



HAL
open science

Approaching Spatial Audio Processing by Means of Decorrelation and Ring Modulation in Ambisonics

Paul Goutmann, Alain Bonardi

► **To cite this version:**

Paul Goutmann, Alain Bonardi. Approaching Spatial Audio Processing by Means of Decorrelation and Ring Modulation in Ambisonics. Sound and Music Computing, Université Jean Monnet, Saint-Étienne; GRAME; INRIA, Jun 2022, Saint-Étienne, France. 10.5281/zenodo.6797487 . hal-03973343

HAL Id: hal-03973343

<https://hal.science/hal-03973343>

Submitted on 4 Feb 2023

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

Approaching Spatial Audio Processing by Means of Decorrelation and Ring Modulation in Ambisonics

Paul Goutmann

ArTeC
Université Paris 8
Musidanse/CICM
paul@goutmann.net

Alain Bonardi

Université Paris 8
Musidanse/CICM
alain.bonardi@univ-paris8.fr

ABSTRACT

This paper presents an approach to spatial audio processing in ambisonics as a part of musical practice. We present the control of two treatments - the decorrelation and ring modulation in ambisonic - to express how the design of digital objects feeds into the composition of spatial attributes of sound and the musical idea of space. The decorrelation factor gives the possibilities rangy from localizable sources to diffuse sound field, and the modulation factor enables the variation between static situations to complex movement in space. These parameters encapsulate a multitude of operations on the signal as well as ways to think about interactions with spatial attributes of sound. Based on an experimental methodology of research-creation combining audio experiments with development, this approach shows how rich it is to explore different layers of our compositional environments to propose new paradigms of spatial audio. This has resulted in a new version of the decorrelation and ring modulation inside the Faust library *abc*.

1. INTRODUCTION

The spatial dimension of sound is essential in the field of contemporary music creation, especially in electroacoustic and electronic music. Today, concert halls and studios are equipped with multipoint diffusion systems. Since the 1990s, software implementing sound spatialization models have been created and shared by the community of composers, computer music designers and audio engineers [1]. The development of these technologies feeds new paradigms of representation and ways to think about space in music [2]. Most of the spatialization techniques are based on point sources in close relationship with spatial hearing models [3] [4] [5]. Accessible musical operations for creators are the placement of sources, designing movements and room acoustic simulation. However, other approaches of interaction with spatial attributes of sound exist [6] [7] [8]. The project *HOALibrary* contributed to those approaches in synergy with many years of research and creation about the spatiality of sound at CICM. The *Cen-*

tre d'Informatique et de Création Musicale created by Horacio Vaggione in 1992 accommodates researches which have interactions¹ between creation, technologies and computer sciences especially spatialization since the beginning of the 2000s [9]. Developed by J. Colafrancesco, P. Guillot and E. Paris between 2012 and 2015 in the context of the Labex Arts H2H projects, *HOALibrary* is a library of spatial audio processing tools in ambisonics [10]. One of the particularities of this library for Max and Pd is the insertion of signal processes (well known in electroacoustic and mixed music) such as granulation, ring modulation or decorrelation into the ambisonic model. Unfortunately the Pd version is obsolete since 2019. In order to formalize twenty years of work on mixed music and to address the obsolescence of *HOALibrary*, Bonardi began to transcribe treatments of the object *hoa.process~* in Faust (as well as ambisonic basics and multichannel processes) grouped in the library *abc.lib* [11]². In this context, we began new developments and fine-tuning around the decorrelation and the ring modulator in ambisonics lean on a research-creation practice³. The musical interest of those two treatments is in part linked to the fact that they allow complex sensitive results through the manipulation of just one parameter, the factor. The decorrelation produces progressively diffuse sound fields from one mono source (Sound 6.1) and the ring modulation creates the sensation of complex movement from one mono source (Sound 6.2). We refer to sound examples all along the article which are indexed in the appendix (6) and available on the platform *HAL* of the *Centre pour la Communication Scientifique* of the CNRS.

