plusieurs approches ont été établies au cours des développements de la bibliothèque HOA. Le cadre de cette thèse ne permet pas de revenir en détail sur l'ensemble de celles-ci, mais il est possible de présenter brièvement un aperçu des expérimentations, des problèmes rencontrés et des solutions proposées pour arriver au résultat actuel.

La première mise en œuvre est fondée de manière totalement empirique. L'idée centrale était d'offrir une approche où les contributions associées à chaque degré diminuaient les unes après les autres de telle sorte que les variations de la résolution angulaire passent par les équivalents des ordres de décomposition entiers intermédiaires. En considérant un ordre de décomposition

---

<sup>164</sup> Il est important de noter qu'en raison d'une erreur de normalisation, cette loi n'est pas valable sur les signaux associés aux harmoniques sphériques dans la version 2.2 de la mise en œuvre de la bibliothèque HOA. Néanmoins, il est possible d'utiliser un correctif que j'ai publié récemment sous le nom de *Patch SN3D* disponible sur le site du projet HOA et sur le répertoire Github du projet.

$N$ , certains états doivent correspondre exactement aux ordres de décomposition  $N-1$ ,  $N-2$ , ...,  $1$ ,  $0$ . Par exemple pour l'ordre de décomposition 3, avant d'arriver à l'ordre de décomposition virtuel 0, l'approche doit simuler les ordres de décomposition virtuels entiers 2 puis 1. Et entre deux ordres de décomposition virtuels entiers afin d'offrir une variation continue de la résolution angulaire, le système offre l'ensemble des ordres de décomposition virtuel fractionnaires, tels que par exemple les ordres de décomposition virtuels 2,1 ou 2,7 entre 2 et 3.

Plusieurs problèmes ont été rencontrés lors de la mise en œuvre cette première approche, générant notamment de nombreux artefacts plus ou moins importants, notamment dus à la création d'un algorithme s'adaptant à n'importe quel ordre de décomposition. Il a été relativement difficile de mettre en place une échelle cohérente où les contributions du signal associé au premier harmonique omnidirectionnel ne prennent pas trop rapidement d'importance. La compensation de la perte d'énergie du champ sonore n'était pas linéaire et de ce fait des variations apparaissaient selon les variations de la résolution angulaire. Et selon l'ordre de décomposition réel maximal, l'énergie du champ sonore omnidirectionnel pouvait être totalement en dessous de son énergie initiale. La simulation d'ordres de décomposition entiers et fractionnaires tels que présentés, génère des figures de directivité ne variant pas non plus de manière linéaire. Aussi, les lobes secondaires apparaissent et disparaissent, amenant une oscillation des contributions des haut-parleurs qui ne sont pas dans la direction de la source sonore. Cet artefact est en réalité prévisible au regard des figures de directivité associées à différents ordres de décomposition. Il est en effet possible de remarquer que la figure directive associée à un ordre de décomposition 3 augmente sensiblement les contributions dans des directions différentes de la source, par rapport à l'ordre de décomposition 6. Enfin un dernier problème de cette approche, mais non des moindres, réside dans la complexité de l'algorithme et dans sa grande demande de calculs du processeur, rendant parfois problématique l'utilisation de nombreux processus en parallèle<sup>165</sup>.

---

<sup>165</sup> Des moyens ont pu être trouvés pour l'implémentation C++ de l'algorithme au détriment de l'allocation de la mémoire. En effet, l'algorithme nécessitait d'allouer de larges matrices où, pour chaque harmonique, les gains étaient pré-calculés pour un grand nombre de valeurs prédéfinies, afin d'éviter des interpolations. Néanmoins cet artifice implique d'autres problèmes dus à l'allocation de larges matrices de mémoire.



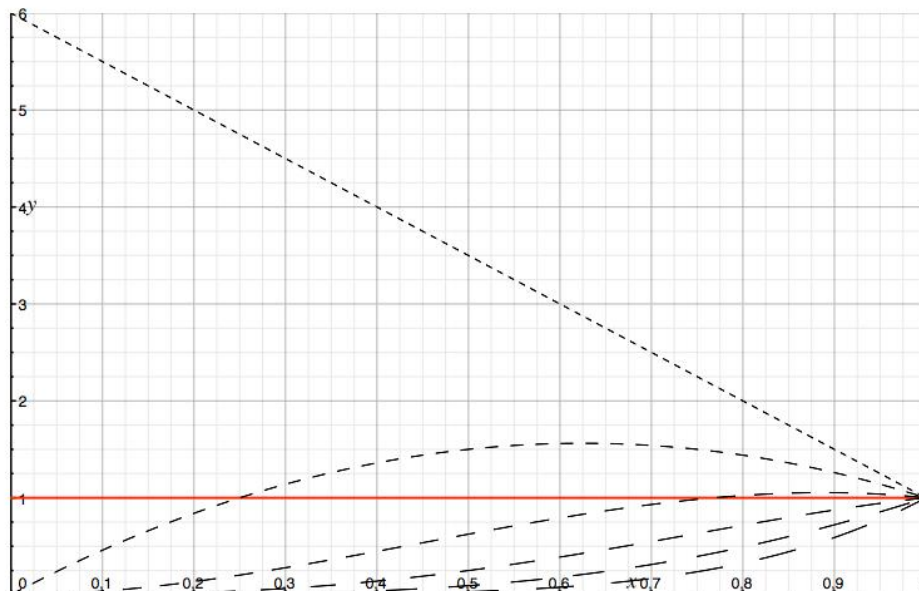


Figure 4.3. Représentation des gains des signaux associés aux harmoniques sphériques lors de la variation de la résolution angulaire pour un ordre de décomposition 5. Les courbes en tirets représentent les gains des signaux associés respectivement de haut en bas (ou des petits tirets au grands tirets) aux degrés de 0 à 5. La ligne rouge représente l'énergie globale du champ sonore pondérée par le nombre de degrés.

Plusieurs propositions ont été réalisées par la suite. L'idée était alors de s'inspirer des optimisations *max-r<sub>e</sub>* et *in-phase* proposées respectivement par M. A. Gerzon et D. G. Malham, qui diminuent conjointement les contributions des signaux associées à plusieurs degrés. Ces optimisations avaient pour but de concentrer l'énergie dans la direction de la source sonore, impliquant réciproquement une diminution de la résolution angulaire. Il s'agit en quelque sorte d'une simulation de multiples ordres de décomposition fractionnaires. Par exemple, l'optimisation *max-r<sub>e</sub>* peut être vue pour le deuxième ordre de décomposition comme la combinaison des deux ordres fractionnaires 0,755 et 1,4. Une approche idéale aurait d'ailleurs pu offrir des résolutions angulaires équivalentes pour des valeurs spécifiques à ces deux optimisations. Cependant, aucun système général permettant de regrouper ces deux approches n'a été trouvé. Mais différentes tentatives ont permis de mettre en œuvre un système toujours non-idéal diminuant cependant énormément les artefacts de la première proposition. Cette approche est celle proposée dans la version actuelle de la bibliothèque HOA. Néanmoins de plus récentes recherches ont permis de définir une dernière démarche répondant aux spécifications. Cet algorithme ne se trouve pour le moment que sur la dernière version de développement de la bibliothèque HOA, mais il a déjà pu être testé et validé<sup>166</sup>. Il consiste à

<sup>166</sup> L'approche est déjà disponible dans la version C++ en ligne de la bibliothèque ainsi que sous forme d'abstraction destinée à être utilisée avec l'objet *hoa.2/3d.process~*.

définir un gain pour chaque signal associé à un harmonique, selon le degré  $l$  de celui-ci, de l'ordre de décomposition  $N$  et en fonction du facteur de résolution angulaire  $x$  tel que

$$G_l(x) = x^l(1 + (1 - x)(N - l)).$$

Équation 4.3. Gains des signaux associés aux harmoniques sphérique pour simuler une variation de la résolution angulaire du champ sonore avec  $N$  l'ordre de décomposition,  $l$  le degré associé aux harmoniques sphériques compris entre 0 à  $N$  et  $x$  le facteur de résolution angulaire compris entre 0 et 1 et

Cette approche permet de complètement corriger la variation globale de l'énergie du champ sonore présentée précédemment, tout en utilisant une approche similaire aux optimisations de champ sonore en modifiant conjointement les contributions de l'ensemble des degrés [Figure 4.3]. Aussi les calculs nécessaires sont bien plus simples et bien moins coûteux en termes d'utilisation du processeur que les propositions faites jusqu'alors. Enfin, les variations de la résolution angulaire semblent relativement linéaires avec le facteur de contrôle. Il est cependant envisagé pour de futurs travaux d'analyser ces résultats au regard des vecteurs énergies et vitesses. Leurs normes étant représentatives de la résolution angulaire du champ sonore, si pour tout ordre de décomposition donné et un système de restitution optimal associé, les normes des vecteurs sont égales au facteur de résolution angulaire alors la proposition devrait être totalement validée.

La mise en œuvre de cette opération du champ sonore a permis de mettre en avant certaines spécificités de l'approche ambisonique présentée dans cette thèse. Il s'agit d'une part, d'offrir un traitement dont la « qualité » dépend de l'ordre de décomposition, mais qui s'adapte aussi à cet ordre de décomposition pour offrir un effet similaire, quel que soit celui-ci. Cette approche relativement classique est intéressante. Elle permet une première appropriation des mécanismes mis en jeu, par la décomposition en harmoniques sphériques, notamment via la considération des contributions de chaque degré d'harmoniques sur la résolution angulaire de la représentation de l'espace. Cette opération de l'espace est disponible via l'objet *hoa.2/3d.wider~* ou via le patch *hoa.fx.wider~* instancié avec l'objet *hoa.2/3d.process~*. C'est aussi l'opération mise en œuvre afin de simuler le rapprochement des sources sonores ponctuelles au sein de l'objet *hoa.2/3d.map~*.

## 4.2. Synthèse par décorrélation temporelle

Une des approches des plus expressives, mais aussi une des plus simples à appréhender dans un premier temps pour le musicien, est la synthèse de champ sonore par décorrélation temporelle des signaux associés aux harmoniques sphériques. En effet, facile d'accès, le traitement sur lequel elle repose est fortement ancré dans les pratiques et relativement simple à mettre en œuvre, ainsi qu'à formuler.

### 4.2.1. Approche aléatoire

Cette opération d'encodage de champ sonore n'est donc pas construite sur une pondération directionnelle des signaux associés aux harmoniques sphériques, mais sur l'utilisation de retards dont les temps diffèrent, afin de créer des relations complexes. Mais aussi afin de faire émerger un champ sonore où le son ne semble pas provenir d'une direction précise de l'espace et qui offre un caractère enveloppant à l'auditoire.

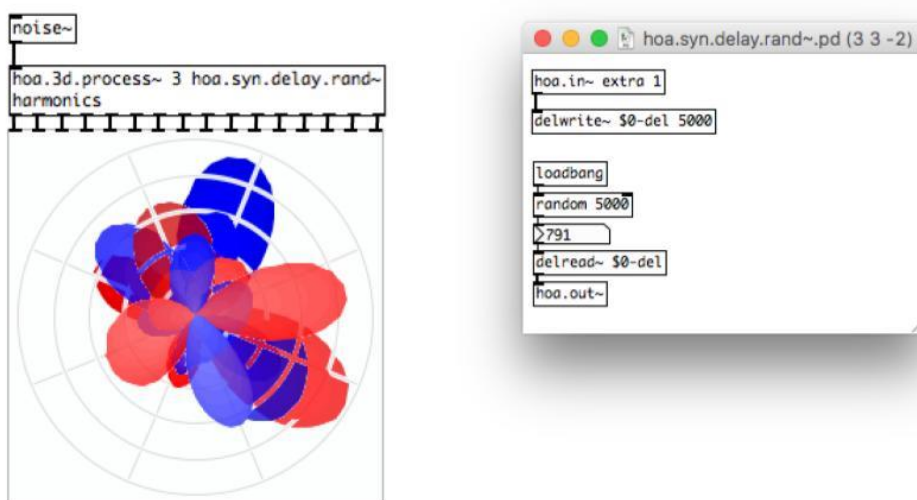


Figure 4.4. Patch Pure Data permettant de synthétiser un champ sonore où les signaux associés aux harmoniques sphériques sont des versions retardées du signal source. Et dont les temps de retard sont définis de manière aléatoire. A gauche, le patch principal avec l'objet *hoa.3d.process~* permettant d'instancier l'abstraction *hoa.syn.delay.rand~* pour chaque signal associé à un harmonique et la représentation du champ sonore synthétisé dans le temps via la superposition du rendu de l'objet *hoa.3d.scope~* à trois instants. Les caractéristiques spatiales sont associables à un champ sonore diffus. A droite, l'abstraction *hoa.syn.delay.rand~* composée d'une ligne à retard dont la valeur de temps est générée aléatoirement à l'instanciation.

*A priori*, les temps de retard peuvent être définis de manière aléatoire. Les relations de pondération entre les différents harmoniques n'étant plus respectées dans le sens directionnel du terme, le champ sonore ainsi synthétisé devrait arriver vers de multiples directions et selon

aucune logique apparente, de manière quelque peu analogue à ce que pourrait être un champ diffus<sup>167</sup> [Figure 4.4]. Ce type d'approche, bien qu'offrant un aperçu du potentiel musical de ce procédé, est difficilement prévisible. Elle n'offre en soi aucun contrôle, que ce soit sur le plan de la morphologie sonore via le traitement même de la décorrélation temporelle, que sur le plan spatial via l'adaptation de ce traitement à l'espace. Il s'agit de définir une méthodologie permettant d'appréhender ce traitement dans l'espace. L'objectif est de comprendre comment chaque signal associé à un harmonique, contribue au champ sonore lorsqu'il est synthétisé à partir d'un module de décorrélation temporel. Pour, par la suite, tenter d'étendre l'analyse au champ sonore dans sa globalité.

#### 4.2.2. Concevoir le traitement dans l'espace

L'approche qui suit tente donc, dans un premier temps, de mettre en exergue le contrôle de la spatialité du son, pour comprendre comment cela influence aussi sur sa morphologie intrinsèque. Pour cela, il s'agit tout d'abord de considérer un système de restitution tridimensionnel en hexaèdre ou cube, comprenant donc six haut-parleurs  $LS_{front}$ ,  $LS_{left}$ ,  $LS_{back}$ ,  $LS_{right}$ ,  $LS_{top}$  et  $LS_{bottom}$  positionnés de manière régulière, respectivement à l'avant, à gauche, à l'arrière, à droite, en haut et en bas du champ sonore. Pour une telle configuration, il est d'usage de procéder à une décomposition de l'espace au premier ordre, aboutissant à quatre signaux associés aux harmoniques sphériques  $Y_{0,0}$ ,  $Y_{1,-1}$ ,  $Y_{1,0}$  et  $Y_{1,1}$ .



Figure 4.5. Synthèse des signaux associés aux harmoniques sphériques d'ordre de décomposition 1 par décorrélation temporelle selon des temps de retard croissants tels que  $t_{0,0} < t_{1,-1} < t_{1,0} < t_{1,1}$ . Les représentations via l'objet *hoa.3d.scope~* représentent les contributions du champ sonore à ces quatre temps.

<sup>167</sup> Evidemment l'effet est bien plus « surréaliste » que le champ diffus défini par les réflexions tardives d'un espace réel.

De manière arbitraire, un temps  $t$  dépendant du degré et de l'ordre azimutal des harmoniques est appliqué à chaque signal tel que  $t_{0,0} < t_{1,-1} < t_{1,0} < t_{1,1}$ . Dans ce contexte, utiliser en tant que signal source une impulsion dont la durée est négligeable au regard des différences de temps  $t$ , génère un champ sonore constitué exclusivement des contributions des signaux associés à l'harmonique  $Y_{0,0}$  à  $t_{0,0}$  puis des contributions de l'harmonique  $Y_{1,-1}$  à  $t_{1,-1}$  puis des contributions de l'harmonique  $Y_{1,0}$  à  $t_{1,0}$  et enfin des contributions de l'harmonique  $Y_{1,1}$  à  $t_{1,1}$  [Figure 4.5].

En considérant alors les représentations spatiales des harmoniques circulaires, il est aisément possible de prévoir le champ sonore généré par le système de haut-parleurs à chaque impulsion. Ainsi à  $t_{0,0}$ , le champ sonore est omnidirectionnel et l'ensemble des haut-parleurs contribue de manière identique, en générant une première fois l'impulsion du signal source. Par la suite, à  $t_{1,-1}$ , le champ peut être associé à une figure-de-huit orientée sur l'axe des abscisses et cette fois, par la figure de directivité du champ sonore, seuls les haut-parleurs  $LS_{front}$  et  $LS_{back}$  génèrent pour la deuxième fois, une impulsion du signal sonore. Enfin à  $t_{1,0}$  puis à  $t_{1,1}$  ce dernier phénomène spatial se reproduit selon les directions définies par chacun des harmoniques sphériques auxquels ces temps sont associés, et à la paire de haut-parleurs impliqués dans leurs restitutions, qui sont respectivement définies par l'axe de hauteur puis l'axe des ordonnées associées aux haut-parleurs  $LS_{top}$  et  $LS_{bottom}$  puis  $LS_{left}$  et  $LS_{right}$ <sup>168</sup>. Dans de tels cas, en négligeant les oppositions de phase entre les haut-parleurs, résultant de la polarité des lobes en figure-de-huit, l'auditeur perçoit la source sonore comme apparaissant depuis toutes les directions puis simultanément sur l'axe avant-arrière, puis haut-bas et enfin gauche-droite. Cette approche peut être facilement étendue et généralisée aux ordres de décomposition supérieurs.

### 4.2.3. Extension de l'approche aux ordres élevés

Dans l'approche qui vient d'être présentée, les temps des retards sont choisis de manière arbitraire. Cela peut s'avérer laborieux à réaliser pour des ordres de décomposition élevés, surtout s'il s'agit d'offrir un traitement dynamique qui s'adapte aux différents ordres de décomposition. L'enjeu est alors de définir le temps  $t$  de retard relatif à chaque signal, associé à un harmonique sphérique comme fonction de l'ordre de décomposition  $N$ , de degré  $l$  et de l'ordre azimutal  $m$  de l'harmonique sphérique et d'un temps global à l'ensemble des harmoniques sphériques noté  $t_g$ . Aussi, le temps  $t_{l,m}$  relatif à un harmonique peut être exprimé mathématiquement tel que

---

<sup>168</sup> Afin de visualiser les figures directives, il est ici possible de se référer au deuxième chapitre [Figure 2.2] précédent. De même que pour la suite de la présentation.

$$t_{l,m} = t_g f(N, l, m).$$

Équation 4.4 : Relation entre le  $t_{l,m}$  relatif à un harmonique et le temps global  $t_g$  selon l'ordre de décomposition, le degré et l'ordre azimutal de l'harmonique défini par une fonction  $f$ .

La nature de la relation, ici définie par la fonction  $f$  selon l'ordre de décomposition, le degré et l'ordre azimutal de chaque harmonique, permet de définir sur le plan temporel l'émergence dans le champ acoustique des caractéristiques spatiales des harmoniques circulaires. Une infinité d'approches peuvent être mises en œuvre afin de définir cette relation. Deux approches, particulièrement démonstratives de la méthodologie employée dans ces expérimentations et faisant preuve de clarté dans la manière d'appréhender l'espace, sont exposées dans ce travail.

De manière général dans ces exemples, afin de définir une homogénéité dans les usages, il a été défini que le temps  $t_g$  fait référence à la valeur maximale de retard qui peut être appliquée au signal d'un harmonique sphérique et que la valeur minimale de retard soit définie à 0 millisecondes. Aussi, pour des raisons de transparence, il est préférable d'utiliser la notation  $t_{max}$  au lieu de  $t_g$ .

La première approche vise à étendre la relation temporelle, présentée précédemment. Dans ce cas, bien que restreint au premier ordre de décomposition, il est possible de dégager que les temps de retard des signaux associés aux harmoniques sphériques augmentent avec leurs degrés. Et qu'au sein d'un degré, les temps de retard augmentent, de même avec l'ordre azimutal de telle sorte que

$$t_{0,0} < t_{1,-1} < t_{1,0} < t_{1,1} < t_{2,-2} < t_{2,-1} < \dots < t_{N,N-2} < t_{N,N-1} < t_{N,N}.$$

Cette relation est similaire à l'ordre de numérotation proposée par l'ACN. En estimant que l'approche accepte des ordres de grandeur linéaires entre les temps de retards, il est possible de définir mathématiquement le temps de retard relatif à un harmonique, selon le nombre total d'harmoniques sphériques  $H$  défini à partir de l'ordre de décomposition tel que

$$t_{l,m} = \frac{ACN}{H-1} t_{max}$$

Équation 4.5. Temps de retard associés aux harmoniques sphériques pour la synthèse de champ sonore par décorrélation temporelle, suivant la logique définie par le format ACN.

L'ACN allant de 0 à  $H-1$ , la fonction de relation permet d'assurer que le temps de retard minimal soit 0 millisecondes et le temps de retard  $t_{max}$ . Selon les équations de l'ACN définies dans le chapitre précédent [Equation chapitre 2 partie 3], il est possible de développer l'équation à la fois pour une représentation bidimensionnelle et pour une représentation tridimensionnelle telles que :

$$t_{l,m} = \frac{l}{N} + \frac{m}{2Nl} t_{max} \text{ en } 2D$$

Et

$$t_{l,m} = \frac{l \times (l + 1) + m}{(N + 1)^2 - 1} t_{max} \text{ en } 3D.$$

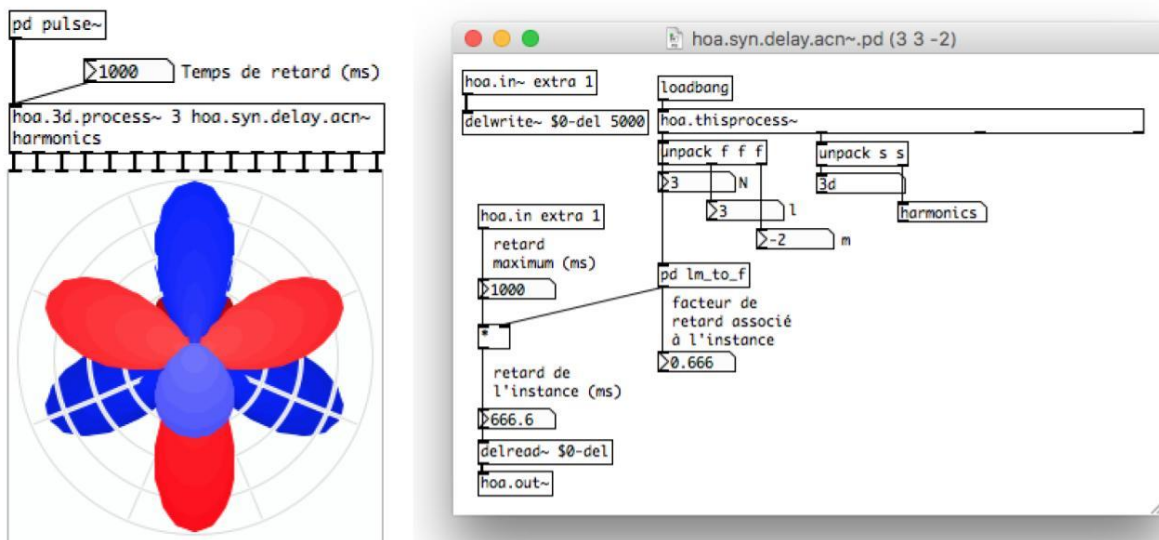


Figure 4.6. Patch Pure Data permettant de synthétiser un champ sonore où les signaux associés aux harmoniques sphériques sont des versions retardées du signal source dont les temps sont organisés selon une approche de type ACN. A gauche le patch principal avec l'objet *hoa.3d.process~* qui a instancié l'abstraction *hoa.syn.delay.acn~* et la représentation dans le temps via le superposition du rendu de l'objet *hoa.3d.scope~* lors de l'apparition de la contribution de l'harmonique  $Y_{3,2}$ . A droite, l'abstraction *hoa.syn.delay.acn~* qui calcule le temps de retard de l'instance en fonction d'un temps de retard maximal et d'un facteur associé à l'instance dépendant de l'organisation définie par l'ACN.

Cette approche, définie sous le terme de relation de type ACN, peut être mise en œuvre via une abstraction destinée à être utilisée dans l'objet *hoa.2/3d.process~*. L'opération consiste à définir une entrée spécifique pour le temps de retard maximum  $t_{max}$  et à calculer le temps de retard spécifique à l'abstraction. Cela en fonction des informations récupérées via l'objet *hoa.thisprocess~*, qui sont relatives à l'harmonique auquel l'abstraction est associée [Figure 4.6].

Selon cette relation, et toujours en assumant le signal source comme une impulsion de durée inférieure à la différence de temps minimum entre les retards, il est possible de prévoir l'émergence successive d'autant d'impulsions qu'il y a de composantes du champ sonore. Et ces impulsions sont définies dans l'espace selon les figures directives des harmoniques sphériques auxquelles elles sont associées. Aussi, en considérant une représentation tridimensionnelle, les impulsions émergent tout d'abord selon une figure omnidirectionnelle associée à l'harmonique de degré 0. Puis trois impulsions apparaissent successivement selon des figures-de-huit associées aux harmoniques de degré 1, sur un plan horizontal le long de l'axe des ordonnées pour l'harmonique  $Y_{1,-1}$  puis sur un plan vertical le long de l'axe de la hauteur pour l'harmonique  $Y_{1,0}$  et de nouveau sur un plan horizontal le long de l'axe des abscisses pour l'harmonique  $Y_{1,1}$ . Ce schéma se répète alors pour les signaux associés à chaque degré, des figures comportant de plus en plus de lobes directionnels à mesure que le degré augmente. Et à chaque degré  $l$ , les impulsions apparaissent premièrement sur le plan horizontal selon la figure associée à l'harmonique  $Y_{l,-l}$  pour se redresser progressivement jusqu'au plan vertical selon la figure associée à l'harmoniques  $Y_{l,0}$  pour redescendre progressivement vers le plan horizontal selon la figure associée à l'harmonique  $Y_{l,l}$ , avec un décalage azimutal des figures dû au changement de signe de l'ordre azimutal.

Sur un plan perceptible, le résultat d'une telle approche est forcément fonction du dispositif de haut-parleurs utilisé pour restituer le champ sonore, mais aussi de l'ordre de décomposition. Il est attendu une série d'impulsions aux caractères directionnels de plus en plus en « fin », partant d'une figure omnidirectionnelle et augmentant conjointement avec les degrés auxquels les signaux sont associés. Est attendue aussi une figure spatiale se répétant, oscillant de plus en plus lentement entre deux extrêmes associés au plan horizontal et au plan vertical. En somme, cette approche joue sur les symétries et les relations d'organisations des harmoniques sphériques, via



le nombre de lobes qu'ils possèdent et leurs directions, pour faire émerger des figures rythmiques de l'espace.

#### 4.2.4. Renversement de l'approche

Cette première approche a été établie de manière arbitraire en s'inspirant de l'organisation des signaux associés aux harmoniques sphériques définis par l'ACN, mais bien d'autres figures spatiales peuvent être envisagées. À titre d'exemple, il est possible de proposer une deuxième approche jouant sur la symétrie des harmoniques sphériques d'une toute autre manière, et en retournant le problème. Plutôt que de partir d'une relation arbitraire entre les temps des signaux associés aux harmoniques, une approche plus musicale serait de définir un motif temporel et spatial d'apparition des impulsions. L'enjeu proposé est le suivant : il s'agit de retourner quelque peu le processus spatial en partant du plan horizontal pour redresser les figures directionnelles petit à petit sur le plan vertical, et finir sur un champ sonore omnidirectionnel. L'objectif pourrait alors être de trouver un « chemin » cohérent dans l'apparition des figures directives associées aux harmoniques sphériques, afin d'obtenir le motif spatial et temporel voulu. Afin de limiter les possibilités, il est défini de manière arbitraire qu'une figure directive associée à un harmonique ne peut apparaître qu'une seule fois dans le cheminement. Une proposition peut alors être construite ainsi :

- Etant donné que les figures directives initiales doivent être coplanaires avec le plan horizontal, il est logique de commencer avec les figures horizontales associées aux harmoniques sphériques dont la valeur absolue de l'ordre azimutal est égale au degré, tel que  $|m| = l$  défini dans la partie 2 du chapitre 1.
- Etant donné que les figures directives finales doivent être coplanaires avec le plan vertical, il est logique de terminer avec les figures verticales associées aux harmoniques d'ordre azimutal.
- Selon les deux propositions précédentes et afin que les figures spatiales soient désaxées petit à petit du plan horizontal vers le plan vertical, il est nécessaire de faire apparaître successivement les figures spatiales associées aux harmoniques de degré  $|m| = l - 2$ , puis  $|m| = l - 2$ , etc. en évitant les cas où  $|m| = 0$ .

- La dernière figure spatiale devant apparaître est censée être omnidirectionnelle donc associée à l'harmonique  $Y_{0,0}$ . Il est aussi logique d'assumer que la fin du cheminement parte des degrés élevés pour finir sur ce dernier degré 0. Aussi les temps de retard des signaux associés aux harmoniques peuvent être organisés selon la relation

$$t_{N,0} < t_{N-1,0} < t_{N-2,0} < \dots < t_{2,0} < t_{1,0} < t_{0,0}$$

- Ce cheminement final part de la figure spatiale de l'harmonique de degré le plus élevé équivalent à l'ordre de décomposition  $N$  et d'ordre azimutal 0, la figure directionnelle précédente doit aussi être associée à l'harmonique de degré  $N$  mais d'ordre azimutal  $|m| = N - 1$ . Et il est défini arbitrairement qu'il doit être d'ordre azimutal positif.

Finalement, la relation temporelle des signaux associés aux harmoniques peut être représentée par

$$t_{N,-N} < t_{N-1,-(N-1)} < t_{N-2,-(N-2)} < \dots < t_{1,-1} < t_{1,1} < \dots < t_{N-2,N-2} < t_{N-1,N-1} < t_{N,N} < t_{N,N-1} < t_{N-1,N-2} < \dots < t_{2,1} < t_{2,-1} < \dots < t_{N-1,-(N-2)} < t_{N,-(N-1)} < t_{N,0} < t_{N-1,0} < \dots < t_{1,0} < t_{0,0}.$$

Pour un ordre de décomposition 3 donné et selon l'ACN, afin de faciliter la représentation du parcours sur la [Figure 2.2], le cheminement peut être défini par

$$9 \rightarrow 4 \rightarrow 1 \rightarrow 3 \rightarrow 8 \rightarrow 15 \rightarrow 14 \rightarrow 7 \rightarrow 5 \rightarrow 10 \rightarrow 11 \rightarrow 13 \rightarrow 12 \rightarrow 6 \rightarrow 2 \rightarrow 0$$

Figure 4.7. Représentation de l'ordre d'apparition des figures spatiales associées aux harmoniques via les indices ACN.

Cette approche montre bien d'une part la nécessité de concevoir dans l'espace les harmoniques sphériques, ainsi que d'en comprendre les spécificités et les relations qui les régissent. Et cela aussi bien, si ce n'est plus, dans son contexte musical, que mathématique. D'autre part, elle permet d'entrevoir un réel potentiel musical pour penser de manière originale l'espace et le son. Il s'agit de concevoir des motifs de variations de la figure de directivité du champ sonore. Evidemment, cette approche qui fait émerger successivement des figures de directivité associées aux représentations spatiales des harmoniques sphériques, ne peut être valide que si à tout instant seulement un seul signal contribue au champ sonore. Aussi, dès lors que le signal

source n'est plus une impulsion mais un signal continu, ou du moins suffisamment long au regard des différences de temps de retard assignés aux signaux associés aux harmoniques, les mécanismes en jeu se complexifient et avec eux l'analyse et l'anticipation des résultats escomptés.

#### 4.2.5. Analyse de l'approche

Les procédés de décorrélation temporelle présentés jusque-là s'appuient sur l'utilisation d'une impulsion comme signal source. Cette approche permet de faire abstraction du temps comme quelque de chose continu et d'analyser le champ sonore à des instants donnés par ces impulsions pour lesquels la représentation spatiale des harmoniques est suffisante à la compréhension des mécanismes en jeu. Cependant comme suggéré précédemment, dans la pratique, le signal original peut prendre de multiples formes. Il est alors nécessaire de ne plus concevoir le champ acoustique restitué comme une succession d'impulsions plus ou moins multidirectionnelles, mais de généraliser l'approche de la décorrélation temporelle en considérant le champ sonore sous la forme d'un ensemble de signaux continus, dont les contributions se superposent et interagissent en elles de manières complexes. Aussi pour cela, la représentation du champ sonore, via la projection à un instant donné, ne semble plus suffire à l'appréhension des mécanismes. Il faut, en outre, une autre manière de concevoir les implications des opérations dans le domaine des harmoniques sphériques.

L'encodage d'une source sonore directionnelle prend la forme d'une opération de matricage. Chaque signal associé à un harmonique sphérique correspond au signal source pondéré par l'harmonique sphérique auquel il est associé en fonction de la direction d'encodage. En projetant ces signaux associés sur une sphère via une deuxième opération de matricage, équivalent à un nouveau jeu de pondération de signaux, il est possible de représenter la figure directive du champ sonore à un instant donné. Il est possible également d'y associer le signal source – tel que le fait l'objet *hoa2/3d.scope~*. En considérant une projection simple, c'est à dire sans optimisation du champ sonore ou compensation de l'irrégularité du système de restitution, il est aisément possible d'appréhender le résultat à partir de la figure de directivité obtenue, car elle correspond directement aux gains du signal sonore d'origine, selon l'ensemble des directions. Définir par exemple comment contribue un haut-parleur, revient à trouver sur cette figure l'amplitude associée au signal d'origine dans la direction du haut-parleur. Aussi dans un tel cas, l'ambisonie peut être considérée comme une simple fonction de panoramique d'amplitude, qu'il est relativement aisé d'envisager selon la direction d'encodage. Le problème étant justement que dans le cas de la décorrélation temporelle des signaux associés aux

harmoniques sphériques, il ne s'agit plus de considérer un jeu de pondération mais de considérer une opération qui est fonction du temps, un filtre linéaire. Dès lors, une représentation de la projection à un instant donné ne semble plus convenir, ou du moins être suffisante, pour rendre part des conséquences des opérations sur le champ sonore. Aussi est-il nécessaire de trouver une autre façon d'appréhender le rendu spatial et sonore.

Une solution proposée est de changer, et même d'inverser les domaines de représentation. La représentation d'une projection des signaux associés aux harmoniques sphériques sur la sphère est réalisée de manière continue sur un plan directionnel mais à un instant donné. C'est à dire que cette représentation est tel un instantané des amplitudes dans l'ensemble des directions de la sphère composant le champ sonore. C'est seulement la succession de ces représentations qui peut offrir une idée du mouvement et des variations qui se déroulent dans le champ sonore. Cela explique aussi comment cette représentation est fonction des directions et que l'encodage directionnel est aussi fonction de la direction. Cela amène à définir que la synthèse par décorrélation temporelle- soit un encodage « temporel » -, nécessite une représentation qui est fonction du temps, quitte à ne considérer qu'une direction pour mieux concevoir les mécanismes en jeu.

Une méthode qui semble dès lors efficace, consiste à définir mathématiquement le résultat dans une direction de l'opération de synthèse des signaux associés aux harmoniques, par décorrélation temporelle. Le signal associé à un harmonique sphérique  $S_{l,m}$  peut être exprimé à un échantillon  $n$  donné selon le signal d'origine  $S$  retardé et un temps  $t_{l,m}$  spécifique à ce signal défini aussi en échantillons tel que

$$S_{l,m}[n] = S[n + t_{l,m}].$$

En utilisant l'équation de la projection définie en [Équation 2.11], le signal associé à une direction donnée du champ sonore peut être exprimée mathématiquement, en remplaçant les amplitudes associées aux harmoniques par l'expression du signal  $S_{l,m}$  tel que

$$y_{\theta,\phi}[n] = \sum_{l=0}^N \sum_{m=-l}^l (2l+1)Y_{l,m}(\theta, \phi)S[n + t_{l,m}].$$

Équation 4.6. Signal dans une direction donnée du champ sonore synthétisé par décorrélation temporelle du signal source.

Afin de mieux envisager les opérations en jeu, cette expression peut être développée en considérant simplement un ensemble de coefficients  $a$  qui sont fonctions de la direction tel que :

$$y[n] = a_{0,0}S[n + t_{0,0}] + a_{1,-1}S[n + t_{1,-1}] + a_{1,0}S[n + t_{1,0}] + \dots + a_{N,N}S[n + t_{N,N}].$$

Équation 4.7 : Développement de l'Équation 4.6] sous la forme d'une série de filtre en peigne.

Cette équation démontre que le signal associé à une direction donnée peut être considéré comme la somme pondérée du signal encodé, passant au travers d'une série de filtres en peigne. Les coefficients de pondérations  $a$  étant les sorties des harmoniques normalisés dans la direction. L'enjeu est alors de concevoir les valeurs de coefficients selon la formulation des harmoniques sphériques.

Un exemple peut alors consister à utiliser un système bidimensionnel de quatre haut-parleurs  $LS_{front}$ ,  $LS_{left}$ ,  $LS_{back}$  et  $LS_{right}$  placés de manière régulière respectivement à l'avant, à gauche, à l'arrière et à droite du champ sonore et d'une décomposition du champ sonore à l'ordre  $l$  avec donc trois harmoniques circulaires. Chaque signal associé à un harmonique est une version retardée du signal  $S$  d'origine tel que le retard du signal associé à l'harmonique de degré 0 est nul, le retard du signal associé à l'harmonique d'ordre azimutal  $-l$  est de 10 échantillons tout comme le retard signal associé à l'harmonique d'ordre azimutal  $l$  [Figure 4.8]. Il est possible d'exprimer les signaux associés aux haut-parleurs par

$$S_{front}[n] = S[n] + S[n + 10],$$

$$S_{left}[n] = S[n] + S[n + 10],$$

$$S_{back}[n] = S[n] - S[n + 10],$$

$$S_{right}[n] = S[n] - S[n + 10].$$

Il est possible d'étendre l'approche à une représentation tridimensionnelle similaire au premier exemple de décorrélation temporel et de définir que :

$$S_{top}[n] = S[n] + S[n + 10],$$

$$S_{bottom}[n] = S[n] - S[n + 10].$$

Équation 4.8. Signaux associés aux haut-parleurs pour un système de six haut-parleurs en cube pour un champ sonore décomposé à l'ordre 1 et synthétiser par décorrélation temporel du signal d'origine.

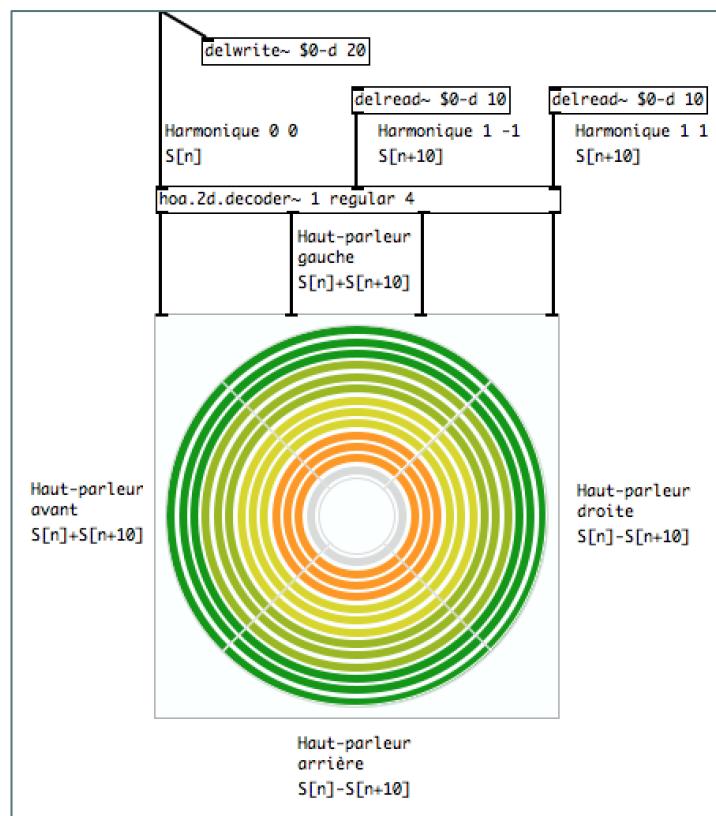


Figure 4.8. Patch Pure Data représentant les filtrages au niveau de quatre haut-parleurs régulièrement répartis sur un cercle obtenu suite à la projection des signaux associés aux trois premiers harmoniques circulaires et dont les deux associés au degré 1 sont retardés de 10 échantillons.

A l'ordre de décomposition 1, la synthèse par décorrélation temporelle peut être exprimée, dans certaines directions spécifiques, comme l'expression d'un filtre en peigne simple dont les temps de retard sont ceux définis aux signaux associés aux harmoniques sphériques. Et dont les coefficients sont définis par les amplitudes des harmoniques dans ces directions spécifiques. En augmentant l'ordre de décomposition, cette approche aboutit à un système complexe de multiples filtres en peigne en parallèle, dont les coefficients sont définis en fonction des caractéristiques spatiales des harmoniques dont dépendent les signaux, mais qui sont toujours exprimables sous une forme mathématique [Équation 4.7].

L'analyse mathématique de la projection dans une direction donnée de ce procédé de synthèse par décorrélation temporelle du signal source, permet dès lors de rattacher cette approche à certaines techniques pour la synthèse des réverbérations tardives, qui sont généralement associées à un champ diffus [Figure 4.9]. Il est à cet effet intéressant de noter la correspondance naturelle réalisée par les approches purement sensibles qui dénomment l'espace sonore de la même manière : « champ diffus ». En effet sur le plan perceptif, bien que le ressenti soit subjectif et plus difficile à définir, le résultat est un champ sonore enveloppant l'auditeur de manière quelque peu chaotique, et offrant une impression d'élargissement de l'espace. Finalement, l'utilisation de la représentation intermédiaire du champ sonore sous des signaux associés aux harmoniques sphériques a permis de concevoir un traitement de décorrélation temporelle, qui agit à la fois sur la morphologie du son mais aussi sur sa spatialité.

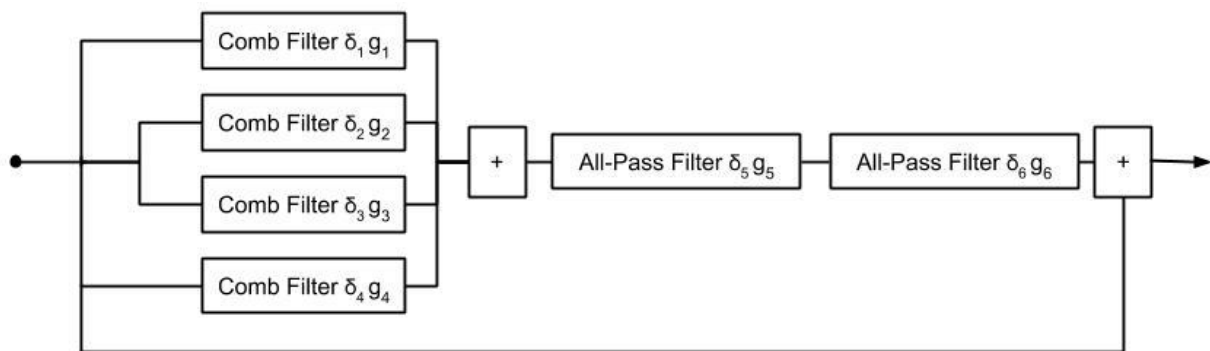


Figure 4.9. Module de réverbération artificielle présentée par M. R. Schroeder [Schroeder, 1962] où le signal passe notamment au travers d'un ensemble de filtres en peigne en série.

Et en un sens, cette représentation intermédiaire du champ sonore a été appréhendée en se libérant petit à petit des implications techniques et matérielles, et en considérant le champ sonore de manière quelque peu abstraite. L'idée du champ sonore concret est en réalité absente et ne revient qu'à la fin de l'analyse pour appréhender le rendu réel, ou comme cela sera présenté par la suite, lors de la restitution. Bien que ces deux approches puissent être tout à fait équivalentes, ce sont les représentations spatiales des harmoniques qui ont amené à penser et concevoir ce traitement. En somme, l'un des éléments important, dans cette méthodologie, est de considérer la décomposition en harmoniques sphériques comme un outil structurant la pensée musicale de l'espace.

Enfin, deux approches différentes ont été proposées pour organiser temporellement les retards des signaux associés aux harmoniques sphériques. L'une fondée sur l'organisation des signaux ambisoniques définis par l'ACN est issue d'un choix relativement direct et pragmatique visant une formulation mathématique simple, pour prendre en main le système. L'autre est fondée sur une organisation plus originale suivant des considérations spatiales en prenant en compte les figures de directivité associées aux harmoniques. Il est donc intéressant de mettre en œuvre d'autres opérations dans le domaine des harmoniques sphériques, pour discerner comment leurs représentations spatiales sont appréhendées, en relation avec les mécanismes en jeu.

### 4.3. Synthèse par granulation

L'une des premiers traitements de l'espace et du son proposé dans la bibliothèque HOA a été construit autour des techniques dites de granulation. L'opération mise en œuvre est alors relativement identique à la synthèse de champ sonore par décorrélation des signaux associés aux harmoniques sphériques. Il s'agit, en effet, de synthétiser un champ sonore en utilisant un module de synthèse granulaire quasi-synchrone (QSGS) [Roads, 1996], [Roads, 2001] pour générer les signaux associés aux harmoniques sphériques à partir d'un signal source.

#### 4.3.1. Synthèse granulaire quasi-synchrone

La QSGS génère des grains à intervalles plus ou moins réguliers, en utilisant une ligne à retard avec réinjection et une enveloppe d'amplitude. Une telle approche offre les avantages d'être peu coûteuse en terme de calculs du processeur et d'être applicable en temps-réel<sup>169</sup>. La mise en œuvre de cette technique de granulation peut donner lieu à de nombreux paramètres de contrôle. Cependant, afin d'offrir à l'utilisateur un outil simple à prendre en main, le module de synthèse présenté restreint ces paramètres au nombre de quatre : la taille moyenne des grains en millisecondes, le temps de retard en millisecondes, le pourcentage de réinjection des grains et le pourcentage de raréfaction des grains<sup>170</sup> [Figure 4.10]. Ces différents paramètres sont variables en temps-réel et permettent d'obtenir de nombreuses variations liées à la « masse » sonore et au timbre. Aussi, afin de créer des variations inhérentes au système, les valeurs de la

---

<sup>169</sup> Les échantillons en entrée sont directement traités et apparaissent transformés en sortie. D'autres approches de la synthèse granulaire « piochent » des grains dans une mémoire tampon préalablement rempli.

<sup>170</sup> À titre indicatif, un cinquième paramètre essentiel est la fonction générant l'enveloppe d'amplitude. Celle-ci peut, théoriquement, prendre n'importe quelles formes. Les usages favorisent des fonctions de type gaussienne, sinus, et sinus cardinal offrant des rendus sonores « lisses », comme le propose C. Roads [Roads, 2001]. Il n'est cependant pas prohibé d'utiliser toutes sortes de fonctions dont les variations possibles sont autant de possibilités de rendu sonore. Il faut néanmoins noter que, pour la mise en œuvre actuelle, modifier ce paramètre en temps-réel engendre des artefacts sonores.



taille des grains, du temps de retard et de la raréfaction intègrent du bruit, ou de l'aléatoire, dans leurs utilisations.

Hors du domaine ambisonique, dans une approche plus « classique », les modules de synthèse granulaire sont généralement utilisés en parallèle, afin d'obtenir plusieurs « ruisseaux » de grains, donc une masse plus importante, mais plus de flexibilité et de marge quant aux contrôles. Cette approche qui génère chaque signal associé à un harmonique via un module de synthèse granulaire est alors d'autant plus intéressante. Car au-delà d'accroître la masse, chaque ruisseau de grain est alors défini spatialement, selon l'harmonique sphérique duquel il dépend.

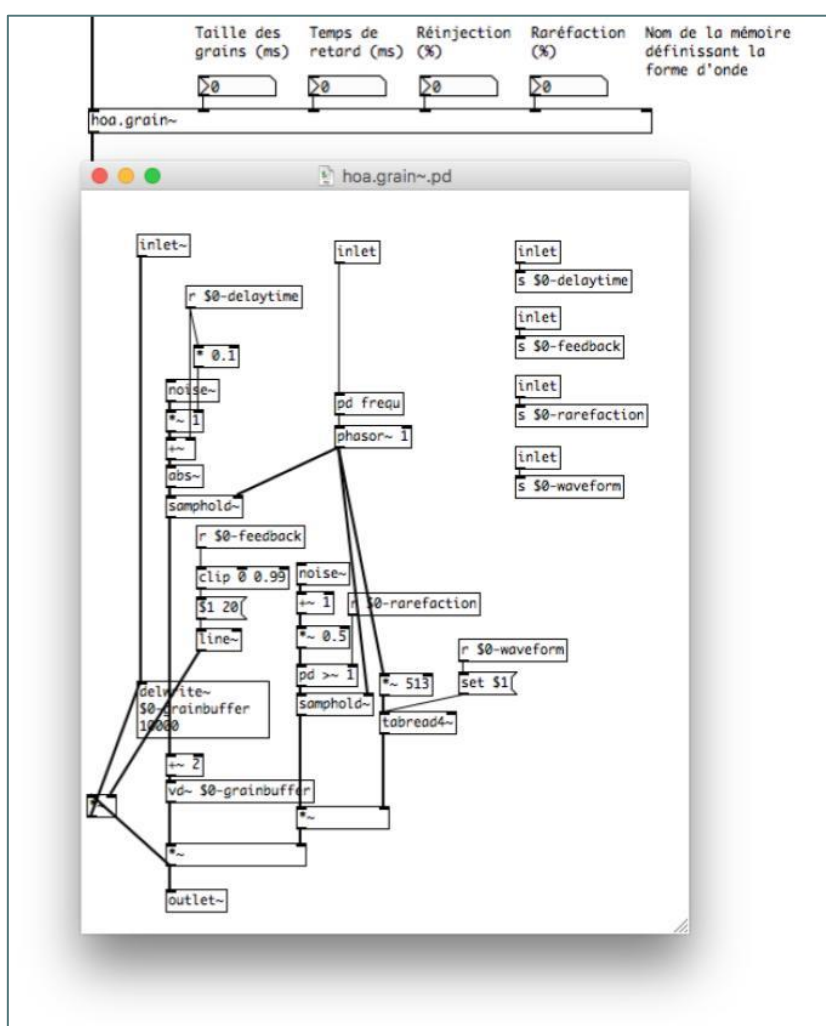


Figure 4.10. Patch *hoa.grain~* dans Pure Data associé à un module de synthèse granulaire quasi-synchrone avec les quatre paramètres principaux la taille des grains, le temps de retard, la réinjection et la raréfaction.

### 4.3.2. Masse sonore et ordre de décomposition

Cette première caractéristique de l'adaptation du module de synthèse dans le domaine des harmoniques circulaires met déjà en évidence une spécificité liée à l'ordre de décomposition. En effet, la masse sonore dépend, dans cette mise en œuvre, du nombre d'harmoniques sphériques et donc de l'ordre de décomposition. Ainsi, un ordre de décomposition peu élevé ne présentera que peu de modules de synthèse granulaire et n'offrira qu'un champ sonore de « faible masse ». En dehors de toute considération spatiale, il faut donc prendre cette caractéristique en considération lors de l'utilisation d'une telle approche. Sur-décomposer un champ sonore par rapport au système de restitution, en vue d'augmenter la résolution angulaire des sources directionnelles n'est dans la pratique que peu pertinent, voire inutile. Comme il a été remarqué précédemment, le modèle mathématique sous-jacent de l'opération de projection, implique que le système de restitution ne puisse *à priori* pas, selon l'approche classique de projection, rendre compte des contributions associées aux harmoniques d'ordre élevé. L'approche granulaire présentée met clairement en évidence une restriction analogue. Si l'opération de projection ignore les contributions associées aux harmoniques d'ordre élevé, alors les ruisseaux de grains qui les définissent ne peuvent se retrouver dans le champ sonore final. Il est possible de voir en cela une simple règle logique, qui définit ceci : de manière analogue à l'approche directionnelle où la résolution angulaire dépend de l'ordre de décomposition, la « qualité » des traitements utilisés en ambisonie dépend aussi de l'ordre de décomposition. Néanmoins, il reste à définir si les implications mises sous la dénomination de « qualité » de traitement ne sont pas plus importantes que celles de la résolution angulaire. Il n'est pas défini que diminuer le nombre de ruisseaux dans un champ sonore ne soit pas plus important que la diminution de la résolution angulaire dans une approche musicale. Un des enjeux de la mise en œuvre de traitements originaux pour la synthèse de champ sonore en ambisonie, est de définir de nouvelles stratégies pour compenser cette contrainte, et cela notamment au niveau de l'opération de projection.

### 4.3.3. Adaptation aux harmoniques sphériques

Au-delà des questions liées à la masse sonore, l'enjeu de la mise œuvre est de pouvoir contrôler les paramètres de chaque module de synthèse granulaire afin si possible, de faire ressortir ou plutôt de jouer sur les caractéristiques spatiales des harmoniques sphériques. Et, par la même occasion, de pouvoir modifier la morphologie sonore des ruisseaux de grains. Une approche minutieuse pourrait définir la valeur de chaque paramètre de chaque module indépendamment.

Cependant, l'enjeu étant de définir une méthodologie de la création de traitements dans le domaine des harmoniques sphériques, de manière identique à la synthèse par décorrélation temporelle, il paraît intéressant de définir une approche générale. L'objectif est alors de concevoir une relation entre les représentations spatiales des harmoniques et les caractéristiques sonores des modules de synthèse qui leurs sont associées. Et faire ainsi en sorte de préserver la nature de ces relations<sup>171</sup>. Tout comme les approches présentées lors de la synthèse de champ sonore par décorrélation temporelle, l'idée revient à définir des macro-paramètres que chaque module puisse adapter aux quatre paramètres énoncés – la taille de grain, le temps de retard, la raréfaction et la réinjection – selon les caractéristiques de l'harmonique auquel il est associé. La méthodologie y est très similaire, de même que les temps de retard des signaux associés aux harmoniques sont définis selon les caractéristiques spatiales des harmoniques sphériques. Il est donc possible d'adapter les macro-paramètres de la granulation selon des considérations spatiales.

Il a été défini arbitrairement que la forme d'onde générant l'enveloppe d'amplitude, la raréfaction et la réinjection, sont spécifiés de façon globale à l'ensemble des modules associés aux harmoniques. Aussi, bien que cela puisse être envisagé, leurs valeurs respectives au sein de chaque module associé à un harmonique sont identiques. Seuls les tailles des grains et les temps de retard sont définis selon les caractéristiques des harmoniques, cela en fonction du degré, de l'ordre azimutal de ceux-ci, ainsi que de l'ordre de décomposition. La raison sous-jacente est aussi que ces paramètres semblent *a priori* être les plus appropriés à une première adaptation dans l'espace, car ils sont les plus évidents à entendre. De plus, la mise en œuvre de l'utilisation de l'enveloppe d'amplitude n'étant pas adaptée à des changements en temps-réel, il n'a pas semblé opportun d'intégrer ce paramètre à l'approche spatiale. La raréfaction est un paramètre faisant varier la masse sonore ainsi que le timbre, mais de manière plus subtile. Il est en outre envisageable de jouer dessus spatialement, ceci permettant de plus ou moins raréfier les grains selon le plan horizontal ou vertical. Cependant, les premiers essais étant restreints à une représentation bidimensionnelle de l'espace et donc au plan horizontal, les possibilités de mise en espace sont plus limitées. Ainsi, il a semblé préférable de laisser ce paramètre global, permettant de se fait de garder un contrôle plus simple de la masse. De plus, limiter les variations spécifiques à chaque harmonique à un petit nombre, permet entre autres de mieux

---

<sup>171</sup> Cette approche possède aussi l'avantage de réduire grandement le nombre de paramètres, en n'offrant via la factorisation de l'approche, pas plus de paramètres à l'utilisateur qu'il n'en existe dans un seul module, facilitant ainsi l'utilisation d'un tel procédé de synthèse.

orienter l'écoute et de mieux concevoir les implications perceptives des approches. C'est donc pour une raison similaire que la réinjection est elle aussi globale. Car les variations de timbre et de masse que ce paramètre entraîne ne sont, selon les tailles des grains ou les temps de retard, que peu perceptibles.

Afin de simplifier les choix et parce que cette présentation ne peut recouvrir toutes les possibilités, il a été défini que le temps des lignes à retard utilisées par les modules de synthèse granulaire, sont adaptés au domaine des harmoniques sphériques de manière identique à la synthèse par décorrélation temporelle, en suivant l'organisation définie par l'ACN. Néanmoins, il est à espérer que le traitement utilisé pour la synthèse étant de nature différente, le résultat perceptif impliqué par cette approche s'en trouve varié par rapport à la synthèse de champ sonore par décorrélation temporelle. L'adaptation des tailles des grains au domaine des harmoniques sphériques a été mise en œuvre selon une approche qui prend le contre-pied de celle-ci, et qui pourrait être dite de type ACN-inverse. Dans ce cas, il est possible de définir que les tailles des grains  $s$  des modules de synthèse granulaire associés aux harmoniques sphériques, diminuent inversement avec leurs degrés. Et qu'au sein d'un degré, les tailles des grains diminuent de même, inversement avec l'ordre azimutal de telle sorte que :

$$s_{0,0} > s_{1,-1} > s_{1,0} > s_{1,1} > s_{2,-2} > s_{2,-1} > \dots > s_{N,N-2} > s_{N,N-1} > s_{N,N}.$$

L'approche nécessite cependant une légère subtilité d'adaptation, en raison du fait qu'il n'est pas souhaitable que les tailles des grains soient nulles. En estimant de même que l'approche accepte des ordres de grandeur linéaires par rapport aux tailles des grains, il est possible de définir mathématiquement la taille de grain relative à un harmonique, selon le nombre total d'harmoniques sphériques  $H$  défini à partir de l'ordre de décomposition, et en fonction d'une taille de grain maximum  $s_{max}$  définie en millisecondes tel que :

$$s_{l,m} = \left(1 - \frac{ACN + 1}{H}\right) s_{max}$$

Équation 4.9. Taille des grains associés aux harmoniques sphériques pour la synthèse de champ sonore par synthèse granulaire quasi-synchrone suivant la logique inverse définie par le format ACN.

Cette fonction de relation permet quant à elle d'assurer que la taille minimale des grains soit de  $\frac{1}{H} t_{max}$  millisecondes et il est aussi possible de développer l'équation à la fois pour une représentation bidimensionnelle et pour une représentation tridimensionnelle tel que

$$s_{l,m} = \left(1 - \frac{(2l)^2 + m}{2Nl + l}\right) s_{max} \text{ en } 2D$$

Et

$$s_{l,m} = \left(1 - \frac{l \times (l + 1) + m + 1}{(N + 1)^2}\right) s_{max} \text{ en } 3D.$$

Selon toute vraisemblance, cette approche devrait permettre d'obtenir des grains de grande taille, relativement à la taille maximale de grain, dans la composante omnidirectionnelle du champ sonore et des grains de plus en plus petits à mesure que les degrés des harmoniques associés aux signaux augmentent. Les plus petits grains sont donc associés spatialement à de nombreux lobes plus fins et plus directifs. En réalité cette mise en relation n'est pas si anodine, il s'agit implicitement de jouer sur le contenu fréquentiel des masses sonores avec les caractéristiques spatiales offertes par la décomposition en harmoniques sphériques. L'idée étant que les grains de petites tailles agissent comme des filtres, étant donné que le module de synthèse granulaire peut être assimilé à un modulateur en anneau associé à une ligne à retard avec réinjection. L'approche consiste donc à générer des grains aptes à préserver le contenu fréquentiel original dans les signaux associés aux harmoniques dont les figures spatiales sont peu directionnelles. Et inversement, les grains favorisant la distorsion, notamment l'émergence de fréquences plus aiguës et plus graves, sont générés dans les signaux associés aux harmoniques dont les figures spatiales sont plus directionnelles. En considérant cette caractéristique spatiale conjointement avec la gestion des temps de retards spécifiques de chaque module selon leur degré et leur ordre azimutal, il est possible d'appréhender un champ sonore où les premiers sons émergent sous forme de grains relativement grands. Mais aussi avec une faible résolution angulaire et en préservant le contenu fréquentiel original. Ainsi, au fur et à mesure, de plus petits grains émergent en dénaturant de plus en plus le signal d'origine. Lors d'un signal source caractérisé par une forte attaque et une extinction, utilisé avec un important facteur de réinjection et un temps de retard maximum relativement grand, il devient aisé de distinguer à la fois le premier grain omnidirectionnel et la profusion de petits grains dénaturés. Ces derniers perdurent plus longtemps et proviennent de toutes les directions, offrant un effet d'impact et de diffusion pertinent musicalement, autant sur le temporel que spatial.

Comme pour les approches précédemment présentées, les enjeux sont doubles. La méthodologie employée vise à jouer sur les caractéristiques spatiales des harmoniques sphériques associés aux signaux composant le champ sonore, mais aussi à définir des approches permettant de tirer parti des différents ordres de décomposition, en rajoutant notamment de la « masse » dans le champ sonore. Tout en adaptant les traitements et modules de synthèse à l'ordre de décomposition, selon les degrés et les ordres azimutaux des harmoniques. Aussi, dans ce contexte, tout comme pour la synthèse de champ sonore par décorrélation temporelle, certains paramètres sont mis à l'échelle afin de préserver la caractéristique spatiale du rendu sonore. Un élément qui interpelle donc dans cette mise en œuvre est la taille minimale des grains qui est de  $\frac{1}{H} t_{max}$  millisecondes. Contrairement à la taille maximale des grains qui reste fixe, la taille minimale est dépendante de l'ordre de décomposition. Aussi par exemple, pour une taille maximale de 100 millisecondes et au premier ordre de décomposition, la taille minimale des grains est d'environ 25 millisecondes. Et dès lors que l'ordre de décomposition augmente, cette taille minimale diminue, avec par exemple 6,25 millisecondes pour le troisième ordre de décomposition. D'un tel écart peuvent résulter de grandes variations du champ sonore, notamment sur le plan fréquentiel. Aussi, une approche plus adaptable consisterait à envisager une taille minimale des grains telle que

$$s_{l,m} = \left(1 - \frac{ACN}{H}\right) (s_{max} - s_{min}) + s_{min}$$

Équation 4.10. Taille des grains associés aux harmoniques sphériques pour la synthèse de champ sonore par synthèse granulaire quasi-synchrone suivant la logique inverse définie par le format ACN avec prise en compte d'une taille minimale.

Que ce soit via la synthèse des signaux associés aux harmoniques sphériques par décorrélation temporelle du signal source ou via l'utilisation de modules de synthèse granulaire, les champs sonores engendrés possèdent des caractéristiques spatiales pour le moins originales et intéressantes sur un plan musical, mais néanmoins chaotiques. Il est, quelles que soient les valeurs de paramètres utilisées, impossible de pouvoir revenir vers une approche plus directionnelle permettant à l'auditeur de définir le champ sonore comme le résultat d'une source. Les opérations mises en œuvre sont radicales, notamment par leurs utilisations de la ligne à retard, et impliquent une perte de contrôle sur cette approche de l'espace. Un des champs de recherche qu'il est possible d'explorer, est de définir des approches permettant de passer linéairement d'une source directionnelle à un champ sonore diffus ou chaotique.

## 4.4. Synthèse par pondération aléatoire

L'encodage directionnel d'une source sonore dans le domaine des harmoniques sphériques revient à synthétiser les signaux associés à chacun de ces harmoniques, en créant autant de versions pondérées du signal source. Et dont les gains respectifs sont fonctions des angles d'incidence – azimut et élévation – ainsi que des degrés et des ordres azimutaux des harmoniques sphériques. Par le principe même de cette approche, il existe logiquement des rapports essentiels entre les gains associés aux harmoniques, en considérant que tous les harmoniques sphériques prennent les mêmes angles comme variables. Cela permettant de créer une figure directive et de simuler pour l'auditeur une source sonore provenant d'une direction donnée de l'espace, lorsque tous les signaux associés aux harmoniques sphériques sont projetés sur un système de restitution. Sur ce constat élémentaire, une expérimentation originale du modèle ambisonique peut être de jouer sur cet ensemble de gains afin de brouiller les relations, pour générer de nouvelles figures de directivité. Dès lors, il pourra être envisagé de se tourner à nouveau vers un système directionnel, en revenant linéairement vers des coefficients d'encodage classiques. Afin de faciliter la mise en œuvre, la présentation se limite à une approche bidimensionnelle de l'espace. Cela permet notamment d'éviter les calculs des polynômes associés de Legendre, nécessitant une approche récursive difficile à mettre en œuvre sous la forme d'un patch Pure Data pour tous les ordres de décomposition. Cependant, les principes présentés restent valides pour une représentation tridimensionnelle. De plus, la question de la faisabilité technique de la mise en œuvre des calculs des coefficients d'encodage tridimensionnels dans Pure Data ou Max sera soulevée par la suite.

### 4.4.1. Génération de gains aléatoires

Dans ce contexte, la première approche présentée consiste à générer aléatoirement et de manière indépendante les gains à appliquer aux signaux associés harmoniques circulaires. La méthode équivaut à réaliser une opération de synthèse d'une source sonore virtuelle, mais dont la figure de directivité est aléatoire et comporte un nombre variable de lobes plus ou moins proéminent. Et qui dépend du nombre de signaux, donc de l'ordre de décomposition et des relations de gains entre eux, au lieu d'une figure unidirectionnelle permettant la synthèse d'une source sonore ponctuelle telle qu'une cardioïde ou hypercardioïde [Figure 4.11].

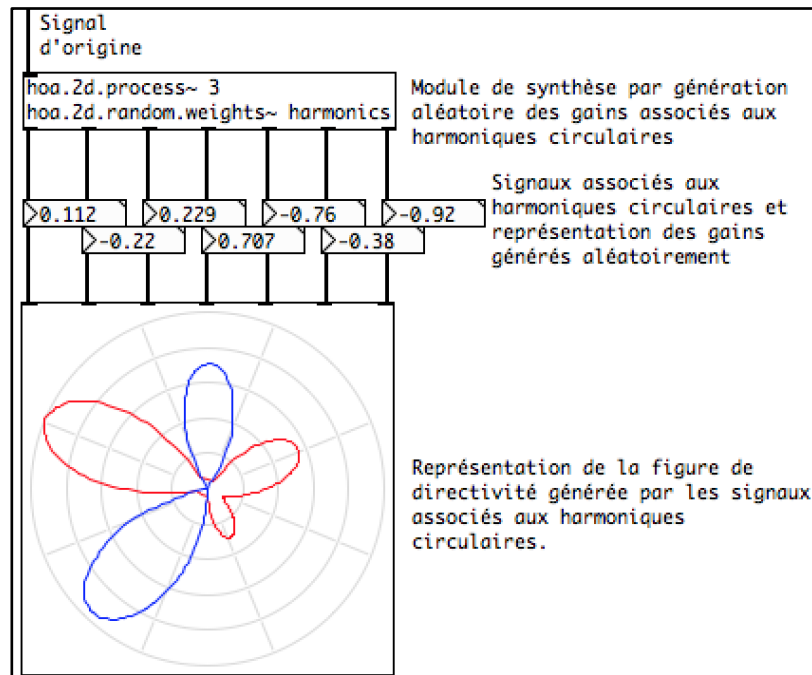


Figure 4.11. Abstraction *hoa.2d.random.weights~* dans le logiciel Pure Data permettant de synthétiser un champ sonore dans le domaine des harmoniques circulaires selon une série de gains aléatoires. Les signaux associés aux harmoniques circulaires sont des versions du signal d'origine, pondérées aléatoirement. La figure de directivité associée au champ sonore possède de nombreux lobes dépendant des gains.

Cette approche permet d'obtenir un espace sonore composé d'une source aux qualités spatiales souvent multidirectionnelles, dépendant du jeu de gains générés. L'espace sonore escompté est composé d'une source sonore atypique, pouvant être assimilée à une série de sources sonores ponctuelles identiques sur le plan du contenu fréquentiel<sup>172</sup>, mais positionnées autour de l'auditoire de manière aléatoire et avec des variations de gains telles que définies par les différents lobes de la figure de directivité. La restitution de l'espace sonore dépend de la méthode de projection des harmoniques circulaires, de l'optimisation de l'espace sonore et du système de restitution. Ainsi, le résultat perceptif pour un auditoire donné implique des variations ou des altérations de cet espace sonore escompté. L'artefact majeur est la fusion perceptive des lobes adjacents en un seul lobe plus proéminent, résultant en une source sonore à la directivité élargie au lieu de multiples sources sonores ponctuelles plus directives. Il est difficile de prévoir et de formaliser cet effet, car il dépend aussi de la nature spectrale du son d'origine. À titre d'exemple la restitution sur un système de huit haut-parleurs placés de manière régulière sur un cercle suite à une projection directe et sans optimisation des harmoniques circulaires, résulte en une source située à gauche de l'auditoire et recouvrant plus d'un demi-cercle de l'espace sonore. Les contributions des deux petits lobes situés sur la partie droite de

<sup>172</sup> Les sources sont possiblement en opposition de phase mais dans la pratique, cette caractéristique est négligeable, notamment en haute fréquence.



l'espace sonore sont négligeables et les contributions des trois autres lobes sont combinées pour ne former qu'une seule et large source sonore.

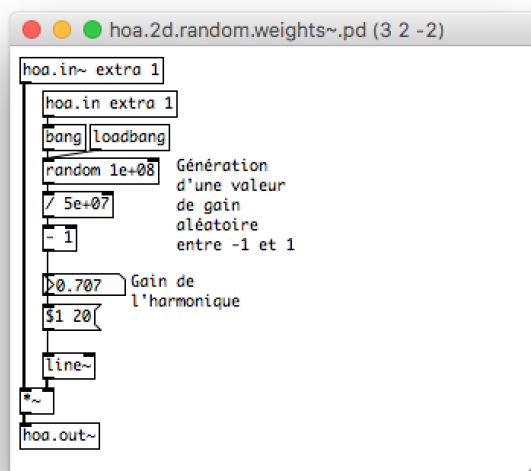


Figure 4.12. Instance du patch Pure Data *hoa.2d.random.weights~* associée à l'harmonique de degré 2 et d'ordre azimutal -2 utilisée pour la génération de gains aléatoires. La valeur de gain est générée entre -1 et 1 à l'instanciation de l'abstraction.

Cette approche a été réalisée via l'abstraction *hoa.2d.random.weights~* pour le logiciel Pure Data, qui synthétise les signaux associés aux harmoniques circulaires en pondérant le signal d'origine selon des gains définis aléatoirement à l'instanciation, et pouvant recouvrir des valeurs allant de  $-1$  à  $1$  tel que présenté sur la [Figure 4.12]. La plage de valeurs des gains est définie de manière arbitraire, suivant la règle élémentaire en traitement du signal numérique selon laquelle si le signal d'origine est normalisé et que la valeur absolue de l'amplitude dépasse  $1$ , alors cela amène à la saturation du signal sonore. Cependant, cette approche est établie sans considération du degré, ni de l'ordre azimutal des harmoniques auxquels ils sont associés, pas plus encore que de l'ordre de décomposition.