¹ The interactions take place in research projects where the activity of creation permits to test in sensitive and qualitative ways the results of audio digital processes. <https://cicm.mshparisnord.fr> Accessed on: April 6, 2022

² The library is currently compiled for Max and Pd and available here with the Faust code: <https://github.com/alainbonardi/abc.lib> Accessed on: April 6, 2022

³ A state of the art of 'research-creation' or 'artistic research' is beyond the scope of this paper. 'Research-creation' refers to a large set of practices and methods. For this paper, its doing research in interaction with creation in an academic context. Both part emerges from and merges with the other. Composing is here a manner to handle tools (theoretical and technological) and to produce knowledge in music, musicology and computer music sciences. Composing as a research activity is at the core of CICM and ArTeC. We explain in fact the methodology in section 3.

2. INTERACTING WITH SPATIALITY

In the practices of electroacoustic, mixed and electronic music, the work on the spatial dimension of sound cannot be reduced to spatialization [12] [13] [14] [15]⁴. Designing treatments as a way to design instruments is at the core of musical practices in digital environments. The *HOALibrary* project approached space in terms of spatial process of sound. The library has been conceived to give access to musicians to the operations in the framework of the HOA (High Order Ambisonic) model. In this section we will first present the way to create treatments in *HOALibrary* and their transcriptions in the Faust language. Then we will present the design of macro parameters as controllers for digital processes. Finally, we will go further and expose how the design of functions is an important step to give access to operations.

2.1 Spatial Processing

The B-Format proposed by M.Gerzon and HOA proposed by J.Daniel are techniques to record, synthesize, process and reproduce acoustic fields. These techniques are based on three steps: encoding, optimisation and decoding. The encoding principle is based on the description of the spherical functions (in 3D) or circular functions (in 2D) corresponding to the acoustic field at one point of space using an ensemble of reference functions. These reference functions are called 'harmonics', their number depends on the encoding order. The more harmonics are used, the more precise the encoding will be. In a 2D encoding there are $2(N+1)$ harmonics and in 3D encoding there are $(N+1)^2$ harmonics. They are organized in order m (line) and degree l (row) as $-l \leq m \leq +l$ and can be represented as in the Figure 1.

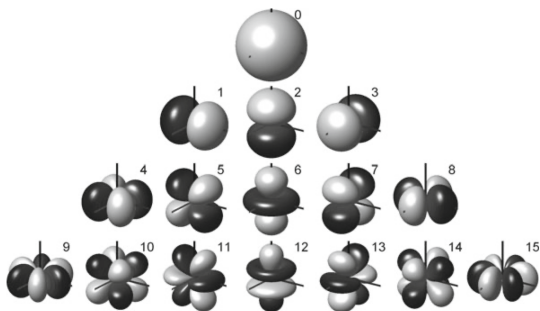


Figure 1. Spherical harmonics indexed by the ACN (*Ambisonic Channel Number*) in [16].

The decoding stage consists of calculating the projection of the spherical function on the loudspeaker setup. In other words, calculating the gain sent to each speaker depending of a diffusion system which needs to include more speakers (set on a circle or a dome) than the number of harmonics. Processes in *HOALibrary* consist of applying digital audio treatment to each harmonic. Inspired by the object *poly~*,

⁴ These four papers are written by composers (except the last) which cover electroacoustic, mixed music, electronic music and musical installation. They are focused on their own approach to the spatial dimension of sound, and gives an idea of the range of approaches.

the objects *hoa.2d.process~* or *hoa.3d.process~* - takes as first argument the order of encoding - instances a patch of treatment for each harmonic [17]. For example, the object *hoa.2d.process~* in the Figure 2, instances the patch *hoa.fx.decorrelation~* (as in Figure 3 for Y_0) fifteen times corresponding to the numbers of harmonics in HOA 2D 7th order. The process parameters are generally proportional to an harmonic coefficient which depend on the order, the degree and the number of harmonics. Since 2020 patches instantiated by the *hoa.process~* are grouped in *abc.lib* in Faust. This functional language for signal processing gives the opportunity to clear operations especially with the primitive function 'par' which creates indexed operations in parallel. Multiple windows encapsulated in Max or Pd correspond to a few lines in Faust that can be edited using an embedded compiler.