Dans le cas présent, l'approche ne tient pas compte des lois de l'énergie du champ sonore en ambisonie, telles que définies par [Équation 4.1] et [Équation 4.2]. Il en résulte que l'énergie globale des contributions de chaque degré de l'espace sonore, fait qu'elles ne sont pas toutes égales. Cela implique donc une accentuation de certains degrés par rapport à d'autres, ainsi

qu'une variation de l'énergie globale de l'espace sonore en fonction du jeu de gains générés<sup>173</sup>. Une des conséquences concrètes d'une telle approche est la possibilité d'un espace sonore de faible amplitude sonore, voire inaudible. Une autre des conséquences reste l'inanité des contributions de certains degrés par rapport à d'autres, pouvant générer des artefacts tels qu'un espace sonore peu directionnel, et même omnidirectionnel lorsqu'un degré contribue de manière très importante comparé aux autres.

Un aspect marquant de cette approche est évidemment le manque de mouvement. La génération de poids aléatoires au sein de l'abstraction *hoa.2d.random.weights~* n'utilise aucune variable d'entrée et n'offre pas de contrôle, contrairement à la synthèse de source sonore ponctuelle qui utilise des angles d'incidences variables afin de créer du mouvement<sup>174</sup>.

#### 4.4.2. L'encodage aléatoire

Une approche plus pertinente et permettant de mieux concevoir le potentiel de cette approche voudrait donc pouvoir moduler l'effet dans le temps et créer des variations et du mouvement. La façon la plus simple, tout en utilisant une génération aléatoire de poids, est de générer de nouveaux poids à une fréquence donnée et de glisser linéairement entre eux. Plusieurs modifications peuvent être envisagées telles que le changement de fonction glissante par une fonction exponentielle ou logarithmique afin de créer des dynamiques différentes de mouvements. Cependant afin de garder une présentation relativement concise et étant donné que l'idée principale est simplement de créer des variations continues, une simple fonction linéaire est utilisée. Le résultat perceptif est une source sonore plus ou moins directionnelle et plus ou moins « dupliquée » dont la vitesse varie en fonction du paramètre de fréquence. Le champ sonore n'est pas proprement parlant diffus, mais peut être qualifié de chaotique.

Afin de répondre au problème de normalisation, une manière simple pour obtenir une série de gains aléatoires tout en respectant la normalisation ambisonique est de générer aléatoirement un angle différent par degré puis de calculer les gains qui leur sont associés suivant la formulation de l'encodage directionnel classique [Figure 4.13]. Cette approche permet alors de respecter la loi de l'énergie du degré, et donc logiquement du champ sonore, par l'utilisation purement et simplement des équations dont ces lois découlent tout en offrant un caractère

---

<sup>173</sup> Evidemment l'énergie globale de du champ sonore dépend aussi de la source sonore d'origine.

<sup>174</sup> Il est à noter ici un aspect intéressant des approches construites sur l'utilisation de lignes à retard qui en agissant dans le domaine temporel créent directement des variations spatiales.

chaotique en donnant en entrée de ces fonctions des valeurs différentes selon les degrés. En somme, les relations de gains permettant d'assurer une directionnalité cohérente au champ sonore restent cohérentes pour les signaux associés aux harmoniques d'un même degré mais sont rompues dès lors que l'ensemble des signaux associés à un degré sont comparés à ceux associés à un autre degré. Cette approche offre alors un champ sonore multidirectionnel et de nature chaotique mais préservant une énergie globale stable et tirant à tout moment partie des contributions de l'ensemble des composantes qui le compose permettant d'éviter une trop grande variation de la résolution angulaire.

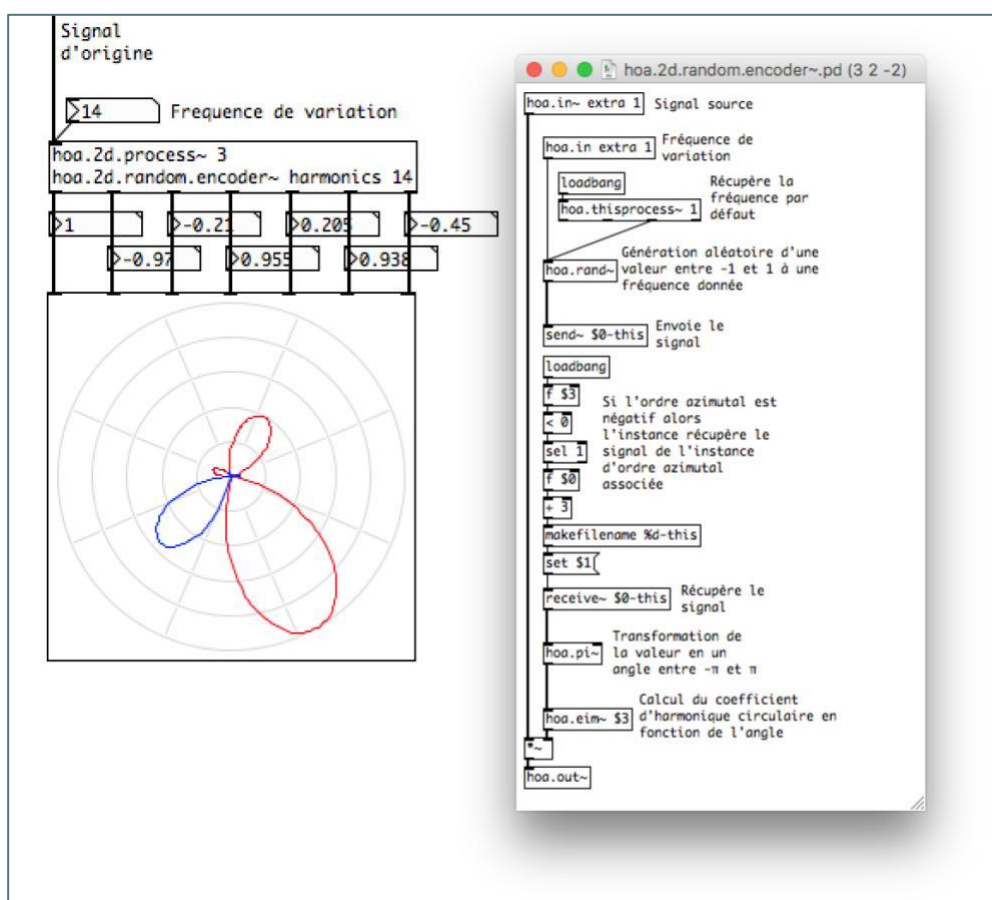


Figure 4.13. Abstraction *hoo.2d.random.encoder~* dans le logiciel Pure Data permettant de synthétiser un champ sonore dans le domaine des harmoniques circulaires selon une série de gains aléatoires, normalisés et modulant à une fréquence donnée. Les signaux associés aux harmoniques circulaires sont des versions du signal d'origine pondérées. Les gains de pondérations sont modulés linéairement et aléatoirement selon la fréquence donnée en entrée. La figure de directivité associée au champ sonore possède de nombreux lobes dépendant des gains. A droite, l'instance de l'abstraction associée à l'harmonique de degré 2 et d'ordre azimutal -2 révélant le système permettant de normaliser les harmoniques. Un système d'objet *send~* et *receive~* permet aux instances associées à un même degré de partager le même angle puis de définir le gain en calculant le coefficient d'encodage correspondant.

Il est néanmoins possible d'envisager une approche où la résolution angulaire varie aussi de manière aléatoire tout en préservant l'énergie globale du champ sonore. L'enjeu serait alors de rompre les relations de pondération au sein des signaux associés à un même degré mais

d'assurer que la somme des énergies de tous les degrés reste égale à l'ordre de décomposition. Pour cela, la mise en œuvre nécessiterait de calculer successivement les gains des signaux associés à chaque harmonique aléatoirement, mais toujours normalisés entre -1 et 1, à l'exception d'un dernier gain ou d'un ensemble de derniers gains pour lesquels l'enjeu devra être de compenser la perte potentielle d'énergie du champ sonore. En considérant que le processus de calcul descend des signaux associés aux ordres élevés jusqu'au signal associé à l'harmonique omnidirectionnelle alors le gain de ce dernier pourrait être calculé tel que

$$g_0 = 2 - \sum_{l=1}^N E_l.$$

Néanmoins une telle approche a, dans le cas présent, tendance à favoriser la contribution de l'harmonique omnidirectionnelle réduisant le caractère chaotique du champ sonore. Et si le choix se porte sur un autre signal ou un autre ensemble de signaux afin de compenser l'énergie du champ sonore, alors le problème persiste et spatialement la figure de directivité associée à ces composantes étant de même favorisée, elle crée potentiellement un déséquilibre dans le champ sonore vers les directions associées. Aussi une telle approche ne semble pas au final souhaitable.

De plus, en reprenant la précédente proposition il est aisément possible d'augmenter la mise en œuvre afin de prendre en compte la nécessité de retrouver un champ directionnel.

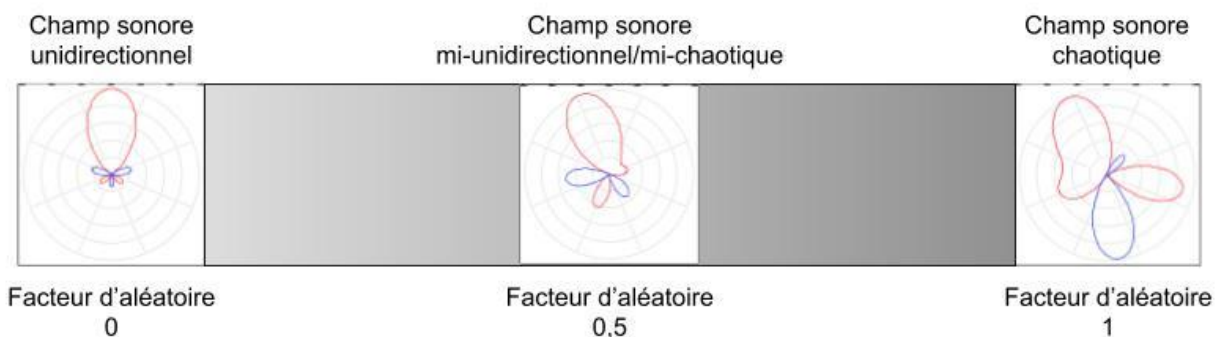


Figure 4.14. Représentation d'une échelle linéaire entre deux états de champ sonore. Avec à gauche un champ sonore unidirectionnel associé à un facteur d'aléatoire de 0, à droite un champ sonore chaotique associé à un facteur d'aléatoire de 1 et au centre l'état « parfaitement » intermédiaire mi-unidirectionnel/mi-chaotique associé à un facteur d'aléatoire de 0,5.

### 4.4.3. Le facteur d'aléatoire

L'idée est de proposer une nouvelle approche qui offre la capacité de synthétiser les signaux associés aux harmoniques circulaires, à la fois de manière directionnelle et de manière chaotique. Il est intéressant de remarquer que la mise en œuvre de cette approche permettant de passer d'un état directionnel à un état chaotique, peut être définie selon des spécifications techniques et fonctionnelles très similaires à celles établies pour la variation de la résolution angulaire. Les principes sous-jacents sont d'ailleurs identiques à ceux présentés précédemment, indépendamment bien sûr de l'objectif final.

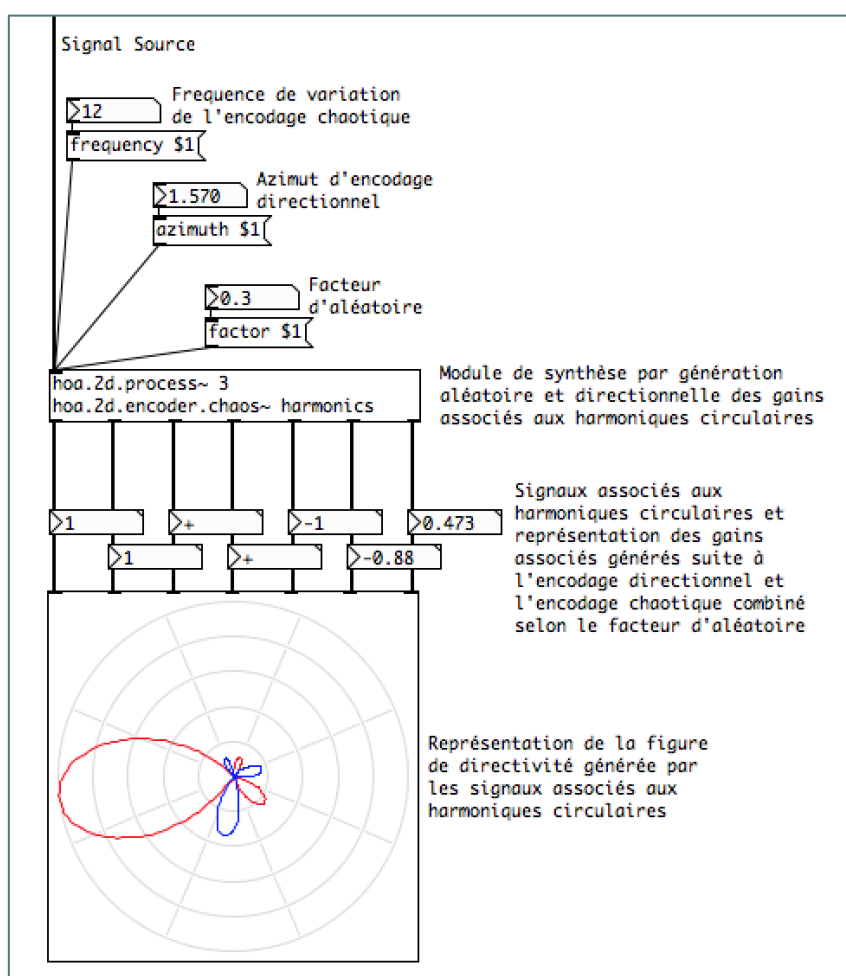


Figure 4.15. Abstraction *hoa.2d.encoder.chaos~* dans le logiciel Pure Data permettant de synthétiser un champ sonore dans le domaine des harmoniques circulaires selon une série de gains aléatoires, normalisés et modulant à une fréquence donnée ainsi que selon un encodage directionnel. Le facteur d'aléatoire étant à 0,3, le champ sonore est sensiblement plus directionnel que chaotique. Aussi les gains des signaux associés aux harmoniques circulaires restent corrélés sur un plan directionnel jusqu'au degré 2 inclus, mais les gains des signaux associés aux harmoniques circulaires de degré 3 rompent la corrélation et en étant issus d'un angle généré aléatoirement.

Dans ce contexte, il est souhaitable de pouvoir générer l'ensemble des possibilités intermédiaires entre les deux états extrêmes. De même, il est aussi envisagé de pouvoir offrir

une variation linéaire, en évitant les phénomènes d'inégalité des contributions associées aux harmoniques sphériques ou circulaires en fonction de leurs degrés. En somme, l'objectif est de pouvoir glisser d'un état à un autre en donnant l'impression que le champ sonore devient de plus en plus chaotique ou de plus en plus unidirectionnel. Enfin, comme cela a été réalisé dans les exemples précédents, l'enjeu est ici d'offrir un traitement de synthèse de l'espace et du son qui s'adapte aux différentes résolutions du champ sonore, et donc de l'ordre de décomposition. Il s'agit en l'occurrence et de manière similaire à la mise en œuvre de la variation angulaire, de définir une échelle « juste » sur le plan perceptif au centre de laquelle le champ sonore serait interprété comme semi-chaotique. Cette approche propose alors un facteur dit d'aléatoire, ou un facteur chaotique, compris entre 0 et 1. À 0 le champ sonore est unidirectionnel, à 1 le champ sonore est chaotique et multidirectionnel, et les valeurs intermédiaires recouvrent le spectre des possibilités entre ces deux états de telle sorte qu'à 0.5 le champ sonore soit mi-unidirectionnel/mi-chaotique, en considérant que cet état d'équilibre existe [Figure 4.14].

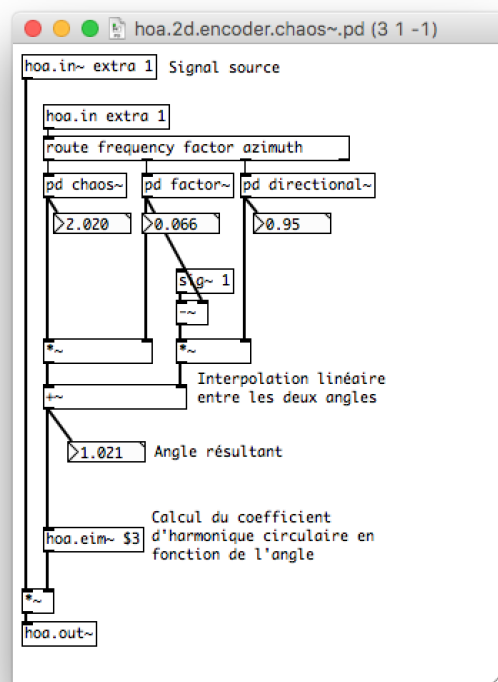


Figure 4.16. Instance du patch *hoa.2d.encoder.chaos~* associée à l'harmonique de degré 1 et d'ordre azimutal -1. Le patch fait une interpolation entre l'angle d'encodage et l'angle généré aléatoirement, suivant un facteur de transition adapté en fonction du degré de l'harmonique associé à l'instance.

Sur le plan des spécifications, une telle approche nécessite donc de recevoir en plus du signal source en entrée, une direction permettant de définir la position du signal lors de l'encodage

directionnel ainsi qu'une fréquence permettant de contrôler la génération de gains aléatoires lors de l'état « chaotique ». Enfin, un dernier paramètre doit permettre de définir l'état entre ces deux champs sonores [Figure 4.15] et [Figure 4.16].

L'approche consiste alors à réaliser une interpolation entre un encodage chaotique et un encodage directionnel. Il est alors possible d'envisager une interpolation directe entre deux champs sonores, l'un synthétisé via une opération d'encodage directionnel et l'autre synthétisé selon l'approche qui vient d'être présentée. C'est-à-dire réaliser une interpolation terme à terme des signaux respectifs des champs sonores. Mais dans le cas présent, ce serait ne pas tirer parti de la mise en œuvre de l'encodage chaotique. En effet, étant donné que le calcul des coefficients d'encodage directionnel est intégré à l'opération de synthèse de champ sonore chaotique et aléatoire, il est possible d'utiliser cette opération à la fois pour les deux approches afin de limiter les calculs et de simplifier la mise en œuvre. L'aspect chaotique ne reposant que sur la génération aléatoire des angles azimutaux, il peut être envisagé que le passage d'un état chaotique à un état unidirectionnel se fasse uniquement par l'interpolation entre la génération de ces angles et l'angle d'encodage directionnel<sup>175</sup> [Figure 4.16].

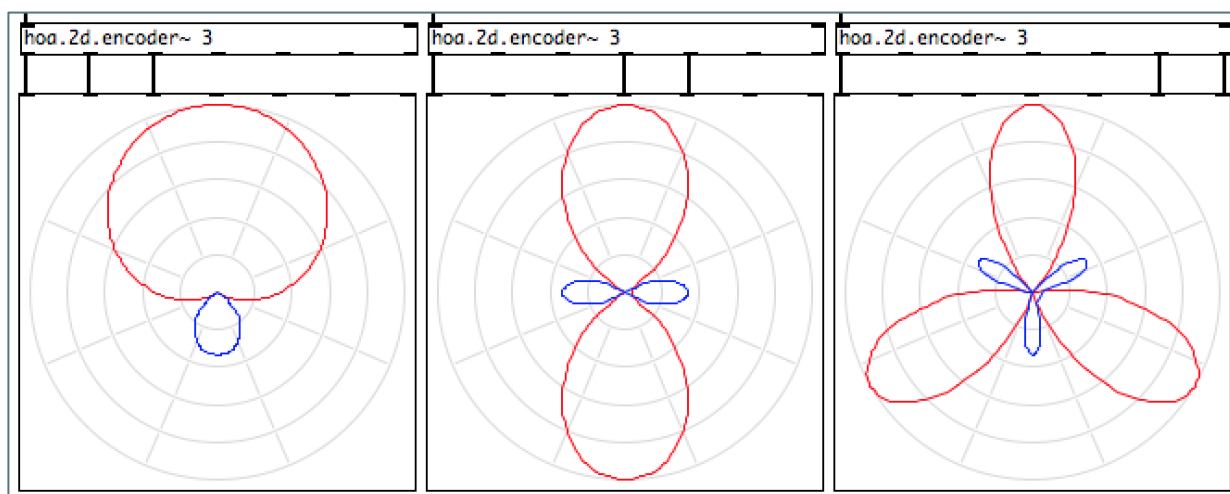


Figure 4.17. Représentation de l'importance de la relation entre l'harmonique de degré 0 et les harmoniques des autres degrés pour la construction de l'uni-directionnalité. La combinaison des signaux associés aux harmoniques de degrés 0 et 1 préserve l'unidirectionnalité alors que la combinaison des signaux associés aux harmoniques de degrés 0 et 2 ou 0 et 3 génère des figures multidirectionnelles.

<sup>175</sup> Il est important de noter que la mise en œuvre présentée réalise une interpolation linéaire directement entre l'angle généré aléatoirement et l'angle reçu sans prendre en compte l'aspect cyclique des angles résultant en une variation possiblement fautive de l'écart angulaire dépasse  $\pi$ . Une meilleure mise en œuvre tiendrait compte de cette caractéristique des angles en réalisant par exemple une interpolation sur des coordonnées cartésiennes pour par la suite revenir sur des coordonnées polaires ou sphériques.

La complexité réelle réside alors dans la mise en œuvre de la fonction d'interpolation. Il a été présenté que les harmoniques sphériques et circulaires ne contribuent pas de manière identique à la représentation de l'espace et que les degrés élevés ne permettent d'augmenter que faiblement la résolution angulaire. Il est aussi nécessaire de rappeler qu'au-delà d'une simple question de résolution angulaire, la justesse de l'opération de décomposition repose sur la hiérarchie des harmoniques sphériques et circulaires. Aussi, rompre la relation d'amplitude entre l'harmonique de degré 0 et les harmoniques de degré 1 implique de perdre automatiquement tout le processus directionnel, même si des harmoniques d'ordre supérieur restent corrélées dans un sens directionnel, avec l'harmonique de degré 0. A l'opposé, rompre la relation d'amplitude entre l'harmonique de degré 0 et les harmoniques de degrés élevés implique seulement une déformation du processus directionnel, tant qu'au minimum les harmoniques de degré 1 restent corrélés dans un sens directionnel avec l'harmonique de degré 0. Pour bien concevoir ce mécanisme, il est possible d'utiliser la spécificité de la décomposition en harmoniques sphériques permettant d'additionner les représentations et donc de considérer un champ sonore comme la somme d'un ensemble de champs sonores. Dès lors, il est possible de représenter un champ sonore composé d'une source directionnelle pour un ordre de décomposition donné  $N$  comme la somme de  $N$  champs sonores, associés respectivement à chaque degré. Et dont les contributions ne sont préservées que dans les signaux associés à ces degrés, ainsi qu'au degré 0 [Figure 4.17]. Aussi est-il possible de remarquer que seul le champ sonore possédant des contributions dans les signaux associés aux degrés 0 et 1 préserve un caractère directionnel, car cela revient à une décomposition d'ordre 1. Les autres champs sonores sont alors associés selon leurs contributions à des champs sonores multidirectionnels. Ainsi, rompre exclusivement les relations de pondération des signaux associés aux degrés élevés tout en préservant celles des signaux associés aux degrés 0 et 1, revient à considérer le champ sonore comme l'addition d'un champ sonore généré par un encodage directionnel au premier ordre de décomposition avec un champ sonore multidirectionnel. Ce champ sonore étant possiblement aléatoire et dont les contributions sont restreintes aux signaux associés aux degrés élevés. Le champ sonore résultant reste donc directionnel mais possède des caractéristiques chaotiques liées aux ordres élevés. Mais rompre les relations de pondération des signaux associés aux degrés 0 et 1, même en préservant celles des signaux associés aux degrés élevés, revient à considérer un champ sonore comme l'addition de deux champs sonores multidirectionnels.

Suite à cette analyse, il paraît essentiel que le facteur préserve la relation de pondération entre les signaux associés aux premier et deuxième degrés tant que la nécessité de mettre en valeur



une direction apparente est nécessaire. Aussi, une approche similaire à celle proposée pour la variation de la résolution angulaire qui modifie simultanément les gains de l'ensemble des signaux associés aux harmoniques, n'est pas souhaitable dans ce contexte. L'enjeu est alors, dans le cas d'un passage d'un champ sonore directionnel à un champ sonore chaotique, de moduler au départ les gains des signaux associés au degré le plus élevé, afin qu'il ne soit progressivement plus issu d'un encodage directionnel mais d'un encodage aléatoire et ce de manière continue et linéaire. Une fois que ces gains sont générés de façon totalement aléatoire, le processus consiste à réaliser la même opération pour les signaux associés au degré inférieur. Et cela jusqu'à atteindre les signaux associés au degré 1 lorsque le facteur d'aléatoire est de 1, car leur modification fait perdre directement le caractère directionnel du champ sonore. Le signal associé à l'harmonique de degré 0 reste inchangé, car l'approche aléatoire définit implicitement que son gain est 1 tout comme lors d'un encodage directionnel [Figure 4.18]. Et le passage d'un champ sonore chaotique à un champ sonore directionnel nécessite simplement le mécanisme inverse.

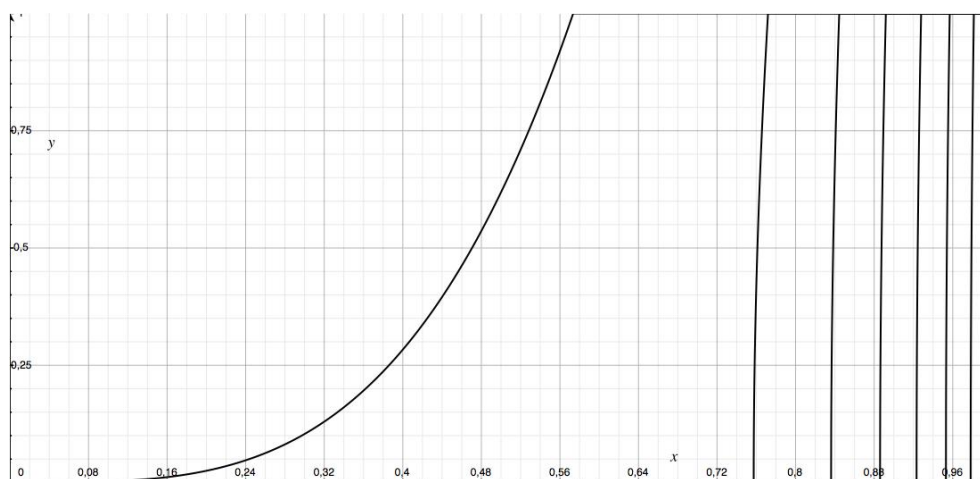


Figure 4.18. Représentation de la courbe du facteur d'interpolation pour un ordre de décomposition 7. De gauche à droite, les courbes sont associées aux degrés 7 à 1.

Enfin il n'est pas nécessaire de prendre en compte une diminution de l'énergie du champ sonore. Comme le système de chaque module de synthèse est fondé sur une interpolation entre deux gains, ou plutôt deux angles, chaque signal contribue au champ sonore quel que soit le facteur d'aléatoires. Aussi, le facteur d'interpolation, définissant l'amplitude de l'angle généré aléatoirement et réciproquement de l'angle d'encodage directionnel, doit être compris entre 0 et 1 en fonction du facteur d'aléatoire  $x$  et selon l'ordre de décomposition et le degré de l'harmonique auquel il est associé tel que

$$f_l(x) = \sqrt{x^N - \frac{(N-l)}{N}} \quad N \begin{cases} f(x) < 0 \Rightarrow f(x) = 0 \\ f(x) > 1 \Rightarrow f(x) = 1 \end{cases}$$

Équation 4.11 : Facteur d'interpolation définissant l'amplitude de l'angle généré aléatoirement, et réciproquement de l'angle d'encodage directionnel, pour un degré donné et un ordre de décomposition donné en fonction du facteur d'aléatoire.

Au final, ce traitement offre plusieurs caractéristiques intéressantes. Il s'agissait de partir d'un champ sonore chaotique, puis la création d'une courbe d'interpolation a permis de générer des transitions permettant de transformer linéairement ce champ sonore en une source sonore directionnelle. Cette courbe d'interpolation a été conçue selon les caractéristiques spatiales des harmoniques et les principes de création de directionnalité du champ sonore. Enfin, la mise en œuvre a été conçue selon des principes déjà établis tels que l'adaptation à toutes les résolutions variantes du champ sonore, c'est à dire à différents ordres de décomposition. Cela via la mise en place d'un facteur d'aléatoire qui est compris entre 0 et 1, mais aussi grâce à l'utilisation de l'ensemble des composantes du champ sonore.

#### 4.4.4. Généralisation de l'approche

Partant de cette approche sur un champ sonore chaotique, il est possible d'appliquer le même type de procédé sur d'autres techniques de synthèse de champ sonore. Dans le cadre de ces recherches, une approche similaire a été mise en place pour créer un traitement permettant de passer linéairement vers une source sonore directionnelle, mais à partir d'un champ sonore diffus synthétisé par une décorrélation des signaux associés aux harmoniques sphériques.

Cependant, il est apparu que la mise en œuvre du facteur d'interpolation telle que proposée par l'[Équation 4.11] peut ne pas répondre correctement à toutes les approches. Selon l'équation, pour un ordre de décomposition 7 et un facteur d'aléatoire de 0,5, seuls les signaux associés aux harmoniques de degrés 6 et 7 sont décorrélés. Or, certaines approches peuvent nécessiter des variations plus ou moins radicales. En somme, il serait intéressant de pouvoir contrôler le caractère plus ou moins linéaire de l'émergence des variations selon les degrés des harmoniques [Figure 4.19]. Aussi, afin de répondre à ce besoin, un indice de linéarité  $p$  a été introduit dans l'équation tel que

$$f_l(x, p) = \sqrt{x^{N+p} - \frac{(N-l)}{N}} \quad N \begin{cases} f(x) < 0 \Rightarrow f(x) = 0 \\ f(x) > 1 \Rightarrow f(x) = 1 \end{cases}$$

Équation 4.12. Adaptation du facteur d'interpolation à de multiples approches selon un indice de linéarité.

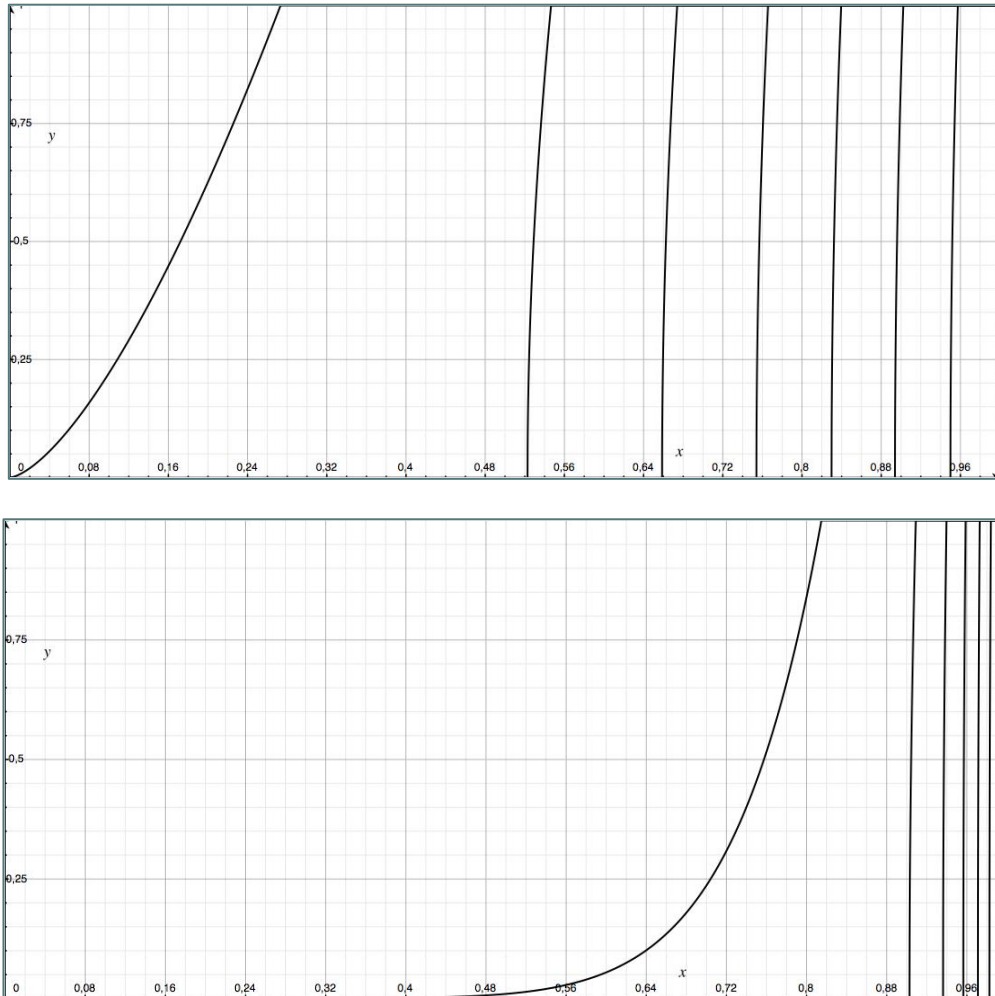


Figure 4.19. Représentation de la courbe du facteur d'interpolation prenant en compte un indice de linéarité pour un ordre de décomposition 7. De gauche à droite, les courbes sont associées aux degrés 7 à 1. En haut, l'indice de linéarité est de -4 et en bas, l'indice de linéarité est de 12.

Cette approche permet d'obtenir des changements plus abrupts dès lors que l'indice de linéarité diminue. Aussi, les variations sur les signaux associés aux harmoniques de degrés faibles apparaîtront plus rapidement. Ce type d'approche est particulièrement utilisé dans le cas d'un traitement ou d'une opération qui ne modifie pas énormément le caractère directionnel du champ sonore, tant que cela reste appliqué aux harmoniques de degrés élevés<sup>176</sup>. A l'opposé, un indice positif permet de faire en sorte que les variations apparaissent très tardivement sur les

<sup>176</sup> Il a été mis en œuvre des effets de saturation et de distorsion qui ont cette propriété.

signaux associés aux harmoniques de degrés faibles. Réciproquement, cette approche est utile dans le cas d'un traitement qui distord énormément les signaux et casse rapidement le caractère directionnel du champ sonore<sup>177</sup>.

## 4.5. Flanging ambisonique

Lors de la présentation de la synthèse de champ sonore par décorrélation temporelle, la méthodologie consistait à définir l'opération de base puis à l'appliquer telle quelle sur les signaux associés aux harmoniques sphériques. Puis il s'agissait de comprendre le résultat de cette expérimentation dans le champ sonore. En somme, ce procédé permet de voir comment la représentation du champ sonore sous la forme de signaux associés aux harmoniques sphériques permet de transformer une opération en une autre. L'enjeu d'une telle approche est de discerner les modifications sur le résultat même de l'opération que le changement de domaine implique. Dans le cas de la synthèse de champ sonore par décorrélation temporelle, une simple ligne à retard devient dans le domaine des harmoniques sphériques, une opération équivalent à un système de multiples filtres en peigne en parallèles, soit des lignes à retard avec réinjection et dont les coefficients sont définis dans l'espace selon les caractéristiques des harmoniques sphériques. À partir de ces expérimentations, de leurs aboutissants et des conclusions qu'il a été possible d'en tirer, il est à présent envisageable de procéder selon une méthodologie inversée, en tentant de retrouver une opération ou un traitement spécifique via le domaine des harmoniques sphériques. Il s'agit dans un premier temps de définir ce que devrait être ce traitement dans le domaine ambisonique. En effet, il ne s'agit pas de définir le traitement spatialement afin qu'il préserve ses caractéristiques initiales. L'application d'une distorsion, bien que répandue dans l'espace, devrait toujours sonner à l'oreille comme une distorsion. Dans un second temps, l'enjeu est de retrouver l'opération correspondante au niveau des signaux associés aux harmoniques sphériques pour recréer cet effet.

Le *flanging* semble particulièrement adapté à une extension en ambisonie vers le domaine des harmoniques sphériques. En effet, il s'agit d'un traitement qui a été défini dans un premier temps de manière monophonique pour, par la suite être étendu à une approche stéréophonique<sup>178</sup>. Dans cette approche, il ne s'agit donc pas d'appliquer un module de *flanger* sur chacun des signaux associés aux harmoniques, mais de fonder l'approche au regard des

---

<sup>177</sup> A ce titre la décorrélation temporelle des signaux ou la synthèse granulaire sont de bons exemples.

<sup>178</sup> Et parce que c'est un traitement que j'apprécie particulièrement notamment en raison du bagage historique qu'il apporte avec lui, du fait de son utilisation par les Beatles et Jimi Hendrix.

observations et résultats obtenus suite à la décorrélation temporelle dans le domaine des harmoniques sphériques. L'effet souhaité dans le champ sonore devrait donc être une sorte de « *multi-flanging spatial* ». Afin de définir le *flanging ambisonique* il est nécessaire de procéder par étape en commençant par le traitement initial puis en complexifiant petit à petit l'approche.

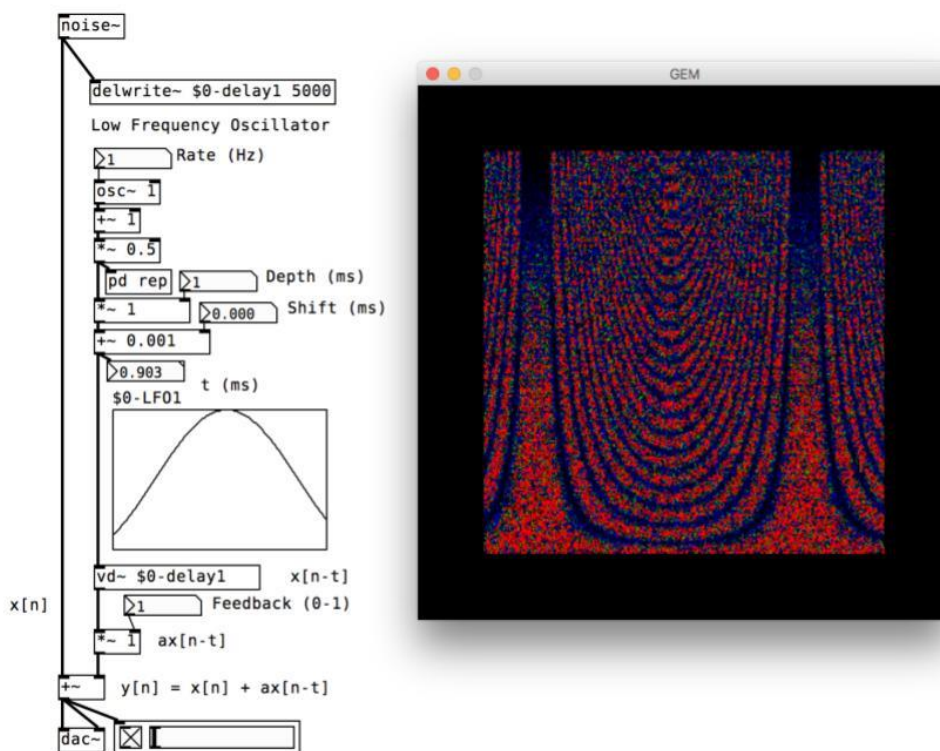


Figure 4.20. Patch Pure Data réalisant un flanging sur un bruit blanc. Le résultat de l'effet est représenté à droite dans la fenêtre GEM où le balayage des fréquences est visible sur le sonagramme.

#### 4.5.1. Flanging monophonique

Le *flanger* est un module traitement où le signal original est mixé avec une version retardée de lui-même et dont le temps de retard est contrôlé par un *LFO* (oscillateur basse fréquence) [Smith, 2010] produisant l'effet de *flanging* [Figure 4.20].

Cet effet est le produit  $y$  d'un filtre en peigne qui balaye le spectre fréquentiel du signal original  $x$  selon la fréquence  $f$  du LFO, une profondeur  $p$  et un retard initial  $o$  en temps ainsi qu'un coefficient  $a$  de réinjection entre 0 et 1 tel que

$$y[n] = x[n] + ax \left[ n - \frac{\cos(2\pi fnT) + 1}{2} p + o \right].$$

Équation 4.13. Effet de flanging.

## 4.5.2. Flanging stéréophonique

L'idée derrière le *flanging stéréophonique* est de générer un espace sonore où le balayage fréquentiel se balance entre l'image spatiale de gauche et l'image spatiale de droite [Figure 4.21]. Cet effet de l'espace et du son peut être créée via deux canaux ou plus et pour deux haut-parleurs ou plus. De même, le son d'origine peut être monophonique ou avoir déjà subi un traitement stéréophonique se trouvant alors sur deux canaux ou plus. L'exemple présenté offre en sortie deux canaux et prend en entrée un canal.

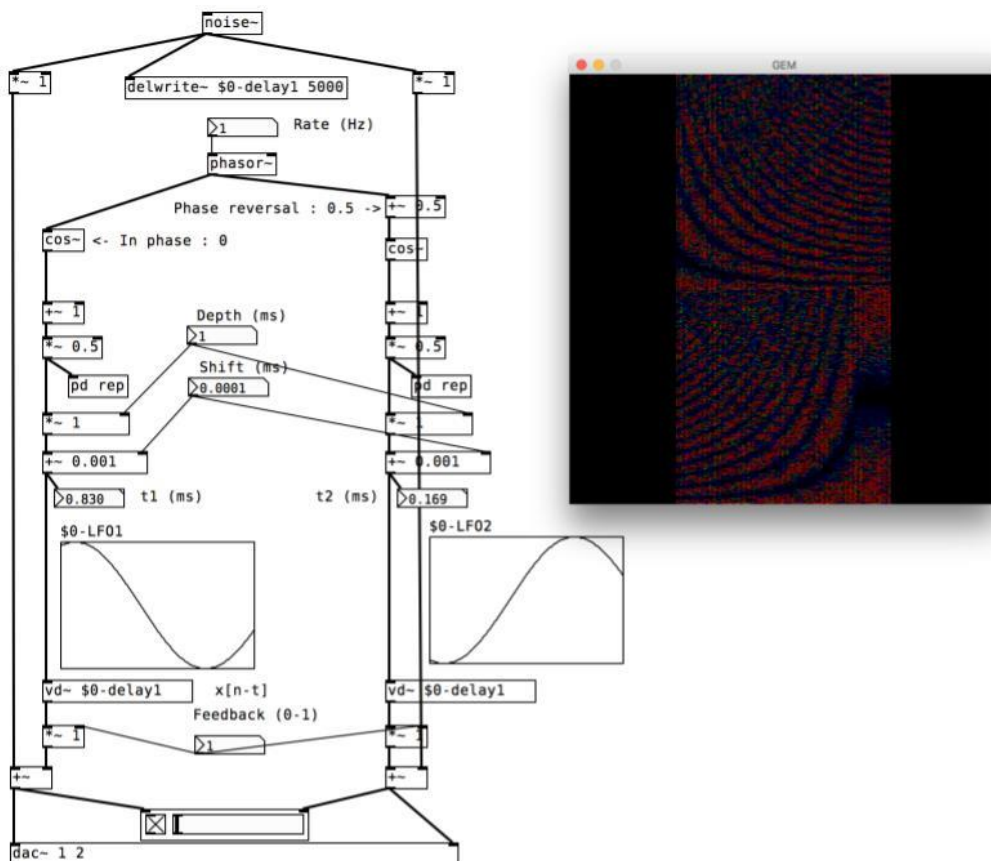


Figure 4.21. Patch Pure Data réalisant un flanging stéréophonique sur un bruit blanc. Le résultat de l'effet est représenté à droite dans la fenêtre GEM où les balayages inversés des fréquences selon les deux canaux sont visibles sur le sonagramme.

L'effet de balancement est créé en décalant respectivement les phases  $\varphi$  initiales des LFO contrôlant les temps de retard du signal réinjecté tels que

$$y[n] = x[n] + ax\left[n - \frac{\cos(2\pi fnT + \varphi) + 1}{2}p + o\right]$$

Équation 4.14. Effet de flanging avec contrôle de la phase du LFO.

Aussi, en considérant un LFO de référence, modifier la phase d'un LFO associé à une autre voix peut permettre de plus ou moins décaler les oscillations des temps de retard. Lorsque les deux LFO sont en décalage de phase, le *flanging* donne alors un effet de balancement des balayages fréquentiels entre les canaux. Généralement associé à deux canaux, gauche et droite, l'effet attendu sera un balayage bas-haut-bas-haut-etc. sur un canal en même temps qu'un balayage haut-bas-haut-bas-etc. sur le deuxième canal, donnant à l'effet une dimension spatiale. A l'opposé, lorsque les LFO sont en phase, le *flanging* retrouve les mêmes caractéristiques spatiales que l'approche monophonique.

### 4.5.3. Flanging enveloppant

Partant de ce principe et avant d'essayer d'étendre l'approche au domaine des harmoniques sphériques, il est intéressant dans un premier temps de considérer une approche similaire, pour un système de restitution enveloppant l'auditeur. De nombreuses approches peuvent être envisagées, mais l'idée est que l'ouverture de l'espace sonore pour une approche stéréophonique, dépend des décalages de phase des LFO des modules de *flanger* associés à chaque canal. Aussi, extrapoler cette approche selon un système immersif enveloppant l'auditeur pourrait permettre d'associer à chaque canal de restitution, un module de *flanger* dont la phase initiale serait fonction de la direction associée au canal.

Afin de relativement limiter les possibilités, il faut restreindre l'approche à une représentation bidimensionnelle de l'espace et envisager un système de haut-parleurs répartis sur un cercle. Une première expérience pourrait être de faire en sorte que les LFO des modules de *flanger* associés à des haut-parleurs diamétralement opposés soient en opposition de phase, et que les phases des LFO des modules de *flanger* associés à des haut-parleurs intermédiaires soient des valeurs intermédiaires entre ces deux états. Aussi est-il simplement envisageable d'associer la direction d'un haut-parleur en radians directement à la phase du LFO du module de *flanger* qui lui est associé. Dans ce type d'approche l'effet serait, selon le dispositif de haut-parleurs bien

entendu, une faible variation de l'effet de balayage pour des haut-parleurs relativement proches, mais un effet de balancement pour les haut-parleurs diamétralement opposés, tels que pour des binômes avant-arrière ou gauche-droite. Une approche un peu plus dynamique pourrait aussi permettre de déplacer dans l'espace les phénomènes d'opposition de phase des LFO des modules de flanging, en intégrant un indice de multiplication  $i$  au calcul de la phase  $\varphi$  à partir de la direction  $\theta$  d'un haut-parleur tel que  $\varphi = i\theta$  amène à concevoir le signal  $y$  émit par un haut-parleur tel que

$$y[n] = x[n] + ax[n - \frac{\cos(2\pi fnT + i\theta) + 1}{2}p + o]$$

Équation 4.15. Effet de flanging avec contrôle de la phase en fonction de l'azimut et d'un facteur de multiplication.

Cette approche permettrait par exemple que l'opposition de phase des LFO n'apparaisse plus à 180 degrés mais à 90 si  $i$  est égal à 2 ou 45 degrés si  $i$  est égal à 4. Aussi, un indice nul permettrait de remettre tous les haut-parleurs en phase et de perdre la notion de *flanging* spatial.

#### 4.5.4. Flanging ambisonique

L'approche suivante est de se demander en quoi consisterait alors un *flanging* ambisonique. L'idéal serait, comme pour les autres approches de synthèse et de traitement dans le domaine des harmoniques sphérique, de pouvoir faire émerger au sein du traitement les caractéristiques spatiales associées à cette décomposition spécifique de l'espace, tout en préservant les caractéristiques sonores du *flanging*. L'idée est que la mise en œuvre d'une simple décorrélation temporelle dans le domaine des harmoniques sphériques, via l'utilisation uniquement d'une ligne à retard, engendre une série de filtres en peigne dont les coefficients de réinjections et les temps de retard varient dans l'espace selon les propriétés des harmoniques sphériques. Alors une opération simple de mise en œuvre dans le domaine des harmoniques sphériques pourrait amener un effet de *flanging* spatial au moment de la restitution. Aussi, la mise en place d'un *flanger* ambisonique apparaît être un choix judicieux en vue de l'objectif final. En effet, comme cela vient d'être présenté, l'effet de *flanging* peut être assimilé à un filtre en peigne à retard variable, dont les variations dépendent d'un LFO. Etant donné que la mise en œuvre d'une décorrélation temporelle des signaux génère un filtre en peigne dans l'espace, envisager de contrôler les variations du temps de retard par un LFO devrait générer à la fin un effet de *flanging*.



Suite à l'analyse de la mise en œuvre de la synthèse par décorrélation temporelle des signaux associés aux harmoniques sphériques, il a été remarqué que le signal associé à l'harmonique de degré 0, par son caractère omnidirectionnel, contribue de manière homogène sur la sphère. Et dès lors qu'il est associé à un temps de retard nul, il permet d'obtenir une constante dans toutes les directions telle que le signal projeté soit défini par :

$$y_{0,0}^{\theta,\phi}[n] = Y_{0,0}(\theta, \phi)x[n] = x[n].$$

Équation 4.16. Contribution sur la sphère du signal associé à l'harmonique de degré 0 pour un traitement de *flanging* ambisonique.

Aussi, une telle approche attribuant le signal original au signal associé à l'harmonique de degré 0, permet de construire le premier terme de l'équation du filtre en peigne dans toutes les directions. Il ne s'agit plus que d'utiliser les autres signaux associés aux harmoniques pour définir les termes successifs consistant en une version pondérée et retardée du signal original. En reprenant la formulation du *flanging* défini précédemment [Équation 4.13], il est possible de définir que les signaux associés aux harmoniques de degrés non nuls doivent donc, suite à leur projection dans une direction donnée, offrir un signal tel que :

$$y_{l,m}^{\theta,\phi}[n] = ax \left[ n - \frac{\cos(2\pi fnT) + 1}{2} p + o \right].$$

Équation 4.17. Contribution sur la sphère d'un signal associé à un harmonique de degré non nul pour un traitement de *flanging* ambisonique.

Cette formulation définie, il s'agit alors de concevoir que ce signal doit être le résultat du signal associé à un harmonique sphérique pondéré par l'amplitude de cet harmonique sphérique dans la direction de projection. L'enjeu est alors d'intégrer l'amplitude de l'harmonique  $Y_{l,m}(\theta, \phi)$  dans l'équation. En reprenant l'approche proposée lors de la création de la synthèse de champ sonore par décorrélation temporelle, il est possible d'envisager que le coefficient  $a$  de réinjection soit défini par l'amplitude  $Y_{l,m}(\theta, \phi)$  de l'harmonique sphérique, selon la direction de projection permettant de redéfinir la contribution du signal associé à un harmonique sur la sphère par

$$y_{l,m}^{\theta,\phi}[n] = Y_{l,m}(\theta, \phi)x \left[ n - \frac{\cos(2\pi fnT) + 1}{2}p + o \right].$$

Équation 4.18. Contribution sur la sphère d'un signal associé à un harmonique de degré non nul pour un traitement de *flanging* ambisonique en remplaçant le coefficient de réinjection par l'amplitude de l'harmonique dans la direction de projection.

Cette approche offre alors la propriété intéressante de faire varier dans l'espace le coefficient de réinjection du filtre en peigne, en fonction des caractéristiques des harmoniques sphériques. Aussi, pour l'exemple, il peut déjà être envisagé que la contribution, résultant du signal associé à l'harmonique de degré 1 et d'ordre azimutal 1, possède, pour une élévation nulle, un coefficient de réinjection de 1 à l'azimut 0, 0 à l'azimut 90 degrés, -1 à l'azimut 180° et de nouveau 0 à l'azimut 270 degrés.

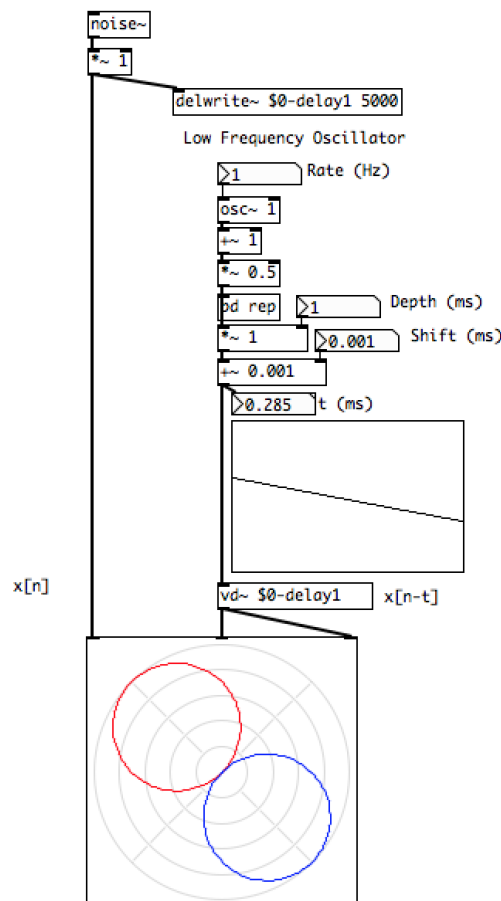


Figure 4.22. Patch Pure Data réalisant un *flanging* dans le domaine des harmoniques circulaires au premier ordre de décomposition.

Avant de complexifier l'approche, en essayant d'y intégrer les spécificités de la démarche stéréophonique, il est intéressant d'appréhender le résultat d'une telle mise en œuvre. Afin de

faciliter la compréhension, le contexte se limite à une représentation bidimensionnelle et à un ordre de décomposition de 1. Il s'agit alors d'associer le signal original au signal associé à l'harmonique de degré 0, et le signal original retardé aux signaux associés aux harmoniques de degré 1 [Figure 4.22]<sup>179</sup>. Alors le signal résultant d'une projection peut être facilement défini sous la forme d'une combinaison linéaire de signaux associés aux harmoniques sphériques tels que

$$y^\theta[n] = x[n] + (Y_{1,-1}(\theta) + Y_{1,1}(\theta))x\left[n - \frac{\cos(2\pi fnT) + 1}{2}p + o\right].$$

Équation 4.19. Contribution sur la sphère d'un champ sonore synthétisé par un *flanging* ambisonique au premier ordre de décomposition.

Suivant cette approche, le résultat est un effet de *flanging* dont le coefficient de réinjection est égal à la somme des amplitudes des harmoniques sphériques dans cette direction. Dès lors, en additionnant au préalable les amplitudes des harmoniques sphériques, il est possible d'obtenir une fonction sur la sphère représentant le coefficient de réinjection dans toutes les directions, telle que représentée sur la [Figure 4.22].

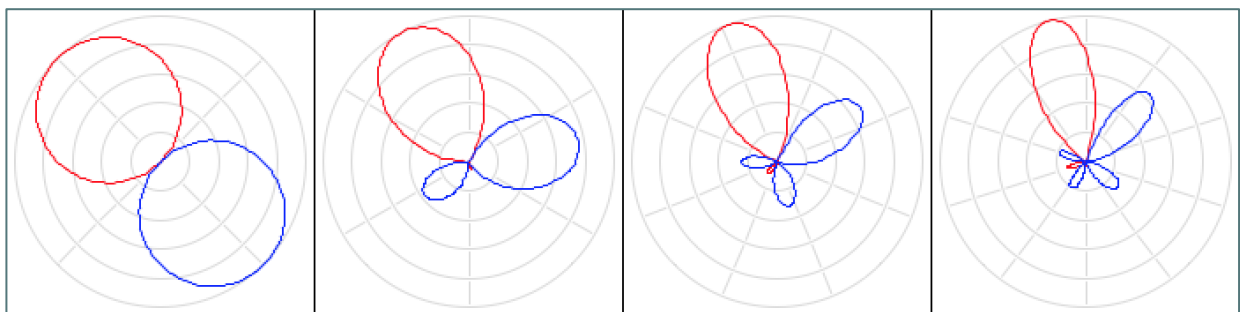


Figure 4.23. Représentation des figures des coefficients de réinjection définie selon l'approche du *flanging* ambisonique et une restitution circulaire. Avec de gauche à droite les figures associées aux ordres de décomposition 1, 2, 3 et 4.

Néanmoins, force est de constater que l'approche ambisonique du *flanging* offre un résultat spatial relativement pauvre. Selon les figures, en reprenant l'exemple du premier ordre de décomposition, il est possible de prévoir qu'un haut-parleur situé à 45 degrés offrira un effet de *flanging* marqué. Des haut-parleurs à 135 ou 315 degrés quant à eux n'offriront aucun effet de

<sup>179</sup> Dans cette approche, la normalisation du signal associé à l'harmonique de degré 0 est éludée afin de simplifier l'explication.

*flanging*. Un haut-parleur à 225 degrés offrira un effet de *flanging* similaire à celui de 45 degrés, mais avec un filtrage inversé<sup>180</sup>. Ce type de procédé ne permet pas d’espérer un effet spatial très marqué, surtout que les figures ont tendance à engendrer moins de variations dès lors que l’ordre de décomposition augmente. Il paraît donc nécessaire d’améliorer la « spatialité » de l’approche.

#### 4.5.5. Amélioration de l’approche

Afin d’améliorer le rendu spatial du *flanging* ambisonique, il est possible de s’inspirer de la démarche proposée par le *flanging* stéréophonique. L’idée pour créer de la spatialité est de jouer sur les décalages de phases des LFO de chacun des modules de *flanging*. En élargissant cette approche au domaine des harmoniques sphériques, il est possible de définir un *flanging* ambisonique qui associe un module de *flanging* à chaque signal associé à un harmonique sphérique et dont la phase initiale de chaque LFO est indépendante à chaque module.

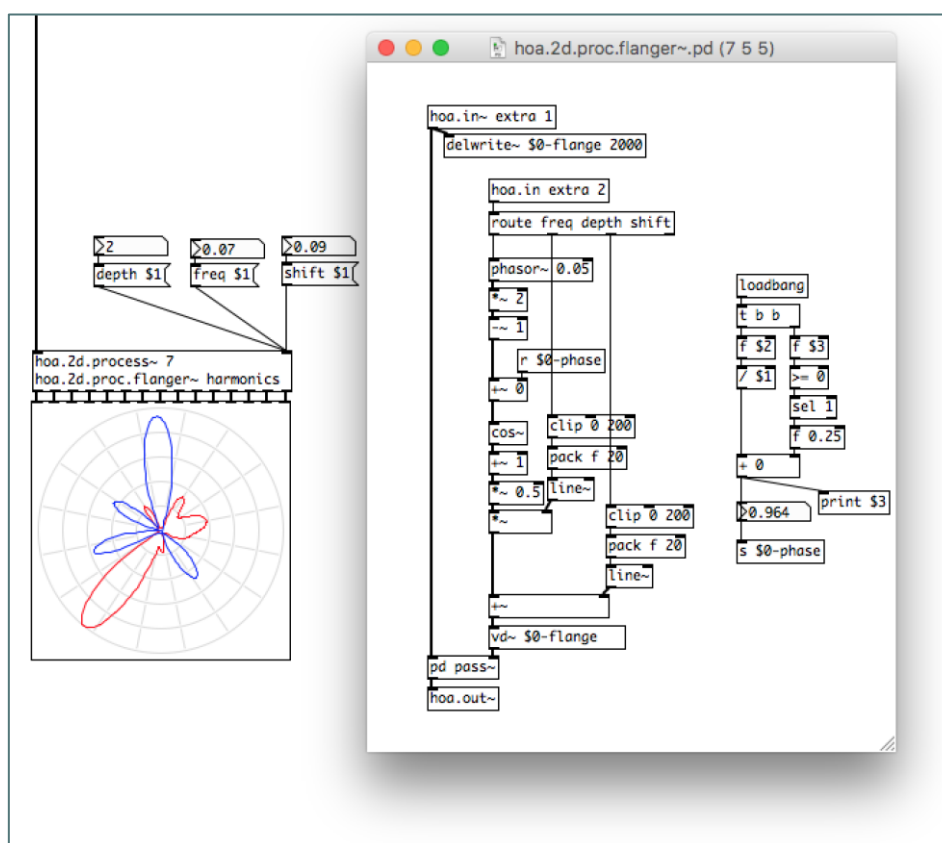


Figure 4.24. Patch Pure Data de synthèse de champ sonore par *flanging* ambisonique dans le domaine des harmoniques circulaires à l’ordre de décomposition 7 avec l’abstraction *hoa.2d.proc.flanger~* chargée dans l’objet *hoa.2d.process~*. A droite, l’instance d’abstraction *hoa.2d.proc.flanger~* associée à l’harmonique de degré 5 et d’ordre azimutal 5.

<sup>180</sup> Les bandes de fréquences atténuées sont augmentées et inversement les bandes de fréquences augmentées sont atténuées.

Comme pour les précédentes approches, l'enjeu est de définir un ensemble de variables, en l'occurrence les phases des LFO des modules de *flanging*, selon les caractéristiques des harmoniques sphériques. L'approche s'inspire alors de la proposition du flanging *enveloppant* consistant à associer à une direction de l'espace, sous la forme d'un angle en radians et un décalage de phase lui aussi en radian. Evidemment, il n'est pas question d'associer une direction avec un harmonique mais au degré et à l'ordre azimutal [Figure 4.24] tels que

$$\varphi_{l,m} = \frac{2\pi l}{N} - \begin{cases} m < 0 \Rightarrow \frac{\pi}{2} \\ m \geq 0 \Rightarrow 0 \end{cases}$$

Équation 4.20. Phase du LFO d'un module associé à un harmonique sphérique.

Cette approche, pensée pour une restitution bidimensionnelle, associe le signe de l'ordre azimutal à un décalage de phase de  $-\frac{\pi}{2}$  permettant de considérer la fonction cosinus du LFO en une fonction sinus en écho à la partie azimutale de la formulation des harmoniques sphériques [Équation 2.8]. Le rapport de  $\frac{l}{N}$  fait aussi écho à cette même formulation, en référence à l'indice de multiplication dépendant de degré  $l$ . Enfin, le signal résultant d'une projection peut être facilement défini sous la forme d'une combinaison linéaire des signaux associés aux harmoniques sphériques tels que :

$$y^{\theta,\phi}[n] = x[n] + \sum_{l=1}^N \sum_{m=-l}^l Y_{1,1}(\theta, \phi) x \left[ n - \frac{\cos(2\pi f n T + \varphi_{l,m}) + 1}{2} p + o \right]$$

Équation 4.21 : Contribution sur la sphère d'un champ sonore synthétisé par un *flanging* ambisonique.

Le signal résultant de la projection dans une direction est donc un *multi-flanger*, dans le sens où au signal original est réinjecté plusieurs versions retardées de lui-même. Le nombre de versions retardées de ce signal correspond au nombre d'harmoniques sphériques moins un. Leurs temps de retard varient selon des LFO dont les phases initiales dépendent de la formulation des harmoniques sphériques. Et les coefficients de réinjection dépendent des amplitudes de ces harmoniques dans la direction de la projection. Le résultat est alors un champ

diffus possédant des caractéristiques fréquentielles du *flanging*<sup>181</sup>. Une autre approche pourrait être de remplacer directement la fonction sinusoïdale du LFO par l'harmonique sphérique associé. Une telle approche permettrait que les balayages fréquentiels soient moins linéaires, et que leurs fréquences augmentent avec les degrés des harmoniques. Enfin, selon l'approche précédente, il est possible d'envisager de passer linéairement d'une source directionnelle à un champ diffus.

## 4.6. Bilan

La première partie de ce chapitre consacré à l'appropriation musicale de l'approche ambisonique, a permis de présenter plusieurs mises en œuvre originales réalisées dans le cadre de cette thèse. Ces mises en œuvre sont sous la forme de modules de synthèse et de traitement de champs sonores dans le domaine des harmoniques sphériques, et permettent des opérations telles que la variation de la résolution angulaire, la synthèse de champ sonore par décorrélation temporelle ou par granulation, l'encodage chaotique et le *flanging* ambisonique. Ces présentations ont permis de mettre en avant une approche particulière du modèle ambisonique et de révéler les possibilités musicales qui en découlent.

Aussi, la mise en œuvre d'un traitement de variation de la résolution angulaire du champ sonore a permis de présenter une première prise en main du modèle des harmoniques sphériques, où l'un des principaux enjeux était de concevoir comment chaque harmonique contribue à la représentation de l'espace, selon le degré auquel il est associé. La création d'un module de synthèse de champ sonore par décorrélation temporelle a quant à elle servi à concevoir comment jouer sur les propriétés spatiales des harmoniques sphériques. Il s'agissait d'élaborer la mise en œuvre, au regard des relations de symétrie des harmoniques sphériques, de leurs caractéristiques directionnelles et des plans ou axes auxquels ils peuvent être associés, tels que le plan horizontal ou l'axe vertical. Cette approche a aussi soulevé les premières questions relatives à la méthodologie à appliquer pour mettre en œuvre un traitement original dans le domaine des harmoniques sphériques. Notamment en posant la question de la manière de concevoir et envisager le résultat à la fois spatial et sonore. Ce travail a été continué par la suite dans la mise en œuvre d'un module de synthèse granulaire quasi-synchrone, en élaborant une approche différente de l'usage des propriétés spatiales des harmoniques en relation avec les

---

<sup>181</sup> Il est à noter que les temps de retard utilisés pour le *flanging* sont relativement petits afin de créer des filtres fréquentiels engendrant un effet d'espace plus restreint que l'utilisation de la synthèse par décorrélation temporelle. Aussi des temps de retard plus grands amèneraient à un effet plus proche d'un *chorus*.

caractéristiques sonores du traitement. Un autre point essentiel a aussi été soulevé plus en détails : celui de la qualité du rendu spatial et sonore, lié notamment à l'ordre de décomposition. La notion de « qualité du rendu sonore » est pour les champs sonores diffus et musicaux ce qu'est la « résolution angulaire » pour les sources sonores directionnelles. Ce qui amène alors à la quatrième expérimentation, qui vise à générer un champ sonore original en réalisant une pondération aléatoire des signaux associés aux harmoniques sphérique. Celle-ci offre la possibilité de convertir le champ sonore chaotique en une source sonore directionnelle. Au-delà du potentiel artistique, cette approche aura aussi permis de révéler la relation hiérarchique impliquée par la décomposition en harmoniques sphériques, dans l'établissement d'une directivité dans le champ sonore. Enfin la dernière approche, forte de l'expérience acquise, a proposé de prendre le contre-pied méthodologique. Via la mise en œuvre d'un *flanging* ambisonique, l'objectif n'était plus de partir d'une opération pour voir par la suite quel traitement de l'espace et du son sa mise en œuvre dans le domaine des harmoniques sphérique occasionne. Il s'agissait plutôt de fixer un résultat sonore, et de voir quelle opération dans le domaine des harmoniques sphériques doit être mise en œuvre pour y arriver.

Finalement, ces approches ont permis de révéler deux méthodologies distinctes de la création de traitements et de synthèses dans le domaine des harmoniques sphériques. La première, plus expérimentale et incertaine, consiste à prendre une opération de traitement du signal, la mettre en œuvre dans le domaine des harmoniques sphériques. Puis il s'agit de comprendre et analyser sur le plan du son et de l'espace, les mécanismes en jeu. La deuxième méthodologie, plus rationnelle et rigoureuse, consiste à définir un résultat sur le plan sonore et de le mettre en espace via le modèle ambisonique, en prenant en considération les caractéristiques particulières de la décomposition de l'espace qui lui est associée.

Ces méthodologies ont permis de définir deux spécifications principales. La première spécification établit qu'un traitement doit pouvoir s'adapter à tout ordre de décomposition. En effet, il doit pouvoir reproduire des mécanismes similaires, quel que soit l'ordre de décomposition. Mais il doit aussi tirer parti de l'ensemble des composantes offertes. Aussi, la « qualité » du traitement doit dépendre de l'ordre de décomposition au même titre que la résolution angulaire en dépend. D'ailleurs, la résolution angulaire peut être considérée comme la « qualité » du traitement d'encodage directionnel. Cette propriété a été présentée pour l'ensemble des traitements, que ce soit la décorrélation temporelle, la variation de la résolution angulaire, le *flanging* ambisonique, etc. La deuxième spécification est qu'un traitement doit pouvoir s'adapter à une représentation de l'espace à deux et trois dimensions. Cela est *a priori*

vrai pour l'ensemble des traitements mis en œuvre pour une restitution tridimensionnelle et il est possible d'envisager que cela le soit aussi pour les autres. Néanmoins les contraintes sont ici plutôt d'ordres techniques liés à la difficulté de mettre en œuvre certaines opérations comme les calculs récursifs dans un patch, freinant l'utilisation des polynômes de Legendre par exemple.

Au-delà des spécifications, ces approches liées aux caractéristiques spatiales des harmoniques sphériques offrent de nouvelles perspectives quant à l'appropriation musicale du modèle ambisonique. Il s'agit alors de jouer selon les opérations mises en œuvre, en prenant en compte les caractéristiques énumérées des harmoniques sphériques et de la décomposition qui leur est associée, telles que leurs contributions à la résolution angulaire selon leurs degrés, leurs figures de directivité, etc. Evidemment, les jeux et relations proposées ne sont que des exemples, d'ailleurs, à chaque mise en œuvre, d'autres propositions sont aussi envisagées. En somme, les possibilités de jeu sont ici infinies. Cela amène à se poser la question, si dans ce contexte, via l'appropriation expérimentale et musicale de l'ambisonie, le réel intérêt de ce modèle ne s'en trouve pas changé. Et s'il ne réside pas à présent dans le potentiel musical et spatial qu'offre la décomposition en harmoniques sphériques. Comme révélé, lors de la mise en œuvre de la décorrélation, l'un des éléments importants dans cette approche est de considérer la décomposition en harmoniques sphériques comme un outil structurant la pensée musicale de l'espace. Cela amène à considérer si l'utilisation de la représentation du champ sonore via la décomposition en harmoniques sphériques définie en ambisonie pour répondre à des problématiques matérielles, acoustiques et psychoacoustiques est réellement pertinente dans ce contexte. Il s'agit de voir si en soi, cette représentation intermédiaire de l'espace n'est pas le centre réel du sujet. Dans ce cas, les problématiques auxquelles elle est censée répondre peuvent paraître obsolètes.

Enfin, ces approches originales du modèle ambisoniques ont surtout défini une manière de penser le son et l'espace conjointement, permettant de créer de nouveaux espaces sonores décorrélés, granulés, chaotiques, etc. Ces nouveaux espaces sonores amènent alors de nouvelles questions quant à leur analyse, leur critique et leur restitution et nécessite peut-être, comme cela vient d'être suggéré de modifier le modèle ambisonique ou du moins, d'y apporter un nouveau regard et de nouveaux outils.



## 5. Adaptation des outils

Dans la partie précédente, un certain nombre de traitements originaux de l'espace et du son en ambisonie ont été présentés. Ce travail de recherches et d'expérimentations a permis au travers de deux méthodologies distinctes, de faire ressortir des spécifications fonctionnelles quant aux traitements mis en œuvre. Cela a aussi permis de mettre en avant un certain nombre de méthodes, faisant office de suggestions permettant de répondre aux questions soulevées. Ces nouvelles approches ont créé de nouveaux champs sonores, qui ne sont plus seulement composés de sources sonores directionnelles, mais aussi de masses, de grains, de sources aux figures de directionnalité variées et aux caractères hétéroclites.