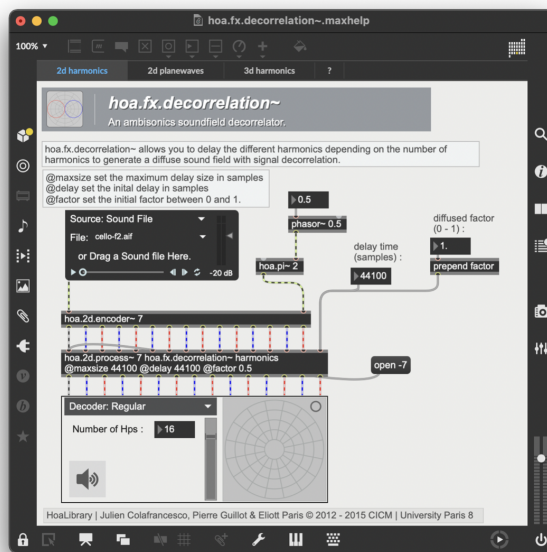


Figure 2. Help patch of the *hoa.fx.decorrelation~* spatial process.

2.2 The Factor

Designing treatments is a way to build a part of our 'composable space'. This notion originally developed by Horacio Vaggione is the environment allowing composers to generate and control sound morphologies. Composable spaces are multidimensional and are themselves morphologies that can be designed [14]. The composable space offers controllers to operate on and compose sound. In *HOALibrary* and *abc.lib*, the controllers which are designed for decorrelation and ring modulation are accessible parameters, and notably the 'factor'. The 'factor' parameter is a meta-parameter which controls the depth of the treatment. The idea of Guillot and Paris was to begin to process higher harmonics for low values of the 'factor' up to the harmonic 0 for the maximum of the 'factor'. When these treatments are applied after a mono source encoding, the result can be described by a transition from point source to a diffuse sound field with a color chart behind

the works of Kendall and Vaggione [20] [21], the decorrelation in *HOALibrary* (and then in *abc.lib*) is based on the dephasing of the harmonics after the encoding process with the insertion of delay lines. Inserting some delays between harmonics breaks the correlation of post-encoding signals and gives a rich perceptive result of a diffuse sound field. The decorrelation has initially two parameters: 'delay', 'factor'. We add two more: type of distribution of the delays ('functiontype') and feedback ('fdbk'). The factor determines how many harmonics will be delayed, the delay is the maximum delay time applied to the harmonics, the distribution of delay is the function used to spread the delays between the harmonics. The decorrelation is defined for each spatial component among:

$$P = \begin{cases} 2 \cdot N + 1 & \text{in 2D,} \\ (N + 1)^2 & \text{in 3D} \end{cases} \quad (1)$$

where P is the number of harmonics at the ambisonic order N , by the delay d_i :

$$d_i = \begin{cases} 0 & \text{if } fa < 1 - \frac{i+1}{P}, \\ \text{delay} \cdot f\left(\frac{i+1}{P}\right) & \text{else} \end{cases} \quad (2)$$

where $f(x)$ is the function of distribution called *abc_{th}* in the library. The blockdiagram of the decorrelation in Figure 4 shows that we hear the signal delayed or not according to the factor and the index of the harmonic, thanks to the two envelopes *env1* and *env1c* - when one is equal to zero the other is equal to one.