Aussi une telle approche implique forcément des répercussions sur le modèle ambisonique même. Le modèle et les outils ambisoniques ont été construits sur des principes acoustiques et psychoacoustiques, visant à assurer une reconstruction cohérente pour l'auditeur d'une source sonore directionnelle. Mais dès lors que d'autres approches sont mises en œuvre dans le but de concevoir différents types de sources sonores et de nouveaux champs sonores, alors les outils proposés initialement par le modèle peuvent ne plus répondre correctement à ce contexte. Il est donc nécessaire dans ce cas d'en concevoir de nouveaux.

Cette partie propose de revenir sur les outils et les opérations mathématiques définies en ambisonie et permettant d'analyser, de critiquer et de restituer des champs sonores. L'enjeu est de comprendre comment ces outils et ces opérations ont été élaborés d'un point de vue technique et comment ils répondent aux principes acoustiques et psychoacoustiques établis en ambisonie. Il sera alors possible de confronter ces outils et opérations aux nouvelles approches proposées dans la partie précédente et de critiquer dans ce contexte leur validité vis à vis des nouvelles problématiques apparues dans le cadre de ces recherches. Au regard de ces résultats, de nouveaux outils et de nouvelles opérations permettant l'analyse, la critique et la restitution du champ sonore pourront être envisagées afin de répondre à ces problématiques.

Ce travail sera réalisé selon deux axes, à savoir l'analyse du champ sonore et la restitution du champ. Le premier axe se penche sur les outils permettant de mettre en relief et de critiquer les opérations proposées dans le deuxième axe. Enfin, il aurait été intéressant d'introduire un troisième axe d'étude : celui de l'adaptabilité des processus mis en œuvre. Il s'agirait de pouvoir présenter des solutions permettant de répondre aux problèmes de changements de

représentation. Notamment, l'adaptation des mises en œuvre destinées à une représentation bidimensionnelle de l'espace vers une représentation tridimensionnelle ou l'adaptation inverse. Une première étude et des propositions ont été réalisées à ce sujet notamment avec A. Bonardi dans le cadre de sa pièce *Pianotronics 2* [Bonardi & al, 2015], mais il s'agit encore de généraliser l'approche. Ce domaine de recherche recoupe aussi la question de l'adaptabilité des mises en œuvre et des champs sonores à différents ordres de décomposition. Des réponses à ce vaste problème ont déjà été apportées dans le cadre des expérimentations présentées dans la partie précédente. D'autres propositions permettront aussi de résoudre ce problème lorsque la question de la restitution de champ sonore sera soulevée.

## 5.1. Analyse du champ sonore

L'analyse de la restitution d'un système de spatialisation du son est une étape essentielle de la mise en œuvre d'un tel système, car elle permet, au-delà de valider ou invalider une approche, de pouvoir en révéler certaines caractéristiques et certains défauts. Selon les résultats, il est alors possible de critiquer l'approche sur des aspects spécifiques, permettant alors d'envisager des solutions ou de contourner les problèmes afin d'améliorer le rendu sonore. Dans le cadre de ces recherches, offrir des outils d'analyse critique permettant de mieux prendre en main les propositions originales de synthèse et transformation de champ sonore, est un élément essentiel.

L'analyse et la critique de la restitution d'un système de spatialisation du son peuvent être réalisées de différentes façons. Soit de façon empirique, suite à l'expérimentation et l'écoute, si possible en essayant de recouper les avis afin d'offrir des résultats plus objectifs. Soit de façon systématique, en utilisant un modèle d'analyse construite sur des données mesurées et rationnelles. Il n'est aucunement question dans ce travail de prendre parti pour une approche plutôt qu'une autre car ces deux approches peuvent d'ailleurs se compléter. Les expérimentations qui se sont déroulées dans le cadre de cette thèse ont été confrontées d'une part à une analyse et une critique empirique et personnelle, puis ont été confrontées à d'autres avis. Les résultats semblaient satisfaisants, notamment via la publication des outils et des traitements amenant à leurs utilisations possibles dans des œuvres musicales. Et d'autre part, ces expérimentations ont été analysées d'une manière plus systématique, en utilisant des modèles mathématiques, acoustiques et psychoacoustiques, afin de tenter de caractériser de manière rationnelle le champ sonore. L'idée est de pouvoir déterminer, dans le cas d'une expérience satisfaisante, ce qui fait que cette approche soit pertinente sur un plan formel, et dans le cas contraire, d'essayer de déterminer les raisons de sa faillite. Cependant, il est très difficile de présenter la méthode d'analyse par les sens. Bien que la critique et la validation

d'une expérimentation soient les résultats de l'observation et de l'écoute de personnes multiples, il faut avouer que dans le cadre de ces recherches, les analyses ont toujours eu lieu de manière très subjective et leurs fondements restent relativement vagues. En somme, cela reviendrait à définir la méthode selon laquelle un panel d'auditeurs finisse par indiquer à propos d'un morceau de jazz que « Ça tourne ! » ou non. Retrouver les implications de telles analyses, s'il est possible de les nommer ainsi, dépasse le contexte de ces recherches. Aussi est-il plus intéressant et pertinent ici de définir une méthode d'analyse systématique permettant, si ce n'est de prévoir le rendu sonore avec exactitude, au moins d'offrir des outils représentant certains éléments caractéristiques du champ sonore. De plus, les approches musicales et expérimentales présentées dans cette thèse prennent place dans des contextes de restitution très variables, que ce soit par les lieux, les systèmes de restitution ou par l'auditoire. Là, une méthode empirique ne serait valide que pour un contexte donné, ou à la rigueur pour d'autres systèmes très similaires. La méthode systématique offre par sa nature une capacité d'adaptation beaucoup plus grande, qui convient mieux à ce contexte.

La question importante est alors de définir une ou des méthodes d'analyse systématique qui pourraient répondre à la fois aux problématiques « classiques » de simulation de source sonore directionnelle et potentiellement ponctuelle, mais aussi aux nouvelles approches de synthèse de champ sonore, présentées dans le cadre de ces recherches. Pour cela, ce travail propose dans un premier temps de revenir sur les outils psychoacoustiques que sont les vecteurs énergie et vitesse, permettant d'analyser et de critiquer un système de spatialisation construit autour de la notion de sources directionnelles. L'enjeu étant alors de contextualiser leur utilisation et d'essayer de les améliorer, afin de proposer de nouveaux outils plus adaptés à l'approche musicale et expérimentale. Tout en restant dans le paradigme des sources directionnelles. Par la suite, en mettant en relation ces propositions et les recherches originales réalisées sur la synthèse de champ sonore, l'objectif sera de soumettre une nouvelle utilisation des vecteurs énergie et vitesse, en relation avec les problématiques soulevées par l'approche musicale de l'ambisonie<sup>182</sup>.

---

<sup>182</sup> L'ensemble des livrables sont disponibles pour Pure Data dans la bibliothèque HoaLibrary-Analyser. Cependant, il est important de noter que ces outils sont expérimentaux et qu'ils ne sont pour le moment destinés qu'à une parution sous la forme d'une version *alpha* ou *bêta*. Pour cette raison et parce que la bibliothèque HOA ne possède aucune dépendance à ces outils, proposer ces outils sous la forme d'une bibliothèque annexe semble plus judicieux.

### 5.1.1. Vecteurs énergie et vitesse

Comme présenté précédemment, il existe deux outils proposés par M. A. Gerzon permettant de décrire l'effet de localisation dans le domaine des basses-fréquences avec le *vecteur vitesse*  $\vec{V}$  et dans le domaine des hautes-fréquences avec le *vecteur énergie*  $\vec{E}$  définis mathématiquement en fonction des gains  $g$  et des directions  $\vec{u}$  associés aux haut-parleurs tels que

$$\vec{V} = \frac{\sum_{i=1}^N g_i \vec{u}_i}{\sum_{i=1}^N g_i}$$

et

$$\vec{E} = \frac{\sum_{i=1}^N g_i^2 \vec{u}_i}{\sum_{i=1}^N g_i^2}.$$

Équation 5.1. Équations des vecteurs vitesse et énergie.

Le vecteur vitesse revient donc à calculer la somme des directions des haut-parleurs pondérées par leurs gains respectifs, et à normaliser ce résultat par la somme des gains des haut-parleurs. Alors que le vecteur énergie revient à calculer la somme des directions des haut-parleurs pondérées par leurs énergies respectives, le carré de leurs gains revient à normaliser ce résultat par la somme des énergies des haut-parleurs. Sur un plan pratique, il s'agit de représenter les directions associées aux haut-parleurs sous une forme cartésienne<sup>183</sup> et de calculer indépendamment l'abscisse, l'ordonnée et la hauteur des vecteurs tel que<sup>184</sup>

$$x_{\vec{V}} = \frac{\sum_{i=1}^N g_i x_{\vec{u}_i}}{\sum_{i=1}^N g_i}, y_{\vec{V}} = \frac{\sum_{i=1}^N g_i y_{\vec{u}_i}}{\sum_{i=1}^N g_i} \text{ et } z_{\vec{V}} = \frac{\sum_{i=1}^N g_i z_{\vec{u}_i}}{\sum_{i=1}^N g_i}$$

et

$$x_{\vec{E}} = \frac{\sum_{i=1}^N g_i^2 x_{\vec{u}_i}}{\sum_{i=1}^N g_i^2}, y_{\vec{E}} = \frac{\sum_{i=1}^N g_i^2 y_{\vec{u}_i}}{\sum_{i=1}^N g_i^2} \text{ et } z_{\vec{E}} = \frac{\sum_{i=1}^N g_i^2 z_{\vec{u}_i}}{\sum_{i=1}^N g_i^2}.$$

<sup>183</sup> Il est nécessaire de procéder aux calculs avec un système de coordonnées cartésiennes plutôt que polaire ou sphérique, afin d'éviter les effets de bord dus au caractère périodique des angles. En effet, calculer la moyenne de deux angles peut aboutir à des erreurs si leur écart absolu est supérieur à un angle de  $\pi$  radians.

<sup>184</sup> Cette opération est disponible dans la bibliothèque HoaLibrary-Analyzer via les abstractions `hoa.2/3d.proc.vector.velocity~` et `hoa.2/3d.proc.vector.energy~` destinées à une utilisation dans l'objet `hoa.2/3d.process~` ou les objets `hoa.2/3d.vector.velocity~` et `hoa.2/3d.vector.energy~`.

Comme défini dans le premier chapitre de cette thèse, le but de tout système de spatialisation pour un encodage directionnel est de parvenir à des vecteurs énergie et vitesse possédant une norme unitaire, et une direction identique à celle de l'encodage. Dès lors, il est nécessaire de transformer ces coordonnées cartésiennes en coordonnées polaires ou sphériques, selon l'approche présentée dans le chapitre précédent, afin d'obtenir les informations nécessaires à l'analyse du champ sonore<sup>185</sup>.

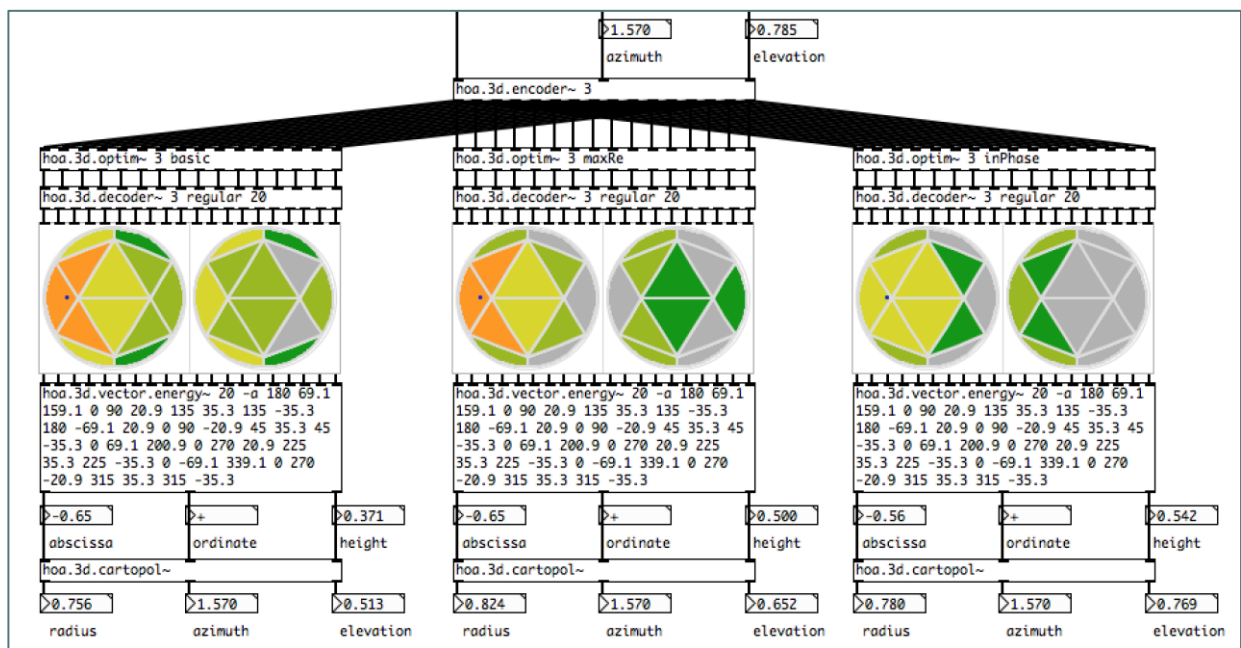


Figure 5.1. Patch Pure Data comparant les résultats du vecteur d'énergie, associé au domaine hautes-fréquences, pour une source sonore encodée à  $\frac{\pi}{2}$  radians d'azimut et  $\frac{\pi}{4}$  d'élévation à un ordre de décomposition 3 et restituée sur un système sphérique de vingt haut-parleurs formant un icosaèdre selon les trois modes d'optimisation *basic*, *max-re* et *in-phase*.

Dès lors, l'analyse consiste à comparer les coordonnées sphériques des vecteurs avec les coordonnées d'encodage de la source [Figure 5.1]. En considérant le domaine fréquentiel associé à chaque vecteur, la norme est représentative de la résolution angulaire avec une valeur de 1 équivalente à une onde plane, donc une résolution optimale, et une valeur de zéro équivalente à un champ sonore omnidirectionnel<sup>186</sup>. Les angles, azimut et élévation, permettent

<sup>185</sup> Cette conversion peut être réalisée avec l'abstraction *hoa.2/3d.cartopol~* de la bibliothèque d'analyse *Hoalibrary-Analyser*. Et l'opération réciproque peut être réalisée avec l'abstraction *hoa.2/3d.poltocar~*.

<sup>186</sup> Comme suggéré précédemment, cela en fait un outil idéal pour critiquer les algorithmes de variation de la résolution angulaire.

quant à eux de révéler des déformations respectivement sur le plan horizontal et le plan vertical. Ainsi l'azimut d'un vecteur révèle, dans le domaine fréquentiel associé, la position perçue de la source sonore sur le plan horizontal. Il en est de même avec l'élévation pour le plan vertical. Aussi, une variation des angles des vecteurs par rapport à ceux de l'encodage traduit une distorsion due soit à un système de restitution sous-optimal, soit à une opération sur le champ sonore. En exemple en [Figure 5.1], dans le domaine hautes-fréquences, le champ sonore ne semble pas déformé dans la direction d'encodage sur un plan horizontal, quelle que soit l'optimisation. Car les azimuts des vecteurs énergie sont tous égaux à l'azimut d'encodage. Mais le champ sonore est déformé sur le plan vertical, particulièrement avec les optimisations *basic* et *max-r<sub>e</sub>* et relativement faiblement avec l'optimisation *in-phase*, car les élévations des vecteurs énergie sont toutes plus ou moins inférieures à l'élévation d'encodage. Cependant cette optimisation *in-phase* diminue la résolution angulaire du champ sonore, au même titre que l'optimisation *basic* et de façon moindre pour l'optimisation *max-r<sub>e</sub>*. Les normes (ou rayons) des vecteurs énergies sont en effet tous inférieurs à 1. De ce fait dans cette direction, le système de restitution ne semble pas optimal mais l'optimisation *max-r<sub>e</sub>* semble préserver au mieux la résolution angulaire alors que l'optimisation *in-phase* limite les déformations sur le plan horizontal.

Aussi, en utilisant un système discrétisant parfaitement l'espace, ou de manière optimale dans le cas d'une représentation tridimensionnelle de l'espace et au-delà de l'ordre de décomposition 3, il est possible d'utiliser les coordonnées des vecteurs pour critiquer les opérations mises en œuvre. Réciproquement, en considérant une opération ou un ensemble d'opérations sur le champ sonore ainsi qu'un système de restitution existant, il est possible d'utiliser les informations des vecteurs. Cela permet de trouver la meilleure approche pour restituer le champ sonore, et d'adapter l'opération de décodage tout en appliquant une ou des optimisations différentes selon les domaines fréquentiels.

Un deuxième exemple en [Figure 5.2], monte que la première approche de restitution du champ sonore utilise une projection directe des composantes ambisoniques sur le système de haut-parleurs. La résolution angulaire est grandement diminuée et la source sonore est déplacée vers la scène frontale. La deuxième approche de restitution du champ sonore tente de compenser ces artefacts dus au système de restitution irrégulier, en appliquant une optimisation *in-phase* et un décodage dit *irrégulier*. La résolution angulaire est grandement améliorée, mais la source sonore est toujours déplacée cette fois-ci vers la scène arrière. Aussi dans ce contexte, le

jugement de ces deux analyses permet d'orienter le choix vers la deuxième approche, tout en continuant possiblement de l'améliorer en soumettant les analyses à d'autres optimisations.

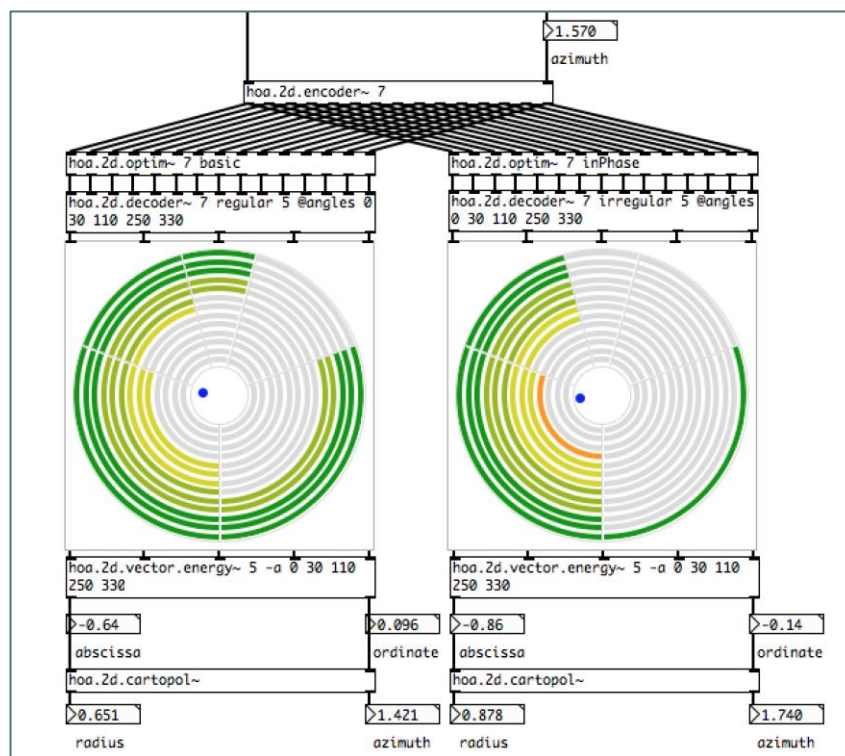


Figure 5.2. Patch Pure Data comparant les résultats du vecteur d'énergie, associé au domaine hautes-fréquences, pour une source sonore encodée à  $\frac{\pi}{2}$  radians d'azimut à un ordre de décomposition 7 et restituée sur un système circulaire de type 5.1 (le canal des fréquences basses étant ignoré) selon deux approches distinctes.

### 5.1.2. Valeurs significatives

Les vecteurs énergie et vitesse sont des outils puissants mais restant néanmoins à un trop bas niveau. Ils ne sont à eux seuls pas significatifs du rendu perceptif du champ sonore, et nécessitent d'être mis en relation avec une direction de référence. Puis, ils doivent être analysés en fonction de celle-ci et en considération des échelles sur lesquels les informations sont définies. Le problème est similaire sur celui d'une échelle de puissance qui ne prend pas en compte la perception humaine, pour représenter le niveau d'un signal sonore. Une solution dans ce cas est le décibel, une unité représentant un rapport de puissance entre la grandeur mesurée et celle d'une valeur de référence associée à un son inaudible et qui est construite sur une échelle logarithmique. Aussi, dans le cadre de ces recherches, trois outils mathématiques sont proposés afin d'offrir des valeurs plus justes de représentation du champ sonore. Ces outils qui peuvent être considérés comme des outils de macroanalyse construits sur les vecteurs énergie et vitesse,

permettent de définir la déviation angulaire, la diminution de la résolution et la variation globale de précision<sup>187</sup>.

La déviation angulaire  $\Delta_{\theta,\phi}$  est définie à partir de la distance du grand cercle<sup>188</sup> entre la direction d'encodage et la direction du vecteur. Elle permet de représenter, sous la forme d'un pourcentage, les distances combinées sur le plan horizontal et vertical. Avec sur le plan horizontal, la distance entre l'azimut de la direction escomptée  $\theta_e$  et l'azimut de la direction perçue  $\theta_v$ . Et avec sur le plan vertical, la distance entre l'élévation de la direction escomptée  $\phi_e$  et l'élévation de la direction perçue  $\phi_v$ . La déviation angulaire est définie telle que

$$\Delta_{\theta,\phi} = \frac{200}{\pi} \text{asin} \sqrt{\sin^2 \left( \frac{\phi_e - \phi_v}{2} \right) + \cos \phi_e \cos \phi_v \sin^2 \left( \frac{\theta_e - \theta_v}{2} \right)}$$

Équation 5.2. Equation de la déviation angulaire du champ sonore fondée sur le vecteur énergie ou le vecteur vitesse.

La diminution de la résolution  $\Delta_r$  équivaut simplement à la différence entre la norme du vecteur  $\rho_v$  et le rayon de la sphère ou du cercle ambisonique, considéré comme une unité de 1, et permet de représenter sous la forme d'un pourcentage la diminution de la résolution angulaire par rapport à une restitution optimale telle que

$$\Delta_r = 100(1 - \rho_v)$$

Équation 5.3. Equation de la diminution de la résolution du champ sonore fondée sur le vecteur énergie ou le vecteur vitesse.

La variation globale de précision  $\Delta_g$  équivaut à la distance euclidienne<sup>189</sup> entre la position associée au vecteur et la position d'encodage. En somme, cette variable prend en compte à la

---

<sup>187</sup> L'ensemble de ces outils sont disponibles via l'abstraction *hoa.2/3d.vector.analyser* dans la bibliothèque d'analyse HoaLibrary-Analyser.

<sup>188</sup> Le calcul de la distance du grand cercle est réalisé via l'abstraction *hoa.2/3d.distance.angular* dans la bibliothèque d'analyse HoaLibrary-Analyser.

<sup>189</sup> Le calcul de la distance euclidienne est réalisé via l'abstraction *hoa.2/3d.distance.euclidean* dans la bibliothèque d'analyse HoaLibrary-Analyser.



fois la déviation angulaire et la diminution de la résolution sous la forme d'un seul outil défini de même en pourcentage tel que

$$\Delta_g = \sqrt{(x_e - x_v)^2 + (y_e - y_v)^2 + (z_e - z_v)^2}$$

Équation 5.4 : Equation de la variation globale de précision du champ sonore fondée sur le vecteur énergie ou le vecteur vélocité.

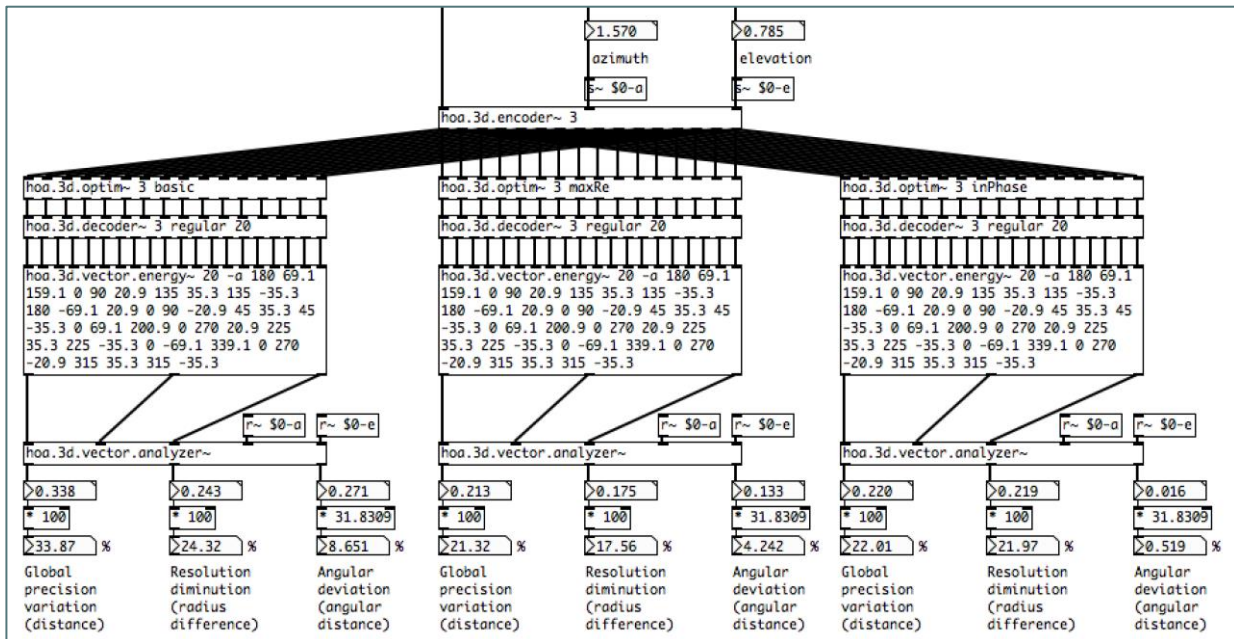


Figure 5.3. Patch Pure Data calculant les valeurs significatives d'analyse de champ sonore selon les mêmes configurations que définies par la [Figure 5.2].

En reprenant le dispositif présenté en [Figure 5.2], il est possible de définir bien plus facilement la nature des variations et les implications d'une approche ou d'une autre. Aussi, en utilisant exclusivement la valeur de la variation globale de la précision, il est possible de restreindre le choix de l'approche à utiliser afin d'offrir une restitution optimale sur l'optimisation *max-r<sub>e</sub>* ou *in-phase* dans le domaine hautes-fréquences. Il s'agit alors de comparer la diminution de la résolution et la déviation angulaire, afin que les avantages respectifs de ces deux dernières optimisations pour ce système de restitution soient clairs. Et facilitent ainsi les choix possibles [Figure 5.3]. Un auditoire élargi et recouvrant toute la zone de restitution souffrira sûrement plus de la déviation angulaire qu'un auditoire restreint au centre de la zone de restitution et qui sera plus sensible à une variation de la résolution angulaire. Aussi, le premier auditoire amène à utiliser une optimisation *in-phase* alors que le deuxième auditoire amène à utiliser une optimisation *max-r<sub>e</sub>*.

### 5.1.3. Analyse complète

Bien que pertinents pour l'analyse, ces valeurs significatives dérivent des vecteurs énergie et vitesse et ne sont valables que pour une direction donnée. Or, afin de pouvoir envisager dans sa globalité la restitution d'un champ sonore, l'utilisateur doit procéder à l'analyse dans de multiples directions. Cette opération peut se révéler extrêmement longue et laborieuse, donc peu propice à l'expérimentation de nouvelles approches. En considérant qu'il faille partir d'un système initial, la méthodologie revient à analyser le champ sonore selon les valeurs significatives, dans un ensemble de directions suffisamment important afin d'offrir une représentation de l'espace assez précise. Si les résultats se révèlent peu satisfaisants, il est alors nécessaire de mettre en œuvre une nouvelle approche telle que l'application d'une optimisation du champ sonore, ou encore une amélioration de l'opération de décodage. Ce changement nécessite alors de procéder de nouveau à une analyse du champ sonore dans de multiples directions. Puis de comparer ces résultats avec les premiers établis et de modifier l'approche en conséquence. Aussi, les opérations peuvent continuer pendant très longtemps, rendant la mise en œuvre d'un système optimal complètement chronophage. Ainsi, il semble particulièrement nécessaire d'offrir un outil permettant d'automatiser cette opération. A cet effet, dans le cadre de ces recherches a été élaborée une abstraction *hoa.2/3d.vector.recorder~*<sup>190</sup> permettant d'analyser le champ sonore de façon précise, sur l'ensemble du cercle ou de la sphère.

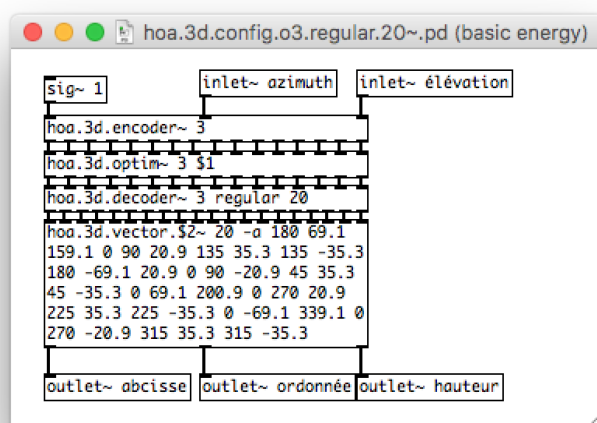


Figure 5.4. Exemple type d'un patch qui synthétise une source sonore directionnelle, décode le champ sonore et l'analyse pouvant être utilisée dans le système d'analyse globale du champ sonore *hoa.3d.vector.recorder~*.

<sup>190</sup> Cette abstraction, bien que fonctionnelle, n'est pas finalisée et peut potentiellement changer afin d'offrir une meilleure ergonomie.

Cet outil charge le patch permettant de synthétiser et décoder le champ sonore, et analyse le champ sonore suite à l'encodage d'une source dans un grand nombre de directions. L'enjeu est d'offrir un outil qui s'adapte à toutes les opérations de synthèse, de transformation et de décodage ainsi qu'à tous les ordres de décomposition. En même temps, l'approche nécessite que le patch chargé ait un nombre préétabli d'entrées et de sorties. Car *a priori*, le logiciel Pure Data ne permet pas plus de dynamisme. Aussi, pour le moment et compte tenu de la nature de l'approche, deux entrées doivent être présentes afin de recevoir respectivement les azimuts et l'élévation. Et trois sorties doivent être utilisées pour envoyer les coordonnées d'un des vecteurs selon la direction donnée. Le premier outil de la chaîne doit donc être un encodeur directionnel et le dernier outil de la chaîne doit calculer l'un des vecteurs du champ sonore. Entre ces deux étapes, le logiciel permet de réaliser toutes sortes d'opérations. Enfin, afin d'offrir un outil un peu plus dynamique, l'abstraction `hoa.2/3d.vector.recorder~` transmet ces deux derniers arguments au patch chargé permettant de les utiliser, afin de définir l'optimisation ou le vecteur à calculer [Figure 5.4].

```
hoa.3d.vector.recorder~ hoa.3d.config.o3.regular.20~ 64 64
$0-values basic energy
```

Figure 5.5. Exemple type de l'utilisation de l'abstraction `hoa.3d.vector.recoder~` avec comme argument le patch chargé `hoa.3d.config.o3.regular.20~`, le nombre d'angles discrétisant l'espace sur le plan horizontal et sur le plan vertical `64` et `64`, le nom de la mémoire tampon dans laquelle écrire les résultats `$0-values`. Et enfin les deux arguments supplémentaires et facultatifs donnés au patch chargé.

Les premiers arguments utilisés par l'abstraction définissent quant à eux le nom du patch à charger, puis le nombre d'angles discrétisant l'espace sur le plan horizontal et sur le plan vertical. Ils définissent aussi le nom de la mémoire tampon dans laquelle écrire les résultats. Afin de pouvoir écrire tous les résultats, il est d'ailleurs nécessaire que la taille de la mémoire tampon soit supérieure ou égale au nombre d'angles discrétisant l'espace sur le plan horizontal, multiplié par le nombre d'angles discrétisant l'espace sur le plan vertical. Des valeurs de 64 angles pour chacun semblent largement suffisantes à une bonne discrétisation. Néanmoins il est possible d'augmenter ou de diminuer ces grandeurs, afin de faire varier la résolution temporelle avec laquelle le champ sonore est analysé. Par la suite, en envoyant le message `global`, `resolution` ou `angular`, l'abstraction lance l'analyse respectivement de la variation globale de précision, de la diminution de la résolution ou de la déviation angulaire sur l'ensemble de l'espace. Les valeurs obtenues, mises sur une échelle entre 0 et 1 pour en faciliter l'usage, sont accessibles via la mémoire tampon.

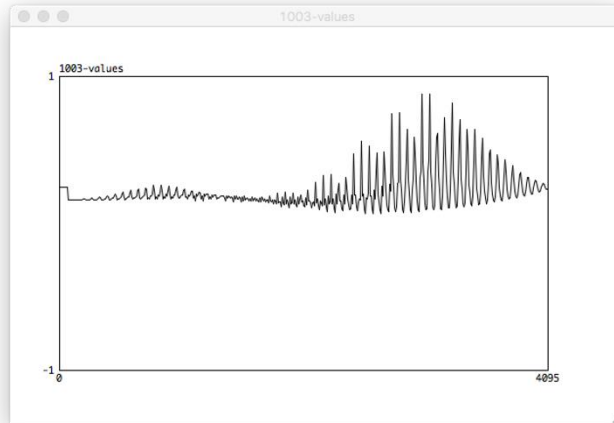


Figure 5.6. Résultats de l'analyse de la résolution sur l'ensemble de la sphère et enregistrés dans la mémoire tampon pour hémisphérique de seize haut-parleurs placés de manière irrégulière pour un champ sonore synthétisé au cinquième ordre de décomposition.

Bien qu'il soit possible de discerner globalement certains artefacts de la restitution, notamment par des variations de l'amplitude des résultats [Figure 5.6], force est de constater qu'ils sont peu utilisables en tant que tels. Il est nécessaire en effet d'associer les valeurs aux angles spécifiques desquels ils découlent. Aussi, il faut considérer que les calculs sont réalisés en partant du sommet vers le bas et que pour chaque élévation, les calculs sont réalisés dans le sens antihoraire. Il est alors possible de retrouver l'azimut et l'élévation en fonction de l'indice  $I$  de la valeur dans la mémoire tampon, du nombre  $N_\theta$  d'azimuts discrétisant le plan horizontal, et du nombre  $N_\phi$  d'élévation discrétisant le plan vertical tel que

$$\theta = 2\pi \frac{I \% N_\phi}{N_\theta} \text{ et } \phi = -\pi \left( \frac{I}{N_\phi - 1} - 0.5 \right).$$

Équation 5.5. Azimut et élévation selon l'indice d'une valeur d'analyse du champ sonore dans un tableau.

De là, il est possible de définir des valeurs statistiques telles que le minimum, le maximum ou la moyenne. Cependant, autant la moyenne est représentative de l'ensemble du champ sonore car elle prend en compte l'ensemble des résultats de l'analyse, le minimum et le maximum quant à eux peuvent prêter à l'erreur. Le minimum est censé représenter la direction dans laquelle l'analyse obtient le meilleur résultat, et réciproquement le maximum est censé représenter la direction dans laquelle l'analyse obtient le moins bon résultat et inversement. Cependant, ces résultats ne sont pas représentatifs d'une zone générale. Aussi, une meilleure approche

consisterait à appliquer un filtre passe-bas sur les résultats, tout en tenant compte de leur organisation sous forme matricielle. D'autres valeurs statistiques peuvent être aussi envisagées, telles que l'écart type ou la variance, qui pourraient donner des informations sur le rendu du champ sonore. Mais cela n'a malheureusement pas pu être encore mis en œuvre, dans le cadre de ces recherches.

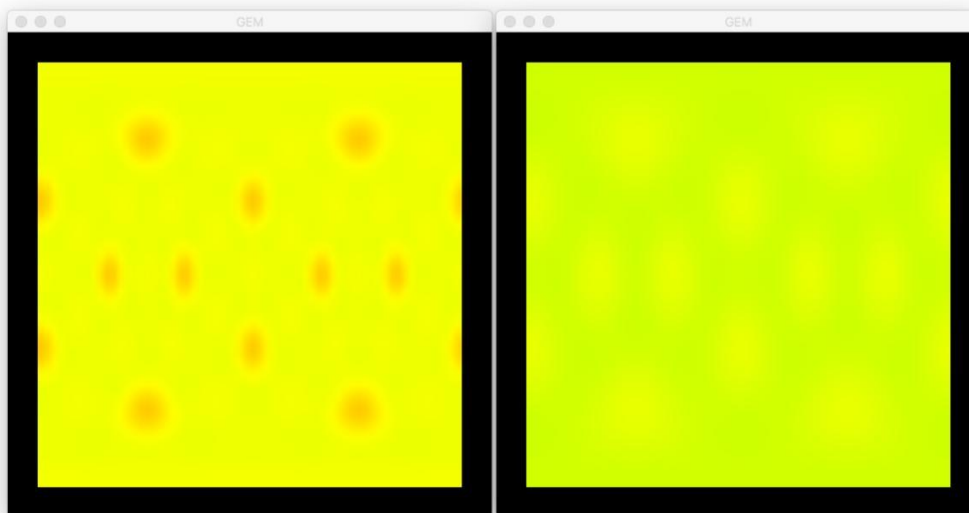


Figure 5.7. Représentation graphique de l'analyse de la résolution angulaire d'un système sphérique de vingt haut-parleurs formant un icosaèdre sous la forme d'un planisphère avec à gauche une optimisation de type *basic* et à droite une optimisation de type *in-phase*. En vert sont représentées les zones où la résolution angulaire est optimale et en rouge sont représentées les zones où la résolution angulaire est dégradée.

Enfin, pour faciliter la compréhension des résultats, un système de représentation graphique des valeurs a été créé. L'abstraction *hoa.2/3d.vector.plot* permet d'afficher dans une fenêtre GEM [Danks, 1996] l'amplitude des valeurs en représentant les analyses sur une surface telle un planisphère. Ce type de représentation est d'une grande efficacité pour analyser et comparer les résultats de manière globale, sur l'ensemble de la sphère. L'utilisation du module GEM était *a priori* nécessaire pour permettre un rendu graphique tridimensionnel sous la forme d'une sphère. Cependant il s'est avéré que cette représentation était moins informative, car elle nécessite de faire pivoter la sphère pour avoir une représentation de l'ensemble. Aussi un planisphère, malgré les déformations que cela implique au niveau des pôles, reste plus adapté. Dès lors, une mise en œuvre dans le système graphique natif de Pure Data [Puckette, 2002] est à présent envisagée.

#### 5.1.4. Analyse temporelle

Enfin, bien que ces données d'analyses soient tout à fait pertinentes pour la restitution de champs sonores composés de sources sonores directionnelles, elles le sont moins dans le cadre d'approches plus originales présentées dans le cadre de ces recherches, telles que la synthèse de champ sonore par décorrélation temporelle ou par granulation quasi-synchrone. Aussi, il est nécessaire de mettre en œuvre des outils répondant au mieux à ces besoins. Il s'agit pour cela de pouvoir caractériser les champs sonores du point de vue de leur caractère diffus.

En mettant en œuvre la synthèse par décorrélation temporelle des signaux associés aux harmoniques, il a été remarqué, de même, que la représentation du champ sonore sous la forme graphique via la projection des composantes sur un cercle ou sphère, ne semble pas convenir. Car elle ne prend justement pas en compte l'aspect temporel. La solution proposée était alors ne plus concevoir l'espace dans son ensemble à un instant donné, mais de concevoir dans le temps un point donné de l'espace. La meilleure représentation des mécanismes en jeu se trouvait être la formulation mathématique des opérations combinées de synthèse et de projection pour ce point de l'espace. Cependant ici, bien que les nécessités ne soient pas les mêmes, il s'agit de préserver l'idée d'intégrer le temps au modèle d'analyse.

La caractéristique essentielle des pratiques présentées, est d'offrir des champs sonores enveloppants où le son semble provenir de multiples directions à la fois. Cela au-delà de la variation de la résolution angulaire qui, comme suggéré, peut être analysée avec les outils classiques venant d'être présentés ici, la définition recoupe, comme cela a déjà été établi, celle du champ diffus : « [...] un ensemble d'ondes temporellement dense, d'incidences équiréparties dans toutes les directions, cette propriété d'isotropie étant observée en tout point de la salle »<sup>191</sup>. L'élément essentiel semble être la notion d'isotropie, à savoir que les propriétés du champ sonore ne changent pas en fonction de la direction. Aussi, définir si cette propriété du champ sonore est respectée est un bon moyen d'analyser à la fois les opérations originales proposées, et la qualité de restitution du champ sonore qui leurs sont associées. Evidemment à un instant donné, il n'est pas possible d'avoir une bonne valeur représentative de l'isotropie. En acoustique des salles, cette propriété est généralement le résultat des réflexions diffuses telles que définies par B-I. Dalenback, M. Kleiner et P. Svensson, [Dalenback & al., 1993]. Il faut donc considérer le champ diffus comme un très grand nombre de sources sonores réparties dans

---

<sup>191</sup> [Daniel, 2001], p. 29.

toutes les directions, mais qui sont décorrélées sur un plan temporel. Amenant dans l'ensemble, à percevoir un champ sonore homogène dans toutes les directions. Cette définition sonne alors en écho avec les procédés mis en jeu, et invite à procéder à une analyse dans le temps.

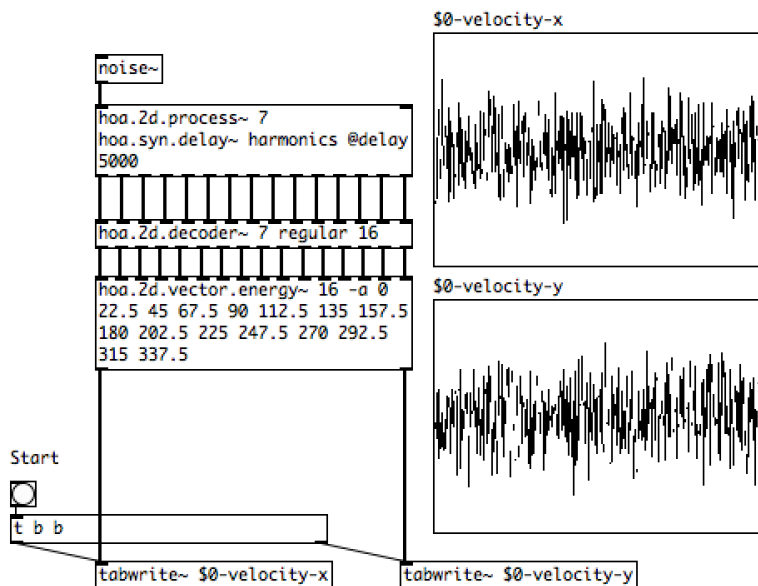


Figure 5.8. Patch Pure Data permettant d'enregistrer les coordonnées du vecteur énergie sur 512 échantillons résultant de la synthèse d'un champ sonore diffus par décorrélation temporelle et d'un décodage pour seize haut-parleurs placés de manière régulière sur un cercle.

En reprenant l'idée de multiples sources sonores réparties dans toutes les directions, cette analyse dans le temps peut être tirée en calculant dans une durée définie les vecteurs vitesse et énergie à de multiples instants. Cette analyse donne alors un ensemble de points associés aux vecteurs représentant le champ sonore sur la durée [Figure 5.8]. Afin d'assurer une bonne isotropie, l'objectif est que ces points recouvrent de la manière la plus homogène possible l'espace sonore. Il est donc possible à partir de là de définir un certain nombre de valeurs tirées des outils statistiques, et permettant de rendre plus évident cette caractéristique.

En considérant que les sons doivent provenir de manière homogène de l'ensemble des directions, il est possible de définir que la moyenne des directions se trouve à l'origine, au centre du système. Afin de critiquer le rendu d'un champ sonore diffus, il est possible d'utiliser

la moyenne des directions d'un ensemble de  $N$  vecteurs calculés à de multiples instants sur une durée<sup>192</sup> telle que

$$\bar{x}_v = \frac{1}{N} \sum_{i=1}^N x_v^i, \bar{y}_v = \frac{1}{N} \sum_{i=1}^N y_v^i \text{ et } \bar{z}_v = \frac{1}{N} \sum_{i=1}^N z_v^i.$$

Équation 5.6. Moyenne dans le temps de la direction d'un vecteur énergie ou vitesse.

Aussi, plus les coordonnées cartésiennes de la moyenne des vecteurs se rapprochent de zéro, plus le champ sonore est censé être isotrope. Néanmoins, cette loi est aussi caractéristique des sources sonores omnidirectionnelles ou encore d'une source sonore pivotant de manière régulière à une période multiple de la durée sur laquelle sont réalisés les calculs. Aussi est-il nécessaire de proposer une autre valeur caractéristique du champ diffus. Afin d'assurer que ces deux exemples de champ sonore répondant aussi à des vecteurs d'énergie et de vitesse nuls, puissent être confondus, il est possible d'utiliser la variance des directions des vecteurs dans le temps. La variance caractérise la dispersion telle que

$$V_v(X) = \frac{1}{N} \sum_{i=1}^N (x_v^i - \bar{x}_v)^2, V_v(Y) = \frac{1}{N} \sum_{i=1}^N (y_v^i - \bar{y}_v)^2 \text{ et } V_v(Z) = \frac{1}{N} \sum_{i=1}^N (z_v^i - \bar{z}_v)^2.$$

Équation 5.7. Variance dans le temps de la direction d'un vecteur énergie ou vitesse.

La variance permet de représenter la dispersion des valeurs, aussi est-elle un outil d'analyse idéal pour compléter la description du champ sonore diffus, après la moyenne. Et là où la direction moyenne dans le temps des vecteurs devait être nulle, la variance dans le temps des vecteurs doit être positive. Néanmoins, l'isotropie du champ implique que cette valeur doit rester faible. Ainsi, une trop grande variance serait le résultat d'une plus grande dispersion des valeurs. Cela implique alors un phénomène se rapprochant d'une onde plane, et donc d'un encodage directionnel dans plusieurs directions. De même que pour les autres approches, une abstraction a été élaborée pour faciliter les calculs. Elle nécessite de définir un temps d'analyse en millisecondes et doit recevoir en entrée les signaux des coordonnées du vecteur énergie ou vitesse. En lançant l'opération, l'abstraction enregistre les signaux sur le laps de temps défini,

---

<sup>192</sup> En utilisant l'approche présentée plus haut en [Figure 5.8], le calcul de la moyenne est simple à réaliser dans Pure Data en utilisant l'objet *array sum*.



et renvoie la moyenne et la variance dans le temps de la direction du vecteur présenté en entrée<sup>193</sup>.

Synthèse	Directionnelle	Temporelle	Granulaire
Moyenne (x y z)	0 0,9 0	0 0 0	0 0 0
Variance (x y z)	0 0 0	0,03 0,03 0,03	0,02 0,02 0,02

Tableau 5.1. Résultats de l'analyse de la moyenne et la variance dans le temps sur 5000 millisecondes de la direction du vecteur énergie pour trois types de synthèses (la source directionnelle est dirigée vers 0 radians d'azimut et d'élévation), directionnelle, décorrélation temporelle et granulaire. La source directionnelle est dirigée vers 0 radians d'azimut. Le champ sonore est décomposé à l'ordre 3 pour un système de vingt haut-parleurs formant un icosaèdre sur la sphère.

Ces valeurs statistiques sont relativement significatives, mais elles ne sont représentatives que d'un premier pas vers l'analyse des champs musicaux. Il est en effet possible d'envisager de nombreux autres outils d'analyses. De par l'analogie qu'il existe entre la décomposition en harmoniques sphériques et cette approche fondée sur des valeurs statistiques, l'idée émerge de transposer le travail réalisé par M. Malt et E. Jourdan pour la description spectrale des sons via la bibliothèque pour Max *zsa.descriptors* [Malt & al., 2008], [Malt & al., 2009]. L'idée serait alors de compléter les analyses avec par exemple, la pente spatiale ou le déploiement spatial, qui seraient les équivalents de ce que M. Malt et E. Jourdan appellent la pente spectrale<sup>194</sup> et le déploiement spectral<sup>195</sup>. De plus, il peut être envisagé de compléter ces informations avec une représentation graphique telle un spectrogramme pour l'espace où il s'agirait de représenter par des points sur un planisphère l'ensemble des coordonnées des vecteurs calculés sur une durée. Cette approche permettrait *a priori* de représenter de manière éloquente le champ, sonore en révélant la répartition plus ou moins homogène ou éparse des sources sonores. Cette représentation peut aussi d'ailleurs intégrer des figures représentant des valeurs statistiques pertinentes.

---

<sup>193</sup> Ces calculs sont réalisés via l'abstraction *hoa.3d.vector.analyser.diffuse~* dans la bibliothèque d'analyse HoaLibrary-Analyser.

<sup>194</sup> *Spectral slope* en anglais.

<sup>195</sup> *Spectral roll-off* en anglais.

Cette approche temporelle et statistique des vecteurs énergie permet donc de caractériser et comparer les opérations de synthèse de champ diffus présentées dans le cadre de ces recherches. Certains résultats sont donnés en exemple par le [Tableau 5.1]. Néanmoins les approches présentées offrent la possibilité de créer des champs sonores qui se trouvent par moment entre ces deux états que sont la source directionnelle et le champ diffus, notamment via la synthèse par pondération aléatoire ou le *flanging* ambisonique. Dès lors, il serait intéressant de voir jusqu'à quel point, les outils présentés suffisent à analyser et critiquer de telles approches. Ces recherches sont des perspectives, qu'il conviendra aussi de poursuivre. Enfin, au-delà de critiquer une approche de synthèse en tant que telle, il est important de souligner que ces vecteurs énergie et vélocité peuvent servir à analyser les opérations liées à la restitution d'une source ponctuelle pour un système de haut-parleurs. Il est aussi possible de se servir de ces outils pour analyser et critiquer ces mêmes opérations – décodage, optimisation, etc. – pour un champ diffus synthétisé par les approches expérimentales présentées.

## 5.2. Restitution du champ sonore

L'approche ambisonique admet une représentation intermédiaire du champ sonore, via sa décomposition en harmoniques sphériques. Cette représentation, comme déjà mentionné dans le premier chapitre de cette thèse, offre de nombreux avantages. Dont notamment le fait de posséder l'ensemble des informations nécessaires à sa restitution. L'enjeu de cette étape consiste à tirer au mieux parti de ces informations pour reproduire à l'oreille de l'auditeur l'effet qu'aurait donné le champ sonore, s'il avait vraiment existé. Afin de répondre à cette problématique, deux opérations primaires sont proposées : l'optimisation de champ sonore et le décodage de champ sonore, aussi appelé « projection », en référence à l'opération mathématique dont elle découle. Ces opérations sont souvent présentées sous une forme groupée, relative à l'étape de restitution dans son ensemble. Mais que le contexte modulaire des logiciels tels que Pure Data et Max amènent spécifiquement à penser de manière séparée, afin de pouvoir expérimenter sur leurs usages et considérer pleinement leurs implications respectives. Tel que présenté dans le premier chapitre de ce travail, ces opérations en ambisonie sont construites sur un modèle acoustique, notamment sur une représentation locale du champ sonore sous la forme d'un gradient de pression au centre de la tête de l'auditeur, des principes psychoacoustiques, notamment de localisation des basses et hautes fréquences, et un modèle mathématique, celui de la décomposition en harmoniques sphériques et qui permet de faire le lien entre les différents domaines. Cependant, suite aux expérimentations dans le domaine des harmoniques sphériques présentées en première partie de ce chapitre, la nature même de la représentation du champ sonore en ambisonie est remise en jeu. Il ne s'agit plus à proprement

parler de vouloir rendre un champ sonore cohérent d'un point de vue directionnel. La mise en œuvre d'opérations de synthèse fondées sur la granulation ou la décorrélation temporelle, amène à concevoir le champ sonore de manière diffuse, au sens large du terme. L'enjeu de la directionnalité a été remplacé comme cela a été montré, par la question de la masse sonore, notamment lors de la synthèse de champ sonore granulaire. Mais aussi par l'élargissement de l'espace, lors de la synthèse de champ sonore par décorrélation temporelle, offrant comme la réverbération, un effet d'échos multiples et multidirectionnels. Ces nouveaux enjeux ont été caractérisés par une notion d'isotropie temporelle lors de l'analyse de ces champs sonores. Aussi les enjeux n'étant plus les mêmes, les objectifs des opérations de restitution s'en trouvent également changés. Il s'agit donc de considérer cette question au regard des nouvelles approches.

Sur un plan méthodologique, à présent que les outils d'analyse sont définis sur un plan formel et technique, l'enjeu consiste de même, à définir techniquement les différentes opérations qu'il est possible de mettre en jeu, afin d'améliorer la restitution du champ sonore. De manière identique à l'approche proposée pour prendre en main les outils d'analyse que sont les vecteurs vitesse et énergie, il est important de comprendre comment fonctionnent les différentes étapes de la restitution du champ sonore. Cela va permettre de mettre en perspective leur avantages et leurs défauts, non seulement au regard de l'approche « classique » de synthèse de sources sonores directionnelles mais aussi au regard des approches originales de synthèse de champs sonores diffus ou chaotiques, présentées dans ces recherches. Construite autour des connaissances acquises, l'enjeu sera alors de présenter une approche plus adaptée aux pratiques musicales présentées. Ce travail sera donc réalisé en deux étapes correspondant aux deux opérations en jeu dans la restitution du champ sonore : l'optimisation de champ sonore et le décodage de champ sonore.

### 5.2.1. Optimisations des sources directionnelles

L'optimisation du champ sonore est une opération qui vise à consolider la restitution selon différents critères, dont notamment le domaine fréquentiel et la disposition de l'auditoire. Les différentes optimisations proposées dans la littérature ont été présentées dans le chapitre précédent selon une approche relativement pragmatique consistant à déterminer dans l'ensemble quelles optimisations appliquer sur quel auditoire et pour quel domaine fréquentiel. Néanmoins, ces principes ont été élaborés par M. A. Gerzon, D. G. Malham et J. Daniel en considérant un champ sonore composé de sources sonores directionnelles. L'enjeu de ces approches est d'assurer un bon maintien des caractéristiques directionnelles des sources sonores

pour l'auditoire. Sur un plan plus technique, l'objectif de ces opérations est d'assurer que les vecteurs vitesse et énergie, dont la mise en œuvre et les possibilités viennent d'être explorées, restent colinéaires avec la direction de la source, tout en préservant au maximum une norme unitaire relative à la résolution angulaire. Cependant, les recherches viennent de montrer que dans le cas d'un champ diffus, ce ne sont plus ces caractéristiques de colinéarité qui sont importantes, mais une question d'isotropie temporelle des vecteurs vitesse et énergie.

Comme suggéré précédemment, les optimisations sont des opérations de pondération des signaux associés aux harmoniques sphériques, en fonction de leurs degrés et de l'ordre de décomposition. Selon les définitions généralisées définies par J. Daniel, les optimisations peuvent être définies sous la forme d'un ensemble de gains  $g$  associés aux signaux.

L'optimisation  $max-r_e$  telle que définie par M. A. Gerzon afin de consolider l'énergie dans la direction de la source sonore, et implicitement maximiser le vecteur énergie associé aux hautes-fréquence, a été généralisée aux ordres élevés par J. Daniel [Daniel, 2001], p. 183, en deux dimensions telles que

$$g_l^{Re2D} = \cos \frac{l\pi}{2N + 2}.$$

Équation 5.8. Optimisation  $max-r_e$  pour une restitution bidimensionnelle.

En trois dimensions, l'opération est plus complexe mais elle a été approximée par E. Zotter et M. Frank [Zotter & al., 2012] selon l'approche généralisée présentée par J. Daniel telle que

$$g_l^{Re3D} = P_l(R_e) \approx P_l \left( \cos \frac{137,9\pi}{180(N + 1,51)} \right).$$

Équation 5.9. Optimisation  $max-r_e$  pour une restitution tridimensionnelle.

L'optimisation *in-phase* telle que définie par D. G. Malham minimise les gains des haut-parleurs s'éloignant de la direction de la source sonore et les annule dans la direction diamétralement opposée. Cette technique a aussi été généralisée aux ordres élevés par J. Daniel [Daniel, 2001], p. 185, en deux dimensions telles que

$$g_l^{in-phase2D} = \frac{N!^2}{(N+l)!(N-l)!}$$

Équation 5.10. Optimisation *in-phase* pour une restitution bidimensionnelle.

Ainsi que pour restitution tridimensionnelle telle que

$$g_l^{in-phase3D} = \frac{N!(N+1)!}{(N+l+1)!(N-l)!}$$

Équation 5.11. Optimisation *in-phase* pour une restitution tridimensionnelle.

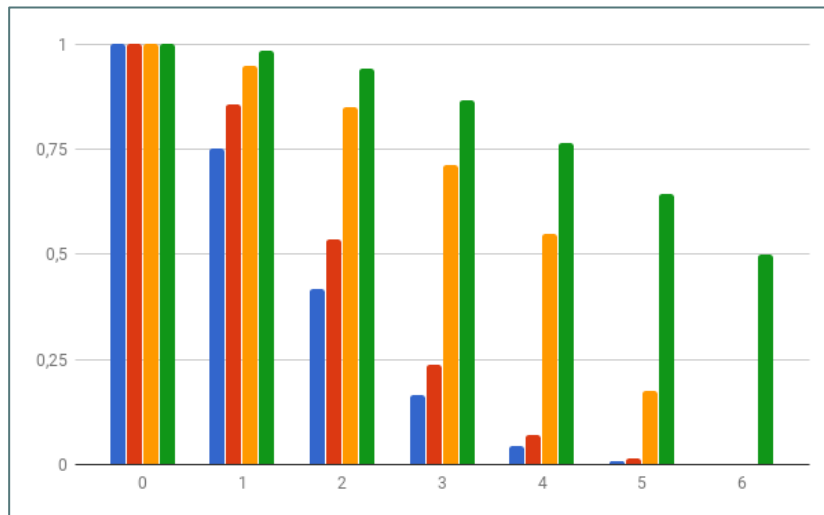


Tableau 5.2. Gains des signaux associés à chaque degré de l'ordre de décomposition 6 avec en bleu les gains d'un optimisation *in-phase* pour représentation tridimensionnelle, en rouge les gains d'une optimisation *in-phase* pour représentation bidimensionnelle, en orange les gains d'une optimisation *max-re* pour représentation tridimensionnelle et en vert les gains d'une optimisation *max-re* pour représentation bidimensionnelle.

En appliquant la formulation pour un ordre de décomposition donné, il est donc possible de définir un gain à appliquer à chacun des signaux selon le degré auquel il est associé. Ce gain diminue les contributions associées aux degrés élevés par rapports aux premiers degrés, tel que représenté en exemple pour les optimisations *in-phase* et *max-re* sur le [Tableau 5.2]. Il est possible, selon cette approche, de réaliser deux remarques. L'une étant qu'en dehors du gain du signal associé au degré 0, tous les gains sont inférieurs à 1. Cela implique une diminution de l'énergie du champ sonore. En réponse à ce problème, J. Daniel propose un coefficient permettant de préserver l'énergie [Daniel, 2001], p. 183-185. L'autre étant que ce procédé opère

telle une diminution de la résolution angulaire, comme suggéré lors de la mise en œuvre de ce traitement de l'espace. Les optimisations diminuent les composantes associées aux degrés supérieurs par rapport aux composantes associées aux degrés inférieurs. L'approche implique alors un élargissement du lobe directionnel principal, afin notamment d'atténuer les lobes secondaires, et donc de faire en sorte que plus les haut-parleurs sont éloignés de la direction de la source sonore moins ils contribuent à sa restitution [Figure 5.9]. Les résultats montrent alors que la consolidation de la reproduction directionnelle des sources, implique de réduire la résolution angulaire.

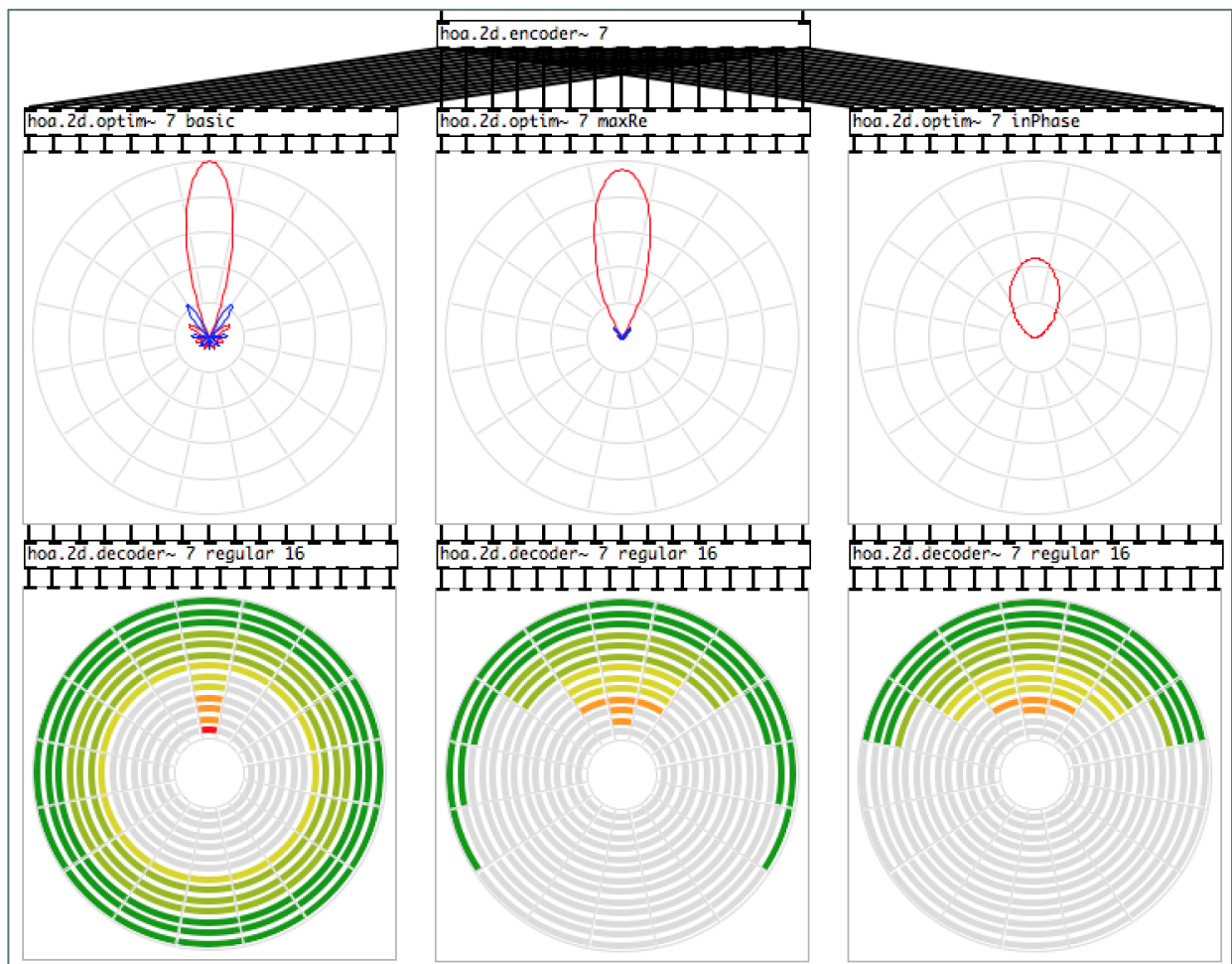


Figure 5.9. Représentation des figures directives d'une source sonore encodée à l'ordre de décomposition 7 pour une restitution bidimensionnelle après les optimisations *basic*, *max-re* et *in-phase*. Les contributions respectives de trois systèmes de seize haut-parleurs régulièrement répartis sur un cercle sont montrées suite à un décodage régulier.

L'enjeu est alors d'appliquer les optimisations en fonction de la configuration de haut-parleurs, de l'auditoire et selon les domaines fréquentiels. L'idée originale proposée par M. A. Gerzon, ce que J. Daniel appelle le décodage psychoacoustique, est de considérer deux bandes de fréquences associées respectivement au vecteur vitesse, en-dessous de 700 hertz, et au vecteur

énergie, au-dessus de 700 hertz. L'opération consiste alors à appliquer un *shelf-filter* (filtre en étage) sur l'ensemble des composantes du champ sonore à cette fréquence de coupure, pour obtenir deux champs sonores associés respectivement aux deux bandes de fréquence mais néanmoins concordants<sup>196</sup>. À chacun des champs sonores, l'optimisation correspondante est, par la suite, appliquée. À savoir l'optimisation *basic* pour les basses fréquences et l'optimisation *max-r<sub>e</sub>* pour les hautes fréquences. Aussi dans le premier cas, les informations de phases sont préservées afin de satisfaire l'approche psychoacoustique définie par M. A. Gerzon. Dans le second cas, les informations d'énergie sont optimisées toujours selon cette même approche. Néanmoins, il est possible de constater que cette approche est construite autour de représentation du champ acoustique sous la forme d'un gradient de pression, au centre de la tête de l'auditeur. Dès lors que l'auditeur se déplace dans la zone de restitution ou qu'un auditoire élargi remplace cet auditeur l'approche n'est plus valide.

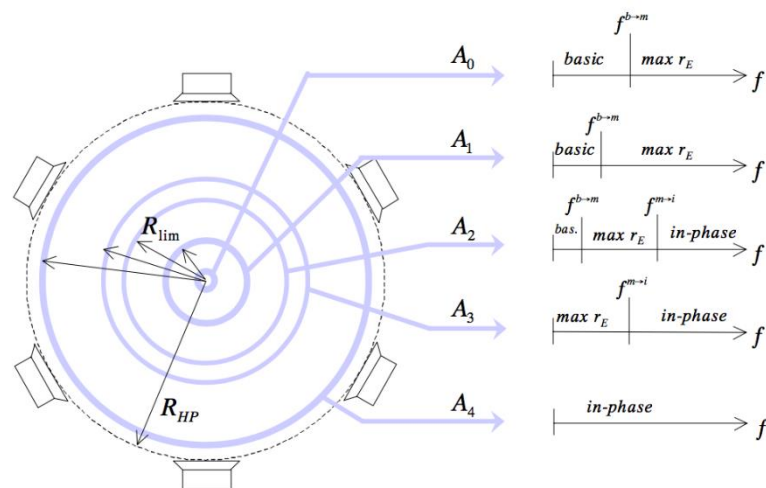


Figure 5.10. Application des optimisations sur le champ sonore selon la disposition de l'auditoire et les domaines fréquentiels tiré de [Daniel, 2001], p. 160. Les fréquences de transition entre les optimisations *basic* et *max-r<sub>e</sub>* sont notée  $f^{b \rightarrow m}$  et les fréquences de transition entre les optimisations *basic* et *max-r<sub>e</sub>* sont notée  $f^{m \rightarrow i}$ .

Afin de répondre à cette restriction, J. Daniel propose une adaptation du dispositif pour différentes conditions d'écoute, de l'auditeur idéalement centré à l'auditoire recouvrant l'ensemble de la zone de restitution [Daniel, 2001], p. 159. L'enjeu est alors, selon ces conditions, de scinder le champ sonore selon un, deux ou trois domaines fréquentiels et d'appliquer l'une des trois optimisations à chacun des champs sonores résultant, de manière identique à l'approche proposée par M. A. Gerzon en y intégrant l'optimisation *in-phase* de D.

<sup>196</sup> À savoir que, si les deux champs sonores résultant sont additionnés, alors le champ sonore initial est retrouvé.

G. Malham. La condition d'écoute optimale est celle définie par M. A. Gerzon. Et l'optimisation de champ sonore qui offre la condition d'écoute la plus flexible, allant jusqu'à pouvoir placer l'auditoire en dehors du système de restitution, est celle proposée par D. G. Malham. Elle consiste à appliquer sur l'ensemble du domaine fréquentiel l'optimisation *in-phase*.

Pour le moment, la bibliothèque n'offre que des outils permettant d'appliquer ces optimisations séparément. Elle invite ainsi implicitement l'utilisateur à réaliser lui-même la mise en œuvre des filtres et la parallélisation des procédés. Cette opération est relativement laborieuse à mettre en œuvre et il n'est pas rare de procéder à l'application directe de l'optimisation de D. G. Malham, alors qu'un système intermédiaire pourrait potentiellement améliorer la qualité de restitution. Aussi serait-il judicieux dans le futur d'offrir sous la forme d'une abstraction destinée à l'objet *hoa.2/3.process*~ un système plus dynamique. L'abstraction nécessiterait de calculer les différents gains d'optimisation selon l'ordre azimutal, le degré et l'ordre de décomposition associés à son instance et de les appliquer selon des domaines fréquents définis par l'utilisateur<sup>197</sup>. Enfin, comme expliqué précédemment, dans le contexte de la musique expérimentale et de recherche, les systèmes de restitution et les auditoires sont si variés qu'il n'est que peu pertinent de présenter des résultats pour un système spécifique – contrairement à de systèmes de type *home-cinema* – car il est peu vraisemblable de rencontrer les conditions identiques. C'est pourquoi l'enjeu est plutôt de favoriser les expérimentations spécifiques à chaque condition.

### 5.2.2. Optimisations des champs diffus

L'enjeu à présent est de définir comment optimiser, un champ sonore diffus tel que présenté dans le cadre de ce travail, selon les outils d'analyse établis précédemment. Pour cela, l'approche propose d'analyser deux champs diffus définis par leurs types de synthèse, la décorrélation temporelle et la granulation. Afin d'offrir un système optimal et de pouvoir mettre en perspective des résultats associés à des ordres de décomposition élevés, les analyses sont réalisées sur une représentation bidimensionnelle de l'espace pour un ordre de décomposition de 7.

---

<sup>197</sup> Faute de temps et en raison de la complexité du calcul des polynômes de Legendre, cet outil n'a pas encore pu être mis en œuvre.



Synthèse		Optimisation	basic	max- $r_e$	in-phase
Temporelle	Moyenne		0,011 0,002	0,016 -0,003	0,021 -0,016
	Variance		0,061 0,054	0,083 0,069	0.127 0.127
Granulaire	Moyenne		-0.007 -0.001	-0.005 -0.006	0.006 -0.002
	Variance		0.0501 0.055	0.065 0.065	0.090 0.086

Tableau 5.3. Résultats de l'analyse de la moyenne et la variance dans le temps sur 2000 millisecondes de la direction du vecteur énergie pour la synthèse de champ sonore par décorrélation temporelle ou granulation en fonction des trois optimisations *basic*, *max- $r_e$*  et *in-phase*.

Les résultats de ces analyses montrent, que ce soit pour la synthèse de champ sonore par décorrélation temporelle ou pour la granulation des signaux associés aux harmoniques et qu'importe l'optimisation appliquée, que les vecteurs vitesses offrent des résultats relativement similaires. C'est ici une lacune des outils d'analyse. Il s'avère que le vecteur vitesse est peu représentatif des phénomènes de champs diffus, en raison des trop grandes variations générées par les inversions de la phase. En somme, une telle approche n'implique pas nécessairement que les optimisations n'aient aucun impact dans le domaine des basses fréquences. Aussi, de plus amples recherches devraient être portées à trouver un outil capable de représenter les phénomènes de champ diffus dans ce domaine fréquentiel. D'un autre côté, les résultats des analyses du vecteur énergie [Tableau 5.3], peuvent être interprétés. Les variations des valeurs autour de la moyenne sont relativement faibles. Cela amène à définir que les multiples sources sonores qui composent le champ diffus doivent provenir de l'ensemble des directions, quelle que soit l'optimisation utilisée. En cela, ces approches n'opèrent pas d'artefacts et le champ sonore devrait *a priori* être toujours immersif. Ces résultats sont en adéquation avec le point de vue selon lequel la perception de la direction est renforcée en augmentant les contributions des signaux associés aux harmoniques de degrés peu élevés. Les variances varient quant à elles selon les optimisations, de manière significative. Que ce soit pour la synthèse par décorrélation temporelle des signaux ou par granulation des signaux, la variance est plus importante avec une optimisation *max- $r_e$*  qu'avec une optimisation *basic*. Et plus important aussi avec une optimisation *in-phase* qu'avec une optimisation *max- $r_e$* . Compte tenu du fait que l'optimisation *max- $r_e$*  diminue les contributions des signaux associés aux harmoniques de degrés élevés et que l'optimisation *in-phase* les diminuent d'autant plus, il est possible de déduire que la variance augmente d'autant plus que les pondérations des degrés

élevés diminuent. Une grande variance peut être comprise comme le résultat de multiples sources directionnelles et dont la directivité se rapproche d'une onde plane. En d'autres termes, plus la variance est grande, plus les sources sonores auront tendance à se rapprocher d'un modèle de champ libre. Or l'intérêt des techniques de synthèse présentées est d'offrir un phénomène proche des réflexions diffuses de l'acoustique des salles. Aussi, cela amène à penser que l'application d'optimisations sur le champ sonore appauvrit la qualité du rendu spatial.

Cet artefact du champ sonore recoupe dans la pratique le problème déjà soulevé lors de la mise en œuvre de la synthèse de champ sonore par granulation des signaux associés aux harmoniques sphériques. L'application d'une optimisation peut être assimilée à une variation de la résolution angulaire. En effet, en diminuant les contributions associées aux harmoniques de degrés élevés, le processus peut être vu comme une opération de multiples ordres fractionnaires, lors de la mise en œuvre de ce premier traitement. Or, il a été remarqué lors de l'expérimentation de la synthèse de champ sonore par granulation, que l'ordre de décomposition dépend aussi de la masse sonore, et donc de la qualité du rendu sonore. Ce principe a d'ailleurs été mis en relation aussi avec la synthèse de champ sonore par décorrélation temporelle, où un ordre de décomposition plus élevé augmente le nombre d'écho et donc amplifie l'effet voulu. Inversement, une diminution de l'ordre diminue la qualité de cet effet. Afin de mettre en exergue l'ampleur d'une diminution non-souhaitée<sup>198</sup> de l'ordre de décomposition, il est possible de considérer un traitement où chaque composante spatiale et sonore soit associée à une voix particulière d'une œuvre. Pour cela, il suffit par exemple de reprendre l'exemple proposé par A. Bonardi dans *Pianotronics 1* d'un *harmonizer* où chaque voix correspond au signal associé à un harmonique. Dans ce cas, diminuer l'amplitude des signaux associés aux degrés d'ordre élevés ne change pas seulement la caractéristique spatiale de la pièce. Cela change aussi sa nature profonde, en faisant varier l'amplitude voix qui peut être importante sur le plan harmonique. Dans ce contexte, l'application d'une optimisation sur cette approche risque de faire perdre une grande partie de l'œuvre. En reprenant le [Tableau 5.2] et en considérant une optimisation *in-phase*, il est évident qu'à partir de là, les voix correspondantes aux signaux associés à l'ordre 4 deviennent très peu audibles et celles correspondantes aux signaux associés à l'ordre 5 et 6, disparaissent.

---

<sup>198</sup> Contrairement à une diminution de l'ordre de décomposition prédéfinie, selon laquelle un traitement peut s'adapter selon différentes stratégies présentées dans la première partie de ce chapitre, ici la diminution de l'ordre de décomposition est une conséquence non désirable de la consolidation de l'effet directionnel des sources sonores.

Aussi, ces remarques consolident l'idée que l'application d'une optimisation du champ sonore, destinée à la restitution de sources sonores directionnelles, appauvrit le rendu des synthèses originales de champs diffus. Pour de futurs travaux, il serait alors intéressant de procéder à des recherches permettant de proposer une approche qui optimiserait ce type de champ sonore. Ces résultats amènent aussi à se poser la question de comment optimiser un champ sonore composé à la fois de sources sonores directives et de champ diffus. Cela soulève aussi la question de l'approche à utiliser lors de la mise en œuvre de traitements originaux, permettant de passer linéairement d'une source directionnelle à un champ diffus ou chaotique. Dans le cadre de cette thèse, il n'est malheureusement pas possible de répondre à toutes ces questions, mais de premières expérimentations ont déjà été réalisées.

### 5.2.3. Contraintes du décodage et solutions

L'enjeu, suite à l'optimisation du champ sonore, est de définir une approche permettant de générer les signaux destinés aux haut-parleurs à partir de la représentation du champ sonore sous la forme de signaux associés aux harmoniques sphériques. Cette opération est sujette à une restriction qui définit que le nombre de haut-parleurs nécessaire à une restitution optimale du champ sonore doit être supérieur ou égal au nombre d'harmoniques. Cela définit aussi que ces haut-parleurs offrent une discrétisation parfaite de la sphère entourant l'auditoire, ou du cercle dans le cas d'une restitution bidimensionnelle. Pour un système bidimensionnel, ce type de configuration, bien qu'envisageable, reste relativement rare. Les systèmes personnels étant plus généralement des systèmes de type *home cinéma*, ou du moins restreints, les systèmes dits réguliers les plus courants comportent en générale quatre, six ou huit haut-parleurs mais rarement plus. Cela implique une limitation de l'ordre de décomposition assez draconienne et généralement aux ordres 1, 2 ou 3. Pour un système tridimensionnel, ces configurations sont d'autant plus limitées qu'il n'existe que cinq configurations optimales. Dans la pratique, de tels systèmes sont exceptionnels car ils nécessitent que des haut-parleurs soient situés sous l'auditoire. Les lieux permettant une telle installation nécessitent de très grands volumes pour un large auditoire système. Et dans le cas contraire, le système est destiné à un seul auditeur. Dès lors, pour une restitution en trois dimensions, il n'existe vraisemblablement que quelques dispositifs au monde acceptables, et les usages sont par défaut sous-optimaux. Aussi est-il nécessaire de proposer une approche permettant de corriger au mieux les artefacts dus aux problèmes techniques et matériels.

Avant de soumettre une solution, il est nécessaire de revenir de façon plus technique sur le problème, afin de pouvoir discerner les tenants et les aboutissants de cette restriction liée à

l'opération de projection des harmoniques sphériques sur la sphère. Cette analyse sera réalisée au regard des deux modes opératoires définis dans le cadre de ces travaux, à savoir la synthèse de sources sonore directionnelles et la synthèse de champs diffus. A partir de là, ce travail de recherche ne pourra encore malheureusement pas mettre à l'épreuve toutes les approches qui ont été proposées dans la littérature. Il est donc nécessaire de rappeler les principales raisons impliquant qu'un tel objectif soit irréalisable et improductif. En effet, beaucoup d'approches sont spécifiques à des configurations de haut-parleurs qui sont relativement peu pratique à mettre en œuvre, notamment pour l'expérimentation. Or, il est nécessaire de pouvoir confronter les modèles à la réalité, surtout en ce qui concerne l'amélioration de la qualité de la restitution. Dans ce contexte, de telles approches auraient été réduites à l'utilisation et la critique de données et de valeurs mathématiques. De plus, ces travaux visent une approche générique. Les systèmes de restitutions sont multiples et la pertinence d'une approche dédiée à un système spécifique reste faible. Enfin, nombre des approches présentées sont construites sur l'utilisation d'optimisations ou d'une pseudo-inverse de la matrice de décodage – qui peut être assimilée à une optimisation locale du champ sonore – ce qui implique nécessairement une dégradation non-souhaitable<sup>199</sup> de la restitution des champs sonores. Aussi pour le lecteur souhaitant néanmoins plus d'information sur le sujet, la thèse de J. Daniel recouvre les approches disponibles jusqu'au début des années 2000. Plus récemment, deux nouvelles approches proposées conjointement par F. Zotter, M. Frank, H. Pomberger et A. Sontacchi semblent être intéressantes : le EPAD (*Energy-Preserving Ambisonic Decoding*), [Pomberger & al., 2012], [Zotter et al., 2013] et le AllRAD (*All-round Ambisonic Decoding*), [Zotter et al., 2010], [Zotter et al., 2012] et [Zotter et al., 2013]<sup>200</sup>. Pour des raisons de priorités et de temps, ces recherches n'ont pas encore pu investiguer le EPAD, qui semble à ce sujet plus complexe à mettre en œuvre que les autres approches. Afin de couper court au suspense, la solution acceptée dans le cadre de ce travail, est celle du AllRAD. Cette approche revient, à quelques variantes près, à la mise en œuvre proposée dans la bibliothèque depuis 2014 à partir d'une idée originale et similaire<sup>201</sup>, mais est plus établie dans la littérature. L'enjeu sera donc de présenter comment cette approche permet de répondre aux différents problèmes rencontrés lors du décodage primaire, qui consiste en une projection des signaux associés aux harmoniques sphériques dans la direction des haut-parleurs.

---

<sup>199</sup> Ce type d'approche sera néanmoins brièvement évoquée afin d'en comprendre les principes.

<sup>200</sup> L'approche s'inspire d'une approche inverse proposée par J-M. Batke et F. Keiler [Batke & al., 2010].

<sup>201</sup> N'ayant au moment de cette mise en œuvre pas connaissance du AllRAD, j'ai trouvé la solution très judicieuse. À mon grand dam et désespoir, il convient de donner l'antériorité à F. Zotter, M. Frank, H. Pomberger et A. Sontacchi ☺ !

## 5.2.4. Approche pratique de la projection

Afin de mettre en avant les problèmes du décodage de champ sonore par projection, il est nécessaire de concevoir sur un plan pratique en traitement du signal, en quoi consiste l'opération mathématique présentée dans le chapitre précédent. Une des complexités de ce procédé réside simplement en sa formulation. L'approche s'inscrivant initialement dans un domaine continu, les opérations de décomposition en harmoniques sphériques et de projection sont souvent représentées par des intégrales dans la littérature. Or, la formulation pour une représentation discrète de l'espace proposée est déjà bien plus intelligible et adaptée à ce contexte. Dans ces cas, en raison d'autres opérations intégrées aux mécanismes de décodage, la projection est souvent représentée sous la forme d'une multiplication matricielle, telle que l'utilisation d'une matrice inverse. Etant donné que ce type de procédé est à proscrire, il est possible de simplifier l'approche en la décomposant afin d'en faciliter la compréhension. Il s'agit alors de considérer l'opération pour un signal destiné à un haut-parleur tel que présentée mathématiquement avec l'[Équation 2.11]. Cette opération sera alors répétée pour l'ensemble des haut-parleurs. En éclaircissant le mécanisme global de l'opération projection de la sorte, il est alors possible de rendre plus intelligible comment chaque signal associé à un harmonique contribue à chaque haut-parleur et de ce fait à l'ensemble du champ sonore. Alors les problèmes de configurations irrégulières de haut-parleurs pourront être mieux appréhendés.

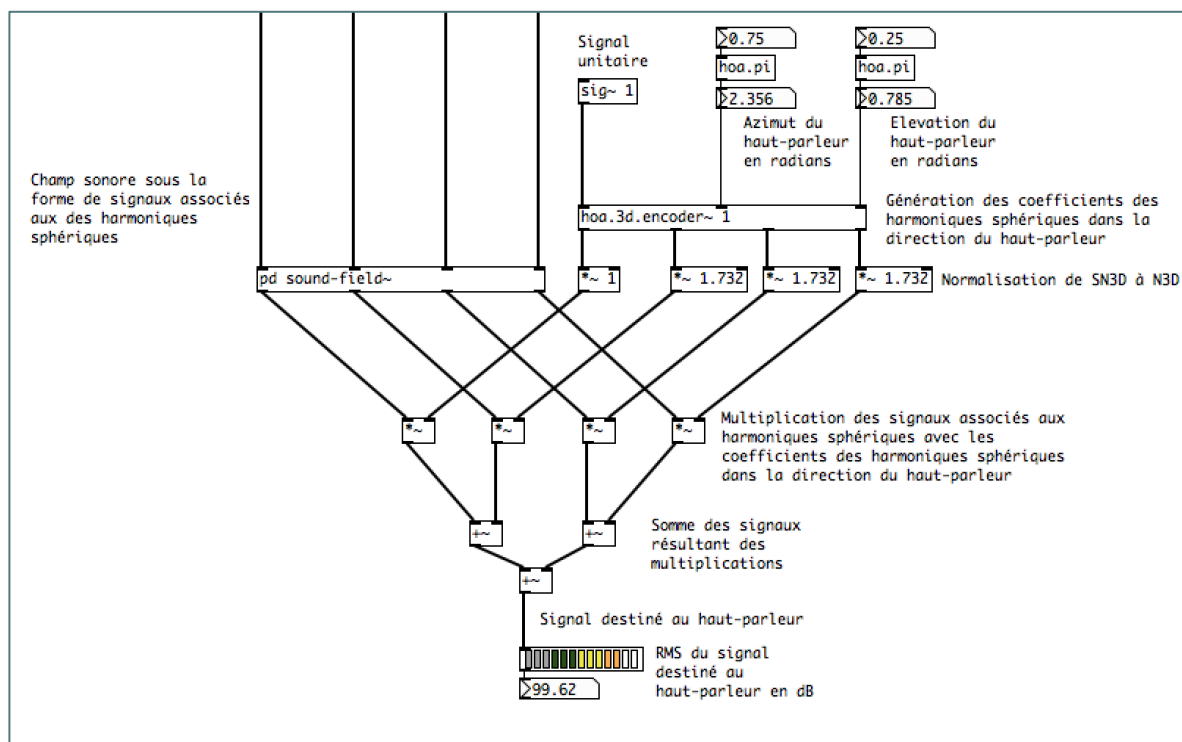


Figure 5.11. Patch Pure Data représentant la projection d'un champ sonore dans la direction d'un haut-parleur au premier ordre de décomposition.

Aussi, l'opération de projection du champ sonore sous la forme de signaux associés aux harmoniques sphériques vers un haut-parleur se résume à calculer la somme de ces signaux pondérés par les coefficients résultant des harmoniques sphériques qui leurs sont associés dans la direction de ce haut-parleur. Cette opération peut aussi être comprise comme la multiplication terme à terme des signaux associés aux harmoniques sphériques du champ sonore, avec les signaux normalisés associés aux harmoniques sphériques de l'encodage d'un signal unitaire dans la direction du haut-parleur, puis la somme des signaux résultants [Figure 5.11]. Cette opération possède de base la propriété d'être extensible à tout ordre de décomposition, et dans toutes les directions. Il suffit pour cela de varier l'angle d'encodage de signal unitaire.

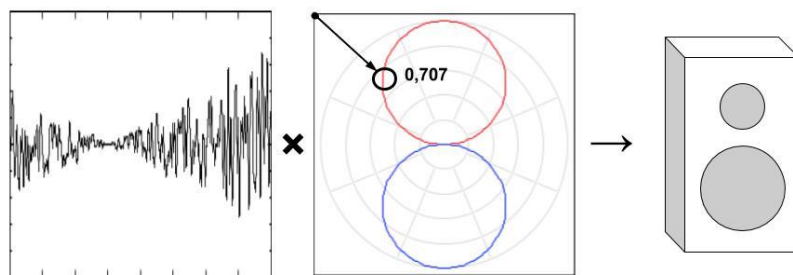


Figure 5.12. Représentation de la contribution du signal de l'harmonique de degré 1 et d'ordre azimutal 1 pour un haut-parleur à 45 degrés sur la gauche.

Il est alors, selon cette approche, plus aisé de concevoir comment chaque signal associé à un harmonique, contribue au champ sonore. Dans une direction donnée, le signal associé à un harmonique sphérique est envoyé dans un haut-parleur en fonction de l'amplitude de cet harmonique sphérique, dans la direction du haut-parleur [Figure 5.12]. Aussi, la règle sous-jacente à la bonne projection du champ sonore est alors que pour chaque signal associé à un harmonique, au moins un haut-parleur contribue à restituer ce signal. Cela implique aussi que la somme des contributions de l'ensemble des haut-parleurs pour cet harmonique soit nulle, en dehors des signaux associés à l'harmonique de degré zéro, car c'est sur celle-ci que repose la construction des figures spatiales.

Cette règle est la réciproque de celle définissant que le système de haut-parleur doit discrétiser parfaitement l'espace, selon un nombre de haut-parleurs supérieur ou égal au nombre d'harmoniques sphériques utilisés pour représenter cet espace. Il peut être démontré qu'en dessous d'un certain nombre de haut-parleurs, certains signaux associés aux harmoniques

sphériques ne contribuent plus, car les figures de directivités qui leur sont associées passent « entre les mailles du filet ». Cette caractéristique peut être aisément représentée telle que sur la [Figure 5.13] pour une restitution bidimensionnelle. L'espace sonore discrétisé par quatre haut-parleurs, associés à quatre points, ne peut recevoir la contribution associée à l'harmonique d'ordre azimutal -2. Aussi une rapide vérification des amplitudes de chaque harmonique dans les directions des points montre que leur somme est nulle, en dehors du premier harmonique. Cette relation entre les différentes amplitudes peut-être aussi vérifiée selon d'autres discrétisations, tant que celle-ci est régulière et donc avec trois points ou plus.

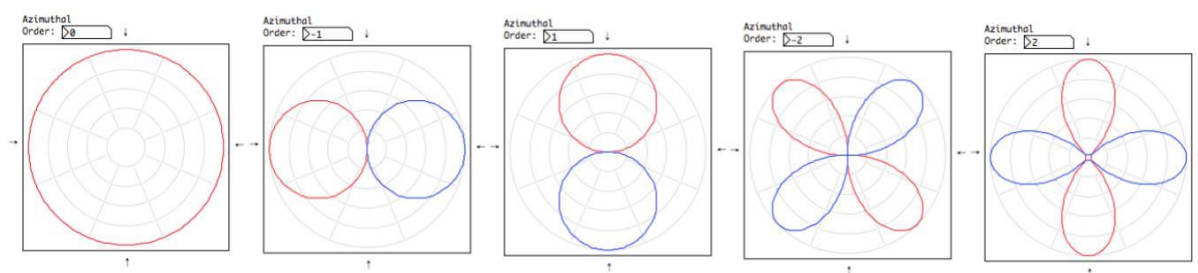


Figure 5.13. Représentation des contributions des cinq premiers harmoniques circulaires pour quatre points et représentant une discrétisation optimale d'un cercle pour l'ordre de décomposition 1 mais sous-optimale pour un ordre de décomposition 2.

En considérant les figures spatiales des harmoniques circulaires, pour une source sonore encodée à 45 degrés vers la gauche à l'ordre de décomposition 2, seul le signal associé à l'harmonique d'ordre azimutal -2 contribue pour les signaux associés au degré 2. L'amplitude du signal associé à l'harmonique d'ordre azimutal 2 étant nul pour une source sonore encodée 45 degrés. Or le signal associé à l'harmonique d'ordre azimutal -2 ne contribue pas à la restitution sur les quatre haut-parleurs. Aussi faut-il prévoir un champ sonore relatif à une décomposition à l'ordre 1, car seuls les signaux associés à ces harmoniques contribuent. Cet exemple est aussi applicable pour un système tridimensionnel.

En somme, sous-discrétiser l'espace par rapport à un ordre de décomposition peut impliquer une diminution effective de l'ordre de décomposition. Seulement si ce n'est pas le seul artefact. En effet, selon l'ordre de décomposition et le système de restitution, le champ sonore peut perdre tout caractère directionnel, tel que le prouve la [Figure 5.14] pour un champ sonore bidimensionnel à un ordre de décomposition 3. Dans ce cas, un système de restitution de quatre haut-parleurs génère une déformation lorsque la source est située à 22,5 degrés. Notamment dans les basses fréquences selon si le vecteur vitesse se trouve dans la direction opposée vers

-22,5 degrés alors que le vecteur énergie est plus proche de 0 degré que de la direction d'encodage. Lorsque la source est située à 45 degrés, et cette analyse est similaire pour 135, 225 et 315 degrés, la restitution est bien plus critique étant donné que l'ensemble des haut-parleurs contribue de manière égale, amenant à un champ sonore omnidirectionnel. Et cela s'applique aussi bien dans le domaine des basses ou hautes fréquences selon les vecteurs vitesse et énergie.

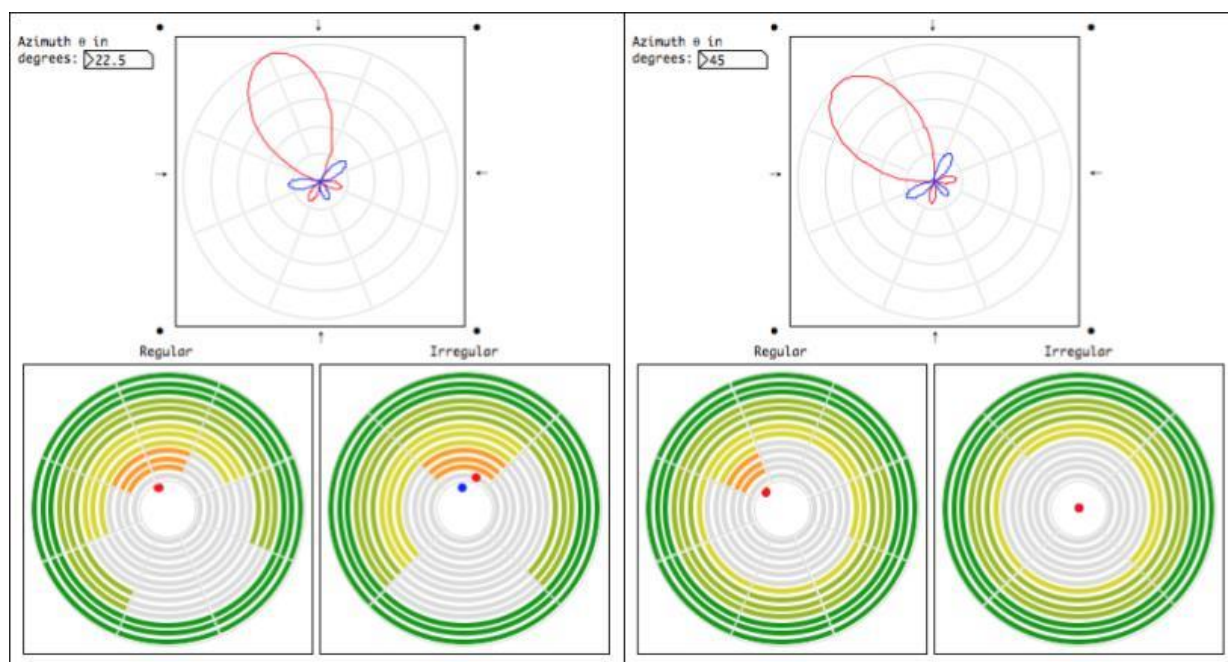


Figure 5.14. Patch Pure Data représentant la déformation du champ sonore lors de l'utilisation d'un système de restitution sous-optimal de quatre haut-parleurs par rapport à un système de restitution optimal de huit haut-parleurs pour une représentation en deux dimensions et une décomposition du champ sonore à l'ordre 3. Les flèches représentent les haut-parleurs du système sous-optimal et les ronds représentent les haut-parleurs manquants, nécessaires au système optimal. Les représentations sous le terme *regular* représentent les contributions des haut-parleurs du système optimal et les représentations sous le terme *irregular* représentent les contributions des haut-parleurs du système sous-optimal. A gauche, l'angle d'encodage est de 22,5 degrés et de 45 degrés à droite. Le vecteur énergie est représenté en bleu et le vecteur vitesse est représenté en rouge.

Aussi est-il possible de retrouver des problèmes similaires pour d'autres systèmes de restitution et d'autres ordres de décomposition. Cette présentation des artefacts de la restitution est cependant limitée à des systèmes réguliers mais sous optimal. Dans l'absolu, il serait envisageable de procéder à de telles méthodes avec des systèmes irréguliers, dans le sens où la discrétisation de l'espace n'est pas régulière, telle que la réalise un système 5.1 par exemple. Néanmoins, les possibilités sont infinies et offrir un aperçu général dépasserait le cadre de cette thèse. Aussi l'essentiel étant de comprendre les mécanismes en jeu et les causes de la déformation du champ sonore à la restitution. De plus, les configurations régulières mais sous optimales par rapport à un ordre de configuration peuvent être considérées comme un sous



ensemble, ou des cas particuliers des configurations irrégulières. Aussi est-il possible de procéder de manière identique à l'approche présentée, en plaçant les haut-parleurs dans l'espace de manière irrégulière et d'analyser de la même manière comment chacun des signaux associés aux harmoniques sphériques ou circulaires contribue à la restitution du champ sonore.

Enfin dans ce type d'approche, la méthodologie s'est évertuée jusqu'à présent à analyser le résultat pour des sources sonores directionnelles. Mais les règles présentées sont toutes aussi valides pour des champs diffus bien que les implications soient différentes. Comme il a été à nombreuses fois remarqué, l'enjeu lors de la restitution d'un champ sonore diffus n'est donc pas forcément de préserver la directionnalité, mais de préserver un jeu de pondération constant entre les signaux associés aux harmoniques. L'enjeu lors de l'analyse des techniques d'optimisation était d'éviter de diminuer le gain de certains signaux par rapport aux autres, et de diminuer la « qualité » des techniques de synthèse de champ sonore, que ce soit la masse des nappes de grains ou l'ouverture spatiale de la décorrélation temporelle. Et ces principes ont été mis en exergue avec l'exemple de l'utilisation d'un *harmonizer* et d'une œuvre où chaque signal associé à un harmonique correspond à une voix. Aussi, l'enjeu est ici similaire. Si le signal associé à un harmonique ne contribue pas, tel que présenté pour le signal associé à l'harmonique d'ordre -2 dans la [Figure 5.13], alors une voix est perdue par l'opération de projection. Et dans une moindre mesure, considérer la règle selon laquelle la somme des contributions d'un signal associé à un harmonique dans tous les haut-parleurs est nulle, permet d'assurer une constance dans la restitution du signal associé à chaque harmonique dans le champ sonore. A cette règle, dans le cas d'un champ diffus, il est possible d'ajouter que les sommes des valeurs absolues des contributions des signaux associés aux harmoniques soient identiques, afin d'assurer que tous les signaux associés aux harmoniques contribuent de la même façon. Comme le prouve cette analyse, les champs diffus, malgré leur caractéristique non directionnelle, n'en restent pas moins sujets aux artefacts d'une configuration de haut-parleurs sous-optimale ou irrégulière.

### Projection optimisée

Il existe, comme évoqué précédemment, de nombreuses approches dans la littérature afin de compenser au mieux les artefacts, en raison d'une disposition irrégulière ou sous optimale des haut-parleurs. Pour cela le lecteur peut se référer à [Daniel, 2001], p. 187-200, qui propose un grand nombre de solutions pour des configurations allant de semi-régulières, comme un rectangle au lieu d'un carré, à irrégulières, comme les formats 3/2 ou 3/4 définis par l'ITU<sup>202</sup>

---

<sup>202</sup> *International Telecommunication Union.*

[ITU, 2012] pour une restitution bidimensionnelle et des configurations hémisphériques ou pyramidales pour une restitution tridimensionnelle. Le but étant pour ces configurations de consolider au mieux la colinéarité des vecteurs énergie et vitesse, avec le vecteur de la source sonore afin de diminuer les déformations directionnelles. Pour cela, plusieurs approches sont donc proposées. Le but de cette présentation n'étant pas de revenir sur leurs mises en œuvre respectives, qui selon les dispositifs de haut-parleurs peuvent prendre des formes variées et complexes, mais simplement d'en ressortir les idées maîtresses. A ce niveau d'abstraction, sans considérer la manière de les mettre en œuvre, des approches principales émergent.

M. A. Gerzon et G. J. Barton proposent une approche consistant en une déformation de la perspective [Gerzon & al., 1992]. Le principe est simple et part du principe que le système de haut-parleurs déforme la scène sonore. A savoir que la source sonore aura tendance une fois restituée à être perçue de manière non linéaire sur le cercle ou la sphère entourant l'auditeur. La déformation aura en effet tendance à agir comme donnant plus d'importance à la scène frontale sur les systèmes de type 5.1. Aussi, une source orientée vers la gauche sera plutôt perçue vers l'avant-gauche. Une manière de compenser cette déformation est alors d'appliquer à l'avance une déformation inverse du champ sonore, ce que M. A. Gerzon appelle *forward dominance*. Cependant, comme le note J. Daniel cette approche nécessite « des méthodes d'optimisation non-linéaire »<sup>203</sup>. Aussi est-elle difficilement réalisable. De plus, les exemples proposés par M. A. Gerzon et G. J. Barton sont restreints à l'ordre 1, posant la question de la validité aux ordres élevés<sup>204</sup>.

La deuxième approche<sup>205</sup> pourrait être considérée comme une projection à résolution variable. Il s'agit donc de réduire les contributions des signaux associés aux harmoniques de degré élevé, dans les directions où le système de haut-parleurs n'est plus assez dense pour satisfaire une bonne directionnalité. L'exemple présenté sur la [Figure 5.15] construit son approche en utilisant le traitement de variation de la résolution angulaire présentée dans la première partie de ce chapitre. Le principe est ici aussi relativement simple. Il s'agit de définir les endroits où

---

<sup>203</sup> [Daniel, 2001], p. 188.

<sup>204</sup> Il est cependant intéressant que j'ai mis en œuvre une approche similaire limitée à une restitution bidimensionnelle mais pour les ordres élevés dans la version 1.2 de la bibliothèque. Cependant, l'approche à dire vrai n'était pas très bien formalisée et effective que dans de rares conditions aussi a-t-elle été enlevée. Néanmoins une trace écrite du procédé mis en œuvre est proposée en [Guillot, 2012] et de plus amples recherches dans cette direction pourraient offrir des perspectives intéressantes.

<sup>205</sup> Cette approche est généralisée pour différents types de configuration selon des opérations d'optimisation matricielle mais il s'agit dans le cas présent d'en présenter l'idée.

des haut-parleurs sont absents. En comparaison du système optimal possédant huit haut-parleurs, le système sous-optimal de droite nécessite quatre haut-parleurs à 45, 135, 225 et 315 degrés. Ces haut-parleurs manquants correspondent aux endroits où la restitution est la plus critique [Figure 5.14]. Ce système de haut-parleurs sous-optimal pour le troisième ordre de décomposition correspond néanmoins à un système de haut-parleurs optimal au premier ordre de décomposition. L'idée est de diminuer la résolution angulaire autour de ces directions critiques associées aux haut-parleurs manquant pour atteindre une résolution sensiblement équivalente à celle du premier ordre de décomposition. Aussi s'agit-il de définir la résolution angulaire selon la direction telle que

$$\Delta_{\theta} = 0.3|\cos 2\theta| + 0.7$$

Équation 5.12. Variation de la résolution angulaire en fonction de la direction pour l'optimisation d'un système quadraphonique à l'ordre de décomposition 3.

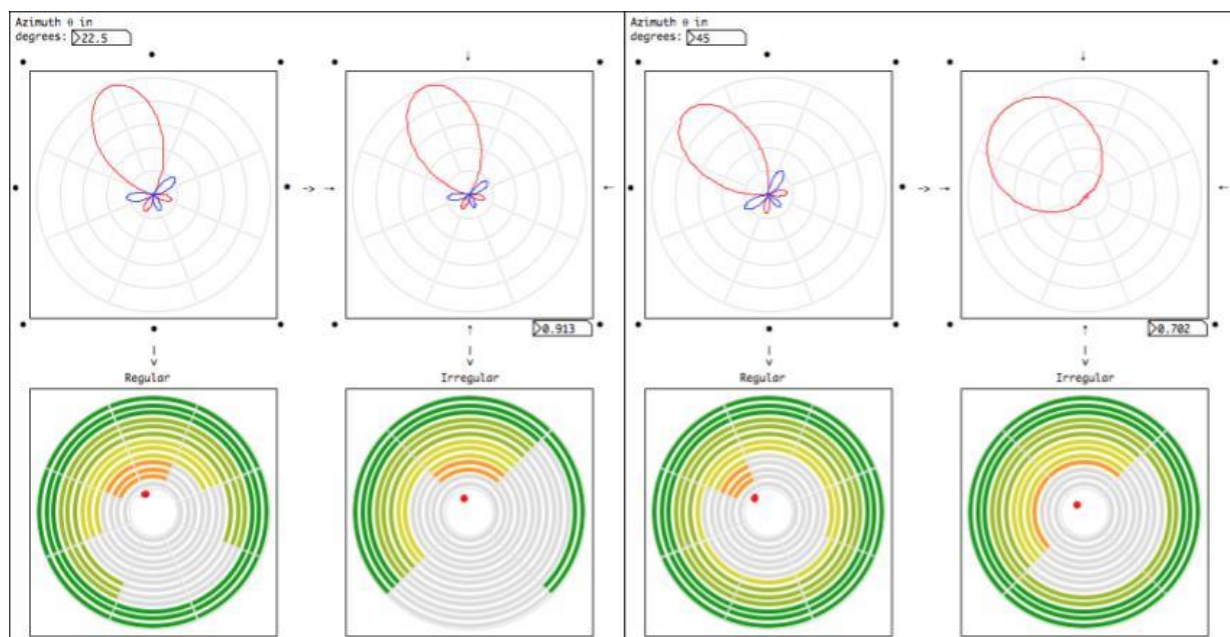


Figure 5.15. Patch Pure Data représentant l'optimisation du champ sonore par variation de la résolution angulaire lors de l'utilisation d'un système de restitution sous-optimal de quatre haut-parleurs par rapport à un système de restitution optimal de huit haut-parleurs pour une représentation en deux dimensions et une décomposition du champ sonore à l'ordre 3. Les flèches représentent les haut-parleurs du système sous-optimal et les ronds représentent les haut-parleurs manquants nécessaires au système optimal. Les représentations sous le terme *regular* représentent les contributions des haut-parleurs du système optimal et représentations sous le terme *irregular* représentent les contributions des haut-parleurs du système sous-optimal. A gauche l'angle d'encodage est de 22,5 degrés et de 45 degrés à droite. Le vecteur énergie est représenté en bleu et le vecteur vitesse est représenté en rouge. Les valeurs sous les figures directives de droite sont associées à la diminution de la résolution angulaire.

Dans le cas présent, compte de tenu de la régularité du système et de l'aspect périodique des haut-parleurs manquants, l'approche est relativement simple à définir sur un plan angulaire. La

difficulté consiste simplement à définir la plage de variation de la résolution angulaire afin de préserver la colinéarité des vecteurs énergie et vitesse avec la direction de la source sonore.

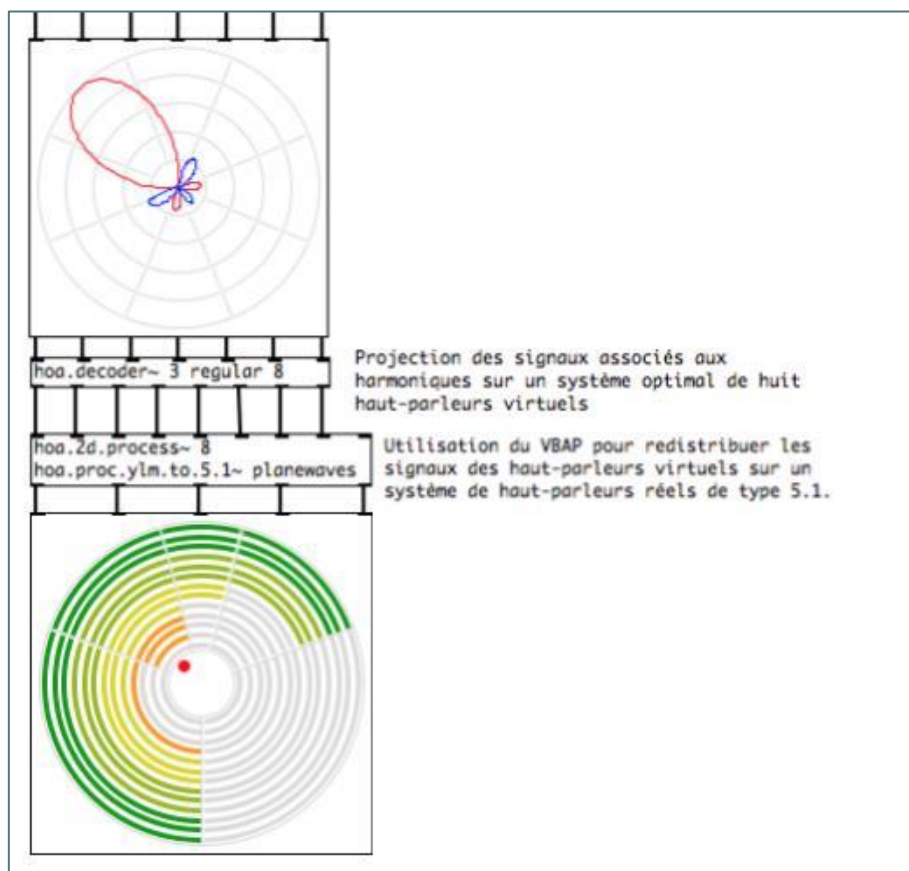


Figure 5.16. Patch Pure Data représentant le décodage combiné ambisonie-VBAP aussi nommé AllRAD d'un champ sonore au troisième ordre de décomposition, pour une restitution sur un système de haut-parleurs de type 5.1.

Néanmoins, une telle approche ne peut convenir aux techniques de synthèse de champs diffus présentées dans ces travaux, pour des raisons évidentes de diminution des contributions associées aux harmoniques de degré élevé. L'approche qui offre une solution acceptable à la fois pour des champs composés de sources sonores directionnelles et des champs diffus, part du principe qu'une restitution préservant les contributions de l'ensemble des harmoniques nécessite, quoi qu'il arrive une configuration de haut-parleurs optimale. Dès lors, une autre approche est intéressante : la combinaison ambisonie-binaurale. Cette approche peut être envisagée sous la forme de deux opérations distinctes. A savoir une première opération de décodage des composantes ambisoniques dans les directions des HRIR. Puis une deuxième opération d'utilisation des signaux résultats comme des sources sonores pour la synthèse binaurale, dans les directions des HRIR. En s'inspirant de cette approche, il est possible d'envisager une approche similaire mais adaptée à une restitution sur haut-parleur et non au

casque. Il s'agit dans un premier temps de projeter les signaux associés aux harmoniques sphériques sur un ensemble de points discrétisant parfaitement la sphère ou du moins, de manière optimale. Ces points peuvent alors être considérés comme des haut-parleurs virtuels résultant d'une prise de son optimale dans le champ sonore. Par la suite, il suffit donc de considérer ces prises de son comme des sources sonores d'une technique de spatialisation moins sujette au dispositif irréguliers, et de spatialiser les sources dans les directions qui leurs sont associées. Selon F. Zotter et M. Frank, la technique qui semble la plus adaptée alors est le *Vector Base Amplitude Panning* (VBAP) de V. Pullki [Pullki, 1997], qui offre une grande flexibilité dans la disposition des haut-parleurs et la possibilité de restituer des champs sonores en deux et trois dimensions [Figure 5.16]. Aussi proposent-ils une solution pour la restitution de champ sonore sous le nom de AllRAD, basée sur cette approche [Zotter & al., 2012].

Dans la bibliothèque HOA, une approche similaire est proposée, mais seulement pour une représentation bidimensionnelle avec l'option *irregular* de l'objet *hoa.2d.decoder~*. L'approche, comme suggéré, a été élaborée en parallèle et sans connaissance des travaux de F. Zotter et M. Frank. Aussi n'est-elle pas construite autour d'une fonction de panoramique d'amplitude telle que proposée par le VBAP, mais sur une approche plus personnelle inspirée par C. Roads [Roads, 1996], p. 460<sup>206</sup>. Il en résulte alors une diffusion légèrement plus importante dans la version de la bibliothèque HOA. De plus, étant donné qu'aucun système de haut-parleurs ne semblait optimal au-delà de l'ordre de décomposition 3 en trois dimensions et du fait que la fonction de panoramique d'amplitude utilisée ne soit pas définie pour une restitution tridimensionnelle, il est préférable pour le moment de restreindre l'approche à une restitution bidimensionnelle.

Afin d'offrir une meilleure solution, il est, à présent, envisagé d'utiliser l'approche AllRAD telle que présentée en [Figure 5.16]. L'idée étant de garder le processus de projection et de VBAP séparé<sup>207</sup> afin d'offrir plus de flexibilité. Pour le moment, les tests réalisés sont adaptés à chaque nouvelle configuration, néanmoins une bibliothèque a été créée afin de permettre de réaliser du VBAP<sup>208</sup> avec des abstractions Pure Data. Ce qui permet d'envisager un système

---

<sup>206</sup> L'approche étant plus de l'ordre de l'algorithme de programmation, intégrant des conditions qui rendent sa formulation mathématique indigeste, il n'a pas semblé nécessaire de la retranscrire ici.

<sup>207</sup> Il faut noter que la mise en œuvre pour le logiciel MAX n'offre pas cette possibilité car le décodage par projection est restreint à des configurations régulières. Cette limitation sera enlevée dans les futures versions.

<sup>208</sup> La bibliothèque est disponible sur le répertoire Github <https://github.com/pierreguillot/vbap>, consulté en septembre 2017. Elle offre en outre la possibilité de réaliser du VBAP et du MDAP. Les fonctionnalités sont disponibles sous la forme de bibliothèque indépendante en C, d'objets compilés et d'abstractions pour Pure Data.

plus dynamique, toujours en utilisant l'objet `hoa.2/3d.process~`. Cette approche permet de réaliser des tests en combinant d'autres techniques telles que le *Distance Based Amplitude Panning* (DBAP) [Lossius & al., 2009]<sup>209</sup>. Il pourrait aussi être envisagé de procéder de la sorte pour la synthèse binaurale, afin d'offrir notamment plus de flexibilité quant au module de synthèse binaurale utilisé, et au choix des HRIR. Ce type d'approche est dans l'ensemble bien plus extensible et offre en outre une meilleure pérennité. Afin de compenser la restriction de l'ordre de décomposition pour une restitution tridimensionnelle, il est proposé d'utiliser la discrétisation de la sphère avec cinquante haut-parleurs virtuels, tels que présentés par F. Zotter et M. Frank [Zotter & al., 2012].

Aussi cette approche offre une solution optimale pour la restitution de sources sonores directionnelles, qui donne toujours la possibilité d'améliorer leurs rendus avec les optimisations *max-r<sub>e</sub>* ou *in-phase*. Elle offre aussi une solution optimale pour la restitution de champs diffus. Cette approche nécessiterait cependant de réaliser plus de tests pour une restitution tridimensionnelle, mais les premiers résultats semblent convaincants. Étant donné que l'approche mise en œuvre pour le moment dans le cadre d'une restitution en deux dimensions a été largement validée sur un plan pratique par de nombreux usages, tout prêté à penser que l'extension à une troisième dimension ne devra pas générer plus d'artefacts. Il faut néanmoins noter qu'une telle approche agit en additionnant dans le signal d'un haut-parleur de nombreux signaux normalement destinés à plusieurs haut-parleurs. Aussi, cet effet de superpositions laisse envisager des phénomènes d'opposition de phases. Il pourrait dès lors être intéressant de confronter les résultats d'une analyse dans le temps du champ sonore après la projection, mais avant le VBAP et après le VBAP pour tenter de discerner les artefacts sur le champ sonore.

### 5.3. Analyse de l'approche

Dans ce dernier chapitre, deux des outils originaux et élémentaires du modèle ambisonique ont été présentés. Ces outils sont destinés à répondre à deux problèmes spécifiques. D'une part la critique des techniques ou des systèmes de spatialisation, et d'autre part la restitution d'un champ sonore pour un système de haut-parleurs. La critique d'une technique ou d'un système de spatialisation se fait par l'analyse du champ sonore, restitué sur le système de haut-parleurs via les vecteurs énergie et vitesse. Ces vecteurs étant des outils mathématiques élaborés sur des principes acoustiques et psychoacoustiques, ils permettent de quantifier la qualité de la

---

<sup>209</sup> Il semble évident de par la caractéristique du DBAP que cette technique soit moins adaptée. Ces recherches reviennent un peu plus sur cette technique dans le troisième chapitre de cette thèse.

restitution d'une source sonore directionnelle. L'élaboration des techniques de restitution consiste à définir les opérations permettant d'optimiser et de tirer au mieux parti d'un système de haut-parleurs et des informations du champ sonore, afin de valider les critères d'évaluation du champ sonore. Aussi existe-t-il une dépendance des techniques de restitution vis-à-vis des critères d'analyse du champ sonore. L'ensemble de ce système a été pensé, selon des critères acoustiques et psychoacoustiques ayant pour but ultime de reproduire au mieux les effets directionnels d'un champ sonore. Aussi, l'ensemble des mécanismes et des outils du modèle ambisonique tournent autour d'une notion de source directionnelle. Et malgré l'extension du modèle pour intégrer la notion de distance par la prise en compte des effets du champ proche ou des phénomènes de réverbération de l'acoustique des salles, l'objectif reste toujours centré autour de l'idée d'une source sonore considérée comme un point, et qui dans ce cas, pourra être plus ou moins loin<sup>210</sup>.

Dans le cadre de cette thèse, la notion de source directionnelle ou ponctuelle n'est pas remise en cause, loin de là. Mais son statut d'hégémonie en tant qu'objet élémentaire du champ sonore est contesté<sup>211</sup>. Par l'élaboration d'outils destinés à la spatialisation du son via les techniques ambisoniques adaptées à une approche expérimentale – notamment en donnant accès aux signaux associés aux harmoniques sphériques mais aussi en offrant un système particulièrement dynamique – il a été possible de prendre en main et de faire des expérimentations à partir de ce modèle. Ces travaux ont alors abouti à la mise en œuvre de nouvelles approches originales de synthèse de champ sonore et de traitements, dans le domaine des harmoniques sphériques. Il s'agit par exemple de la résolution angulaire, de la synthèse de champ sonore par décorrélation temporelle, par granulation ou par pondération aléatoire des signaux associés aux harmoniques ou encore de l'application d'un *flanging* dans le domaine des harmoniques sphériques. Plutôt que de viser exclusivement à synthétiser des sources sonores directionnelles ou ponctuelles ou des phénomènes acoustiques en relation avec cette notion de source sonore tels que les effets de salles, cette approche particulière du modèle ambisonique a permis de faire émerger des champs sonores aux caractéristiques spatiales et fréquentielles variées. Ces champs sonores possèdent des propriétés quelque peu analogues aux champs diffus associés notamment aux

---

<sup>210</sup> La source directionnelle étant aussi un point mais infiniment loin.

<sup>211</sup> Il est alors important de rappeler que cette prise de position n'est pertinente que par le contexte artistique et expérimental dans lequel se déroulent ces recherches. Ce travail n'a pas la prétention de contredire des années de recherches en acoustique et psychoacoustique mais justement de profiter de celles-ci en y portant un regard un peu différent.

réflexions tardives d'une réverbération, mais totalement libérées des contraintes de réalisme<sup>212</sup>. Et c'est à partir de ces nouveaux champs sonores que le renversement s'opère.

Tout le modèle ambisonique est construit autour de la notion de source sonore directionnelle ou ponctuelle, et concevoir un autre modèle de champ sonore implique irrémédiablement de remettre en cause les outils et les opérations définis originalement par l'approche ambisonique. En effet, ce ne sont plus seulement les effets directionnels d'un champ sonore qu'il faut reproduire, mais aussi des effets de masse, de décorrélation, d'élargissement de l'espace sonore, etc. Ce but principal est alors complété par un ensemble d'objectifs plus vastes et variés. Aussi, de nouveaux outils d'analyse ont été proposés afin de pouvoir critiquer les techniques de spatialisation visant à créer et restituer les champs sonores originaux présentés dans ces travaux. De façon plus juste, il a été présenté de nouvelles utilisations des outils déjà présents, afin de combler les pratiques établies. L'objectif étant d'offrir une approche plus générale permettant à la fois de s'adapter au modèle des sources sonores tel que défini originellement par le modèle ambisonique, et au nouveau modèle de champs sonores. A la suite de la présentation de ces outils d'analyses, ce sont les opérations permettant la restitution du champ sonore qui sont examinées. L'objectif est de proposer des manières d'optimiser et de resituer les champs sonores, cette fois-ci encore, plus adaptées au modèle de champ de champs sonores proposés dans ce travail. Tout en préservant dans cette démarche, la notion de source sonore originale. En considérant que le but principal de l'approche ambisonique est de s'élargir et qu'en conséquence les opérations et les outils se trouvent transformés, il est alors possible de se demander si cette approche est toujours de l'ambisonie.

Mais il reste très difficile de répondre à cette question. En effet, l'ambisonie est un ensemble de techniques, mais aussi de pratiques construites autour de principes acoustiques, psychoacoustiques et d'une représentation mathématique du champ sonore formant un noyau central et original, mais qui n'a eu de cesse de se modifier et de se développer. De même que certains voient en l'ambisonie une technique de panoramique d'amplitude [Neukom, 2007], [Neukom & al., 2007] et d'autres une approche holophonique<sup>213</sup> [Nicol & al., 1998], il est vain de vouloir restreindre l'approche ambisonique. Aussi, l'enjeu serait plutôt de définir le cœur

---

<sup>212</sup> A moins qu'une salle puisse donner l'impression de se trouver dans une tornade de grains de voix, mais j'en doute.

<sup>213</sup> A ce propos J. Daniel définit que « On peut ainsi voir en la restitution ambisonique un cas particulier de reconstruction holographique du champ ("*Holophonie*") », [Daniel, 2001], p. 116.



des expérimentations présentées dans le cadre de ces recherches. Il s'agit de définir l'élément du modèle ambisonique sur lequel les nouvelles approches de synthèse et de traitement du champ sonore ont été construites, et de voir sur un plan conceptuel les conséquences de cette appropriation. Au regard des mécanismes présentés lors des expérimentations, il paraît évident que le noyau central sur lequel ces mises en œuvre ont été élaborées, est la décomposition en harmoniques sphériques du champ sonore. L'élément intéressant est alors le changement notable opéré autour de cet outil mathématique dans son utilisation et dans sa finalité même.

En ambisonie, les harmoniques sphériques et l'opération de décomposition qui leur est liée permet d'offrir une représentation du champ sonore qui se révèle pertinente pour plusieurs raisons. Des raisons techniques et matérielles tout d'abord, car cette représentation peut être mise en relation avec un dispositif de prise de son. Bien qu'il soit aujourd'hui difficile de concevoir un microphone correspondant à la figure de directivité d'un encodage au dixième ordre de décomposition, la relation était originellement évidente lorsque l'ambisonie était restreinte au premier ordre de décomposition. Puis, il existe des raisons acoustiques et psychoacoustiques, car les signaux associés aux harmoniques sphériques correspondent à une représentation locale – au centre de la tête de l'auditeur – du champ sonore et qu'ils possèdent les informations nécessaires à une reconstruction optimale de ce champ sonore selon les mécanismes de la perception auditive humaine. Enfin, il faut souligner les raisons fonctionnelles intrinsèques à l'outil mathématique même, car la décomposition en harmoniques sphériques permet des transformations de type rotation ou déformation, mais aussi et surtout une utilisation à différents ordres de décomposition. Aussi originellement, la décomposition en harmoniques sphériques est l'outil mathématique soutenant le modèle ambisonique en permettant de répondre aux spécifications techniques, matérielles et fonctionnelles, tout en préservant les critères acoustiques et psychoacoustiques sur lesquels le modèle repose.

Dans le cadre des expérimentations présentées, la relation de la décomposition en harmoniques sphérique du champ sonore avec un dispositif matériel est tout simplement ignorée. Cela ne veut pas dire que le dispositif de restitution n'est pas pris en compte, mais que la question est annexe et arrive par la suite. En somme, l'élaboration même des traitements de synthèse et de transformation du champ sonore est libérée de cette relation. Suivant la méthodologie parcourue, c'est suite à la création des champs sonores que la question de la restitution se pose. Et il en est de même pour les raisons acoustiques et psychoacoustiques, qui animent l'approche ambisonique. Bien que ces questions perdurent au final dans l'approche, force est de constater que les expérimentations ne prennent pas en considération la question de savoir si les

informations générées sont en adéquation avec les mécanismes de la perception auditive humaine. Le rôle réel alors de la décomposition en harmoniques sphériques est d'offrir une représentation de l'espace simplement intéressante. Et bien de l'espace et non d'un champ sonore. Ou bien d'un champ sonore, mais libéré du formalisme acoustique. Il s'agit de concevoir une représentation de l'espace sur laquelle des sons et des traitements sont appliqués et sur laquelle il est possible de jouer. Comme cela a été évoqué précédemment, dans cette approche particulière de la spatialisation du son, ce sont les possibilités musicales et créatives qu'elle inspire qui rendent la décomposition en harmoniques sphériques pertinente. L'élément intéressant réside donc plutôt dans le modèle mathématique même des harmoniques sphériques et de leurs représentations spatiales. C'est en cela même que cette représentation peut être considérée comme abstraite. Par la suite, et comme pour de nombreuses approches, il reste néanmoins à restituer l'espace sonore. Et c'est en cela que cette représentation est intermédiaire.

Partant de cette analyse, il est possible d'envisager d'autres représentations de l'espace, offrant chacune leur lot d'opérations et de traitements de l'espace et du son et ouvrant chacune de nouvelles perspectives musicales.

## Partie III – Généralisation de l’approche ambisonique

Suite aux expérimentations dans le domaine des harmoniques sphériques et à l’appropriation musicale des techniques ambisoniques, les principes initiaux sur lesquels reposent ces techniques de spatialisation du son ont été remis en question. Ces approches ont notamment permis de révéler le potentiel musical de la représentation du champ sonore sous la forme de signaux associés aux harmoniques sphériques. Mais, de manière plus générale, elles ont amené à changer le regard porté sur la décomposition en harmoniques sphériques, en mettant en avant le potentiel inspirant et créatif de la représentation spatiale qui en découle. Celle-ci s’est trouvée libérée des éléments techniques, matériels, acoustiques et psychoacoustiques qui ont amené à son utilisation et s’est révélé être un outil puissant pour concevoir des espaces sonores. Cette représentation intermédiaire de l’espace est redevenue abstraite, dans le sens où elle n’est plus liée explicitement à une représentation rationnelle du champ sonore. Et par ce biais, elle laisse entrevoir de multiples autres possibilités de représentations abstraites de l’espace, pouvant être utilisées de manière intermédiaire, afin de générer des espaces sonores.

C’est à cette spéculation que cette partie tente de répondre. De nombreuses propositions concrètes peuvent dès lors apparaître. Mais le cadre de ces recherches pourrait à peine en recouvrir une infime partie, et, s’il était possible d’en présenter suffisamment pour offrir un aperçu du large éventail des possibilités, celles-ci ne seraient traitées que superficiellement. Aussi, le choix a été fait de concentrer les travaux sur une représentation particulière de l’espace. Cette approche a été préférée pour plusieurs raisons. D’une part, cela permet de présenter une approche en profondeur et de façon plus complète. Mais aussi de proposer des outils adaptés et surtout de concevoir des expérimentations et des utilisations concrètes, permettant de valider une proposition qui n’était jusqu’à présent que conceptuelle. D’autre part, l’enjeu par cette présentation est de concevoir en filigrane une méthodologie de l’approche qui puisse être réadaptée pour d’autres propositions de représentation de l’espace. En rapprochant cette approche de l’appropriation musicale des techniques ambisoniques, l’objectif est de définir des spécifications conceptuelles, techniques et fonctionnelles de la spatialisation du son fondées sur une représentation intermédiaire et abstraite de l’espace. Aussi, bien que la présentation au cœur de cette partie reste très centrée sur une approche particulière qui sera présentée par la suite, les enjeux auxquels elle tente de répondre sont bien plus généraux. Et de même, la méthodologie employée et les outils proposés se veulent les plus génériques possibles.

## 6. De nouveaux outils

Pour proposer de nouvelles approches et de nouvelles techniques de spatialisation fondées sur des représentations abstraites et intermédiaires de l'espace, il est nécessaire d'avoir des outils logiciels permettant de mettre en œuvre, mais surtout d'expérimenter, de telles approches. Cependant, aucun de ces nouveaux processus de mise en espace du son n'est pour le moment défini. Cela rend alors impossible la définition d'un ensemble de spécifications précises et d'outils dédiés à une technique particulière. Ce problème peut alors paraître sans fin. Aussi, pour répondre à ce sujet *a priori* sans issue, ce travail de recherche propose de porter sur l'analyse et la critique de la bibliothèque HOA, afin de déterminer les approches qui pourraient être généralisées et de définir ainsi certaines spécifications. Suite à cette étude, une solution concrète au problème sera proposée, sous la forme d'un petit ensemble d'outils logiciels pour le logiciel Pure Data. Cet ensemble d'outils tentera de répondre au mieux aux spécifications établies précédemment, en offrant un système flexible et des fonctionnalités qui pourront s'adapter à de nombreuses propositions. Avant de procéder à une première suggestion d'un nouveau système de spatialisation fondé sur une représentation intermédiaire et abstraite de l'espace, ces outils nécessitent d'être d'abord confrontés à un usage réel permettant de les critiquer. Il s'agira alors de mettre en œuvre via ces outils, certains traitements déjà réalisés avec la bibliothèque HOA afin de définir à l'avance leurs contraintes et leurs limites et de pouvoir proposer des solutions ou dans le cas contraire, en tenir compte pour les futures mises en œuvre.

### 6.1. Spécifications

L'objectif est définir les spécifications techniques et fonctionnelles d'un ensemble d'outils génériques qui pourraient s'adapter à de multiples représentations intermédiaires et abstraites de l'espace, toutes aussi différentes que variées. La méthodologie proposée est de revenir sur la mise en œuvre de la bibliothèque HOA afin de critiquer les approches, pour pouvoir envisager des fonctionnalités intéressantes répondant à de futurs besoins. L'enjeu est, en quelque sorte, de définir ce qui fait la spécificité de cette mise en œuvre, afin de la généraliser.

Selon les travaux présentés dans la première partie de cette thèse, la bibliothèque HOA tente de répondre à plusieurs problématiques. Il s'agit, d'une part, de répondre aux enjeux d'optimisation et d'utilisabilité des processus, notamment pour le contexte de musiques temps-réel. D'autre part, les approches proposées visent à satisfaire des enjeux didactiques et

pédagogiques, notamment en tentant de rendre intelligibles et transparents les mécanismes sous-jacent aux processus mis en jeu. Enfin, le point le plus important est d'offrir un système relativement flexible et adapté à l'approche ambisonique, qui permette de mettre en œuvre de nouvelles expérimentations de l'espace et du son via la création de traitements originaux dans le domaine des harmoniques sphériques. Aussi, afin de répondre à ces problématiques, trois classes d'objets ont été proposés : les objets de type « boîte » de traitement du signal, les interfaces graphiques, et les outils de vectorisation des processus dont l'élément central est l'objet *hoa.2/3d.process~*. Pour chacune de ces classes, en confrontant les mises en œuvre des objets avec les utilisations qui en sont faites, cela au regard des problématiques qui viennent d'être rappelées, il sera possible d'offrir un aperçu général des objectifs atteints ou, au contraire, des problèmes persistants. A partir de ces observations, le but est de définir les atouts des outils proposés, afin de reproduire les mécanismes en vue d'une utilisation plus générale, mais aussi, à l'opposé, afin de révéler leurs points faibles.

### 6.1.1. Traitement du signal

Comme cela a été énoncé pendant leurs présentations, les objets de traitement du signal, bien qu'offrant des avantages certains sur le plan de l'utilisabilité et de l'optimisation, ne semblent *a priori* pas la meilleure approche. Afin de justifier cette affirmation, il semble pertinent de revenir sur les deux opérations les plus communes de l'approche ambisonique, à savoir l'encodage et le décodage.

Les outils d'encodage sont au nombre de deux, l'objet *hoa.2/3d.encoder~* et l'objet *hoa.2/3d.map~*. Mais étant donné que le premier est une version restreinte du second<sup>214</sup>, il est possible de ne considérer que ce dernier. Cet outil permet d'encoder un nombre dynamique de sources sonores à n'importe quel ordre de décomposition. Il est donc à remarquer une certaine ergonomie de l'outil par son aspect dynamique et offrant la possibilité de s'adapter à plusieurs situations. De plus, les mécanismes mis en jeu sont relativement complexes de par les calculs des amplitudes des harmoniques sphériques. Les opérations sont réalisées à des fréquences relativement élevées, à savoir la fréquence d'échantillonnage audionumérique. Il est donc possible d'estimer que cette approche offre un avantage vis-à-vis de l'optimisation. Cependant, les mécanismes restent opaques à l'utilisateur. Afin de répondre à ce problème, un tutoriel spécifique à la question de l'encodage et des processus mis en œuvre a d'ailleurs été réalisé. Et,

---

<sup>214</sup> Il est d'ailleurs suggéré dans la première partie de cette thèse que le premier objet soit totalement remplacé par le second.

bien qu'elle réponde aux questions d'optimisation et d'utilisabilité, l'approche ne satisfait que très peu les enjeux didactiques et pédagogiques et il en est de même pour les enjeux expérimentaux. Selon les besoins actuels, il est évident que la nécessité d'offrir un système flexible, adaptable et transparent afin de rendre les mécanismes intelligibles, est une priorité. De ce fait, cette approche n'est que peu convaincante. Il a été remarqué lors de la mise en œuvre du *flanging* ambisonique, que l'opération d'encodage peut être réalisée sous forme de patch, rendant ainsi envisageable une autre approche, plus ouverte, bien que cela soit au détriment de l'optimisation. Mais dans ce contexte cette considération peut être négligée. Il est néanmoins nécessaire de déterminer si une telle approche reste suffisamment dynamique et flexible. De plus, cette approche présentée dans la mise en œuvre du *flanging* ambisonique est restreinte à une représentation en deux dimensions. Aussi les mécanismes en jeu évitent les calculs récursifs des polynômes associés de Legendre, facilitant la mise en œuvre. Pour avoir réalisé cette même opération pour une représentation tridimensionnelle de l'espace sous forme de patch, notamment à destination des tutoriels, force est de constater que l'approche est alors bien plus complexe et bien moins adaptée à une utilisation musicale. Et cela d'autant plus dans un contexte de temps-réel.

De l'autre côté, les trois opérations de décodage offertes par l'objet *hoa.2/3d.decoder~* définissent elles aussi un système relativement optimisé et flexible, en permettant de s'adapter à différents ordres de décomposition. Et comme cela a été présenté auparavant, un travail a été réalisé afin d'optimiser les calculs matriciels, notamment pour la restitution binaurale. Néanmoins, force est de constater que cette volonté d'optimiser les processus n'est pas sans conséquences. Cette approche a multiplié la charge de travail en utilisant tout d'abord une bibliothèque tierce qui n'était pas sous licence libre – un choix évidemment non idéal que ce soit en terme d'extensibilité et modularité, mais surtout pour des questions d'accessibilité des mécanismes en mis en jeu – puis en programmant manuellement les opérations. De plus, les processus du décodage binaural ou du décodage irrégulier peuvent être assimilés à une double opération, consistant premièrement en la projection du champ sonore puis, soit en une synthèse binaurale soit en du *vector base amplitude panning*. Or, une partie des optimisations réside dans l'utilisation d'algorithmes qui réunissent ces deux étapes en une seule. Par cette caractéristique l'approche est encore, d'une part, bien moins intelligible et d'autre part, restrictive vis-à-vis de la capacité d'adaptabilité et d'extension du système. Ce qui limite les possibilités d'expérimentations possibles. Une approche bien plus en adéquation avec ce contexte de recherche consisterait à ne proposer qu'une opération de projection qui puisse être réalisée pour n'importe quel système de haut-parleurs, à savoir sans limitation due au fait que la discrétisation

de l'espace puisse être irrégulière ou sous optimale<sup>215</sup>. Puis cette approche mettrait en œuvre par la suite les systèmes de spatialisation secondaires permettant une restitution au casque, via la synthèse binaurale, ou sur un système de haut-parleurs irrégulier, via le VBAP. Envisager un tel système permettrait aussi d'expérimenter l'approche avec d'autres systèmes de spatialisation secondaires tels que le *Distance Based Amplitude Panning* (DBAP) [Lossius & al., 2009], [Kostadinov & al., 2010] ou le *Virtual Microphone Control* (ViMiC) [Braasch, 2005]. Un autre avantage serait aussi de changer plus facilement les HRIR de la synthèse binaurale permettant vraisemblablement d'offrir un rendu bien plus optimal pour de nombreuses personnes. Enfin de manière générale, l'approche serait plus pédagogique encore si l'opération de projection était réalisée sous la forme d'un patch. Il est important de noter que ce procédé a déjà été mis en œuvre pour les tutoriels, afin de présenter la restitution bidimensionnelle. Et compte tenu du caractère statique de l'opération, le résultat était relativement optimisé. Contrairement à l'opération d'encodage où les angles des sources sonores virtuelles varient en temps réel, ici, les positions des haut-parleurs sont fixes. De même, un module de VBAP adapté à l'approche ambisonique a aussi été mis en œuvre sous forme de patch pour une restitution tridimensionnelle sur le système de haut-parleurs du *Stevens Institute of Technology* dans le cadre de la Pd Convention 2016. Il est donc envisageable que la deuxième opération du décodage soit aussi réalisée sous forme de patch.

### 6.1.2. Interfaces graphiques

Que ce soit pour l'opération d'encodage ou pour l'opération de décodage, tout amène à penser qu'une approche plus flexible et adaptée à un contexte expérimental puisse être envisagée. Partant de ce principe, il peut être envisagé de restreindre drastiquement l'utilisation d'objets de traitement du signal compilés sous forme de boîte. L'analyse identique est plus difficile à réaliser pour interfaces graphiques utilisateurs. Autant les interfaces graphiques de contrôle du champ sonore – l'objet *hoa.2/3.map* ou l'objet *hoa.space* – se sont révélées très peu utiles dans le cadre de ces recherches<sup>216</sup>, autant les interfaces de représentation du champ sonore, l'objet *hoa.2/3d.scope~* et l'objet *hoa.2/3d.meter~*, sont à tous points de vue indispensables. En effet, les présentations de la création et de la mise en œuvre des traitements dans le domaine des

---

<sup>215</sup> Cette spécification touche spécifiquement la mise en œuvre pour le logiciel Max qui limite la projection pour un système régulier lors d'une représentation bidimensionnelle du champ sonore.

<sup>216</sup> A vrai dire, la question de la source ponctuelle n'a été traitée qu'au travers d'autres approches. Et la mise en œuvre de traitements associés au domaine des ondes planes et de haut-parleurs virtuels tels que le filtrage spatial pour lequel l'objet *hoa.space* est relativement pertinent, n'a pas été traitée car elle est quelque peu hors de propos dans ce sujet (une première présentation de ces traitements est réalisée dans [Guillot, 2013]). Aussi cette affirmation peut sembler *a priori* gratuite. Cependant, il est évident que de tels outils visent à améliorer l'ergonomie et à faciliter la prise de la bibliothèque. Mais ils n'offrent que très peu d'utilité dans un cadre expérimental.

harmoniques sphériques ont révélé à la fois leur fort potentiel didactique, mais aussi leur grande utilité afin d'appréhender les résultats des opérations conjointes du son et de l'espace. Aussi est-il nécessaire de pouvoir offrir des outils analogues permettant la représentation du champ sonore. Il faut néanmoins remarquer un défaut pédagogique dans ces approches. En effet, la complexité des représentations et la réutilisation du code originel des objets pour le logiciel Max a orienté les mises en œuvre sous la forme d'objets compilés. Or ce type d'approche, comme cela a été remarqué pour les objets de traitement du signal de type « boîte » rendent opaques les mécanismes internes. Ainsi, autant pour l'objet *hoa.2/3d.meter~* les mécanismes en jeu peuvent paraître évidents en raison du caractère relativement usuel d'un *peak programme meter*, autant pour l'objet *hoa.2/3d.scope~* cela peut ne pas l'être. Pour beaucoup, il n'est pas clair que la représentation résulte d'une simple projection des signaux associés aux harmoniques sphériques, sur un nombre important de points dans l'espace. Or, une fois ce mécanisme assimilé, la lecture de l'interface devient beaucoup plus claire. En outre, comme pour les objets de traitement du signal, définir un système où les mécanismes soient transparents serait une bonne amélioration<sup>217</sup>. Cependant, la complexité des calculs de projection comme du rendu graphique, peut vraisemblablement empêcher la mise en œuvre sous forme de patches de tel sorte que cela puisse être envisagé avec le système de structures et de données graphique natif de Pure Data. Il faut à ce sujet rappeler qu'une bibliothèque tierce a été élaborée, pour faciliter la mise en œuvre de ces interfaces dans le logiciel Pure Data.