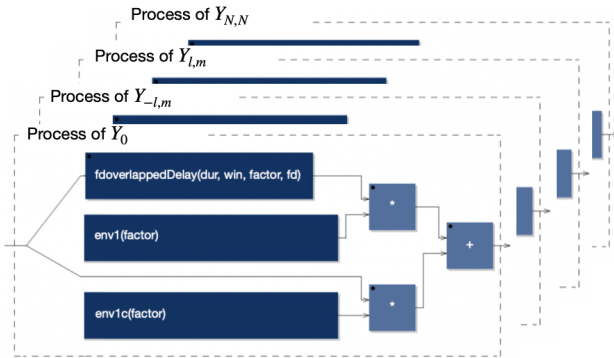


Figure 4. Blockdiagram of the decorrelation in *abc.lib*

3.1.1 The stairs problem

In this situation, if the factor value is in the interval $[0, \frac{1}{P}]$ the last harmonic is delayed with the maximum delay time and all the others won't be delayed. Therefore, as the factor increases, the second to last harmonic will be delayed starting from a factor equal to $\frac{2}{P}$, the third to last will be delayed starting from $\frac{3}{P}$, etc. The sensitive result isn't a progressive generation of a diffuse sound field but a progressive insertion of static delay lines (Sound 6.3). Figure 5 illustrates the coefficient of delay time applied to each harmonic when the factor increases in 2D 3th order: when the factor increases we first hear the last harmonic delayed ($Y_{-3,3}$), then second to the last ($Y_{3,3}$) and this until the first (Y_0).

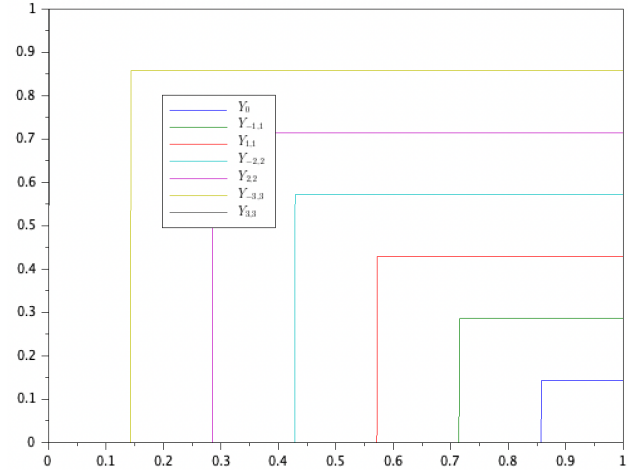


Figure 5. Plot of the delay coefficient per harmonic versus the factor.

In consequence, we decided to introduce the factor in the delay equation of each harmonic. That leads to the reduction to the gap between the delays of harmonics already delayed and harmonics that begin to be delayed when the factor increases. Moreover, the delays progressively increase as the factor increases as illustrated in the Figure 6 and that amplifies the effect of progressive generation of diffuse sound field (Sound 6.4). We have chosen to duplicate each signal corresponding to the harmonics and add a gate as a binary wet/dry controlled by the factor. The delay time is equal to 0 when the factor is under a threshold and the feedback introduces coloration and saturation. So when the factor is under the threshold we hear the signal without delay line and when the factor is higher the threshold we hear the signal with a delay line.

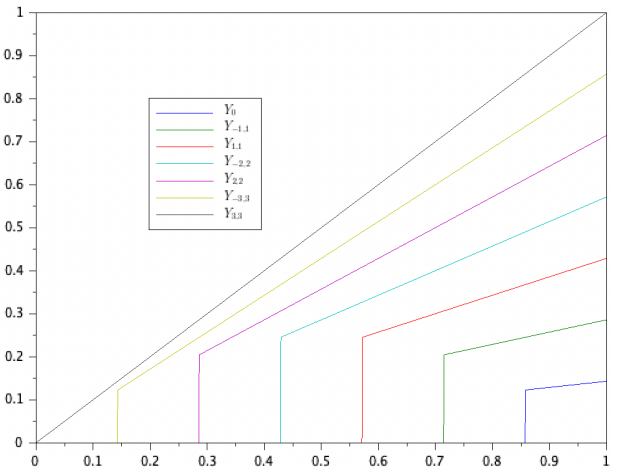


Figure 6. Plot of the delay coefficient per harmonic versus the factor when including the factor in the delay equation.