### 6.1.3. Outils d'expérimentation

Enfin, et cela semble évident, les objets de vectorisation des processus sont les outils qui ont répondu au mieux à l'ensemble des exigences de la bibliothèque. Ils sont ainsi au centre des expérimentations et mises en œuvre dans le cadre de ces recherches. L'objet *hoa.2/3d.process~* et les objets de gestions des entrées et des sorties offrent un système qui s'adapte particulièrement à l'approche multicanal nécessaire en ambisonie par la gestion dynamique de multiples instances d'une abstraction. Le système offre par ailleurs une gestion des informations relatives au contexte ambisonique. Par exemple, l'objet *hoa.thisprocess~* en fait un outil relativement flexible, en permettant notamment de mettre en œuvre des abstractions qui puissent s'adapter au nombre de dimensions spatiales, à l'ordre de décomposition et bien sûr aussi spécifiquement à une instance via le degré et l'ordre azimutal de l'harmonique sphérique

---

<sup>217</sup> Il est néanmoins évident que de simplement donner accès aux mécanismes internes ne suffit pas en soi à satisfaire tous les plans didactiques, mais cela y contribue énormément.



associé<sup>218</sup>. De plus, ce système est particulièrement adapté aux enjeux didactiques. En incitant les mises en œuvre des traitements sous la forme de patches, le système offre une représentation des mécanismes internes qui peut être tout à fait compréhensible par chacun<sup>219</sup>. Deux remarques sont néanmoins à faire. D'une part, il est fortement envisagé que de nombreuses opérations soient plus optimisées sous la forme d'un objet compilé. D'autre part, comme cela a été évoqué, certaines opérations semblent difficiles voire impossibles à mettre en œuvre. L'exemple récurrent est celui de l'encodage de sources sonores directionnelles pour une représentation tridimensionnelle de l'espace. Cette opération nécessite de calculer les polynômes associés de Legendre et l'approche la plus dynamique, et donc la plus adaptée, semble être l'utilisation des formules récursives<sup>220</sup>. L'enjeu est alors de calculer le polynôme associé de Legendre pour une élévation donnée dans le degré zéro puis d'utiliser ce résultat pour calculer les polynômes associés de Legendre de degré 1 et d'utiliser l'ensemble des résultats pour calculer les polynômes associés de Legendre de degré 2, etc. Pour le moment, les objets de communication sans connexions natives de Pure data ont suffi à répondre au problème, mais la lisibilité des abstractions en a souffert. Aussi, l'approche nécessiterait un système de communication bien plus évolué entre les différentes instances. Il faut aussi remarquer que ce procédé est d'autant plus compliqué dès lors que ces calculs sont réalisés en signal audio. Une telle approche nécessiterait de pouvoir partager, de la même manière que pour les messages de contrôle, des signaux audio entre les instances. Et cette approche demanderait surtout de définir l'ordre d'exécution de chaque instance afin d'assurer la justesse des calculs. En effet, dans la mise en œuvre actuelle, l'ordre d'exécution des instances est défini selon l'ACN inverse<sup>221</sup>. Alors que les calculs relatifs à un harmonique sphérique sont en cours, les résultats qui leur sont nécessaires et qui sont les résultats des harmoniques sphériques de degrés inférieurs, ne sont pas encore réalisés. Aussi, le système récursif s'en trouve bloqué. Ce problème se retrouve aussi pour la mise en œuvre de l'opération de rotation sous la forme d'une abstraction où chaque signal associé à un harmonique est résultat d'une combinaison avec d'autres signaux. Ainsi, il est nécessaire au sein d'une instance de pouvoir accéder à certains signaux associés à d'autres

---

<sup>218</sup> De plus l'objet *hoa.2/3d.thisprocess~* est aussi utilisable dans le domaine des ondes planes en permettant de mettre en œuvre des abstractions qui puissent s'adapter au nombre maximum d'onde plane discrétisant l'espace et aux indices associés aux ondes planes.

<sup>219</sup> Sous réserve bien entendu qu'un travail de lisibilité soit réalisé dans ce sens.

<sup>220</sup> Une autre approche peut être de prédéfinir des équations simplifiées jusqu'à un certain point, afin d'éviter les calculs récursifs. Mais en plus d'offrir un système moins dynamique, cette approche implique potentiellement une moins bonne lisibilité de l'abstraction réalisant l'encodage et forcément une moins bonne visualisation et compréhension des relations et de la hiérarchisation entre les harmoniques sphériques.

<sup>221</sup> Ce n'est pas le résultat d'une spécification particulière mais le résultat de l'ordre inversé par lequel les instances sont chargées.

instances, alors que celles-ci n'ont toujours pas été exécutées dans la chaîne de traitement du signal.

Compte tenu de ces différentes analyses, il est envisageable que le système le plus adapté réside en une série d'outils logiciels similaires, permettant le chargement dynamique de plusieurs instances d'une même abstraction, et qui puisse s'adapter à différentes représentations de l'espace. Les questions d'optimisation ou de calculs complexes peuvent être investiguées à la suite des premières expérimentations et difficultés de mise en œuvre rencontrées. Un élément essentiel de ces analyses est qu'au-delà de l'élaboration de nouveaux outils, celles-ci peuvent évidemment servir aux futurs développements de la bibliothèque HOA.

## 6.2. La bibliothèque *spam*

La bibliothèque *spam*<sup>222</sup> est un petit ensemble d'objets pour le logiciel Pure Data dont le développement a commencé début 2016. L'objectif est d'offrir un outil similaire à l'objet *hoa.2/3d.process~* de la bibliothèque HOA, mais offrant des fonctionnalités moins spécifiques à l'approche ambisonique, afin de pouvoir répondre à un plus grand nombre d'utilisations possibles.

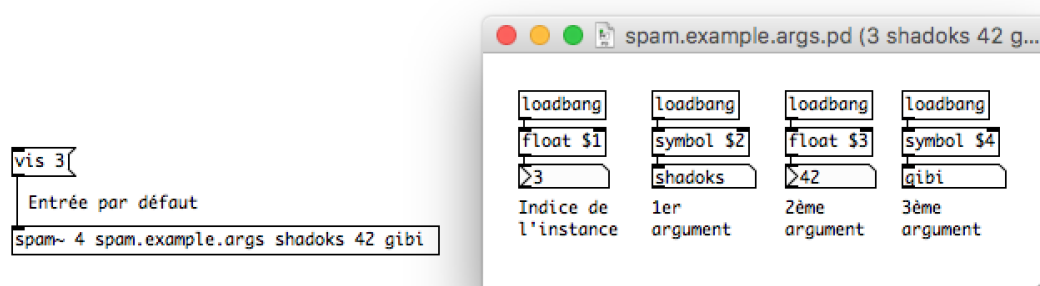


Figure 6.1. Exemple type du mécanisme de l'objet *spam~*. L'abstraction *spam.example.args* est instanciée quatre fois avec les trois arguments supplémentaires « shadoks 42 gibi ». A droite, la quatrième instance chargée et correspondant à l'indice 3 est visible. Le premier argument d'une instance est son indice, pouvant dans ce cas aller de 0 à 3, et les autres arguments sont les arguments supplémentaires, dans ce cas « shadoks 42 gibi ».

<sup>222</sup> La bibliothèque est disponible en ligne sous licence BSD sur le répertoire Github <https://github.com/pierreguillot/spam> et consultée en septembre 2017. Des versions précompilées des objets sont disponibles pour les systèmes Linux, MacOS et Windows en 32 et 64 bits.

La bibliothèque se forme d'un objet principal *spam~*, l'analogue de l'objet *hoa.2/3process~* de la bibliothèque HOA. Une caractéristique importante est qu'il n'existe pas de version de l'objet dépendante du nombre de dimensions spatiales. Ce parti pris implique de ne pas mettre en œuvre de mécanismes spécifiques au nombre de dimensions et de laisser l'utilisateur gérer lui-même cette partie du système. Bien que potentiellement plus complexe, il s'agit là d'offrir un système suffisamment flexible pour s'adapter à toutes les pratiques. Aussi, s'il convient à l'utilisateur de mettre en œuvre un système à seulement une dimension ou à plus de trois dimensions<sup>223</sup>, cela est en théorie réalisable. De manière similaire aux objets *hoa.2/3process~*, *clone* de Pure Data ou *poly~* de Max, l'objet *spam~* permet d'instancier autant de fois que voulu une abstraction. Le premier argument de l'objet est le nombre d'instances à charger et le deuxième argument est le nom de l'abstraction à charger [Figure 6.1]. Suite à ces deux arguments, il est possible de définir autant d'arguments supplémentaires que souhaité.

### 6.2.1. Gestion des arguments

Alors que la bibliothèque HOA utilise un mécanisme de gestions des arguments et des attributs identique à celui du logiciel Max<sup>224</sup>, le choix a été fait pour la bibliothèque *spam* de n'entretenir que le mécanisme natif de gestion des arguments de Pure Data. Ce choix a été réalisé pour deux raisons. D'une part, le mécanisme de gestion des attributs du logiciel Max est relativement peu adapté au logiciel Pure Data car il ne fait pas partie des usages classiques. D'autre part, cette bibliothèque étant développée exclusivement pour Pure Data, du moins pour le moment, il n'est pas nécessaire de faciliter la compatibilité des approches entre les deux logiciels. Enfin, les dernières versions du logiciels Pure Data offrent un mécanisme de gestion des messages via la suite d'objets *text* relativement avancé, et qui jusqu'alors, a suffi à l'ensemble des expérimentations. Selon le mécanisme de la bibliothèque *spam*, chaque instance possède comme premier argument son indice, en commençant par 0, puis les arguments supplémentaires donnés à l'objet *spam~*. Il est donc à noter que le nombre total d'instances chargées n'est pas défini en tant qu'argument dans les instances. Il s'agit de répondre aux mêmes spécifications qu'établies précédemment, afin de ne pas orienter les usages, car cela implique possiblement d'en restreindre d'autres. Il est, en effet, possible d'envisager un système où ce nombre total d'instances n'est pas nécessaire et qui dans ce cas, peut prêter à confusion. Aussi, si l'utilisateur

---

<sup>223</sup> La pertinence d'une telle approche peut être critiquée mais si déjà, sur un plan conceptuel cela inspire le compositeur, il est intéressant de proposer cette possibilité.

<sup>224</sup> L'utilisation de ce système est notamment disponible via l'objet *hoa.thisprocess~* de la bibliothèque HOA.

envoie ce nombre total d'instances à chacune des instances chargées, il doit le définir une seconde fois dans l'objet *spam~* en tant qu'argument supplémentaire<sup>225</sup>.

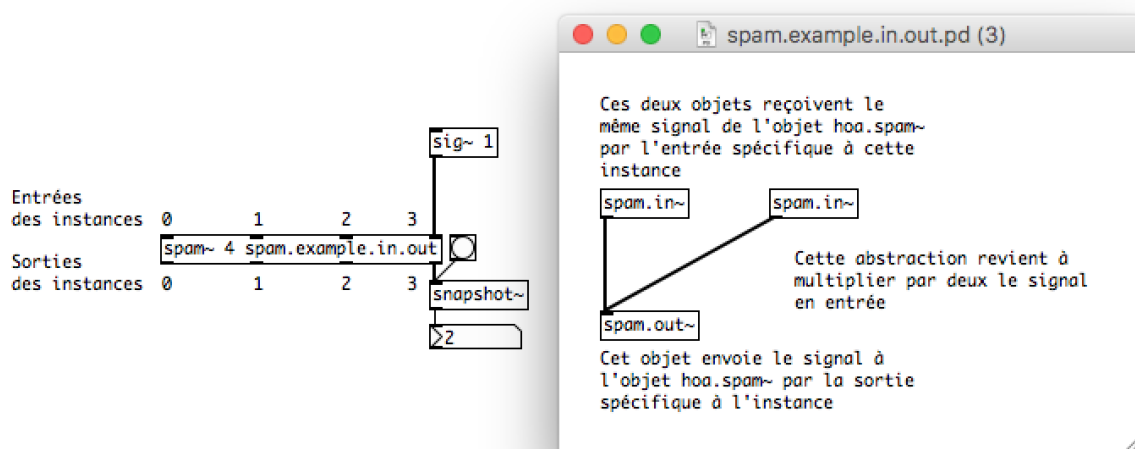


Figure 6.2. Exemple de la gestion des entrées et sorties *signal* spécifiques à chaque instance par l'objet *spam~*. A droite, la quatrième instance chargée est visible. Les deux objets *spam.in~* correspondant à la même entrée de l'objet *spam~* renvoie le même signal à l'objet *spam.out~*. Aussi le signal en sortie de l'objet *spam~*, est le double du signal en entrée.

## 6.2.2. Gestion des entrées et sorties spécifiques

La gestion des entrées et des sorties est relativement similaire à celle de la bibliothèque HOA. La bibliothèque *spam* offre quatre objets qui peuvent être utilisées dans les abstractions afin de gérer la liaison entre l'instance et l'objet *spam~* dans lequel celle-ci est chargée. Deux objets *contrôle*, *spam.in* et *spam.out*, sont disponibles pour la réception et l'envoi de messages et deux objets *signal*, *spam.in~* et *spam.out~*, sont disponibles pour la réception et l'envoi de signaux audio. Sans argument, les objets *signal* sont spécifiques à chaque instance. Aussi si une abstraction possède un objet *spam.in~* sans argument, une entrée est créée sur l'objet *spam~* pour chaque instance. Et de même si une abstraction possède un objet *spam.out~* sans argument, une sortie est créée sur l'objet *spam~* pour chaque instance. Suite aux pratiques de la bibliothèque HOA, il a été remarqué qu'il n'est jamais nécessaire d'offrir plus d'une entrée ou d'une sortie spécifique par instance. Aussi si une abstraction possède plusieurs objets *spam.in~*

<sup>225</sup> Il est aussi possible de définir ce nombre par d'autres mécanismes, via un objet *text* par exemple, comme suggéré précédemment, et dont le nom, lui est donné en tant qu'argument supplémentaire.

ou plusieurs objets *spam.out~*, ceux-ci ne sont associés en réalité qu'à un seul et même signal [Figure 6.2].

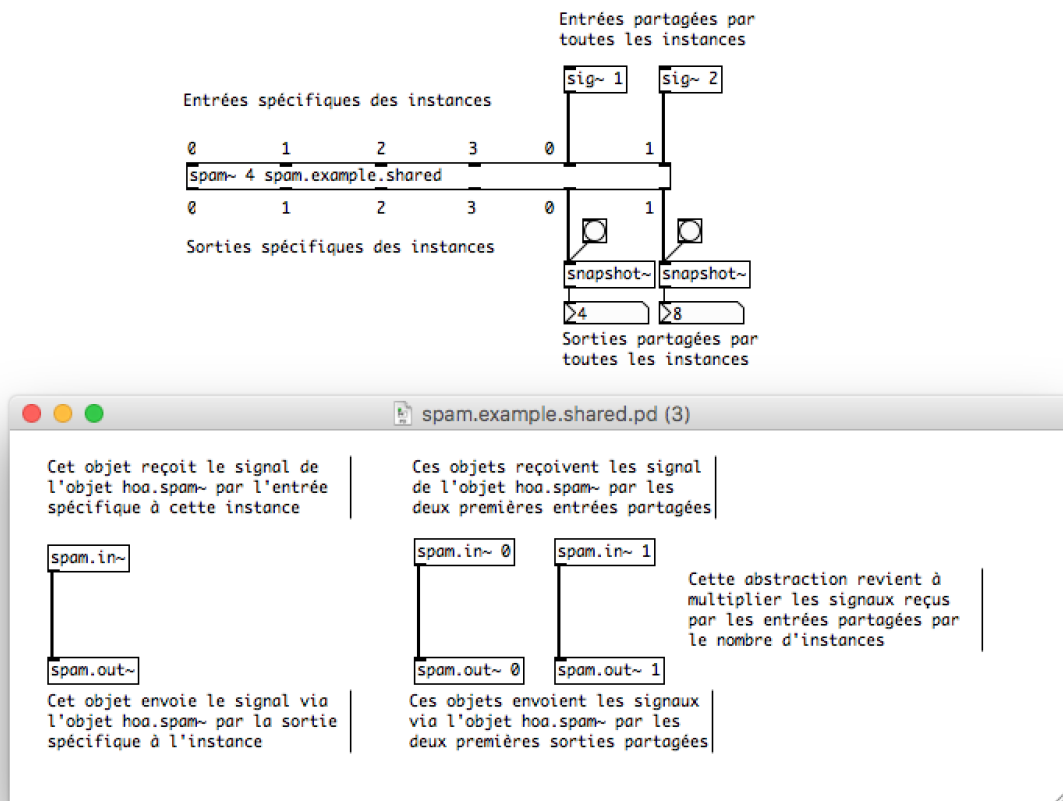


Figure 6.3. Exemple de la gestion des entrées et sorties *signal* partagées par toutes les instances par l'objet *spam~*. En dessous, la quatrième instance chargée est visible. Chaque objet *spam.in~* possédant un indice reçoit le signal par l'entrée partagée relative à cet indice de l'objet *spam~*. Chaque objet *spam.out~* possédant un indice signal additionne le signal qu'il reçoit et le résultat est envoyé à la sortie partagée relative à cet indice de l'objet *spam~*. Aussi comme il y a quatre instances, les signaux des sorties partagées de l'objet *spam~*, sont les quadruples des signaux reçus par les entrées partagées.

### 6.2.3. Gestion des entrées et sorties partagées

Les objets signal peuvent aussi prendre en argument un indice, en commençant par 0, les définissant comme des entrées et des sorties partagées et dont leurs positions dans l'objet *spam~* sont définies par la valeur de l'indice. Les entrées et les sorties partagées sont toujours après les sorties et les entrées spécifiques à chaque instance [Figure 6.3]. Toutes les instances reçoivent par les objets *spam.in~* possédant un indice, le même signal que l'entrée relative à cet indice de l'objet *spam~*. Et les objets *spam.out~* possédant le même indice additionnent les signaux reçus par toutes les instances et envoient le résultat via la sortie partagée relative à cet indice de l'objet *spam~*. En soi, le mécanisme est ici similaire à celui des entrées et sorties partagées offert par

les objets *hoa.in~* et *hoa.out~*, à l'exception de l'argument *extra* qui n'est plus nécessaire car l'ajout d'un indice comme argument, défini automatiquement les entrées et sorties partagées. Ce mécanisme n'était pas possible dans la bibliothèque HOA car les entrées et les sorties spécifiques pouvaient dépendre aussi d'un indice.

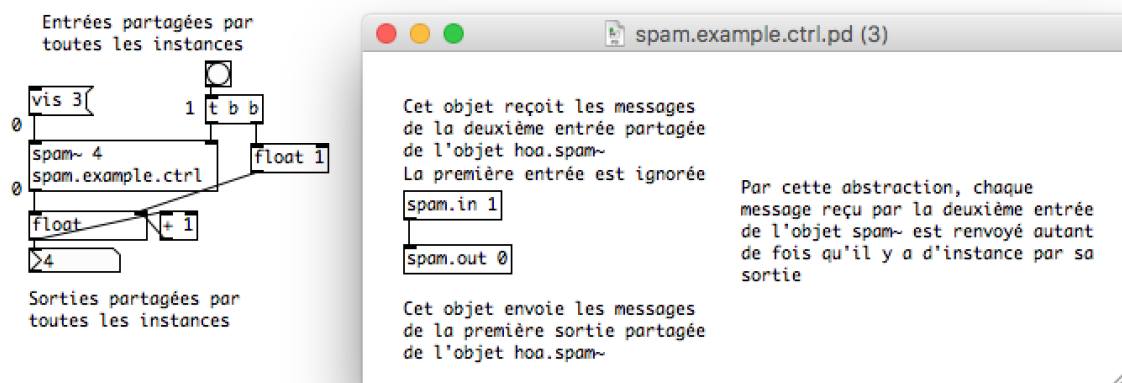


Figure 6.4. Exemple de la gestion des entrées et sorties *contrôle* par l'objet *spam~*. A droite, la quatrième instance chargée est visible. L'objet *spam.in 1* reçoit les messages de la deuxième entrée, relative à l'indice 1, de l'objet *spam~* et envoie ce message à l'objet *spam.out 0* qui renvoie le message par la sortie de l'objet *spam~*. Aussi comme il y a quatre instances et que les entrées et sorties *contrôle* sont partagées, chaque message reçu est renvoyé quatre fois. Dans le patch principal, l'envoi d'un message *bang* comptabilise quatre messages *bang* en sortie.

Les entrées et les sorties *contrôle* fonctionnent de manière identique aux entrées et sorties *signal* à la différence que celles-ci ne peuvent pas être spécifiques à chaque instance mais sont obligatoirement partagées par l'ensemble des instances [Figure 6.4]. Suite aux usages dans la bibliothèque HOA, il a été remarqué que l'utilisation d'entrées et de sorties de *contrôle* spécifiques à chaque instance était très rare. En effet, cela s'explique car ces entrées et sorties *contrôle* spécifiques nécessitent par la suite de générer ou de traiter indépendamment chaque message relatif à une instance. Cela peut être laborieux et n'apporte pas forcément plus de flexibilité à l'approche. Avec des entrées *contrôle* spécifiques, il est possible d'imaginer un système où chaque instance nécessite de recevoir une valeur spécifique, comme une fréquence pouvant être associée à un oscillateur, pour avoir un exemple concret. Alors l'utilisateur peut créer une entrée spécifique à chaque instance et connecter une boîte nombre à chacune des entrées pour envoyer les différentes valeurs, et contrôler indépendamment chaque oscillateur spécifique à chaque instance. Cependant, un fonctionnement exactement similaire peut être obtenu en utilisant une entrée partagée et en envoyant une liste de valeurs où chaque valeur est associée à une instance. Chaque instance de l'abstraction peut alors sélectionner en fonction de

son indice la valeur qui lui est spécifique dans la liste. En somme, les entrées et les sorties *contrôle* spécifiques à chaque instance ne permettent pas plus de possibilités de mises en œuvre que les entrées et les sorties *contrôle* partagées. Aussi ce système ne fait que complexifier les mécanismes en proposant une deuxième approche, et en augmentant grandement le nombre d'entrées et de sorties. Ainsi, il paraît plus judicieux de simplement enlever ce mécanisme.

L'avantage de la gestion d'un système simple et restreint tel qu'il est offert par la bibliothèque *spam*, est double. En effet, il permet d'une part d'assurer une meilleure pérennité des outils. Les mises en œuvre étant plus simples et plus en adéquation avec les mécanismes internes et les usages conventionnels du logiciel Pure Data, le système devrait être moins sujet à des problèmes de maintenance. D'autre part, en limitant l'aspect spécifique des mécanismes, celle-ci devrait offrir une meilleure adaptation aux différentes approches. Aussi, l'enjeu est à présent de procéder à une première mise en contexte des objets, afin d'en envisager les possibilités et les limites. Et l'objectif est de déterminer si cet ensemble d'outils logiciels offre une flexibilité acceptable et une capacité d'adaptation suffisante pour pouvoir mettre en œuvre des opérations similaires à celles présentées dans le cadre de ces recherches. Il paraît aussi intéressant de déterminer si ce type d'approche peut aussi s'étendre à des opérations matricielles plus complexes ou si elle permet de mettre en œuvre des interfaces graphiques de représentation d'espace sonore sous la forme d'un système multicanal.

### 6.3. Mise en contexte

Afin de vérifier que le système puisse satisfaire au maximum ces exigences préétablies, l'approche propose de mettre en œuvre de deux objets critiques, une opération de projection des signaux associés aux harmoniques sphériques, l'équivalent sous la forme d'une abstraction de l'objet *hoa.2d.decoder~* en mode *regular*<sup>226</sup>, et une interface graphique de représentation du champ sonore dans le domaine des harmoniques sphériques, l'équivalent sous la forme d'une abstraction de l'objet *hoa.2d.scope~*.

---

<sup>226</sup> Compte tenu des précédentes analyses, les autres modes de décodage ne semblent pas pertinents dans cette approche.

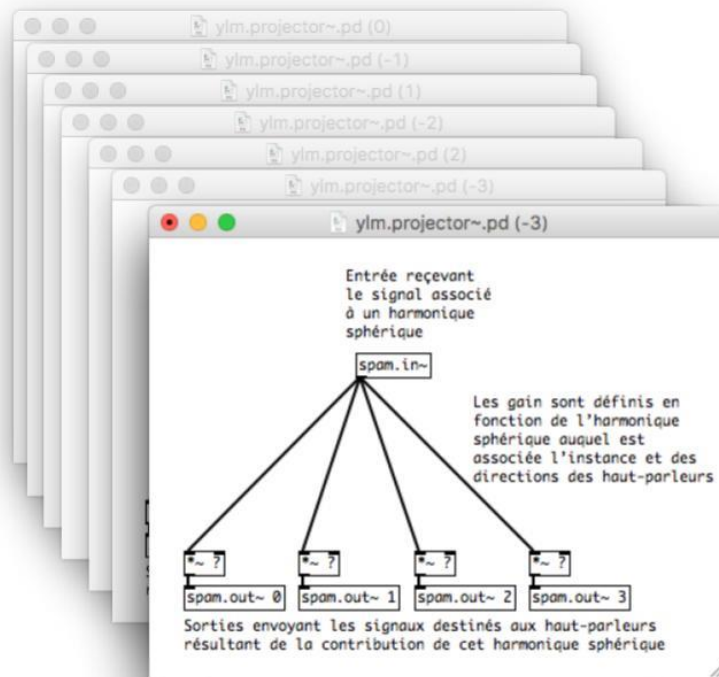


Figure 6.5. Ensemble de patches Pure Data, non fonctionnels, représentant le système de matricage pour la projection des signaux associés aux harmoniques sphériques vers un système de quatre haut-parleurs. Chaque instance du patch *ylm.projector~* est associée à un harmonique sphérique et génère les contributions du signal adjoint pour le système de haut-parleurs.

### 6.3.1. Approche matricielle

La première approche tente de réaliser la projection des signaux associés aux harmoniques sphériques pour un ensemble de haut-parleurs. Cette opération globale peut, comme cela a été vu, être décomposée en une série d'opérations élémentaires consistant à projeter chaque signal associé à un harmonique sphérique dans la direction de chaque haut-parleur. En somme cette opération revient à un matricage des canaux consistant à envoyer le signal de chaque entrée, associée à un harmonique sphérique, vers l'ensemble des sorties, associées aux haut-parleurs, selon des pondérations différentes. Le nombre d'entrées doit donc dépendre de l'ordre de décomposition et le nombre de sorties du nombre de haut-parleurs. Dès lors, il est possible d'envisager un système où chaque instance est spécifique à un harmonique sphérique et possède alors une entrée spécifique, pour le signal associé à l'harmonique sphérique, et autant d'entrées partagées qu'il y a de haut-parleurs, pour les signaux destinés aux haut-parleurs<sup>227</sup>. L'opération consiste à envoyer le signal d'entrée dans chaque sortie selon un gain défini en fonction de

<sup>227</sup> Le système inverse aurait aussi été envisageable en considérant que chaque instance est spécifique à un haut-parleur et possède alors une entrée spécifique, pour le signal destiné au haut-parleur, et autant d'entrées partagées qu'il y a de harmoniques sphériques, pour les signaux associés aux harmoniques sphériques.



l'harmonique sphérique auquel est associée l'instance et la direction du haut-parleur associé à la sortie [Figure 6.5].

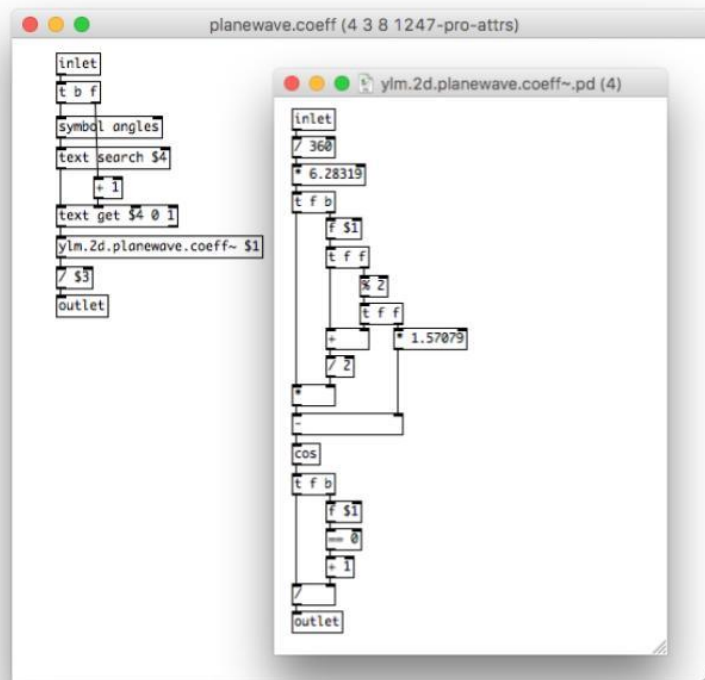


Figure 6.6. Patch *planewave.coeff* qui récupère l'angle d'un haut-parleur via un objet *text* donné en argument et l'envoi à l'abstraction *ylm.2d.planewave.coeff~.pd* qui calcule le coefficient d'un harmonique sphérique défini par son ACN donné en argument dans cette direction.

Le premier enjeu est donc de générer la série de gains associée aux haut-parleurs pour un harmonique sphérique. Afin d'offrir un système dynamique et sans limite du nombre de haut-parleurs, l'approche propose que les angles soient définis dans un objet *text* sous la forme d'une série de valeurs successives sur une ligne, commençant par le symbole *angles*. Par ce système, l'abstraction n'a besoin que du nom associé à l'objet *text* pour récupérer la liste des angles, offrant la possibilité que celle-ci soit de taille variable. Le coefficient de l'harmonique sphérique peut alors être calculé pour chaque direction selon le degré et l'ordre azimutal de l'harmonique sphérique. En considérant que l'indice de chaque instance équivaut à l'ACN, il est possible de redéfinir ces informations selon l'Équation 3.1]. Aussi, l'approche propose de définir l'ordre de décomposition, le nombre de haut-parleurs et le nom de l'objet *text* en tant qu'arguments

supplémentaires de l'objet *spam~*<sup>228</sup>. Une fois que l'abstraction permettant de réaliser la projection est chargée dans l'objet *spam~*, chaque instance de cette abstraction possède comme argument son indice correspond à l'ACN de l'harmonique. Puis cette série d'arguments permet de calculer les gains associés à chaque haut-parleur.

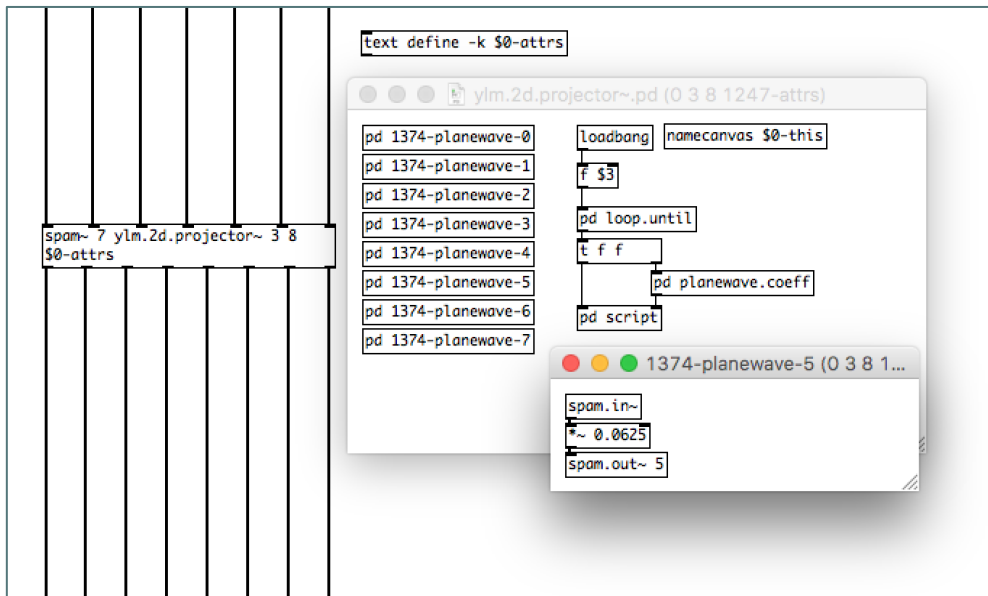


Figure 6.7. Abstraction *ylm.2d.projector~* chargée par l'objet *spam~* permettant de projeter un champ sonore bidimensionnel dans le domaine des harmoniques sphériques vers un système de haut-parleurs. Dans cet exemple, l'ordre de décomposition est 3 et le système de restitution comporte huit haut-parleurs dont les angles sont définis dans l'objet *text* se nommant *\$0-attrs*. Le premier argument est le nombre total d'instances qu'il est nécessaire de spécifier. La première instance de l'abstraction est visible à droite et possède huit sous-patch générés dynamiquement afin de réaliser l'opération de matricage nécessaire à la projection.

La dernière difficulté réside alors en la génération automatique dans chaque instance d'une série de sorties partagées associés aux haut-parleurs. Pour cela, l'approche propose d'utiliser un système de *patching* dynamique permettant de générer autant de sous-patches qu'il y a de haut-parleurs. Et où chaque sous-patch génère la contribution associée à l'harmonique sphérique en fonction de la direction du haut-parleur [Figure 6.7]. Grâce au système de gestion des entrées et sorties de la bibliothèque, ce sous-patch peut prendre une forme très simple, se résumant en un objet *spam.in~* récupérant le signal associé à l'harmonique. Ce signal est alors multiplié par le coefficient défini selon la direction du haut-parleur, et envoyé en sortie via l'objet *spam.out~* dont l'argument est l'indice du haut-parleur.

<sup>228</sup> Il est cependant possible de remarquer que dans ce contexte, l'ordre de décomposition n'est pas réellement nécessaire pour calculer le degré et l'ordre azimutal d'un harmonique. De même, le nombre de haut-parleurs peut être déduit de la taille de liste définissant les angles des haut-parleurs. Cependant, ces informations sont préservées afin d'assurer une homogénéité des pratiques et des usages avec celles d'autres abstractions destinées à des usages conjoints.

Au final, cette approche permet de réaliser un système tout aussi dynamique que l'objet compilé *hoa.2d.decoder~*, offrant la possibilité de s'adapter à tout ordre de décomposition et toute configuration de haut-parleurs. Un système de restitution tridimensionnelle pourrait tout aussi bien être envisagé. En effet, étant donné que les coefficients des harmoniques sphériques sont calculés seulement à l'initialisation sur des données de contrôle, la mise en œuvre est grandement simplifiée. Enfin, comme les opérations sur le signal se résument à des multiplications par des scalaires, la charge d'utilisation du processeur reste relativement faible et laisse envisager un bon fonctionnement dans un contexte de musique temps-réel.

### 6.3.2. Approche graphique

La deuxième approche propose de réaliser une interface graphique de visualisation du champ sonore, dans le domaine des harmoniques sphériques. L'enjeu est alors similaire, il s'agit de préserver la lisibilité des mécanismes en jeu, en évitant l'utilisation d'un langage compilé. Aussi, l'objectif est de réaliser à la fois les calculs et les dessins sous la forme d'une abstraction. Pour la représentation, afin d'offrir des outils relativement pérennes et multiplateformes, l'approche propose d'utiliser le système de structure de données graphiques natif de Pure Data.

Le premier mécanisme à mettre en œuvre est simplement un système permettant de transformer le signal de chaque harmonique en une valeur de contrôle selon une fréquence donnée<sup>229</sup>. Pour ce type de représentation, il est nécessaire de prélever la valeur d'un échantillon dans le signal, à un instant donné. Mais il n'est pas souhaitable de réaliser une discrimination des valeurs selon leurs amplitudes, comme le ferait un PPM par exemple. Cela impliquerait, par exemple, que la valeur prélevée sur le signal associé à un harmonique ne corresponde pas sur le plan temporel à la valeur prélevée sur le signal associé à un autre harmonique. Auquel cas, les valeurs représentant les amplitudes des signaux associés aux harmoniques ne seraient pas corrélées et la figure représentant le champ sonore serait déformée. L'enjeu est donc d'assurer que les valeurs soient prises de manière synchrone sur l'ensemble des signaux associés aux harmoniques. Cette opération peut être réalisée indépendamment dans chaque instance d'une abstraction chargée par l'objet *spam~* et simplement à l'aide de l'objet *snapshot~* et d'un objet *metro* [Figure 6.8]. L'avantage de Pure Data est qu'il assure dans ce cas un bon fonctionnement

---

<sup>229</sup> En effet, il ne s'agit pas de réaliser les calculs et de dessiner à la même fréquence que le traitement du signal. Cela serait extrêmement coûteux en termes de charge d'utilisation du processeur et ne servirait à rien étant donné que la fréquence de rafraîchissement de l'écran d'un ordinateur est bien plus petite.

du système. Sauf lors de la mise en place d'une utilisation vraiment spécifique, notamment avec l'objet *pd~*, Pure Data ne fonctionne que sur un seul fil d'exécution du processeur assurant que les objets *metro*, génèrent des messages *bang* de façon synchrone tant que leurs fréquences restent identiques et que leur déclenchement est de même synchrone. Et cela, bien que les messages soient répartis sur différentes instances<sup>230</sup>. Or, c'est bien le cas dans cette mise en œuvre de l'objet *spam~*.

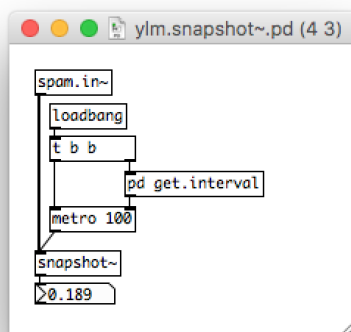


Figure 6.8. Patch utilisé conjointement avec les objets de la bibliothèque *spam~* permettant de transformer le signal associé à chaque harmonique en une valeur de contrôle selon une fréquence donnée.

L'enjeu consiste à projeter les valeurs associées aux harmoniques sphériques dans un nombre de directions suffisamment important, afin d'offrir une représentation relativement précise visuellement de l'espace sonore. De manière analogue à l'approche qui vient d'être proposée pour la projection du champ sonore vers un système de haut-parleurs, chaque instance réalise la projection relative à l'harmonique auquel elle est associée. La complexité réside alors dans le fait d'additionner terme à terme les résultats des projections relatives à chaque harmonique sphérique. Le problème revient donc à pouvoir communiquer les différentes instances. Pour cela, l'approche propose d'utiliser une mémoire tampon dont le nom est défini par un des arguments supplémentaires de l'abstraction. Aussi, afin d'assurer que ce nom soit unique, l'argument choisi est le nom du patch graphique, généralement un *Graph-on-Parent* (GOP), dans lequel la représentation sera dessinée et auquel l'indice de l'instance est rajouté. Dans la pratique, il existe donc une mémoire tampon par instance, mais seule celle relative à la première instance est utilisée [Figure 6.9]. L'opération consiste alors pour chaque instance à incrémenter

---

<sup>230</sup> Dans Pure Data et pour cette utilisation, la notion de simultanéité est relative à l'exécution des blocs de traitement du signal audio.

dans chaque direction la valeur associée dans la mémoire tampon, par la valeur résultant de la projection relative à l'harmonique dans cette direction.

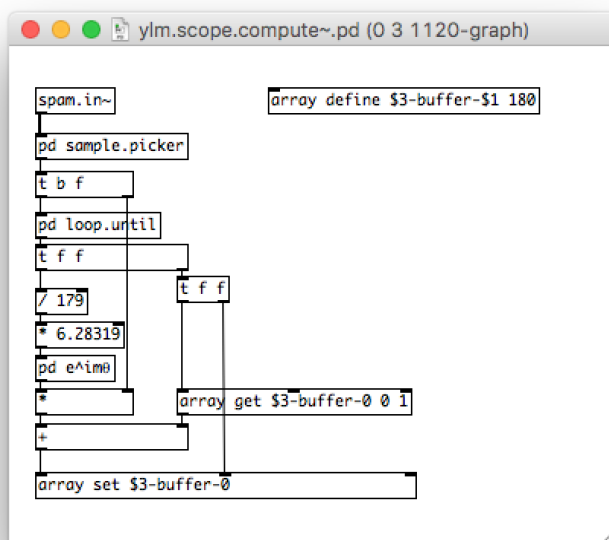


Figure 6.9. Patch utilisé conjointement avec les objets de la bibliothèque *spam~* permettant de transformer le signal associé à chaque harmonique en une valeur de contrôle selon une fréquence donnée, puis de projeter la valeur dans 180 directions discrétisant un cercle et d'additionner les résultats dans une mémoire tampon globale à l'ensemble des instances chargées.

La mise en œuvre actuelle de Pure Data assure que les chaînes de traitement audio des patches soient réalisées dans l'ordre inverse de leurs chargements. Etant donné que l'objet *spam~* charge chaque instance suivant leur indice, il est possible d'assumer que les chaînes de traitement audio des instances soient réalisées dans l'ordre inverse de leurs indices<sup>231</sup>. Un tel système permet de définir que la dernière instance à réaliser les calculs pour un ensemble de valeurs, est la première. Aussi, il convient à cette instance de réaliser le dessin dans le patch graphique, via le système de structure de données graphiques de Pure Data. Puis il s'agit de remettre l'ensemble des valeurs de la mémoire tampon à zéro pour pouvoir réaliser de nouveau les calculs, lorsque les nouvelles valeurs d'amplitudes des harmoniques sphériques sont prises. La représentation revient à dessiner une série de cercles concentriques permettant d'offrir une échelle relative de l'espace, puis d'y ajouter deux polygones : un bleu associé aux valeurs de projections négatives, et un rouge associé aux valeurs de projections positives. Les figures

---

<sup>231</sup> Etant donné que cette spécificité n'est pas une caractéristique clairement définie de Pure Data, il serait préférable pour des questions de pérennité d'assurer ce mécanisme par l'objet *spam~* même, plutôt que de reposer l'approche sur l'ordre implicitement établi par Pure Data.

offrent l'impression d'une courbe, par le nombre important de points qui les définissent et dont les coordonnées cartésiennes sont obtenues à partir des angles discrétisant l'espace et les valeurs absolues des résultats de la projection dans ces directions [Figure 6.10].

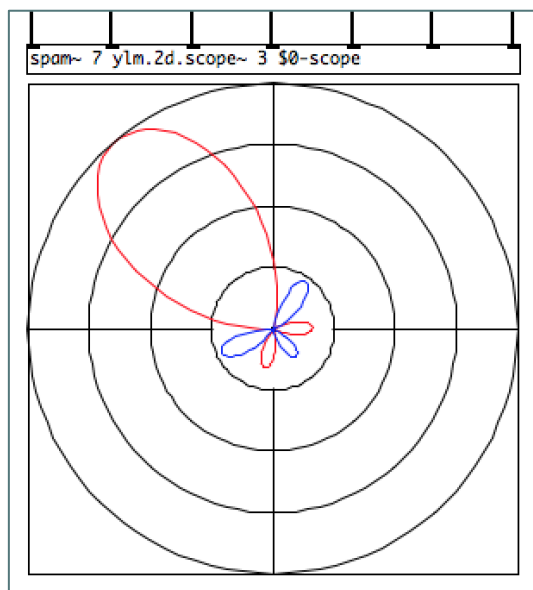


Figure 6.10. Abstraction *ylm.2d.scope~* chargée par l'objet *spam~* permettant de dessiner un champ sonore bidimensionnel dans le domaine des harmoniques sphériques, dans un patch graphique. Dans cet exemple, l'ordre de décomposition est 3 et le patch graphique porte le nom *\$0-scope*. Le premier argument est le nombre total d'instances qu'il est nécessaire de spécifier.

Malgré quelques astuces de mise en œuvre nécessitant une assez bonne connaissance des mécanismes internes du logiciel Pure Data, la réalisation suit une approche relativement lisible et compréhensible, sous condition évidente que le calcul des coefficients des harmoniques soit une opération assimilée par l'utilisateur<sup>232</sup>. D'un point de vue strictement utilisateur, l'abstraction réalisée permet une utilisation similaire à l'interface disponible par la bibliothèque HOA. En somme, via l'objet *spam~*, le système peut directement être raccordé aux signaux associés aux harmoniques sphériques, et l'abstraction peut s'adapter à différents ordres de décomposition. Il peut être remarqué que pour le moment, certains attributs tels que ceux permettant la gestion des couleurs ne sont pas intégrés à la mise en œuvre. Mais cela est envisageable selon les mêmes mécanismes déjà établis pour les attributs d'intervalle de temps entre chaque dessin et de valeur de gain, valeur permettant d'agrandir la figure dessinée. Ces valeurs étant récupérées selon un mécanisme identique à la récupération des angles des haut-parleurs dans la mise en œuvre de l'opération de projection sous la forme d'une abstraction.

---

<sup>232</sup> Evidemment sans cela, les mécanismes en jeu paraissent abstraits, mais il en est aussi de même pour l'ensemble des opérations réalisables en ambisonie, aussi cette condition peut être négligeable.

La réalisation d'interfaces graphiques pour représenter des champs sonores complexes sous la forme d'abstraction semble donc possible, à l'aide de la bibliothèque *spam*~. Il reste cependant à définir les limites d'un tel système vis à vis des performances, notamment s'il s'agit de réaliser des interfaces graphiques pour des représentations tridimensionnelles de l'espace. Néanmoins, force est de constater, comme cela a été évoqué précédemment, que même via des objets compilés, la demande en ressources du processeur est relativement importante. Aussi n'est-il pas raisonnable, quoiqu'il arrive, d'utiliser ce genre d'outils dans le cadre d'une représentation temps-réel. Enfin, un grand avantage de la mise en œuvre des interfaces graphiques via des abstractions est d'offrir à l'utilisateur de modifier les mécanismes et de changer le mode opératoire de la représentation. À ce sujet, il est possible d'envisager d'offrir une représentation linéaire plutôt que circulaire, ou de superposer les représentations. Étant donné que le patch graphique est indépendant du module de calcul, ce type d'approche est facilité.

### 6.3.3. Bilan

Le travail présenté dans cette partie tente de répondre à la nécessité de nouveaux outils logiciels moins spécifiques aux techniques de spatialisation du son ambisonique. Aussi, le premier objectif était de définir un ensemble de spécifications techniques et fonctionnelles, pouvant répondre à des enjeux plus génériques et amener à l'élaboration d'outils adaptés à de multiples représentations intermédiaires et abstraites de l'espace. Cela a été réalisé en confrontant les outils déjà disponibles et proposés par la bibliothèque HOA, aux pratiques et usages qui en ont été fait. Il a alors été défini que l'essence même de l'approche réside dans le système de vectorisation des opérations audionumériques et des procédés mis en œuvre. Partant de ce constat, la bibliothèque *spam* pour le logiciel Pure Data a été proposée. Cette bibliothèque est un ensemble d'objets dont l'approche est similaire à la suite d'outils logiciels gravitant autour de l'objet *hoa.2/3d.process*~. La particularité de la bibliothèque *spam* est, d'une part, qu'elle est bien moins spécifique et tente de réduire à l'essentiel l'approche de la bibliothèque HOA. D'autre part, elle vise principalement à satisfaire une approche expérimentale et l'intelligibilité des procédés mis en œuvre. Avant de tenter de proposer de nouvelles approches originales de spatialisations, il semblait important de valider ces outils en les utilisant pour des premières mises en œuvre « tests » d'un système de spatialisation. La confrontation avec des problématiques réelles de la mise en œuvre d'une opération audionumérique matricielle et de la mise en œuvre d'une interface de représentation graphique du champ sonore, a permis de valider les outils logiciels de la bibliothèque *spam*. Il est à ce titre aussi important de remarquer que ces deux présentations ne sont qu'une petite partie des mises en œuvre réalisées avec cette

bibliothèque. Ces phases de mise à l'épreuve des outils ont été aussi l'occasion de réaliser de nouvelles mises en œuvre de traitements déjà établis, et de généraliser des approches.

Aussi, dans la distribution de la bibliothèque *spam*, un ensemble d'abstractions destinées à la spatialisation ambisonique, est intégré. Ces abstractions portent toutes un nom commençant par *ylm*, en référence à la formulation mathématique et intègrent donc les deux abstractions présentées. Mais ces outils permettent, de plus, des opérations telles que la synthèse de source sonore directionnelle, la synthèse de champ sonore par décorrélation temporelle, par granulation ou modulation en anneau ou encore un *flanger* et un *phaser* ambisonique. Ces approches peuvent être intéressantes car elles généralisent certains procédés présentés dans la partie précédente. Aussi, pour chaque mise en œuvre, certains paramètres sont définis spécifiquement à chaque instance selon les caractéristiques des harmoniques auxquels elles sont associées. L'ensemble des approches permettent de passer linéairement d'une source sonore ponctuelle<sup>233</sup> à un champ sonore diffus ou chaotique. Ou encore, la méthodologie utilisée pour la mise en œuvre du *phaser* ambisonique est similaire à celle présentée dans ce travail pour la mise en œuvre du *flanger* ambisonique. Un autre ensemble plus réduit comportant seulement pour le moment deux abstractions, mais visant à être élargi, est plutôt destiné à une représentation de l'espace par un ensemble de points pouvant être assimilés à des haut-parleurs. Aussi, ces outils dont les noms commencent par *zpt* offrent un système d'analyse de champ sonore, via entre autre les vecteurs vitesse et énergie, et une interface de représentation du niveau de crête d'un système multicanal. Cette abstraction est l'équivalent de l'objet *hoa.2/3d.meter~*, mais les positions des haut-parleurs ne sont pas restreintes à la direction et peuvent être représentées sur l'ensemble de l'espace. Aussi, cela en fait un outil plus adapté pour des approches de type DBAP ou ViMiC.

Au-delà du simple aspect de contextualisation et de la mise à l'épreuve des outils de la bibliothèque *spam*, l'ensemble de ces réalisations sous la forme d'abstractions a permis de définir et de présenter une méthodologie, afin de mettre en œuvre des opérations matricielles et des interfaces graphiques associées à l'approche multicanal. Aussi est-il à présent envisageable de proposer de nouvelles représentations de l'espace dans le contexte de la spatialisation du son, en vue de proposer de nouveaux outils usuels et opérationnels.

---

<sup>233</sup> Les mises en œuvre prennent en compte la distance selon une approche similaire à celle proposée par l'objet *hoa.2/3d.map* mais utilisent l'algorithme d'ordre fractionnaire présenté dans la partie précédente.



## 7. Introduction à la représentation matricielle

Grâce aux outils logiciels plus génériques présentés dans la partie précédente et destinés à la mise en œuvre ainsi qu'à l'expérimentation de techniques de spatialisation du son, il est à présent possible de proposer de nouvelles approches de la spatialisation. Ces dernières sont fondées sur des représentations originales de l'espace. Aussi, comme cela a été abordé précédemment, pour des raisons évidentes de concision et de temps, les présentations préfèrent se limiter à une approche spécifique, afin d'offrir un exposé plus complet et plus intéressant que la simple énumération d'un ensemble de procédés seulement envisagés<sup>234</sup>. Cette représentation de l'espace qui est au centre de cette partie est construite sur une décomposition matricielle de l'espace. Ce choix n'est pas anodin, car ce type de représentation est déjà ancré dans de nombreux usages et laisse donc entrevoir de nombreuses possibilités. Aussi, ce chapitre est dédié dans un premier temps à la définition des origines et des principes sur lesquels repose cette approche. Notamment afin de mettre en avant les possibilités qui peuvent être offertes par une transposition dans le domaine musical. Par la suite, une première mise en œuvre de l'approche sera proposée afin de confronter les idées à un contexte musical concret. Ces mises en œuvre pourront alors être poursuivies dans le chapitre suivant.

### 7.1. Origines, principes fonctionnels et techniques

Avant de présenter les principes théoriques et fonctionnels de cette nouvelle approche, il est intéressant de montrer les origines de la décomposition matricielle de l'espace, et les concepts qui y sont rattachés. Notamment pour mettre en évidence certaines spécificités liées à ces concepts. Comme présenté précédemment, les approches plus classiques de la spatialisation, au sens physique du terme, s'évertuent à démontrer une cohérence sur un plan acoustique avec les modèles de propagation des ondes et/ou sur un plan psychoacoustique liés à notre perception des sons dans l'espace<sup>235</sup>. Quant à elles, les expérimentations réalisées dans le domaine des harmoniques sphériques ont permis de définir, ou plutôt d'ajouter, de nouvelles aspirations. Ces nouvelles exigences vis-à-vis du modèle de spatialisation amènent à combiner les approches acoustiques et psychoacoustiques avec des considérations porteuses de sens, sur un plan compositionnel. L'approche matricielle de l'espace est construite sur l'idée de *pixel* sonore et sur le Pixel Art. L'enjeu en revenant sur ces concepts est de mettre en exergue le potentiel créatif du modèle de spatialisation et de définir comment ces idées de bases amènent à

---

<sup>234</sup> Et cependant, bon nombre des possibilités de cette approche ne devront être qu'évoquées malgré tout.

<sup>235</sup> La première partie de cette thèse en est l'exemple.

concevoir de nouveaux traitements de l'espace et du son. Mais aussi de se demander comment les musiciens et compositeurs peuvent considérer et s'approprier ce modèle et la représentation de l'espace qui en découle. Cela soulève aussi d'ores et déjà, des questions pragmatiques liées à l'utilisabilité du système.

### 7.1.1. Origines et l'approche

Le terme de *pixel* provient de l'assemblage des mots anglais *picture* et *element* signifiant élément d'image. Les premières publications à ce sujet remontent à 1975 par F. C. Billingsley [Lyon, 2006], mais l'utilisation du *pixel* remonte bien avant, dans le domaine du traitement d'image et vidéo. Ce terme désigne l'élément fondamental utilisé dans les images numériques matricielles<sup>236</sup>, et correspond à la plus petite unité qu'un ordinateur puisse dessiner. Le *pixel* est généralement associé à un petit carré comportant des informations colorimétriques telles que des niveaux de gris ou des niveaux couleurs – généralement le rouge, le vert, le bleu et possiblement l'alpha selon les formats d'image. Ces couleurs sont liées à des informations de position, qui en image matricielle sont généralement définies sous la forme d'indices tels que des indices de ligne et de colonne. L'ensemble de ces petits carrés de couleurs positionnés les uns à côté des autres selon leurs informations spatiales, permet de représenter dans leur globalité une image cohérente<sup>237</sup> [Figure 7.1].



Figure 7.1. Image de l'icône du logiciel Pure Data avec un zoom sur une petite partie de l'image afin de mettre en évidence un ensemble de *pixels*.

### 7.1.2. Spécificité de l'approche

En passant du visuel au sonore, un peu d'imagination amène à concevoir un système qui, par analogie au *pixel*, est construit autour d'un élément sonore de base, appelé *soxel* : la contraction de *sound* et *element*. A ces *soxels* sont liées des informations de positionnement dans l'espace. Un ensemble de *soxels* peut alors être regroupé dans une matrice, afin de représenter un espace

---

<sup>236</sup> Il existe d'autres types d'images numériques telles que les images vectorielles.

<sup>237</sup> Non pas au sens figuratif mais dans l'idée de préserver la cohérence de l'image en elle, que les éléments de l'image restent bien positionnés les uns par rapport aux autres.

sonore. De ce concept, plusieurs éléments fondamentaux liés aux usages des *pixels* et à la pratique du Pixel Art, paraissent particulièrement intéressants pour une projection du son dans des éléments matriciels associés à l'espace, tant sur un plan concret et technique que sur un plan idéal relatif aux usages et à l'inspiration :

- L'image matricielle, lorsqu'elle représente une vue du dessus d'un espace, offre une représentation simple et familière où chaque *pixel* correspond à une zone aisément définissable. Cette approche de l'espace peut être comparée au quadrillage d'un plan où chaque case comporterait son lot d'informations : la case A1 contient les Arènes de Lutèce, la case A2 le campus de Jussieu, ..., la case C3 contient la Grande Mosquée de Paris, etc. Une telle représentation de l'espace peut être aisément comprise et assimilée par les musiciens et compositeurs notamment, comparé à la complexité de la représentation de l'espace sous la forme d'harmoniques circulaires ou sphériques.
- Il est implicitement admis que le nombre de *pixels* utilisés en Pixel Art est très limité et peut même être réduit à des matrices de 16x16 ou même 8x8. En partie pour l'enjeu et le défi de créer avec si peu d'éléments mais surtout pour son aspect nostalgique, car cette approche résulte une certaine esthétique liée aux images numériques des années 80 et 90. Bien que dans la pratique, il n'existe aucune règle stricte à ce sujet, cette caractéristique du Pixel Art est une des principales raisons de sa popularité. Il ne s'agit nullement de réaliser une analogie avec le domaine de la spatialisation du son, car seules des musiques échantillonnées en 8-bits pourraient *a priori* espérer provoquer une nostalgie similaire à une image de 8x8 *pixels*. Mais l'idée est, qu'une fois transposée dans un contexte sonore, il soit aussi possible de restreindre les approches à un petit nombre de *soxels*, tout en offrant suffisamment d'éléments pour que le résultat soit pertinent<sup>238</sup>. L'enjeu caché derrière cette affirmation réside dans la conjecture que les traitements sonores utilisés dans un espace matriciel peuvent très rapidement dépasser les limites de calculs des ordinateurs personnels courants. En effet, en considérant une petite matrice de 8x8 équivalent à 64 *soxels* et en présupposant qu'à chaque *soxel* est associé un traitement du son dont la complexité des calculs est analogue à ceux mis en œuvre dans le cadre des expérimentations en ambisonie, cela revient à mettre en œuvre un traitement dans le domaine des harmoniques sphériques à l'ordre de décomposition

---

<sup>238</sup> Il est, bien entendu, encore trop tôt pour définir si cette hypothèse est validée, mais les exemples concrets valideront, par la suite, cette proposition.

7. Ce qui, par expérience, est excessivement coûteux en terme de calculs et peut ne pas être supporté par des ordinateurs personnels.

- Avec l'arrivée des moteurs graphiques 3D, le *pixel* a été augmenté d'une troisième dimension spatiale, pour donner le *voxel* [Hughes & al., 2013] où la syllabe *pi* a été remplacée par la syllabe *vo* pour *volume*. Le *voxel* permet de stocker des informations colorimétriques associées à des informations spatiales tridimensionnelles qui donnent la possibilité de définir la notion de profondeur<sup>239</sup>. Cette extension du procédé a elle aussi amené à un certain nombre de pratiques et d'usages qui ont été très rapidement pris en main et acceptés. La construction d'espace peut être réalisée par blocs, de façon très simple et ressemblant fortement à l'utilisation de briques de Lego dont l'exemple de popularité le plus marquant est le jeu Minecraft, développé initialement de manière indépendante par Markus Persson<sup>240</sup>. Ainsi, il est fortement envisageable que l'extension des *soxels* à une troisième dimension spatiale soit réalisable. Surtout qu'elle implique, elle aussi, de nouvelles approches.
- Enfin une des propriétés les plus importantes de cette représentation de l'espace réside dans le nombre d'opérations matricielles disponibles et la simplicité de leur formulation sur un plan algébrique. En élaborant des approches fondées sur les calculs matriciels, il est aisément possible d'envisager toute une série d'opérations, au même titre que le sont la rotation, la distorsion de la perspective, ou encore la focalisation dans le domaine des harmoniques sphériques. Bon nombre d'entre elles sont relativement simples, et leurs propriétés dans le domaine des traitements de l'espace et du son peuvent être clairement présumées, telles que l'addition, la soustraction, la transposition de matrices ou encore la multiplication de matrices par un scalaire. D'autres peuvent quant à elles, être de véritables champs d'expérimentations telles que l'inversion de matrices, la réduction de matrices ou encore l'application de matrices de convolution utilisées en traitement de l'image pour générer des flous, ou encore effectuer de la détection de contours, etc.<sup>241</sup>

---

<sup>239</sup> Ou de hauteur si les *pixels* étaient associés à une vue du dessus.

<sup>240</sup> Le jeu est aujourd'hui distribué par plusieurs plateformes et compte plus de cent millions d'utilisateurs. Plus d'informations sont disponibles sur ce site consulté en juin 2017 [http://minecraft-fr.gamepedia.com/Minecraft\\_Wiki](http://minecraft-fr.gamepedia.com/Minecraft_Wiki).

<sup>241</sup> Certains de ces différents procédés sont étudiés plus en profondeur par la suite.

Mettre en œuvre un système des traitements de l'espace et du son fondé sur l'idée de *soxels* et à une représentation matricielle de l'espace, implique de définir le fonctionnement intrinsèque de ce modèle de décomposition de l'espace. A partir de ces principes et en utilisant les observations et les notions développées précédemment dans l'approche ambisonique, il est possible de définir une série de spécifications fonctionnelles nécessaires à l'orientation et la mise en œuvre des outils.

### 7.1.3. Les matrices mathématiques

Une matrice équivaut à un tableau à deux dimensions au sens informatique du terme. Elle est caractérisée par une taille définie en nombre de lignes  $L$  multipliée par un nombre de colonnes  $M$  qui est usuellement noté  $L \times M$  en mathématique, car il y a  $L$  éléments correspondant au nombre de lignes dans une colonne, et  $M$  éléments correspondant au nombre de colonnes dans une ligne résultant en un nombre total de  $L \times M$  éléments au sein de la matrice [Tableau 7.1]. En mathématiques, la taille de la matrice est aussi fréquemment appelée dimension, mais pour des raisons évidentes de confusion avec les dimensions spatiales, ce terme sera évité.

Matrice $L \times M$				
$E_l^m$	M colonnes $\rightarrow$			
L lignes ↓	$E_0^0$	$E_0^1$	$E_0^2$	$E_0^3$
	$E_1^0$	$E_1^1$	$E_1^2$	$E_1^3$
	$E_2^0$	$E_2^1$	$E_2^2$	$E_2^3$
	$E_3^0$	$E_3^1$	$E_3^2$	$E_3^3$

Tableau 7.1. Représentation de l'organisation des éléments d'une matrice de taille  $L \times M$ .

Dans un espace cartésien, afin de rester cohérent avec les représentations mathématiques des matrices, l'approche définit de manière arbitraire qu'une ligne est parallèle à l'axe des abscisses et qu'une colonne est parallèle à l'axe des ordonnées. Chaque élément d'une matrice est associé à deux indices  $l$  et  $m$ , qui définissent sa position au sein de celle-ci, et qui correspondent à son indice dans une ligne et son indice dans une colonne. En présupposant une extension du système à une troisième dimension spatiale, la matrice est remplacée par un tableau à trois dimensions.

La nouvelle dimension possède une taille notée  $N$  et est parallèle à l'axe de la côte – la hauteur – et l'élément d'un tableau possède alors un troisième indice  $n$  caractérisant la position dans cette dernière dimension.

#### 7.1.4. L'espace matriciel

L'enjeu est à présent d'associer à un espace, les outils mathématiques que sont les matrices. Un certain nombre de particularités et de spécificités découlent implicitement de cette mise en relation. Contrairement à l'approche ambisonique où l'espace est circulaire ou sphérique, l'approche matricielle de l'espace implique nécessairement que celui-ci soit rectangulaire<sup>242</sup>. L'espace est alors décomposé par une série de lignes et de colonnes et leurs intersections définissent les *soxels* qui coïncident avec les éléments d'une matrice. Ainsi, un *soxel* est la composante fondamentale de l'espace, tel que l'harmonique sphérique en ambisonie. Et les *soxels* correspondent à des rectangles de dimensions similaires. Un ensemble de *soxels* sous une forme matricielle représente l'espace sonore sur lequel il est possible d'opérer. Ainsi l'espace sonore opérable est limité par sa taille, à savoir  $L \times M$ , définissant le nombre de lignes et le nombre de colonnes utilisées pour le décomposer. Les *soxels* sont définis par des coordonnées sous la forme d'indices matriciels tel que  $S_{l,m}$  avec  $l \in [0, L - 1]$  et  $m \in [0, M - 1]$ . Logiquement, l'ordre de décomposition dans l'approche matricielle de l'espace est défini par le nombre de lignes et de colonnes utilisées. Et de manière analogue à l'approche ambisonique où plus l'ordre de décomposition est important plus il y a d'harmoniques sphériques pour représenter l'espace, plus il y a de lignes et de colonnes plus il y a de *soxels* pour représenter l'espace.

Cet espace matriciel est, de manière analogue à l'appropriation musicale des harmoniques sphériques en ambisonie, abstraite et n'est donc aucunement associée à un système de haut-parleurs ou plus généralement à aucun système de restitution. Nécessairement, le *soxel* et la portion de l'espace à laquelle il est associé, sont de même abstraits. Le *soxel* n'est régi que par sa position par rapport aux autres *soxels* et son organisation au sein de l'espace matriciel.

---

<sup>242</sup> L'espace est un rectangle lorsque l'approche est bidimensionnelle. Il devient logiquement un parallélépipède rectangle lorsque l'approche est tridimensionnelle.

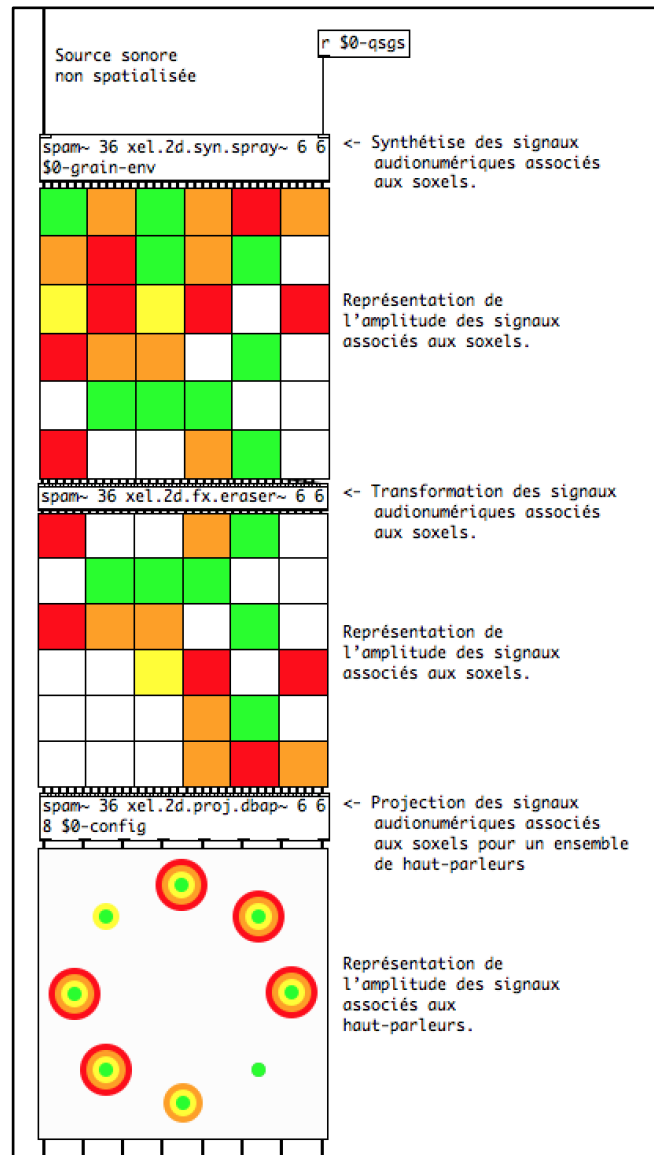


Figure 7.2. Représentation de la chaîne de traitement au sein de la bibliothèque xel

### 7.1.5. Traitements audionumériques par *soxel*

Tout comme en ambisonie où à un harmonique sphérique est associé un signal audionumérique, à chaque *soxel* est associé un signal audionumérique, et l'ensemble de ces signaux audionumériques permettent de représenter l'espace sonore sous une forme matricielle. En définissant des opérations sur ces signaux, il est alors possible de synthétiser, transformer et projeter l'espace sonore. Comme pour l'approche ambisonique, la chaîne de traitement audionumérique dans l'approche matricielle comporte donc deux étapes principales, la synthèse et la projection :

- La synthèse consiste à synthétiser les signaux audionumériques associés aux *soxels* en considérant, ou non<sup>243</sup>, leurs indices et leurs positions relatives au sein de l'espace.
- La projection consiste à reproduire au mieux l'espace sonore pour un système de restitution sonore, en utilisant l'ensemble ou une partie des signaux associés aux *soxels* en fonction de leurs propriétés spatiales et du système utilisé.

Ces deux étapes dans la chaîne de traitement audionumérique offrent possiblement trois domaines d'interaction. Le premier domaine se situe avant la synthèse où il est possible d'opérer avec une ou des sources sonores non spatialisées correspondant à un ou des signaux audionumériques sans informations spatiales. Le troisième domaine se situe après la projection où il est possible d'opérer avec un ou des signaux audionumériques associés à un système de restitution. Et au centre, se trouve le deuxième domaine où il est possible d'opérer des transformations de l'espace sonore sous sa forme matricielle, en utilisant les signaux associés aux *soxels* et notamment leurs informations spatiales [Figure 7.2].

### 7.1.6. Fonctionnement et aspects génériques des outils

La bibliothèque est composée d'un ensemble d'abstractions pour le logiciel Pure Data, dont le fonctionnement repose sur la bibliothèque d'objets externes *spam* présentée dans le chapitre précédent.

La bibliothèque suit la même syntaxe définie pour la mise en œuvre des abstractions de traitements spatiaux de la bibliothèque HOA. Elle offre des avantages quant à l'organisation et la flexibilité au vu de futurs développements – notamment pour une représentation tridimensionnelle de l'espace – et augmente les morphèmes utilisés pour prendre en compte les opérations qui sont à présent sous la forme d'abstraction. Toutes les abstractions commencent par le terme *xel* faisant référence à la bibliothèque, puis un terme définit le nombre de dimensions spatiales à savoir  $2d$  (mais potentiellement  $3d$  dans le futur<sup>244</sup>), suivi d'un terme

---

<sup>243</sup> Comme pour l'approche ambisonique, il est bien sûr plus intéressant sur un plan formel de prendre en considérations les propriétés spatiales de chaque composante fondamentale issue de la décomposition de l'espace. Mais rien n'empêche de réaliser des mises en œuvre et des choix empiriques libérés de ces contraintes spatiales.

<sup>244</sup> Il serait aussi possible d'envisager un système à une dimension. Bien que l'intérêt soit relativement limité, une telle approche pourrait convenir à certaines configurations de haut-parleurs tels que ceux utilisés en WFS, par exemple.



définissant la fonction générale – *syn* pour synthèse, *fx* pour traitement dans le domaine des *soxels*, *proj* pour l’opération de projection et *gui* pour les interfaces graphiques utilisateur.

La création des abstractions au sein de Pure Data avec la bibliothèque d’objets externes *spam* nécessite de mettre en œuvre certaines pratiques et opérations génériques qui seront utilisées de manière récurrente au sein de la plupart des abstractions. Il a été défini de manière arbitraire que les deux premiers arguments supplémentaires donnés à l’objet – c’est à dire après le nom du patch à charger donc les troisième et quatrième arguments – définissent la taille de l’objet avec les valeurs  $L$  le nombre de ligne et  $M$  le nombre de colonnes. L’objet *spam~* nécessite le nombre total d’instances défini alors par  $L \times M$  [Figure 3.2].

```
spam~ 12 xel.2d.syn.spray~ 3 4 $0-grain-env
```

Figure 7.3. Représentation de la gestion des arguments de la bibliothèque *xel* avec l’objet *spam~*. Le premier argument 12 est le nombre total d’instances, le deuxième argument *xel.2d.syn.spray~* est le nom de l’abstraction chargée, le troisième argument 3 est le nombre de ligne  $L$ , le quatrième argument 4 est le nombre de colonnes  $M$  et le dernier argument *\$0-grain-env* est spécifique à l’abstraction.

Afin d’assurer le bon fonctionnement des différentes abstractions entre elles ainsi qu’une certaine normalisation de l’approche, de manière analogue à la numérotation des canaux en ambisonie via l’ACN, il est nécessaire de définir une méthode afin d’organiser les instances associées aux *soxels* au sein de l’objet *spam~*. Cela notamment afin de faire concorder entre elles les entrées et sorties des différentes abstractions. Le principe est donc – de manière analogue à l’ambisonie où à partir de l’ACN il est possible de définir le degré  $l$  et l’ordre azimutal  $m$  des harmoniques – de pouvoir définir le numéro de la ligne  $l$  et le numéro de la colonne  $m$  auxquelles appartient un *soxel* en fonction de l’indice  $i$  de l’instance qui lui est associé et réciproquement. La logique simple amène à définir que  $l = \frac{i}{M}$  et  $m = i - l$ . Ainsi, les premières instances correspondent à la première ligne définie par  $l = 0$ , en commençant par  $m = 0$  jusqu’à  $m = M - 1$  et à toutes les nouvelles lignes donc toutes les  $M$  instances, l’indice  $l$  est incrémentée jusqu’à atteindre  $l = L - 1$ . La première instance est associée au *soxel* d’indices  $l = 0$  et  $m = 0$  et le dernier *soxel* est associé au *soxel* d’indices  $l = L - 1$  et  $m = M - 1$  [Tableau 7.2]. Réciproquement, à partir des indices  $l$  et  $m$ , il est possible de retrouver l’indice de l’instance tel que  $i = l \times M + m$ .

$0_0^0$	$1_0^1$	$2_0^2$	$3_0^3$
$4_1^0$	$5_1^1$	$6_1^2$	$7_1^3$
$8_2^0$	$9_2^1$	$10_2^2$	$11_2^3$

Tableau 7.2. Représentation de l'organisation des instances associées aux *soxels* avec les indices des instances  $i$ , les indices des lignes  $l$  et les indices des colonnes  $m$  tel que  $i_l^m$ .

A la suite de la présentation des origines et des sources d'inspiration de l'approche, les principes fonctionnels d'ordres généraux de la bibliothèque ont été exposés. Il peut alors être envisagé de confronter cette approche, qui n'est pour le moment que théorique, à une réelle mise en œuvre. L'enjeu est de valider la pertinence sur un plan artistique et conceptuel, de la décomposition matricielle de l'espace pour la spatialisation du son. Pour cela, la présentation propose de rester dans un cadre encore relativement abstrait, via la création d'une approche qui peut être qualifiée de Soxel Art. Suite à cela, il sera possible d'envisager d'étendre l'approche et de proposer des opérations répondant à des questions d'ordres plus classiques, telles que l'encodage de sources directionnelles. Et répondant aussi à des questions plus pratiques, telles que la mise en œuvre d'une technique de projection de l'espace sonore sur un système de haut-parleurs.

## 7.2. Soxel Art

Dans cette sous partie, est présentée comment le Pixel Art, qui est une approche particulière du dessin numérique, peut être envisagée dans le domaine de l'espace sonore. La question est aussi de savoir comment cette approche peut créer de nouvelles pratiques de composition de l'espace et du son, pouvant être qualifiées de Soxel Art<sup>245</sup>. La première réflexion part des caractéristiques de cette pratique picturale, pour définir les concepts clés du Soxel Art, sans pour autant faire des analogies directes ou rentrer dans des considérations liées à la synesthésie et au-delà de l'aspect matriciel. Pour cela, le Pixel Art est mis en relation avec des approches compositionnelles dans le domaine de l'audio numérique générale, mais aussi spécifiques à l'espace sonore afin de faire émerger une première approche de la composition d'espace sonore. Ce premier cheminement amène à concevoir l'outil central de cette pratique, permettant de synthétiser les signaux associés aux *soxels*. A la suite de la mise en œuvre de ce premier module de synthèse, émergent de nouvelles exigences visant à améliorer et à faciliter la prise en main et l'appropriation de ces procédés. Pour tenter de répondre à ces besoins, de nouveaux outils

<sup>245</sup> Suite au *soxel*, ce terme est défini majoritairement pour son aspect ludique.

sont élaborés, d'après l'étude des logiciels disponibles pour la création d'images en Pixel Art. Enfin, une première critique et analyse de l'approche est réalisée.

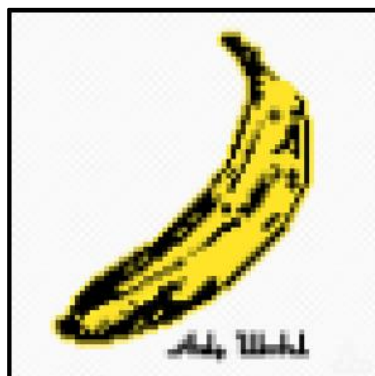


Figure 7.4. Interprétation en Pixel Art par Jaebum Joo de la pochette de l'album Velvet Underground & Nico initialement réalisée par Andy Warhol<sup>246</sup>.

### 7.2.1. Principes

Du *pixel* et de son utilisation, découle le Pixel Art, une pratique picturale consistant à générer des images, souvent de basse résolution, en dessinant *pixel* par *pixel* [Figure 7.4]. Le terme de Pixel Art a pour la première fois été publié en 1982 par A. Golberg et R. Flegal [Golberg & al., 1982] mais son concept remonte lui aussi bien avant avec par exemple le logiciel SuperPaint créé par R. Shoup en 1972, permettant de produire des images en Pixel Art pour la télévision [Shoup, 2001]. Malgré les énormes progrès qu'il y a eu depuis en termes d'image numérique [Hughes & al., 2013], [Gonzales & al., 2007], le Pixel Art est très populaire aujourd'hui par sa simplicité et son ancrage nostalgique dans une esthétique des années 1980 et 1990 comme le démontre la forte popularité des jeux dont l'univers graphique est fondé sur cette approche [Grahm, 2013].

Cette approche peut s'apparenter, en utilisant les mots de Horacio Vaggione à propos de la pratique de la décorrélation micro-temporelle [Vaggione, 2002], à une pratique manuelle du compositeur, en opposition à une pratique algorithmique<sup>247</sup>. Il explique à ce sujet que « la décorrélation peut être aussi réalisée graphiquement, en coupant ou dessinant des formes d'onde avec l'aide d'un éditeur de sons ayant des bonnes facilités de zooming ». Et que cette « [...]

---

<sup>246</sup> L'image est en sous licence Creative Common pour un usage non-commercial (CC BY-NC-ND-2.0) et disponible en juin 2017 à l'adresse <https://www.flickr.com/photos/joojaebum/>.

<sup>247</sup> L'approche algorithmique est plutôt à associer aux systèmes qui seront présentés par la suite.

possibilité n'est ni si laborieuse ni si naïve qu'elle le paraît : c'est une attitude de compositeur que d'écrire de la musique « manuellement », note-à-note, ou partiel-par-partiel, ou grain-par-grain, et il n'a pas à rugir [sic] d'utiliser cette possibilité de son artisanat ». Une première possibilité de création d'espace sonore<sup>248</sup> via un système de représentation matricielle de l'espace, peut s'inspirer de cette approche manuelle de la composition et des pratiques du Pixel Art, en envisageant une réalisation minutieuse où chaque *soxel* de l'espace puisse être créé, édité et modifié individuellement. Il est nécessaire de rappeler à ce sujet que c'est aussi ce type de procédé mis en œuvre par G. Wakefield – où les signaux associés aux harmoniques spatiaux sont synthétisés individuellement [Wakefield, 2006] – qui a amené aux premières expérimentations ambisoniques au sein de la bibliothèque HOA, pour par la suite développer des approches algorithmiques.

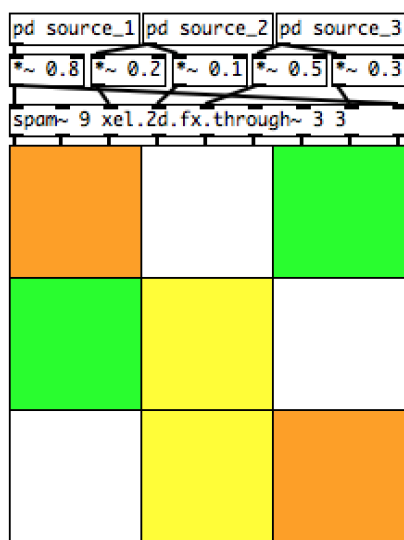


Figure 7.5. Capture d'écran d'un patch Pure Data où l'espace sonore matriciel de dimension  $3 \times 3$  est synthétisé par trois sources sonores dont les signaux sont envoyés directement aux *soxels* après des pondérations appliquées manuellement. L'interface graphique générée par l'abstraction *xel.2d.gui.scope~* représente par des couleurs l'amplitude *RMS* des signaux associés aux *soxels*.

## 7.2.2. Approche manuelle de la synthèse sonore

La première approche envisagée est des plus rudimentaires. Elle vise à synthétiser directement les signaux associés aux *soxels*, de la même manière que l'aurait offert l'utilisation d'un outil permettant exclusivement d'appliquer une couleur à un seul *pixel* à la fois. C'est d'ailleurs l'usage et la fonctionnalité de base au sein du logiciel Aseprite. Afin de faciliter et de clarifier

<sup>248</sup> La création d'espace sonore recouvre ici, comme suggéré précédemment, la synthèse des signaux sans foncièrement considérer l'opération de projection et le rendu sur un système de restitution. L'idée est de rester dans un domaine abstrait.

un tant soit peu la mise en œuvre de cette approche, une simple abstraction *xel.2d.fx.through~* permettant seulement de laisser passer les signaux audio numériques est utilisée dans l'objet *spam~*, pour offrir un objet de raccordement initial [Figure 7.5]. Cette approche, bien évidemment potentiellement pertinente sur un plan musical, est néanmoins excessivement laborieuse à mettre en œuvre, et les possibilités d'interaction et de contrôle sont extrêmement limitées. Dans l'exemple présenté, les connexions entre les sources sonores et les *soxels* sont statiques, tout comme les gains qui leurs sont associés. Et il est impossible de les modifier en temps-réel, au risque de générer des artefacts. Il est néanmoins possible de modifier l'approche afin d'offrir ces possibilités. Pour cela chaque source doit être raccordée à autant de modules de contrôle de l'amplitude adaptés à un usage temps-réel, qu'il y a de *soxels*. Puis les signaux associés à un *soxel* doivent être additionnés. Néanmoins, cette approche est incommode, redondante et surtout très peu dynamique. Elle nécessite donc d'être redéfinie à chaque nouvel usage. Le premier outil qui s'impose suite à cette approche doit pouvoir offrir une gestion facile et flexible de la synthèse des signaux associés aux *soxels*, à partir des sources sonores et des séries de gains associés.

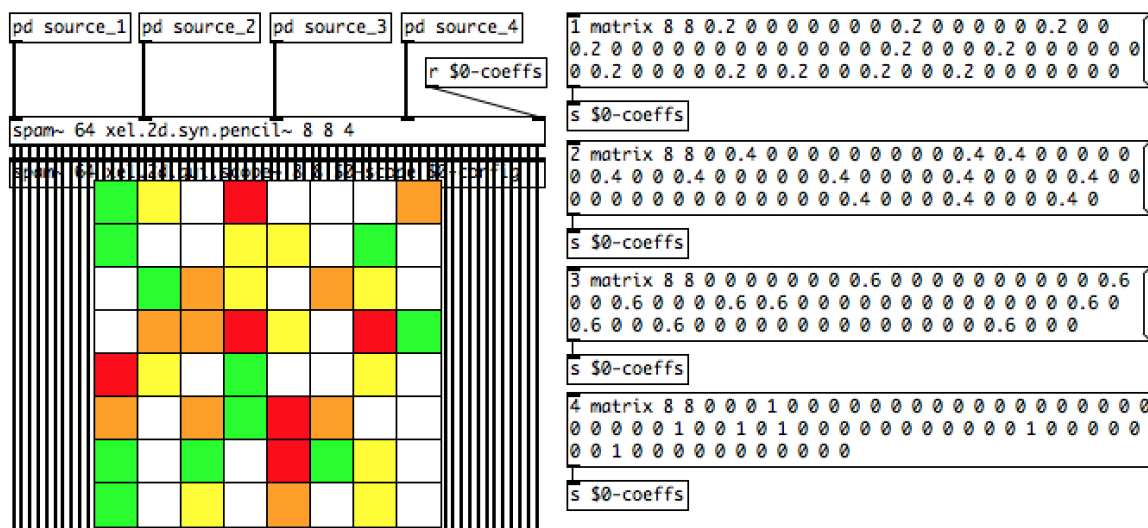


Figure 7.6. Capture d'écran d'un patch Pure Data où l'espace sonore matriciel de dimension  $8 \times 8$  est synthétisé par quatre sources sonores dont les signaux sont envoyés aux *soxels* via l'abstraction *xel.2d.syn.pencil~* suivant des matrices de coefficients définies par les quatre messages à droites. Les *soxels* générés par la source 1 sont associés à une amplitude de 0.2 et sont en vert. Les *soxels* générés par la source 2 sont associés à une amplitude de 0.4 et sont en jaune, les *soxels* générés par la source 3 sont associés à une amplitude de 0.6 et sont en orange. Et les *soxels* générés par la source 4 sont associée à une amplitude de 1 et sont en rouge.

### 7.2.3. Approche dynamique de la synthèse sonore

Afin de répondre à ces spécifications, un premier module de synthèse des signaux audio numériques a été élaboré sous la forme d'une abstraction pour l'objet *spam~*.

L'abstraction *xel.2d.syn.pencil~* est un module de synthèse d'espace sonore matriciel, qui reste néanmoins à bas-niveau et en adéquation avec une approche « manuelle », dans le sens où il est nécessaire de définir de façon explicite chaque coefficient d'amplitude associé aux *soxels*. L'avantage de cet outil est double : d'une part il accepte les matrices de coefficients au format *iemmatrix*<sup>249</sup> et d'autre part, il permet de synthétiser en une combinaison plusieurs couches spatiales sonores, au sein d'une seule opération. Ces couches spatiales sont respectivement définies par une source sonore et une série de coefficients matriciels [Figure 7.6].

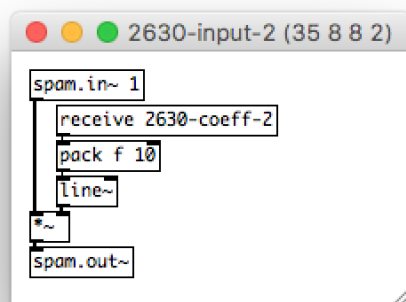


Figure 7.7. Représentation du sous patch permettant de générer le signal de la deuxième couche spatiale sonore du *soxel* d'indice linéaire 35. Le signal source de la deuxième couche est reçu par l'objet *spam.in~*, le coefficient de gain spécifique à la couche et à l'instance du *soxel* est reçu via l'objet *receive* et le signal pondéré est envoyé via l'objet *spam.out~* à la sortie spécifique à l'instance du *soxel*.

L'abstraction est utilisée au sein de l'objet *spam~* et est donc instanciée autant de fois qu'il y a *soxels* dans l'espace sonore. Les entrées des signaux sont communes à l'ensemble des instances, et permettent de récupérer les sources sonores relatives à chaque couche spatiale, et sont générées dynamiquement via un script selon le nombre défini par le troisième argument<sup>250</sup>. À chaque entrée signal est associé un facteur de gain qui est relatif à l'instance du *soxel* et à une couche spatiale sonore spécifique tel que représenté dans la [Figure 7.7]. Une fois l'abstraction instanciée dans l'objet *spam~*, cette dernière offre autant d'entrées qu'il existe de couches spatiales sonores, plus une entrée de contrôle permettant de définir les matrices de coefficients.

<sup>249</sup> L'usage du format de message représentant des matrices de coefficients établies par la bibliothèque *iemmatrix* est expliqué et décrit par la suite.

<sup>250</sup> Les deux premiers arguments définissent la taille matricielle pour des raisons de cohérence entre les différentes abstractions, bien que ces arguments ne soient pas utilisés.

Chaque liste définissant les coefficients matriciels d'une couche sonore doit être envoyée à l'objet selon l'indice relatif à cette couche<sup>251</sup>, tel que représenté dans la [Figure 7.6].

Enfin, un dernier aspect notable de cette mise en œuvre réside en l'utilisation de l'objet *line~* permettant de transformer les coefficients de contrôle en signaux audionumériques selon un temps d'interpolation de dix millisecondes, tel que présenté dans la [Figure 7.7]. Une fonctionnalité intéressante pourrait être de pouvoir régler ce temps, afin de passer de manière plus lente d'un état de l'espace sonore à un autre<sup>252</sup>.

Une lacune majeure apparaît rapidement dans la pratique, au niveau de l'utilisabilité pour générer et modifier les matrices de coefficients. Celles-ci étant sous la forme de listes de valeurs numériques, elles sont quelque peu abstraites vis à vis de la représentation de l'espace, et surtout difficilement manipulable. Il est donc nécessaire de fournir des outils permettant d'en faciliter à la fois l'édition et la manipulation, mais aussi d'offrir un système permettant de les manier dans le temps et une interface permettant de les représenter dans l'espace.

#### 7.2.4. Manipulation de matrices

Le Pixel Art transposé dans le domaine de la composition de l'espace sonore offre un fort potentiel inspirateur. Comme évoqué précédemment, l'un des atouts majeurs de ce style est son enracinement dans cette esthétique principalement associée aux arts vidéoludiques et un peu datée. Mais au-delà de cet aspect, le fait de rencontrer une résonance avec des outils et des pratiques connues – et même familières – amène à envisager de nouvelles solutions permettant de faciliter la prise en main de la composition d'espace sonore matriciel.

---

<sup>251</sup> Pour des raisons d'utilisabilité, les coefficients de la première couche peuvent être définis directement selon la syntaxe matricielle sans indice de couche.

<sup>252</sup> Bien entendu, l'idée de contrôler ces gains directement par des signaux audionumériques revient à réaliser l'approche exclusivement manuelle et quelque peu laborieuse, présentée précédemment. Cela rend obsolète l'usage d'une abstraction.

Les outils mis en œuvre s’inspirent des fonctionnalités primaires mais simples, disponibles dans les premières versions du logiciel Microsoft Paint<sup>253</sup> ou plus récemment du logiciel Aseprite<sup>254</sup> dédié à la création de Pixel Art [Figure 7.8].

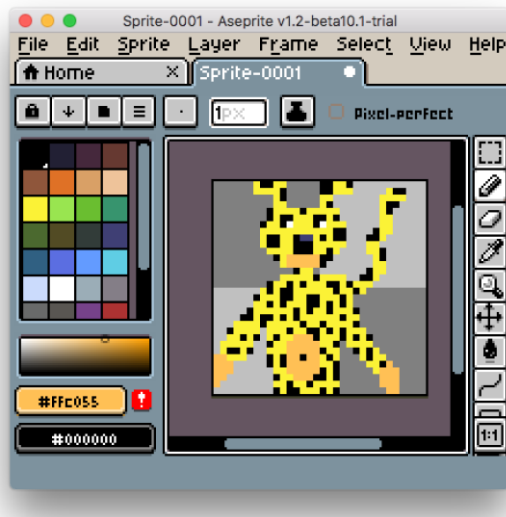


Figure 7.8. Interface du logiciel Aseprite dans lequel le Marsupilami est dessiné en Pixel Art par un admirateur de Franquin. L’interface offre des fonctionnalités primaires mais essentielles de dessin.

La toute première nécessité est de pouvoir générer des matrices uniformes composées de coefficients de valeurs identiques, afin de pouvoir modifier de manière générale l’ensemble des contributions de l’espace sonore et surtout de les annuler en définissant comme valeur zéro. Puis, très rapidement, le besoin s’est imposé de définir la valeur d’un coefficient spécifique. À cela s’ajoute certaines fonctionnalités simples, comme la définition des coefficients de l’ensemble d’une ligne ou de l’ensemble d’une colonne avec une valeur. L’ensemble de ces approches est réalisable via la bibliothèque d’objet pour Pure Data *iemmatrix*<sup>255</sup> développée par I. m zmoelnig, T. Musil et F. Zotter à l’IEM<sup>256</sup> et qui a notamment été utilisée pour réaliser une mise en œuvre de l’ambisonie et du VBAP [Kaiser, 2001].

<sup>253</sup> Le logiciel fut publié la première fois en 1985 sous le nom de PaintBrush et est toujours développé et disponible en ligne en juin 2017 à l’adresse <https://support.microsoft.com/>.

<sup>254</sup> Le logiciel est libre et développé par David Capello depuis 2007 et est disponible en ligne en juin 2017 à l’adresse <https://www.aseprite.org/>. Les sources sont disponibles en ligne en juin 2017 à l’adresse <https://github.com/aseprite/aseprite>.

<sup>255</sup> La bibliothèque est disponible via le système de gestion d’objets externes Deken, et le code source est disponible en ligne en juin 2017 à l’adresse <https://git.iem.at/pd/iemmatrix>.

<sup>256</sup> *Institut für elektronische musik und akustik* à Graz en Autriche.



Cette bibliothèque possède un avantage certain, car elle offre de nombreuses possibilités pour générer et opérer sur des matrices. Sans rentrer pour autant dans une présentation exhaustive, en plus des fonctionnalités énoncées plus haut, elle offre la possibilité de générer des matrices identités et de réaliser un certain nombre d'opérations arithmétiques terme à terme, comme l'addition ou la multiplication<sup>257</sup>. Afin d'éviter tout changement de format des messages pour passer d'une bibliothèque à une autre, les matrices de coefficients utilisées dans la bibliothèque *xel* empruntent leur syntaxe à la bibliothèque *iemmatrix* et sont constituées du symbole *matrix* suivi de la dimension de la matrice – le nombre de lignes et le nombre de colonnes – puis d'une liste facultative de coefficients qui doit respecter la taille matricielle donnée comme présentée en exemple sur la [Figure 7.9].

```
matrix 4 4 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1
```

Figure 7.9. Exemple d'un message représentant une matrice d'identité de dimension  $4 \times 4$  dont les valeurs ont été générées via l'objet *mtx* de la bibliothèque *iemmatrix*.

### 7.2.5. Visualisation et édition graphique

Une des premières mises en œuvre permet de visualiser les matrices. L'abstraction *xel.2d.gui.mat* représente les coefficients des matrices via un patch graphique, grâce au système de données et de structures graphiques de Pure Data [Puckette, 2002]. L'abstraction nécessite en argument le nom du patch graphique dans lequel elle doit représenter la matrice, suivi de deux premiers arguments facultatifs définissant la taille matricielle par défaut et les deux derniers arguments facultatifs définissant le minimum et le maximum des coefficients de la matrice.

L'abstraction peut recevoir le message *matrix*, accompagné d'une taille matricielle, du nombre de ligne et du nombre de colonnes, afin de redimensionner la matrice à afficher. Si ce message est suivi par une série de valeurs numériques, alors il est interprété comme les coefficients matriciels à afficher [Figure 7.10]. Sans coefficients, l'abstraction définit les valeurs à zéro, car elle offre aussi à l'utilisateur la possibilité de les modifier par la suite. Cette modification se fait

---

<sup>257</sup> Afin d'éviter toutes dépendances trop restrictives et contraignantes, l'utilisation de ces outils est limitée aux cas d'exemples, et n'apparaît jamais au sein d'une des abstractions de la bibliothèque *xel* néanmoins son usage est grandement facilité. La bibliothèque étant sous la forme d'une série d'abstractions, la compatibilité avec tous les systèmes supportant Pure Data est pratiquement assurée alors que la bibliothèque *iemmatrix* est majoritairement constituée d'objets compilés potentiellement moins pérennes.

soit en les définissant à l'aide du clavier numérique, soit en cliquant et glissant. C'est d'ailleurs principalement pour cet usage que les valeurs minimum et maximum sont utiles, afin d'éviter de générer de trop grandes valeurs<sup>258</sup>. Bien qu'il soit possible d'afficher des valeurs décimales via des messages, il est impossible dans Pure Data d'en générer via l'interface graphique. Le choix a été fait de multiplier par cent puis de tronquer ces valeurs à l'affichage : ainsi une valeur de 0.59 est affichée 59. Lorsque l'abstraction reçoit une matrice ou que celle-ci est modifiée via sa représentation graphique, elle envoie sous forme de message cette matrice, permettant ainsi de contrôler le module de synthèse. La sortie automatique des valeurs peut être évitée, en ajoutant le message *set* aux matrices entrantes, et les matrices sortantes peuvent ainsi être renvoyées à l'abstraction et affichées sans créer de boucles infinies. Cette fonctionnalité est utile afin d'appliquer des modifications aux coefficients, tel qu'un modulo sur les valeurs par exemple. Cette fonctionnalité permet aussi de prendre en compte ce changement au sein de l'interface graphique, sans créer de boucle infinie.

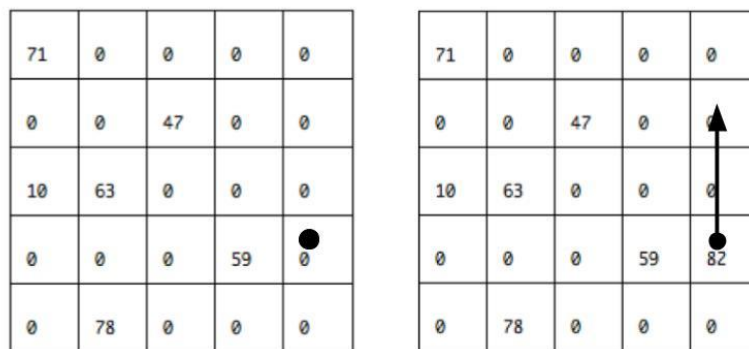


Figure 7.10. Représentation du rendu d'une matrice de taille  $5 \times 5$  dans un patch graphique avec l'abstraction *xel.2d.gui.mat*. Le point et la flèche représentent l'interaction permettant de modifier la valeur d'un coefficient associé à une cellule. La valeur est de 0 au clic à gauche et en glissant le curseur vers le haut, tel que représenté à droite, la valeur est incrémentée.

Cette approche qui permet de définir précisément les coefficients d'une matrice, au centième près, est relativement pratique dans un contexte de temps différés. Mais l'opération peut vite s'avérer redondante et laborieuse, notamment dans un contexte temps-réel. Afin de répondre à ce problème, il a été envisagé d'offrir un mode de jeu où par simple clic sur une cellule associée à un *soxel*, il serait possible de modifier les coefficients de la matrice<sup>259</sup>. Cette approche

<sup>258</sup> L'étendue de valeurs jouable est aussi modifiable après la création, avec le message *range* suivi d'une valeur minimum et d'une valeur maximum.

<sup>259</sup> La possibilité de pouvoir cliquer et glisser sur le patch graphique de cellule en cellule a été considérée, afin de modifier leurs valeurs – de manière similaire aux logiciels de dessin – mais la mise en œuvre actuelle de l'interface graphique, par l'usage de sous-structures de données graphique, rend impossible cette fonctionnalité.

consisterait à définir une valeur de coefficient et tous les éléments de la matrice sur lesquels se pose le curseur, prendraient cette valeur. Cette fonctionnalité est avantageuse dès lors que l'utilisateur souhaite définir rapidement à un ensemble de cellule une même valeur, mais ce cas d'usage est relativement limité et un problème apparaît rapidement dès qu'il s'agit de générer une matrice disparate aux valeurs hétérogènes. En effet, cela implique beaucoup d'actions de la part de l'utilisateur. Le principe d'interagir par simple clic de la souris semble pertinent, aux vues des usages des logiciels de dessin, mais aussi des interfaces de mise en espace du son telles que proposées dans la bibliothèque HOA. Cependant, il semble nécessaire de concevoir une approche plus ouverte et plus flexible que la simple affectation d'une valeur à une cellule. Pour cela, l'abstraction possède une deuxième sortie, qui envoie les indices de ligne et de colonne d'une cellule lorsqu'elle reçoit un clic de la souris. Cela permet de mettre en œuvre selon ces informations, des opérations plus adaptées à l'approche compositionnel de l'espace sonore.

### 7.2.6. Édition par motifs

Un aspect intéressant des mises en œuvre des logiciels de dessin tel que GIMP<sup>260</sup>, réside en le faible nombre d'outils destinés au dessin à proprement parler. En effet, seulement trois outils sont disponibles : le crayon, le pinceau et l'aérographe<sup>261</sup>. Le crayon offre un tracé net et pixelisé, le pinceau offre un tracé plus flou<sup>262</sup> et l'aérographe permet de contrôler le rendu en fonction de la vitesse du mouvement. Ces trois outils sont, en réalité, des variantes d'une même approche consistant à remplir un certain nombre de *pixels* d'une image suivant un motif, plus communément appelé « brosse ». Cette brosse peut être commune à l'ensemble des outils, car le crayon n'implique pas forcément que les *pixels* appliqués soient collés selon une figure rectangulaire. En effet, ils peuvent être espacés selon une figure ressemblant d'ailleurs plus à l'image faite d'un aérographe. L'approche de base pourrait être le crayon et le pinceau en serait une version filtrée, l'aérographe quant à lui serait une version du pinceau où les qualités du filtre dépendraient de la manière de dessiner. L'idée importante est qu'en dehors du traitement de rendu ou du mode de contrôle, ce sont des outils similaires qui font appliquer une brosse à un motif de *pixels*. En ce qui concerne l'action de dessiner à proprement parler dans les logiciels, par sa flexibilité et le nombre illimité de variantes, l'utilisation des brosses semble répondre à la majeure partie des usages. Cette fonctionnalité a été notamment reprise dans le logiciel Aseprite, qui permet de sélectionner un ensemble de *pixels* pour créer rapidement une nouvelle

---

<sup>260</sup> Logiciel de dessin et d'édition d'image libre, gratuit et très complet disponible sur internet à l'adresse <https://www.gimp.org/fr/> (consulté en juin 2017).

<sup>261</sup> Il existe aussi l'outil de remplissage et de dégradé mais l'idée ici est de penser dessin dans le sens du tracer.

<sup>262</sup> L'effet de flou sera présenté dans les chapitres suivant, qui traitent des filtres.

brosse<sup>263</sup>. Cette approche par motifs, transposée dans le logiciel Pure Data, est donc intéressante car elle permet d'envisager un outil générique recouvrant un nombre très important d'opérations.

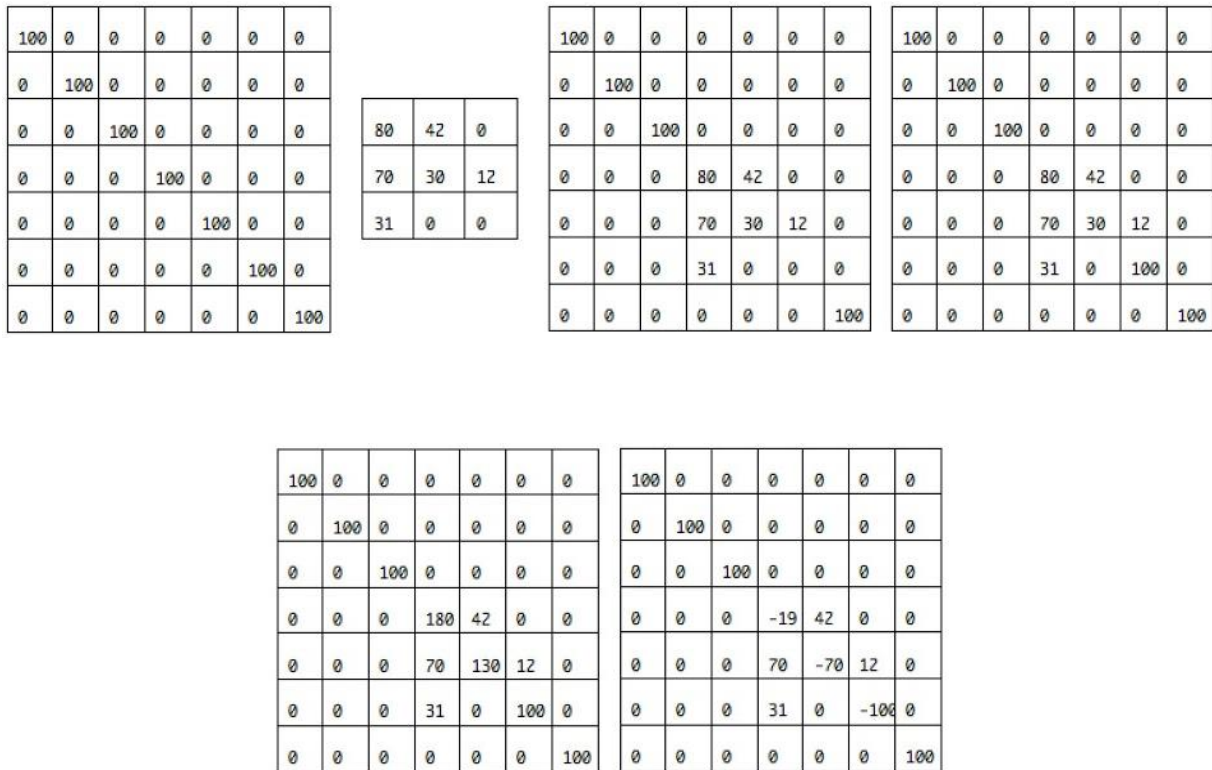


Figure 7.11. Représentation de l'application d'un motif avec l'abstraction *xel.2d.mat.brush* en position [3 3] avec à gauche la matrice porteuse et à sa droite le motif de coefficients. À droite et en bas, sont représentées les quatre matrices résultant des différents modes d'applications. En haut à gauche, se trouve la matrice résultant du mode « remplacement » où les neuf coefficients centraux sont remplacés par le motif de coefficients. En haut à gauche, se trouve la matrice résultant du mode « remplissage » où les neuf coefficients centraux sont remplacés par le motif de coefficients, à l'exception du coefficient en position [5 5] qui n'est pas modifié car il correspond à une valeur 0 dans le motif de coefficients. En bas à gauche se trouve la matrice résultante du mode « ajout » où les neuf coefficients centraux sont incrémentés par les valeurs contenues dans le motif de coefficients. Et en bas à droite, se trouve la matrice résultante du mode « retranchement » où les neuf coefficients centraux sont décrémentés par les valeurs contenues dans motif de coefficients.

L'abstraction *xel.2d.mat.brush* permet de générer et modifier des matrices de coefficients, en s'inspirant du système de motifs des logiciels présentés, visant à offrir des outils adaptés à la création d'espace sonore matriciel et notamment pour l'édition de matrices. L'abstraction nécessite de définir une matrice porteuse sur laquelle opérer, et un motif de coefficients à appliquer. La matrice porteuse peut être définie soit en envoyant sous forme de message une

<sup>263</sup> Cette approche se retrouve vraisemblablement dans d'autres logiciels, mais étant donné que le nombre d'actions sont relativement limitées dans ce logiciel comparé à GIMP ou Photoshop, cela fait de cette pratique un élément central.

taille matricielle qui génère une matrice vierge<sup>264</sup>, soit en envoyant une matrice de coefficients permettant de définir un état initial avancé. Le motif de coefficients peut être défini en envoyant une matrice de coefficient sous forme de message, où le terme *matrix* est remplacé par *motif*<sup>265</sup>. Les paramètres de fonctionnement initiaux étant établis, l'abstraction permet d'appliquer ce motif de coefficients à un sous ensemble de la matrice porteuse, selon une position définie par un indice de ligne et un indice de colonne, et suivant quatre modes originaux d'application afin de générer de nouvelles matrices [Figure 7.11] :

- Remplacement : ce mode, associé au symbole *replace*, permet d'appliquer le motif de coefficients en remplaçant les valeurs existantes dans la matrice porteuse par les valeurs du motif de coefficients.
- Remplissage : ce mode, associé au symbole *fill*, permet d'appliquer le motif de coefficients en remplaçant les valeurs existantes dans la matrice porteuse, par les valeurs du motif de coefficients lorsque celles-ci sont non-nulles [Figure 5].
- Ajout : ce mode, associé au symbole *add*, permet d'appliquer le motif de coefficients en ajoutant les valeurs du motif de coefficients aux valeurs existantes dans la matrice porteuse.
- Retranchement : ce mode, associé au symbole *subtract*, permet d'appliquer le motif de coefficients en soustrayant les valeurs du motif de coefficients aux valeurs existantes dans la matrice porteuse.

L'application d'un motif à une position donnée peut être réalisée selon une dernière valeur facultative correspondant à un facteur de multiplication des coefficients du motif, et permettant de gérer le taux d'application. Cette approche peut être réalisée en multipliant par un facteur la matrice correspondant au motif, avant de l'envoyer à l'abstraction *xel.2d.mat.brush*. Cependant cette approche est d'une part plus compliquée à mettre en œuvre, et plus coûteuse en ressources de calcul. Intégrer cette fonctionnalité au sein de l'abstraction permet de contrôler à la fois la position et le taux d'application du motif, avec une simple liste de valeurs facilitant la mise en œuvre de processus complexes [Figure 7.12]. Après l'application d'un motif de coefficients, la matrice résultante devient la matrice porteuse, et peut recevoir une nouvelle application du

---

<sup>264</sup> Il est aussi possible d'utiliser les arguments pour définir par défaut une matrice vierge d'une taille spécifique.

<sup>265</sup> Par défaut, le motif de coefficients est une matrice d'un élément de valeur 1. Ces matrices porteuses de motifs peuvent donc être créées « à la main », sous la forme d'un message, et être générées via les objets de la bibliothèque *iemmatrix* ou encore grâce à l'abstraction graphique présentée précédemment.

motif. L'utilisation du symbole *set* avant les messages modifiant la matrice porteuse, peut aussi être utilisée pour empêcher l'abstraction d'envoyer la matrice transformée. Cela permet de générer une matrice avec l'abstraction, de la modifier par un processus tel que l'utilisation d'un objet de la bibliothèque *iematrix* et de redonner cette matrice à l'abstraction, sans pour autant ici encore créer de boucle infinie dans la chaîne.

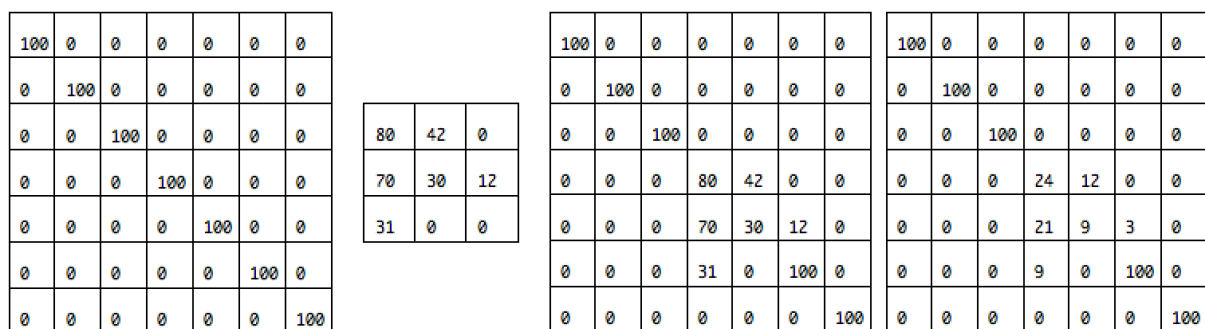


Figure 7.12. Représentation de l'application d'un motif avec l'abstraction *xel.2d.mat.brush* en position [3 3] avec à gauche la matrice porteuse, au centre le motif de coefficients et à droite les deux matrices résultant du mode remplissage avec un taux d'application de 1 à gauche et un taux d'application de 0.3 à droite. Les coefficients modifiés s'en trouvent diminués.

En combinant l'abstraction *xel.2d.mat.brush* avec l'abstraction graphique *xel.2d.gui.mat*, il est possible de créer un système récursif où le clic sur la représentation de la matrice de coefficients permet de déclencher sa modification à la position donnée, selon le motif de coefficients [Figure 7.13]. Cette approche plus complexe de l'édition graphique peut servir de base ou d'inspiration à de nouvelles approches en remplaçant l'abstraction *xel.2d.mat.brush* par un autre dispositif. Une mise en œuvre relativement simple pourrait par exemple permettre de modifier les coefficients d'une ligne de cellules définies, par deux clics consécutifs sur deux cellules. L'idée pour ce premier exemple est d'offrir un outil inspiré des logiciels de dessin numérique qui soit suffisamment flexible pour convenir à un large panel de pratiques et d'usages. Cette approche simplifie l'édition et la modification de matrices de coefficients selon des motifs complexes. Mais elle reste néanmoins cloisonnée à un usage dans un contexte de temps différé. L'étape finale est de concevoir des stratégies, afin de faciliter la modification et l'édition de l'espace sonore dans le temps.

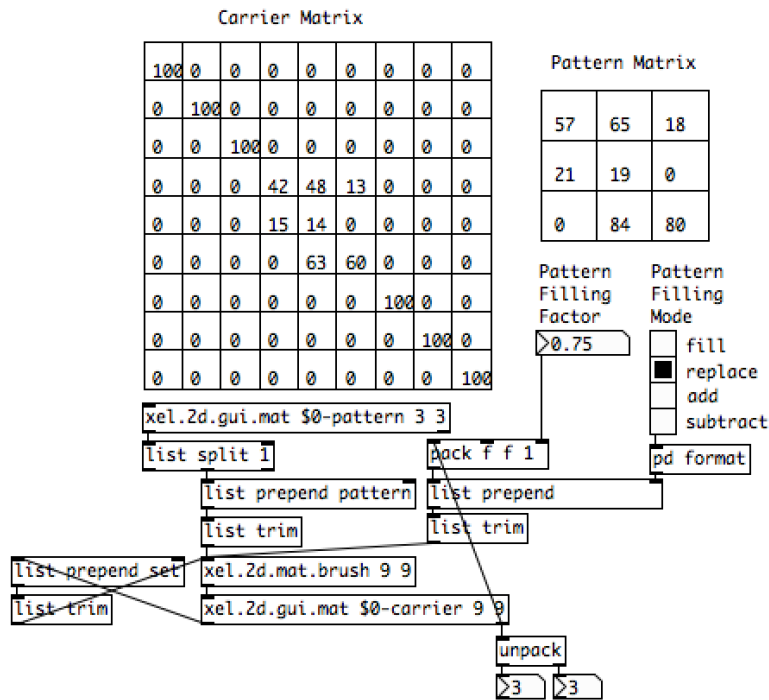


Figure 7.13. Patch Pure Data où un motif de coefficients de taille 3 × 3 représenté graphiquement en haut à droite est appliqué grâce à l'abstraction *xel.2d.mat.brush* à une matrice porteuse de taille 9 × 9 représentée graphiquement en haut à gauche selon un mode « remplacement », un facteur d'application de 0.75 et à la position définie par les indices [3 3] représentés en bas et générés par l'abstraction *xel.2d.gui.mat* suite à un clic de la souris sur l'interface de la matrice porteuse.

## 7.3. Utilisation, analyse et perspectives

L'édition en temps réel ou l'exécution dans le temps d'un ensemble de modifications sur les matrices de coefficients, et implicitement sur l'espace sonore matriciel, est la dernière question essentielle de l'utilisation de l'espace matriciel sur un plan compositionnel et musical. A partir des outils mis en œuvre, le contrôle des variations de l'espace sonore dans le temps peut recouvrir deux modes opératoires : l'édition en temps réel ou la transition entre des états prédéfinis de l'espace sonore<sup>266</sup>.

### 7.3.1. Gestion temporelle des motifs

L'édition en temps réel de l'espace sonore sous-entend de générer et modifier des matrices de façon suffisamment fluide pour être adaptée à une approche musicale. L'enjeu est de définir comment les outils présentés peuvent être utilisés, afin de varier et jouer sur l'espace sonore. L'obstacle majeur rencontré réside en la génération de motifs de coefficients à appliquer à des

<sup>266</sup> Evidemment ces deux modes opératoires peuvent être combinés dans la pratique. L'enjeu ici est de clairement délimiter les deux axes de recherches.

matrices porteuses, qui nécessite de multiples opérations de la part de l'utilisateur. En effet il est nécessaire pour cela de définir une taille matricielle pour le motif, les valeurs des coefficients, le mode d'application et possiblement un facteur d'application. Ce problème se retrouve aussi dans les logiciels de dessin numériques, car il n'est en effet pas envisageable d'obliger l'utilisateur à définir le motif de sa brosse à chaque fois qu'il lui est nécessaire d'en changer. La solution proposée consiste à pouvoir créer et sauvegarder des bibliothèques de brosses, afin de pouvoir par la suite les charger et les modifier rapidement. Dans le cadre de la bibliothèque de spatialisation matricielle, une abstraction graphique *xel.2d.gui.mat.brush*, librement inspirée de la fenêtre des options des outils du crayon dans le logiciel GIMP [Figure 7.14], a été mis en œuvre afin de répondre de manière similaire, à ce problème d'édition.

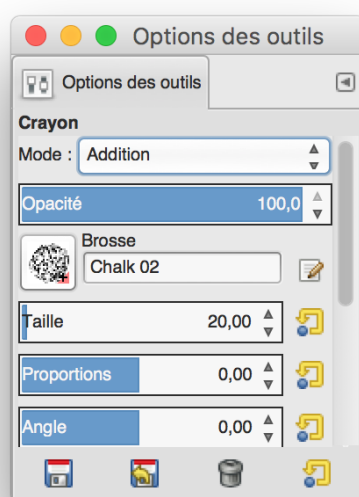


Figure 7.14 : Fenêtre des options des outils du crayon dans GIMP, où la brosse est définie selon un motif de type craie. Le mode d'application est de type « addition », l'opacité est de 100% et la taille est de 20 *pixels*.

Le fonctionnement de l'abstraction est légèrement différent de celui proposé dans le logiciel de dessin. Le fonctionnement est séparé en deux parties distinctes : l'une permettant l'édition des motifs de coefficients et l'autre permettant la gestion des motifs [Figure 7.15]. L'interface d'édition est relativement similaire à celle présentée en [Figure 7.13], à l'exception de la taille matricielle qui n'est plus définie que par une seule valeur<sup>267</sup> et du facteur d'application mis sur

---

<sup>267</sup> Le motif de coefficient est de manière arbitraire défini selon une matrice carrée dont la taille peut être définie selon la plus grande valeur entre le nombre de lignes et le nombre de colonnes d'une matrice rectangulaire, pour que l'approche soit équivalente.



une échelle de 0 à 100 pour offrir une interaction plus précise et correspondre à la valeur d'opacité utilisée dans les logiciels de dessin. L'interface de gestion des motifs est le réel apport de la mise en œuvre. Cette partie du module utilise une abstraction intermédiaire *xel.preset* qui offre un système basé sur les fonctionnalités de l'objet *text* de Pure Data et qui permet de gérer un ensemble de listes de messages. Via cet outil, il est possible de sauvegarder des listes de messages, pouvant correspondre à des états ou à un préréglage d'une matrice, mais dont la syntaxe est libre, en les associant à des entrées sous la forme de symboles. L'abstraction permet par la suite de recharger les listes de messages ou potentiellement de les supprimer, via leurs symboles d'entrées. L'interface graphique de l'abstraction permet simplement de faciliter la manipulation de cet outil, dans le cadre de matrice de coefficients. Cela notamment en représentant visuellement les symboles des entrées, et en offrant la possibilité de recharger les états associés de la matrice de coefficients par simple clic ou encore en facilitant l'enregistrement de nouveaux états ou leurs suppressions. La liste est elle-même enregistrable sous la forme d'un fichier textuel, pour une réutilisation ultérieure. A chaque modification du motif de coefficients ou de son mode d'application, l'abstraction renvoie les modifications qu'il est alors possible d'utiliser, de manière similaire à celle présentée dans la [Figure 7.13], et qui offre une solution adaptée à un usage en temps réel.

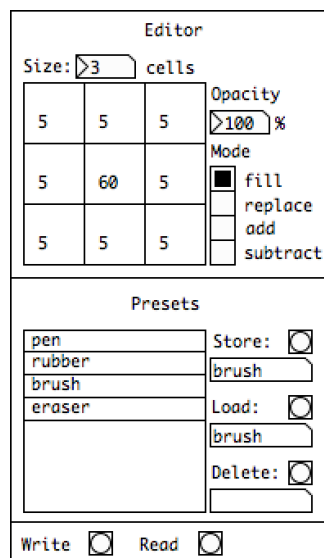


Figure 7.15. Abstraction graphique *xel.2d.gui.mat.brush* offrant quatre motifs prédéfinis, « pen », « rubber », « brush » et « eraser ». Le motif « brush » est choisi avec une taille de  $3 \times 3$ , une opacité de 100% et un mode d'application « remplissage ».

L'utilisation dans un contexte temps-réel est bien plus exigeante sur le plan fonctionnel, car les outils offerts doivent répondre rapidement et facilement aux volontés de l'utilisateur. Lors d'une utilisation en temps différé, l'utilisateur ne souffre que très peu de devoir utiliser des

opérations détournées ou intermédiaires, pour aboutir à ses fins. Alors qu'en temps-réel, s'il lui est nécessaire de réaliser deux actions là où une seule aurait pu être suffisante, cela révèle un vice de fonctionnement ou du moins une perspective d'amélioration. C'est le cas en ce qui concerne deux points que l'utilisation temps-réel des outils a permis de révéler. D'une part la relation entre la position d'application et le motif de coefficients peut, dans certain cas, être inadaptée. D'autre part, les modes d'application ne semblent pas suffire pour certaines opérations *a priori* simples.

La mise en œuvre actuelle de l'abstraction d'édition par motif, considère implicitement que la position d'application, définie par des indices de ligne et de colonne, correspond à la première cellule du motif. Or dans la pratique, il semble plus adapté d'associer cette position à la cellule centrale du motif, car les valeurs sont généralement réparties de façon radiale autour de cette cellule, tel que représenté sur la [Figure 7.15]. Cela empêche notamment d'appliquer directement la valeur centrale aux cellules des bordures du haut et de gauche. Il peut être facilement envisagé d'introduire un décalage de la position d'application en fonction de la taille matricielle du motif de coefficient. Cependant, aucune approche n'est réellement idéale et cette solution implique d'autres restrictions, comme le fait que la taille matricielle doit impérativement être impaire. Une approche dynamique telle qu'une option de compensation de la taille matricielle, qui soit potentiellement aussi configurable par l'utilisateur, semble donc être une solution acceptable.

L'édition par motif de coefficient suggère implicitement deux types d'opérations permettant l'augmentation et la diminution des gains associés aux *soxels*. Ces deux types d'opération correspondent dans les logiciels de dessin, à peindre et effacer et sont disponibles via deux outils distincts : le crayon (ou pinceau) et la gomme. La nécessité de deux outils peut se justifier par la plus grande complexité d'une matrice de *pixel*, due au fait qu'un *pixel* contient des informations de couleurs – généralement rouge, vert et bleu – rendant impossible ou très complexe pour l'utilisateur l'action d'effacer un ensemble de *pixels* avec l'outil crayon. L'approche sonore présentée n'associe le *soxel* qu'à un coefficient de gain. Dès lors, la nécessité d'un deuxième outil analogue à la gomme n'est pas nécessaire, car il est possible de réaliser les deux opérations d'augmentation et de diminution des gains, en jouant sur les valeurs du motif de coefficients et sur son mode d'application. Une application en mode « remplacement » ou « remplissage », avec des valeurs d'amplitude importante, permet de facilement augmenter les coefficients de la matrice porteuse, et donc les gains associés aux *soxels*. Et inversement, des valeurs d'amplitude faible permettent de diminuer les coefficients de la matrice porteuse et

donc les gains associés aux *soxels*. Cependant dans la pratique, ces deux approches peuvent paraître trop drastiques, car ces modes définissent les coefficients à des valeurs, sans prendre en compte les valeurs précédentes de la matrice. C'est à cet effet qu'ont été définis les modes « ajout » et « retranchement », permettant d'augmenter et de diminuer graduellement les valeurs des coefficients de la matrice porteuse. Utilisée conjointement avec une fonction de limitation des valeurs<sup>268</sup>, une opération de type crayon pour l'un des types de gomme, est aisément réalisable. L'usage voudrait qu'une opération de type gomme sur les coefficients, amène les valeurs de gain à 0. Or, en considérant que les valeurs des coefficients puissent être négatives, un mode de type « retranchement » qui soustrait des valeurs aux valeurs présentes des coefficients, impliquerait qu'au bout d'un moment, les valeurs positives dépassent 0 et deviennent négatives<sup>269</sup>. Aussi, il serait intéressant d'envisager d'autres modes d'insertion, permettant de réaliser d'autres opérations arithmétiques simples répondant à ce genre de problèmes, telles que la division ou la multiplication.

### 7.3.2. Gestion temporelle des matrices

Une deuxième approche de gestion temporelle des matrices de coefficients, consiste à pouvoir définir à l'avance un ensemble d'états. La mise en œuvre peut dans ce contexte s'inspirer du logiciel Aseprite, qui offre la fonctionnalité de créer des animations GIF à l'aide d'une succession d'images (*frames*). Le logiciel offre un cadre permettant la gestion temporelle de l'image, en associant à une valeur numérique un état de l'image, une *frame*, correspondant en soit à une nouvelle image, généralement issue d'une variation de l'image originale. L'objectif étant, de manière identique à la vidéo, de créer du mouvement en faisant défiler successivement les *frames*. Aussi, de manière identique, il est possible d'associer l'état d'une matrice de coefficients à une *frame* et d'offrir un système permettant de passer facilement d'un état à un autre.

---

<sup>268</sup> En utilisant le quatrième et le cinquième argument de l'abstraction *xel.2d.gui.mat* par exemple.

<sup>269</sup> Donc le signal sonore serait non nul.

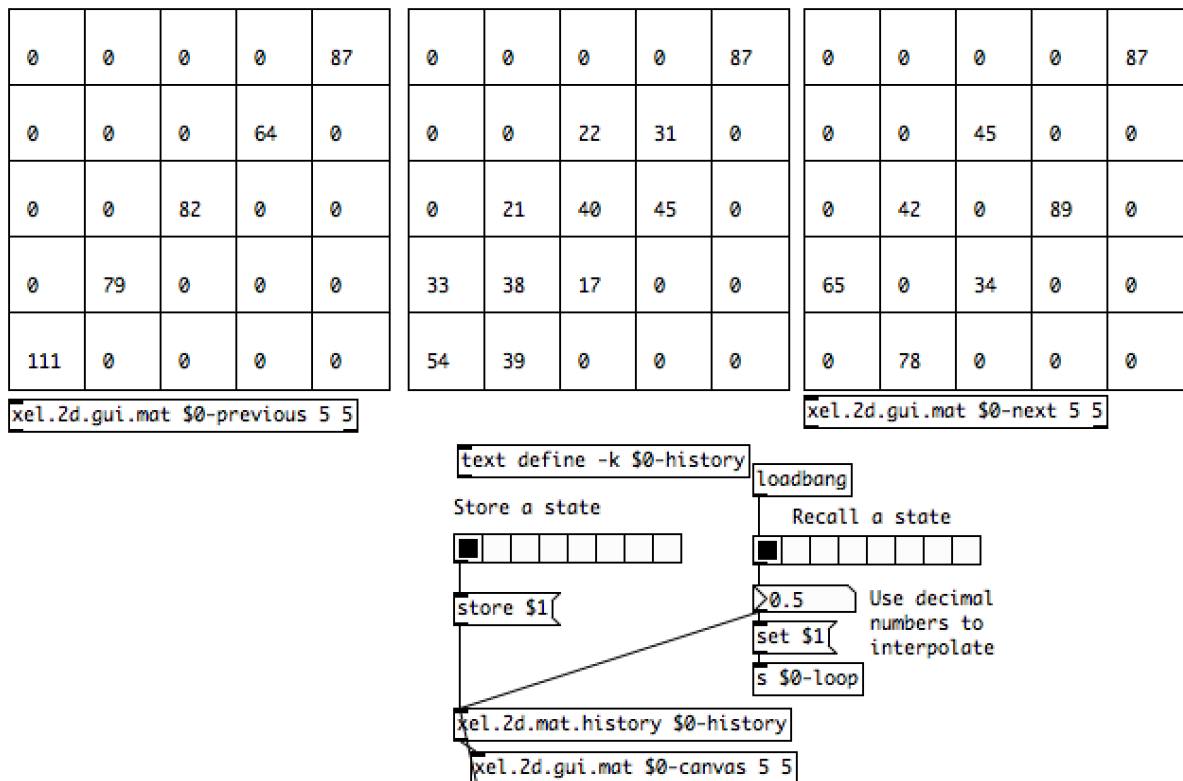


Figure 7.16. Abstraction graphique *xel.2d.mat.history* permettant de gérer des états de matrices de coefficients et d'interpoler entre ces états. Dans cet exemple, à gauche est représentée la matrice de coefficients associée à l'état de l'entrée 0 et à droite est représentée la matrice de coefficients associée à l'état de l'entrée 1. Au centre est représentée la matrice de coefficients associée à l'état intermédiaire 0,5 généré par l'abstraction.

L'abstraction *xel.2d.mat.history* permet de générer des états de matrices selon un fonctionnement similaire à l'abstraction *xel.preset*. En somme, elle permet d'enregistrer des listes de messages et de recharger ces listes en utilisant l'objet *text* de Pure Data, afin de contenir les données. La première différence avec l'objet plus générique *xel.preset*, est que les entrées sont définies par un nombre. Aussi, son fonctionnement se prête plus à une utilisation telle que celle qui vient d'être présentée avec le logiciel Aseprite. En incrémentant un compteur dont la valeur définit les indices d'entrée des états de la matrice, il est possible de renvoyer successivement les différents états préalablement enregistrés, afin de créer des variations. Cependant, il est évident que la temporalité des images n'est pas similaire à la temporalité de son ou de l'espace sonore. Mais de manière plus générale, le contexte n'étant pas le même, les enjeux sont différents. Aussi, il ne s'agit pas dans ce contexte de spatialisation du son, de donner une impression de mouvement tel qu'un personnage se déplaçant ou encore d'enchaîner les images à une vitesse définie selon la perception visuelle humaine nécessitant 24 ou 25 images par secondes. L'enjeu est plutôt de définir des éléments clés, qu'il serait possible d'associer à des temps sur une ligne temporelle, mais aussi à tout autre types d'évènements et de pouvoir glisser entre deux états. Pour cela, l'abstraction propose un système permettant de réaliser une

interpolation entre deux états de la matrice de coefficients, en envoyant une valeur décimale intermédiaire [Figure 7.16]. Cette opération est réalisée par l'abstraction, en récupérant tout d'abord les deux matrices de coefficients associées aux deux états voisins, et dont les entrées sont la valeur entière en-dessous et la valeur entière au-dessus de la valeur décimale donnée. Puis, coefficient par coefficient, l'abstraction réalise une interpolation linéaire entre ces deux matrices de coefficients, selon la partie décimale donnée.

### 7.3.3. Bilan et perspectives

En s'inspirant de la pratique du Pixel Art, une première approche de l'espace sonore matriciel a été proposée. Élaborée pour une prise en main relativement bas niveau, cette approche vise à éditer et à transformer l'espace sonore *soxel* par *soxel*. Partant de cette restriction, un nombre néanmoins relativement important de procédés ont pu être envisagés. L'enjeu étant de concevoir de pratiques, tout en répondant aux nouvelles problématiques posées par la transposition de l'approche dans le contexte de la spatialisation du son. Cependant, de nombreuses solutions ont été proposées, en s'inspirant des logiciels de dessin numérique. Le résultat est un ensemble d'outils sous la forme d'abstractions, offrant des fonctionnalités variées et répondant à différentes problématiques telles que la synthèse même de l'espace sonore, l'édition graphique de l'espace sonore ou encore la gestion temporelle. Un certain nombre d'outils ont été évoqués mais n'ont pas pu être pleinement présentés, tels que l'abstraction *xel.2d.gui.scope~* permettant de représenter la RMS de l'amplitude des signaux associés aux *soxels* [Figure 7.5]. Et de nombreuses améliorations et nouvelles fonctionnalités sont aussi envisagées, comme l'ajout de couleurs à l'interface graphique *xel.2d.gui.brush*, afin d'offrir justement un retour visuel plus explicite et en adéquation avec l'interface graphique *xel.2d.gui.scope~*. Il est aussi envisagé d'offrir un mode où l'édition puisse être réalisée de manière plus identique aux logiciels de dessins, en faisant glisser la souris pour éditer les coefficients associés à plus de *soxels*. Enfin, il est aussi envisagé de réaliser une interface enveloppant les abstractions *xel.2d.gui.mat* et *xel.2d.mat.history* et permettant de gérer temporellement plusieurs couches de matrices de coefficients, afin d'offrir un outil plus adapté à l'abstraction *xel.2d.syn.pencil~*. L'ensemble de ces procédés ne restent pour le moment qu'à un état d'expérimentation, mais ils possèdent un fort potentiel musical et proposent des approches pour le moins originales.

Enfin une dernière approche est envisagée liée à l'utilisation conjointe des abstractions *xel.2d.gui.mat* et *xel.2d.mat.brush*. Il s'agit de réappliquer avec un petit temps de retard sur la

matrice de coefficients, les motifs de coefficients appliqués, mais selon un mode inversé. L'idée est de réaliser un effet de « crayon qui s'efface », en faisant disparaître toutes les nouvelles contributions après un temps défini. Une autre possibilité serait même de faire disparaître les contributions, seulement à chaque nouvelle contribution. Et ce type d'approche amène à une conception peut être plus classique de l'espace sonore où le motif de coefficient serait associé à une source sonore plus ou moins ponctuelle et pouvant se déplacer dans l'espace. Aussi serait-il intéressant, à présent, de réaliser des propositions moins spécifiques sur ce qu'est le Soxel Art, mais qui répondraient à des problématiques plus communes, telles que la synthèse de source sonore ponctuelles. L'enjeu serait aussi de définir comment projeter cet espace sonore sur un système de restitution, ou encore de définir quelles seraient les opérations et transformations réellement spécifiques à cette représentation matricielle de l'espace.

## 8. Développement de la représentation matricielle

L'approche matricielle de l'espace sonore a permis d'élaborer un certain nombre d'outils répondant à un ensemble de pratiques, regroupées sous le terme de Soxel Art. L'enjeu était alors de définir ce qui fait l'originalité de cette proposition, en définissant concrètement le modèle de représentation de l'espace et en l'associant aussi bien à un système de pensée, qu'à des outils mathématiques et une pratique artistique. Bien qu'inspirante, cette approche reste cependant cloisonnée au domaine des idées, aussi est-il nécessaire de pouvoir l'intégrer dans un contexte plus concret. Et au-delà, il semble important de développer la mise en espace du son fondée sur une représentation matricielle de l'espace, autour de nouvelles approches. Pour cela, ce chapitre propose d'offrir des solutions à la restitution de l'espace sonore sur un système de haut-parleurs. Mais aussi de développer les opérations autour du cadre mathématique, notamment via l'élaboration de filtres spatiaux. Enfin l'objectif final est de proposer des approches plus systématiques de synthèse de l'espace sonore, répondant aux opérations d'encodage généralement mises en œuvre.

### 8.1. Projection des soxels

Au-delà de permettre la restitution de l'espace sonore sur un système de haut-parleurs, la mise en œuvre d'une technique permettant de projeter l'espace sonore matriciel répond à un double objectif. Il s'agit, d'une part, de montrer par une approche pratique comment l'expérience acquise des traitements du son spatialisés via les techniques ambisoniques peut être généralisée, notamment via la question de l'adaptation de l'opération de projection des signaux associés aux harmoniques vers un système irrégulier. Il s'agit aussi de voir comment cette approche peut

permettre de concevoir l'opération de projection des composantes de la représentation matricielle de l'espace sonore. D'autre part, l'objectif est aussi de définir une méthodologie permettant de généraliser ce type de mises en œuvre à d'autres représentations de l'espace que celle matricielle.

Le cheminement commun lors de la création d'une technique de spatialisation est de partir d'une problématique liée à un système de restitution ou de contraintes pratiques et matérielles, à savoir par exemple un système de haut-parleurs hémisphériques ou le placement spécifique de l'auditoire. Puis, il s'agit d'établir un système répondant au mieux à ces contraintes et de mettre en œuvre une technique ou un algorithme. C'est notamment l'approche suivie pour la mise en œuvre du VBAP [Pulkki, 1997] ou du DBAP [Lossius & al., 2009] mais aussi, à moindre mesure, pour l'approche ambisonique<sup>270</sup>. L'approche expérimentale présentée dans ce travail réalise le parcours inverse en partant donc d'une représentation intermédiaire et abstraite de l'espace pour définir, par la suite, les systèmes de restitution avec lesquels elle peut être utilisée et les dispositions de l'auditoire acceptées dans ce cadre. Car évidemment, tous les systèmes de restitution ne conviennent pas à toutes les représentations abstraites de l'espace. Et il en est de même pour la disposition de l'auditoire et logiquement, pour l'algorithme à mettre en œuvre pour réaliser la projection des composantes spatiales de l'espace sonore vers les systèmes de restitution. Cette spécificité de cette approche inverse est donc intéressante et à la fois complexe, car elle est affranchie de tout système de restitution réel. Il est néanmoins nécessaire de définir par la suite les contextes dans lesquels utiliser la représentation abstraite de l'espace sonore, ainsi que les outils nécessaires et les contraintes matérielles que cela implique.

Définir la technique de projection selon cette approche, implique plusieurs étapes. Dans un premier temps, il faut définir les choix conceptuels liés notamment à la manière de mettre en relation l'espace sonore abstrait et l'espace réel de restitution. Il s'agit là de concevoir un espace réel tel une entité opaque, sans considérer un système de restitution et un auditoire spécifiques. L'objectif est de comprendre comment l'espace sonore abstrait et ses composantes doivent être représentés dans un espace réel. Par la suite, il sera possible de réaliser des techniques permettant de répondre à ce choix conceptuel. Pour cela, l'approche s'inspirera des opérations

---

<sup>270</sup> Comme présenté en première partie de cette thèse, l'approche ambisonique émerge de la nécessité d'offrir un meilleur système de spatialisation du son que ceux proposés par les techniques quadriphoniques classiques associées à des configurations de haut-parleurs spécifiques. C'est pour un auditeur placé au centre du dispositif que l'approche a été établie.

proposées en ambisonie, à savoir des combinaisons avec des techniques tierces comme le VBAP ou la synthèse binaurale. Enfin, une mise en œuvre pourra être proposée sous la forme d'une abstraction. Il sera alors possible de présenter sa réalisation et son fonctionnement.

### 8.1.1. Choix conceptuels

L'analyse de l'approche ambisonique a permis de définir que projeter l'espace sonore sous sa forme abstraite, nécessite de le mettre en relation avec un espace réel. Plus concrètement, mettre en place l'opération de projection nécessite de mettre en relation un espace réel de restitution, associé à un auditoire et/ou un système de restitution, avec la représentation abstraite de l'espace sonore. Deux approches opposées sont envisageables, afin d'orienter les choix à réaliser. L'une définit, de manière analogue à l'ambisonie, que les sources sonores sont infiniment lointaines. Ce qui élude de manière implicite la notion de distance. L'autre définit que les sources sonores sont proches de l'espace sonore, ce qui intègre au système de manière tacite la notion de distance. Chacune de ces alternatives amène à envisager deux sous-possibilités quant à la relation entre l'espace sonore abstrait et l'espace réel de restitution. La première possibilité suppose que l'espace réel de restitution soit complètement dissocié de l'espace sonore abstrait, et que les deux espaces ne se recoupent pas. La deuxième assume quant à elle que l'espace réel de restitution est superposé à l'espace sonore abstrait. L'opération consiste alors à calquer un espace sur l'autre. Aussi, quatre possibilités apparaissent selon que la distance est ou n'est pas prise en compte dans le modèle de relation et que les deux espaces, réel et abstrait, se superposent ou non. L'objectif est alors de définir les implications de chaque possibilité, afin de définir le choix de relation semblant le plus en adéquation avec la représentation matricielle de l'espace sonore :

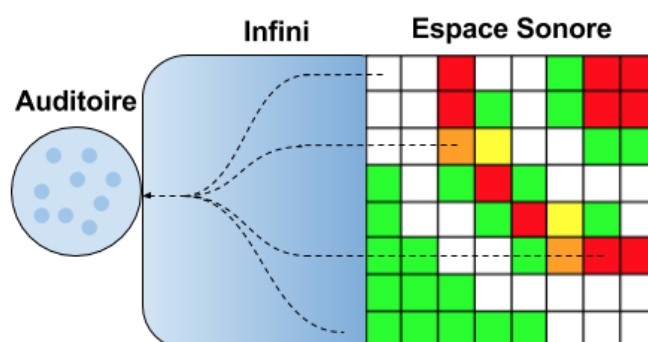


Figure 8.1. Représentation de la première possibilité de relation entre l'espace sonore sous sa forme matricielle et un auditoire associé à l'espace réel de restitution. Les flèches représentent la manière dont les *soxels* doivent être projetés vers l'auditoire. Les deux espaces sont dissociés et séparés d'une distance infinie, aussi l'ensemble des contributions des *soxels* apparaît sous la forme d'un signal provenant d'une direction unique.



1. La première possibilité assume que les deux espaces sont dissociés et que la distance est considérée comme infinie. Dans ce cas, l'espace réel de restitution est infiniment loin dans une direction donnée de l'espace sonore abstrait. L'espace sonore abstrait peut alors être considéré comme un élément ponctuel infiniment petit, et en-dehors de l'espace réel de restitution. Et l'ensemble des composantes spatiales, les *soxels*, sont alors considérés comme des sources sonores provenant de ce point [Figure 8.1]. Par cette approche, les informations spatiales des *soxels* deviennent inutiles, car ceux-ci se superposent tous en ce point. Et la seule information disponible pour l'opération de projection est la direction du point par rapport à l'espace réel de projection. Aussi, que la décomposition de l'espace soit réalisée par un *soxel* ou un millier revient au même. Et qu'un *soxel* soit avant ou après un autre n'implique finalement aucune variation dans leur mode de projection. Donc, sur un plan pratique, l'opération revient à sommer directement sans même un jeu de gain, tous les signaux associés aux *soxels*, sans considération de leurs positions matricielles. Puis l'opération consiste à les restituer en tentant au mieux de reproduire une onde plane. Au-delà d'un rendu qu'il est possible de présupposer comme peu intéressant sur le plan spatiale, l'approche entraîne une énorme perte d'informations et manque totalement de cohérence avec la représentation de l'espace sonore. Aussi, cette approche peut d'ores et déjà être abandonnée.



Figure 8.2. Représentation de la deuxième possibilité de relation entre l'espace sonore sous sa forme matricielle et un auditoire associé à l'espace réel de restitution. Les flèches représentent la manière dont les *soxels* doivent être projetés vers l'auditoire. Les deux espaces se superposent et la distance est infinie. Aussi l'espace réel de restitution apparaît comme infiniment loin « au centre » de l'espace sonore matriciel. Alors seules les informations directionnelles des *soxels* par rapport au centre de l'espace sonore matriciel, sont considérées dans l'opération de projection.

2. La deuxième possibilité assume que les deux espaces se recoupent et que la distance est toujours considérée comme infinie. Il est alors nécessaire de faire preuve d'une bonne faculté d'abstraction, car cette approche suppose que l'espace sonore matriciel entoure

l'espace réel de restitution, tout en étant infiniment éloigné de celui-ci. La relation entre ces deux espaces peut être appréhendée en supposant une  $n^{\text{ième}}$  dimension spatiale, à savoir une courbure de l'espace. Ou de façon plus pragmatique, un système macroscopique associé à l'espace sonore matriciel et un système microscopique associé à l'espace réel de restitution. Aussi, les notions de distances étant relatives à chaque système, une distance quantifiable pour le premier devient une distance infinie pour le second [Figure 8.2]. Contrairement à la première approche, cette possibilité possède l'avantage d'offrir un champ sonore enveloppant. Cependant, les rapports de relations spatiales entre les *soxels* se retrouvent réduits à une variation de leurs incidences respectives vis à vis de l'auditoire. Cette approche implique une faible variation angulaire entre les *soxels*, et même une superposition angulaire pour certains *soxels* se trouvant sur les axes verticaux et horizontaux médians ou les axes diagonaux. Aussi, cette approche manque de cohérence avec la représentation de l'espace matricielle de l'espace sonore. Et, bien que cela soit dans une moindre mesure par rapport à l'approche précédente, cette relation entre les deux espaces implique aussi une perte d'informations non négligeable.