3.1.2 The distribution choice

The calculation of the delay time for each harmonic depends on the factor, the maximum delay time and the coefficient of the harmonic. The higher the order of the harmonic is, the deeper the treatment. But the delay time

applied if the factor is higher than the threshold increases linearly from the first harmonic to the last. Based on experiments on graphic user interfaces to control the decorrelation [18], we proposed to control the type of distribution with 21 functions called abc_{th} that are in the range $[0, 1]$ when $x \in [0, 1]$ (Sound 6.5).

3.2 Ring Modulation

Ring modulation consists in modulating a bipolar signal by another bipolar signal. In ambisonics, ring modulation consists in modulating each harmonic with a sinusoidal signal and has two parameters: the factor of modulation and the frequency of the carrier. If the factor is greater than the threshold:

$$\frac{(P - i - 1)}{P} \quad (3)$$

the i_{th} ring modulator is active with a carrier c of:

$$c = f_0 \cdot \frac{(i + 1)}{P} \quad (4)$$

where P is the number of harmonics. If the factor is under the threshold, the original signal is provided. Therefore, ring modulators are progressively revealed when the factor increases. For example, after an encoding at third order in 2D, there are seven harmonics. The harmonic 0 will have a modulation with a carrier of $\frac{f_0}{7}$ and will be modulated between a factor of $[\frac{6}{7}, 1]$. Continuing the example, if the frequency of modulation is 10Hz the frequency of modulation of each harmonic will take the value shown in the Table 1 (Sound 6.2).

Y_{ith}	$freq_{mod}$
Y_0	1.42 Hz
$Y_{-1,1}$	2.85 Hz
$Y_{1,1}$	4.28 Hz
$Y_{-2,2}$	5.71 Hz
$Y_{2,2}$	7.14 Hz
$Y_{-3,3}$	8.57 Hz
$Y_{3,3}$	10.0 Hz

Table 1. Modulation frequency per harmonic for 2D 3th order

The sensitive result of this process ranges from complex movements to timbre modification thanks to the behaviour of the ring modulation with a frequency modulation higher than 20 Hz.

3.3 FX/SYN

There are two modes of applying in *abc.lib*: FX and SYN. FX mode applies the process to the harmonics post-encoding. The idea is to be able to control the depth of the treatment on the sound field from a directional sound field to a diffuse sound field. The idea of the SYN mode is that it doesn't matter if the signal is not ambisonic encoded before applying the treatment because the process distorts the coherence of the intermediate representation of space.

SYN mode is thought to synthesize directly a diffuse sound field from one mono source. But when the factor is equal to zero, the same signal is sent to all the decoder inputs as if a localized sound field was synthesized the source is localized in one direction ⁶: 45°. For example, in the first order case, if the same signal is sent to all the entries of a decoder it get close to an encoding of a source at 45° (when the source angle is 45°, the output of $Y_{-1,1}$ and $Y_{1,1}$ are 0.707) and the intermediate representation of space (circular function) has shown in the Figure 7.