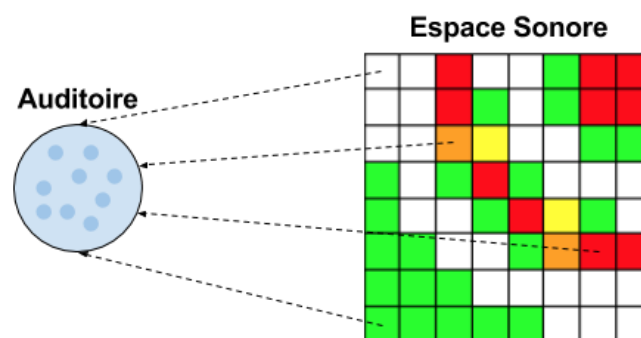


Figure 8.3. Représentation de la troisième possibilité de relation entre l'espace sonore sous sa forme matricielle et un auditoire associé à l'espace réel de restitution. Les flèches représentent la manière dont les *soxels* doivent être projetés vers l'auditoire. Les deux espaces sont dissociés et séparés d'une distance finie. Aussi, les contributions des *soxels* apparaissent sous la forme d'un ensemble de signaux provenant de multiples directions, limitées cependant à un demi-cercle.

3. La troisième possibilité assume que les deux espaces sont dissociés, mais que la distance est à présent considérée comme finie. De manière identique à la première approche, il faut alors considérer l'espace sonore abstrait comme provenant globalement d'une direction donnée vers l'espace réel de restitution. Mais cette approche assume que l'espace sonore abstrait se situe suffisamment proche de l'espace réel de restitution, pour prendre en considération les distances des *soxels*. Cette prise en compte de la distance des *soxels* permet de considérer l'ensemble des informations relatives à leurs

positions au sein de l'espace sonore matriciel. Cela impliquant une variation de leurs projections respectives, selon le procédé mis en œuvre par la suite. L'inconvénient de cette approche est que l'étendue angulaire maximum sur laquelle peut être projeté l'espace sonore est comprise dans un demi-cercle [Figure 8.3]. Cette étendue dépend notamment de la distance globale de l'espace sonore abstrait par rapport à l'espace réel de restitution, et de leurs tailles relatives. Un espace sonore abstrait considéré comme sensiblement plus petit que l'espace réel de restitution, sera associé à une étendue angulaire bien plus réduite que s'il était considéré comme aussi grand, voire plus, que l'espace réel de restitution. Cette caractéristique implique qu'il est fort vraisemblable d'aboutir, là encore, à de faibles variations angulaires entre les *soxels* par rapport à l'espace réel de restitution, impliquant une superposition de certaines contributions, notamment pour les *soxels* associés à l'axe horizontal médian. Cette relation entre l'espace sonore matriciel et l'espace réel de restitution offre potentiellement une perte d'information minime comparée aux précédentes approches. Cela amène possiblement à une projection plus cohérente des *soxels*, mais néanmoins sous-optimale.



Figure 8.4. Représentation de la troisième possibilité de relation entre l'espace sonore sous sa forme matricielle et un auditoire associé à l'espace réel de restitution. Les flèches représentent la manière dont les *soxels* doivent être projetés vers l'auditoire. Les deux espaces sont superposés et la distance entre les *soxels* et l'espace réel de restitution est finie. Aussi, les contributions des *soxels* apparaissent sous la forme d'un ensemble de signaux provenant de multiples directions enveloppant l'espace réel de restitution, à des distances différentes.

4. La troisième et dernière possibilité assume que les deux espaces se superposent, et que la distance est à présent considérée comme finie. Dans cette approche, l'espace réel de restitution recouvre l'espace sonore matriciel, sans pour autant définir que cette couverture soit partielle ou réalisée de manière homogène. Aussi est-il possible que l'espace réel de restitution ne recouvre qu'une partie de l'espace sonore abstrait, ou inversement qu'il recouvre intégralement l'espace sonore abstrait et en dépasse les

limites. Bien que ces changements de perspectives n'impliquent pas de variations dans les mécanismes de projection, une approche pratique aura néanmoins tendance à centrer, dans la limite du possible, les deux espaces [Figure 8.4]. Cette possibilité offre l'avantage d'ouvrir énormément la répartition angulaire des *soxels* vis-à-vis de l'espace réel de projection. En définissant un point d'écoute central ou un auditoire centré dans l'espace réel de restitution, les *soxels* recouvrent alors angulairement l'ensemble du cercle définissant cet espace réel et amènent, de fait, à considérer un champ sonore enveloppant. Le risque que les contributions des *soxels* se chevauchent lors de leurs projections est donc moindre. De plus, la prise en compte de leurs distances respectives comme une information importante dans l'opération de projection, permet de créer des variations sensiblement intéressantes et de dissocier leurs contributions respectives. Même si plusieurs *soxels* sont proches angulairement, une particularité notable de cette approche réside dans le fait que certains *soxels*, ou possiblement l'ensemble même des *soxels*, peuvent être considérés comme provenant de l'intérieur de l'espace réel de restitution. Cette dernière proposition de relation entre l'espace sonore matriciel et l'espace réel de restitution, offre le moins de perte d'informations, en préservant au maximum les informations à la fois directionnelles et de distances. Selon les choix réalisés pour la mise en œuvre de projection, cette approche devrait donc offrir une représentation plus fidèle de l'espace sonore, que celles résultant des propositions précédentes.

De l'analyse de ces quatre propositions, il est possible d'affirmer que les modèles qui prennent en compte la distance pour projeter l'espace sonore abstrait sur un système de restitution, semblent les plus adaptées à une représentation en adéquation avec l'approche matricielle. Dans une approche contraire, le risque est de perdre un nombre important d'informations spatiales, engendrant dès lors une restitution non cohérente de l'espace sonore matriciel. Enfin, envisager une superposition des deux espaces, abstrait et réel, permet de faciliter la prise en compte des informations angulaires et de distances. Aussi la quatrième proposition est, de ce fait, la plus adaptée pour élaborer une approche pratique.

### 8.1.2. Choix techniques

L'enjeu à présent, est de définir l'opération permettant de répondre à la relation entre l'espace sonore abstrait et l'espace réel de projection, en transformant les signaux associés aux *soxels* en un ensemble de signaux destinés à un système de restitution. L'une des solutions avancées en ambisonie afin de restituer un espace sonore pour un système de haut-parleurs, est de projeter

l'espace sonore sur un ensemble de haut-parleurs virtuels et intermédiaires, pouvant être associés à de nouvelles sources sonores virtuelles. Puis de redistribuer ces sources sonores virtuelles sur le système de haut-parleurs réels, par l'intermédiaire d'une technique de spatialisation du son tierce. Plusieurs approches ont été proposées, toutes n'ont pas été traitées, mais comme démontré précédemment, le choix de l'utilisation du VBAP [Batke & al., 2010], [Zotter & al., 2010, 2012, 2013] est particulièrement adapté à l'approche ambisonique. La principale raison sous-jacente à ce choix est que l'ambisonie assume dans cette approche, que les signaux des haut-parleurs virtuels sont associés à des ondes planes. Aussi, seules leurs informations de direction sont à prendre en compte. Ainsi, des techniques de spatialisation du son tierces, dont les approches sont élaborées autour d'autres informations comme la distance ou l'élargissement de la source sonore virtuelle, ne sont évidemment pas optimales et favorisent donc l'approche directionnelle du VBAP<sup>271</sup>. En un sens, l'opération de projection des signaux associés aux harmoniques sphériques sur une sphère, peut être associée au choix qui vient d'être réalisé, permettant de définir la relation entre l'espace sonore abstrait et l'espace réel de restitution. Cette opération de projection définit implicitement une relation entre les signaux associés aux harmoniques définissant l'espace sonore sous une forme abstraite et la sphère sur laquelle ils sont projetés et qui peut être associée à l'espace réel de restitution. Aussi, l'utilisation combinée du VBAP peut être associée au choix pratique et technique permettant de réaliser l'opération de projection sur un système de haut-parleurs réels. Dans le contexte de la représentation matricielle de l'espace sonore, la projection des *soxels* nécessite alors de trouver la ou les techniques de spatialisation les plus adaptées. Notamment selon la relation établie entre cet espace sonore abstrait et l'espace réel de restitution. Aussi, le choix doit être réalisé en considérant les avantages et les contraintes inhérentes à cette relation. Il n'est pas possible de considérer toutes les approches pratiques envisageables pour réaliser la projection des *soxels* sur un système de haut-parleurs, car il est vain de vouloir énumérer toutes les techniques de spatialisations existantes. L'objectif est donc d'envisager un ensemble restreint de techniques de spatialisation clairement établies sur un plan conceptuel et fonctionnel, afin de limiter les possibilités et le champ d'investigation, tout en permettant vraisemblablement d'offrir une solution adéquate.

---

<sup>271</sup> Il est cependant intéressant de remarquer que via son extension sous la forme du MDAP (*multiple-direction amplitude panning*) il est possible de contrôler l'élargissement des sources sonores. Mais bien que ce système puisse avoir un intérêt certain lorsqu'il est combiné avec l'approche ambisonique, cette fonctionnalité reste additionnelle et n'est pas nécessairement à prendre en compte dans le choix de la technique.

Selon une approche similaire à la combinaison ambisonie-VBAP mais aussi ambisonie-binaurale, la projection des *soxels* consiste concrètement à utiliser les informations spatiales de chaque *soxel* comme coordonnées de sources sonores virtuelles et les signaux qui leurs sont associés comme les signaux de ces sources sonores virtuelles. Partant de ce principe, différentes méthodes peuvent être proposées. Bien entendu, aucun système n'est idéal et comme le remarque M. Poletti, « dans tout système pratique, des compromis sont inévitables »<sup>272</sup>. Il s'agit donc de définir les critères qui semblent les plus judicieux à respecter au regard de la représentation de l'espace sous forme matricielle, et des techniques de projections connues en considérant leurs champs d'applications. Et bien entendu en considérant de manière pragmatique le système, ou plutôt les types de systèmes de restitution pour lesquels ils sont destinés. Des techniques de spatialisation seront donc critiquées selon les possibilités qu'elles offrent pour la spatialisation de sources sonores ponctuelles et les libertés qu'elles permettent pour le placement des haut-parleurs. Ou de manière générale, le système de restitution qu'elles permettent d'envisager, et enfin les contraintes qu'elles impliquent vis-à-vis de la répartition de l'auditoire :

- Le VBAP [Pulkki, 1997] offre la possibilité de restituer des sources sonores pour un système de haut-parleurs relativement flexible, mais devant être placé sur un cercle ou sur une sphère, dans le cas d'une restitution tridimensionnelle. La technique est destinée à un auditeur centré, bien que dans la pratique, un auditoire peut recouvrir une large étendue de l'espace défini par le système. Cette technique offre l'avantage de pouvoir *a priori* être associée aux troisième et quatrième schémas de relation entre l'espace sonore et l'auditoire. En effet, l'approche convient aussi bien à un système restreint pour un champ sonore frontal qu'à un système élargi adapté à un champ sonore enveloppant. Cependant, le système n'admet pas la distance de la source sonore, ainsi il est de fait inapproprié à une restitution de l'espace sonore sous une forme matricielle.
- La synthèse binaurale [Møller, 1992] serait potentiellement plus adaptée. En effet, il semble envisageable d'associer pour chaque *soxel*, un jeu de deux HRIR, les réponses impulsionnelles relatives à la tête correspondant chacune respectivement à une oreille. L'approche consisterait alors à réaliser une synthèse binaurale en filtrant les signaux associés aux *soxels* avec leurs HRIR respectives, et pour chaque oreille, d'additionner

---

<sup>272</sup> [Poletti, 2007], p. 598, traduit de l'anglais : « In any practical system some compromise are inevitable. »

les signaux qui leurs sont destinés. Bien que le nombre de réponses impulsionnelles nécessaires puisse être relativement limité comparé au nombre de réponses impulsionnelles disponibles dans des bibliothèques telles que LISTEN de l'IRCAM ou CIPIC de UC Davis [Algazi & al., 2001], les positions qui devraient leur être associées sont particulièrement incongrues pour une approche binaurale. Dans la pratique, de tels jeux de réponses impulsionnelles n'existent pas. Le cadre de ces recherches ne peut se permettre de tels enregistrements et expérimentations, qui seraient un sujet de recherche en soi. Aussi cette approche n'est pas envisageable dans ce contexte<sup>273</sup>.

- L'ambisonie permet de synthétiser et transformer des champs sonores pour un système de haut-parleurs répartis également sur un cercle ou une sphère, et enveloppant l'auditeur. De nombreuses variations du système permettent d'éviter la restriction du placement des haut-parleurs, notamment en associant le VBAP à l'ambisonie. Cependant, cette technique étant elle-même écartée, l'utiliser n'est pas approprié<sup>274</sup>. Cette technique peut, à propos du système de restitution, être associée à la quatrième proposition de relation entre l'espace sonore abstrait et l'espace réel de restitution<sup>275</sup>. Bien que valide pour un seul et unique auditeur, l'espace de restitution peut être élargi à un auditoire relativement large, au moyen d'optimisations. Cependant, tout comme pour le VBAP, l'ambisonie n'offre pas, dans son modèle initial, la prise en compte la distance des sources sonores<sup>276</sup>.

---

<sup>273</sup> Cependant, une solution temporaire peut être d'envisager l'utilisation d'une deuxième technique de spatialisation intermédiaire, entre l'approche matricielle et l'approche binaurale, afin de faire concorder des sources sonores ponctuelles et virtuelles, avec les positions d'enregistrements associées à des réponses impulsionnelles binaurales existantes.

<sup>274</sup> De manière générale, il serait nécessaire d'étudier plus en profondeur la possibilité de combiner à une technique de projection une troisième ou n<sup>ïèmes</sup> techniques, afin de projeter l'espace sonore. Mais dans ce contexte d'espace sonore matriciel restitué via l'ambisonie, l'approche semble peu pertinente et trop compliquée pour une première approche.

<sup>275</sup> Sur un plan fonctionnel, l'approche ambisonique pourrait aussi être associée à la troisième proposition, cependant la moitié du champ sonore – celui opposé à l'espace sonore – étant vierge, l'installation de haut-parleurs entourant l'auditeur est peu adaptée.

<sup>276</sup> Deux possibilités sont néanmoins envisageables pour pallier ce problème de distance. Celle proposée dans le cadre de la bibliothèque HOA et celle proposée par Daniel [Daniel & al., 1998]. La première proposée dans le cadre de la bibliothèque HOA et qui est fondée sur la simulation d'ordres fractionnaires pour le rapprochement de la source sonore, semble peu adaptée. Notamment si l'espace réel de restitution recouvre une large partie de l'espace sonore, car les ordres fractionnaires s'approchent d'un champ sonore omnidirectionnel et impliquent une perte d'information spatiale des *soxels*, offrant alors une représentation peu cohérente de l'espace sonore abstrait. La deuxième approche proposée par Daniel, nécessiterait quant à elle d'être déjà validée dans la pratique. Aussi, les présentes recherches ne peuvent se permettre d'étendre son usage à l'opération de la projection de la représentation matricielle de l'espace sonore.

- La WFS [Berkhout & al., 1993] permet de synthétiser des sources sonores, via une ligne de haut-parleurs très rapprochés. Cette technique permet notamment de pouvoir simuler la distance des sources sonores, leur éloignement mais aussi leur rapprochement, lorsque la source est devant le système de restitution. Cela en fait une technique *a priori* optimale pour projeter les *soxels*, notamment selon la relation entre l'espace sonore abstrait et l'espace réel de restitution défini par la dernière proposition. Il est possible d'envisager un système de haut-parleurs devant l'auditoire, dans l'espace sonore abstrait ou en dehors. Chaque *soxel* pourrait alors être restitué virtuellement à l'endroit où il devrait apparaître, même si c'est entre deux auditeurs. Le problème majeur d'une telle technique est d'établir le système de haut-parleurs nécessaire à une restitution convenable de l'espace sonore. Cela en fait une approche idéale sur le papier<sup>277</sup>, mais dans la pratique les conditions permettant de la mettre en œuvre sont rares.
- Le DBAP [Lossius & al., 2009] est une technique fondée sur la distance des sources sonores par rapport aux haut-parleurs. De ce fait, elle permet un placement libre des haut-parleurs et de même, un placement libre des sources sonores virtuelles dans ce dispositif de haut-parleurs. Aussi, cela en fait une approche particulièrement adaptée à l'approche matricielle de l'espace. Et de plus, qui est en adéquation avec la proposition privilégiée où l'espace sonore abstrait et l'espace réel de restitution se superposent, et où la distance est finie. Enfin, l'auditoire peut *a priori* recouvrir la même étendue que le dispositif de restitution, faisant ainsi du DBAP une approche relativement flexible. Néanmoins, l'approche reste assez floue sur la façon dont l'auditoire perçoit une source en un point donné, et les travaux critiques [Kostadinov & al, 2010] sur le sujet ne semblent pas encore suffisants pour définir la pertinence de l'approche sur un plan perceptif. En outre, l'atout majeur de cette approche réside essentiellement sur sa capacité d'adaptation à des systèmes de restitution variés<sup>278</sup>.

Compte tenu des différents avantages et inconvénients de chaque technique de spatialisation du son, deux solutions sont suffisamment en adéquation avec la représentation matricielle de

---

<sup>277</sup> Un autre avantage de la combinaison de cette technique avec la représentation matricielle de l'espace est que celle-ci offre des sources sonores pour la WFS qui sont par défaut statiques. Par ce principe, d'une part, la grande demande en ressource du processeur nécessaire aux calculs de la spatialisation par la WFS est restreinte et d'autre part, la limitation du système où une source sonore ne peut croiser la ligne de haut-parleurs n'est jamais rencontrée.

<sup>278</sup> L'approche offre aussi une mise en œuvre relativement simple à réaliser, ou du moins peu coûteuse en ressources de calcul du processeur.



l'espace sonore, dans les choix conceptuels définis précédemment : la WFS et le DBAP. Bien qu'en principe la WFS semble plus apte à offrir une restitution cohérente de chaque *soxel* dans l'espace, le système de haut-parleurs nécessaire à l'approche rend la pratique trop peu généralisable <sup>279</sup> pour que sa mise en œuvre soit réellement pertinente dans ce contexte. Aussi, le choix a été fait de mettre en œuvre un module de projection des *soxels* élaboré selon le DBAP. Enfin, il est nécessaire de rappeler que d'autres approches telles que le ViMiC [Braasch, 2005] aurait pu être intéressante à critiquer, mais le cadre de ces recherches il n'a pas encore été permis d'envisager d'autres hypothèses.

### 8.1.3. Mise en œuvre

Le principe de projection des *soxels* consiste à considérer chaque *soxel* comme une source sonore, et d'utiliser le DBAP pour définir comment cette source sonore est distribuée sur un ensemble de haut-parleurs données. Pour cela, il s'agit dans un premier temps de définir des coordonnées pour chaque *soxel*. Afin d'offrir un système qui puisse s'adapter à n'importe quelle taille matricielle, il est nécessaire que les coordonnées soient mises sur un espace aux dimensions fixes. L'enjeu est ici de préserver une cohérence entre les espaces sonores de tailles différentes, de manière similaire à l'adaptabilité des opérations en ambisonie, selon différents ordres de décomposition en harmoniques sphériques. L'objectif est que le *soxel* central sur un espace sonore matriciel de taille 3x3, corresponde dans l'espace réel à la même position que le *soxel* central sur un espace sonore matriciel de taille 5x5 ou 7x7 etc.

$0_{0,0} = [-1, 1]$	$1_{1,0} = [0, 1]$	$2_{2,0} = [1, 1]$
$3_{0,1} = [-1, 0]$	$4_{1,1} = [0, 0]$	$5_{2,1} = [1, 0]$
$6_{0,2} = [-1, -1]$	$7_{1,2} = [0, -1]$	$8_{2,2} = [1, -1]$

Tableau 8.1. Coordonnées cartésiennes des *soxels* pour un espace matriciel de 3X3.

Afin de faciliter l'approche et de coïncider avec les représentations généralement admises dans le domaine de la spatialisation du son, il a été défini de manière arbitraire que le système de

---

<sup>279</sup> De plus cette approche ne peut pas être testée dans le studio du CICM à la MSH Paris Nord.

coordonnées est cartésien, et que l'espace matriciel définit les limites telles que la première ligne soit associée à l'ordonnée -1, la dernière ligne à l'ordonnée 1, la première colonne à l'abscisse -1 et la dernière colonne à l'abscisse 1 [Tableau 8.1]. Il s'agit alors de définir à l'intérieur de ce système les coordonnées de chaque *soxel* telles que

$$x_s = -\frac{2\left(m - \frac{M}{2} + 0.5\right)}{M - 1} \text{ et } y_l = -\frac{2\left(l - \frac{L}{2} + 0.5\right)}{L - 1}.$$

Équation 8.1. Coordonnées cartésiennes des *soxels* selon leurs indices.

Cette approche assume que les coordonnées des haut-parleurs sont données relativement à cet espace défini par les *soxels*. Cependant, comme suggéré précédemment, rien n'établit que les haut-parleurs doivent nécessairement se trouver dans l'espace. L'enjeu sera, lors de l'utilisation de trouver l'échelle la plus optimale pour une restitution cohérente de l'espace.

Une fois, les coordonnées des *soxels* définies, il s'agit de définir les gains associés à chaque haut-parleur. Cette opération est réalisée dans la bibliothèque *xel*, par l'abstraction *xel.2d.proj.dbap*. Cette abstraction accepte en argument le nom d'un objet *text* qui est utilisé, de manière similaire à l'abstraction de projection proposée en ambisonie, pour définir les positions des haut-parleurs en coordonnées cartésiennes dans une ligne commençant par le symbole *positions*. À ce titre, deux autres paramètres peuvent être définis. Le flou spatial permet de faire que lorsqu'un *soxel* se trouve à la même position qu'un haut-parleur, celui-ci ne soit pas le seul à contribuer à sa restitution<sup>280</sup>. La pente d'extinction du son en décibels sur le double de la distance, permet de contrôler de manière inverse la manière dont une source sonore se répand dans l'espace. Une pente d'extinction faible peut être associée à une source sonore très répandue, et une pente d'extinction forte associe la source sonore à un point. En envoyant des coordonnées à l'abstraction, celle-ci renvoie une liste de gains associés aux haut-parleurs. Ces gains peuvent alors être utilisés pour générer les signaux associés aux haut-parleurs pour chaque *soxel*.

---

<sup>280</sup> Pour cela il faut définir une ligne dans l'objet *text* commençant par le terme *blur*. Ce paramètre est initialement présent dans le DBAP. Mais il ne semble pas réellement pertinent dans cette approche de favoriser le flou spatial. Aussi, une valeur relativement faible de l'ordre de 0,01 préserve mieux la restitution. Néanmoins, dans certains cas, ce paramètre permet de « lisser » les variations spatiales du son.

La projection d'un *soxel* pour un ensemble de haut-parleurs est réalisée par l'abstraction *xel.2d.proj.dbap~* qui, pour chaque gain associé à un haut-parleur, instancie un sous patch dynamiquement, afin de générer la contribution du *soxel* pour ce haut-parleur [Figure 8.5]. Sur le plan de la mise en œuvre, l'approche est alors similaire à celle présentée pour la projection des signaux associés aux harmoniques sphériques vers un système de haut-parleurs. En utilisant cette abstraction dans l'objet *spam~*, cela permet de générer pour tous les haut-parleurs les contributions de l'ensemble de l'espace sonore. Comme toutes les abstractions de la bibliothèque, l'abstraction *xel.2d.proj.dbap~* prend comme argument la taille de la matrice, le nombre de lignes et le nombre de colonnes, mais elle nécessite aussi un nombre de haut-parleurs comme troisième argument et le nom de l'objet *text* décrivant les caractéristiques de l'utilisation du DBAP.

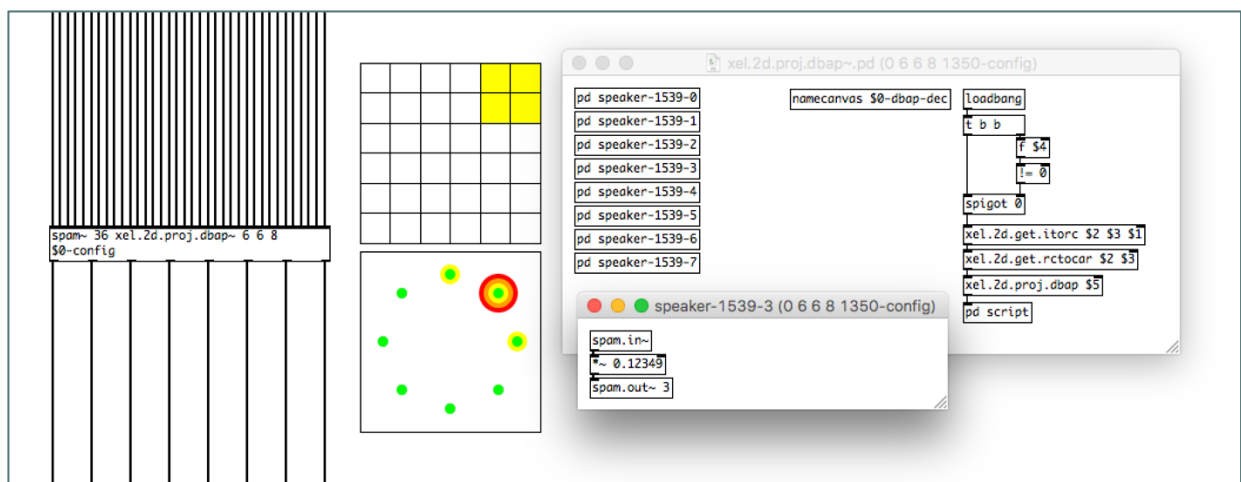


Figure 8.5. Projection d'un espace sonore matriciel de taille 6x6 vers un système de huit haut-parleurs circulaires avec l'abstraction *xel.2d.proj.dbap~* située à gauche. L'espace sonore est représenté au centre en haut, et les niveaux sonores des haut-parleurs sur l'interface graphique au centre, en bas. A droite, l'abstraction réalisant la projection du premier *soxel* pour les haut-parleurs est affichée avec au-dessus le sous-patch réalisant la contribuant associée au quatrième haut-parleur.

Cette première mise en œuvre offre donc la possibilité de s'adapter à n'importe quelle configuration de haut-parleurs, et cela quelle que soit la taille matricielle de l'espace sonore. Elle possède de plus l'avantage de ne réaliser qu'une succession d'opérations arithmétiques simples d'additions et de multiplications, permettant d'être relativement peu demandeuse en ressource de calculs du processeur. Laissant ainsi envisager une bonne utilisation dans un contexte de musique temps-réel<sup>281</sup>. Enfin, les recherches n'ont pas pu pour le moment mettre

<sup>281</sup> La WFS étant mise en œuvre via l'utilisation de lignes à retard, devrait être bien plus couteuse en termes d'utilisation du processeur.

en œuvre d'autres approches pour la projection de l'espace sonore matriciel. Aussi, il est impossible de critiquer la technique du VBAP par la comparaison. De même, les recherches restant axées principalement pour le moment sur des mises en œuvre d'opérations originales de l'espace, le temps n'a pas été pris pour critiquer la restitution avec des outils d'analyses tels que proposés en ambisonie. Aussi ces deux approches seraient intéressantes à développer dans de futurs travaux. Néanmoins, selon un point de vue subjectif, le rendu spatial et sonore semble cohérent avec les espaces sonores abstraits. Il est donc d'ores et déjà possible d'explorer de nouvelles synthèses d'espaces sonores matriciels et de nouvelles transformations, construites autour de cette représentation matricielle de l'espace.

## 8.2. Filtres spatiaux

Une des opérations les plus évidentes offerte par la représentation matricielle de l'espace sonore est le filtrage spatial. Cette opération pourrait être considérée pour cette représentation de l'espace, comme l'équivalent de ce qu'est la rotation pour la représentation de l'espace via un ensemble d'harmoniques sphériques. En effet, les harmoniques sphériques sont particulièrement utiles pour résoudre des problèmes invariants en rotation, ce qui amène implicitement à cette pratique. Comme cela sera présenté, la représentation matricielle amène à l'utilisation de filtres qui, dans le domaine de l'espace sonore, peuvent être considérés comme un filtre spatial. Cette sous-partie propose de définir ce qu'est un filtre en traitement de l'espace sonore, sous sa représentation matricielle.

Un filtre est un outil relativement difficile à définir, car sa fonction peut changer selon les domaines et selon ses utilisations. En traitement du signal numérique, l'application d'un filtre est une opération très répandue, utilisée pour faire varier certaines parties du spectre sonore du signal. Cela permet par exemple, d'égaliser les énergies sur un plan fréquentiel, afin de faire ressortir ou diminuer certaines caractéristiques sonores ou encore de simuler des phénomènes acoustiques tels que la réverbération, lors de l'utilisation d'un système complexe de filtres. En traitement d'images, l'application d'un filtre est, de manière analogue au traitement du signal numérique, une opération permettant de transformer les informations contenues dans une image, offrant de nombreux traitements tels que le flou, l'augmentation des contrastes, le renforcement et la détection des contours, la réduction du bruit, etc. De manière générale en traitement du signal, un filtre est un système transformant et modifiant le contenu d'un signal [Figure 8.6]. Mais les mises en œuvre et les principes restent encore très différents

selon les contextes et les domaines d'applications. Il est donc intéressant dans ce cadre, d'étudier ce qu'est un filtre de l'espace sonore matriciel<sup>282</sup> et quelles sont ses possibilités.



Figure 8.6. Représentation la plus générale de l'opération d'un filtre en traitement du signal.

Dans un premier temps, les principes mathématiques sur lesquels cette opération repose et ses mises en œuvre dans les domaines du traitement du signal numérique et du traitement d'images sont présentées, afin de mettre en évidence le fonctionnement et les implications qu'un tel procédé puisse avoir. En conséquence, une première mise en œuvre d'un système permettant d'utiliser des filtres dans le domaine de l'espace sonore matriciel est proposée, afin de pouvoir réaliser de premières expérimentations et de soulever les premiers enjeux techniques et fonctionnels. Par la suite, différents filtres inspirés notamment du traitement d'images sont mis en œuvre dans le domaine de l'espace sonore matriciel. Puis leurs fonctionnements et leurs résultats sont analysés et comparés, notamment par rapport à leur dessein d'origine. Enfin, une première analyse de l'approche est réalisée. Les différents problèmes et restrictions de la mise en œuvre actuelle sont présentés et des perspectives de développement sont envisagées.

### 8.2.1. Produit de convolution

Comme évoqué précédemment, les filtres peuvent être utilisés à des fins et des objectifs différents et variés. Aussi, leurs mises en œuvre peuvent être réalisées de multiples manières. Les filtres peuvent être utilisés par exemple dans des systèmes à une dimension, ou dans des systèmes à plusieurs dimensions et possèdent des réponses impulsionnelles finies ou infinies. Sans recouvrir l'ensemble du sujet, il s'agit de replacer celui-ci dans le contexte de ces recherches, en présentant les deux domaines centraux que sont le son et les matrices. De plus, les présentations se limitent aux filtres à réponse impulsionnelle finie, où le résultat provient uniquement des valeurs du signal en entrée et non sur des valeurs résultantes antérieures.

En traitement du signal numérique, les filtres opèrent sur des signaux à une dimension, en l'occurrence le temps. Et les effets de l'application du filtre sont naturellement associés au

---

<sup>282</sup> Le filtre de l'espace présenté auparavant en ambisonie sur les ondes planes est encore différent.

domaine temporel, car ils agissent sur des phénomènes périodiques. Ce filtre peut être représenté mathématiquement par le produit  $y$  de convolution d'un signal audionumérique  $x$ , avec une réponse impulsionnelle  $h$  de taille  $N$ , qui revient pour tout instant  $n$  à calculer la somme des  $N$  derniers échantillons pondérés par les valeurs de la réponse impulsionnelle  $h$  tel que :

$$y_n = \sum_{i=0}^N x_{n-i} h_i = x_n h_0 + x_{n-1} h_1 + x_{n-2} h_2 + \dots + x_{n-N+1} h_{N-1}.$$

Équation 8.2. Produit de convolution sur un signal à une dimension.

Néanmoins, il est possible d'appliquer des filtres sur des signaux n'étant pas forcément temporels, comme c'est le cas en traitement d'images. Les filtres n'agissent plus en fonction du temps, mais en fonction des indices des *pixels*. En traitement d'images, le filtre est, de manière analogue au traitement du signal audionumérique, une opération qui consiste à appliquer à une image une réponse impulsionnelle aussi appelée noyau ou *kernel* en anglais. En fonction du noyau utilisé, cette opération de filtrage permet d'obtenir un grand nombre de transformations telles que le flou, l'augmentation des contrastes, le renforcement et la détection des bords, ou encore l'élargissement. En étendant le produit de la convolution à deux dimensions, cette opération revient pour chaque élément d'indices  $i$  et  $j$  d'une matrice  $x$ , à calculer la somme des valeurs de tous les autres éléments de la matrice, pondérées par les valeurs correspondantes dans la réponse impulsionnelle  $h$  de taille  $L \times M$  tel que

$$y_i^j = \sum_{l=-L}^L \sum_{m=-M}^M x_{i+l}^{j+m} h_l^m.$$

Ou encore

$$y_i^j = x_{i-L}^{j-M} h_{-L}^{-M} + x_{i-L+1}^{j-M+1} h_{-L+1}^{-M+1} + \dots + x_i^j h_0^0 + \dots + x_{i+L-1}^{j+M-1} h_{L-1}^{M-1} + x_{i+L}^{j+M} h_L^M.$$

Équation 8.3. Produit de convolution sur un signal à deux dimensions.

Cette opération qui peut sembler *a priori* abstraite, revient à calculer la valeur de chaque élément de la matrice, en réalisant une moyenne des valeurs de l'ensemble des éléments adjacents pondérées selon des coefficients définis par le noyau.

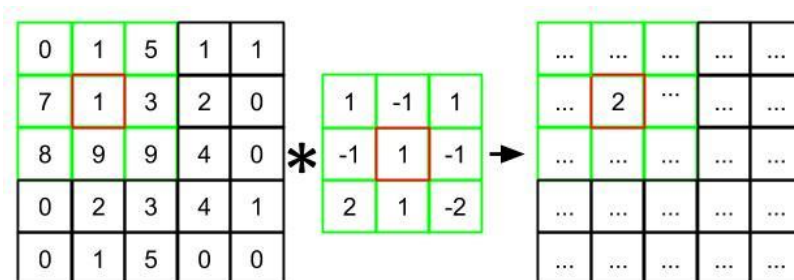


Figure 8.7. Représentation du calcul d'un élément d'indices [1][1] lors du calcul du produit  $y$  de la convolution à droite de la matrice  $x$  à gauche avec le noyau  $h$  au centre. Cette opération peut être décomposée par  $y_{1,1} = x_{0,0}h_{0,0} + x_{0,1}h_{0,1} + x_{0,2}h_{0,2} + x_{1,0}h_{1,0} + x_{1,1}h_{1,1} + x_{1,2}h_{1,2} + x_{2,0}h_{2,0} + x_{2,1}h_{2,1} + x_{2,2}h_{2,2}$ .

Sur un plan mathématique, et en traitement du signal de manière générale, les opérations sont transposables quel que soit le nombre de dimensions. Et à chaque type de réponse impulsionnelle à une dimension, il existe un équivalent dans le système à deux dimensions. Tout comme il est possible de définir les coefficients d'une réponse impulsionnelle audionumérique pour réaliser des filtres passe-bas, passe-haut, etc., en traitement de l'image, il est possible de réaliser des noyaux offrant des résultats similaires sur un plan mathématique. Cependant, le changement de domaine amène des résultats différents d'un point de vue perceptif. L'enjeu est alors de comprendre à quoi correspond sur un plan graphique, une transformation bien définie sur un plan audionumérique. Et il s'agit par la suite de transposer les résultats au domaine de l'espace sonore, ce qui vraisemblablement devrait impliquer des modifications et des adaptations de l'approche.

### 8.2.2. Mise en œuvre

L'abstraction permettant d'appliquer un noyau par convolution est `xel.2d.fx.kernel~` qui ne nécessite aucun argument, si ce n'est la taille matricielle de l'espace sonore. Le défi à relever afin de réaliser la mise sous forme d'abstraction pour l'objet `spam~` est de faire que le signal associé à un `soxel` soit la combinaison des signaux de plusieurs `soxels`, tels que représentés dans la [Figure 8.7]. Cela afin de pouvoir réaliser une opération mathématique similaire à l'application d'un filtre en traitement d'image.

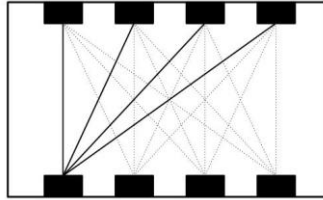


Figure 8.8. Représentation de l'association des entrées et sorties de l'objet *spam~* par le patch *xel.2d.fx.kernel~* pour réaliser un filtre de convolution.

Ce problème a déjà été soulevé lors la mise en œuvre d'une abstraction de projection, où les signaux résultants destinés à un système de haut-parleurs, sont l'association de l'ensemble des signaux associés aux composantes spatiales. La solution consiste à générer dynamiquement des séries de sous-patches dans chaque instance, ce qui permet de gérer la contribution de cette instance pour l'ensemble des instances tel que présenté dans la [Figure 8.8]. Sur un plan pratique, une instance génère autant de sous-patches qu'il y a d'instances. Chaque sous-patch reçoit le signal d'entrée de l'instance dans laquelle il a été instancié via l'objet *spam.in~* et multiplie ce signal selon la valeur de pondération reçue via un objet *receive*, spécifique à la fois à l'instance globale et à l'instance du sous-patch. Il renvoie le signal à la sortie associée à l'instance du sous-patch via l'objet *spam.out~* [Figure 8.9]. Selon ce principe, chaque instance génère des signaux pour toutes les sorties et c'est l'objet *spam~* qui accumule pour chacune des sorties, les signaux respectifs qui leurs sont associés.

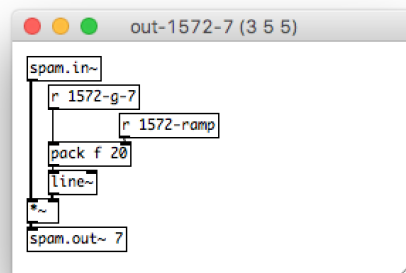


Figure 8.9. Sous-patch de la 3<sup>ème</sup> instance de l'abstraction *xel.2d.fx.kernel~* associé à la 7<sup>ème</sup> sortie équivalent à générer la contribution du *soxel* de ligne 0 et colonne 3 pour le *soxel* de ligne 1 et de colonne 2.

L'abstraction possède une entrée supplémentaire de contrôle, permettant de recevoir notamment le noyau de convolution. Lorsque l'abstraction reçoit un noyau, dans un premier temps sa taille est comparée au nombre de coefficients qu'il possède. Et elle vérifie que le nombre de lignes et le nombre de colonnes sont bien impairs. Dans le cas contraire, les calculs



sont évités et un message d'erreur est envoyé à l'utilisateur<sup>283</sup>. Les opérations consistent alors simplement à retrouver, en fonction de l'indice de l'instance courante, à quelle instance de sous-patch un coefficient doit être envoyé. En acceptant que l'élément central du noyau correspond à l'instance courante d'indices  $l_i$  et  $m_i$ , les indices matriciels  $l_c$  et  $m_c$  sont définis selon la taille matricielle  $L_n \times M_n$  du noyau et les indices matriciels  $l_n$  et  $m_n$  du coefficient au sein du noyau par :

$$l_c = l_n - \frac{L_n - 1}{2} + l_i \text{ et } m_c = m_n - \frac{M_n - 1}{2} + m_i.$$

Équation 8.4. Calculs des indices dans la matrice d'origine d'un élément associé à un coefficient du noyau.

Il s'agit alors de définir les indices linéaires à partir des indices matriciels, [Tableau 7.1] afin d'envoyer les coefficients aux instances de l'objet *receive* correspondantes. Forcément, selon la taille du noyau, une partie des sorties sont associées à des instances qui ne sont pas prises en compte par le noyau et dans ce cas leurs coefficients sont remis à zéro. L'opération revient implicitement à définir la taille du noyau à deux fois la taille de la matrice d'origine, par complétion de zéro (*zero-padding*).

Plusieurs aspects de la mise en œuvre sont intéressants. Comme représentés dans la [Figure 8.9], les valeurs de pondérations reçues en contrôle sont lissées et transformées en signal audionumérique via l'objet *line~* dont le temps pour glisser à la nouvelle valeur peut être défini, de manière globale, à l'ensemble des instances. Cette fonctionnalité permet donc de passer d'un état linéaire à un autre état dans un temps donné, comme par exemple passer linéairement en quelques secondes, d'un espace sonore intact à un espace sonore subissant un flou<sup>284</sup>. De plus, le fait que chaque *soxel* contribue à l'ensemble des *soxels* résultant, revient à pouvoir utiliser indifféremment n'importe quelle taille de noyau<sup>285</sup>. Et le fait que les coefficients des instances n'étant pas pris en compte soient remis à zéro permet de varier la taille des noyaux et de réaliser des transitions.

---

<sup>283</sup> Le fait que les tailles du noyau doivent être impaires est une condition impérative pour avoir un élément central dans le noyau et donc pouvoir associer cet élément aux éléments de la matrice originale.

<sup>284</sup> Cette opération est présentée par la suite.

<sup>285</sup> Tant que le nombre de lignes et le nombre de colonnes reste impair.

Enfin, un élément important de la convolution qui n'est pas directement pris en compte dans la mise en œuvre actuelle, est la gestion des bordures. Quand le *soxel* initial est à moins de la moitié de la taille du noyau par rapport au bord de l'espace sonore, une partie du noyau dépasse les limites de l'espace sonore et n'est donc pas pris en compte dans le résultat du *soxel* final. Dans la pratique, la convolution d'une matrice par un noyau nécessite d'agrandir la matrice initiale de la moitié de la taille du noyau<sup>286</sup>. Plusieurs procédés existent pour contrecarrer le manque d'informations, mais aucune approche n'est idéale et leurs utilisations vont dépendre majoritairement du type de traitement appliqué, et donc du type de noyau utilisé. Les procédés les plus courants sont définis par [Hughes & al., 2013], p. 589, et comportent : l'extension par zéro (qui revient à l'approche implémentée), l'extension par réflexion, l'extension par des constantes, la limitation de la taille de la matrice reconstruite ou l'ajustement du filtre près des bords. Une amélioration notable et même essentielle de l'abstraction serait donc de pouvoir offrir certaines de ces fonctionnalités supplémentaires. Néanmoins, par cette abstraction, il est déjà possible de réaliser un grand nombre de filtrages spatiaux.

### 8.2.3. Noyaux et opérations

Une infinité de noyaux générant une infinité de traitements de l'espace sonore sont envisageables. Il est donc nécessaire de mettre en lumière la conception et l'utilisation de certains principes, afin de mieux envisager les réelles possibilités et implications liées à leurs manipulations. Pour cela, une fois n'est pas coutume, il est possible de construire les approches sur des outils et pratiques déjà établies en traitement d'images. Les présentations actuelles se limitent à un nombre restreint de procédés, à savoir le flou et l'accentuation de l'image, car ils sont très bien exposés dans la littérature sur un plan formel. De plus, ils sont suffisamment répandus dans les usages pour en définir clairement les tenants et aboutissants. L'objectif est donc de mettre en avant le fonctionnement des filtres sur un plan mathématique et dans le domaine de l'image, pour envisager leurs résultats et leurs adaptations dans l'espace sonore. Ce qui permet à terme de concevoir et mettre en œuvre d'autres procédés. Cependant, pour une utilisation directe des filtres, de nombreux noyaux sont relativement simples à réaliser pour des tailles fixes de noyaux. De plus, ces filtres permettent d'expérimenter rapidement de nombreuses opérations dans le domaine de la spatialisation du son tels que le repoussage, le renforcement des bords, la détection des bords, etc.

---

<sup>286</sup> En réalité comme la taille du noyé est impaire, c'est la moitié de la taille moins un.

## Le Filtre de Boîte

Le filtre de boîte ou *box filter* en anglais [Shapiro & al., 1995], [Hugh & al., 2013], est une technique où chaque composante est le résultat d'une combinaison linéaire d'elle-même et de ses composantes voisines.

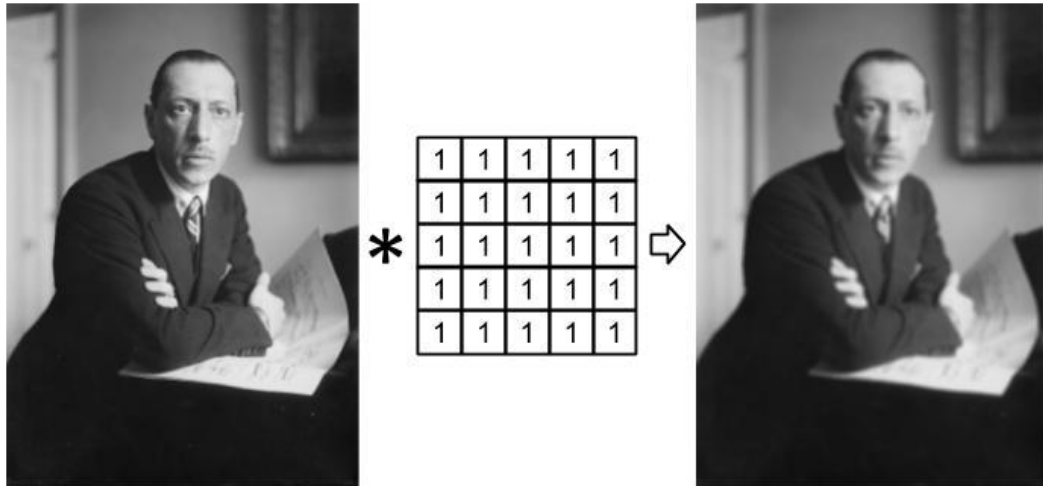


Figure 8.10. Représentation de l'application d'un filtre de boîte au centre, sur une image à gauche et son résultat à droite.

L'approche présentée dans la [Figure 8.10] revient à calculer la moyenne de la valeur de gris de chaque *pixel* de l'image avec la valeur de gris de leurs vingt-quatre *pixels* avoisinants, en pondérant le résultat par le nombre de *pixels* pris en compte au total, afin d'éviter la saturation tel que :

$$y[i][j] = \frac{\sum_{l=-2}^{+2} \sum_{m=-2}^{+2} x[i+l][j+m]}{25}.$$

Équation 8.5. Filtre de boîte réalisé par un noyau de taille 5x5.

Ce type de noyau correspond à une fonction porte et revient à réaliser un filtre passe-bas dont la transformée de Fourier est un sinus cardinal [Figure 8.15], correspondant en traitement d'image à un flou. L'opération amoindrit les variations importantes de valeurs<sup>287</sup> entre les

---

<sup>287</sup> Le terme de valeur est utilisé de manière quelque peu abstraite mais il peut définir un niveau de gris pour des images en noir ou se référer à l'ensemble des couches de couleurs primaires de la synthèse additive.

*pixels* : si un *pixel* est noir et qu'un *pixel* adjacent est blanc, ces deux *pixels* deviennent d'un gris intermédiaire dont le niveau dépend de la taille du noyau.

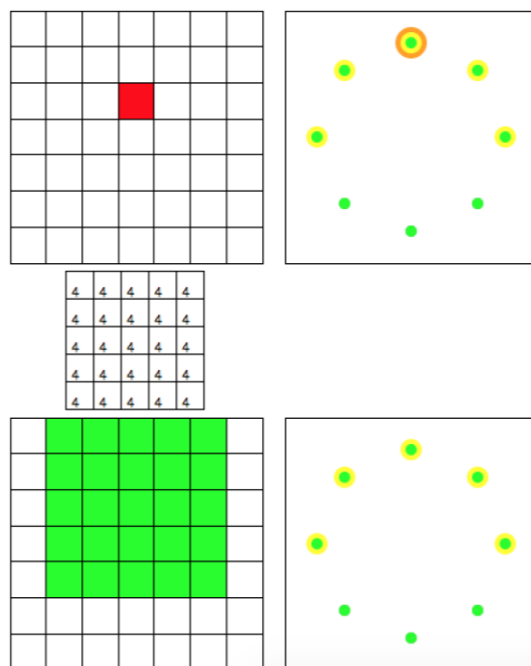


Figure 8.11. Interfaces graphiques d'un patch Pure Data réalisant une convolution d'un espace sonore de taille 7x7 où seul le *soxel* [2][3] contribue fortement (en haut à gauche) avec un noyau de taille 5x5 dont les coefficients sont tous égaux à 0.04 (au centre à gauche) pour donner un espace sonore où les *soxels* compris recouvrant l'espace d'indices allant de [0][1] à [4][5] contribuent faiblement et uniformément (en bas à gauche). Les contributions d'un système de huit haut-parleurs placés en cercle, sont représentées suite à la projection de l'espace avant la convolution (en haut à droite) et après la convolution (en bas à droite).

L'abstraction *xel.2d.mat.blur.box* permet de générer automatiquement un noyau de convolution pour le filtre de boîte dont les valeurs sont égales et normalisées en fonction de la taille de celle-ci. Lorsque ce noyau est utilisé dans le domaine spatial, il en résulte que chaque signal associé à un *soxel* se retrouve réparti de manière homogène, sur un ensemble de signaux associés aux *soxels* avoisinants selon la taille du noyau, tel que représenté sur la [Figure 8.11]. Evidemment, la géométrie visuelle n'est pas identique à la géométrie auditive, ainsi le noyau a beau être abrupt sur un plan auditif, cette opération ne donne pas à entendre un large carré de son uniforme et clairement défini dans l'espace. De même que la contribution d'un seul *soxel* ne laisse pas entendre une source sonore ponctuelle exactement en ce point<sup>288</sup>. Il est important de prendre en considération l'opération de projection des *soxels* pour un système de restitution sonore, afin de concevoir réellement les caractéristiques de cette opération. Dans l'espace

<sup>288</sup> Sauf évidemment si un haut-parleur se trouve placé à l'exact centre du *soxel* dans l'espace réel et que l'opération de projection par DBAP possède un paramètre de flou négligeable et une pente relativement forte.

virtuel, les contributions plus ou moins importantes des *soxels* se retrouvent étalées, diminuant alors les particularités de l'espace au profit d'une homogénéité. Et dans l'espace réel, les haut-parleurs contribuent de manière plus similaire, les signaux sont plus équilibrés et tendant à un espace sonore omnidirectionnel et étendu. L'effet permet notamment, dans le cas de sources sonores ponctuelles, d'élargir leurs diffusions.

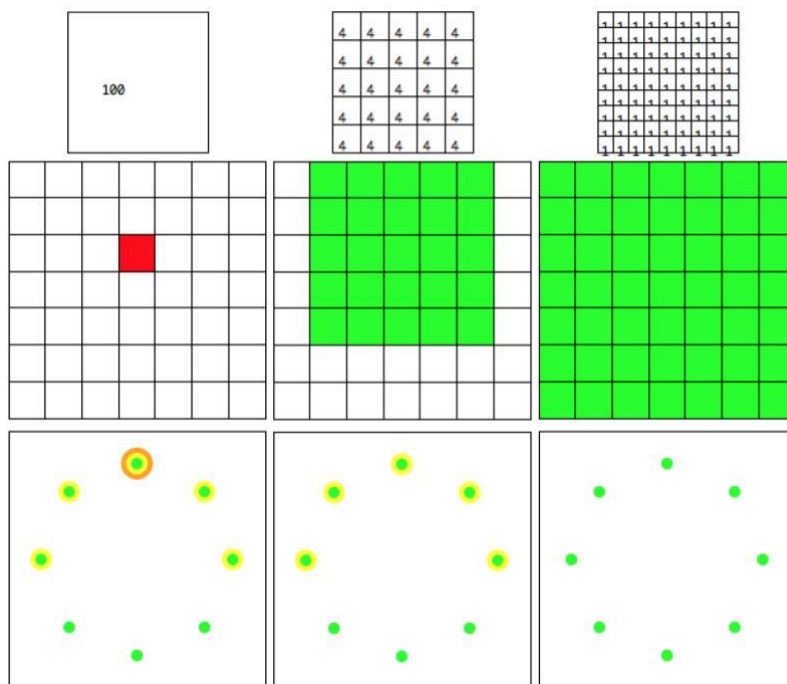


Figure 8.12. Interfaces graphiques d'un patch Pure Data représentant différents état d'un espace sonore (figures de la ligne du milieu) et les contributions d'un système de haut-parleurs circulaires après projection (figures de la ligne du bas), suite à l'utilisation d'un filtre de boîte pour différentes tailles de noyau (figures de la ligne d'en haut). Les tailles de noyau utilisées sont de gauche à droite 1x1, 5x5 et 9x9.

Bien que l'approche soit relativement simple, elle reste efficace sur un plan perceptif, car les variations de contributions des haut-parleurs sont évidentes<sup>289</sup>. L'un des avantages réside aussi en la possibilité d'interpoler entre des noyaux de tailles différentes [Figure 8.12], ce qui permet d'éviter des changements abrupts de rayonnement, notamment en utilisant une valeur de temps d'interpolation suffisamment grande pour glisser d'un noyau à un autre de taille plus grande. De plus, la variation des gains des sources sonores est compensée par le fait que les gains associés aux *soxels* contribuant à une source sonore diminuent alors que de nouvelles lignes et de nouvelles colonnes de *soxels* contribuent à l'espace sonore. Un autre aspect intéressant est la possibilité de générer des noyaux rectangulaires, afin de créer des flous plus ou moins

<sup>289</sup> Du moins pour le système de huit haut-parleurs circulaires du CICM à la Maison des Sciences de l'Homme Paris-Nord, sur lequel le traitement a été testé.

proéminents suivant l'axe des abscisses ou l'axe des ordonnées. Il est aussi possible d'envisager des flous orientés dans des directions intermédiaires, amenant la nécessité de réaliser un noyau de flou dont les valeurs ne sont pas identiques.

### Le Filtre Gaussien

Le filtre Gaussien ou *Gaussian filter* en anglais [Shapiro & al., 1995], [Hugh & al., 20013] est une technique où chaque composante est le résultat d'une combinaison pondérée d'elle-même et de ses composantes voisines, selon la densité de probabilité de la loi normale [Laplace, 1814], [Fuchs, 1995].

4	12	4
12	32	12
4	12	4

2	10	2
10	48	10
2	10	2

Figure 8.13. Exemples de noyaux de taille 3x3 permettant de réaliser des filtres Gaussien avec des écarts types de 0.72 (à gauche) et 0.57 à droite.

Ce filtre peut être envisagé comme une version améliorée du filtre de boîte. En traitement d'images, l'opération revient pour chaque *pixel* de l'image à calculer une moyenne de la valeur de ce *pixel* et des valeurs de ses *pixels* avoisinants. Cette approche diffère par la pondération des valeurs, selon des coefficients qui sont non-uniformes. Ces coefficients sont définis par les noyaux [Figure 8.13] et calculés à partir d'une fonction Gaussienne en deux dimensions, dont la formulation la plus simple possède en paramètre un écart type  $\sigma$  et où l'espérance est ignorée tel que :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Équation 8.6. Fonction Gaussienne en une dimension.

La transformée de Fourier de cette fonction est aussi une fonction Gaussienne, possédant la forme d'une cloche centrée en zéro. Ce qui revient à réaliser un filtre passe-bas correspondant en traitement d'image à un flou [Figure 8.15]. Elle est continue et possède des limites infinies qui tendent vers zéro, c'est à dire que lorsque  $x$  ou  $y$  s'éloignent de l'origine, les valeurs

diminuent [Figure 8.14]. Appliquer une courbe Gaussienne à un noyau, implique que l'élément central possède la valeur la plus importante et que les valeurs diminuent lorsque les éléments se rapprochent des limites. Plus l'écart type diminue, plus le lobe central de la courbe devient proéminent, et inversement plus l'écart type augmente plus le lobe s'élargit. De même, appliqué à un noyau, l'écart type permet d'augmenter ou de diminuer la différence entre la valeur centrale et les valeurs circulaires. Plus l'écart type diminue et se rapproche de zéro, alors plus l'élément central contribue de manière importante comparée aux éléments adjacents. Et le noyau se rapproche alors d'un noyau unitaire de taille  $1 \times 1$  qui équivaut à un Dirac, et son application préserve l'image originale. Réciproquement, plus l'écart type augmente et tend vers l'infini, plus les éléments contribuent de manière homogène. Le noyau se rapproche alors d'un noyau d'un filtre de boîte, et son application revient à réaliser un flou sur l'image.

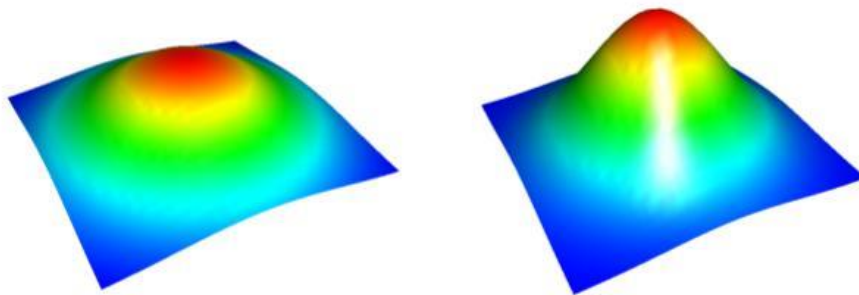


Figure 8.14. Représentation 3D de la fonction Gaussienne en deux dimensions [Équation 8.6] pour un écart type de 0.72 (à gauche) et un écart type de 0.57 (à droite) dont les limites sont définies entre -1 et 1 sur les axes des abscisses et des ordonnées et entre 0 et 1 sur l'axe de la hauteur.

L'abstraction `xel.2d.mat.blur.gaussian` permet de générer automatiquement des noyaux de convolution pour un filtre Gaussien, selon une taille donnée. L'enjeu de cette mise en œuvre est de faire coïncider un domaine discret et limité, à savoir une matrice, avec un domaine continu et infini, à savoir la fonction gaussienne. En analysant les différents outils offerts par les logiciels tels que GIMP ou Photoshop, ainsi que les matrices proposées dans la littérature [Shapiro & al., 1995] [Hugh & al., 2013], il a été remarqué que les limites sur lesquelles la fonction est considérée ne sont jamais utilisées en tant que paramètres pour générer le noyau, mais qu'elles sont fixes. La raison implicite semble être que, lors de l'application de la fonction gaussienne à un noyau, augmenter ses limites revient à resserrer le lobe au centre du noyau. Ce qui correspond à une diminution de l'écart type, à un facteur de gain près. Et la relation réciproque est aussi valable. Etant donné que les valeurs sont normalisées, utiliser les limites sur lesquelles la fonction est considérée revient donc à utiliser un paramètre caduc, qui peut être

remplacé par l'écart type. Dans la mise en œuvre de l'abstraction, les limites sont donc définies de manière arbitraire entre  $-1$  et  $1$ , sur les axes des abscisses et des ordonnées. Le calcul de la valeur de chaque élément du noyau est réalisé en utilisant des coordonnées virtuelles des éléments transposés dans cet espace cartésien, et qui sont définis selon la même méthode que celle utilisée pour la projection des *soxels* [Équation 8.1].

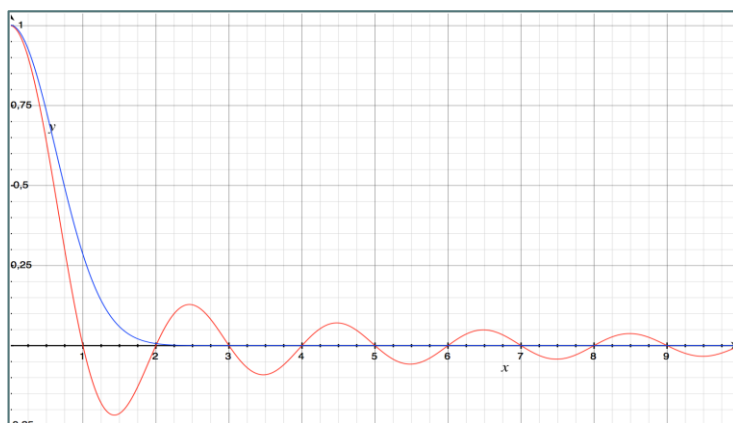


Figure 8.15. Représentation d'une fonction sinus cardinal en rouge et d'une fonction Gaussienne en bleu correspondant respectivement aux transformées de Fourier des fonctions porte et Gaussienne.

Appliqués au domaine de l'espace sonore, les principes sont relativement similaires à ceux d'un filtre de boîte. Chaque signal associé à un *soxel* est redistribué sur un ensemble de signaux associés aux *soxels* avoisinants, mais selon la géométrie de la distribution normale. Cependant, sur le plan auditif, notamment par l'opération de projection, le rendu de cette opération est très proche de celui offert par l'application d'un filtre de boîte. Cette remarque est évidente pour des valeurs de l'écart type très élevées permettant de générer des noyaux similaires à ceux d'un filtre de boîte, mais c'est aussi le cas pour des noyaux *a priori* relativement différents. Dans Computer Graphics, il est d'ailleurs défini qu'un « flou bien meilleur que [...] le filtre de boîte est réalisé en réalisant une convolution avec un filtre Gaussien [...] »<sup>290</sup>. Cependant, comme suggéré précédemment, la géométrie auditive n'est pas identique à la géométrie visuelle. Et à cela il faut ajouter que les tailles matricielles sur lesquelles sont exécutées les opérations ne sont pas non plus identiques. Les fonctions sur lesquelles reposent ces deux filtres, à savoir une fonction porte et une fonction Gaussienne, ont des transformées de Fourier qui sont respectivement une fonction sinus cardinal et une fonction Gaussienne [Figure 8.15]. Dans le domaine des basses fréquences, les caractéristiques de ces fonctions sont donc extrêmement

<sup>290</sup> [Hugh & al., 2013], p. 543, traduit de l'anglais : « A much nicer blur than the one we saw with the box filter comes from convolving with a Gaussian filter, that is, samples of the function. »



semblables et le réel avantage de l'utilisation d'une fonction Gaussienne par rapport à une fonction porte, est de pouvoir éviter les artefacts de celle-ci dans les hautes fréquences. La représentation matricielle de l'espace sonore, qui implique un nombre relativement restreint de *soxel*, correspond à un faible échantillonnage de l'espace. Cela explique alors l'extrême similitude des rendus. En somme, l'approche est similaire et revient à définir la pertinence de varier le type de flou sur une image de seulement quelques *pixels*. Les techniques de flou étant des filtres passe-bas, il est normal de douter du fait que différentes approches puissent offrir des résultats différents et perceptibles, tant que la résolution spatiale reste faible.

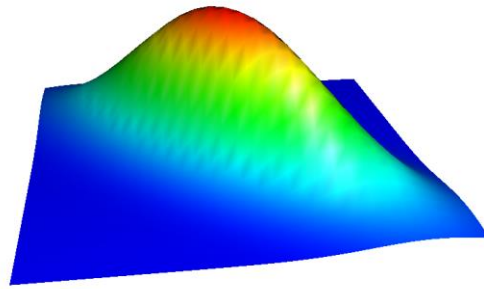


Figure 8.16. Représentation 3D de la fonction Gaussienne en deux dimensions [Équation 8.7] pour un écart type contrôlant la largeur de 0.27 et un écart type contrôlant la longueur de 0.73. Les limites sont définies entre -1 et 1 sur les axes des abscisses et des ordonnées et entre 0 et 1 sur l'axe de la hauteur.

Cette approche selon une fonction Gaussienne offre cependant certains avantages. D'une part, l'utilisation de l'écart type comme paramètre de contrôle du flou de l'espace – ou de la diffusion des sources sonores – est bien plus opérationnel que l'utilisation de matrices de différentes tailles, et peut possiblement s'adapter plus aisément à différentes tailles matricielles. D'autre part – bien que cela n'est pas encore été mis en œuvre – il est possible de reformuler la fonction Gaussienne en deux dimensions, afin d'étirer le lobe dans une direction  $\theta$  donnée [Figure 8.16], et de contrôler indépendamment la largeur du lobe avec l'écart type  $\sigma_w$  et la longueur du lobe avec l'écart type  $\sigma_l$  tel que

$$z(x, y) = e^{-(ax^2 + 2bxy + cy^2)}$$

$$a = \frac{\cos^2\theta}{2\sigma_w^2} + \frac{\sin^2\theta}{2\sigma_l^2}, b = -\frac{\sin(2\theta)}{4\sigma_w^2} + \frac{\sin(2\theta)}{4\sigma_l^2} \text{ et } c = \frac{\sin^2\theta}{2\sigma_w^2} + \frac{\sin^2\theta}{2\sigma_l^2}.$$

Équation 8.7. Fonction Gaussienne en deux dimensions.

Enfin, la fonction Gaussienne possède une dérivée seconde, utilisée pour la détection de contours en images : le filtre LOG [Shapiro & al., 1995].

### Le Filtre LOG

En considérant  $h$  comme le noyau de flou de type Gaussien et  $x$  l'image originale, alors  $x * h$  est une version floutée de  $x$ . En termes plus précis,  $h$  est un filtre passe-bas, et une convolution avec ce noyau enlève les hautes fréquences de l'image  $x$ , il amoindrit les variations importantes entre les valeurs des *pixels*. En traitement du signal audionumérique, une opération courante est de soustraire à un signal original une version filtrée par un filtre passe-bas de ce signal, afin de générer un filtre passe-haut. Il est possible d'envisager la même opération afin de générer un filtre passe-haut dans le domaine de l'image, en soustrayant à l'image originale, une version filtrée d'elle-même telle que  $(1 + \alpha)x - \alpha(h * x)$  où  $\alpha$  peut être considéré comme un taux de soustraction permettant de contrôler l'amplitude du filtre. Et où  $(1 + \alpha)$  permet de compenser l'obscurcissement général de l'image, tel que défini dans [Hugh & al., 2013].

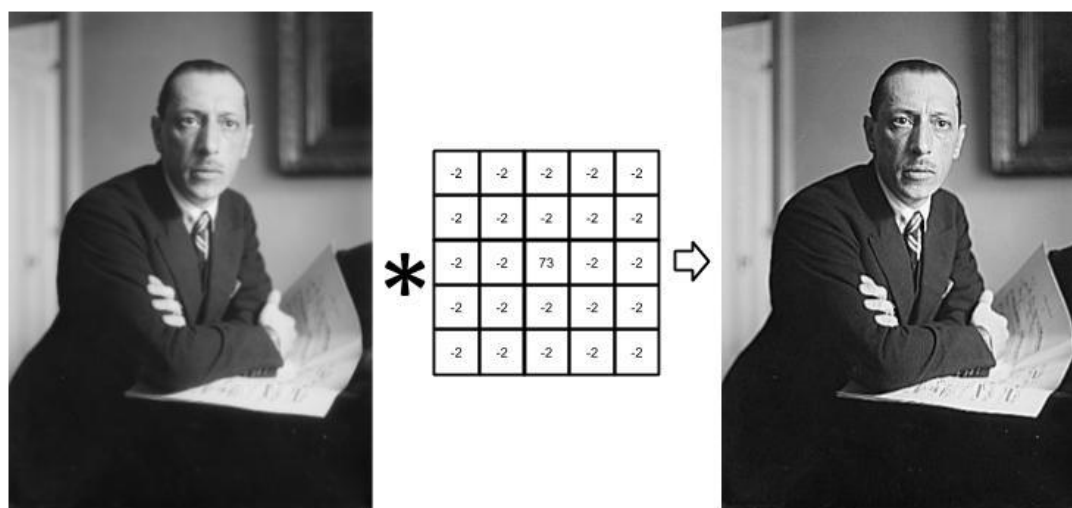


Figure 8.17. Représentation de l'application d'un filtre passe-haut (au centre) sur une image (à gauche) et son résultat (à droite).

En remplaçant  $h$  par un filtre de boîte, l'équation permet de générer un noyau où la valeur centrale est importante et positive, les valeurs annexes quant à elles, sont de faible amplitude et négatives [Figure 8.17]. En considérant trois *pixels* associés à des niveaux de gris qui varient graduellement du gris clair au gris foncé, l'application de ce type de noyau transforme les valeurs de gris des *pixels* de façon à ce que le premier *pixel* devienne blanc, le deuxième *pixel* devienne d'un gris foncé et le dernier *pixel* devienne noir. Ce filtre accentue les variations des valeurs entre les *pixels*, et correspond donc à un filtre passe-haut.

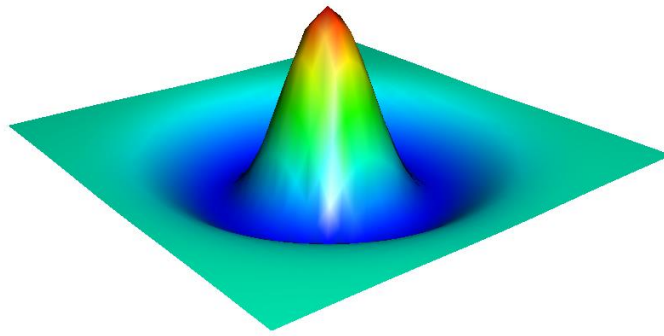


Figure 8.18. Représentation 3D de l'inverse de la dérivée seconde normalisée de la fonction Gaussienne en deux dimensions [Équation 8.8] pour un écart type 0.2. Les limites sont définies entre -1 et 1 sur les axes des abscisses et des ordonnées et entre 0 et 0.5 sur l'axe de la hauteur.

Tout comme le filtre boîte, cette approche du filtre passe-haut peut être en quelque sorte améliorée en utilisant le filtre LOG, pour *Laplacian of Gaussian* en anglais, [Shapiro & al., 1995] dont le tracé en forme de sombrero [Figure 8.18] est calculé à partir de l'inverse de la dérivée seconde d'une fonction Gaussienne tel que

$$f(x, y) = - \left( \frac{x^2 + y^2}{\sqrt{2\pi}\sigma^5} - \frac{1}{\sqrt{2\pi}\sigma^3} \right) e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Équation 8.8. Inverse de la dérivée seconde de la fonction Gaussienne.

La mise en œuvre au sein de la bibliothèque d'un filtre passe-haut, a été réalisée selon cette approche pour des raisons similaires à celles qui poussent à utiliser l'approche Gaussienne pour l'application d'un filtre passe-bas<sup>291</sup>. L'utilisation de l'écart type est plus adaptée aux changements de la taille matricielle et permet le contrôle de la largeur du lobe central et donc des caractéristiques de filtrage. Cette approche permet aussi d'envisager une modification de l'équation, afin de contrôler indépendamment l'écart type relatif à sa largeur et l'écart type relatif à sa longueur, ce qui permet de modifier la directivité au lobe. L'abstraction permettant de générer le noyau est `xel.2d.mat.fltr.log`. La mise en œuvre implique les mêmes principes de mise à l'échelle du noyau sur un plan cartésien, que le filtre Gaussien. Cependant, la différence

---

<sup>291</sup> La première approche basée sur un filtre boîte peut néanmoins être très facilement appliquée à l'aide de la bibliothèque `iemmatrix`.

de signe des valeurs centrales et des valeurs périphériques est un élément essentiel à la caractéristique passe-haut du filtre. Et la discrétisation de la fonction continue implique que la somme des valeurs négatives ne compense pas nécessairement de manière totale la somme des valeurs positives. Pour corriger cet artefact de l'échantillonnage de la fonction, deux normalisations respectives aux valeurs positives et négatives, sont appliquées de telle sorte que la somme des valeurs absolues de chaque partie est alors égale à  $\frac{1}{2}$ .

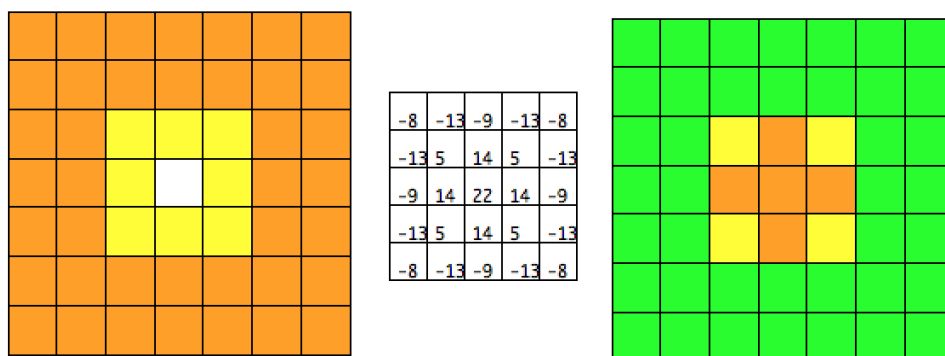


Figure 8.19. Capture d'écran des interfaces graphiques d'un patch Pure Data réalisant une convolution d'un espace sonore de taille 7x7 (à gauche) où les *soxels* contribuent de manière presque homogène. A l'exception de neuf *soxels* centraux allant de [2][2] à [4][4], qui contribuent plus faiblement jusqu'à une contribution nulle pour le *soxel* [3][3] avec un noyau de taille 5x5 (au centre) dont les coefficients sont générés à partir d'un filtre LOG d'écart type 0.66 pour donner un espace sonore (à droite) où les *soxels* centraux contribuent plus fortement qu'à l'origine et les *soxels* périphériques plus faiblement.

Ce type de noyau en traitement d'images est utilisé afin d'accentuer les variations des valeurs des *pixels* – de rendre les images moins floues d'une certaine façon – et l'une des principales applications de ce procédé est de permettre la détection de contours [Shapiro & al., 1995]. Appliqués au domaine de l'espace sonore les principes en jeu sont similaires, l'opération permet de diminuer les contributions des *soxels* lorsque ceux adjacents possèdent une forte corrélation et réciproquement d'augmenter la contribution des *soxels* lorsque ceux adjacents sont peu corrélés [Figure 8.19]. Sur un plan pratique, cette approche fait ressortir les particularités ou les irrégularités de l'espace sonore et diminue son aspect uniforme. Ce type de traitement de l'espace et du son offre par exemple la possibilité de faire disparaître ou d'amoindrir les contributions d'une source uniforme recouvrant une large étendue ou l'entièreté de l'espace.

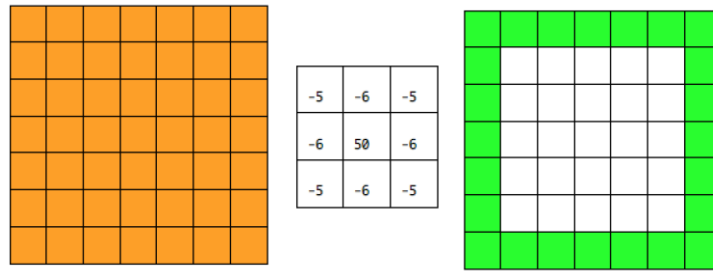


Figure 8.20. Interfaces graphiques d'un patch Pure Data réalisant une convolution d'un espace sonore de taille 7x7 (à gauche) où les *soxels* contribuent de manière homogène et uniforme sur l'ensemble de l'espace avec un noyau de taille 3x3 (au centre) dont les coefficients sont générés à partir d'un filtre LOG d'écart type 0.57 pour donner un espace sonore (à droite) où seuls les *soxels* limitrophes contribuent.

La prise en compte des bordures telle qu'elle est actuellement mise en œuvre pour l'opération de convolution, n'opère que peu d'artefacts sur les filtres de type passe-bas – permettant ainsi de jouer sur la diffusion des sources sonores composant l'espace – mais cette approche peut s'avérer peu adaptée aux filtres de type passe-haut. Ignorer les *soxels* dépassant les limites de l'espace sonore revient implicitement à étendre l'espace sonore par des lignes et des colonnes de *soxels* virtuels, aux contributions nulles. En considérant un espace sonore totalement uniforme où les contributions sont homogènes, les variations des contributions sont de fréquence nulle et l'application d'un filtre passe-haut devrait les réduire voire les annuler, et générer un espace sonore vide. Mais étant donné que l'opération de convolution prend en compte de manière tacite des lignes et des colonnes de *soxels* virtuels aux contributions nulles, cela génère des variations dans l'espace créant des artefacts qui se répercutent aux extrémités [Figure 8.20]. Cette approche nécessite de pouvoir utiliser une autre gestion des bordures, telle que l'extension par constante, qui duplique virtuellement les *soxels* périphériques, afin de préserver une homogénéité de contributions des *soxels* dans l'espace sonore.

Les possibilités de filtres spatiaux pourraient être encore bien plus développées, cependant les recherches présentées, qui possèdent déjà de nombreuses directions à investiguer, ont dû s'arrêter à ces trois exemples. Mais ces approches permettent d'ores et déjà d'établir des conclusions, et d'entrevoir des perspectives de développement. Deux types de filtres ont été proposés : à savoir passe-bas avec le filtre de boîte et le filtre Gaussien, et passe-haut avec le filtre LOG. Ces deux filtres offrent déjà sur l'espace sonore, des transformations intéressantes. La première approche permet d'appliquer un flou spatial sur l'espace sonore, qui peut correspondre à un élargissement des particularités sonores de l'espace et donc à une diminution de celles-ci. La deuxième approche permet, quant à elle, d'augmenter les particularités en accentuant les variations importantes dans l'espace sonore. Mais en diminuant et même

annulant les espaces sonores homogènes. Au-delà de ces possibilités, de nouvelles perspectives sont envisagées, comme par exemple, une meilleure gestion des bordures, tel que présenté précédemment. Mais aussi l'utilisation de couche *alpha* permettant de restreindre l'application d'un filtre à une partie de l'espace. Enfin, utiliser conjointement un filtre de type flou avec l'effet de « crayon qui s'efface », présenté dans le chapitre précédent, permettrait de simuler des sources sonores ponctuelles. Aussi serait-il intéressant d'élargir la synthèse d'espace sonore matriciel, selon des approches plus systématiques que celle proposée par le Soxel Art.

### 8.3. Synthèse de l'espace sonore

La mise en œuvre actuelle de la bibliothèque *xel* propose un ensemble de trois abstractions destinées à la synthèse d'espaces sonores matriciels, dont les mécanismes ont été élaborés de manière plus rationnelle. La mise en œuvre de ces outils s'inspire des approches originales proposées en ambisonie par la bibliothèque HOA, et notamment du travail présenté dans la partie précédente de cette thèse. L'enjeu consiste à utiliser des traitements relativement usuels en traitement du signal audio numérique, afin de synthétiser les signaux associés aux *soxels*. Cependant, le changement de représentation de l'espace sonore des harmoniques sphériques vers un système matriciel, implique de reconsidérer toutes les approches. Et c'est en cela que réside l'intérêt de chaque modèle de représentation de l'espace. Il s'agit de faire émerger de nouveaux procédés, et de nouveaux traitements à la fois de l'espace et du son. Afin d'orienter les approches, les mises en œuvre se sont encore ici inspirées des pratiques et des outils disponibles dans les logiciels de dessins numériques. L'approche présentée pour la mise en œuvre des filtres de l'espace, transpose directement une opération du traitement de l'image numérique au traitement de l'espace sonore, puis analyse les résultats. Et contrairement à cette approche, la réalisation des modules de synthèse tente de préserver l'idée des outils des logiciels de dessins numériques, en proposant un équivalent dans le domaine de la synthèse d'espace. L'enjeu n'est en aucun cas de définir s'il est possible de reproduire un coup de pinceau dans l'espace sonore, mais simplement d'utiliser les outils existants dans le domaine du dessin numérique, afin d'orienter les mises en œuvre et les propositions d'outils dans le domaine de l'espace sonore. En raison de la complexité et du temps que cela prendrait de présenter en détails les mécanismes internes de ces modules de synthèse, cet exposé préfère se limiter à un aperçu général du fonctionnement de ces outils, et des possibilités qu'ils offrent.

### 8.3.1. Le pinceau de l'espace sonore

Le premier outil mise en œuvre est l'abstraction *xel.2d.syn.brush~* qui synthétise les signaux des cellules d'une matrice à partir d'un signal d'origine et en fonction de coordonnées. Aussi, en un certain sens, l'objectif est de permettre la synthèse de source sonore directionnelle. Mais les paramètres de contrôle offerts par cette approche amènent à considérer l'outil comme l'équivalent d'un pinceau.

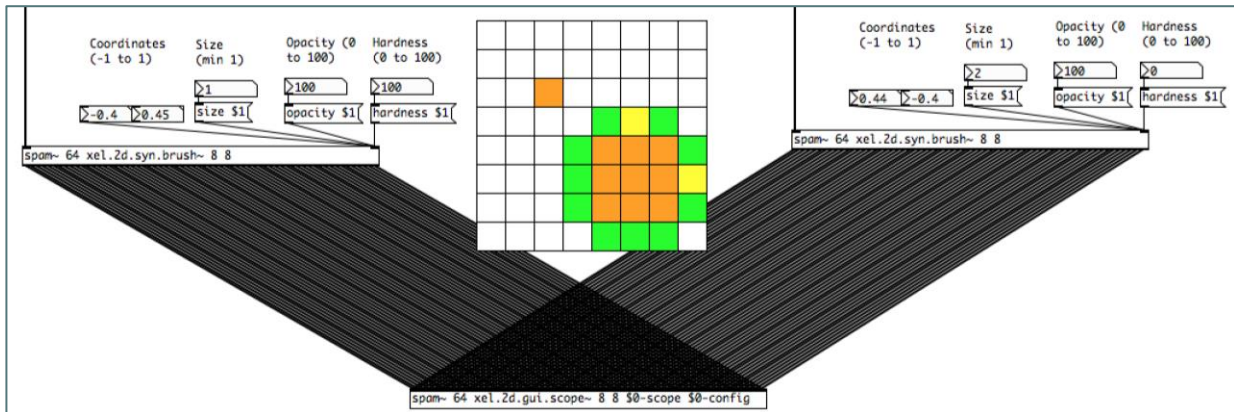


Figure 8.21. Patch Pure Data où deux abstractions *xel.2d.syn.brush~* synthétisent deux espaces sonores additionnés sur l'interface graphique au centre. Le premier espace sonore est synthétisé à gauche avec un pinceau d'une taille de 1 *soxel* et d'une dureté de 100% aux coordonnées cartésiennes -0,45 et 0,45 résultant en la contribution seulement du *soxel* [3][3]. Le deuxième espace sonore est synthétisé à droite avec un pinceau d'une taille de 2 *soxels* et d'une dureté de 0% aux coordonnées cartésiennes 0,44 et -0,45 résultant en la contribution de plusieurs *soxels* autour d'un *soxel* central [6][6] et dont les contributions évoluent avec la distance.

L'opération en jeu pour synthétiser les signaux associés aux *soxels*, est de générer des versions pondérées du signal d'origine. Pour cela, le gain du signal dans chaque *soxel* est calculé en utilisant la distance entre les coordonnées de synthèse et les coordonnées associées à chaque *soxel*. Aussi, afin d'offrir une cohérence tout au long de la chaîne de traitements, les coordonnées sont définies telles que présentées pour la projection de l'espace sonore [Équation 8.1]. Cette approche assure que la position relative aux haut-parleurs ne devrait pas changer en fonction de la taille matricielle de l'espace sonore. Le gain associé à un *soxel* diminue en fonction de sa distance avec les coordonnées de synthèse. Et lorsque cette distance atteint un seuil maximal, le gain est nul. Les variations du gain via une fonction sinusoïdale – afin d'offrir un rendu qui semble plus fluide et par défaut la distance maximale – est définie par la distance minimale entre deux cellules. Ainsi, le signal ne peut pas contenir plus de cinq cellules et lorsque les coordonnées de synthèse sont identiques à celles d'un *soxel*, alors seulement ce *soxel* contribue à l'espace sonore [Figure 8.21]. Néanmoins, le paramètre de taille, en référence à la taille du pinceau en *pixel*, permet d'augmenter le nombre de *soxels* qui contribuent à l'espace

sonore en diminuant virtuellement la distance entre la source et les *soxels* permettant d'élargir la zone de synthèse, ou la source sonore ponctuelle. Le paramètre de dureté modifie la distance maximale après laquelle les contributions deviennent nulles, ce qui permet d'obtenir une délimitation plus ou moins abrupte entre les *soxels* contribuant à l'espace sonore, et les autres. Enfin, le paramètre d'opacité agit comme un gain global sur l'espace sonore.

Selon une approche classique et usuelle de la spatialisation du son, cette abstraction permet de générer des sources sonores ponctuelles et de créer des trajectoires, en variant les coordonnées. Néanmoins, au-delà de cette approche, cela permet déjà de mettre en place des opérations intéressantes, comme la variation de la taille des sources sonores et leur manière de se répandre dans l'espace.

### 8.3.2. Le pot de peinture de l'espace sonore

Le deuxième outil mis en œuvre est l'abstraction *xel.2d.syn.bucket~*. Cette approche s'inspire du terme « pot de peinture » des logiciels de dessins. Evidemment, une opération exactement similaire de type « pot de peinture » serait intéressante pour la synthèse d'espace sonore, dans le sens où elle consisterait à appliquer de manière homogène, un signal à l'ensemble des *soxels*. L'approche dans ce contexte s'inspire en réalité plutôt de l'effet de recouvrement qu'offrirait un réel pot de peinture versé dans un espace. Celui-ci s'écoulerait depuis le point de versement, dans toutes les directions.

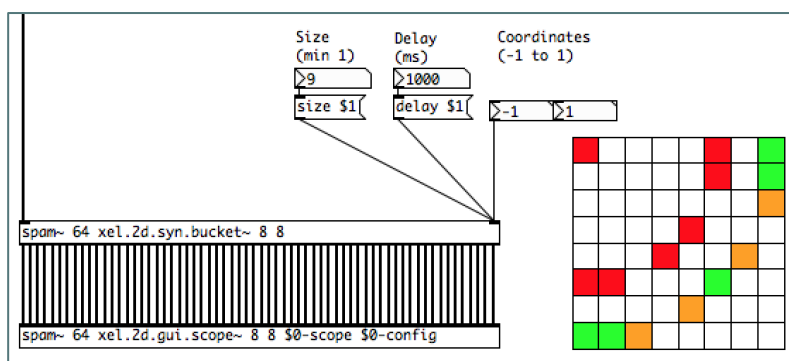


Figure 8.22. Patch Pure Data où l'abstraction *xel.2d.syn.bucket~* synthétise à gauche un espace représenté sur l'interface graphique et à droite à partir d'un signal d'origine constitué de plusieurs impulsions. Les coordonnées de synthèse sont -1 et 1, la taille de 9 *soxels* et un temps de retard de 1000 ms. Aussi, le résultat apparaît alors sous la forme de vagues de sons, qui se propagent de manière radiale du coin avant-gauche de l'espace sonore pour atteindre le coin opposé en 1000 ms.



L'opération en jeu pour synthétiser les signaux associés aux *soxels*, est de générer des versions retardées du signal d'origine. Pour cela, le retard du signal dans chaque *soxel* est calculé en utilisant, de manière identique à l'approche précédente, la distance entre les coordonnées de synthèse et les coordonnées associées à chaque *soxel*. De même, les coordonnées sont définies telles que présentées pour la projection de l'espace sonore [Équation 8.1]. Le retard associé à un *soxel* augmente de manière proportionnelle à la distance de ce *soxel* avec les coordonnées de synthèse. Cela permet ainsi de donner l'impression que le son se répand radialement, à partir d'un point de l'espace [Figure 8.22]. Le temps de retard associé à chaque *soxel* est donc défini en fonction des coordonnées de synthèse, mais aussi d'un temps de retard maximal et d'une taille maximale en *soxel*, après laquelle les *soxels* ne contribuent plus. En effet, au fur et à mesure que la distance augmente, de manière similaire à la synthèse par pinceau, les gains associés aux *soxels* diminuent.

Cette approche originale permet de générer des espaces sonores complexes, dont les résultats peuvent être associés aux champs diffus synthétisés avec la décorrélation des signaux associés aux harmoniques sphériques. Cependant, cette similitude est relativement superficielle. De par la représentation particulière de l'espace et la mise en œuvre spécifique de la ligne retard sur les signaux associés aux *soxels*, l'espace sonore offre une diffusion au caractère bien différent, mais toujours aussi difficilement définissable. L'une des caractéristiques majeures pour concevoir les différences entre les deux approches peut être – de manière similaire à l'approche ambisonique – d'utiliser des impulsions relativement courtes comme signal d'origine, comparé au temps de retard. Dans l'approche matricielle qui vient d'être présentée, l'effet peut alors être assimilé à un phénomène de vagues ou d'ondes se répandant dans l'espace. De manière, plus subjective, l'impression est celle d'une source sonore se déplaçant du point défini par les coordonnées de synthèse dans de multiples directions de l'espace sonore.

### 8.3.3. L'aérosol de l'espace sonore

Le troisième outil mis en œuvre est l'abstraction *xel.2d.syn.spray~*. Cette approche s'inspire, encore ici librement, de l'effet d'aérosol des logiciels de dessins. L'idée est donc de générer autour d'un point central, un effet de répartition stochastique des contributions associées aux *soxels*, telle une pulvérisation des contributions dans l'espace. L'objectif est que la contribution d'un *soxel* apparaisse pendant que celle d'un autre disparaît. Et ainsi, que la probabilité que les *soxels* contribuent de manière générale, diminue en fonction de leur éloignement du point central.

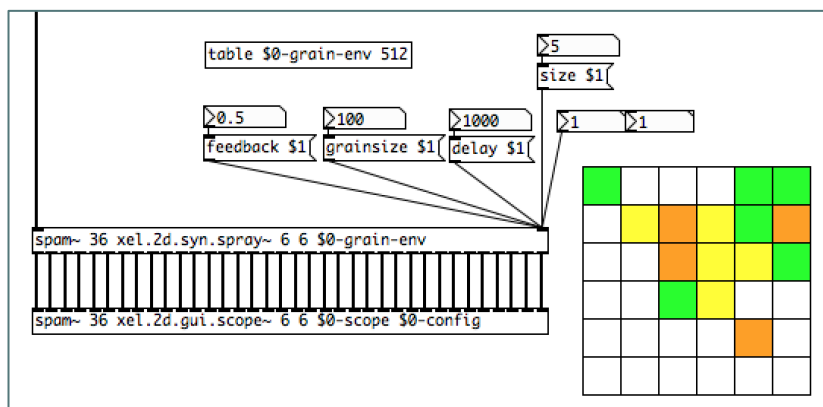


Figure 8.23. Patch Pure Data où l'abstraction *xel.2d.syn.spray~* synthétise à gauche un espace représenté sur l'interface graphique à droite à partir d'un signal d'origine. Les coordonnées de synthèse sont 1 et 1 et la taille de 5 *soxels*. Les *soxels* contribuent de manière décorrelée en temps et en amplitude. Les *soxels* du coin avant-droite contribuent de manière importante et les contributions diminuent au fur et à mesure que les *soxels* s'éloignent de ce point, pour être quasiment nuls à la diagonale avant-gauche /arrière-droite de l'espace.

Pour obtenir un effet similaire, l'abstraction synthétise les signaux associés aux *soxels* avec un module de synthèse granulaire quasi-synchrone, permettant d'obtenir des grains sonores décorrelés à la fois temporellement et en amplitude, dans l'ensemble des harmoniques. L'aspect intéressant de l'approche reste les quatre paramètres de la QSGS à savoir la taille des grains, le temps de retard, la réinjection et la raréfaction. Cette mise à l'échelle de ces paramètres, dépend des coordonnées de synthèse et de la taille de la pulvérisation. Avec la taille par défaut, le grain est concentré autour des coordonnées d'encodage, car la raréfaction des cellules éloignées de ces coordonnées est de cent pour cent. Alors que la taille augmente, la raréfaction de ces cellules diminue, et les grains sont répartis dans l'espace. De plus, la taille du grain diminue et le temps de retard augmente avec la distance. Ainsi, si la taille de la pulvérisation est relativement importante, les grains principaux apparaissent d'abord autour des coordonnées de synthèse et se propagent ensuite dans l'espace, sous la forme de grains plus petits [Figure 8.23]. Seule la valeur de réinjection est globale à l'ensemble des *soxels*.

De par l'utilisation de lignes à retard et d'un temps qui augmente en fonction la distance des *soxels* par rapport à la position de synthèse, l'effet pourrait se rapprocher de la synthèse par pot de peinture. Mais le caractère aléatoire et la génération de grain transforme radicalement le résultat. Cette approche permet de générer des masses de grains mais contrairement à l'approche ambisonique où ceux-ci recouvrent l'entièreté de l'espace, l'intérêt de l'approche matricielle de l'espace réside en la possibilité de contrôler leurs répartitions. Cela se fait en utilisant des coordonnées permettant de définir le centre de gravité des grains, une taille en *soxels* permettant de définir l'étendue de leurs disséminations et un temps en millisecondes.

A ces trois modules de synthèse d'espace sonore qui permettent de « dessiner », il est possible d'associer un dernier module, qui s'inspire de l'outil de gomme des logiciels de dessin. Ce module offre de diminuer ou d'annuler les contributions de l'espace sonore. L'approche et les paramètres y sont sensiblement similaires à ceux du pinceau de l'espace sonore, mais évidemment le résultat est opposé.

L'enjeu dans le cas présent, est d'offrir un aperçu des possibilités. Aussi, trois approches de la synthèse d'espace sonore ont été présentées mais, comme pour l'approche ambisonique, de nombreuses autres peuvent être envisagées. Cependant, il est déjà possible de remarquer certaines propriétés récurrentes de ces mises en œuvre, qui mettent en avant les caractéristiques de la représentation matricielle de l'espace. La propriété peut-être la plus importante est que ces traitements de synthèse d'espace sonore sont construits sur des calculs de distances. Ils amènent donc à jouer de manière bien différente sur l'espace comparé, par exemple, à l'approche plus directionnelle de l'ambisonie. Un autre élément essentiel est la possibilité de varier le caractère plus ou moins diffus des espaces sonores, dans le cas présent, en jouant notamment sur les paramètres de taille des modules de synthèse. Dans ce cas, l'approche fait écho à l'objectif proposé par la synthèse chaotique – ou plus récemment dans les expérimentations ambisoniques associées à la bibliothèque *spam* – qui propose un paramètre de diffusion permettant de passer linéairement d'un espace sonore diffus à une source ponctuelle. Aussi est-il possible de remarquer qu'en certains points, l'approche musicale de l'ambisonie et la pratique des *soxels* se recoupent mais, qu'en d'autres elles se dissocient. Et bien qu'il soit intéressant de poursuivre les recherches, aussi bien en ambisonie par l'utilisation des *soxels* ou sur de nouvelles représentations de l'espace, il semble déjà pertinent de confronter ces deux approches.

## 8.4. Perspectives

Il est difficile, si tant est que cela soit possible, de valider de façon objective le raisonnement qui a conduit à envisager les représentations abstraites et intermédiaires de l'espace sonore, comme des outils pertinents pour la création musicale. Dans cette dernière partie, la tentative de justification de cette approche a été réalisée en proposant premièrement un ensemble d'outils logiciels permettant d'expérimenter la mise en espace du son avec de nouvelles représentations de l'espace sonore. Mais surtout, par la suite, en mettant en place une série de propositions et de réalisations concrètes. Cette dernière approche a consisté à définir le modèle de

représentation de l'espace, en l'associant à un système de pensée, à des outils mathématiques et à une conception spécifique de l'espace sonore. Par la suite, une première proposition de prise en main de cette représentation de l'espace a été proposée. L'enjeu n'est pas simplement de présenter comment cette représentation peut être utilisée, mais bien de montrer par les opérations offertes, le potentiel musical et compositionnel sous-jacent à cette représentation. Evidemment il est, dans ce cas, à chacun d'en juger et une critique un tant soit peu objective des outils offerts ne pourra être réalisée que si les possibilités qu'ils offrent se retrouvent dans des approches musicales. De façon identique, par exemple, à la validation des procédés mis en œuvre dans la bibliothèque HOA. Aussi, afin de dépasser le domaine des idées, d'autres axes de critiques seront proposés dans ce dernier chapitre. En effet, l'approche a été développée afin d'offrir des solutions à la restitution de l'espace sonore sur un système de haut-parleurs. Mais aussi afin de développer les opérations autour du cadre mathématique, notamment via le filtrage spatial. Et enfin de façon à proposer des approches plus systématiques de synthèse de l'espace sonore, faisant écho aux pratiques élaborées précédemment en ambisonie.

A présent qu'une deuxième approche originale de la spatialisation du son construite sur une représentation intermédiaire et abstraite de l'espace, a été non seulement proposée mais relativement bien développée, il serait nécessaire de la confronter avec la première. L'objectif serait de tirer parti des expérimentations – réalisées en ambisonie et sur la représentation de l'espace sonore par des *soxels* – pour déterminer un ensemble de spécifications conceptuelles, techniques et fonctionnelles. Ces spécifications permettant d'orienter les méthodologies qui mettent en œuvre des approches originales de la spatialisation du son, fondées sur une représentation abstraite de l'espace, mais aussi d'offrir des outils permettant de critiquer les approches. Cette généralisation de l'approche pourrait valider de manière plus rationnelle le raisonnement présenté dans cette thèse, au-delà des outils et des opérations déjà présentées. Au même titre que M. A. Gerzon propose une métathéorie de la localisation auditive afin d'élaborer et de critiquer des approches visant à reproduire les effets directionnels des sons dans l'espace, l'enjeu serait, à terme, de proposer une métathéorie<sup>292</sup> de la représentation intermédiaire et abstraite de l'espace, afin de pouvoir élaborer et critiquer des approches plus musicales, telles que présentées dans cette thèse.

---

<sup>292</sup> Oui, n'ayons pas peur des grands mots ! Car le gouvernail dont les voiles, sous l'impulsion du magnifique cheval de trait indigène, entraînent, sur la route toute droite de la prospérité, le champagnacien qui ne craint pas ses méandres car il sait qu'en serrant les coudes il gardera les deux pieds sur terre, afin de s'élever à la sueur de son front musclé, vers des sommets... [Franquin & al., 1958]

# Conclusion

Cette thèse défend que les représentations abstraites et intermédiaires de l'espace sonore permettent d'élaborer de nouvelles approches originales de la mise en espace du son. Le raisonnement amenant à cette affirmation commence par l'appropriation musicale de l'approche ambisonique. Mettre en œuvre de nouveaux traitements de l'espace et du son en ambisonie amène, à utiliser de manière originale les signaux associés aux harmoniques sphériques et à concevoir différemment les relations qui les régissent et leur hiérarchisation. La particularité de ces approches expérimentales et les caractéristiques singulières des champs sonores générés, nécessitent de concevoir de nouveaux outils – théoriques et pratiques – pour l'analyse et la restitution des champs sonores. Les changements opérés sont tels, qu'il est possible de libérer cette approche des enjeux techniques et matériels initiaux en ambisonie. Mais surtout à s'émanciper des modèles psychoacoustiques et acoustiques sur lesquels les techniques ambisoniques reposent originellement. Dans ce contexte, les signaux associés aux harmoniques sphériques ne sont plus nécessairement une représentation rationnelle du champ sonore mais deviennent une représentation abstraite de l'espace sonore possédant en soi, un potentiel musical. Cette thèse peut alors être validée en généralisant cette approche et en proposant des traitements originaux de l'espace et du son, spécifiques à de nouvelles représentations de l'espace sonore.

Ce raisonnement a été développé dans la deuxième partie de cette thèse suite à une première partie qui présentait l'ensemble des connaissances et les outils nécessaires à l'élaboration du raisonnement. La troisième partie, quant-à-elle, a tenté de valider la problématique. Chacune de ces parties soulève des enjeux spécifiques à leurs sujets et ont été divisées en plusieurs chapitres accompagnant la réflexion.

Dans le premier chapitre de cette thèse, un état de l'art a été établi, retraçant les principaux développements des techniques ambisoniques de ses origines à nos jours. A travers cela, l'enjeu était de faire ressortir les différentes problématiques auxquelles les techniques ambisoniques tentent de répondre, telles que la reproductibilité du champ sonore ou l'adaptation à différents dispositifs de restitution. D'autre part, cette présentation a aussi permis de comprendre comment l'ambisonie répond à ces problématiques. Aussi, l'objectif était de faire ressortir la structure du modèle ambisonique, et comment s'agencent les différentes notions auxquelles ce modèle fait appel. Après un regard macroscopique sur le sujet, le deuxième chapitre de cette thèse a proposé de revenir en détail sur ces notions importantes. En se penchant sur les domaines

mathématiques, acoustiques et psychoacoustiques sur lesquels repose l'ambisonie, l'enjeu a été de pouvoir, par la suite, se les approprier dans une approche musicale. Mais aussi d'offrir les outils permettant de comprendre et d'analyser les implications que cette utilisation originale des techniques ambisoniques peut engendrer sur son modèle. Le troisième chapitre de cette thèse a présenté les différents outils logiciels nécessaires à l'expérimentation, dans le domaine des harmoniques sphériques. L'enjeu n'était alors pas seulement de pouvoir comprendre les expérimentations présentées par la suite sur un plan pratique. Cela a aussi permis de présenter les mises en œuvre et les spécifications techniques et fonctionnelles, afin de pouvoir – suite à leur usage – les critiquer et proposer des outils plus adaptés à l'approche expérimentale proposée dans cette thèse.

A la suite de la présentation du modèle ambisonique et des outils théoriques et logiciels, le quatrième chapitre a proposé de revenir sur cinq opérations originales du champ sonores. L'enjeu est de révéler les caractéristiques spécifiques de l'appropriation musicale du modèle ambisonique, et notamment de la représentation du champ sonore sous la forme de signaux associés aux harmoniques sphériques. Pour n'en citer qu'une par opération : la variation de la résolution angulaire permis de présenter un jeu sur les contributions relatives à chaque degré des signaux associés aux harmoniques sphériques. La synthèse de champ sonore par décorrélation temporelle a proposé des approches originales sur les symétries spatiales de ces signaux. La synthèse de champ sonore par granulation a amené à considérer la qualité du champ sonore en fonction de l'ordre de décomposition. L'encodage de champ sonore chaotique, de son côté, a soulevé la question de la création de la directionnalité. Enfin, le *flanging* ambisonique a amené la problématique de la transcription spatiale d'un effet spécifique. Le cinquième chapitre a permis alors d'analyser les expérimentations présentées, d'établir les conséquences que cela implique sur le modèle ambisonique et de proposer de nouveaux outils pour répondre aux problématiques actuelles, qui ont été soulevées. L'enjeu a alors été de construire le raisonnement menant à la problématique de cette thèse, tout en suivant une approche pratique. Pour cela, les outils d'analyse critique du champ sonore proposés originellement en ambisonie – via les vecteurs vitesse et énergie – ont été confrontés aux spécificités de l'appropriation musicale de l'ambisonie. Partant du constat qu'ils se sont révélés caduques, la théorie psychoacoustique sur laquelle repose l'approche ambisonique – via notamment la description des mécanismes de localisations auditives basses et hautes fréquences – est apparue de même inadaptée ou insuffisante pour appréhender les approches originales de cette thèse. Et de nouveaux outils d'analyse ont alors été proposés. Selon une approche similaire, l'opération de décodage a été remise en question et reformulée afin de satisfaire les

nouvelles exigences liées aux champs sonores synthétisés, telles que la qualité du caractère diffus ou chaotique des champs sonores. Les changements opérés sur le modèle ambisonique ont amené à concevoir les signaux associés aux harmoniques sphériques comme des outils de mise en espace du son, construits sur une représentation abstraite et intermédiaire de l'espace sonore. Et dont les utilisations ont pu être généralisées pour concevoir de nouvelles approches.

Avant de valider sur un plan pratique la thèse, le sixième chapitre a proposé de revenir confronter les outils logiciels de la bibliothèque HOA, aux usages qui en ont été fait. L'enjeu était alors de définir un certain nombre de spécifications techniques et fonctionnelles, pouvant répondre aux usages futurs de plusieurs représentations abstraites et intermédiaires de l'espace sonore. Le résultat est un livrable – la bibliothèque *spam* – qui est un ensemble d'outils pour le logiciel Pure Data répondant à des problématiques expérimentales dans des systèmes multicanaux, et favorisant la modularité ainsi que le dynamisme des opérations. Cette bibliothèque répond aussi à des enjeux didactiques favorisant la lisibilité des mécanismes mis en œuvre. Le septième chapitre de cette thèse a alors proposé une représentation particulière de l'espace, sous une forme matricielle. L'objectif consistait à associer un système matriciel à un espace sonore, pour en faire un outil de spatialisation. Pour cela, l'espace sonore abstrait a été décomposé par un ensemble d'éléments fondamentaux nommés des *soxels*, et sur lesquels un ensemble de pratiques ont été élaborées. En proposant des opérations de l'espace pertinentes ou du moins originales, l'enjeu était de valider cette thèse sur un plan conceptuel. Enfin, dans le huitième chapitre, l'opération de projection des signaux associés aux *soxels* a été présentée, permettant ainsi de soulever les problèmes et les enjeux liés à son élaboration. Cette approche de la mise en espace du son a aussi été développée autour de son formalisme mathématique – offrant des opérations de filtrages spatiaux. Et des techniques de synthèse d'espace sonore – par décorrélation temporelle, granulation ou pondération – ont été mise en œuvre par l'adaptation des méthodologies utilisées en ambisonie. L'objectif de ce dernier chapitre a été de confronter cette approche aux considérations concrètes d'un système de restitution afin notamment de valider cette thèse sur un plan pratique. Mais l'objectif a surtout été d'établir de nouvelles pistes de réflexion quant à l'élaboration d'un système de mise en espace du son.

Ce travail de recherche et d'expérimentation aura amené de nombreuses propositions à la fois théoriques et conceptuelles – vis-à-vis des modèles de spatialisation – mais aussi pratiques, avec son lot de mises en œuvre et d'outils logiciels. Et ces propositions auront, selon moi, permis de valider la problématique de cette thèse en définissant la représentation intermédiaire et abstraite de l'espace comme outil de spatialisation du son. Cependant, le travail ne peut en

rester là, car de nombreux thèmes ont été traités, ouvrant à de vastes domaines de recherche. Plusieurs points mériteraient donc d'être approfondis. Mais avant cela, il me semble important de présenter les problèmes majeurs que j'ai pu rencontrer pendant l'élaboration de ce travail de thèse.

Le premier est lié à des considérations principalement techniques, dues à la mise en œuvre des outils logiciels. Je ne suis pas programmeur de formation, et bien que je sois arrivé à travailler sur les projets HOA en possédant des notions convenables de programmation en C et C++, bon nombre de mes connaissances dans ce domaine ont été apprises au cours de ces projets. Donc en pratiquant, bien sûr, mais aussi en faisant par là même des erreurs. Aussi les développements strictement informatiques de la bibliothèque HOA ont été une expérience à la fois extrêmement enrichissante mais aussi chronophage. Je ne compte plus le nombre de refontes de la bibliothèque à partir de zéro, ou le temps passé à mettre en œuvre des algorithmes tels que le diagramme de Voronoï sphérique pour la représentation de l'objet *hoa.3d.meter~*. Et à ce titre, je remercie tout particulièrement Elliott Paris qui m'a toujours suivi, notamment dans ces folles restructurations. Et j'espère avoir été tout aussi enthousiaste et compréhensif, lorsque lui aussi – à son tour – s'engouffrait dans les méandres sans fins des classes C++<sup>293</sup>. Finalement, j'en suis sorti enrichi, et malgré le fait que de nombreux problèmes résident encore dans les mises en œuvre – notamment sur des questions de compatibilités avec les différentes versions des logiciels pour les lesquels la bibliothèque a été mise en œuvre – je pense désormais éviter de nombreux pièges dans mes récents développements.

Le deuxième problème que j'ai rencontré est d'autant plus important qu'il touche au nerf de cette thèse. Peut-être était-ce totalement absurde de ma part, mais j'en suis venu à travailler sur une technique de spatialisation – en l'occurrence l'ambisonie – reposant sur des principes acoustiques, psychoacoustiques et faisant appel à des notions mathématiques relativement complexes. Or, il est à noter que j'étais alors complètement néophyte dans les deux premiers domaines – voire inculte – et légèrement mieux qu'ignorant dans le troisième domaine. Aussi ai-je dû faire preuve d'énormément de patience et de persévérance<sup>294</sup>, afin de pouvoir cerner les enjeux et les implications liés à ces domaines, mais afin de me sentir un peu légitime. Ces

---

<sup>293</sup> Nous avons formé, il est vrai, un beau duo de programmeurs bras cassés mais l'envie était là et c'est ce qui compte !

<sup>294</sup> À cela, il faut rajouter une méthode d'apprentissage relativement brutale consistant à lire vingt fois un ouvrage pour assimiler son contenu. Méthode effectivement sous optimale mais que je m'efforce bizarrement à préserver dans certains cas.



lacunes sont logiques, car j'ai suivi un cursus en musique et composition assistée par ordinateur, et non en acoustique. Et au regard du sujet de cette thèse et de l'orientation donnée à l'approche ambisonique, elles me paraissent aussi une chance. En effet, avoir parcouru ce chemin tardivement et de manière autonome me permet, je crois, de porter un regard original sur ces domaines et de pouvoir faire preuve de pédagogie. Car le caractère didactique de mes présentations et la pédagogie sont au cœur de ma démarche. Que ce soit dans ce travail de thèse, que j'espère avoir pu rendre accessible à un grand nombre de personnes, musiciens, compositeurs ou autres. Au travers des tutoriels que j'ai réalisés pour la bibliothèque HOA. Ou dans le cadre du séminaire de Master « Créations d'espaces sonores » que j'eus la chance de pouvoir concevoir et dispenser pendant deux ans à l'Université Paris 8. Et à ce titre, je remercie encore particulièrement Anne Sèdes et Alain Bonardi de m'avoir fait confiance et donné une véritable liberté pour mener à bien ces séminaires.

Au-delà de ces problèmes, et peut-être des tracasseries propres à l'ensemble des doctorants dans quelque domaine qu'il soit, je souhaite rappeler que le contexte de travail au CICM a été particulièrement agréable et enrichissant. Aussi, je désire à présent envisager des perspectives à mon sujet de thèse, mais aussi remettre en cause certains aspects de cette approche.

En bien des endroits, cette approche peut être questionnée. Notamment au sujet de la source directionnelle ou ponctuelle. En effet, si les procédés originaux proposés en ambisonie permettent encore de synthétiser des sources directionnelles, alors il peut sembler logique de considérer que les signaux associés aux harmoniques restent cohérents avec la représentation locale du champ sonore, telle que définie originellement en ambisonie. Dans ce cas, le raisonnement proposé serait donc mis en cause. Aussi, afin de répondre à cette question, il convient de partir de la représentation matricielle de l'espace sonore, et d'accepter qu'un *soxel* peut tout aussi bien être considéré comme une source sonore ponctuelle, selon bien sûr certaines approximations. Même en assumant que, sur un plan acoustique et psychoacoustique, la distance ne peut pas être convenablement simulée via une technique de projection telle que le VBAP, un *soxel* apparaît néanmoins, comme une source plus ou moins directionnelle lors de la restitution. Et il semble raisonnable d'assumer que cette approximation n'est ni pire, ni mieux, que celle faite en ambisonie pour simuler une source sonore directionnelle. L'encodage directionnel en ambisonie peut être vu simplement comme un jeu de gains particulier sur les signaux associés aux harmoniques sphériques qui, lorsqu'il suit une certaine logique, permet de concentrer l'énergie dans une direction. Cela donne ainsi l'impression d'une source sonore provenant de cette direction. En somme, l'encodage directionnel ne serait qu'une opération au

milieu de tant d'autres. Et de même que certaines lois et relations sont nécessaires pour faire émerger un espace sonore diffus ou un espace sonore chaotique, certaines lois et relations sont nécessaires pour faire émerger la direction. Aussi, l'appropriation musicale de l'ambisonie proposée dans cette thèse peut être considérée comme une approche générale de la spatialisation ambisonique, fondée sur la décomposition de l'espace sonore en harmoniques sphériques. Elle intégrerait dans certains cas les notions de directionnalité pour lesquels les principes acoustiques et psychoacoustiques seraient valides. Mais ces notions de directionnalités ne seraient pas générales à l'approche. Et par là même, ces principes acoustiques ne seraient pas valides pour l'ensemble des possibilités offertes par la décomposition de l'espace sonore en harmoniques sphériques. En somme, dans ce cas, l'ambisonie ne serait qu'une approche restreinte d'une approche plus générale de la spatialisation du son fondée sur la décomposition de l'espace sonore en harmoniques sphériques, de manière analogue à ce qu'est la théorie de la relativité restreinte, vis-à-vis de la théorie de la relativité générale. Aussi, tout comme les lois de la mécanique de Galilée-Newton ne satisfont plus l'ensemble des expériences modernes et nécessitent d'être étendues, les principes acoustiques et psychoacoustiques qui sous-tendent l'ambisonie, nécessitent d'être élargis. A ce sujet, une première ébauche de travail a été proposée dans le cinquième chapitre de cette thèse. Ce chapitre s'intéresse en effet à la caractérisation et la description des espaces sonores originaux synthétisés, en utilisant dans le temps les vecteurs énergie et vitesse. Mais ce travail nécessite d'être approfondi et de plus, de nombreuses autres perspectives peuvent être envisagées dans cette approche.

Enfin, pour continuer sur cette lancée visant à étendre et généraliser les approches, il serait intéressant, dans le cadre de futures recherches, d'explorer de nouvelles représentations de l'espace. Procéder de la sorte permettrait d'autant plus de valider la problématique de cette thèse. La liste des possibilités est infinie et à titre personnel, une d'entre elle me tient particulièrement à cœur. Il s'agit de l'utilisation d'une triangulation de Delaunay sur des points générés aléatoirement, permettant d'envisager des résultats continuellement différents. Car les représentations intermédiaires et abstraites de l'espace n'impliquent aucune règle ou restriction préalable. Aussi toutes sortes d'approche peuvent être envisagées. Néanmoins, il semble évident que toutes n'offrent pas le même potentiel et ne sont pas forcément pertinentes. Il serait intéressant dans ce cas, de proposer un certain nombre de critères permettant d'évaluer les représentations intermédiaires et abstraites de l'espace sonores, et les techniques qui en découlent. L'objectif étant de pouvoir définir des spécifications, des enjeux et des objectifs élémentaires auxquels répondre pour élaborer des techniques de spatialisation du son, fondées sur des représentations abstraites et intermédiaires de l'espace. Pour cela une approche qui

paraît *a priori* efficace consiste à recouper et confronter les opérations et les mises en œuvre issues de représentations spécifiques de l'espace sonore. Il s'agit de faire ressortir les éléments qui semblent pertinents dans une approche liée à une représentation de l'espace sonore. Et de déterminer ensuite si ces caractéristiques se retrouvent dans d'autres approches liées à d'autres représentations de l'espace sonore. Il est déjà possible, d'ailleurs, de procéder de la sorte entre les techniques offertes par la décomposition en harmoniques sphériques de l'espace et celles offertes par la décomposition matricielle. Un élément important et évident reste l'ordre de décomposition définissant le nombre de signaux associés aux harmoniques sphériques, ou le nombre de signaux associés au *soxels* permettant de représenter l'espace sonore. Il serait alors possible de questionner l'adaptabilité des systèmes aux différents ordres de décomposition, et d'en faire un critère pour l'élaboration des techniques et leurs mises en œuvre. Une autre caractéristique forcément récurrente de la représentation est qu'elle est intermédiaire et implique les opérations d'encodage et de décodage. La méthodologie utilisée pour concevoir ces opérations liées à une représentation spécifique de l'espace sonore pourraient, ici aussi, se révéler adaptée à d'autres représentations de l'espace sonore. Ce qui a d'ailleurs été le cas pour la mise en œuvre de la projection des *soxels* via le *distance based amplitude panning*, qui s'inspire de l'approche combinée de l'ambisonie et du *vector base amplitude panning*. Aussi, cela revient à élaborer cette métathéorie de la spatialisation du son fondée sur la représentation abstraite et intermédiaire de l'espace.

# Bibliographie

- [Adriaensen , 2006], F. Adriaensen, « Near Field filters for Higher Order Ambisonics », Note technique disponible accessible en ligne à l'adresse <http://kokkinizita.linuxaudio.org/papers/> (consulté en septembre 2017)
- [Algazi & al., 2001], V. R. Algazi, R. O. Duda, D. M. Thompson et C. Avendano, « The CIPIC HRTF database », Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, p. 99-102, Etats-Unis, New Paltz, octobre 2001
- [Askey, 2005], R. Askey, « The 1839 paper on permutations: its relation to the Rodrigues formula and further developments », dans. « Mathematics and social utopias in France: Olinde Rodrigues and his times », par Simón L. Altmann et E. L. Ortiz, History of mathematics, 28, Providence, R.I.: American Mathematical Society, p. 105–118, 2005
- [Batke & al., 2010], J-M. Batke et F. Keiler, « Using VBAP-derived Panning Functions for 3d Ambisonics Decoding », Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics, France, Paris, mai 2010
- [Bauer & al., 1971], B. B. Bauer, D. W. Gravereaux, et A. J. Gust, « A Compatible Stereo-Quadraphonic (SQ) Record System », Journal of the Audio Engineering Society, vol. 19, p. 638-646, septembre 1971
- [Bennett & al., 1985], J. C. Bennett, K. Barker et F. O. Edeko, « A New Approach to the Assessment of Stereophonic Sound System Performance », Journal of the Audio Engineering Society, vol. 33, no. 5, p. 314-321, may 1985
- [Berkhout, 1988], A. J. Berkhout, « A Holographic Approach to Acoustic Control », Journal of the Audio Engineering Society, vol. 36, no. 12, p. 977-995, décembre 1988
- [Berkhout & al., 1991], A. J. Berkhout, D. de Vries et P. Vogel, « Acoustic Control by Wave Field Synthesis », Journal of the Audio Engineering Society, vol. 93, no. 5, p. 2764-2778, mai 1993
- [Blanco & al., 1997], M. A. Blanco, M. Florez et M. Bermejo, « Evaluation of the rotation matrices in the basis of real spherical harmonics », journal of Molecular Structure (Theochem), vol. 419 p. 19–27, janvier 1997
- [Blauert, 1997], J. Blauert, « Spatial hearing: the psychophysics of human sound localization », MIT Press, Cambridge, États-Unis, Massachusetts, rev. edition, 1997

[Blumlein, 1933], A. D. Blumlein, « British Patent Specification 394,325 (Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems) », reproduit dans le Journal of the Audio Engineering Society, vol. 6, no. 2, p. 91-99, avril 1958

[Boone & al., 1995], M. M. Boone, E. N. G. Verheijen et P. F. Van Tol, « Spatial Sound-Field Reproduction by Wave-Field Synthesis », Journal of the Audio Engineering Society, vol. 43, no. 12, p. 1003-1012, décembre 1995

[Bonardi & al., 2015], A. Bonardi et P. Guillot, « Concevoir des traitements sonores ambisoniques en 2d et en 3d – l'exemple de Pianotronics 2", Actes des Journées d'Informatique Musicale, Canada, Montréal, 2015

[Braasch, 2005], J. Braasch, « A loudspeaker-based 3D sound projection using Virtual Microphone Control (ViMiC) », Proceedings of the 118th AES Convention, Preprint 6430, Barcelona, Spain.

[Busson, 2006], S. Busson, « Individualisation d'Indices Acoustiques Pour La synthèse Binaurale », Thèse de Doctorat en Mécanique, Physique Et Modélisation, Université de la méditerranée Aix-Marseille II, 2006

[Chapman & al., 2009], M. Chapman et P. Cotterell, « Towards A Comprehensive Account of Valid Ambisonic Transformations », Ambisonics Symposium, Graz, Autriche, juin 2009

[Chowning, 1977], J. M. Chowning, « The simulation of moving sound sources ». Computer Music Journal, vol. 1, no. 3, p. 48-52, juin, 1977

[Chowning & al., 1986], J. M. Chowning et D. Bristow, « FM theory and applications by musicians for musicians », Yamaha Music Foundation, October 1986

[Clark & al – 1958], H. A. M. Clark, G. F. Dutton et P. B. Vanderlyn, « The stereosonic recording and reproducing system », Journal of the Audio Engineering Society, vol. 6, no. 2, p. 102-117, décembre 1958 reprint from Proceedings of the Institution of Electrical Engineers, septembre 1957

[Colafrancesco, 2012], J. Colafrancesco, « L'ambisonie D'ordre Supérieur Et Son Appropriation Par Les Musiciens », Actes des Journées d'Informatique Musicale, Belgique, Mons, 2012

[Colafrancesco & al., 2013], J. Colafrancesco, P. Guillot, E. Paris, A. Sedes et A. Bonardi, « La bibliothèque HOA, bilan et perspectives », actes des journées d'informatique musical, France, Saint-Denis, 2013

[Cooper, 1970], G. Cooper, « Tetrahedral Ambiophony », Studio Sound, vol. 12, no. 6, p. 233-234, juin 1970

- [Cooper & al., 1972], D. H. Cooper et T. Shiga, « Discrete-Matrix Multichannel Stereo », *Journal of the Audio Engineering Society*, vol. 20, no. 5, p. 346-260, juin 1972
- [Cooper & al., 1989], D. H. Cooper et J. L. Bauck, « Prospects for Transaural Recording », *Journal of the Audio Engineering Society*, vol. 37, no. 1/2, janvier/février 1989
- [Craven & al., 1977], P. G. Craven et M. A. Gerzon, « Coincident Microphone Simulation Covering Three Dimensional Space and Yielding Various Directional Outputs », US Patent 4042779, août 1977
- [Dalenback & al., 1993], B-I. Dalenback, M. Kleiner et P. Svensson, « A Macroscopic View of Diffuse Reflection », Preprint of The 95th Audio Engineering Society Convention, New-york, Etats-unis, octobre 1993
- [de Boer, 1940], K. de Boer, « Stereophonic sound production », *Philips Technical Review*, vol. 5, p. 107-144, 1940
- [Daniel & al., 1998], J. Daniel, J-B. Rault et J-D. Polack, « Ambisonic Encoding of Other Audio Formats for Multiple Listening Conditions », Preprint of the 105th Audio Engineering Society Convention, San Francisco, Californie, Etats-Unis, septembre 1998
- [Daniel, 2001], J. 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, 2001
- [Daniel, 2003], J. Daniel, « Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format », *Proceedings of the Audio Engineering Society, 23rd International Conference*, Danemark, Copenhague, mai 2003
- [Danks, 1996], M. Danks, « The Graphics Environment for Max », *Proceedings of the International Computer Music Conference (ICMC 1996)*, Hong Kong, Chine, 1996
- [Davis, 2003], M. F. Davis, « History of Spatial Audio », *Journal of the Audio Engineering Society*, vol. 51, no. 6, p. 554-569, juin 2003
- [Eargle, 1971], J. M. Eargle, « Multichannel Stereo Matrix Systems: An Overview », *Journal of the Audio Engineering Society*, vol. 19, p. 552-559, juillet/août 1971
- [Einstein, 2001], A. Einstein, « La relativité – Théorie de la relativité restreinte et générale – La relativité et le problème de l'espace », traduction par Maurice Solovine, paru initialement aux éditions Gauthier-Villars en 1956 en langue française et publié pour cette présente édition aux Éditions Payot & Rivages à Paris en 2001

[Elen, 1991], R. Elen, « Whatever Happened to Ambisonics? », originellement publié dans AudioMedia Magazine en novembre 1991, accessible en ligne en septembre 2017 à l'adresse <http://www.ambisonic.net/>.

[Farina, 1998], Angelo Farina, « Software Implementation of B-Format Encoding and Decoding », Preprint of the 104th Audio Engineering Society Convention, Amsterdam, Pays-Bas, mai 1998

[Feldman, 1971], L. Feldman, « On Matrix Quadraphonic Systems », Audio Magazine, vol. 55, p. 20-28, octobre 1971

[Fellgett, 1974], P. B. Fellgett, « Ambisonic Reproduction of Directionality in Surround Sound Systems », Nature, vol. 252, p.534-538, décembre 1974

[Fellgett, 1975], P. B. Fellgett, « Ambisonics Part 1 – General system description », Studio Sound, Vol. 17, p. 20- 40, août 1975

[Fletcher, 1934], H. Fletcher, « Auditory Perspective – Basic Requirements », Electrical Engineering, p. 9-11, janvier 1934

[Fletcher, 1941], H. Fletcher, « Stereophonic Sound Film System – General Theory », Journal of Acoustical Society of America, vol. 13, 1941, p. 88-99

[Franquin & al., 1958], A. Franquin, M. Greg et J. De Mesmaeker, « Le Prisonnier du Bouddha », Le Journal de Spirou du n° 1048 au n° 1082, 1958

[Fuchs, 1995], A. Fuchs, « Plaidoyer Pour La Loi Normale », paru dans le supplément de la revue « Pour la Science », p. 17, 1995

[Gardner, 1995] W. Gardner, « Efficient Convolution Without Input-Output Delay », Journal of the Audio Engineering Society, vol. 43, vol. 3, p. 127-136, mars 1995

[Gaskell & al., 1977], P. S. Gaskell et P. A. Ratliff, « Quadraphony Developments in Matrix H Decoding », BBC Research Department Report, février 1977

[Gerzon, 1970a], M. A. Gerzon, « The Principles of Quadraphonic Recording Part One – Are Four Channels Really Necessary? », Studio Sound, Vol. 12, p. 338-342, août 1970

[Gerzon, 1970b], M. A. Gerzon, « The Principles of Quadraphonic Recording Part 2 – The Vertical Element », Studio Sound, Vol. 12, p. 300-384, septembre 1970

[Gerzon, 1971a], M. A. Gerzon, « Experimental Tetrahedral Recording: Part One », Studio Sound, Vol. 13, p. 396-398, août 1971

[Gerzon, 1971b], M. A. Gerzon, « Experimental Tetrahedral Recording: Part Two », Studio Sound, Vol. 13, p. 472-475, septembre 1971

- [Gerzon, 1971c], M. A. Gerzon, « Experimental Tetrahedral Recording: Part Three », *Studio Sound*, Vol. 13, p. 510-515, octobre 1971
- [Gerzon, 1973], M. A. Gerzon, « Periphony With-height Sound Reproduction », *Journal of the Audio Engineering Society*, vol. 21, no. 10, janvier 1973
- [Gerzon, 1974], M. A. Gerzon, « Surround-sound psychoacoustics, Criteria for the design of matrix and discrete surround-sound systems », *Wireless World*, décembre 1974
- [Gerzon, 1975a], M. A. Gerzon, « The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound », Preprint of the 50th Audio Engineering Society Convention, Londres, Royaume-Uni, mars 1975
- [Gerzon, 1975b], M. A. Gerzon, « Dummy Head Recording », *Studio Sound*, Vol. 17, p. 42-44, mai 1975
- [Gerzon, 1975c], M. A. Gerzon, « Ambisonics Part 2 – Studio techniques », *Studio Sound*, Vol. 17, p. 24- 40, août 1975
- [Gerzon, 1977], M. A. Gerzon, « Criteria for evaluating surround sound systems », *Journal of the Audio Engineering Society*, vol. 25, no. 6, p. 400-408, juin 1977
- [Gerzon, 1981], M. A. Gerzon, « Sound Reproduction Systems », UK Patent 2,073,056A, octobre 1981
- [Gerzon, 1985], M. A. Gerzon, « Ambisonics in Multichannel Broadcasting and Video », *Journal of the Audio Engineering Society*, vol. 33, no. 11, p. 849-871, novembre 1985
- [Gerzon, 1992a], M. A. Gerzon, « General Metatheory of Auditory Localization », Preprint 3306 of the 92nd Audio Engineering Society Convention, Autriche, Vienne, mars 1992
- [Gerzon & al., 1976], M. A. Gerzon et J. M. Woram, « Blumlein Stereo Microphone Technique », *Journal of the Audio Engineering Society*, vol. 24, no. 1, p. 36-38, février 1971
- [Gerzon & al., 1992], M. A. Gerzon et G. J. Barton, « Ambisonic Decoder for HDTV », Preprint 3345 of the 92nd Audio Engineering Society Convention, Autriche, Vienne, mars 1992
- [Golberg & al., 1982], A. Goldberg et R. Flegal, « ACM President's Letter: Pixel Art », *Communication of the Association for Computing Machinery*, vol. 25, p. 861–862, 1982
- [Gonzales & al., 2007], R.C. Gonzales & R. E. Woods, « Digital image processing, 3rd Edition », Pearson Education Inc., Upper Saddle River, New Jersey, Etats-Unis, juillet 2007
- [Grahn, 2013], Emma Grahn, « Modern Pixel Art Games », Diplôme de candidat en technologie des médias, Blekinge University of Technology, 2013



- [Guillot, 2013], P. Guillot, « Les traitements musicaux en ambisonie », Mémoire de Master 2, Université Paris 8, Saint-Denis, France, 2013
- [Guillot, 2014], P. Guillot, « Une nouvelle approche des objets graphiques et interfaces utilisateurs dans Pure Data », Actes des Journées d'Informatique Musicale, Bourges, France, 2014
- [Guillot & al., 2013], P. Guillot, E. Paris et M. Deneu, « La bibliothèque de spatialisation HOA pour Max/MSP, Pure Data, VST, FAUST... », Revue Francophone d'Informatique Musicale [En ligne], Saint-Denis, France, 2013
- [Hughes & al., 2013], J. F. Hughes, A. van Dam, M. McGuire, D. F. Sklar, J. D. Foley, S. K. Feiner et K. Akeley, « Computer Graphics: Principles and Practice, 3rd Edition », Pearson Education Inc., Upper Saddle River, New Jersey, Etats-Unis, 2013
- [Itoh, 1971], R. Itoh, « Proposed Universal Encoding Standards for Compatible Four-Channel Matrixing », présenté à 41st Convention of the Audio Engineering Society, New York, octobre 1971
- [ITU, 2012], International Telecommunication Union, « Multichannel stereophonic sound system with and without accompanying picture », Rec. ITU-R BS.775-3, 2012
- [Jessel, 1973], M. Jessel, « Acoustique théorique, propagation et holophonie », Masson, Paris, 1973
- [Kaiser, 2011], F. Kaiser, « A hybrid approach for three-dimensional sound spatialization », Séminaire IEM algorithmes en acoustique et informatique musicale, Graz, mai 2011
- [Kohsaka & al., 1972], O. Kohsaka, E. Satoh et T. Nakayama, « Sound- image localization in multichannel matrix reproduction », Journal of the Audio Engineering Society, vol. 20, No. 7, p. 542-548, septembre 1972
- [Kostadinov & al., 2010], D. Kostadinov, J. D. Reiss, V. Mladenov, « Evaluation of Distance Based Amplitude Panning for Spatial Audio », Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal (ICASSP), Dallas, Etats-Unis, mars 2010
- [Kronlachner, 2013] M. Kronlachner, « Ambisonics plug-in suite for production and performance usage », Linux Audio Conference 2013, Graz, Autriche, mai 2013
- [Kronlachner, 2014a] M. Kronlachner, « Plug-in Suite for Mastering the Production and Playback in Surround Sound and Ambisonics », Gold Award at AES Student Design Competition, Berlin, Allemagne, avril 2014

[Kronlachner, 2014b], M. Kronlachner et F. Zotter, « Spatial Transformations for the alteration of Ambisonic Recordings », Master thesis, University of Music and Performing Arts, Graz, Autriche, juin 2014

[Kronlachner & al., 2014], M. Kronlachner et F. Zotter, « Spatial Transformations for the enhancement of Ambisonic Recordings », 2nd International Conference on Spatial Audio, Erlangen, Allemagne, février 2014.

[Laplace, 1814], P. S. Laplace, « Théorie Analytique des Probabilités », Seconde Édition, M<sup>me</sup> V<sup>e</sup> Courcier, Imprimeur-Librairie pour les Mathématiques et la Marine, Quai des Augustins, n° 57, Paris, 1814

[Larcher, 2001], V. Larcher, « Techniques de spatialisation des sons pour la réalité virtuelle », Thèse de Doctorat en Acoustique, Université Paris VI, 2001

[Lossius & al., 2009], T. Lossius, P. Baltazar, T. de la Hogue, « DBAP Distance-Based Amplitude Panning », Proceedings of the International Computer Music Conference, ICMC 2009, Montreal, Canada, 2009

[Lyon, 2006], R. F. Lyon, « A Brief History of ‘Pixel’ », proceedings the IS&T/SPIE Symposium on Electronic Imaging, États-Unis, Californie, janvier 2006

[Makita, 1960], Y. Makita, « On The Directional Localisation of Sound in The Stereophonic Sound Field », This article is based upon an article read by him at the 12th Meeting of the Technical Committee of the E.B.U., Monte Carlo, octobre, 1960

[Malham, 1990], D. G. Malham, « Ambisonics A Theory for Low Cost High Precision Three-dimensional Sound Diffusion », Proceedings of the ICMC 1990, p. 118-120, Royaume-Uni, Glasgow, 1990

[Malham, 1992], D. G. Malham, « Experience with Large Area 3-D Ambisonic Sound Systems », Proceedings of the Institute of Acoustics Autumn Conference on Reproduced Sound 8, p. 209-216, Royaume-Uni, Windermere, 1992

[Malham, 1999], D. G. Malham, « High Order Ambisonic Systems for the Spatialisation of Sound », Proceedings of the ICMC 1999, p. 484-487, Chine, Pékin, 1999

[Malham, 2003], D. G. Malham, « Higher Order Ambisonic Systems », abstracted from « Space in Music – Music in Space », an Mphil thesis by Dave Malham, submitted to the University of York in April 2003, revised and passed in December 2003, Royaume-Uni, décembre 2003

[Malham & al., 1991], D. G. Malham et R. Orton, « Progress in the Application of Ambisonic Three Dimensional Sound Diffusion Technology to Computer Music » Proceedings of the ICMC 1991, p. 467-470, Canada, Montréal, 1991

[Malham & al., 1995], D. G. Malham et A. Myatt, « 3-D Sound Spatialization using Ambisonic Techniques », Computer Music Journal, vol. 19, no. 4, p. 58-70, hiver 1995

[Malt & al., 2008], M Malt et E. Jourdan, « Zsa.Descriptors- a library for real-time descriptors analysis », Proceedings of the Sound and Music Computing Conference (SMC 2008), Berlin, Allemagne, juillet 2008

[Malt & al., 2009], M Malt et E. Jourdan, « Real-Time Uses of Low Level Sound Descriptors as Event Detection Functions Using the Max/MSP Zsa.Descriptors Library », in SBCM 2009, Recife, Brésil, 2009

[Møller, 1992], H. Møller, « Fundamentals of Binaural Technology », Applied Acoustics, vol. 36(3/4), p. 171-218, 1992

[Moore, 2013], B. C. Moore, « An introduction to the psychology of hearing », 6<sup>ème</sup> édition, Brill, Université de Cambridge, Royaume-Uni, 2013

[N.R.D.C., 1977], « Encoding Standards for NRDC Universal HJ Surround-Sound Encoding System: System UHJ », National Research Development Corporation NRDC/FCC 2, novembre 1977

[Nachbar & al., 2016] C. Nachbar, F. Zotter, E. Deleflie et A. Sontacchi, « AmbiX – A suggested Ambisonics format », Ambisonics Symposium, Lexington, Etats-Unis, juin 2016 (version révisée en 2016).

[Neukom, 2007] M. Neukom, « Ambisonic Panning », Presented at the 123rd Convention of the Audio Engineering Society, New-York, Etats-Unis, octobre 2007

[Neukom & al., 2007] M. Neukom et J. C. Schacher, « Ambisonic Equivalent Panning », Proceedings of the International Computer Music Conference ICMC, Copenhagen, août 2007

[Nicol & al., 1998], R. Nicol et M. Emerit, « 3D Binaural Sound Reproduction using a Virtual Ambisonic Approach », Proceedings of the 6th Int. Conference on Digital Audio Effects (DAFX98), Espagne, Barcelone, novembre 1998.

[Orlarey & al., 2006], Y. Orlarey, A. Graf et G. Kersten, « DSP Programming with Faust, Q and SuperCollider », Linux Audio Conference, 2006

[Paris, 2013], E. Paris, « Des interfaces pour la mise en espace du son avec la bibliothèque HOA », Mémoire de Master 2, Université Paris 8, Saint-Denis, France, 2013

[Poletti, 2005], « Three-Dimensional Surround Sound Systems Based on Spherical Harmonics », Journal of the Audio Engineering Society, vol. 53, No. 11, p. 1004-1025, novembre 2005

[Poletti, 2007], « Robust Two-Dimensional Surround Sound Reproduction for Nonuniform Loudspeaker Layouts », Journal of the Audio Engineering Society, vol. 55, No. 7/8, p. 598-610, juillet/août 2007

[Pomberger & al., 2011], H. Pomberger et F. Zotter, « Warping of 3d Ambisonic Recordings », Ambisonics Symposium, Kentucky, États-Unis, juin 2011

[Pomberger & al., 2012], H. Pomberger et F. Zotter, « Ambisonic panning with constant energy constraint », Fortschritte der Akustik – DAGA 2012, Darmstadt, Allemagne, mars, 2012.

[Puckette, 2002], M. Puckette, « Using Pd as a score language », Proceedings of the International Computer Music Conference (ICMC 2002), p. 184-187, Gothenburg, Suède, 2002

[Pulkki, 1997], V. Pulkki, « Virtual Sound Source Positioning Using Vector Base Amplitude Panning », Journal of the Audio Engineering Society, vol. 45, no. 6, p. 456-466, juin 1997

[Pulkki, 1999], V. Pulkki, « Uniform Spreading of Amplitude Panned Virtual Sources », Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, États-Unis, octobre 1999

[Pulkki & al., 2001], V. Pulkki et M. Karjalainen, « Localization of amplitude-panned virtual sources I – Stereophonic Panning », Journal of the Audio Engineering Society, vol. 49, no. 9, p. 739-752, septembre 2001

[Rayleigh, 1876], J. W. S. (Lord) Rayleigh, « On Our Perception of the Direction of a Source of Sound », Proceedings of the Musical Association, 2nd Session, p. 75-84, avril 1876

[Rayleigh, 1907], J. W. S. (Lord) Rayleigh, « On our perception of sound direction », Philosophy Magazine, vol. 13, p. 214-32, 1907

[Roads, 1996], C. Roads, « The Computer Music Tutorial », MIT Press, Cambridge, États-Unis, Massachusetts, 1996

[Roads, 2001], C. Roads, « Microsound », MIT Press, Cambridge, États-Unis, Massachusetts, 2001

[Schader, 2010], J. C. Schacher, « Seven Years of Icest Ambisonics Tools for Maxmsp – A Brief Report », Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics, France, Paris, mai 2010

- [Schacher & al., 2006] J. C. Schacher, P. Kocher, « Ambisonics Spatialization Tools for Max/MSP », Proceedings of the 2006 International Computer Music Conference, USA, New Orleans, 2006
- [Scheiber, 1971a], P. Scheiber, « Four Channels and Compatibility », Journal of the Audio Engineering Society », vol. 19, p. 267-279, avril 1971
- [Scheiber, 1971b], P. Scheiber, « Suggested Performance Requirements for Compatible Four-Channel Recording », Journal of the Audio Engineering Society, vol. 19, p. 647-650, septembre 1971
- [Schroeder, 1962], M. R. Schroeder, « Natural Sounding Artificial Reverberation », Journal of the Audio Engineering Society, vol. 2, no. 3, p. 219-223, septembre 1962
- [Sèdes, 2015], A. Sèdes, « Approche musicale de la décorrélation microtemporelle dans la bibliothèque HOA », Actes des Journées d'Informatique Musicale, Canada, Montréal, 2015
- [Sèdes & al., 2003a], sous la direction d'A. Sèdes, « Espaces sonores : Actes de recherches », Transatlantiques (éditions musicales), Paris, 2003
- [Sèdes & al., 2003b], A. Sèdes, B Courribet et J-B. Thiébaud, « Egosound, an egocentric interactive and realtime approach of sound space », Proceedings of the 6th Int. Conference on Digital Audio Effects (DAFX-03), Royaume-Uni, Londres, septembre 2003
- [Sèdes & al., 2014], A. Sèdes, P. Guillot et E. Paris, « The HOA library, review and prospect », proceedings of the ICMC-SMC 2014, Grece, Athenes, 2014
- [Shapiro & al., 1995], H. Shapiro C. Glebbeek, « Jimi Hendrix: Electric Gypsy », St. Martin's Griffin, août 1995
- [Shoup, 2001], R. Shoup, « SuperPaint: An Early Frame Buffer Graphics System », IEEE Annals of the History of Computing, p. 30-37, avril 2001
- [Smith, 2007], J. O. Smith, « Introduction to Digital Filters with Audio Applications », <http://ccrma.stanford.edu/~jos/filters/>, livre en ligne accédé en mai 2017, édition 2007
- [Smith, 2010], J. O. Smith, « Physical Audio Signal Processing », <http://ccrma.stanford.edu/~jos/pasp/>, livre en ligne accédé en mai 2017, édition 2010
- [Sontacchi & al., 2001], A. Sontacchi et R. Höldrich, « Further investigations on 3D sound fields using distance coding », Proceedings of the DAFX-01, Limerick, Ireland, décembre 2001
- [Stark, 1971], C. Stark, « One Approach to Four-Channel Sound », Audio Magazine, vol. 55, p. 30, octobre 1971

- [Steinberg & al., 1934], J. C. Steinberg et W. B. Snow, « Auditory Perspective – Physical Factors », *Electrical Engineering*, p. 12-17, janvier 1934
- [Steinke, 1996], G. Steinke, « Surround Sound: The New Phase », Preprint of the 100th Audio Engineering Society Convention, Copenhagen, Danemark, mai 1996
- [Trausmuth & al., 2006], R. Trausmuth, C. Dusek et Y. Orlarey, « Using Faust for FPGA Programming », Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06), Canada, Montréal, septembre, 2006
- [Vaggione, 2001], H. Vaggione, « Composing Musical Spaces by Means of Decorrelation of Audio Signals », proceedings of the DAFx Conference on Digital Audio Effects, Ireland, Limerick, 2001
- [Vaggione, 2002], H. Vaggione, « Décorrélation microtemporelle, morphologies et figurations spatiales », Actes des journées d'informatique musical, France, Marseille, mai 2002
- [Wakefield, 2006] G. Wakefield, « Third-Order Ambisonic Extension for Max/MSP with Musical Applications », Proceedings of the 2006 International Computer Music Conference, p. 123-126, New Orleans, US, 2006.
- [Zotter & al., 2010], F. Zotter, M. Frank et A. Sontacchi, « The Virtual T-Design Ambisonics-Rig Using VBAP », Proceedings of the 1st EAA Congress on Sound and Vibration, Slovénie, Ljubljana, septembre 2010
- [Zotter & al., 2012], F. Zotter et M. Frank, « All-Round Ambisonic Panning and Decoding », *Journal of the Audio Engineering Society*, vol. 60, no. 10, p. 807-820, octobre 2012
- [Zotter & al., 2013], F. Zotter, M. Frank et H. Pomberger, « Comparison of energy-preserving and all-round Ambisonic decoders », *Fortschritte der Akustik, AIA-DAGA*, mars 2013.