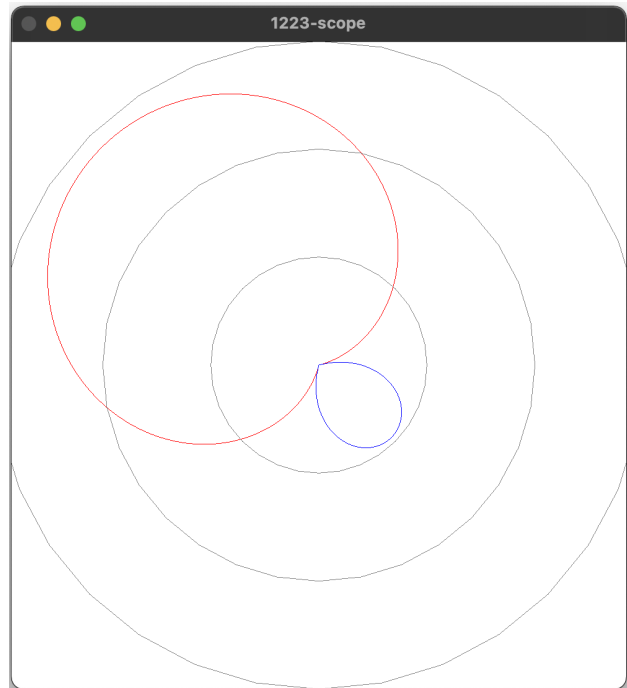


Figure 7. Intermediate representation of space with Y_0 , $Y_{-1,1}$ and $Y_{1,1}$ equal to one. In red the positive and in blue the negative part of the function.

Consequently, we decided to cut all harmonics except the first when the factor is equal to 0. The objective was to create a strong mono when the factor is equal to zero and to go progressively to a diffuse sound field. We apply gain compensation and progressively fade in the other harmonics when the factor increases between $[0, \frac{1}{n}]$. The Figure 8 shows the gain compensation of the Y_0 per ambisonic order (until 7) versus the factor.

4. CONCLUSION

In this article we highlighted the potential of 'treatment' oriented approach for spatial audio processing in *HOALibrary* and *abc.lib*. We presented the richness of a research-creation methodology based on an iterative prototyping loop guided by listening with two examples: decorrelation and ring modulation in ambisonics. In the part 2 'Interacting with spatiality' we first provide a summary on ambisonics and the approach of space in *HOALibrary*. In the section 2.2, we presented a new paradigm to compose and

⁶ In Head-Related Coordinate System, and the angle is given in anti-clockwise

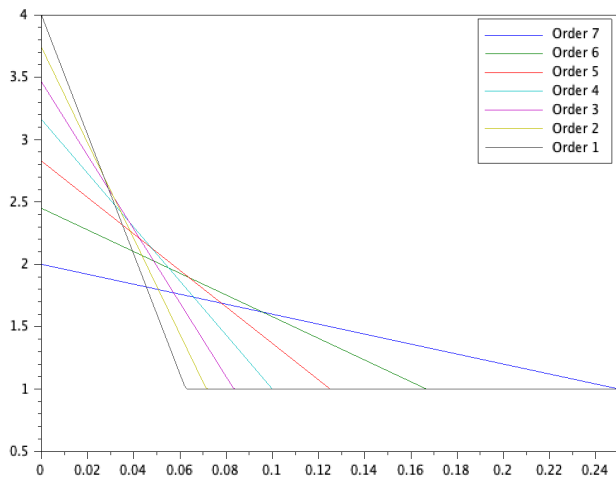


Figure 8. gain compensation per ambisonic order versus the factor

control spatial attributes of sound based on the design of macro-parameters. We then suggested that attention given to the access to the functions (the controllers) take a part in the design of that Vaggione calls our 'composable space'. Based on two examples, we explained our methodology in part 3 'Designing digital objects'. First, we presented in the section 3.1 the fine tunings on the delay time equation, then in the section 3.1.2 the different types of distributions between the delay lines that we use in the library. We presented in the section 3.2 the ring modulation and a way to distribute all the frequencies of modulation to the spatial harmonics. Finally, we presented two modes of treatment in *HOALibrary* and *abc.lib*: FX and SYN ; we explain our choice to use only the first spatial component when the factor is equal to zero for the SYN mode. In future work, we plan to investigate other treatments of the library⁷ using the same methodology, combining tests and musical experiments to improve and tune the parameters/meta-parameters of the processes. Furthermore, we will investigate the particularity of the ambisonic situation in different multi-channel situations. If processes of *abc.lib* are applied to the signals associated with harmonics or other multi-channel signals, it will be the same mechanism. Nevertheless, we could imagine processes working per order and per degree or other logics that we would need to design.

Acknowledgments

This research has been supported by the EUR ArTeC financed by the French National Agency for Research ANR through the PIA ANR-17-EURE-0008. The author would like to thank Atau Tanaka for his advice and proof-reading. The author would like to thank also the *Maison des Sciences de l'Homme Paris Nord* for hosting this work and specially Anne Sèdes for her advice.

⁷ like granulation, wider and rotation of the sound field.

5. REFERENCES

- [1] J. Garcia, T. Carpentier, and J. Bresson, "Interactive-compositional Authoring of Sound Spatialization," *Journal of New Music Research*, vol. 46, no. 1, pp. 74–86, nov 2016.
- [2] D. Fober, J. Bresson, P. Couprie, and Y. Geslin, "Les nouveaux espaces de la notation musicale," in *Journées d'Informatique Musicale*. Montréal, Canada: Faculté de musique de l'Université de Montréal, 2015. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01160759>
- [3] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, Mass: MIT Press, 1983.
- [4] B. Moore, *An Introduction to the Psychology of Hearing*, 6th ed. Brill Academic Pub, 2013.
- [5] J. Colafrancesco, "Spatialisation de sources auditives étendues : applications musicales avec la bibliothèque HOA," phdthesis, Université Paris 8 Vincennes Saint-Denis, 2015.
- [6] T. Pihlajamäki, O. Santala, and V. Pulkki, "Synthesis of spatially extended virtual source with time-frequency decomposition of mono signals," *J. Audio Eng. Soc.*, vol. 62, no. 7/8, pp. 467–484, aug 2014. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=17339>
- [7] D. Kim-Boyle, "Spectral and Granular Spatialization with Boids," in *ICMC*. Michigan Publishing, 2006, pp. 139–142.
- [8] A. Bonardi, "Composer l'espace sonore," in *Revue Francophone d'Informatique Musicale*, no. 7-8 Culture du code, 2020. [Online]. Available: <https://revues.mshparisnord.fr/rfim/index.php?id=624>
- [9] A. Sedes, *Espaces sonores actes de recherche*. Paris: éd. musicales transatlantiques, 2003.
- [10] A. Sèdes, P. Guillot, and E. Paris, "The HOA library, review and prospects," in *International Computer Music Conference — Sound and Music Computing*, ser. Proceedings ICMC—SMC, Athènes, Greece, Sep. 2014, pp. 855–860. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01196453>
- [11] A. Bonardi, "La librairie abclib : un ensemble de codes Faust rassemblant 20 ans de recherche, enseignement et création en musique mixte," in *Journées d'Informatique Musicale 2021*. Visioconférences, France: AFIM, Jul. 2021. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-03313611>
- [12] M. Stroppa, *Accorder musicalement un espace réel et un espace inventé*. Paris: Presses universitaires de France, 2007.
- [13] N. Barrett, "Spatio-musical composition strategies," *Organised Sound*, vol. 7, no. 3, p. 313–323, 2002.

- [14] H. Vaggione, *L'espace composable. Sur quelques catégories opératoires dans la musique électroacoustique*. L'Harmattan, 1998.
- [15] Y. Etienne, *De l'espace sonore*, Les presses du réel, Ed. Strasbourg, France Dijon, France Monts, Frances: Haute école des arts du Rhin, Les Presses du réel Présence graphique, 2014.
- [16] M. F. Franz Zotter, *Ambisonics*. Springer International Publishing, May 2019. [Online]. Available: https://www.ebook.de/de/product/35683429/franz_zotter_matthias_frank_ambisonics.html
- [17] A. Bonardi and P. Guillot, "Concevoir des traitements ambisoniques en 2D et en 3D : l'exemple de Pianotronics 2," in *Journées d'Informatique Musicale*, ser. Actes des Journées d'Informatique Musicale, vol. 2015. Montréal, Canada: Olivier Bélanger and Nicolas Bernier and Julie Delisle and Pierre Michaud and Robert Normandeau and Caroline Traube, May 2015, p. 8. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01199688>
- [18] P. Goutmann, "Traitement spatial du son par décorrélation des signaux en ambisonie d'ordre élevé," in *Journées d'Informatique Musicale 2021*. Visioconférences, France: AFIM, Jul. 2021. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-03313616>
- [19] A. Sèdes, "Approche musicale de la decorrelation microtemporelle dans la bibliothèque HOA," in *Journées d'Informatique Musicale 2015*, université de Montréal, Ed. Montréal, Canada: Université de Montréal, May 2015. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01199014>
- [20] G. S. Kendall, "The decorrelation of audio signals and its impact on spatial imagery," *Computer Music Journal*, vol. 19, no. 4, p. 71, 1995.
- [21] H. Vaggione, "Décorrélation microtemporelle, morphologies et figurations spatiales," in *Journées d'Informatique Musicale*, Marseille, France, May 2002. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-02992872>

6. APPENDIX

List of sound examples quoted in the paper with short descriptions and links to HAL repository. All audio examples are mono sound encoded and processed in 2D 3th order, and decoded in binaural (48kHz, 24bits).

6.1 Sound 1

- source: loop of an audio excerpt (10sec.) from the composition *Compression-Dilatation 6.6*
- description: the sound is presented five times, first without decorrelation (10sec.), second with the factor increasing, third with the factor at 1, fourth with the factor decreasing in 10 seconds and finally the sound without decorrelation (10sec.) ;

- parameter: delay = 1000ms, fonctiontype = 0, factor goes from 0 to 1 in 10 seconds then return to 0 in 10 seconds ;
- <https://hal.archives-ouvertes.fr/hal-03637787>

6.2 Sound 2

- source: loop of a clarinet playing A3 ordinario ;
- description: we hear the effect of the factor of ring modulation ;
- parameter: f0 = 10, factor goes from 0 to 1 in 10 seconds and after 5 seconds returns to 0 in 10 seconds after 5 seconds ;
- <https://hal.archives-ouvertes.fr/hal-03637789>

6.3 Sound 3

- source: looped sample of a doublebass doing a Bartok pizzicato ;
- description: we hear the insertion of static delay line when the factor increases. First one echo, then two, etc., until we heard seven distinct echos (each harmonic - at the third order in 2D) ;
- parameter: delay = 1000ms, fonctiontype = 0, factor goes from 0 to 1 in 25 seconds ;
- <https://hal.archives-ouvertes.fr/hal-03637800>

6.4 Sound 4

- source: looped sample of a doublebass doing a Bartok pizzicato ;
- description: we hear the delay time of each harmonic increasing progressively when the factor increases ;
- parameter: delay = 1000ms, fonctiontype = 0, factor goes from 0 to 1 in 25 seconds ;
- <https://hal.archives-ouvertes.fr/hal-03637802>

6.5 Sound 5

- source: looped sample of a doublebass doing a Bartok pizzicato ;
- description: we hear the effect of the different delay distribution between the ambisonics component ;
- parameter : delay = 1000ms, factor = 1, fd = 0, fonctiontype goes from 0 to 22 ;
- <https://hal.archives-ouvertes.fr/hal-03637803>

6.6 Sound 6

- title: *Compression - Dilatation* ;
- description: this is a sketch of electroacoustic piece composed in the context of the Ph.D research of Goutmann ;
- <https://hal.archives-ouvertes.fr/hal-03642163>